

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

EXTRACTION DE CONNAISSANCES À PARTIR DE TRACES
MULTIDIMENSIONNELLES DE MOUVEMENTS OCULAIRES POUR UN
TRANSFERT D'EXPERTISE PERCEPTIVO-DÉCISIONNELLE DANS LE
DOMAINE MÉDICAL

THÈSE

PRÉSENTÉE

COMME EXIGENCE PARTIELLE DU

DOCTORAT EN INFORMATIQUE COGNITIVE

PAR

KOMI SÉPÉLI SODOKÉ

DÉCEMBRE 2020

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de cette thèse se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.10-2015). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Au terme de ce travail, j'exprime ma reconnaissance à toutes les personnes qui m'ont orienté et soutenu durant mes années d'études doctorales et de rédaction de la thèse. Ce fut plus que des années d'étude académique; mais toute une expérience marquante sur un plan intellectuel, de résilience face aux défis et d'aboutissements personnels.

J'ai eu le bonheur d'être soutenu par ma conjointe Yasmina Ayéva et mes deux enfants Maéva et Maliah Sodoké; suivi par des directeurs extraordinaires et rencontrés des personnes mémorables. Cette partie du manuscrit leur est spécialement dédiée.

Je remercie tout particulièrement mon directeur de thèse, monsieur Roger Nkambou, ma codirectrice madame Aude Dufresne et mon codirecteur Issam Tanoubi. Leur confiance, leur patience, leurs conseils avisés et leur rigueur scientifique ont permis la réalisation de ce travail. Cette thèse s'inscrit dans les activités du laboratoire GDAC en collaboration avec la Cite et le CAAHC. Ces laboratoires m'ont fourni le cadre qui a permis de mener à terme ce travail. Ils ont également mis à ma disposition d'autres chercheurs et collègues dont les suggestions et commentaires ont positivement influencé notre recherche.

Je remercie tous les membres de ma famille pour leurs encouragements constants durant de nombreuses années et qui m'ont motivé à entreprendre mes études doctorales.

À tous les professeurs qui me feront l'honneur de contribuer à l'évaluation de ce travail j'adresse mes gratitude.

TABLE DES MATIÈRES

| | |
|---|------|
| LISTE DES FIGURES..... | viii |
| LISTE DES TABLEAUX..... | xi |
| RÉSUMÉ | xiii |
| ABSTRACT | xiv |
| INTRODUCTION : mise en contexte, problématiques, hypothèses et objectifs de recherches..... | 15 |
| 1.1 Mise en contexte | 16 |
| 1.2 Problématiques | 19 |
| 1.3 Hypothèses de recherches..... | 25 |
| 1.4 Objectifs de recherche | 29 |
| 1.5 Organisation du Manuscrit | 31 |
| CHAPITRE II Les éléments du cadre théorique..... | 35 |
| 2.1 Les Systèmes Tutoriels Intelligents (STI) | 35 |
| 2.1.1 Définitions et objectifs | 35 |
| 2.2 Les modules de l'architecture générique d'un STI..... | 36 |
| 2.2.1 Module de communication..... | 37 |
| 2.2.2 Modèle de l'apprenant..... | 37 |
| 2.2.3 Modèle de l'expert/des connaissances du domaine | 40 |
| 2.2.4 Module pédagogique..... | 43 |

| | | |
|---|---|----|
| 2.3 | Le diagnostic cognitif des connaissances dans les STI | 48 |
| 2.4 | La métacognition dans les STI | 52 |
| 2.5 | Les domaines mal-définis : de l'IA au STI..... | 55 |
| 2.5.1 | Définition et mise en contexte..... | 55 |
| 2.5.2 | Limites des paradigmes classiques des STI pour les domaines mal-définis 57 | |
| 2.6 | L'expertise : définition et étapes d'acquisition..... | 59 |
| 2.6.1 | Définition et caractéristiques..... | 59 |
| 2.7 | Conclusion et discussions | 62 |
| CHAPITRE III Apprentissage machine et réseaux de neurones artificiels | | 66 |
| 3.1 | Introduction..... | 66 |
| 3.2 | Apprentissage machine | 68 |
| 3.2.1 | La tâche T..... | 69 |
| 3.2.2 | L'expérience E | 69 |
| 3.2.3 | La mesure de performance M | 71 |
| 3.3 | Formalisation des notions en apprentissage machine et exemple | 72 |
| 3.4 | Les réseaux de neurones : définition, principes et historique | 73 |
| 3.4.1 | Le neurone biologique..... | 73 |
| 3.4.2 | Du neurone biologique au neurone artificiel (neurone formel) | 74 |
| 3.5 | Historique et évolution | 76 |
| 3.6 | Apprentissage dans les RNA (cas du perceptron multicouche)..... | 79 |
| 3.7 | Réseau de neurones récurrent (RNN)..... | 85 |
| 3.8 | Architectures profondes..... | 86 |
| 3.9 | Discussion et positionnement | 89 |
| CHAPITRE IV Extraction et prédiction de séquences | | 92 |
| 4.1 | Définitions et mise en contexte..... | 92 |
| 4.2 | État de l'art, défis et propositions de solutions..... | 94 |
| 4.3 | Les tâches en fouilles de données séquentielles | 95 |
| 4.4 | L'algorithme CMRules..... | 98 |

| | | |
|---|--|-----|
| 4.5 | Défis des algorithmes de fouille de données séquentielles et propositions de solutions | 100 |
| 4.6 | L'algorithme PhARules | 101 |
| 4.7 | Discussion et positionnement | 102 |
| CHAPITRE V Méthodologie, cadre expérimental et expérimentations | | 106 |
| 5.1 | Mise en contexte | 106 |
| 5.2 | Technologie pour la collecte des données | 107 |
| 5.3 | Participants | 108 |
| 5.4 | Déroulement de l'expérimentation | 109 |
| 5.5 | Collecte et validation des données | 115 |
| 5.6 | Modélisation, caractérisation et traitement des données | 117 |
| 5.7 | Processus générique de traitement des traces de résolution | 118 |
| CHAPITRE VI Analyses et résultats obtenus | | 121 |
| 6.1 | Prétraitement initial | 122 |
| 6.2 | Analyses préliminaires | 123 |
| 6.3 | Analyse comportementale des mouvements oculaires autour des évènements clés | 130 |
| 6.4 | Analyse de la trajectoire globale (<i>scanpath</i>) des fixations et saccades | 135 |
| 6.5 | Conclusion et discussion | 140 |
| CHAPITRE VII Classification des séquences - DeepConvLSTM Vs WKM-1NN | | 144 |
| 7.1 | Introduction | 144 |
| 7.2 | Nature des données | 145 |
| 7.3 | Choix de modèle et classification des séquences | 148 |
| 7.3.1 | Approches fondées des techniques d'apprentissage connexionniste | 149 |
| 7.3.2 | Approches fondées des techniques d'apprentissage statistique | 156 |
| 7.4 | Discussion et positionnement | 160 |
| CHAPITRE VIII Extraction de motifs comportementaux | | 162 |

| | | |
|--|--|-----|
| 8.1 | Introduction..... | 162 |
| 8.2 | Nature des données | 163 |
| 8.3 | Spécificités et contraintes sur les données de mouvements oculaires | 168 |
| 8.4 | Étude et comparaisons de quelques algorithmes d'extraction de motifs..... | 169 |
| 8.4.1 | Algorithme APriori | 170 |
| 8.4.2 | Algorithme FP-Growth | 171 |
| 8.4.3 | Algorithme ECLaT..... | 172 |
| 8.4.4 | Algorithme N-gram..... | 173 |
| 8.5 | Discussions et limites des algorithmes disponibles pour l'extraction de séquence | 174 |
| 8.6 | L'algorithme MDSE (Multidimensional Dynamic Sequence Extraction) | 175 |
| 8.6.1 | Entrée de l'algorithme MDSE..... | 176 |
| 8.6.2 | Sortie de l'algorithme MDSE | 178 |
| 8.6.3 | Déroulement de l'algorithme MDSE | 180 |
| 8.6.4 | Résultats obtenus..... | 183 |
| 8.7 | Discussion et positionnement | 186 |
| CHAPITRE IX Spécifications et fonctionnement du sti EyeLab | | 189 |
| 9.1 | Utilisation de l'oculométrie cognitive | 189 |
| 9.1.1 | Marketing web | 190 |
| 9.1.2 | Médical..... | 190 |
| 9.1.3 | Domaine militaire..... | 191 |
| 9.1.4 | Environnement d'apprentissage en ligne | 191 |
| 9.2 | Oculométrie cognitive et transfert d'expertise | 192 |
| 9.3 | Fondement du cadre pédagogique dans le STI EyeLab..... | 193 |
| 9.4 | Cadre d'utilisation et fonctionnalités du STI EyeLab | 195 |
| 9.5 | Guidage des mouvements oculaires dans le STI EyeLab..... | 197 |
| 9.6 | Discussion et positionnement | 199 |
| CHAPITRE X Conception et Implémentation du STI eyeLab..... | | 201 |
| 10.1 | Quelques défis de l'implémentation de EyeLab..... | 201 |
| 10.2 | Le choix d'architecture | 204 |

| | |
|--|-----|
| 10.3 Quelques étapes du mécanisme de fonctionnement avec les composants associés..... | 206 |
| 10.4 Processus d'extraction des données oculaires | 208 |
| 10.5 Les choix technologiques pour la persistance de données..... | 211 |
| CONCLUSION : Contributions de la thèse, perspectives et travaux futurs | 214 |
| Contribution de la thèse | 214 |
| Quelques limites dans les travaux | 219 |
| Travaux futurs, expérimentations du STI et évaluation de la généricité..... | 221 |
| ANNEXE A Questionnaire post expérimentation | 223 |
| ANNEXE B Extrait de séquences prétraitées pour l'exécution de l'algorithme de Smith-Waterman | 226 |
| ANNEXE C Exemple de motifs extrait par l'algorithme MDSE pour le novice 5 | 227 |
| APPENDICE A Interface d'administration | 229 |
| APPENDICE B Interface de détection des pupilles et du visage | 230 |
| APPENDICE C Interface de détection et calibration de mouvements oculaires.... | 231 |
| APPENDICE D Interface de EyeLab en mode tutorat | 232 |
| APPENDICE E Interface de comparaison des séquences de mouvements oculaires | 233 |
| BIBLIOGRAPHIE | 235 |

LISTE DES FIGURES

| Figure | Page |
|--|------|
| Figure 1 : Les étapes du processus de raisonnement hypothético-déductif | 22 |
| Figure 2 : Catégorisation des dimensions des domaines mal-définis selon Mitrovic et Weerasinghe..... | 28 |
| Figure 3 : Architecture générale d'un STI (adaptée de Nwana, 1990) | 36 |
| Figure 4 : Représentation du modèle du type recouvrement, avec une variante incorporant les connaissances erronées de l'apprenant | 39 |
| Figure 5 : Intégration de la réflexion dans un cadre de base pour un diagnostic cognitif adapté de (Tchétagani, Nkambou et al. 2006) | 51 |
| Figure 6 : Positionnement de l'IA, de l'apprentissage machine et de l'apprentissage profond (deep learning) imbriquées les unes aux autres | 67 |
| Figure 8 : Neurone avec gaine de myéline..... | 74 |
| Figure 9 : Neurone artificiel..... | 75 |
| Figure 10 : Le perceptron multicouche | 80 |
| Figure 11 : Descente de Gradient..... | 82 |

| | |
|---|-----|
| Figure 12 : Exemple d'un réseau récurrent ouvert montrant les entrées séquentielles et l'état interne les sorties | 85 |
| Figure 13 : Photo du moniteur avec l'oculomètre TobiiEyeX® TX300 | 111 |
| Figure 14 : Vidéo de la simulation vue par les participants avec les AOI..... | 113 |
| Figure 15 : Figure montrant les fixations et saccades un participant..... | 116 |
| Figure 16 : Modules et processus génériques de traitement des traces de résolution | 119 |
| Figure 17 : Moyenne de la durée totale de fixation de Novice Vs Expert | 124 |
| Figure 18 : Moyenne du nombre de fixations Novice Vs Expert | 124 |
| Figure 19 : Moyenne de la durée des visites Novice Vs Expert | 125 |
| Figure 20 : Heatmap pour les novices..... | 126 |
| Figure 21 : Heatmap avec pourcentage pour les novices | 126 |
| Figure 22 : Heatmap pour les experts | 127 |
| Figure 23 : Heatmap avec pourcentage pour les experts..... | 127 |
| Figure 24 : Comparaison du nombre de fixations dans les différents AOI pour les novices et les experts..... | 129 |
| Figure 25 : Topologie de l'architecture DeepConvLSTM..... | 153 |

| | |
|---|-----|
| Figure 26 : Performance obtenue par l'architecture DeepConvLSTM..... | 156 |
| Figure 27 : Convergence obtenue par l'architecture DeepConvLSTM | 156 |
| Figure 28 : Graphique à barres illustrant la dynamique de changement dans les AOI pour un ensemble de 1000 points de données avant la suppression des spikes | 166 |
| Figure 29 : Illustration des étapes de prétraitement | 166 |
| Figure 30 : Graphique à barres illustrant la dynamique de changement dans les AOI pour un ensemble de 1000 points de données après la suppression des spikes | 168 |
| Figure 31 : Diagramme des micro-services dans EyeLab..... | 205 |
| Figure 32 : Diagramme des composants de EyeLab..... | 207 |
| Figure 33 : L'approche Haar Cascades avec 68 points de données pour le contour de la face, des yeux du nez et de la bouche | 209 |
| Figure 34 : Processus d'extraction des données oculaires | 211 |

LISTE DES TABLEAUX

| Tableau | Page |
|---|------|
| Tableau 1 : Étapes d'acquisition d'expertise ainsi que la description associée..... | 62 |
| Tableau 2 : Quelques algorithmes classiques de fouille séquentielle | 97 |
| Tableau 3 : Évènements "Clés" de la simulation pour l'analyse comportementale.. | 114 |
| Tableau 4 : Données pour l'analyse..... | 123 |
| Tableau 5 : Heatmap progressif des mouvements oculaires des novices et des experts autour des évènements clés | 131 |
| Tableau 6 : Résumé des observations sur les heatmaps progressifs des mouvements oculaires autour des évènements clés..... | 133 |
| Tableau 7 : Exemple type de trajectoire. Les saccades des novices sont indiquées avec un fond en rouge et celles des experts en bleu | 136 |
| Tableau 8 : Tableau croisé des scores de similitudes des scanpth des novices obtenus via l'algorithme de Smith-Waterman..... | 139 |
| Tableau 9 : Tableau croisé des scores de similitudes des scanpth des experts obtenus via l'algorithme de Smith-Waterman..... | 139 |

| | |
|---|-----|
| Tableau 10 : Illustration d'une séquence de donnée utilisée pour l'apprentissage et la classification..... | 147 |
| Tableau 11 : Nombre de séquences de mouvement oculaires par classe pour l'ensemble des données | 148 |
| Tableau 12 : Résumé des résultats de l'expérimentation | 155 |
| Tableau 13 : Résumé des performances obtenues par le modèle WKM+1NN..... | 160 |
| Tableau 14 : Illustration d'une séquence de données utilisée pour des fouilles de données séquentielles avec les données manquantes et imprécises | 164 |
| Tableau 15: Illustration d'une séquence de donnée après prétraitement | 167 |
| Tableau 16 : Exemple de transactions pour une extraction de motifs..... | 170 |
| Tableau 17 : Exemple du Dimensional Sequence Flattening | 177 |
| Tableau 18 : Un extrait de séquence des mouvements oculaires pour l'expert 5 après l'opération DSF | 178 |
| Tableau 19 : Nomenclature de la représentation des motifs extraits | 179 |
| Tableau 20 : Déroulement de l'algorithme MDSE | 181 |
| Tableau 21 : Exemple de motifs extraits par l'algorithme MDSE pour l'expert 5 ... | 184 |
| Tableau 22 : Exemple de motifs similaires pour les experts..... | 186 |

RÉSUMÉ

L'acquisition d'expertise a été étudiée dans plusieurs domaines et ces recherches démontrent des changements de paradigme dans les différentes phases de l'évolution du stade de novice à celui de l'expert. Ce projet de recherche s'inscrit dans ce cadre général et porte spécifiquement sur l'expertise que nous qualifions de « perceptivo-décisionnelle » dans le domaine médical. L'approche adoptée pour la recherche est constituée de trois phases. La première phase consistera en une étude exploratoire et comparative des aptitudes perceptivo-décisionnelle de cliniciens novices et experts en situation authentique en utilisant une simulation haute-fidélité. Cette recherche se fera en prenant en considération la perception visuelle pour l'exécution d'une tâche cognitive qui est le raisonnement clinique. La deuxième phase consistera d'une part en l'analyse des données pour une classification des séquences oculaires et d'autre part l'extraction de motifs dans le but de comparer et dégager des régularités caractéristiques des mouvements oculaires des novices et des experts. La dernière phase exploitera les extraits des deux premières en vue d'élaborer des spécifications et la conception d'un système tutoriel intelligent basé sur l'analyse de mouvement oculaire via webcam qui permettrait d'offrir des services tutoriels à un clinicien novice afin de structurer graduellement ses mouvements oculaires comme un expert.

Mots clés : systèmes tutoriels intelligents, apprentissage profond, classification de séquences oculaire, extraction de séquences, oculométrie, webcam EyeTracking

ABSTRACT

The acquisition of expertise has been studied in several fields and this research demonstrates the major paradigm shifts during the different phases of evolution from the novice stage the expert. This research project falls within this general scope and focus specifically on the expertise that we qualify as "perceptual-decision-making" in the medical field. The research approach consists of three phases. The first phase will consist of an exploratory and comparative study of the perceptual-decision-making skills of novice and expert clinicians in authentic situations using high-fidelity simulation. This research will be performed by considering eye movements during a clinical reasoning task. The second phase will consist of data analysis for eye movement sequences classification and eye movement pattern extraction to identify characteristic regularities within novices and experts eye movements. The last phase will use the outputs of the first two to design and implement a webcam based EyeTacking intelligent tutoring system that would provide tutorial services to a novice clinician in order to structure his eye sight as an expert gradually.

Keywords : intelligent tutoring system, deep learning, eye movement sequence classification, sequence extraction, webcam EyeTracking

INTRODUCTION : MISE EN CONTEXTE, PROBLÉMATIQUES, HYPOTHÈSES ET OBJECTIFS DE RECHERCHES

1.0 Préambule

L'acquisition d'expertise a été étudiée dans plusieurs domaines et ces recherches démontrent des changements de paradigme dans les différentes phases de l'évolution du stade de novice à celui de l'expert. Ce projet de recherche s'inscrit dans ce cadre général et porte spécifiquement sur l'expertise que nous qualifions de « perceptivo-décisionnelle » dans le domaine médical. L'approche adoptée pour la recherche est constituée de trois phases. La première phase consistera en une étude exploratoire et comparative des aptitudes perceptivo-décisionnelle de cliniciens novices et experts en situation authentique en utilisant une simulation haute-fidélité. Cette recherche se fera en prenant en considération de la perception visuelle pour l'exécution d'une tâche cognitive qui est le raisonnement clinique. La deuxième phase consistera d'une part en l'analyse des données pour une classification des séquences comportementales en utilisant uniquement les données de perception visuelle et d'autre part l'extraction de séquence dans le but de dégager des régularités caractéristiques du novice et de l'expert. La dernière phase exploitera les extraits des deux premières en vue d'élaborer des spécifications et la conception d'un système tutoriel intelligent qui permettrait d'offrir des services tutoriels au novice afin de structurer graduellement ses connaissances et aptitudes comme un expert.

L'expertise perceptivo-décisionnelle se fonde sur les connaissances déclaratives et procédurales du domaine; elle implique toutefois d'autres dimensions qui sont

difficiles à caractériser mais qui influencent la performance. Des capteurs seront donc utilisés pour récupérer de manière objective toutes les modalités de perceptions visuelles impliquées dans les étapes de résolution. Ceci pose des problèmes en raison de la complexité temporelle et multimodale des données. Également, les données doivent être analysées de manière à extraire de façon cohérente des connaissances qui seront formalisées dans une perspective «pro-pédagogique » exploitable pour concevoir un STI. Cet objectif pose d'autres défis, car les traces de résolution proviennent d'une tâche qui présente les caractéristiques d'un domaine mal-définis. Ces problématiques seront présentées ainsi qu'une proposition de solution basée sur une approche novatrice d'extraction de connaissances de nature comportementale dans un domaine. La solution proposée combinera des algorithmes d'apprentissage machine pour l'extraction de motifs qui seront contextualisés sans nécessiter une modélisation cognitive explicite des connaissances comportementales du domaine. Les spécifications pour les services tutoriels se feront à travers de nouvelles perspectives spécifiques aux objectifs en considérant les paradigmes de cognition impliqués.

1.1 Mise en contexte

L'acquisition d'expertise a été étudiée dans plusieurs domaines et ces recherches démontrent des changements de paradigme dans les différentes phases lors de l'évolution du novice à l'expert (Chase and Simon 1973, Benner 1982, Anderson 1987). De nombreux facteurs peuvent expliquer cette situation. Selon le modèle de Dreyfus (Hubert Dreyfus 1986), lors du passage du stade de novice à expert, la notion de contexte spécifique, de vision globale et d'intuition prennent le dessus sur les connaissances théoriques de règles et procédures du domaine. Également, pour certains tenants du paradigme de la cognition incarnée, les interactions sensori-motrices avec l'environnement sont nécessaires pour supporter ce processus d'acquisition d'expertise (Francisco Varela 1993, Feltovich, Ericsson et al. 2006). Le raisonnement clinique (RC)

qui, se définit comme : « l'activité intellectuelle qui synthétise l'information obtenue à partir de la situation clinique, l'intègre aux connaissances et aux expériences antérieures et l'utilise pour prendre des décisions de diagnostic et de prise en charge du patient » (Vleuten, GR et al. 2008) évolue aussi du stade de novice vers le stade d'expert, en intégrant d'une façon inégale et dépendante du contexte, un raisonnement pouvant être analytique ou intuitif (Monteiro and Norman 2013). Spécifiquement, dans certains domaines médicaux comme l'anesthésiologie, un aspect important à développer dans le processus d'acquisition d'expertise porte sur la « sensibilité situationnelle ». Dans un contexte clinique donné, la sensibilité situationnelle se définit comme la perception d'éléments et d'événements environnementaux en ce qui concerne le temps ou l'espace, la compréhension de leur signification et la projection de leur progression future (Endsley 2015). Dans certains domaines de la médecine comme l'anesthésiologie, la perception visuelle est cruciale, car le clinicien a besoin de visualiser adéquatement les signes vitaux du patient tels que la fréquence cardiaque, la saturation artérielle, la pression artérielle, la fréquence respiratoire, etc. afin de développer sa compréhension et interprétation de la situation clinique, et prendre des actions appropriées pour y répondre. La perception visuelle est un facteur clé qui pourrait moduler l'expertise, influencer la pratique et les performances du clinicien. Elle représente la pointe de l'iceberg de la sensibilité situationnelle en interrelation avec d'autres facteurs comme la gestion de la surcharge cognitive et la gestion des distracteurs.

Le projet de recherche porte sur l'expertise que nous qualifions de «*perceptivo-décisionnelle*» dans le cadre de la sensibilité situationnelle en anesthésiologie. C'est l'aptitude du clinicien de réguler adéquatement sa perception visuelle pour capter les informations nécessaires à la compréhension du cas clinique pour supporter son raisonnement et ses décisions cliniques. Une telle expertise est multimodale et combine la perception en interrelation avec un processus décisionnel exploitant des

connaissances du domaine. En anesthésiologie, cette expertise *perceptivo-décisionnelle* n'est pas une aptitude connexe, mais constitue une facette à part entière des aptitudes à acquérir intrinsèquement reliée aux connaissances du domaine ainsi que les actions de prise en charge qui les accompagnent, le tout ne s'exécutant pas nécessairement séquentiellement. D'autres domaines nécessitent une telle expertise *perceptivo-décisionnelle*. À titre d'exemple, nous pouvons citer le pilotage d'avion, car le balayage visuel des instruments du tableau de bord est lié aux connaissances relatives aux contrôles de l'état de l'appareil à différentes phases de vol (Sandeep, Mehran et al. 1995, Remolina, Ramach et al. 2004).

L'expertise perceptivo-décisionnelle dans le domaine médical est empirique et acquise par la pratique dans des situations réelles ou dans un environnement de simulation. Le CAAHC (Centre d'apprentissage des attitudes et habiletés cliniques) issu d'un partenariat en recherche et développement entre la faculté de médecine de l'Université de Montréal et CAE¹ Santé a mis en place une approche originale d'apprentissage des habiletés cliniques par simulation. De plus, le CAAHC accueille de nombreux stagiaires qui sont des étudiants en médecine ou des professionnels. Le CAAHC dispose donc d'un cadre idéal pour nos recherches sur l'expertise perceptivo-décisionnelle, cadre qui nous a été offert pour mener nos études. Dans le cadre de ce projet de recherche, le CAAHC agit également comme partenaire pour la définition et la validation du protocole de recherche ainsi que la recherche de participants aux expérimentations prévues.

¹ <http://www.cae.com/fr/>

1.2 Problématiques

Dans le contexte général du changement de paradigme lors du passage du stade de novice à celui d'un expert, comment s'opère ce changement dans le domaine spécifique de l'expertise perceptivo-décisionnelle en médecine ? Est-il possible à partir de l'analyse des données de perception visuelles et du raisonnement clinique subséquent de dégager des spécificités caractéristiques de l'expert et du novice? Ces caractéristiques justifient-elles en partie les performances supérieures de l'expert ? Si oui, peuvent-elles être modélisées en vue d'être exploitées pour la production de services tutoriels visant à faire acquérir plus rapidement ces expertises à un novice ? Ces questions de recherche soulèvent diverses problématiques. Nous présentons ci-dessous celles que nous adressons dans le cadre de cette thèse.

Problématique 1 : Existe-t-il une différence entre la perception visuelle du novice et de l'expert ?

Le (CNRTL 2012) définit la perception comme étant « *une opération psychologique complexe par laquelle l'esprit, en organisant les données sensorielles, se forme une représentation des objets extérieurs et prend connaissance du réel* ». Il existe plusieurs types de perceptions dont la perception visuelle, auditive, olfactive, tactile, gustative, du temps et de l'espace, etc. Quel rôle joue cette faculté dans des situations de raisonnement? Quel rôle joue-t-elle dans le processus d'acquisition de l'expertise médicale ? La perception visuelle fait-elle partie des aptitudes qui caractérisent et différencient un novice d'un expert ?

La perception visuelle est un processus complexe et proactif qui présente certaines caractéristiques :

- La vision est intrinsèquement liée à d'autres dimensions cognitives comme la mémoire. Nous faisons des liens entre ce qu'on observe et ce qu'on a déjà observé.
- La vision possède la propriété de détecter et de discriminer. L'observation est faite globalement, mais l'accent est généralement mis sur certains objets plus que d'autres. Ce qu'on observe est généralement lié à un contexte et un objectif.
- La vision ne se limite pas à la représentation d'objets; c'est aussi une activité d'exploration, de collecte d'informations et d'acquisition de connaissances.

Ces caractéristiques pourraient toutefois constituer la source de certaines problématiques que nous soulevons dans le cadre de ce projet. En effet, avec la capacité de sélectionner implicitement les informations jugées les plus pertinentes dans son environnement, il est évident que selon un contexte ou selon les objectifs, une même réalité pourrait donc être perçue différemment d'un individu à l'autre, d'une situation à l'autre. Toutefois, cette différence pourrait-elle être caractéristique du niveau d'expertise dans le domaine ou le phénomène est observé ? Si oui, est-ce le cas dans l'expertise perceptivo-décisionnelle ?

Problématique 2 : Existe-t-il une interdépendance entre la perception visuelle et le RC dans le cadre de l'expertise perceptivo-décisionnelle ?

De nombreuses études confirment des différences entre novices et experts dans leur performance dans le raisonnement clinique (Nendaz, Charlin et al. 2005, May, Withers et al. 2010). Une de ces différences est le nombre élevé d'hypothèses par les novices avant d'arriver à un diagnostic final. En général, ce problème se pose moins chez un expert. Selon (Neistadt 1996). Ceci se justifie par le fait que l'expert dans un contexte donné, peut formuler des hypothèses pertinentes plus rapidement; adapter les principes médicaux en réponse à des situations spécifiques et percevoir l'importance relative de toutes les informations. Ces caractéristiques de l'expert, découleraient selon (Gruppen

and Frohna 2000) d'une structure de connaissances forgée par l'expérience qu'ils appliquent inconsciemment et non d'un raisonnement hypothético-déductif formel. Le cycle de conversion de la connaissance de Hirotaka Takeuchi présente ce cadre de manière plus générale.

La figure 1 présente les étapes du processus formel de raisonnement hypothético-déductif. L'acquisition des données cliniques est conditionnée par la génération précoce d'hypothèses diagnostiques (étape 2). Une reconnaissance de patrons permet à l'expert de générer des hypothèses pertinentes plus rapidement et prendre des décisions opportunistes, mais efficaces de prise en charge (Franklin, Liu et al. 2011). L'étape 3 permet la collecte d'évidences qui serviront de prémisses au RC. C'est à cette étape que les difficultés les plus caractéristiques pour un novice sont observées. Ceci provient du fait que ces prémisses (par exemple les symptômes) sont de natures diverses et peuvent être absentes, bruitées ou tout simplement non remarquées par le clinicien. L'observation visuelle du cas clinique constitue un des moyens pour la collecte de ces prémisses. Or, dans la problématique 1 nous avons établi que la vision n'est en fait qu'une représentation mentale de la réalité influencée par une interprétation implicite via la mémoire et les connaissances préalables. Par conséquent, avec l'hypothèse que les prémisses collectées par la perception visuelle du clinicien pourraient être différentes du novice à l'expert, quel impact ceci pourrait-il avoir sur les étapes subséquentes du RC ?

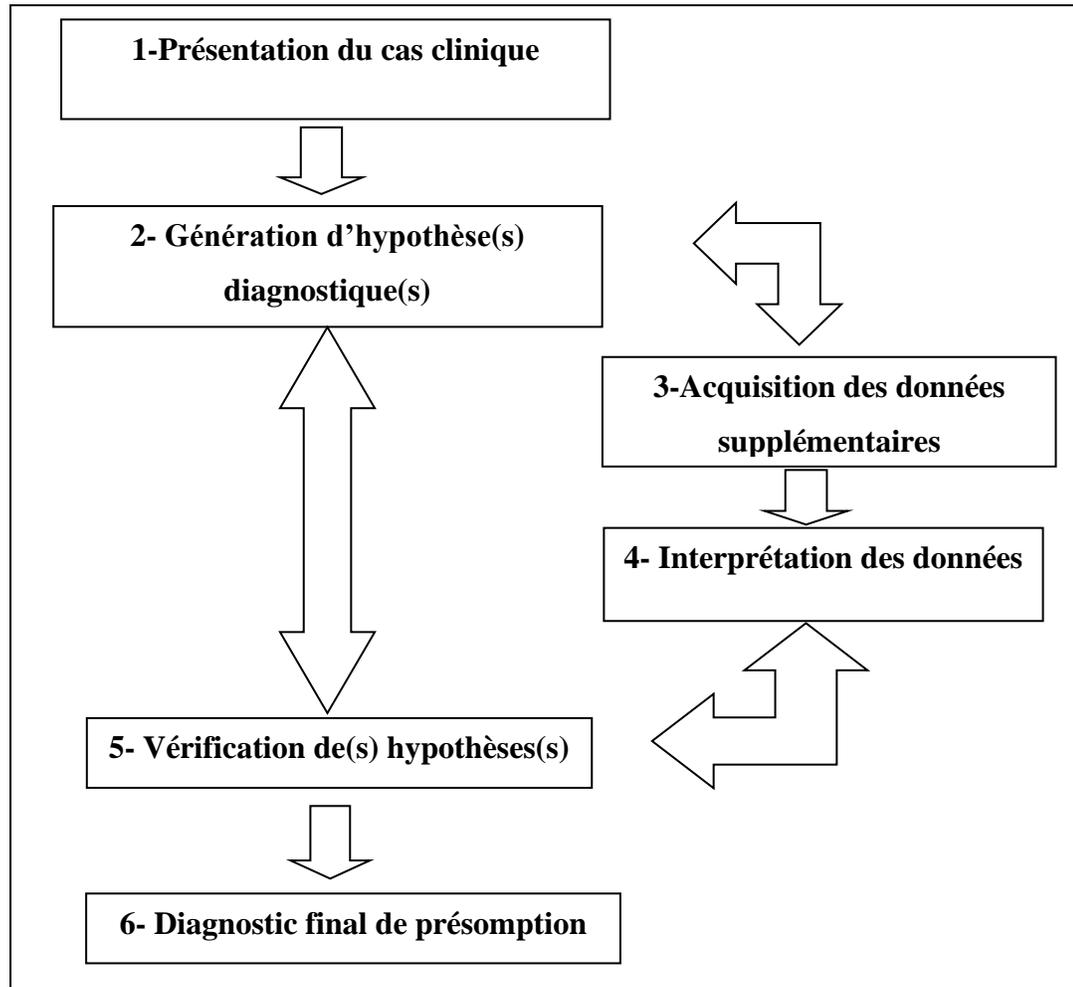


Figure 1 : Les étapes du processus de raisonnement hypothético-déductif

Problématique 3 : Comment peut-on acquérir et modéliser les connaissances du domaine pour favoriser le transfert de l'expertise perceptivo-décisionnelle ?

Existe-t-il des connaissances dans le domaine de l'expertise perceptivo-décisionnelle pour supporter ce processus interdépendant allant de la perception des évidences au raisonnement clinique ? Si oui, une analyse cognitive des tâches permettrait d'établir une description formelle de ces connaissances déclaratives et procédurales ou tout au

moins des techniques utilisées par un expert dans la réalisation des tâches. Ce processus d'élicitation explicite des connaissances peut s'avérer toutefois très coûteux voire impossible, dans certains domaines principalement pour deux raisons. La première c'est qu'il n'est pas toujours facile pour l'expert de verbaliser des expertises acquises par l'expérience, celles-ci étant souvent fondées sur des heuristiques, des connaissances tacites et implicites plutôt qu'explicites. Cette difficulté est connue sous le nom de « *expert blind spot* » (Nathan, Koedinger et al. 2001). La seconde provient du fait que le RC fait partie de problèmes qualifiés de mal-définis. Ces domaines se caractérisent par une ou plusieurs des points suivants : solutions multiples et discutables, absence de théorie formelle du domaine, structure de tâche mal définie, concepts abstraits ou indéterminés, sous-problèmes dépendants (Lynch, Ashley et al. 2006).

Face aux défis de l'acquisition automatique des connaissances, une des solutions les plus utilisées est la théorie d'Anderson (Anderson 1993) formalisée en une architecture cognitive ² nommée : ACT-R (*Adaptive Control of Thought–Rational*). ACT-R présente toutefois une limite majeure provenant du fait que les règles de production qu'elle préconise sont des modélisations rigides de processus de résolution de problème. Ce qui ne les rend pas adaptés aux domaines mal-définis. Des solutions ont été proposées pour pallier ce problème. Par exemple (Fournier-Vigier 2010) propose une approche d'acquisition automatique des connaissances qui exploite les techniques de fouilles de données. De prime abord, ceci pourrait constituer une solution appropriée à la problématique. Toutefois, dans le cadre spécifique de l'expertise perceptivo-décisionnelle, l'extraction des connaissances du domaine nécessitera l'utilisation de

² D'autres architectures cognitives sont également utilisées dont LIDA basée sur la théorie de Baars (1988) qui soutient que la cognition repose sur l'existence d'un espace de travail global dans un système distribué de modules de traitement de l'information.

plusieurs capteurs couvrant les différentes modalités d'interactions perceptuelles et décisionnelles. Ces traces multisources et hétérogènes seront générées indépendamment par différents capteurs, mais représentent un constituant d'une même activité qui est l'évolution d'un cas clinique. Le défi en amont est de parvenir à analyser ces traces d'activité de sorte qu'elles puissent représenter de manière cohérente les différentes modalités des connaissances perceptuelles sous-jacentes.

Problématique 4 : Est-il possible de transférer à un novice certaines aptitudes caractéristiques de l'expertise du domaine (notamment le comportement visuel) afin d'accélérer son cheminement dans l'acquisition d'expertise ?

Selon (Hagedorn 1996), il faudrait environ trois ans pour passer du raisonnement clinique d'un novice à celui d'un expérimenté et dix ans pour atteindre le niveau d'expert. Ce développement d'expertise se passe selon plusieurs axes de compétences. Dans le cadre spécifique de l'expertise perceptivo-décisionnelle, peut-on extraire chez les experts des séquences comportementales caractéristiques de leur performance supérieure ? Si oui, est-il possible de proposer des approches qui permettraient de les expliciter et de les transmettre aux novices afin de contribuer à accélérer leur cheminement vers le niveau expert ou tout au moins lui faire acquérir certaines des aptitudes expertes? Ceci pourrait être possible si nous mettons en place de mécanismes efficaces permettant l'acquisition et la modélisation des comportements distinctifs des experts. En supposant qu'une telle acquisition et modélisation soient réalisées, les résultats obtenus seraient-ils pertinents du point de vue de l'expertise du domaine? De manière spécifique à l'expertise perceptivo-décisionnelle, les connaissances extraites devraient être en lien avec les connaissances (déclaratives et procédurales) du domaine médical et contenir des heuristiques qui supporteront l'approche hypothético-déductive. Finalement, quelles approches tutorielles et paradigme de cognition devront supporter le processus de transfert aux novices ?

1.3 Hypothèses de recherches

Les trois hypothèses de recherches ci-dessous constitueront les fondements sur lesquels seront basées nos recherches. Elles serviront à l'organisation du travail de recherche et constitueront un moyen efficace pour la structuration du cadre expérimental.

Hypothèse 1 : l'analyse des données issues de la perception visuelle pourrait permettre la détection de régularités spécifiques aux novices et aux experts

Plusieurs recherches démontrent que les mouvements oculaires peuvent être un indicateur d'un comportement de l'utilisateur ainsi que des processus cognitifs engagés (Just and Carpenter 1980, Rayner 1998). Ces recherches ont abouti à l'émergence d'une discipline nommée l'oculométrie cognitive qui vise l'étude des comportements sous-jacents à la perception visuelle et l'interprétation des opérations cognitives associées. Les techniques d'oculométrie cognitive ont été utilisées dans de nombreux domaines comme le marketing web (Nielsen and Pernice 2009) , l'éducation (Homof, Cavender et al. 2004), la médecine (Hutchinson, White et al. 1989) et les environnements d'apprentissage en ligne (Conati and Merten 2007). Ces recherches confirment d'une part que différentes caractéristiques de la perception visuelle peuvent être utilisées comme indices des activités cognitives et d'autre part que les yeux peuvent dépasser leur rôle sensoriel passif pour devenir un organe essentiel en situation d'interaction.

Sur la base des résultats des recherches précédemment mentionnées, nous envisageons que l'oculométrie cognitive en combinaison avec la validation du raisonnement clinique nous permettront de caractériser le processus cognitif engagé et même de trouver certaines régularités spécifiques au novice et à l'expert.

Hypothèse 2 : Certaines connaissances de l'expertise perceptivo-décisionnelle peuvent être extraites et modélisées

Dans une perspective générale, l'expertise perceptivo-décisionnelle présente certaines caractéristiques des domaines mal-défini tel que présenté par (Lynch, Ashley et al. 2006, Mitrovic and Weerasinghe 2009). Des solutions ont été proposées pour faire face à la difficulté de formaliser certains domaines procéduraux et mal-définis (Fournier-Vigier 2010). Elles consistent en général en l'utilisation des techniques d'apprentissage machine pour extraire des espaces problèmes partiels à partir des traces de résolution. Les résultats obtenus n'ont toutefois pas la finesse de la modélisation cognitive, mais offrent trois avantages :

1. S'affranchir de l'analyse cognitive de la tâche
2. S'affranchir de la rigueur des règles de production comme celles de ACT-R
3. Obtenir un modèle évolutif des connaissances par un enrichissement progressif à partir de l'exploitation de nouvelles traces de résolution.

Toutefois, comme nous l'avons soulevé dans les problématiques, l'acquisition des données devant servir à l'extraction des connaissances du domaine nécessitera l'utilisation de capteurs couvrant les différentes modalités impliquées dans la perception visuelle. Une considération qui servira de ligne directrice de l'analyse des données issues de la perception visuelle serait qu'il existerait une cohérence entre : la perception visuelle et le RC. L'existence d'une telle cohérence ou interdépendance se traduirait alors par des corrélations entre les sous-séquences des données issues de la perception visuelle et le RC. En somme, il serait possible de donner un contexte clinique à une séquence de perception visuelle; de découvrir des règles d'associations entre ses sous-séquences de perceptions, la condition clinique et les raisonnements qui pourront servir à dégager des régularités caractéristiques de performances supérieures.

Hypothèse 3 : Un STI pourrait être utilisé pour transférer certaines connaissances comportementales au novice pour accélérer son cheminement vers un profil d'expert

Un STI pourrait être utilisé pour forger l'expérience chez le novice en l'amenant progressivement à structurer sa connaissance de plus en plus comme un expert. Cette hypothèse est fondée sur le fait que les STI ont été utilisés avec succès dans d'autres domaines mal-définis. Il s'agit par exemple du CanadaArmTutor (Fournier-Viger, Nkambou et al. 2013) ainsi que de TELEOS (Chieu, Luengo et al. 2010). Ce succès pourrait s'expliquer par l'exploitation dans les STI d'une modélisation pédagogique des connaissances du domaine pour offrir des services tutoriels personnalisés à l'apprenant.

L'expertise perceptivo-décisionnelle en médecine se fonde sur des connaissances (déclaratives et procédurales) dans le domaine médical qui sont relativement bien définies tout en étant en constante évolution. C'est donc d'autres dimensions impliquées dans l'exécution des tâches et qui vont au-delà des connaissances du domaine qui confèrent la caractéristique de domaine mal-défini. En nous référant à la catégorisation des dimensions caractéristiques des domaines mal-définis proposée (Mitrovic and Weerasinghe 2009), l'expertise perceptivo-décisionnelle en médecine qui regroupe des tâches mal-définies exploitant des connaissances d'un domaine bien défini se place dans le cadran D tel qu'illustré sur la figure 2.

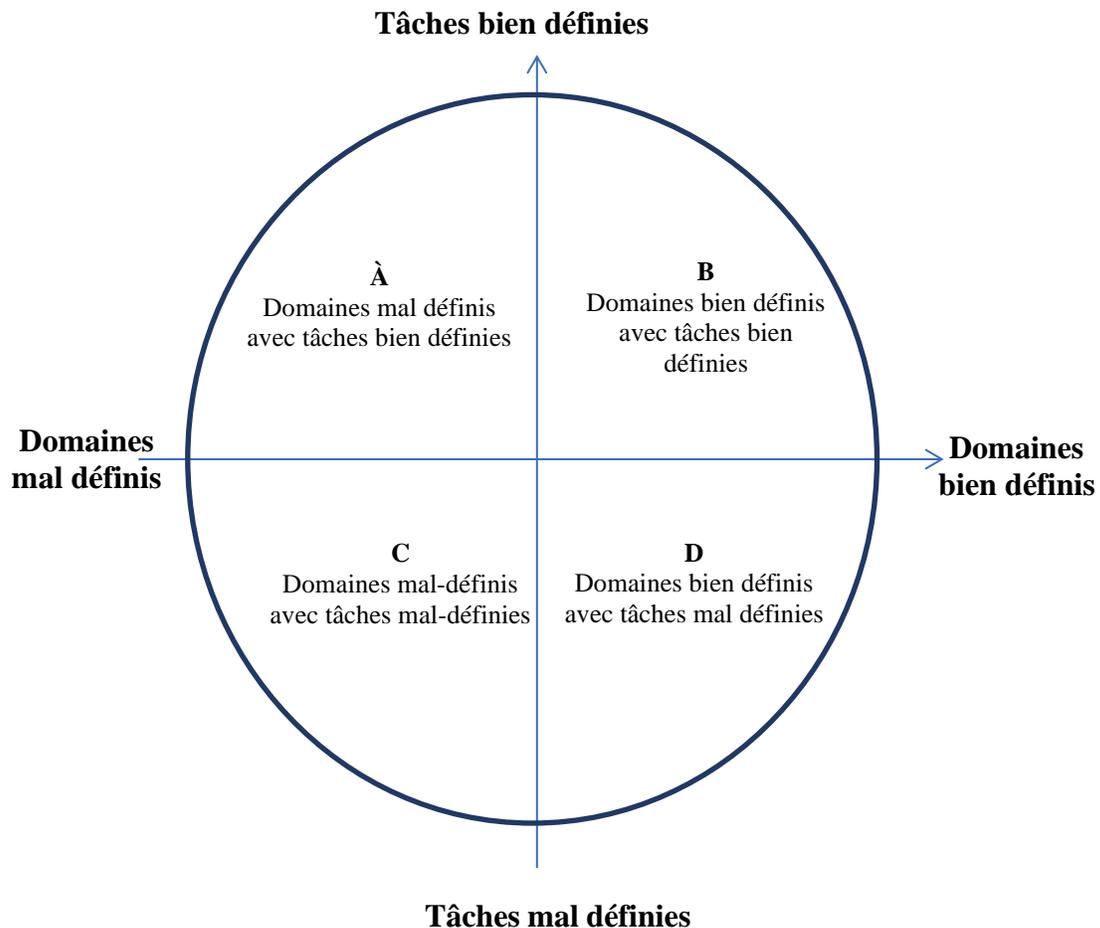


Figure 2 : Catégorisation des dimensions des domaines mal-définis selon Mitrovic et Weerasinghe

Étant donné que l'expertise perceptivo-décisionnelle en anesthésiologie se fonde sur le domaine médical qui dispose de théories et connaissances établies, ces dernières peuvent servir à la validation et la structuration des connaissances qui seront extraites à partir des traces de résolution. Par la suite, une modélisation de ces connaissances dans une perspective pédagogique permettra leur exploitation pour la production de services tutoriels.

1.4 Objectifs de recherche

Les problématiques présentées montrent que le processus d'acquisition de l'expertise perceptivo-décisionnelle dans le domaine médical, les techniques pour l'extraction puis la modélisation de ces connaissances et les mécanismes pour un transfert d'un expert à un novice constituent une démarche complexe qui répond à des exigences variées.

Notre objectif n'est donc pas de répondre à toutes ces problématiques. Dans une perspective à haut niveau, en nous basant sur les hypothèses précédemment énoncées nous visons spécifiquement trois objectifs qui sont présentés ci-dessous.

Objectif 1 : Étude exploratoire et comparative des mouvements oculaires de cliniciens novices et experts

Réaliser une étude exploratoire et comparative des aptitudes perceptivo-décisionnelles de cliniciens novices et experts en situation authentique en salle de simulation. Nous clarifions ici que fondamentalement, l'expertise est un continuum fondé sur les connaissances, les compétences et l'expérience plutôt que sur des catégories prédéfinies. Sur la base de cette clarification et en considérant que nous visons ultimement à offrir des services tutoriels aux cliniciens résidents considérés ici comme les novices, nous choisirons comme experts les membres du personnel hospitalier avec beaucoup d'années d'expérience (minimum huit). Ces deux extrémités du continuum d'expertise permettent d'une part de mieux faire ressortir les différences et d'autre part de la validation des connaissances extraites. Cette étude exploratoire se fera en prenant en considération la perception visuelle grâce à l'analyse des mouvements oculaires durant une activité authentique impliquant le RC. Il s'agira d'un traitement hors ligne (*offline*) de données collectées durant les expérimentations. Les extraits de ce premier objectif nous permettront d'une part de dégager des régularités caractéristiques du novice et de

l'expert et d'autre part d'extraire des connaissances du domaine contenant certaines heuristiques, séquences d'actions caractéristiques de meilleures performances, etc. Cette étude touche un cas spécifique et donc est limitée par rapport aux conclusions générales qui peuvent être tirées à d'autres domaines. Il s'agit d'une étude de cas au sein duquel nous cherchons à développer une méthodologie pour une étude exploratoire et comparative de données oculaires et le développement de recommandations tutorielles pouvant s'appliquer à d'autres domaines.

Objectif 2 : Analyse des données (classification des séquences oculaires, l'extraction de motif)

La deuxième phase consistera d'une part en l'analyse des données pour une classification d'extrait de séquences oculaires et d'autre part l'extraction de séquence dans le but de dégager des régularités caractéristiques du novice et de l'expert. Des séquences comportementales extraites seront formalisées dans une perspective «propédagogique » exploitable par un STI. Le modèle des connaissances envisagé sera hybride avec un modèle explicite si possible ainsi que des espaces problèmes partiels obtenus avec des algorithmes d'apprentissage machine. L'objectif 2 se justifie par le fait que la qualité des services tutoriels d'un STI est tributaire de la qualité de son modèle de connaissances du domaine.

Objectif 3 : Spécifications, conception et implémentation du STI

Il s'agira ici d'exploiter les extrants du deuxième objectif en vue d'élaborer des spécifications pour un STI qui permettrait d'offrir des services tutoriels à un novice afin de structurer graduellement ses connaissances et aptitudes comportementales comme un expert. Les spécifications pour les services tutoriels se feront à travers de nouvelles perspectives spécifiques aux domaines mal-définis. Le STI sera par la suite

utilisé par les cliniciens novices sous l'accompagnement des instructeurs du CAAHC et permettra entre autres :

- L'analyse du processus décisionnel du clinicien novice en fonction des données issues de la cartographie de sa perception visuelle.
- La conception d'outils automatisés visant à aider l'instructeur lors du débriefing post-simulation. Ces outils permettront d'effectuer des validations ainsi que des tests de conformité sur la performance globale réalisée par l'apprenant. Au besoin, une série de mesures de remédiation seront recommandées.
- La proposition de recommandations pour améliorer la sensibilité situationnelle du clinicien.
- L'extraction des règles comportementales pouvant être intégrées dans un agent tutoriel artificiel qui permettront la prévention proactive de situations de difficulté d'apprentissage en relation avec le parcours oculaire de l'apprenant.

1.5 Organisation du Manuscrit

Ce document est organisé en quatre parties (en plus d'une introduction et d'une conclusion) en lien avec les objectifs. Chaque partie contient un ou plusieurs chapitres. À la fin de chaque chapitre nous ferons un sommaire discussion et positionnement. L'introduction nous a permis de décrire une mise en contexte générale de la recherche en détaillant les problématiques, les hypothèses, les objectifs ainsi que la méthodologie qui sera adoptée.

PARTIE 1 - ETAT DE L'ART

Cette partie servira à aborder plusieurs concepts fondamentaux ainsi qu'un état de l'art critique sur des éléments qui vont contribuer à la perspective intégratrice visée par le

projet de recherche. Il s'agit notamment : des STI, des domaines mal-définis, un état de l'art sur les approches en apprentissage machine principalement les réseaux de neurones artificiels, un état de l'art sur les algorithmes de prédiction et de classification des séquences, etc. Cette partie comporte 3 chapitres traitant du cadre théorique (Chapitre 2), de l'apprentissage machine et profond (Chapitre 3) et de l'extraction et prédiction de séquences (Chapitre 4).

PARTIE 2 - ÉTUDE EXPLORATOIRE ET COMPARATIVE DE LA PERCEPTION VISUELLE DE CLINICIENS NOVICES ET EXPERTS.

Cette partie permettra de présenter le protocole expérimental, d'exposer le déroulement des expérimentations ainsi que la description des approches et outils pour la collecte, la validation et l'analyse des données. Elle comporte 2 chapitres dont un, sur la méthodologie et le cadre expérimental (Chapitre 5) et l'autre sur les analyses et résultats obtenus (Chapitre 6).

PARTIE 3 - ANALYSE DES DONNÉES : CLASSIFICATION DES SÉQUENCES, EXTRACTION DE SÉQUENCES.

Cette partie permettra de présenter les approches d'analyses dont les résultats serviront à l'acquisition automatique des connaissances. Elle sera basée sur des architectures d'apprentissage machine pour la classification des séquences comportementales en utilisant uniquement les données des mouvements oculaires. Par la suite, d'autres techniques d'apprentissage machine seront utilisées pour l'extraction de motifs dans le but de dégager des régularités caractéristiques du novice et de l'expert. L'extraction de motifs fera effectuée en considérant les différenciations dans les phases de résolution des séquences suivies d'une contextualisation. Les résultats obtenus serviront alors à émettre des hypothèses et poser les fondements pour les modules de l'architecture du STI. La partie comporte 2 chapitres. Le premier sur la classification des séquences (Chapitre 7) et le second sur l'extraction de motifs comportementaux (Chapitre 8).

PARTIE 4 – SPÉCIFICATIONS, CONCEPTION ET IMPLÉMENTATION DU SYSTÈME TUTORIEL INTELLIGENT.

Cette partie permettra de présenter les spécifications, la conception et l'implémentation du STI nommé *EyeLab*. Par la suite, un cadre d'expérimentations du STI EyeLab sera présenté. Tout cela se fera en deux chapitres : le Chapitre 9 sur la spécification et fonctionnement du STI EyeLab, suivi par le chapitre 10 sur la conception et l'implémentation du STI. Le Chapitre 9 servira à la description d'un protocole pour l'expérimentation du STI EyeLab dans le cadre d'une étude in-situ qui ne sera pas menée dans le cadre de cette thèse compte tenu de son ampleur. Quelques éléments d'une évaluation externe seront décrits pour faire foi d'une validation préliminaire.

Conclusion : Contributions de la thèse, perspectives et travaux futurs.

Elle propose un sommaire des recherches effectuées, des résultats obtenus ainsi que des conclusions pertinentes. Elle met en évidence le caractère bidisciplinaire informatique/sciences cognitives des contributions. Nous y décrivons également les limites des solutions proposées ainsi que les perspectives envisageables pour la suite de la recherche. Les travaux futurs y seront présentés dans une optique de validation longitudinale du STI, et d'évaluation de la généricité des solutions proposées par leurs applications dans d'autres domaines.

PARTIE 1 - ETAT DE L'ART

CHAPITRE II

LES ÉLÉMENTS DU CADRE THÉORIQUE

Dans ce chapitre nous allons présenter plusieurs concepts fondamentaux ainsi qu'un état de l'art sur des éléments qui vont contribuer à la perspective intégratrice visée par le projet de recherche. Il s'agit notamment : des systèmes tutoriels intelligents, des domaines mal-définis et des étapes d'acquisition de l'expertise.

2.1 Les Systèmes Tutoriels Intelligents (STI)

2.1.1 Définitions et objectifs

L'Association américaine pour l'Avancement de l'Intelligence Artificielle (AAAI) définit les STI comme « des logiciels éducatifs contenant des composants d'intelligence artificielle. Ces logiciels tracent le travail des apprenants pour pouvoir leur fournir des rétroactions et un guidage adapté. La collecte d'information relative aux performances d'un apprenant permet aux logiciels éducatifs d'inférer ses points forts et ses points faibles et de lui suggérer des travaux additionnels ». (AAAI, AI Topics/Intelligent Tutoring Systems). Un STI est une plateforme logicielle qui vise à offrir des services tutoriels adaptatifs ainsi qu'une assistance hautement personnalisée aux apprenants.

Pour atteindre cet objectif, le STI doit pouvoir inférer le plus précisément possible le profil (cognitif, affectif, etc.) de l'apprenant à partir de ses performances ainsi que d'autres indicateurs comme son comportement en situation d'apprentissage. Par la suite, le STI planifie et exécute un ensemble d'actions pédagogiques ainsi que leur ordonnancement temporel et/ou hiérarchique. Des métarègles assurent le dynamisme et l'adaptation avec des règles de réorientation dans des contextes pédagogiques définis.

2.2 Les modules de l'architecture générique d'un STI

La figure 2 présente les modules de l'architecture générique d'un STI, les interactions ainsi que les données échangées (Wenger 1987, Nwana 1990).

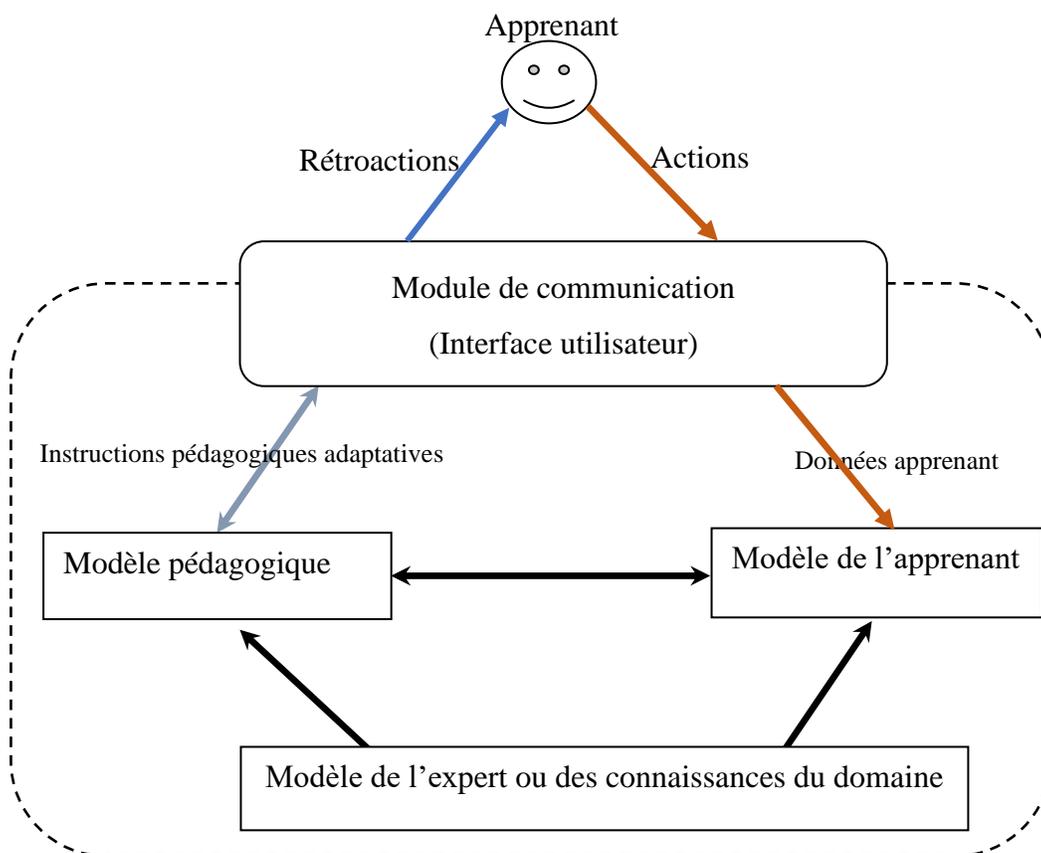


Figure 3 : Architecture générale d'un STI (adaptée de Nwana, 1990)

L'apprenant interagit avec le « module de communication » qui récupère sa performance, son comportement et d'autres indicateurs pertinents dans la situation d'apprentissage. Ses données sont transmises au mécanisme interne du STI qui est constitué de trois modèles en interaction. Finalement, les rétroactions provenant du traitement et de l'interaction entre les modèles internes du STI sont transmises à l'apprenant via le module de communication.

2.2.1 Module de communication

Le module de communication constitue l'interface externe du STI servant aux interactions avec l'apprenant selon les services tutoriels offerts. Il permet de :

1. Récupérer les actions de l'apprenant et les envoyer dans le STI. Ces actions peuvent être de nature épistémique, comportementale, gestuelle (Lallé et Luengo 2011, Aguirre, Lozano-Rodero et al. 2014), etc. Les interfaces permettant la récupération des actions peuvent être très évoluées dépendamment du domaine ainsi que des services tutoriels visés. Ainsi, certains STI intègrent des fonctionnalités d'interactions en langue naturelle (Burton et Brown 1986, Evens, Chang et al. 1997, D'mello et Graesser 2007) et d'autres, des périphériques haptiques destinés aux habiletés physico-motrices (Lallé et Luengo 2011).
2. Envoyer les rétroactions du STI à l'apprenant. Ces rétroactions résultent de l'interaction entre les composants internes du STI en prenant en considération plusieurs paramètres qui évoluent selon le modèle pédagogique adopté.

Le module de communication sert donc à établir un cadre formel du dialogue tutoriel avec l'apprenant avec une assistance proactive tout en stimulant la métacognition.

2.2.2 Modèle de l'apprenant

Le modèle de l'apprenant renferme principalement :

1. Des données de diverses natures sur la performance, le cheminement et les traces de résolution, des données de natures comportementales (latence, temps de réponse) , perceptuelles et affectives.
2. Une représentation déduite par le STI du profil global de l'apprenant.

Le modèle de l'apprenant ne se limite pas qu'aux connaissances et pourrait également contenir des données comportementales et perceptuelles permettant d'inférer l'état affectif (Frasson et Chalfoun 2010). Le modèle de l'apprenant peut donc se concevoir au niveau :

- Épistémique : les données recueillies servent à inférer l'état des connaissances (déclaratives, procédurales, empiriques, etc.) de l'apprenant.
- Comportemental : les données visées servent à inférer la métacognition, à savoir la capacité de l'apprenant à autoréguler son activité d'apprentissage (Conati et Merten 2007) ou d'autres dimensions comportementales liées à l'activité d'apprentissage.

Le contenu des composantes du modèle de l'apprenant varie selon le niveau de granularité de représentation de l'apprenant dans le STI. Il existe plusieurs stratégies de représentation du modèle de l'apprenant. Nous présentons ici les modèles du type recouvrement (Carr et Goldstein 2004) car ils sont les plus couramment utilisés. Dans cette approche, le modèle de l'apprenant est conceptualisé comme un sous-ensemble des connaissances du domaine ou de l'expert. La figure 3.1 représente une illustration du modèle recouvrement. Sur la figure a, toute la connaissance de l'apprenant est un sous-ensemble de la connaissance globale ou d'un expert dans ce domaine. Une modification de ce modèle pour tenir des connaissances erronées que peut avoir l'apprenant est représentée à la figure b. La formulation du contenu du sous-ensemble peut se faire en termes du niveau d'acquisition (probabilités, niveaux discrets, etc.) par rapport à chaque connaissance.

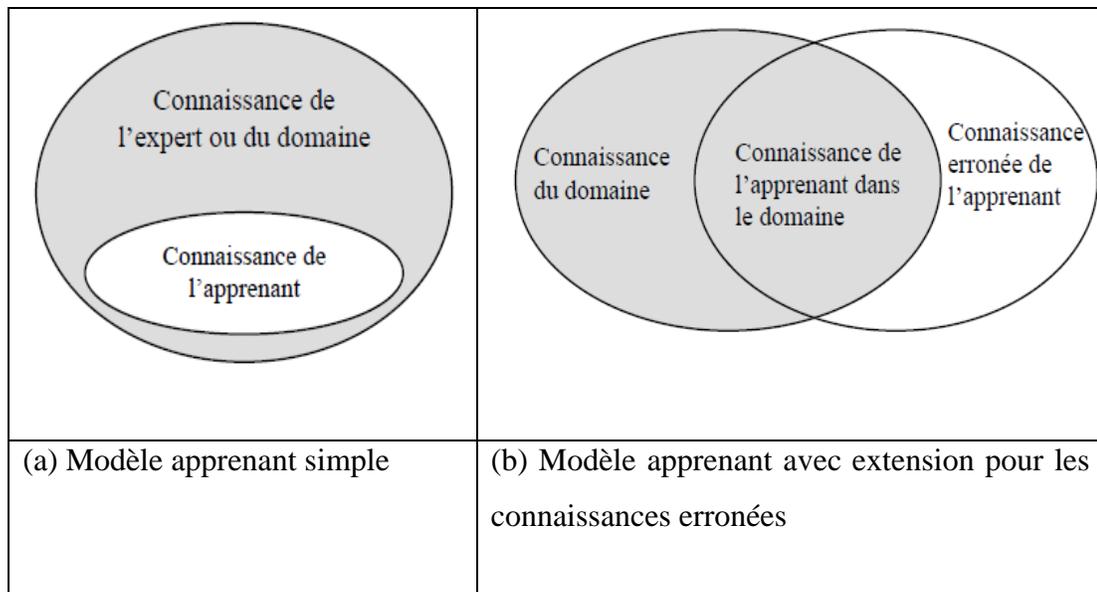


Figure 4 : Représentation du modèle du type recouvrement, avec une variante incorporant les connaissances erronées de l'apprenant

L'utilisation des modèles du type recouvrement nécessite donc d'avoir un modèle bien formalisé des connaissances du domaine ainsi que des techniques pour évaluer et diagnostiquer les connaissances (acquises ou non acquises, validées ou échouées, correctes ou incorrectes, etc.) de l'apprenant au regard des connaissances du domaine. De nombreux techniques sont utilisées pour calculer l'état des connaissances de l'apprenant que nous pouvons regrouper en deux grandes catégories :

1. Les approches déterministes

Ces approches supposent que le domaine peut être modélisé de manière explicite et exhaustive. Les éléments de connaissances (« *knowledge components* ») nécessaires à la réalisation d'une tâche peuvent par conséquent être identifiés et évalués de manière certaine sur une échelle de mesure. Elles utilisent un système à base de règles pour la validation et toutes traces de

résolution non conformes selon une grille d'évaluation est considérée comme incorrecte et les connaissances afférentes, non acquises (Mitrovic, Koedinger et al. 2003).

2. Les approches probabilistes

Ces approches permettent de prendre en compte l'incertitude dans le diagnostic des connaissances de l'apprenant. Elles sont appliquées par le biais de l'exploitation de divers paradigmes de l'intelligence artificielle tels que, la logique floue, la régression logistique, la reconnaissance de plan (« plan recognition ») et les réseaux bayésiens (Woolf, 2010; Tato et al, 2017).

Finalement, dans un but pédagogique visant à encourager l'activité métacognitive de l'apprenant, le concept du Modèle Ouvert de l'Apprenant (Open Learner Model) lui permet de consulter son propre modèle tel qu'il est produit par le STI (Bull et Kay 2010). Nous discuterons également de l'exploitation de la métacognition dans une perspective techno-pédagogique dans les STI.

2.2.3 Modèle de l'expert/des connaissances du domaine

C'est une représentation des connaissances et/ou des compétences d'un domaine pouvant être transmises par des méthodes didactiques et pédagogiques (Nwana 1990) ainsi que des méthodes reliées à leur usage en situation de résolution de problèmes. Il fournit au modèle pédagogique du STI une base de référence pour inférer l'état des connaissances de l'apprenant. La conception de ce modèle repose donc d'une part sur des techniques permettant l'acquisition des connaissances (du domaine ou de l'expert) et d'autre part sur la modélisation propédeutique c'est-à-dire sous forme de faits, principes ou règles structurés permettant une acquisition via un processus d'apprentissage. Le processus d'acquisition des connaissances soulève des

problématiques quant à la qualité, la fiabilité et l'exhaustivité des connaissances extraites. Il peut être relativement facile si les connaissances impliquées sont principalement déclaratives, mais devient complexe si les connaissances sont empiriques et/ou procédurales et dans les domaines qualifiés de mal-définis. Dans ce cas, plusieurs techniques peuvent être utilisées pour l'acquisition des connaissances à savoir : l'analyse cognitive des tâches, l'acquisition (l'extraction) automatique des connaissances, l'élaboration des contraintes du domaine, la fouille de données éducationnelles, etc. Nous décrivons trois de ces approches dans les prochaines sections.

L'analyse cognitive des tâches (*CTA : Cognitive Task Analysis*)

Cette technique vise à obtenir une modélisation formelle des connaissances déclaratives et procédurales ainsi que des étapes dans la réalisation des tâches relatives au domaine. Pour ce faire, il faudrait comprendre la structure des tâches ainsi que leurs propriétés, identifier le processus cognitif nécessaire à leurs réalisations, le niveau de compétence requis, etc. Ceci nécessite donc une collaboration entre un expert du domaine pour l'élicitation des connaissances et un cognicien pour leur modélisation. L'approche « *think aloud* » est généralement utilisée et consiste en la verbalisation à voix haute par l'expert des connaissances et heuristiques mises en jeu tout au long du processus de résolution. L'inconvénient de ce processus est qu'il peut s'avérer très coûteux (temps et effort) voir impossible dans certains cas. Par exemple, il n'est pas toujours facile pour l'expert de verbaliser ses stratégies de résolution en temps réel et acquises de manière tacite par l'expérience. Cette difficulté est connue sous le nom de « *expert blind spot* » (Nathan, Koedinger et al. 2001). Des propositions de solutions ont été faites pour simplifier l'étape de verbalisation. Il s'agit principalement des calculs d'inférence qui permettent d'une part l'enrichissement et la désambiguïsation des informations fournies par l'expert et d'autre part la construction d'un réseau de dépendance à partir de ces informations (Marcus et McDermott 1989). Par la suite, il

s'agit d'extraire de ce réseau un ensemble de règles de production pour la réalisation d'un ensemble de tâches données.

L'acquisition automatique des connaissances du domaine

Le principe de cette technique est de se baser sur des cas clés résolus par les experts en vue d'extraire des connaissances au lieu de chercher à réaliser une formalisation exhaustive des connaissances et expertise du domaine. L'une des théories les plus utilisées pour supporter les travaux dans ce sens est la théorie de (Anderson 1993) formalisée dans l'architecture cognitive nommée ACT-R (*Adaptive Control of Thought–Rational*). Elle suppose que l'apprentissage débute par un ensemble de connaissances déclaratives qui évolue graduellement vers des connaissances procédurales plus ou moins complexes grâce à la résolution de problèmes concrets. Les connaissances déclaratives sont des faits et les connaissances procédurales sont des règles de production exploitant les connaissances déclaratives pour l'accomplissement de tâches. Cependant, il est parfois nécessaire de disposer d'un ensemble préalable de connaissances du domaine qui guideront l'extraction des règles de production.

Nous signalons ici que d'autres travaux ont porté sur l'extraction automatique des connaissances conceptuelles ou déclaratives d'un domaine. Il s'agit par exemple de la construction automatique d'ontologies des connaissances du domaine à partir des textes (Zouaq et Nkambou 2008), etc.

La fouille de données éducationnelles

Il s'agit de l'utilisation des techniques d'apprentissage machine pour l'extraction automatique des connaissances du domaine (Romero et Ventura 2010) sans connaissances préalables sur le domaine. Les connaissances extraites ne sont donc pas nécessairement des règles de production spécifiques au domaine et peuvent être tout simplement des motifs fréquents, corrélations, etc. Les règles ainsi extraites peuvent servir à enrichir le modèle des connaissances du domaine ou à découvrir des contraintes

jaçentes dans le domaine (Ricucci, Carbonaro et al. 2007) . Dans cette perspective, les travaux de (Fournier-Viger, Nkambou et al. 2013) proposent une modélisation hybride qui utilise les techniques de fouilles de données éducationnelles dans le but d'extraire automatiquement des patrons de connaissances fréquents et réutilisables en combinaison avec une modélisation cognitive explicite si elle existe.

Quelle que soit l'approche adoptée, le processus de modélisation des connaissances du domaine et de l'apprenant est une tâche fastidieuse dans les STI. Dans les catégories de STI nommées Cognitive Tutors, des outils comme CTAT (Cognitive Tutor Authoring Tool) de (Koedinger, Aleven et al. 2006) ont grandement facilité le processus. Ces outils fournissent un cadre pour la conception des modèles du STI notamment l'acquisition et la représentation des connaissances du domaine, la modélisation de l'apprenant et la production de services tutoriels. Ils implémentent également les résultats des travaux de (Cen, Koedinger et al. 2006) sur l'analyse de facteurs d'apprentissage (« Learning Factors Analysis »).

2.2.4 Module pédagogique

Le module pédagogique supporte l'offre de services tutoriels (guidage didactique et pédagogique, suggestions d'aides à la résolution de problèmes, etc.) à l'apprenant. Il est également appelé le module tuteur et comprend la formalisation des approches didactiques et pédagogiques devant sous-tendre les services tutoriels proposés. Le mécanisme de fonctionnement du module pédagogique intervient dans une boucle diagnostic/remédiation qui débute généralement par un diagnostic cognitif (DC) des connaissances de l'apprenant. Le DC prend en entrée les traces de résolutions provenant du module de communication; évalue ces traces au regard des connaissances du domaine en vue de déterminer l'état des connaissances de l'apprenant. Par la suite, en se basant sur les résultats du DC, le module pédagogique planifie et exécute un ensemble d'actions pédagogiques selon leur ordonnancement temporel et/ou hiérarchique.

Les approches de représentation du processus d'apprentissage en regard des connaissances de l'apprenant se basent sur l'un des principaux paradigmes suivants: le traçage de modèle (*Model-Tracing*) et la modélisation orientée contraintes (*Constraint-Based Modeling*).

Le traçage de modèle (*Model-Tracing*)

Le traçage de modèle conçoit la progression de l'apprenant comme une acquisition de bonnes procédures de résolution de problèmes. Ces procédures sont donc structurées en étapes que l'apprenant doit graduellement maîtriser pour atteindre les compétences visées. L'évaluation de la progression se fait par un suivi (*tracing*) des étapes de résolution cognitives empruntées par l'apprenant en comparaison avec les étapes attendues provenant de la modélisation du domaine. Il devient alors possible d'une part d'identifier les étapes non réussies ainsi que les connaissances associées et d'autre part de calculer des métriques de similarités ou de divergence par rapport aux étapes attendues.

Le traçage de modèle est principalement utilisé dans les tuteurs cognitifs (*cognitive Tutors*). Le tuteur est dans ce cas le principal acteur des scénarios pédagogiques et décide du cheminement optimal et des interactions avec l'apprenant. Il s'agit d'une approche qui est fondée sur des théories behavioristes et cognitivistes et privilégie l'enseignement explicite de connaissances et procédures. L'évaluation des connaissances est soit formative dans le but d'identifier les connaissances acquises par l'apprenant et au besoin exécuter des actions de remédiation; soit sommative/certificative.

La modélisation orientée contraintes (*Constraint-Based Modeling*)

Cette approche attribue la responsabilité des apprentissages à l'apprenant et se fonde sur le principe que l'apprentissage est une appropriation progressive de connaissances par exploration et par erreurs-corrections (Ohlsson 1996). Les étapes de résolution de

l'apprenant ne sont donc pas comparées et restreintes à un modèle prédéfini d'étapes de résolution. Au contraire, ce dernier réalise l'apprentissage en fonction de ses intérêts, capacités, méthodes d'apprentissage et à son rythme. Dès lors, seules des erreurs à éviter lui sont signalées au besoin selon l'approche pédagogique adaptée. L'objectif est de permettre à l'apprenant de reconnaître ses erreurs et de les éviter progressivement. L'approche orientée contraintes offre deux avantages par rapport au traçage de modèle : La première est qu'elle permet d'éviter la tâche fastidieuse de formalisation de toutes les procédures de résolution des problèmes du domaine en se concentrant plutôt sur les contraintes devant être respectées pour la résolution [Mitrovic *et al.* 2003 ; Martin, 2001]. La seconde est qu'elle favorise l'exploration constructive par chaque apprenant de ces stratégies de résolution dans le respect de ces contraintes. Ces deux avantages font que la modélisation orientée contraintes a été utilisée dans les domaines où il est difficile de définir des stratégies fixes de résolution. Il s'agit d'une approche qui est fondée sur des théories constructivistes qui favorisent l'acquisition des connaissances basées sur les expériences et les situations authentiques de résolution de problèmes permettant à l'apprenant de découvrir lui-même l'objet de ses apprentissages et de lui attribuer ainsi une signification qui lui est propre.

L'évaluation des connaissances ne se fait pas sur des composantes isolées d'une compétence, mais se réalise sur des tâches complètes et contextualisées telles que des projets, des simulations ou des problèmes réels à résoudre.

La modélisation orientée contrôles

Ce modèle issu des travaux de Balacheff (Balacheff 2013) était initialement nommé modélisation CK ϕ . La modélisation CK ϕ se base sur deux théories.

1 - La théorie des situations didactiques de (Brousseau 1998) : considère l'apprenant dans le contexte d'un milieu d'apprentissage avec lequel il interagit. Les connaissances de l'apprenant se définissent alors comme l'équilibre entre ce dernier et les rétroactions avec le milieu.

2 - La théorie des champs conceptuels de (Vergnaud 1990) : vise à fournir un cadre cohérent ainsi que des principes pour l'étude de l'acquisition des compétences complexes. L'acquisition des concepts se fait par un processus d'élaboration pragmatique résultant des situations et problèmes (théoriques ou pratiques) rencontrés et des expériences acquises lors de leurs résolutions. Elles permettent ainsi d'identifier les invariants; de supporter le raisonnement et les inférences et d'aider à l'anticipation des effets des actions ainsi que leur planification et contrôle. Finalement, Vergnaud suppose que les concepts ne sont pas isolés, mais qu'ils coexistent dans un réseau de concepts, auquel il donne le nom de champ conceptuel.

Dans la nomenclature du modèle CK ϕ , le terme : *conception* désigne l'état des connaissances de l'apprenant et la situation dans laquelle cet état est considéré. Formellement, une conception C est un quadruplet (P, R, L, Σ) tel que :

- P, un ensemble de problèmes représentant le domaine de validité de C. En d'autres termes il s'agit de l'ensemble des problèmes que C permet de résoudre
- R, un ensemble d'opérateurs ou actions pouvant agir sur les éléments de P. Un opérateur peut être une règle de production ou un comportement par la visualisation.
- L, un système de représentation ou d'expression des éléments des ensembles P et R
- Σ , une structure de contrôles. Les contrôles représentent des éléments de décision. Ces éléments de décision peuvent être des faits, des contraintes ou encore des connaissances théoriques qui permettent de vérifier la validité de l'emploi d'un opérateur (d'une action) dans le contexte de la conception ou de confirmer l'état résolu ou non résolu d'un problème. En d'autres mots, les éléments permettent d'assurer la cohérence de C.

Un problème p de P est résolu si et seulement si, il existe r de R et s de Σ tel que $s(r(p)) = \text{vrai}$.

La modélisation CK ϕ a été par la suite adaptée pour l'exploitation dans les STI par (Luengo et al. 2006) et nommée modélisation orientée contrôles. La modélisation orientée contrôles considère que l'apprenant et l'environnement dans lequel il opère pour l'acquisition de la connaissance constituent une entité unique, un système qui devrait être pris dans son ensemble pour toute considération cognitive. Le cadre théorique se rapproche plus du paradigme de la cognition incarnée ce qui fait qu'elle est fondamentalement différente du traçage de modèle et de la modélisation orientée contraintes qui sont basés uniquement sur l'évaluation cognitive. Dans les STI la modélisation orientée contrôles a été principalement utilisée pour formaliser les connaissances déclaratives et procédurales dans des domaines nécessitant une interaction sensori-motrice complexe avec l'environnement.

Au-delà des paradigmes fondamentaux présentés, les théories d'apprentissage et d'instruction sont généralement mises à contribution dans le modèle pédagogique afin de guider l'intervention du tuteur. Dans cette perspective, certains travaux ont visé la formalisation de telles théories afin qu'elles soient traitables par la machine. L'ontologie OMNIBUS (Mizoguchi et Bourdeau 2015) est un des résultats les plus significatifs dans ce contexte. OMNIBUS offre une formalisation explicite d'un ensemble de théories d'apprentissage afin de les rendre utilisables non seulement pour des fins de scénarisation pédagogique, mais aussi afin de permettre d'y inférer des principes pouvant orienter un processus d'apprentissage selon une pédagogie particulière.

En résumé le mécanisme de fonctionnement d'un STI est basé sur un mécanisme interne qui est constitué de trois modèles en interactions : modèle de l'apprenant, modèle de l'expert/des connaissances du domaine, modèle pédagogique; ainsi qu'un module de communication pour les interactions avec l'apprenant. Le diagnostic cognitif des connaissances de l'apprenant est un composant essentiel du mécanisme de

fonctionnement du STI, car il permet d'évaluer les connaissances de l'apprenant au regard des connaissances du domaine ce qui est par la suite exploité pour supporter les décisions pédagogiques. Parmi entre autres approches pédagogiques dans les STI, nous avons les suggestions d'aides à la résolution de problèmes pour susciter la métacognition. Dans les deux prochaines sections, nous présentons le diagnostic cognitif des connaissances ainsi que la métacognition pour leur exploitation dans les STI.

2.3 Le diagnostic cognitif des connaissances dans les STI

Les STI effectuent un diagnostic cognitif (DC) qui est exploité pour offrir une assistance hautement personnalisée à l'apprenant dans un domaine d'apprentissage donné. Ceci justifie la nécessité d'une réflexion sur les aspects fondamentaux du DC ainsi que d'une approche holistique et formelle de son exploitation pour offrir les services tutoriels. Cette formalisation impose de clarifier les fondements, les processus et la dynamique interne des constituants et des étapes du DC.

Par son étymologie, le terme "diagnostic" signifie "connaissance à travers les signes". Ainsi, dans un STI, la forme la plus primitive du diagnostic cognitif consiste selon (Wenger 1987), à reconnaître, caractériser, classer et juger les actions observables de l'apprenant. Le diagnostic cognitif peut donc être interprété comme la connaissance de l'état cognitif/affectif de l'apprenant à travers les signes (actions, réponses aux questions, réactions lors d'une activité d'apprentissage, etc.) qu'il présente. Wenger propose trois perspectives de DC:

1- Le diagnostic épistémique : traite de la connaissance des apprenants dans le domaine d'étude ainsi que leurs compétences stratégiques et métacognitives. Il consiste à inférer l'état des connaissances de l'apprenant qui pourrait expliquer sa performance lors de la résolution d'une question ou d'un problème. En prenant comme exemple le diagnostic médical qui fait l'association entre les symptômes et les causes; dans le

contexte d'apprentissage, il fait associer le comportement observable de l'apprenant lors de la résolution d'une question ou d'un problème à l'état cognitif qui expliquerait le mieux ce résultat.

2 - Le diagnostic comportemental : se réfère au comportement des apprenants lors d'une activité de résolution de problème.

3 - Le diagnostic individuel : porte sur des questions telles que la personnalité, les préférences de l'apprenant, les attitudes d'apprentissage et ainsi de suite.

Sur la base de ces perspectives du DC, nous déduisons que dans un STI le DC doit permettre de faire des inférences sur l'état cognitif de l'apprenant, l'interprétation de ses erreurs, la catégorisation de son comportement et la reconnaissance de ces spécificités. Pour d'atteindre ces objectifs dans un STI, (Tchetagni, 2005) propose un processus générique en quatre étapes pouvant être exécutées de façon cyclique:

- **Étape 1:** L'acquisition des indices. Cette acquisition doit être contrôlée par les prémisses qui ont été préalablement fixées sur le genre d'inférence voulu. Ceci permettra de recueillir les données qui sont pertinentes selon les objectifs visés.
- **Étape 2:** La génération d'hypothèses par inférence et interprétation des indices acquis à l'étape 1. Elle consiste à faire une analyse des indices permettant de soutenir les éventuelles hypothèses qui expliqueraient des observations cohérentes considérées comme des indices.
- **Étape 3:** Le diagnostic proprement dit est effectué à cette étape. Il consiste à établir par le biais de tests adéquats quelles sont les hypothèses de l'étape 2 qui sont valides. Cela permet de déterminer les hypothèses qui seront effectivement confirmées comme résultat final du diagnostic.
- **Étape 4:** L'ajustement des résultats du diagnostic afin d'être utilisé pour orienter les stratégies des prochaines acquisitions d'indices.

La dimension procédurale de DC est une exécution des quatre étapes présentées comprenant l'observation, l'interprétation, l'inférence d'hypothèses relatives à l'état cognitif de l'apprenant. Le processus se termine par les ajustements/validation dont les extrants permettent de recommencer le processus.

Nous présentons dans la figure 5 le mécanisme d'implémentation du processus de DC dans un STI. Il débute par une observation et évaluation de l'apprenant dans une situation de résolution de problèmes. Dans un deuxième temps, en fonction des résultats de l'évaluation, le STI déclenche le processus d'inférence qui va générer des hypothèses de diagnostic. Deux principales sources de données sont utilisées pour faire l'inférence à savoir un modèle de l'apprenant et un modèle de tâche qui est une structure qui définit la relation entre le problème et les éléments du domaine de connaissance. La forme et le contenu de ces modèles doivent dépendre de la nature conceptuelle du processus de DC qui peut être épistémique, comportemental ou individuel. Après chaque inférence, le STI peut orienter ses interactions avec l'apprenant pour déclencher la réflexion sur les éléments de connaissance ou les compétences correspondant à l'hypothèse diagnostiquée (la réflexion de l'apprenant). Le STI peut alors utiliser le résultat de la réflexion de l'apprenant afin de confirmer, infirmer ou affiner les hypothèses diagnostiquées.

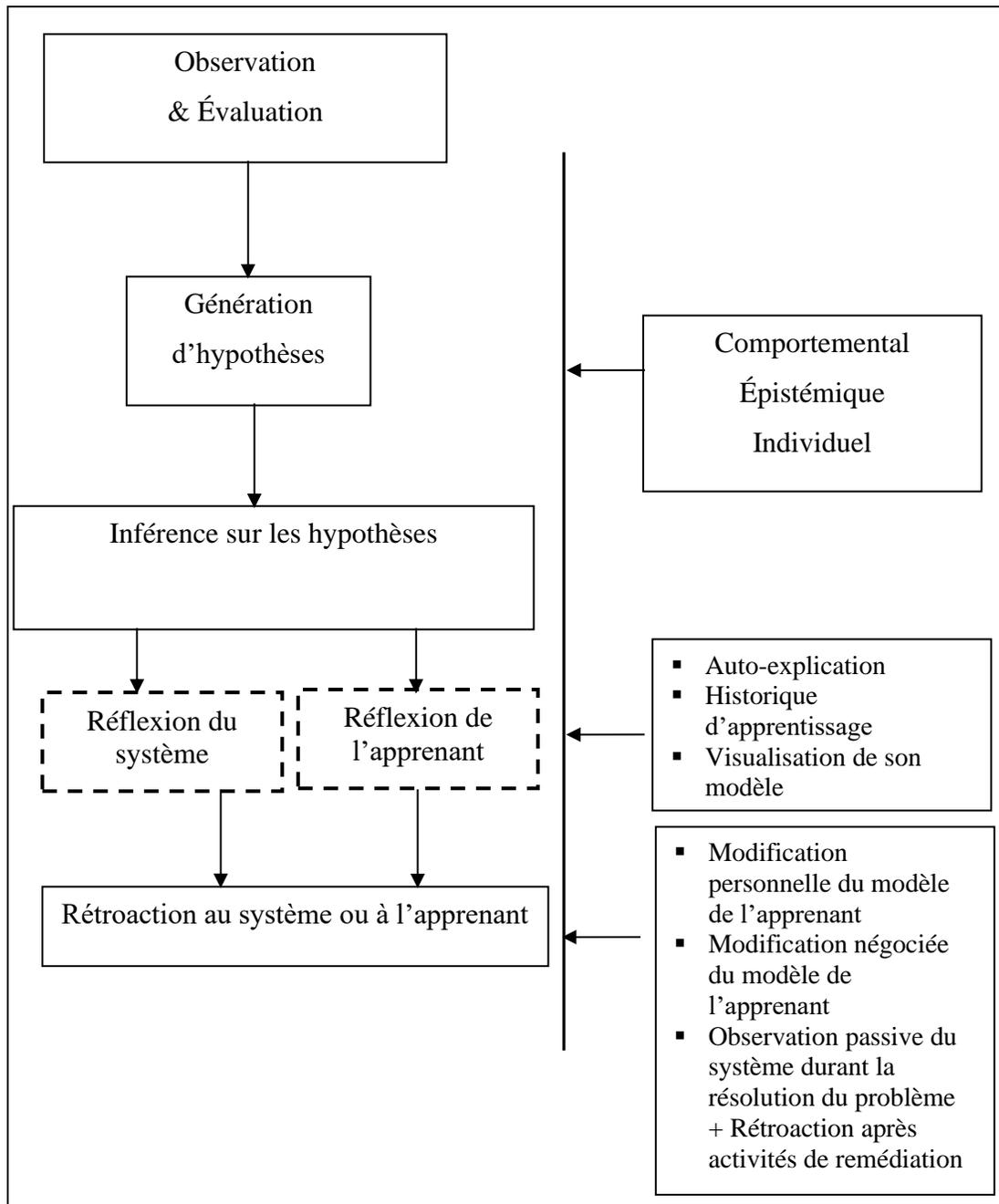


Figure 5 : Intégration de la réflexion dans un cadre de base pour un diagnostic cognitif adapté de (Tchétagni, Nkambou et al. 2006)

Les recherches sur le DC ont été dans un premier temps principalement axées sur la perspective fonctionnelle du DC avec l'application des techniques d'IA qui permettent

au système de déduire avec précision le niveau des connaissances actuel de l'apprenant (Anderson, Corbett, Koedinger, & Pelletier, 1995). Par la suite, les recherches ont pris en considération l'aspect pédagogique du DC dans la conception et l'implémentation des STI (Tchétagni, Nkambou et al. 2006). Cette perspective propédeutique du DC met l'accent sur la façon d'ajuster en permanence les hypothèses du STI en ce qui concerne l'état cognitif de l'apprenant afin d'obtenir des représentations plus précises des besoins de l'apprenant. Cet objectif est tout à fait pertinent dans le contexte des STI. Il faudrait toutefois trouver le chaînon manquant pour établir de manière objective la connexion entre les actions/réactions d'un apprenant pendant un dialogue tutoriel qui témoignent d'une occurrence d'un phénomène recherché. Sans nécessairement expliquer ce phénomène, il faudrait tout au moins le décrire, le caractériser qualitativement en référence à une théorie établie.

Une des théories proposées est celle de la pensée réflexive selon (Dewey 1933). En considérant la réflexion comme un moyen fiable pour l'apprentissage effectif, cette théorie constitue une référence pour établir les fondements visant à promouvoir de manière objective une réflexion dans le dialogue tutoriel. Il ne s'agit pas ici de considérer implicitement que le dialogue tutoriel provoque une réflexion chez l'apprenant, mais de mettre en place un cadre fondé qui s'appuie sur le caractère interactif d'un tel dialogue en favorisant une réflexion explicite avec des résultats observables.

2.4 La métacognition dans les STI

Couramment utilisée dans différents contextes, la métacognition est toutefois difficile à définir de manière formelle et explicite. Prenons l'exemple d'une situation concrète comme un entretien d'embauche, dans lequel il peut être demandé à un candidat de présenter ses qualités et ses faiblesses. Une telle activité requiert de ce dernier

d'accéder aux connaissances qu'il possède sur lui-même (ou à tout le moins selon son estimation), de décider selon la situation, s'il est judicieux ou s'il convient de les modifier pour avoir plus de chances d'être embauché (si oui, de quelle manière).

Par son étymologie, le terme métacognition est constitué de deux mots :

- Méta : du grec méta – au-delà de – après – qui indique le changement – la postérité, la supériorité, le dépassement.
- Cognition : qui est le terme scientifique qui sert à désigner l'ensemble des processus mentaux qui se rapportent à la fonction de connaissance tel que la mémoire, le langage, le raisonnement, l'apprentissage, l'intelligence, la résolution de problèmes, la prise de décision, la perception ou l'attention³. Le terme cognition peut également être utilisé pour désigner non seulement les processus de traitement de l'information dits « de haut niveau » tel que le raisonnement, la mémoire, la prise de décision et les fonctions exécutives en général, mais aussi des processus plus élémentaires comme la perception (Flavell 1976) la motricité ainsi que les émotions.

La définition issue de l'étymologie ne fournit toutefois pas des fondements explicites et exploitables à des fins computationnelles. Ceci a poussé plusieurs chercheurs à proposer des définitions dont :

- La métacognition est un domaine qui regroupe : les connaissances introspectives et conscientes qu'un individu particulier a de ses propres états et processus cognitifs ; les capacités que cet individu a de délibérément contrôler et planifier ses propres processus cognitifs en vue de la réalisation d'un but ou d'un objectif déterminé (Gombert 1990).

³ <http://fr.wikipedia.org/wiki/Cognition>

- La métacognition est un processus mental dont l'objet est soit une activité cognitive, soit un ensemble d'activités cognitives que le sujet vient d'effectuer ou est en train d'effectuer, soit un produit mental de ces activités cognitives. La métacognition peut aboutir à un jugement (habituellement non exprimé) sur la qualité des activités mentales en question ou de leur produit et éventuellement à une décision de modifier l'activité cognitive, son produit ou même la situation qui l'a suscitée (Noël 1997).

Ces définitions génériques permettent de dégager certaines composantes principales de la métacognition à savoir:

1. Les connaissances métacognitives qui regroupent les connaissances déclaratives, procédurales et conditionnelles sur la cognition ainsi que les variables liées aux tâches qui influencent la cognition.
2. Les évaluations ou jugements métacognitifs : il s'agit de la conscience métacognitive et des activités métacognitives. On peut inclure ici les processus de jugements de difficulté ou de facilité d'apprentissage, du jugement de l'apprentissage effectué, du sentiment de connaître ou de la confiance en ses productions/réponses.
3. L'autorégulation/contrôle qui est un processus qui consiste à adapter le comportement suite à l'évaluation en cours de la situation par le biais de la mise en place de diverses stratégies. Les stratégies métacognitives comprennent la planification, le contrôle et l'évaluation.

Ces trois composantes de la métacognition nous permettent d'établir que la différence fondamentale entre la cognition et la métacognition vient de l'aspect conscient, intentionnel et contrôlé des processus. La métacognition touche aussi des composantes visant à cerner les composantes plus générales de ce qui est appris, visant à faciliter une éventuelle généralisation ou un transfert de ces connaissances. Elle touche alors sur la méthode de résolution utilisée.

2.5 Les domaines mal-définis : de l'IA au STI

2.5.1 Définition et mise en contexte

En se basant sur l'étude de la résolution de problèmes chez l'humain, (Allen Newell 1972) définit un domaine mal-défini comme un domaine contenant des problèmes possédant au moins une des propriétés suivantes :

- Point de départ flou. Les instructions ou informations nécessaires pour débiter la résolution du problème sont incomplètes ou vagues.
- But imprécis. Le critère qui détermine si le but est atteint est complexe et imprécis. Il peut y avoir plusieurs solutions discutables.
- Absence de stratégies pour obtenir une solution. Il n'y a pas de stratégies précises pour déterminer les prochaines actions à effectuer à chaque étape du processus de résolution.

Cette définition peut être appliquée dans le contexte des STI en considérant que résoudre un exercice offert par un STI c'est comme résoudre un problème. Toutefois, après une revue de la littérature des travaux sur les STI ainsi que de la prise de décision sous incertitude (Lynch, Ashley et al. 2006) proposent une définition spécifique pour les STI. De manière plus spécifique au STI, ils concluent que les domaines mal-définis possèdent une ou plusieurs des caractéristiques suivantes :

- Solutions multiples et discutables : les problèmes peuvent avoir plusieurs solutions discutables et aucune procédure claire pour évaluer les solutions bien que certaines puissent être préférées selon différents critères.
- Aucune théorie formelle et complète du domaine : pour déterminer les étapes de résolution des problèmes et tester leur validité.
- Structure de tâche mal définie : deux catégories sont identifiées. La première porte sur les « domaines de conception » qui offrent des tâches où l'objectif est de créer de nouveaux artefacts (par exemple, écrire une poésie ou composer de la musique). La seconde catégorie est celle des domaines analytiques où il n'y

a pas une façon certaine de bien faire l'analyse à cause de données floues, incomplètes, fausses provenant d'un environnement dynamique.

- Concepts abstraits ou indéterminés : sont problématiques quand ils doivent être appliqués dans des situations concrètes. Ceci inclut la plupart des domaines utilisant le langage naturel à cause de l'ambiguïté des mots et expressions.
- Sous-problèmes dépendants : qui sont difficiles à décomposer pour faciliter la résolution.

En complément à la définition de (Lynch, Ashley et al. 2006), (Mitrovic et Weerasinghe 2009) considèrent deux dimensions pour les domaines mal-définis : les tâches et les domaines. Ces combinaisons donnent le quadrant précédemment illustré à la figure 2.

Le domaine médical dispose de théories et connaissances bien établies et en constante évolution. C'est un domaine relativement bien défini. Cependant, de nombreuses recherches (Arocha, Wang et al. 2005, Kiesewetter, Ebersbach et al. 2016) confirment que d'autres dimensions impliquées dans l'exécution des tâches et qui vont au-delà des connaissances du domaine confèrent des caractéristiques de tâches mal-définies. En prenant l'exemple d'une tâche de raisonnement clinique, tout en disposant des mêmes connaissances théoriques qu'un expert, selon (Hagedorn 1996), il faudrait environ trois ans à un novice pour structurer son raisonnement clinique comme celui d'un expérimenté et dix ans pour atteindre le niveau d'expert.

Des STI ont été proposés comme une des solutions pour le support au cheminement novice-expert. Nous présenterons ici deux STI qui ont eu du succès dans le domaine des connaissances médicales et l'apprentissage du raisonnement clinique afin de les positionner par la suite par rapport aux objectifs du STI envisagé.

- **SlideTutor**

SlideTutor est un STI développé à l'université de Pittsburgh qui est destiné à l'apprentissage des pathologies en dermatologie (lésions mélanocytaires) ainsi que des maladies inflammatoires de la peau. Il peut être utilisé par des cliniciens résidents ou professionnels. SlideTutor présente des séquences de contenus multimédias des tissus et cellules comme visualisés sous un microscope. L'interface utilisateur offre des espaces interactifs pour permettre à l'utilisateur de décrire ce qu'il observe ainsi que ses décisions dans une session de tutorat. SlideTutor garde une trace des performances de l'apprenant afin de les exploiter à des fins d'assistance ou de recommandation spécifiques.

- **TELEOS**

TELEOS (Chieu, Luengo et al. 2010) est un simulateur dédié à la chirurgie orthopédique percutanée. Il a été conçu dans l'optique de pallier l'écart entre la formation théorique et pratique dans le métier de chirurgien. L'objectif est d'améliorer la transmission du savoir-faire perceptivo-gestuelle aux novices hors du bloc opératoire, afin qu'ils acquièrent de l'aisance face aux interventions chirurgicales. TELEOS intègre un logiciel de simulation, un bras haptique pour la production des retours haptiques et l'enregistrement des gestes chirurgicaux et un oculomètre.

Les approches présentées dans l'architecture des STI ont de nombreuses limites pour la conception des STI dans les domaines mal-définis. Ci-dessous nous présentons brièvement les limites ainsi que les solutions.

2.5.2 Limites des paradigmes classiques des STI pour les domaines mal-définis

Les approches présentées dans la section 2.2 (module générique de l'architecture des STI) ont en commun l'exploitation d'un modèle des connaissances du domaine et supposent à priori un ordonnancement dans les concepts/habilités à acquérir. Ceci les

rend efficaces et bien adaptées d'un point de vue fonctionnel à des domaines bien définis. Dans le contexte des domaines mal définis caractérisés par l'absence d'un modèle explicite des connaissances du domaine et l'absence de théories claires sur les concepts/habilités à acquérir; l'approche classique pose plusieurs problèmes.

L'approche de traçage de modèle est difficilement applicable, car l'ensemble des connaissances et stratégies de résolutions est impossible à cerner de manière exhaustive. Cela conduit à des imprécisions dans le DC de l'apprenant dans les cas où les éléments de connaissance mis en application par celui-ci ne sont pas présents dans les règles de production modélisées (Mitrovic, Koedinger et al. 2003).

Les modèles orientés contraintes ou contrôles se préoccupent non pas des phases de résolution d'un problème, mais des contraintes à respecter dans le processus. Les actions de l'apprenant sont « valides jusqu'à preuve du contraire », c'est-à-dire jusqu'à ce qu'une contrainte ou un contrôle ne soit pas respecté. De ce fait, les stratégies de résolution erronées n'ayant pas été implémentées seront considérées comme valides (Mitrovic, Koedinger et al. 2003).

Plusieurs approches de solutions ont été proposées pour faire face à ces limites des paradigmes classiques de STI dans les domaines mal-définis. Il s'agit notamment de :

- Modéliser un espace de solutions (Le et Menzel 2008)
- Générer dynamiquement un modèle de référence (Moritz et Blank 2008)
- Adapter les paradigmes classiques aux spécificités du domaine (Ogan, Wylie et al. 2006)
- Adopter une approche hybride : (Fournier-Vigier 2010).

2.6 L'expertise : définition et étapes d'acquisition

2.6.1 Définition et caractéristiques

L'expertise se définit comme un haut niveau de connaissances et de compétences. Les perspectives des chercheurs diffèrent quant à la définition ainsi que la nature intrinsèque de l'expertise. Ainsi, Ericsson et ses collègues caractérisent l'expertise par les performances supérieures dans le domaine. Quant à (Hatano et Inagaki 1986), ils distinguent deux catégories d'experts qui se caractérisent certes par leurs performances supérieures, mais se distinguent par :

1. Une maîtrise de procédures utilisées à plusieurs reprises et peaufinées par l'expérience
2. Une bonne connaissance fondamentale renforcée par l'expérience ce qui leur permettra de justifier pourquoi certaines stratégies ou procédures ne peuvent pas être optimales, déterminer les approches pour supporter l'évolution et l'adaptation en réponse aux changements dans l'environnement, etc.

Hatano et Inagaki appellent la deuxième catégorie les experts adaptatifs et ces derniers représentent l'essence de l'expertise. Quelles que soient les perspectives (Ericsson, Charness et al. 2006) dégagent certaines caractéristiques de l'expertise qui font l'unanimité à savoir :

- Requièrè une appropriation, optimisation et extrapolation par rapport aux habiletés spécifiques d'un domaine grâce à l'expérience.
- Requièrè une meilleure représentation et organisation des connaissances du domaine ce qui améliore la mémorisation et le rappel.
- Émerge d'un cadre et d'une structure d'accompagnement appropriés.
- Se manifeste par des performances supérieures et fiables dans la résolution des problèmes du domaine. Ceci s'explique par la génération de meilleures solutions par rapport à des critères et l'utilisation d'heuristiques de résolution.

- Requière une reconnaissance et crédibilité au sein d'une communauté de pratique.

Il y a toutefois certaines limites/faiblesses caractéristiques de l'expertise. Par exemple, la difficulté de transposer les connaissances expertes dans d'autres domaines, car elles sont généralement spécifiques et contextualisées et aussi l'excès de confiance que pourrait avoir un expert en ses connaissances implicites, ce qui pourrait l'amener à un jugement biaisé ou des prédictions inexactes.

Les étapes caractéristiques de l'évolution du novice à l'expert ainsi que les modes d'acquisition ont fait l'objet de nombreuses recherches au cours des dernières décennies. Sans toutefois faire l'unanimité, ce processus est d'abord analytique, puis global, où l'intuition joue un rôle important au stade expert. Le modèle d'acquisition de compétences de Dreyfus s'inscrit dans cette perspective. Ce modèle est né au début des années 80 à la suite des recherches réalisées à l'université de Berkeley par les frères Dreyfus : Hubert Dreyfus philosophe et Stuart Dreyfus mathématicien. En 1986, dans leur livre "*Mind over Machine*" (Hubert Dreyfus 1986), les deux frères exposent leur modèle d'acquisition des compétences dans lequel ils démontrent que la notion d'intelligence humaine est nettement plus complexe que la manipulation de symbole avec des règles et procédures formelles indépendamment du contexte comme proposé par le paradigme cognitiviste dominant à cette époque. Ils prennent pour preuve la variation des résultats selon qu'un même problème soit présenté dans un contexte abstrait ou dans un contexte pratique. Ils démontrent ainsi que la prise en compte du contexte est l'une des caractéristiques essentielles de l'intelligence humaine et intrinsèquement liée au processus d'acquisition d'expertise. Le modèle de Dreyfus fournit une grille de lecture de l'apprentissage et d'acquisition de nouvelles compétences. Le tableau 1 résume ces étapes d'acquisition ainsi que la description associée.

| Étapes | Description |
|------------------------------|--|
| Novice | <p>La personne a peu ou aucune expérience de situations réelles. Il lui est donc difficile de prévoir la réussite de ses actions. De plus, il a pour objectif premier d'obtenir rapidement des résultats et non de comprendre comment les obtenir. De ce fait, il est facilement perturbé par la moindre difficulté ou imprévu.</p> <p>Le novice attend qu'une personne le guide en lui fournissant des règles 'absolues', faisant abstraction du contexte, à appliquer.</p> |
| Débutant avancé | <p>La personne commence à assimiler les schémas d'action et les règles. Il commence tout juste à prendre conscience du contexte.</p> <p>Le débutant avancé reste concentré sur ce qu'il réalise et est toujours conscient de ses actes. Il n'a pas pris de recul sur ce qu'il a compris et ne peut donc pas tout faire. Il n'a pas encore les bons réflexes et reste en fonction des règles, se retranchant derrière elles au moindre problème.</p> |
| Compétent | <p>La personne commence à classer et à trier les éléments pertinents de ceux qui sont insignifiants créant ainsi des 'perspectives'. Ceci est construit sur la base de ses expériences qui sont plus nombreuses.</p> <p>Le compétent choisit une perspective face à une nouvelle situation. Cela lui permet de réduire le nombre de paramètres à prendre en compte et de faciliter ainsi sa prise de décision.</p> |
| Efficace/ Spécialiste | <p>La personne commence faire place à l'intuition pour reconnaître la situation en lieu et place de la théorie. L'analyse de la situation devient 'intuitive', alors que la prise de décision se fait toujours de manière consciente et planifiée en suivant des règles et des procédures.</p> <p>Le spécialiste essaye de comprendre le contexte général et d'en extraire une perspective d'ensemble ainsi que les éléments pertinents.</p> |
| Expert | <p>La personne ne dépend plus du raisonnement conscient pour passer de la compréhension d'une situation à la prise de décision.</p> <p>L'expert a un bagage imposant d'exemples de situations réelles. Une « reconnaissance » intuitive de la situation et se centre immédiatement sur les</p> |

| | |
|--|--|
| | aspects pertinents sans formuler d'hypothèses non productives. Il voit spontanément ce qu'il doit faire, sans avoir à raisonner de façon consciente. |
|--|--|

Tableau 1 : Étapes d'acquisition d'expertise ainsi que la description associée

Ce tableau est très intéressant, mais il soulève l'évolution du lien entre connaissances et application de celles-ci en contexte. En fait l'expert reconnaît le contexte puis applique les règles dans ce contexte. Au niveau du suivi oculaire, cela implique des différences.

Les caractéristiques des cinq niveaux issus du modèle de Dreyfus montrent que le processus d'acquisition passe graduellement d'un stade caractérisé par un comportement analytique basé sur des connaissances et procédures consciemment suivies à un stade où les notions de contexte, de vision globale et d'intuition prennent le dessus. Un tel modèle qui souligne la nécessité de l'interaction avec l'environnement dans le processus d'acquisition, de maintien et de développement de la connaissance et de l'expertise confère à la cognition incarnée toute sa pertinence comme paradigme. Ce paradigme se distingue des paradigmes de la cognition autonome comme le symbolisme et le connexionnisme en supposant que le cerveau existe dans un corps, le corps existe dans le monde, bouge, agit, se reproduit, rêve, imagine ; c'est de cette activité permanente qu'émerge le sens de son monde et des choses (Francisco Varela 1993).

2.7 Conclusion et discussions

Dans ce chapitre nous avons présenté trois éléments qui interviennent dans les objectifs de recherche de la thèse en insistant particulièrement sur les aspects qui feront l'objet de comparaison et choix:

1. Les STI : après avoir présenté l'architecture générique, nous avons passé en revue les principales approches utilisées pour l'acquisition et la modélisation des connaissances du domaine et de l'apprenant ainsi que leur exploitation dans le module pédagogique. Par la suite, nous avons présenté sommairement des notions qui interviennent dans le fonctionnement des STIs notamment le DC et la métacognition.
2. Les domaines mal-définis : une mise en contexte a permis de faire ressortir les spécificités des domaines mal-définis pour présenter sommairement par la suite les limites des paradigmes classiques des STI pour les domaines mal-définis.
3. L'expertise et ses caractéristiques : un accent particulier a été mis sur la définition, les caractéristiques et les étapes dans le processus d'acquisition d'expertise.

Pris individuellement ces trois points semblent non-reliés, mais leur interaction prend un sens au regard des objectifs de la thèse. Ces objectifs se résument en la conception d'un STI qui permettrait d'offrir des services tutoriels à un clinicien novice en vue de structurer graduellement son aptitude perceptivo-décisionnelle comme un expert pour améliorer la dimension de la perception visuelle dans la sensibilité situationnelle.

Dans une perspective générale, le STI que nous envisageons constitue une contribution dans les STI dans le domaine médical. Il ne cible toutefois pas des connaissances déclaratives et procédurales du domaine, mais plutôt les connaissances comportementales notamment le mouvement oculaire. La généralité des approches envisagées et de la plateforme du STI feront qu'elles peuvent être exploitées dans d'autres domaines dans lesquelles la perception visuelle constitue une facette à part entière des aptitudes à développer et pouvant influencer la performance. À titre d'exemple, nous pouvons citer le domaine du pilotage des avions où le balayage visuel des instruments du tableau de bord est lié aux connaissances relatives à l'état de l'appareil et aux différentes phases de vol; le domaine de certains jeux vidéos, etc.

Les propositions décrites dans ce chapitre représentent des pistes prometteuses pour la conception du STI envisagé. Elles devront toutefois être discutées en vue de relever les limites et soutenir les choix à faire dans la conception du STI. Les aspects préliminaires porteront sur l'utilisation du suivi oculaire et le modèle des connaissances comportementales que cela suppose et aussi le modèle de communication visant à mettre en valeur les contextes et les composantes pertinentes à regarder à un moment donné.

Modélisation des connaissances du domaine/de l'expert :

Le domaine dans notre cas d'étude est médical et dispose de théories et connaissances établies. Cependant, les tâches à exécuter soit la perception visuelle (PV) pour soutenir le RC présentent des caractéristiques de domaines mal-définis. Par conséquent, une approche hybride serait envisageable qui consisterait d'une part en l'acquisition automatique par fouille de données de motifs comportementaux caractéristiques des performances expertes et d'autre part une validation, structuration et contextualisation des motifs extraits basées sur les connaissances du domaine. Ceci permettra une modélisation propédeutique des connaissances extraites pour leur exploitation dans la production de services tutoriels. Cette approche hybride présente dans notre cas de nombreux avantages dont l'enrichissement progressif des connaissances et la validation par les experts.

Modélisation de l'apprenant

Elle sera conceptualisée comme un sous-ensemble des connaissances du domaine ou d'un comportement souhaité en utilisant l'expert comme référence. L'approche overlay avec une extension sur les comportements erronés de l'apprenant sera utilisée et implémentée avec une approche probabiliste qui va utiliser des mesures de similarités comportementales dans les mouvements oculaires. L'apprenant pourra consulter son

propre modèle tel qu'il est produit par le STI. Il s'agit donc d'un modèle ouvert de l'apprenant (Open Learner Model) ayant pour objectif de servir à favoriser la métacognition dans les approches tutorielles. Finalement, les interactions entre le modèle de l'apprenant et les connaissances pour le suivi pédagogique seront basées sur du model-tracing. Le traçage de modèle (model-tracing) requiert que les étapes de résolution de l'apprenant soient comparées à un modèle d'étapes de résolution. Nous ne disposons pas à priori de ces étapes comportementales de résolutions. Toutefois, ces étapes peuvent être extraites des données des experts comme des motifs de comportements souhaités à des moments clés. De plus, ce modèle comportemental n'est pas prédéfini d'autant plus qu'il sera enrichi graduellement via l'extraction de nouveaux motifs comportementaux à partir de nouvelles traces de résolution des experts. Une alternative non pertinente serait la modélisation orientée contrôles. Elle aurait été convenable si les objectifs prenaient en compte l'apprenant et l'environnement dans lequel il opère pour l'acquisition de la connaissance.

Modélisation pédagogique

Les approches didactiques et pédagogiques seront basées sur un paradigme constructiviste. En effet le STI vise à favoriser l'acquisition des comportements en utilisant des situations authentiques (simulation haute fidélité) avec des interventions ponctuelles du tuteur. Ceci permet de laisser à l'apprenant le soin de découvrir lui-même l'objet de ses apprentissages et de lui attribuer ainsi une signification qui lui est propre.

CHAPITRE III

APPRENTISSAGE MACHINE ET RÉSEAUX DE NEURONES ARTIFICIELS

Ce chapitre présente l'état de l'art sur des éléments techniques en apprentissage machine qui vont contribuer à la perspective intégratrice visée par le projet de recherche. Nous présenterons ici principalement les réseaux de neurones.

3.1 Introduction

L'intelligence artificielle (IA) selon l'un des pionniers Marvin Lee Minsky se définit comme « la construction de programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains, car elles demandent des processus mentaux de haut niveau tels que : l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique. »

Les idées fondatrices de l'IA ont commencé assez tôt grâce aux réflexions de certains chercheurs, dont Alan Turing. Dès 1950 dans son ouvrage "Computing Machinery and intelligence" il a élaboré des arguments contre les objections qui font croire que l'intelligence demeure une qualité purement humaine.

Comme discipline fondamentale et par rapport à ses perspectives et objectifs, l'IA se fonde sur des paradigmes permettant de supporter les réflexions théoriques. Ces paradigmes servent de fondements à des approches pratiques d'ingénierie dans les

sous-domaines de l'IA comme l'apprentissage machine. Influencé par des considérations au regard des paradigmes fondamentaux en IA, de nombreuses techniques d'apprentissage machine ont été développées avec un fondement connexionniste (neuro-mimétique) ou symbolique. Après une présentation globale des notions d'apprentissage machine, nous allons mettre l'accent sur les réseaux de neurones qui constituent un des modèles utilisés pour les travaux dans cette thèse. La figure 6 montre le positionnement des notions d'IA, d'apprentissage machine et d'apprentissage profond (deep learning) imbriquées les unes aux autres.

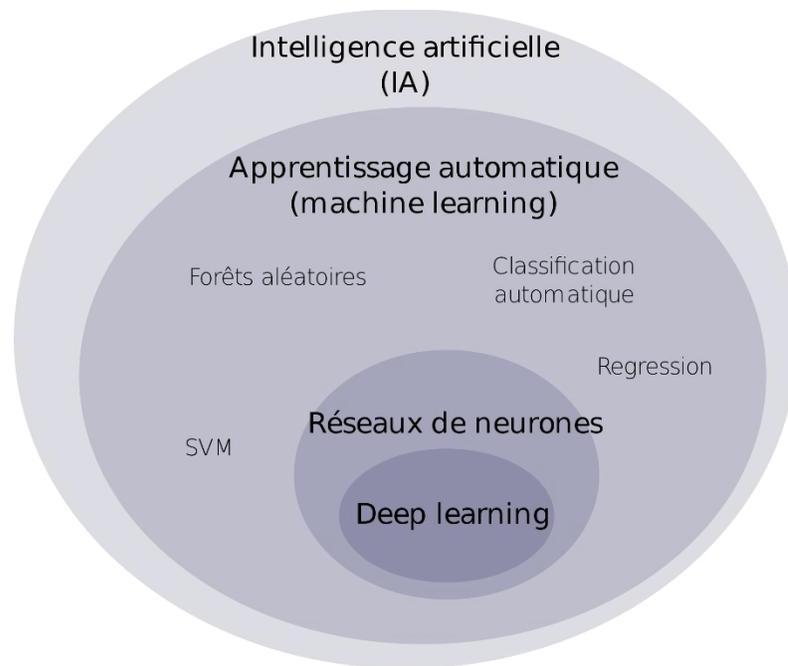


Figure 6 : Positionnement de l'IA, de l'apprentissage machine et de l'apprentissage profond (deep learning) imbriquées les unes aux autres⁴

⁴ Source : https://fr.wikipedia.org/wiki/Intelligence_artificielle#/media/Fichier:Carto_IA_deepLearning.svg

3.2 Apprentissage machine

Plusieurs définitions ont été proposées pour l'apprentissage machine. Ainsi selon,

- Arthur Samuel (1959) : c'est un domaine d'étude qui donne aux ordinateurs la possibilité d'apprendre sans être explicitement programmés (*Field of study that gives computers the ability to learn without being explicitly programmed*)
- Tom Mitchell (1998) : un programme informatique apprendrait de l'expérience E en ce qui concerne une tâche T et une mesure de performance P, si sa performance sur T, telle que mesurée par P, s'améliore avec l'expérience E (*A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E*)

Sur la base de ces définitions génériques, il ressort que l'apprentissage machine est un sous domaine de l'IA qui porte sur le développement, l'analyse et l'implémentation de méthodes automatisables qui permettent à une machine de résoudre un problème grâce à un processus d'apprentissage et ainsi d'exécuter des tâches pour lesquelles il n'est pas explicitement programmé. L'apprentissage machine se concentre donc sur l'élaboration de modèles capables de représenter certaines caractéristiques, d'apprendre certaines propriétés statistiques des distributions des données qu'ils traitent, afin d'accomplir diverses tâches. Les modèles sont construits à partir d'un ensemble d'exemples ou d'observations par un processus d'entraînement ainsi que des mécanismes permettant d'avoir une bonne généralisation. La définition de Tom Mitchell présente un lien étroit entre trois notions fondamentales en apprentissage machine : la tâche à accomplir T, qui est apprise par l'expérience E et évaluée selon une mesure de performance M.

3.2.1 La tâche T

Il s'agit de l'objectif, de l'action que la machine devra effectuer. Il existe plusieurs types de tâches regroupées principalement en trois catégories d'apprentissage avec des variantes :

- L'apprentissage supervisé : la machine apprend par exemple à classifier des entrées, prédire des valeurs en s'appuyant sur des exemples avec les cibles associées. Comme exemple il y a la classification, la régression, etc.
- L'apprentissage non supervisé : la machine apprend par exemple à extraire les informations pertinentes, regrouper, etc. en s'appuyant sur la similarité entre des exemples dont les cibles ne sont pas fournies. Comme exemple il y a la catégorisation, la synthèse et l'estimation de densité de probabilité, etc.
- L'apprentissage par renforcement : qui se base sur un cycle d'expérience/récompense et améliore les performances à chaque itération. Une analogie souvent citée est celle du cycle de dopamine : une "bonne" expérience augmente la dopamine et donc augmente la probabilité que l'agent répète l'expérience.

3.2.2 L'expérience E

Il s'agit de la source des exemples fournis pour le processus d'apprentissage. Ce sont généralement des données provenant de caractéristiques d'un objet ou sa représentation (un document, une image, un enregistrement dans une base de données, etc.) accompagnées ou non de la catégorie à laquelle il appartient. Comme exemple de données, il y a les logs d'un serveur web, le catalogue de produits d'un site de e-commerce, les transactions bancaires, les comportements des utilisateurs d'un site, une banque d'images, des extraits audio/vidéos, etc.

Bien souvent, l'expérience est entièrement décrite dans un ensemble de données statique. Toutefois dans certains cas comme l'apprentissage par renforcement, qui est une technique d'apprentissage qui expérimente avec un environnement; les actions

prises par l'agent influencent ses observations, par-là les données pour l'apprentissage et vice-versa. L'ensemble de données peut donc changer durant l'expérience et dans le cas d'une observabilité partielle par l'agent, certaines caractéristiques de l'objet ou de sa représentation ne sont pas connues.

Les données de l'expérience E , nécessitent généralement un préentraînement avant l'utilisation pour l'apprentissage de la tâche T . Ce préentraînement vise :

- Un réajustement des valeurs des attributs; par exemple par la normalisation, la standardisation et l'uniformisation, etc.
- Une réduction de la dimensionnalité
- Une réduction du bruit
- Une augmentation de données dans les cas où les données ou leurs distributions sont insuffisantes

Après le préentraînement, des techniques d'échantillonnage sont utilisées pour subdiviser les données pour faire d'une part l'apprentissage et d'autre part une validation du modèle appris. Il y existe plusieurs techniques d'échantillonnage dont:

- La séparation en données d'apprentissage (par exemple 80%) et données de test (par exemple 20%)
- La méthode k -fold : on divise l'échantillon original en k échantillons, puis on sélectionne de manière cyclique $k-1$ données pour l'apprentissage le dernier pour la validation.
- La méthode leave-one-out : cas particulier de k -fold où $k = n$ la taille de l'ensemble de données. L'apprentissage se fait avec $k=n-1$ puis la validation est faite sur la n ème observation.

L'apprentissage permet d'obtenir un modèle qui est une description formelle des relations qui existent entre l'ensemble des attributs qui décrivent les données à traiter. Le modèle obtenu peut-être sous forme de régression, de règles ou arbre de décision, de fonction discriminante, de frontières de décision ou hyperplan, etc. Le modèle obtenu est évalué selon des critères de performances donnés.

3.2.3 La mesure de performance M

Il s'agit de méthodes formelles et de métriques qui serviront à évaluer le modèle appris. Pour des tâches de classification avec des sorties discrètes, la précision et la matrice de confusion sont généralement utilisées. Pour des tâches à sortie continue, comme la régression, l'erreur quadratique moyenne est souvent utilisée. Pour une densité de probabilité, l'entropie croisée est souvent utilisée. En somme, la mesure est souvent spécifique à la tâche à accomplir et son choix est très important.

La performance du modèle est évaluée sur de nouvelles données pour mesurer la généralisation. Il s'agit de la capacité du modèle à pouvoir effectuer des prédictions robustes sur de nouvelles données. Deux situations classiques sont à éviter :

- Le surapprentissage (*overfitting*) qui désigne le fait que le modèle soit trop collé aux données d'entraînement.
- Le sousapprentissage (*underfitting*) désigne une situation où le modèle n'est pas assez élaboré pour capturer le phénomène dans son intégralité

La capacité d'un modèle l'aptitude à apprendre précisément tous les détails d'un échantillon d'apprentissage. Plus la capacité est grande, plus la différence entre l'erreur d'entraînement et l'erreur de test augmente. Ce qui a impact sur la généralisation.

Plusieurs solutions fondamentales et heuristiques existent pour obtenir un modèle qui se généralise bien. Des facteurs intrinsèques aux exemples d'entraînement peuvent également influencer significativement la généralisation. Il s'agit notamment de la similarité des exemples, de leurs difficultés (distance d'un exemple en particulier par rapport à la moyenne), de l'entropie de distribution dans les classes, de l'ordre de présentation durant la phase l'entraînement, etc. Une approche de solution est l'apprentissage par curriculum (Bengio, Louradour et al. 2009). En s'inspirant de différents contextes d'apprentissage (éducation de l'enfant, dressage animal, etc.) l'apprentissage par curriculum montre qu'augmenter graduellement la difficulté des exemples d'entraînement améliore la généralisation.

3.3 Formalisation des notions en apprentissage machine et exemple

La notation utilisée dans les chapitres portant sur l'apprentissage machine sont résumées ainsi:

- \mathcal{D} : ensemble de données, contenant les exemples d'entraînement, et possiblement des cibles. Avec les cibles, l'ensemble des données d'entraînement sera alors composé de paires (x_i, t_i) où x_i correspond à une entrée (donnée) et t_i correspond à la cible (réponse au problème posé par l'entrée, par exemple, la classe à laquelle elle appartient)
- \mathcal{D}_{train} : sous-ensemble de \mathcal{D} utilisé pour l'apprentissage
- \mathcal{D}_{valid} : sous-ensemble de \mathcal{D} distinct de \mathcal{D}_{train} , utilisé pour l'optimisation des hyperparamètres
- \mathcal{D}_{test} : sous-ensemble de \mathcal{D} distinct \mathcal{D}_{train} et de \mathcal{D}_{valid} , utilisé pour estimer l'erreur de généralisation
- $h(x_i, \theta)$: La fonction hypothèse/modèle développée sur les données d'entraînement, puis utilisée pour effectuer des prédictions sur de nouvelles données x_i
- θ : Les paramètres modifiables de la fonction h
- L : La fonction de perte qui typiquement compare la prédiction $h(x_i, \theta)$ effectuée sur l'entrée x_i à la cible désirée t_i et retourne un nombre qu'on cherche à minimiser.
- $R(\theta)$: La fonction de régularisation, qui dépend du paramètre θ
- $\mathcal{C}(\theta)$ Le critère d'entraînement, dans un cadre d'apprentissage supervisé

3.4 Les réseaux de neurones : définition, principes et historique

3.4.1 Le neurone biologique

Le cerveau se compose principalement de deux types de cellules : les neurones et les cellules gliales. Il y a environ 86 à 100 milliards de neurones avec 1000 à 10000 synapses (connexions) par neurone. Les cellules gliales forment l'environnement des neurones, procurent des nutriments aux neurones et les protègent en combattant les pathogènes. Elles éliminent également les déchets causés par la mort neuronale et accélèrent la conduction nerveuse en agissant comme gaine isolante de certains axones. Les cellules gliales grâce à une action modulatrice sur la neurotransmission sont impliquées dans le traitement de l'information nerveuse.

Le neurone est une cellule spécialisée composée d'un corps cellulaire (soma) et d'un noyau. Du neurone émanent plusieurs fibres appelées dendrites et une seule longue fibre appelée axone. Les dendrites sont parfois si nombreuses qu'ils sont nommés arborisations dendritiques. Cependant, l'axone est unique et peut s'étaler sur une distance de quelques microns à 1,50 mètre. L'information traitée par le neurone chemine le long de l'axone pour être transmise aux autres neurones formant ainsi un réseau très dense. La jonction entre deux neurones est appelée la synapse. Le nombre de synapses par neurone varie considérablement de plusieurs centaines à une dizaine de milliers. Ainsi, le neurone biologique intègre des influences externes venant d'autres neurones. Si ces influences excèdent un seuil, le neurone déclenche une impulsion de sortie nommée (*potentiel d'action*). Cette impulsion se propage le long de l'*axone* pour être transmise aux autres neurones. Si l'excitation se maintient, le neurone envoie une série d'impulsions excitatrices ou inhibitrices à une fréquence caractéristique et proportionnelle à l'amplitude de l'excitation. L'efficacité de la connexion entre les neurones peut être affectée par le temps et de nombreux autres facteurs, ce qui permet le processus d'apprentissage, de mémorisation, d'oubli, de neuroplasticité, etc.

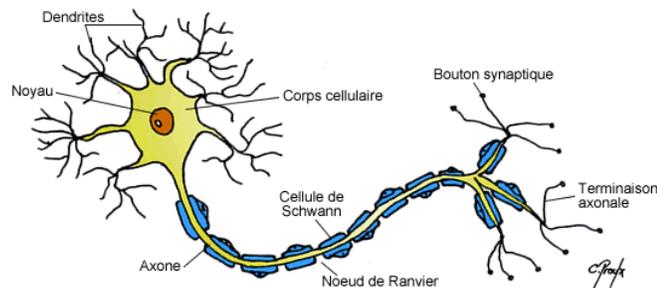


Figure 7 : Neurone avec gaine de myéline⁵

3.4.2 Du neurone biologique au neurone artificiel (neurone formel)

Le neurone artificiel est un modèle inspiré d'une analogie avec le neurone biologique tout en intégrant des principes mathématiques et des influences diverses pour les besoins en ingénierie. Le neurone artificiel intègre certains principes du neurone biologique comme :

- L'activation : le neurone reçoit l'influence d'autres neurones selon certaines règles. Par exemple, quand la somme des signaux reçus en entrée par un neurone dépasse un seuil, il émet à son tour un signal vers les neurones auxquels il est relié selon une fonction d'activation.
- Les connexions synaptiques : caractérise l'influence du neurone source de la connexion sur le neurone cible de la connexion. Un nombre réel est associé pour signifier l'intensité de cette influence.
- La capacité d'apprentissage : reliée à la modification des poids des connexions synaptiques

Ces principes sont à la base de la modélisation neuro-mimétique. Il s'agit donc d'un modèle très simplifié de neurone biologique permettant de faire des calculs simples et

⁵ <http://physique.unice.fr/sem6/2011-2012/PagesWeb/PT/Modelisation/img/myel.gif>

rapides. Comme illustré sur la figure 8, le neurone reçoit à partir des entrées x_j des connexions influencées par des poids synaptiques w_j . Le sommateur calcule la somme pondérée de ses entrées à partir du vecteur d'entrée x et du vecteur de poids w et y rajoutant au besoin un biais. La somme pondérée est par la suite passée à une fonction d'activation qui calculera la sortie du neurone. La fonction d'activation est de manière classique une fonction binaire à seuil, en marche d'escalier ou des sigmoïdes.

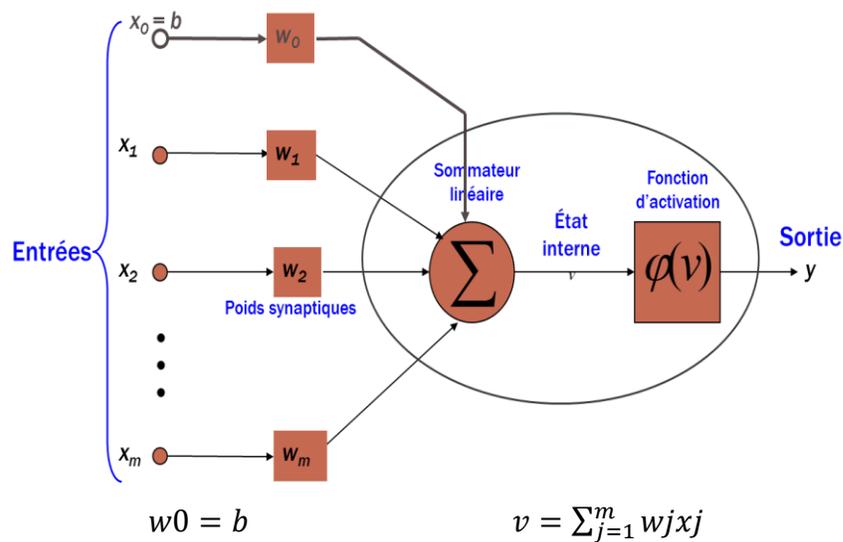


Figure 8 : Neurone artificiel

Le modèle de neurone présenté se base sur l'hypothèse de synapse synchrone, sans perte synaptique et une sortie non impulsive. Avec la sortie, il se comporte comme une fonction discriminante qui détermine si l'entrée x se trouve d'un côté ou de l'autre d'un espace linéairement séparable. Il existe de nombreuses variantes de ce modèle de neurone artificiel présenté.

Un RNA (Réseau de Neurones Artificiels) est un ensemble de neurones artificiels formels dans lequel chaque neurone est une unité de calcul simple ou un nœud processeur connecté aux autres neurones et fonctionnant selon des règles définies par son architecture et l'utilisation souhaitée. Les RNA sont entraînés via des stratégies

d'apprentissage visant à calculer les poids synaptiques. Il existe plusieurs architectures de RNA qui se distinguent par les critères comme :

- Le type de neurone : binaire, à impulsions, cellule, etc.
- La topologie : réseaux à couches, récurrent, etc.
- L'origine : neurobiologie, ingénierie, mathématique, etc.
- Le mécanisme de fonctionnement : propagation avant, convolution, etc.
- Les types de problèmes classiques pour lesquels il est utilisé: classification, catégorisation, optimisation, génération de nouveaux contenus, compression et restitution des entrées, etc.

3.5 Historique et évolution

L'histoire des RNA a été assez mouvementée, influencée par le contexte de la naissance de l'IA, des débats épistémologiques sur le connexionniste vs le symbolisme et l'évolution de la puissance de calcul des ordinateurs. Nous présentons ci-dessous quelques jalons historiques depuis le premier neurone formel jusqu'à aujourd'hui.

- 1943: Dans l'article « A logical calculus of the ideas immanent in nervous activity», Warren S. McCulloch et Walter Pitts présenteraient le premier neurone formel qui porte leurs noms : neurone de McCulloch-Pitts. Ce neurone n'avait pas de mécanisme formel d'apprentissage, mais représentait toutefois un premier jalon dans le paradigme connexionniste de l'IA. Ce paradigme connexionniste est froidement accueilli et même critiqué comme l'illustre cette citation de A. Turing, « Nous pourrions produire des modèles électriques assez précis imitant le comportement du système nerveux, mais il ne semble pas y avoir grand intérêt à le faire. Ce serait comme consacrer beaucoup d'efforts à fabriquer des voitures avançant sur des jambes au lieu de continuer à se servir de roues. »

- 1949 : Hebb Donald (Hebb, 1949) : propose une formulation du mécanisme d'apprentissage sous forme de règles de modification des connexions synaptiques appelées règle de Hebb. Ces règles modélisent le principe de la plasticité synaptique notamment le fait que si des neurones, de part et d'autre d'une synapse, sont activés de façon synchrone et répétée, la connexion synaptique est renforcée. Ceci établit les bases du principe de l'apprentissage connexionniste qui consiste à modifier le poids des connexions entre les neurones. D'autres variantes de la règle de Hebb ont été proposées plus tard par d'autres chercheurs notamment la règle de Widrow-Hoff et la règle de Grossberg. De nos jours de nombreuses règles sont disponibles dépendamment du type de réseau ainsi que des objectifs.
- 1957-1958 : Frank Rosenblatt propose le Perceptron, le premier modèle opérationnel de neurone artificiel avec la capacité d'apprendre par l'expérience pour la reconnaissance d'une configuration et la tolérance aux bruits. Le Perceptron est inspiré du système visuel pour la perception d'où le nom Perceptron. Il dispose d'une couche d'entrées pour la perception ainsi de sorties pour la décision. Étant qu'à cette époque les langages de programmation spécialisés n'étaient pas couramment utilisés, Rosenblatt a implémenté l'idée du Perceptron dans du matériel adapté; et a montré qu'il pouvait être utilisé pour apprendre à classer correctement les formes simples avec des entrées de 20x20 pixels.
- 1969 : Minsky & Papert montrent dans un livre (« Perceptrons ») toutes les limites du perceptron principalement l'impossibilité de classer des configurations non linéairement séparables tels que le célèbre problème du XOR (OU exclusif). Ceci affecte l'enthousiasme des recherches sur les RNA ainsi que l'abandon graduel des financements des recherches sur les RNA.
- 1969 - 1982 : Mise en sommeil des recherches sur les RNA. Elles continuent sous le couvert de domaines divers avec des chercheurs comme Grossberg,

Kohonen, Anderson, Hopfield, etc. Également en 1974, Paul Werbos et d'autres chercheurs proposent un algorithme de rétropropagation de l'erreur, ce qui offre un mécanisme d'apprentissage pour les réseaux multicouches de type Perceptron (MLP : Multi-layer Perceptron). Les MLP pouvaient apprendre des configurations non linéairement séparables et constituent une première réponse à Minsky et Papert.

- 1983 – 1986 : marqué par de nombreux progrès qui redonnent de l'intérêt au RNA. David Rumelhart popularise l'algorithme de rétropropagation du gradient pour l'apprentissage des MLP.
- 1990 : émergence des travaux sur les réseaux à convolution (CNNs). Yann LeCun et ses collaborateurs proposent en 1998 l'architecture LeNet pour la reconnaissance des chiffres manuscrits
- 1997: Proposition de l'architecture LSTM (Long short-term memory) par Sepp Hochreiter and Jürgen Schmidhuber.
- 2014 : Émergence des réseaux GANs (Generative adversarial networks) introduite par Ian Goodfellow. C'est une approche innovante pour l'élaboration de modèles génératifs, c'est-à-dire capables de produire eux-mêmes des données. Les GAN reposent sur l'entraînement non supervisé de deux réseaux de neurones artificiels appelés Générateur et Discriminateur. Ces deux réseaux s'entraînent l'un et l'autre dans le cadre d'une relation contradictoire. Le *Générateur* a la charge de créer des exemples (ex : des images). Le *Discriminateur* reçoit des exemples provenant du générateur et d'une source de données constituée d'exemples réels. Le *Discriminateur* identifie la source de chaque exemple et tente ainsi de deviner si ces exemples sont réels ou s'ils ont été générés par le Générateur. Les deux algorithmes entretiennent donc une relation d'amélioration continue : le Générateur apprend à créer des exemples de plus en plus réalistes et le Discriminateur apprend à identifier de mieux en mieux les exemples réels de ceux provenant du Générateur

- 2017 : Geoffrey Hinton et d'autres chercheurs présentent officiellement le Capsule Networks (CapsNet) permettant une meilleure modélisation des relations hiérarchiques grâce à l'ajout de structures appelées « capsules » à un CNN. Une capsule est un groupe de neurones dont le vecteur d'activité correspond aux paramètres d'instanciation d'un objet ou d'une partie d'objet. Elle indique les caractéristiques de l'objet et sa probabilité. Ces caractéristiques peuvent inclure des paramètres tels que la "pose" (taille, position, orientation), la vitesse, la déformation, l'albédo (réflexion lumineuse), la texture et la teinte. La longueur d'un vecteur capsule indique la probabilité qu'un élément soit présent dans l'image, et la direction du vecteur correspond à ses informations de pose. Le CapsNet utilise les sorties de plusieurs capsules pour former des représentations plus stables par rapport à diverses perturbations apportant une solution à certaines limites connues des CNN. Il s'agit par exemple de l'invariance par rapport à la translation (changement de position ou rotation de l'image) et la représentation interne des données par le CNN qui ne prend pas en compte les hiérarchies spatiales importantes entre les objets simples et complexes. Également, avec des changements au niveau de l'optimisation des poids synaptiques les CapsNet nécessitent une fraction des données que nécessitent un CNN pour l'entraînement et la généralisation.

3.6 Apprentissage dans les RNA (cas du perceptron multicouche)

Le perceptron multicouche (MLP : Multi Layer Perceptron) est un réseau de neurones composé de plusieurs couches de Perceptron : typiquement 3 couches. La couche d'entrée, une couche intermédiaire dite cachée et une couche de sortie.

- La couche d'entrée : elle est connectée aux données brutes ou prétraitées : \hat{x} . Chaque caractéristique est généralement connectée à un neurone de la couche d'entrée.

- La couche cachée : pour composer les relations entre les variables reçues de la couche d'entrée.
- La couche de sortie : retourne la prédiction de la cible \hat{t} associée à \hat{x} effectuée par le réseau.

Formellement, chaque couche du MLP est donc constituée de neurones (Perceptron) que nous allons indiquer a_j , avec $j \in 1, 2, \dots, n$, qui prennent en entrée les valeurs de sortie de la couche précédente comme présenté à la figure 10. Il n'y a pas de connexions d'un neurone vers lui-même ou entre les neurones d'une même couche.

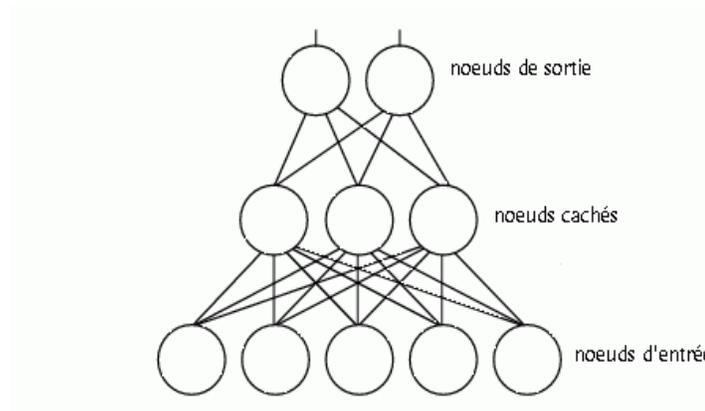


Figure 9 : Le perceptron multicouche

Phase d'apprentissage (ajustement des poids)

L'apprentissage correspond à l'ajustement des poids $w_{i,j}$ du réseau à partir des données fournies durant la phase d'entraînement pour ajuster les réponses obtenues aux réponses désirées. Ainsi, en utilisant le MLP pour une tâche de classification, nous aurons : une entrée \hat{x} , la matrice des poids actuels $w_{i,j}$ et la prédiction $h_w(\hat{x})$ pour la cible \hat{t} effectuée grâce aux valeurs $w_{i,j}$. Différents algorithmes sont utilisés pour trouver les valeurs des poids $w_{i,j}$ qui minimisent la fonction de perte durant le processus d'apprentissage. Le mécanisme est composé de deux phases :

1. Une propagation-avant (feed-forward propagation) : la fonction d'activation de chaque neurone est calculée à l'aide d'une fonction logistique appliquée à la combinaison linéaire pondérée des sorties des neurones de la couche précédente. Le calcul se fait couche par couche, de l'entrée à la sortie d'où le nom propagation-avant. Pour une entrée \hat{x} , à la fin de la propagation-avant, nous obtenons la prédiction \hat{t} . Cette prédiction peut être égale à la sortie désirée ou non. Une fonction de perte est calculée. La fonction de perte vaut 0 dans le cas d'une prédiction parfaite (aucun ajustement des poids n'est alors requis), sinon un ajustement des poids est nécessaire.
2. Une rétropropagation : le gradient de la perte est calculé par rapport à l'activation cible du neurone de sortie, puis cette valeur est rétropropagée à travers le réseau suivant le graphe de flot pour ajuster le poids des connexions.

$$Loss(\hat{t}, h_w(\hat{x})) = -(\hat{t}, h_w(\hat{x}))w \cdot \hat{x} = -(\hat{t}, h_w(\hat{x})) \sum_{i=1}^n (w_i \cdot \hat{x}_i)$$

Équation : Exemple de fonction de perte

$w \cdot \hat{x}$: correspond au produit scalaire entre le vecteur des poids w et le vecteur des données de l'entrée \hat{x} .

Les paramètres de la fonction de perte ne sont pas connus. Cependant la valeur en un point donné peut être calculée. Pour minimiser cette fonction de perte, des techniques basées sur une approche générique par la descente de gradient sont utilisées. Par la suite, une rétropropagation des gradients est faite en suivant le graphe de flot.

Méthode de descente du gradient

Le gradient d'une fonction permet de savoir s'il faut augmenter ou diminuer la valeur de chaque variable si l'on souhaite maximiser ou minimiser cette fonction. Le calcul du gradient est présenté ci-dessous :

$$\nabla f(x, y) = \left[\frac{\partial f(x, y)}{\partial x}, \frac{\partial f(x, y)}{\partial y} \right]$$

Équation : Gradient d'une fonction à deux variables x et y

Étant donné un modèle $h(x, \theta)$, avec un vecteur de paramètres θ de longueur m , la méthode de la descente du gradient peut être utilisée pour chercher à identifier le maximum de vraisemblance, noté θ^* . Elle fonctionne de façon itérative pour trouver un minimum de la fonction de perte sans avoir à calculer explicitement la valeur de θ^* .

Le vecteur des paramètres θ sera mis à jour de manière itérative comme présenté ci-dessous.

$$\theta_{i_{t+1}} = \theta_{i_t} - \lambda \frac{\partial L(\theta_{i_t})}{\partial \theta_i}$$

Cette procédure de mise à jour sera répétée un certain nombre d'itérations en recalculant le gradient de la fonction de perte pour le nouveau vecteur de paramètres, ou encore jusqu'à ce que toutes les composantes de $\nabla L(\theta')$ soient inférieures à un certain seuil très petit $\epsilon > 0$. Il s'agit donc d'une méthode d'optimisation supervisée, itérative et de premier ordre (c'est-à-dire n'utilisant pas de dérivée du second degré).

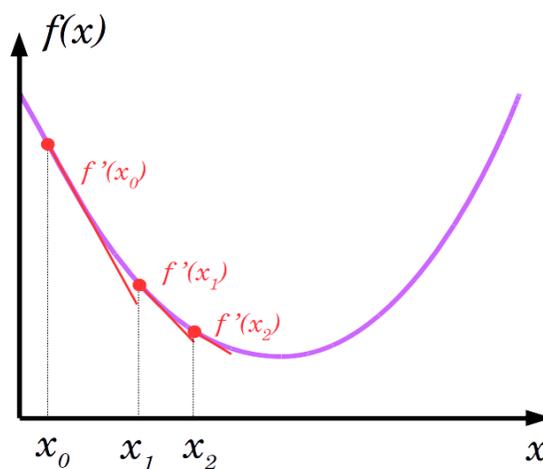


Figure 10 : Descente de Gradient

À terme, la descente du gradient vise à atteindre un minimum local pour la fonction de perte $L(\theta)$. Lorsque la fonction de perte L du problème considéré est convexe, le minimum local identifié est nécessairement global, θ^* sera donc bien identifié au bout d'un certain nombre d'itérations. Comme illustré à la figure 11, un élément déterminant pour atteindre le minimum après un nombre raisonnable d'itérations est l'hyperparamètre λ appelé taux d'apprentissage (*Learning rate*). Avec une grande valeur de λ , les pas des itérations de mis-à-jour risquent de passer un éventuel minimum intéressant et augmenter l'erreur au lieu de la diminuer. Avec une valeur trop petite, la convergence des poids va prendre très longtemps. On cherchera donc à trouver le 'meilleur' λ tout au long de l'entraînement. Il existe des heuristiques qui sont utilisées comme :

- L'heuristique de recuit (annealing), qui correspond à une décroissance de λ proportionnelle à l'inverse du nombre d'itérations λ_t , à partir d'une itération λ_0 donnée :

$$\lambda_t = \min\left(\lambda_0, \frac{\tau_0}{1 + \tau_t}\right)$$

- L'heuristique ADAGRAD, qui calcule un taux d'apprentissage scalaire $\lambda_{t,i}$ à l'itération t pour le paramètre θ_i en divisant un taux d'apprentissage de départ par la somme, depuis t_0 , de la norme 2 du gradient de la perte $L(f_{\theta_i}, z_j)$ par θ_i :

$$\lambda_{t,i} = \frac{\lambda_0}{\sqrt{\alpha_t}}$$

$$\begin{cases} \alpha^{-1} = 1, \\ \alpha_t = \alpha_{t-1} + \sum_{j=t_0}^t \left(\frac{\partial L(f_{\theta_i}, z_j)}{\partial \theta_i}\right)^2 \end{cases}$$

- Les heuristiques de sélection de variables ont souvent une influence critique sur la forme et la taille de l'arbre de recherche. Certaines heuristiques observent le principe first-fail énoncé par Haralick et Elliott : « To succeed, try first where you are most likely to fail » (Pour réussir, essaie d'abord là où il est le plus

probable d'échouer). Les heuristiques dynamiques exploitent souvent les informations sur les domaines et degrés des variables. On retrouve entre autres les heuristiques suivantes : lex, random, Min-domain, dynamic-degree, weighted-degree, etc.

Lors de la phase d'entraînement, le terme « itération » désigne l'application de l'algorithme sur une unité d'entraînement (un x_i de \mathcal{D}_{train}), et « époque » son application sur tout \mathcal{D}_{valid} . L'entraînement dure donc un nombre d'époques : epochs. Quelle que soit l'heuristique utilisée, il faudrait également adopter une stratégie de mise à jour de la valeur des paramètres.

- Si la mise à jour est faite pour chaque x_i de \mathcal{D}_{train} , on parlera de descente de gradient stochastique. Théoriquement, les données d'entraînement seront considérées indépendantes et uniformément distribuées. Pour un \mathcal{D}_{train} très grand, ce processus pourrait s'avérer très coûteux.
- Si la mise à jour est faite selon la moyenne des gradients de f sur tout \mathcal{D}_{train} , on parlera de descente de gradient par lot (batch gradient descent).
- Un compromis entre ces deux méthodes serait des mises à jour des poids qui seront calculées à partir de « mini-lots » (mini-batches). Ce compromis offre plusieurs avantages. Par rapport à la descente stochastique, il y a d'une part l'avantage computationnel et d'autre part la moyenne des gradients atténue le bruit que pourrait engendrer un exemple unique. Par rapport à la descente avec un unique lot, la fréquence plus élevée de mises à jour atténue les biais. Finalement, les mini-lots permettent de profiter de caractéristiques matérielles permettant une exécution des calculs en parallèle.

Mis à part le choix judicieux de l'hyperparamètre λ du taux d'apprentissage et d'une bonne stratégie pour les mises à jour des paramètres, des facteurs intrinsèques aux données peuvent influencer la vitesse de convergence.

3.7 Réseau de neurones récurrent (RNN)

Les réseaux de neurones récurrents ont été initialement introduits par Jeffrey Elman en 1990 dans son article « Finding structure in time ». À la différence des réseaux de neurones classiques, un RNN a une architecture qui admet des boucles. Ainsi, contrairement aux réseaux de neurones à propagation avant, où les entrées des neurones d'une couche ne proviennent que des couches précédentes, la sortie de certains neurones d'un réseau récurrent sera donnée en entrée à des neurones des couches précédentes. Les réseaux récurrents conviennent en particulier pour l'analyse de données séquentielles ou temporelles.

Un réseau de neurones récurrent possède un réseau à propagation avant équivalent. Or, ce dernier sera typiquement bien plus grand que son équivalent récurrent et inclura nécessairement des copies de plusieurs neurones. La figure 12 présente un réseau récurrent ouvert avec des entrées séquentielles dans le temps ($X(t)$, $X(t+1)$, ...), l'état interne ($u(t)$, $u(t+1)$, ...), et les sorties ($y(t)$, $y(t+1)$, ...).

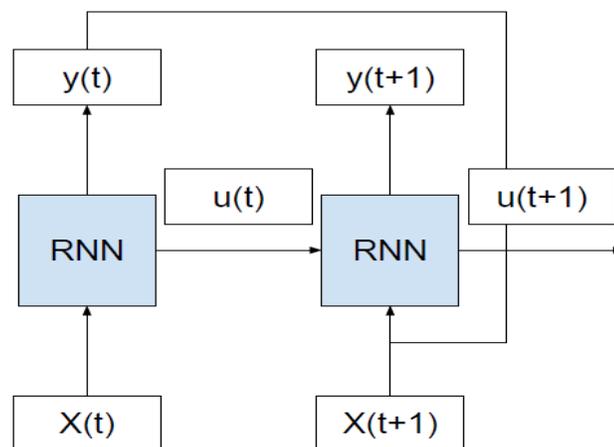


Figure 11 : Exemple d'un réseau récurrent ouvert montrant les entrées séquentielles et l'état interne les sorties

Apprentissage dans un réseau de neurones récurrents et problème de dissipation du gradient

L'apprentissage dans un réseau de neurones récurrents est fait en utilisant une rétropropagation dans le temps nommée BPTT (*Backpropagation through time*). Toutefois, bien que très utile dans certains cas, la présence de boucles de rétroaction dans les réseaux de neurones récurrents peut entraîner une croissance exponentielle du signal (si celui-ci est constamment multiplié par un poids supérieur à 1), ou encore sa disparition dans le cas contraire. Une solution pour régler ce problème de croissance exponentielle ou perte du gradient a été proposée par Hochreiter & Schmidhuber en 1997 avec l'introduction des LSTM (Long - Short Term Memory)

3.8 Architectures profondes

La recherche de solutions à certaines problématiques auxquels faisaient face les architectures disponibles en apprentissage machine et des résultats de recherche sur l'architecture fonctionnelle du cerveau ont permis l'émergence des architectures profondes. Nous présentons deux de ces problématiques ainsi que les motivations principales qui sous-tendent l'apprentissage profond.

Problématique 1 - Complexité des données, non-linéarité et extraction de caractéristiques

La complexité du monde qui nous entoure se retrouve dans les données et exemples pour les tâches en apprentissage machine. Prenons l'exemple de la reconnaissance ou la classification d'une image. L'image peut être représentée par un vecteur contenant les pixels dont elle est constituée ainsi que leur intensité, les différentes régions, les formes particulières, etc. Toutefois, Il ne faudrait pas tous les pixels pour reconnaître l'image, au contraire il s'agira d'utiliser certaines caractéristiques de l'image qui

permettront son identification. Il est donc nécessaire de concevoir des modèles permettant de trouver des caractéristiques, modéliser des non-linéarités et les grouper en abstraction pour la tâche de reconnaissance. Désigné sous le terme anglais de *Feature engineering*, ce processus pour trouver les caractéristiques est long, fastidieux et nécessite une expertise des connaissances du domaine.

Les architectures disponibles en AM avant l'émergence des architectures profondes ont différentes approches pour régler les problèmes. Ainsi en IA symbolique dans les systèmes à base de règle, la base des faits ainsi que les caractéristiques pour supporter le raisonnement étaient conçues manuellement et fournies de manière explicite au système. Des techniques comme la logique floue sont utilisées pour représenter le flou ou l'incertitude sur les entrées et disposent des mécanismes d'inférences appropriés. En apprentissage machine certains modèles offraient des bases théoriques pour supporter un début de l'apprentissage de la représentativité. Par exemple, un PMC avec trois couches est suffisant dans de nombreux cas pour l'approximation de n'importe quelle fonction avec une certaine précision. En augmentant le nombre de neurones de la couche cachée ainsi que les combinaisons de connexion entre la couche d'entrée et la couche cachée nous obtiendrons au niveau de la couche cachée, des regroupements d'entrées qui constitueront des caractéristiques intéressantes pour l'apprentissage. Cependant, une telle technique pourrait engendrer une croissance exponentielle du nombre de neurones de la couche cachée, une représentation non compacte et une limite dans la composition étant donné qu'il y a juste une couche intermédiaire entre l'entrée et la sortie. Dès lors intuitivement, avec une architecture à plusieurs couches cachées, chaque couche d'un modèle hiérarchique pourrait capitaliser sur les représentations successives des couches précédentes pour obtenir des abstractions de plus haut niveau.

Problématique 2 - Architecture profonde du fonctionnement du cerveau

Sur un plan biologique, de nombreuses études du cortex visuel notamment par Hubel et Wiesel (Hubel et Wiesel 1959) montrent qu'il existe une série d'aires impliquées

dans le traitement de l'information venant du nerf optique. Ces aires sont organisées en régions successives et hiérarchisées contenant chacune une représentation de l'entrée (le champ visuel). Les signaux se propagent d'une aire à la suivante et il y a aussi des connexions directes entre aires éloignées ainsi que des chemins parallèles. La représentation de l'information visuelle et son traitement sont donc distribués dans une architecture plus ou moins profonde. Également sur un plan fonctionnel, les humains organisent leurs apprentissages de concepts de façon hiérarchique. Nous apprenons d'abord des concepts simples (chiffres, lettres, etc.), et les combinons ensuite pour représenter des concepts de plus en plus compliqués et abstraits.

Face à ces problématiques, les motivations principales pour étudier les algorithmes d'apprentissage pour des architectures profondes sont entre autres : la découverte de représentation intermédiaire et la compositionnalité.

Motivation 1 : La découverte de représentation intermédiaire

L'apprentissage de représentation vise à extraire plusieurs niveaux de représentations pour mieux capturer les caractéristiques et les spécificités au sein des données. Elle permet de résoudre la tâche non pas sur des données brutes, mais plutôt sur des structures pertinentes créées manuellement ou découvertes automatiquement après plusieurs transformations supervisées et non-supervisées (pour l'extraction des traits caractéristiques discriminants) des entrées. Les représentations intermédiaires peuvent être exploitées pour une extraction graduelle des caractéristiques. Une tâche exécutée non pas sur des données brutes, mais plutôt sur ces caractéristiques offre de nombreux avantages dont une meilleure généralisation ainsi que l'utilisation dans des modèles génératifs.

Un avantage des RNA est qu'on peut empiler plusieurs couches cachées de façon à démultiplier la capacité. En composant les résultats des couches précédentes, chaque nouvelle couche peut tirer avantage des non-linéarités des précédentes pour produire

sa représentation, qui peut donc capter des variations plus complexes, difficilement identifiables manuellement dans de la distribution étudiée.

Motivation 2 : La compositionnalité

Le principe de compositionnalité repose sur le fait que la signification d'une expression complexe, la reconnaissance d'une forme complexe, etc. sont définies par les structures de base et les règles employées pour les combiner. La composition de représentation pour obtenir graduellement des caractéristiques plus abstraites afin de résoudre un problème est utilisée dans différent contexte et domaine aussi bien pour l'apprentissage que la résolution de problème. Mais elle n'était pas intégrée dans les algorithmes d'apprentissage machine. Avec le processus d'apprentissage profond, nous obtenons non seulement des représentations intermédiaires, mais également la composition hiérarchique automatique.

La représentation distribuée permet beaucoup de potentialité de composition dépendamment de la structure compositionnelle de la distribution sous-jacente ayant généré les données. Elle permet également la généralisation à de nouvelles combinaisons de caractéristiques au-delà de celle disponible dans les données d'apprentissage.

3.9 Discussion et positionnement

Les RNA sont très utilisés dans les projets et recherches en apprentissage machine. Ils présentent cependant certaines limites à considérer avant de les choisir comme approche. Entre autres limites il y a :

- L'explication et justification des résultats : dans un RNA les mécanismes pour obtenir les résultats sont encodés dans l'architecture et les valeurs des poids synaptique; ceci ne facilite pas une analyse explicative des résultats obtenus. Les architectures profondes fournissent certes des représentations

intermédiaires, mais ces représentations sont des combinaisons non linéaires complexes et successives sur les entrées des couches. Ces représentations ne sont donc pas nécessairement humainement compréhensibles.

- Le choix de l'architecture convenable : Il est difficile de définir l'architecture optimale du RNA pour un problème donné. Les attributs des données d'entrées peuvent suggérer le nombre de neurones de la couche d'entrées et les sorties désirées peuvent suggérer le nombre de neurones de la couche sortie. La difficulté provient des couches intermédiaires, du choix de la topologie du réseau et la spécification des hyperparamètres pour l'optimisation lors du processus d'apprentissage. Il y a donc une intervention humaine avec de l'expertise ou des heuristiques pour choisir l'architecture convenable et attribuer des valeurs aux hyperparamètres dépendamment du modèle.

Les RNA présentent toutefois de gros avantages par rapport à certains objectifs de la thèse. En effet nous visons entre autres à faire une classification des séquences comportementales en utilisant uniquement les données des mouvements oculaires en vue de déterminer s'il s'agit d'un comportement de novice ou d'expert. Le traitement pour la classification sera fait en ligne (*online*) lors de l'utilisation du STI pour supporter des décisions du module de tutorat. Les possibilités de prétraitement sont donc réduites sur les données oculométriques générées pour supporter la classification. Les RNA n'imposent aucune restriction sur les variables d'entrées (par exemple la distribution); sont résilients aux bruits et sont convenables pour développer des modèles sur des données ayant une hétéroscédasticité (forte volatilité et variance non constante). Ils constituent un modèle à considérer qui sera bien entendu comparé à d'autres modèles possibles sur des critères de performance bien définis. Les données oculométriques qui sont utilisées pour la classification sont d'une part de nature spatiale avec des fixations et saccades sur les images et d'autre part séquentielle en considérant l'ordonnement temporel dans lequel les AOIs sont regardés. Il est possible de combiner adéquatement différentes architectures de RNA pour obtenir une

architecture profonde spécialisée. En prenant en considération les caractéristiques spatiales et temporelles des données oculaires pour effectuer la classification, une architecture profonde avec un réseau de neurones à convolution et un LSTM pourrait être un modèle adéquat.

CHAPITRE IV

EXTRACTION ET PRÉDICTION DE SÉQUENCES

Ce chapitre aborde quelques notions de base importantes en fouille de données séquentielles pour comprendre l'état de l'art ainsi que la contribution envisagée.

4.1 Définitions et mise en contexte

Les fouilles de données séquentielles ont pour objectif la détection de régularités séquentielles ou temporelles observables dans un ensemble de données. Les définitions suivantes décrivent de manière formelle les notions de base en fouille de données séquentielles à savoir : item, itemset et séquence.

Définition : item

Un item i est l'élément atomique constitutif d'une suite d'interaction avec toutes les informations contextuelles associées. Par exemple, Il peut s'agir d'une action, d'une perception, d'une rétroaction du mannequin virtuel ou d'un état de la simulation.

Définition : Itemset

Un itemset I est un ensemble non vide d'items i tel que : $I = \langle i_1, i_2, \dots, i_n \rangle$. Un item ne peut être présent qu'une seule fois dans un itemset, mais peut figurer dans plusieurs itemsets.

Définition : séquence

Une séquence S est un ensemble ordonné d'itemsets tel que $S = \langle I_1, I_2, \dots, I_m \rangle$. La taille de S est le nombre m d'itemsets qu'elle contient. La longueur de S est le nombre $|I_1 \cup I_2 \cup \dots \cup I_m|$ d'items qu'elle contient.

Parmi les tâches généralement effectuées en fouille de données séquentielles, nous avons : la détection de patterns (motifs) fréquents et l'extraction de règles d'association. Ces tâches sont effectuées sur plusieurs familles de motifs comme les itemset, les séquences et dans des cas complexes sur des graphes. Les résultats obtenus peuvent servir à faire une classification, une analyse de similarité/différence, générer de nouvelles données, etc.

Un pattern fréquent (Agrawal et Srikant 1994) est un itemset dont l'occurrence est observée dans un pourcentage de séquences dépassant un seuil donné. Ce seuil est désigné sous le terme de « *support* ». Lorsque l'ordre des items de l'itemset est pris en compte, l'expression « *fouille de patterns séquentiels* » est utilisée. La définition formelle d'un pattern séquentiel est la suivante:

Un pattern séquentiel ps , est un pattern fréquent $ps = \langle t1a1, t2a2, \dots, tman \rangle$
avec $\{t1, t2, \dots, tm\} \in \mathbb{N}$, un ensemble d'estampilles temporelles tel que $t1 < t2 < \dots < tm$

Les règles d'association déterminent les relations de cooccurrence qui existent entre les items d'un ensemble de données. Ces relations sont désignées sous le terme « *associations* ». La définition formelle d'une règle d'association est la suivante:

Soit $I = \{i1 i2 \dots in\}$ un ensemble d'items. Soit $D = (S1, S2, \dots, Sn)$ une base de séquences. Chaque séquence Si de D est un ensemble d'items tel que $Si \subseteq I$. Une règle d'association est une implication de la forme $X \rightarrow Y$, avec X et Y des ensembles d'items non ordonnés, tel que $X \subset I, Y \subset I$ et $X \cap Y = \emptyset$.

Le support d'une règle d'association détermine le pourcentage de séquences dans lesquelles la règle est présente. La confiance d'une règle $X \rightarrow Y$ est le pourcentage de séquences de la base contenant X ainsi que Y .

4.2 État de l'art, défis et propositions de solutions

Cet état de l'art sera axé sur l'utilisation des fouilles de données éducationnelles pour l'acquisition automatique des connaissances du domaine. Dans cette perspective, le principe générique de l'acquisition automatique des connaissances du domaine se base sur l'analyse transversale et comparative de séquences issues des traces de résolution de plusieurs utilisateurs (dont le niveau varie de novice à expert). Afin d'utiliser les algorithmes de fouille de données, les notions de patterns fréquents/séquentiels ainsi

que de règles d'association sont redéfinies dans le cadre spécifique des traces de résolution. Par exemple (Toussaint 2015) définit une trace issue d'interactions perceptuelles et gestuelles comme suit:

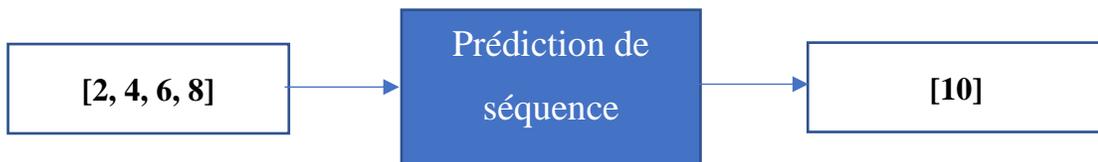
Soit $I = \{i_1, i_2, \dots, i_n\}$ un ensemble d'items. Soit A , un ensemble d'items « actions », avec $A \subset I$; soit G un ensemble d'items « gestes », avec $G \subset I$; soit P un ensemble d'items « perceptions », avec $P \subset I$; tel que $A \cup G \cup P \subseteq I$ et $A \cap G \cap P = \emptyset$. Un pattern perceptivo-gestuel $pa = \langle t_1 a_1, t_2 a_2, \dots, t_m a_m \rangle$ est un pattern séquentiel avec les items $a_1, a_2, \dots, a_m \in A \cup G \cup P$ et t_1, t_2, \dots, t_m , les estampilles temporelles des items.

Par la suite, il s'agit d'extraire des patterns fréquents/séquentiels et règles d'associations à partir de ces traces de résolution, établir des corrélations avec les évaluations à partir de règles expertes pour dégager des éléments de connaissance pouvant être intégrés au module des connaissances du domaine du STI. Ceci inclut les bonnes pratiques, les erreurs fréquentes, les comportements séquentiels souhaités, etc.

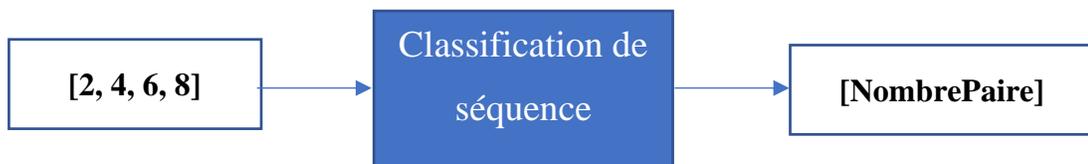
4.3 Les tâches en fouilles de données séquentielles

Les tâches en fouilles de données séquentielles présentent certaines différences par rapport aux autres types de tâches en apprentissage supervisé. En effet, la séquence impose un ordre sur les observations qui doit être préservé lors de la conception des modèles. Il existe principalement quatre types de tâches :

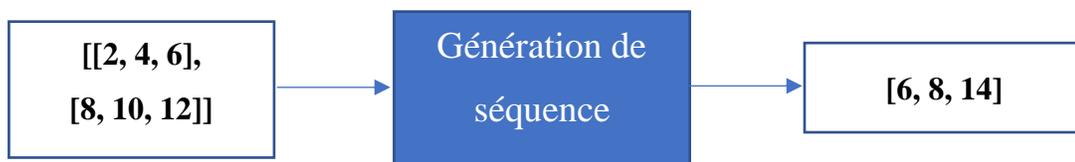
1. Prédiction de séquence : Étant donné une séquence d'entrée, prédire la valeur suivante dans la séquence. Par exemple, une prédiction des prochains achats d'un client, une prévision météorologique, une prévision boursière, etc.



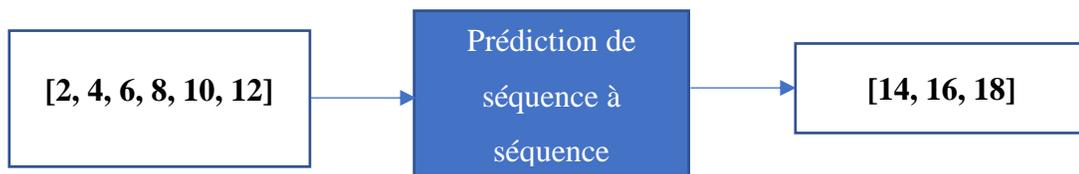
2. Classification de séquence : Étant donné une séquence d'entrée, attribuer une classe à la séquence complète ou à des sous-séquences. Par exemple, la classification de séquence d'ADN, la détection d'anomalie, etc.



3. Génération de séquence : Étant donné des séquences d'entrées, générer une nouvelle séquence de sortie ayant les mêmes caractéristiques que les séquences d'entrées. Par exemple, la génération de nouvelles phrases ou paragraphe d'un ouvrage, la génération de musique pour un chanteur, la génération de sous-titrage d'image, etc.



4. Prédiction de séquence à séquence : À partir d'une séquence d'entrées, générer une séquence de sortie. Cette tâche ressemble à la prédiction de séquence, mais il s'agit de prédire toute une séquence et non un seul item.



Nous présentons dans le tableau 2 certains algorithmes classiques de fouille de données pour le traitement des séquences ainsi que leurs objectifs.

| Exemple d'algorithme | Objectif de l'algorithme |
|---|--|
| Apriori (Agrawal et Srikant 1994) | Extraction d'itemsets fréquent |
| PrefixSpan (Pei, Han et al. 2004) | Extraction de patterns séquentiels |
| OPUS (Webb 1995) | Extraction de règles d'association |
| Non-Redundant Sequential Rules Algorithm (Lo, Khoo et al. 2009) | Extraction de règles séquentielles |
| Frequent Patterns (FP) Growth (Han, Pei et al. 2000) | Extraction d'itemsets fréquent Plus performant que Apriori sans génération de candidat |
| Eclat algorithm(Zaki 2000) | Extraction d'itemsets fréquent Plus performant que Apriori et procède par une recherche en profondeur |
| Variantes du modèle de Markov caché (par exemple N-Gram) | Extraction d'itemsets fréquents |

Tableau 2 : Quelques algorithmes classiques de fouille séquentielle

Les algorithmes classiques présentés n'ont toutefois pas été conçus spécifiquement pour les fouilles de données éducationnelles en vue d'extraire des connaissances pour la modélisation du domaine. Une des contributions dans ce sens a été le projet SPMF⁶ (Sequential Pattern Mining Framework) qui propose en plus de l'implémentation de plusieurs algorithmes de fouille de données classiques, des algorithmes plus spécialisés aux données éducationnelles comme CMRules (Fournier-Viger, Faghihi et al. 2012).

4.4 L'algorithme CMRules

CMRules est un algorithme d'extraction de règles séquentielles communes à plusieurs séquences. L'algorithme prend en entrée une base de séquences D , les seuils $minSeqSup$ et $minSeqConf$ pour produire en sortie l'ensemble des règles séquentielles contenues dans D qui respectent ces seuils. Il procède en considérant d'abord la base de séquences D fournie en entrée comme une base de transactions D' . Dans une transaction, la séquentialité des *itemsets* n'est pas prise en compte. Pour D et D' , chaque règle séquentielle $r = A \Rightarrow B$ de l'ensemble des règles séquentielles R , possède une règle d'association correspondante r' telle que $r' = A \rightarrow B$ dans l'ensemble des règles d'association R' .

$minSup$ et $minConf$: le support et la confiance minimum des règles d'association de l'ensemble R' $minSeqSup$ et $minSeqConf$: le support et la confiance minimum des règles séquentielles dans R . L'algorithme CMRules est présenté ci-dessous suivi des explications :

⁶ <http://www.philippe-fournier-viger.com/spmf>

| L'algorithme CMRules pour l'extraction de règles séquentielles |
|---|
| Entrées : D , $minSeqSup$, $minSeqConf$ |
| <ol style="list-style-type: none"> 1. Obtenir la base de transactions D' en enlevant l'information temporelle de D. 2. Trouver toutes les règles d'association dans la base de transactions en appliquant un algorithme de recherche de règles d'association comme <i>Apriori</i>. Choisir $minsup = minSeqSup$ et $minconf = minSeqConf$. 3. Parcourir la base de séquences D pour calculer la fréquence séquentielle et la confiance séquentielle de chaque règle d'association trouvée à l'étape précédente. Éliminer chaque règle r telle que $seqSup(r) < minSeqSup$ ou $seqConJ(r) < minSeqConf$. 4. Retourner l'ensemble des règles restantes. |
| Sorties : l'ensemble des règles séquentielles contenues dans D qui respectent ces seuils |

Explication

CMRules débute en ignorant l'information temporelle de la base de séquences pour obtenir une base de transactions. Il applique ensuite un algorithme de recherche de règles d'association sur la base de transactions avec les paramètres $minSup = minSeqSup$ et $minConf = minSeqConf$. Puis, l'algorithme calcule la fréquence séquentielle et la confiance séquentielle de chaque règle d'association en parcourant la base de séquences. Finalement à l'étape 4, les règles qui ne respectent pas les seuils $minSeqSup$ et $minSeqConf$ sont éliminées. Les règles restantes constituent l'ensemble des règles séquentielles.

En résumé, les algorithmes génériques de fouille de données séquentielles présentés ainsi que ceux spécifiques aux fouilles de données éducationnelles sont largement

utilisés dans plusieurs plateformes. Il existe toutefois des défis dont certains seront présentés dans la section suivante.

4.5 Défis des algorithmes de fouille de données séquentielles et propositions de solutions

Le défi principal des algorithmes de fouille de données séquentielles provient de l'hétérogénéité interne des données. Ce défi se manifeste par exemple par la fréquence des items.

La fréquence des items peut être variable pour diverses raisons. Prenons par exemple le cas classique du « panier d'épicerie ». Pour une personne donnée, les produits de consommation courante (aliments, cosmétiques) sont des achats plus fréquents que des produits d'utilisation prolongée (électroménagers). La fréquence d'item peut ne pas être intéressante à considérer, mais dans certains cas elle pourrait être déterminante pour comprendre la dynamique et les changements dans les items subséquents. Par exemple, l'achat d'électroménager pourra directement influencer les achats subséquents (lessive à vaisselle, assouplissant, etc.) ce qui se traduira par l'apparition de nouveaux items dans les séquences d'achat. Cette variabilité dans la fréquence pose un dilemme connu sous le nom du « dilemme de l'item rare » (Liu, Hsu et al. 1999) et se manifeste de deux manières : si le support est trop élevé, les itemsets/règles dont la fréquence est faible seront ignorés. Si le support est trop faible, toutes les combinaisons possibles de itemsets/règles fréquents seront explorées ce qui pourra engendrer un long temps de calcul et donner des résultats sans grand intérêt pour l'analyse. Pour pallier la problématique de l'item rare, les principales approches de solutions partent du principe que les items rares peuvent porter des informations essentielles non fournies par les items fréquents et propose :

1. D'utiliser des supports multiples (Liu, Hsu et al. 1999, Kumar, P.V.S.Srinivas et al. 2012)
2. De cibler explicitement les items rares (au détriment des items fréquents) (Szathmary, Napoli et al. 2007)
3. De partitionner les données en sous-ensembles homogènes (Klawonn et Höppner 2006).

Le dilemme de l'item rare se manifeste également avec les algorithmes exploitant les fouilles de données séquentielles pour l'acquisition automatique des connaissances du domaine. L'algorithme PhARules constitue une des solutions proposées.

4.6 L'algorithme PhARules

L'algorithme *PhARules* (*PHase-Aware RULES*) proposé par (Toussaint 2015) apporte une solution au dilemme de l'item rare et complète CMRules par la capacité à prendre en compte l'importance des phases dans les traces de résolution. Ainsi avec *PhARules*, une phase décisive de résolution qui génère peu de données ne sera pas négligée. L'algorithme PhARules prend en entrée une base de séquences D , un ensemble de descripteurs de phases Ω , un support *minSeqSup* et une confiance *minSeqConf* des règles séquentielles à extraire. Une explication des étapes de PhARules est présentée ci-dessous :

1. PhARules débute en parcourant les séquences de la base D pour identifier les phases auxquelles elles appartiennent. Elle génère les sous-ensembles de séquences $D_i, i=1, n$, représentant les n phases identifiées et dans lesquelles les séquences de la base D sont réparties.
2. Pour chaque sous-ensemble D_i , PhARules enlève les informations temporelles des itemsets des séquences générant ainsi les sous-ensembles de transactions $D_i, i=1, n$.

3. Chaque sous-ensemble D_i est ensuite parcouru pour en extraire les règles d'association pour les seuils $minSeqSup$ et $minSeqConf$ fournis en entrée. Ces seuils, définis pour l'extraction des règles séquentielles, sont ainsi utilisés pour l'extraction des règles d'association.
4. Pour chaque règle d'association ainsi obtenue dans D_i , calculer son support $seqSup$ et sa confiance $seqConf$ dans la base de séquences correspondante D_i .
5. Les règles dont le support et la confiance sont inférieures à $minSeqSup$ et $minSeqConf$ sont écartées, les règles restantes constituent l'ensemble des règles séquentielles orientées-phase extraites.

4.7 Discussion et positionnement

L'une des limites classiques de l'utilisation des algorithmes de fouille de données séquentielles pour la détection de motifs fréquents, de motifs séquentiels et des règles d'association est le nombre et la qualité décisionnelle des résultats obtenus. Parmi ces résultats, beaucoup peuvent potentiellement être non pertinents, presque similaires, incomplets ou dépourvus d'intérêt du point de vue des connaissances du domaine. Ces limites ne peuvent donc pas être totalement résolues par l'utilisation de métriques quantitatives comme le support et la confiance. La sélection des meilleurs patterns/règles ou ceux les plus significatifs par les algorithmes de fouille de données séquentielles représentent un grand défi. Cette problématique est plus renforcée pour les algorithmes dédiés aux données éducatives comme CMRules et PhARules. Ceci s'explique par le fait que les motifs et règles extraits doivent servir à supporter d'une part la conception d'un modèle des connaissances du domaine et d'autre part pour le suivi pédagogique des apprenants.

Une autre limite des approches pour l'extraction de motifs/règles réside dans le fait qu'elles sont comportementales c'est-à-dire se limitent uniquement aux

caractéristiques des séquences sans tenir compte du phénomène sous-jacent à leur apparition. Or ces séquences sont des manifestations observables d'un plan de résolution exécuté par un agent (par exemple un clinicien dans notre cas). Le plan de résolution est exécuté en tenant compte des variables contextuelles et contraintes de l'environnement qui, d'une part conditionnent l'émergence d'un pattern/règle observable et d'autre part impose les préconditions pour leur exécution ainsi que postcondition résultant de leurs exécutions.

Finalement, le modèle des connaissances du domaine généré à partir des patterns/règles extraits servent à supporter la production des services tutoriels de nature comportementale. Les services tutoriels épistémiques qui permettent d'aller au-delà d'un comportement observé pour considérer les connaissances qu'elles sous-tendent ne sont généralement pas offerts. Ceci s'explique par le fait que les motifs/règles extraits ne peuvent pas facilement être contextualisés par rapport aux connaissances du domaine. Toutefois, les experts les experts peuvent analyser le contexte des motifs extraits et suggérer une méta connaissance reliée.

En prenant en considération l'état de l'art sur les algorithmes dédiés à l'extraction de patterns/règles à partir de données séquentielles pour la conception d'un modèle des connaissances du domaine, les défis ainsi que les propositions de solutions actuelles, nous avons axés nos travaux sur les points présentés ci-dessous, ce qui servira de bases aux contributions de cette thèse.

- Des techniques pour une visualisation et un prétraitement des données séquentielles issues des mouvements oculaires.
- Un algorithme de fouille de motifs séquentiels de longueur variables en prenant en considération certaines spécificités des séquences de mouvements oculaires comme la durée pour avoir une fixation valide.

- La contextualisation des séquences extraites par rapport aux contraintes de l'environnement. Les contraintes de l'environnement sont fournies par les évènements clés, car ces évènements tiennent compte de la dynamique d'évolution et changement dans les signes vitaux ainsi que les décisions du personnel de prise en charge.

La fouille de motifs séquentiels a été exécutée hors ligne (*offline*) en vue d'extraire les motifs caractéristiques qui permettront de mettre à jour le modèle de l'expert.

**PARTIE 2 - ÉTUDE EXPLORATOIRE ET COMPARATIVE DE LA PERCEPTION
VISUELLE DE CLINICIENS NOVICES ET EXPERTS**

CHAPITRE V

MÉTHODOLOGIE, CADRE EXPÉRIMENTAL ET EXPÉRIMENTATIONS

Ce chapitre présente le protocole expérimental, expose le déroulement des expérimentations ainsi que la description des approches et des outils pour la collecte, la validation et l'analyse des données.

5.1 Mise en contexte

Les cas cliniques pour nos expérimentations portent sur l'anesthésiologie dans des situations critiques périopératoires. Après des discussions avec le CAAHC, les raisons ci-dessous ont motivé ce choix :

- La disponibilité au CAAHC de scénarios pour ces cas cliniques
- L'intérêt des cliniciens novices à améliorer une telle expertise
- La pertinence par rapport à nos objectifs de recherche, car ils requièrent une bonne coordination de la perception visuelle et du RC.
- La forte cohésion interne des compétences pour la résolution et un faible couplage avec d'autres expertises du domaine médical.
- Des observations empiriques de la part des instructeurs du CAAHC qui montrent que généralement la situation est prise en charge d'une façon globale

et rapidement anticipée par les experts, alors que les novices adoptent une approche hypothético-déductive de manière analytique et évolutive.

Le groupe de participants ciblés pour l'expérimentation est constitué des médecins résidents en anesthésiologie, en début de leur résidence et d'un groupe de médecins anesthésiologistes en pratique depuis plus de huit ans. Un accord initial de participation est signé ainsi qu'un engagement de confidentialité. La dimension principale qui est prise en compte pour les expérimentations est la perception visuelle. À ceci s'ajoute une validation qualitative du raisonnement clinique associé. Il est important de choisir d'une part des outils technologiques éprouvés pour la collecte des données oculométriques et l'enregistrement du raisonnement clinique et d'autre part établir un protocole expérimental convenable pour les objectifs de la recherche.

5.2 Technologie pour la collecte des données

La collecte des données a été effectuée en utilisant la suite Tobii Pro⁷. La suite Tobii Pro est une gamme d'outils matériels et logiciels destinée aux chercheurs en milieu académique ou en entreprises impliquées dans la recherche ou le développement de produits exploitants l'oculométrie. L'avantage principal des solutions Tobii est de permettre de mesurer les comportements oculaires de manière objective et automatisée tout en générant plusieurs données descriptives, temporelles et spatiales reliées au mouvement oculaire. Tobii offre également une suite logicielle permettant d'automatiser l'extraction, la compilation et la génération de statistiques descriptives. Ceci permet de réduire la variabilité et les incertitudes liées aux méthodes manuelles

⁷ <http://www.tobii.com>

d'observation, de codage et de traitement. Le processus peut être automatisé de sorte à faciliter et rendre efficace le recueil, le traitement, l'analyse et la comparaison d'une grande quantité de données oculaires. L'étude de la perception visuelle avec le Tobii a permis d'obtenir des résultats intéressants dans de nombreux domaines, dont le marketing web (Nielsen et Pernice 2009), les STI (Conati et Merten 2007), etc.

5.3 Participants

Comme précisé dans l'introduction, l'expertise est un continuum fondé sur les connaissances, les compétences et l'expérience plutôt que sur des catégories prédéfinies. Sur la base de cette clarification, et par rapport aux objectifs de recherche, nous utilisons les définitions ci-dessous pour novice et expert :

- Novice : est un clinicien résident dans la première ou deuxième année du programme de résidence (PGY1 ou PGY2)
- Expert : est un clinicien membre du personnel hospitalier avec plus de 8 ans d'expérience.

Après l'approbation du comité éthique et le consentement des participants, un échantillon de huit experts et de neuf novices du département d'anesthésiologie de l'Université de Montréal a été recruté sur une base volontaire. Les objectifs du projet ont été expliqués aux participants ainsi que le déroulement de l'expérimentation et la durée approximative.

Les participants ont visualisé une simulation portant sur un cas complexe et réaliste tout en verbalisant à haute voix leur processus de résolution et de raisonnement clinique exactement comme s'ils avaient la charge de la situation clinique. La simulation présente une vue globale de la scène, une vue sur les mouvements du personnel, une vue rapprochée sur l'intervention ainsi que les signes vitaux. Spécifiquement, les participants doivent expliquer ce qu'ils ont vu dans les différentes zones d'intérêt (Figure 14) pour supporter leur raisonnement clinique. De plus, les participants doivent

expliquer ce qu'ils ont fait lors de certains évènements clés identifiés tout au long de la simulation. À la fin de l'expérimentation, les participants ont été invités à remplir un questionnaire rétrospectif (Voir Annexe A) visant à collecter :

- Des variables indépendantes telles que l'âge et le sexe
- Les commentaires, etc.
- L'avis sur le réalisme de la situation clinique, de l'expérimentation, le diagnostic final, etc.

La section ci-dessous présente le détail du déroulement de l'expérimentation.

5.4 Déroulement de l'expérimentation

Les étapes ci-dessous présentent le déroulement de l'expérimentation. Les étapes 3.1 (Visualisation de la simulation) et 3.2 (Protocole Think aloud) sont réalisées simultanément.

Étape 1 : Consignes générales au participant

Il s'agit de :

- Éteindre le téléphone portable ou toute source de distraction potentielle.
- Rappeler l'objectif du projet
- Mettre en contexte pour le cas clinique qui sera présenté

Donner les instructions pour le déroulement de l'expérimentation : regarder la simulation comme si vous aviez la charge de la situation clinique, réaliser votre raisonnement clinique, effectuer le protocole Think a Loud, etc.

Étape 2 : Calibration du Tobii par rapport aux yeux du participant

Avant de commencer l'enregistrement du suivi oculaire, chaque participant est soumis à une procédure de calibrage de ses yeux. Au cours de cette procédure, l'oculomètre mesure les caractéristiques des yeux du participant et les utilise avec un modèle interne

d'œil 3D pour calculer les données du regard. Ce modèle comprend des informations sur la forme des yeux, le reflet de la lumière, et les propriétés de réflexion des différentes parties des yeux (par exemple la cornée, l'emplacement de la fovéa, etc.). Pendant le calibrage, le participant devrait regarder des points en mouvement spécifiques sur l'écran. Plusieurs captures des mouvements oculaires sont recueillies et analysées pour effectuer une validation de la précision des données. Lorsque la procédure est terminée, la qualité de l'étalonnage est affichée. Si elle est inférieure à un seuil spécifié, une autre calibration est effectuée.

Étape 3.1 : Visualisation de la simulation

Une simulation haute-fidélité d'une situation d'intubation et de ventilation impossible (CICO : Can not Intubate / Can not Oxygenate) basée sur un algorithme suggéré par la Difficult Airway Society pour gérer l'intubation difficile imprévue chez les adultes a été produite au CAAHC. La durée totale de la simulation est d'environ 13 minutes et a été enregistrée en vidéo haute définition. Le scénario porte sur un patient avec un collier cervical dans l'unité de soins. Tout au long de la simulation, le patient présente une détresse respiratoire croissante aggravant l'insuffisance respiratoire. L'anesthésiste (ou anesthésiologiste) ne parvient pas à ventiler, intuber et insérer un masque laryngé menant à une cricothyroïdotomie. La simulation a été faite pour intégrer diverses complications imprévues tout en étant réaliste. Elle intègre des notions qui nécessitent :

- Une bonne évaluation de la situation et une préparation du matériel nécessaire à la gestion du cas.
- Des essais de sécurisation des voies aériennes et de ventilation grâce aux moyens préparés. Cette gestion du cas clinique va s'avérer infructueuse.
- Des procédures de prise en charge plus complexes comme la trachéotomie.

Chaque participant a regardé un moniteur sur lequel la simulation est diffusée. Il s'agit d'un moniteur HD de 23 pouces (1920px par 1080 px) équipé de TobiiEyeX® TX300, qui un oculomètre qui collecte et agrège automatiquement diverses données de la

perception visuelle. L'oculomètre a été réglé avec une précision de 0,4 degré et mis à un taux d'échantillonnage de 60 Hz. La figure 13 présente une photo du moniteur avec l'oculomètre TobiiEyeX® TX300.



Figure 12 : Photo du moniteur avec l'oculomètre TobiiEyeX® TX300

L'écran est divisé en sept zones, représentant chacune une zone d'intérêt (nommée AOI : Area Of Interest). Ces AOI ne se recouvrent pas, sont fixes pour toute la durée de la simulation et n'étaient pas visibles pour les participants. Ces AOI permettront d'attacher des étiquettes pendant l'extraction des données par l'oculomètre. La figure 14 présente ces différentes AOI soit :

- ZONE 1: Cet AOI montre une vue globale de la scène.
- ZONE 2: Cet AOI montre également une vue globale de la scène en mettant l'accent sur les mouvements du personnel de prise de charge.
- ZONE 3: Geste technique. Cet AOI montre une vue rapprochée du cou du patient avec l'équipe clinique effectuant la prise en charge médicale.

- ZONE 4: Ce AOI montre la fréquence et le rythme cardiaque. C'est le nombre de battements cardiaques (ou pulsations) par unité de temps (généralement en minutes). Chez l'adulte en bonne santé, au repos, la fréquence cardiaque se situe entre 50 (sportif pratiquant l'endurance) et 80 pulsations par minute.
- ZONE 5: Cet AOI montre la saturation (pulsée de l'hémoglobine en oxygène) : c'est le taux d'oxygène transporté par l'hémoglobine, exprimé en pourcentage (%). Des valeurs de la saturation normale sont entre 98 % et 100 %.
- ZONE 6: Cet AOI montre la pression artérielle. Il s'agit de la mesure de la pression ou de la force qu'exerce le sang contre les parois des vaisseaux sanguins appelés artères. La lecture de votre pression artérielle repose sur deux mesures : systolique et diastolique. Le nombre supérieur représente la force exercée lorsque le cœur se contracte et expulse le sang (systolique) et le nombre inférieur est la pression la plus basse, au moment où le cœur se détend entre deux battements (diastolique).
- ZONE 7: Cet AOI montre la fréquence respiratoire. C'est le nombre de cycles respiratoires (inspiration et expiration) par unité de temps, mesurés chez un individu. La fréquence respiratoire normale est de 12 à 20 cycles par minute chez l'adulte et l'adolescent.

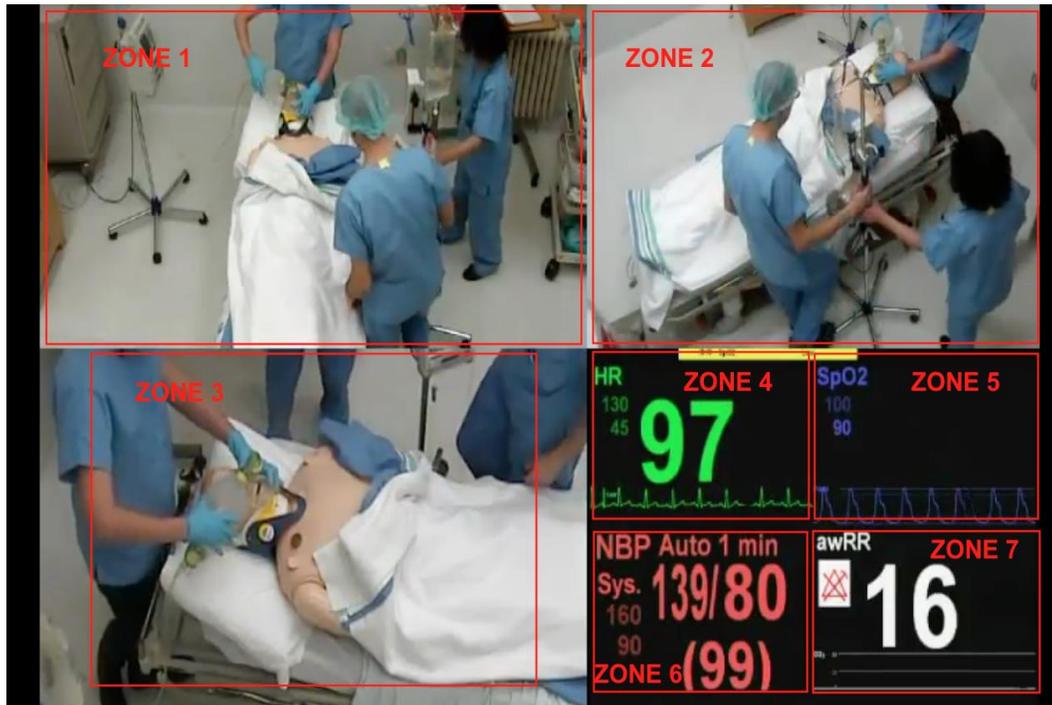


Figure 13 : Vidéo de la simulation vue par les participants avec les AOI

Étape 3.2 : Protocole *Think aloud*

Le *Think aloud* est un protocole très utilisé dans l'étude d'un large éventail d'activités cognitives pour capter les réflexions, observations, etc. provenant du participant à l'expérimentation. Au cours de notre protocole *Think aloud*, les participants sont invités à verbaliser à haute voix leur raisonnement clinique tout en regardant la simulation. Plus précisément, les participants doivent effectuer leur raisonnement clinique à haute voix, expliquer ce qu'ils regardent sur les prises de vue globale de la scène, les gestes techniques, les signes vitaux pour soutenir leur raisonnement clinique ainsi que toute heuristique utilisée. Finalement, les participants doivent expliquer ce qu'ils auraient dû faire dans des événements clés. Ces événements ont été sélectionnés avec précision avec l'aide des instructeurs du CAAHC d'une part pour leur importance

sur le plan des complications cliniques à reconnaître pour une bonne prise en charge et d'autre part pour leur importance comme indices dans la sensibilité situationnelle. Le tableau 3 présente les évènements choisis et présentés dans l'enregistrement de simulation, de même que les temps où ils apparaissent et les acteurs impliqués. Finalement un microphone a été connecté à l'oculomètre (Tobii), ce qui permet d'enregistrer simultanément la voix durant le Think aloud et les mouvements oculaires correspondants.

| Évènements "Clés" | | Temps(mm : ss) |
|---|---|-----------------------|
| Professionnels de la santé : Médecin, Infirmière, Inhalothérapeute | Appel à l'aide | 02 : 41 |
| | Ventilation au Masque | 03 : 35 |
| | Installation de la canule oro-pharyngée | 06 : 41 |
| | Utilisation du vidéolaryngoscope (Glidescope) | 07 :35 |
| | Insertion d'un dispositif supra glottique | 08 : 33 |
| | Déclanchement du Code Bleu | 09 : 39 |
| | Procédure de Cricotomie | 10 : 32 |
| Patient | Diminution réponse verbale | 01 : 10 |
| | Fermeture des yeux | 01 : 25 |
| | Hypoventilation | 02 : 09 |
| Signes vitaux | Diminution de la saturation | 01 : 37 |
| | Bradycardie | 08 : 35 |
| | Disparition du signal de saturation | 10 : 22 |

Tableau 3 : Évènements "Clés" de la simulation pour l'analyse comportementale

5.5 Collecte et validation des données

La validation des données a permis de vérifier si ces dernières pourraient être exploitables pour les objectifs visés. Les données de tous les participants ont été collectées à l'aide du programme Tobii Studio®. Deux critères d'acceptation ont été établis :

1. La qualité des données oculaires : des facteurs environnementaux (par exemple la luminosité), la distance par rapport à l'écran du Tobii, la distraction du participant (regarder en dehors des zones) peuvent affecter l'exactitude des données de suivi oculaire collectées. Un seuil de 75% a été fixé.
2. La conformité au protocole Think a Loud : deux instructeurs indépendants du CAAHC ont visualisé les vidéos avec les mouvements oculaires enregistrés par le Tobii et ont écouté les raisonnements cliniques. L'objectif est de valider que le mouvement dynamique des yeux entre les AOI témoigne d'un processus actif engagé par le clinicien pour résoudre le cas clinique. De ce fait, si un participant ne verbalise pas son RC; ceci constitue un critère de rejet.

Sur la base des deux critères précédemment mentionnés, les données de trois participants ont été rejetées, car elles ne répondaient pas aux critères d'acceptation adaptés : deux novices et un expert. Les données restantes, soit sept novices et sept experts, ont servi aux analyses.

Avec une configuration de 60Hz, l'oculomètre (Tobii) collecte des points de données des mouvements oculaires tous les 17ms. Chaque point de données est identifié par un *timestamp* (estampille temporelle) et les coordonnées : abscisse x et ordonnée y. Ces points sont traités et agrégés dans différentes catégories de mouvements oculaires, à savoir les fixations et les saccades.

- Une fixation est la pause du mouvement des yeux sur une zone spécifique de l'écran.
- Une saccade est un mouvement rapide entre les fixations.

Si deux points sont à une distance minimale prédéfinie l'un de l'autre en fonction de leurs coordonnées, en d'autres termes, les participants n'ont pas bougé les yeux de manière significative entre ces deux points d'échantillonnage, alors ils sont affectés à la même fixation. Graphiquement, comme présenté à la figure 14, les fixations sont représentées par des points (les plus gros points indiquent un temps de fixation plus long), alors que les saccades sont indiquées par des lignes entre les fixations.

Étant donné que les mouvements oculaires sont captés à 17ms d'intervalle, avec une durée d'environ 13 min pour la simulation, nous obtenons un total d'environ **46 000 enregistrements par participant**. Avec les **14 participants retenus**, nous obtenons un total de **645 000 enregistrements**.

Si au cours d'un enregistrement une fixation est à l'intérieur d'un AOI, nous mettons la valeur à 1; sinon, nous fixons la valeur à 0. Puisque les AOI ont été créés pour être mutuellement exclusifs, à priori en un instant t donné, un seul AOI devrait avoir la valeur de 1.

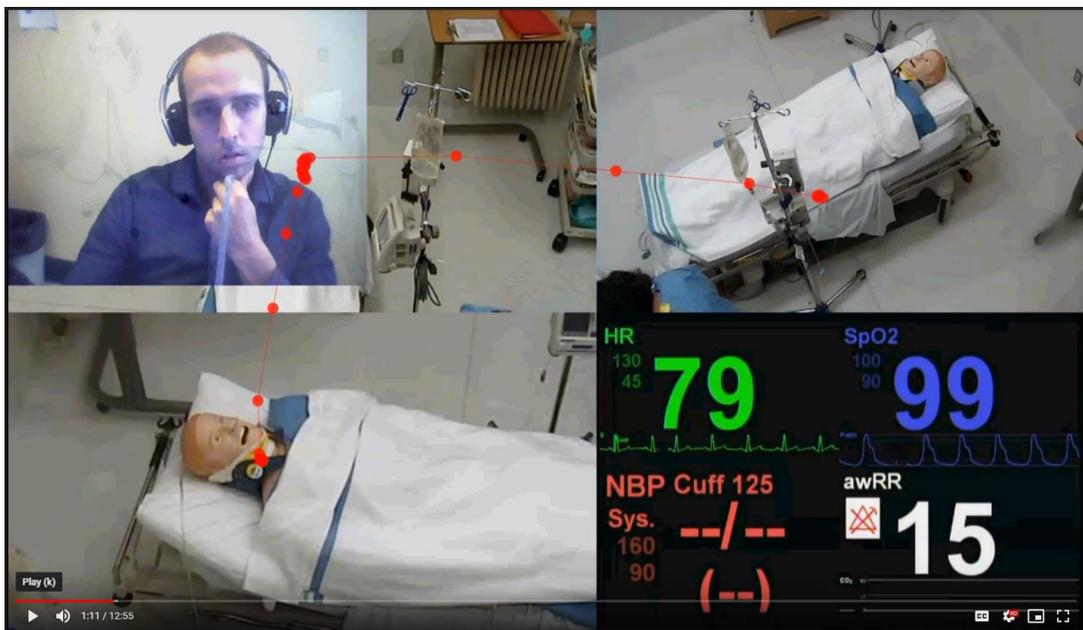


Figure 14 : Figure montrant les fixations et saccades un participant

5.6 Modélisation, caractérisation et traitement des données

Les données oculaires enregistrées par le Tobii, l'évolution des signes vitaux, les prises en charge des professionnels de la santé impliqués, l'enregistrement de la verbalisation du participant constituent une source cohérente et factuelle de données dans un même axe temporel sur l'ensemble les modalités perceptivo-décisionnelles mises en jeu dans la résolution. Il convient donc de traiter ces données multisources de manière conjointe dans le but d'analyser les interactions dans leur globalité en tenant compte de la séquence. En vue de permettre un traitement automatique par les programmes que nous avons développés dans le cadre du projet, il nous a fallu définir un modèle formel et computationnel qui reflète de manière cohérente l'activité multi dimensionnelle sous-jacente. Ce modèle doit avoir certaines caractéristiques :

- Une expressivité pour décrire la situation : rendre compte des interactions perceptivo-décisionnelles tout en conservant leurs séquentialités et occurrences.
- Une représentation computationnelle pour permettre l'exploitation automatique à des fins d'analyse et d'apprentissage machine. L'objectif principal est de pouvoir extraire automatiquement des connaissances comportementales pouvant servir au modèle de connaissances de l'expert.

Les modèles serviront entre autres à :

1. Comparer les séquences : similitudes, distance, etc.
2. Classifier des extraits de séquences comportementales. Par exemple, est-ce qu'un extrait de séquence perceptuelle est un comportement d'un novice ou d'un expert ?
3. Extraire les séquences oculaires spécifiques aux novices
4. Extraire les séquences oculaires spécifiques aux experts
5. Extraire les séquences oculaires mixtes aux deux groupes

Toutes les perceptions visuelles ne sous-tendent pas les mêmes objectifs et ne se traduiront pas par les mêmes traces. Nous distinguerons donc les séquences visuelles qui sont de simples distractions, des séquences visuelles de vérification qui se limitent à une prise d'information sur l'état de l'environnement et les séquences visuelles de décision qui sous-tendent un processus cognitif visant à faire le RC.

5.7 Processus générique de traitement des traces de résolution

Le modèle que nous proposons pour le processus générique de traitement des traces de résolution est une adaptation de l'architecture fonctionnelle du framework PeTRA⁸ (PERceptual-gestural TRACes treatment framework) qui est utilisé essentiellement par des chercheurs en STI et en fouille de données éducationnelles. Ce modèle est présenté dans la figure 15, suivi de quelques brèves explications.

⁸ projet-undertracks.imag.fr

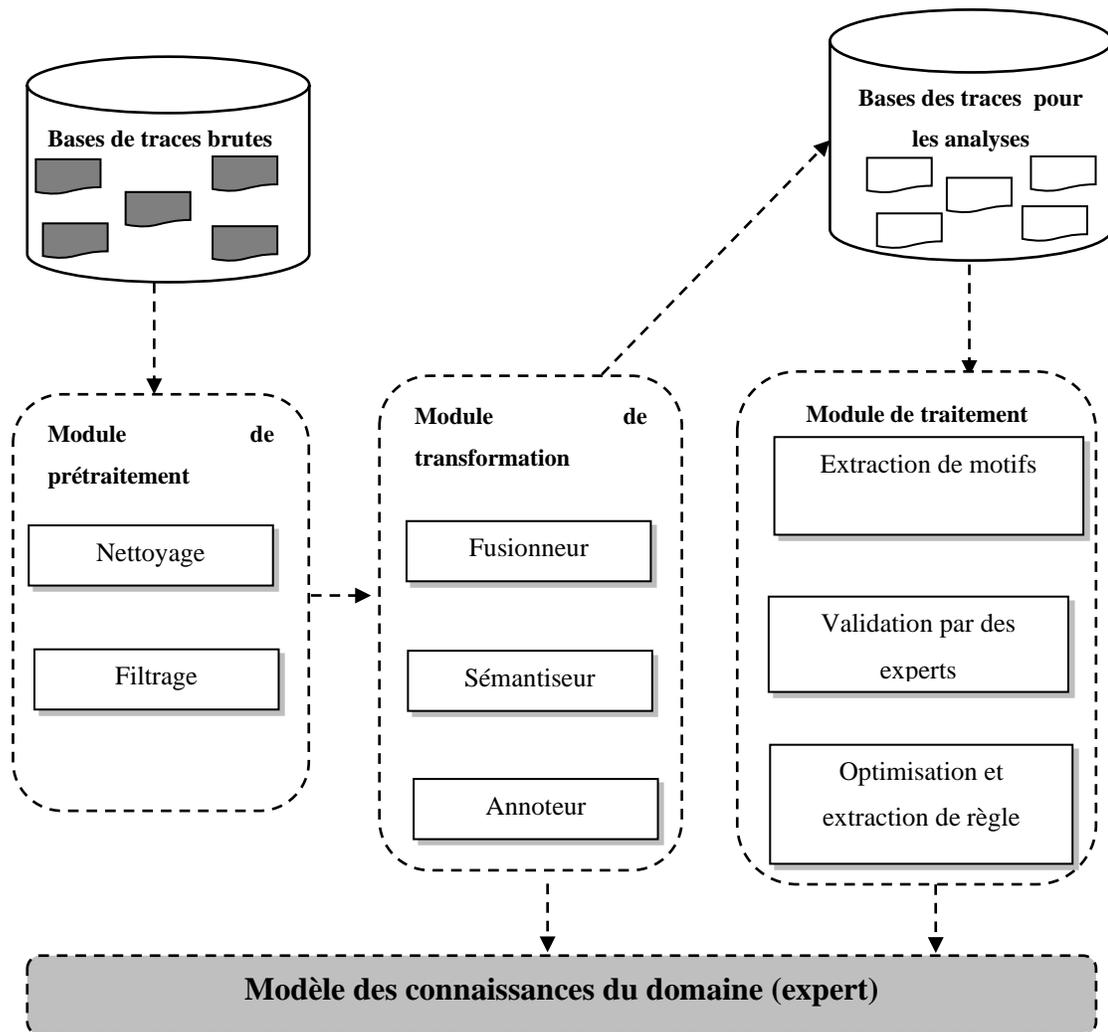


Figure 15 : Modules et processus génériques de traitement des traces de résolution

Module de prétraitement

- Nettoyage : a pour objectif de supprimer des traces brutes fournies en entrée les bruits provenant des données erronées, des erreurs de formatage des données. Si possible, ces erreurs pourront être détectées et corrigées.
- Filtrage : a pour objectif d'enlever des traces nettoyées des séquences ou données jugées non pertinentes pour la suite des traitements.

Module de transformation

- Fusionneur : a pour objectif de joindre les traces des différentes sources provenant d'une même interaction en prenant soin de garder la séquentialité de leurs occurrences afin de conserver la cohérence chronologique du déroulement des interactions.
- Sémantiseur : a pour objectif d'attribuer une dénomination sémantique aux valeurs de certains attributs ou zones. Il s'agit en fait d'une opération de discrétisation en vue de faciliter les manipulations et interprétations.

Le module de traitement

- Extraction de motifs de mouvements oculaires en utilisant les algorithmes de fouille de données séquentielles.
- Validation par les experts des motifs de mouvements oculaires extraits. Cette validation est faite en considérant d'une part les connaissances du domaine et d'autre part la pertinence pédagogique.
- Optimisation à partir des motifs de résolution spécifique et génération de motifs génériques grâce à un regroupement de certains motifs en instances ou spécialisations de comportements plus généraux. L'extraction de règles est faite sur les motifs optimisés en vue d'obtenir des règles d'association.

Modèle des connaissances du domaine (expert)

Ce module exploitera les extraits du module de transformation ainsi que du module de traitement pour construire et faire évoluer les connaissances du domaine. Dans le projet de recherche, les connaissances du domaine sont les connaissances comportementales des experts. En plus des motifs optimisés et validés, le modèle contient des données issues de l'analyse du parcours de résolution, de la validation des phases (événement clés) des statistiques à des fins d'analyse de l'apprentissage, etc.

CHAPITRE VI

ANALYSES ET RÉSULTATS OBTENUS

Ce chapitre présente les analyses préliminaires ainsi que les résultats obtenus. Ces résultats sont des graphiques générés en utilisant le logiciel Tobii Studio et des programmes développés sur mesure pour générer la visualisation des données de nature oculométrique.

Nous avons effectué trois types d'analyses à savoir :

1. Analyse préliminaire : le but est de comparer les novices aux experts en utilisant des statistiques descriptives sur les fixations dans les AOI.
2. Analyse comportementale : le but est de faire une analyse plus fine de la perception visuelle des novices et experts autour des évènements clés.
3. Analyse globale des trajectoires des fixations et saccades (*scanpath*) : le but est de visualiser une représentation de *scanpath* caractéristique des novices et experts. Par la suite, utiliser des algorithmes pour faire des comparaisons croisées des *scanpath*.

Les éléments visuels obtenus permettent d'avoir de manière exploratoire et comparative une compréhension globale à haut niveau de la dynamique dans les données ainsi que de dégager des tendances ou des valeurs inhabituelles. Ceci nous fournit des indices qui servent de base pour des analyses plus poussées avec des techniques d'apprentissage machine.

6.1 Prétraitement initial

Le prétraitement initial des données a consisté à supprimer dans les données générées par l'oculomètre les colonnes non pertinentes aux analyses ou à encoder certaines valeurs (par exemple le nom des participants). Nous obtenons les variables présentées dans le tableau ci-dessous :

| Variable | Description | Valeur |
|----------------------------|---|---|
| ParticipantId | L'identifiant unique du participant | <ul style="list-style-type: none"> ▪ Expert1 ... Expert7 ▪ Novice1... Novice7 |
| Expert | Indique si le participant est expert ou novice | <ul style="list-style-type: none"> ▪ 1 = expert ▪ 0 = novice |
| RecordingTimestamp | Le timestamp en millisecondes depuis le début de l'enregistrement | Timestamp |
| EyeTrackerTimestamp | C'est le temps en microsecondes où un point de données spécifique a été échantillonné par l'oculomètre. | Microsecondes |
| FixationIndex | Une valeur numérique représentant l'ordre dans lequel un événement de fixation a été enregistré. | C'est un index auto-incrémenté commençant par 1 |
| GazeEventType | Type d'événement de mouvement oculaire: | <ul style="list-style-type: none"> ▪ Fixation ▪ Saccade ▪ Unclassified |

| | | |
|------------------------------------|---|--|
| GazeEventDuration | Durée en millisecondes d'un événement de mouvement oculaire | Environ 17 millisecondes |
| AOI 1 - Scene View 1 | Vue globale de la scène 1 | <ul style="list-style-type: none"> ▪ Si une fixation est dans un AOI, la valeur est définie sur 1 sinon, la valeur est 0 ▪ Donnée manquante = impossible d'enregistrer |
| AOI 2 - Scene View 2 | Vue globale de la scène 2 | |
| AOI 3 - Technical gesture | Geste technique | |
| AOI 4 - Heart Rate | Fréquence et rythme cardiaque | |
| AOI 5 - Arterial Saturation | Saturation (pulsée de l'hémoglobine en oxygène) | |
| AOI 6 - Blood pressure | Pression artérielle | |
| AOI 7 - Respiratory rate | Fréquence respiratoire | |

Tableau 4 : Données pour l'analyse

6.2 Analyses préliminaires

Pour cette analyse, nous comparons les novices aux experts en utilisant des statistiques descriptives comme la moyenne sur différentes métriques associées aux fixations. Ces observations servent à dégager des similitudes et différences comparatives sur les caractéristiques des mouvements oculaires dans les AOI. Les graphiques ont été générés en utilisant le logiciel Tobii Studio et nous présentons ci-dessous certains des graphiques.

Graphique 1 : Durée totale de la fixation

Cette métrique mesure la somme de la durée de toutes les fixations dans un AOI. La figure ci-dessous présente la moyenne de la durée totale de fixation pour les novices et les experts.

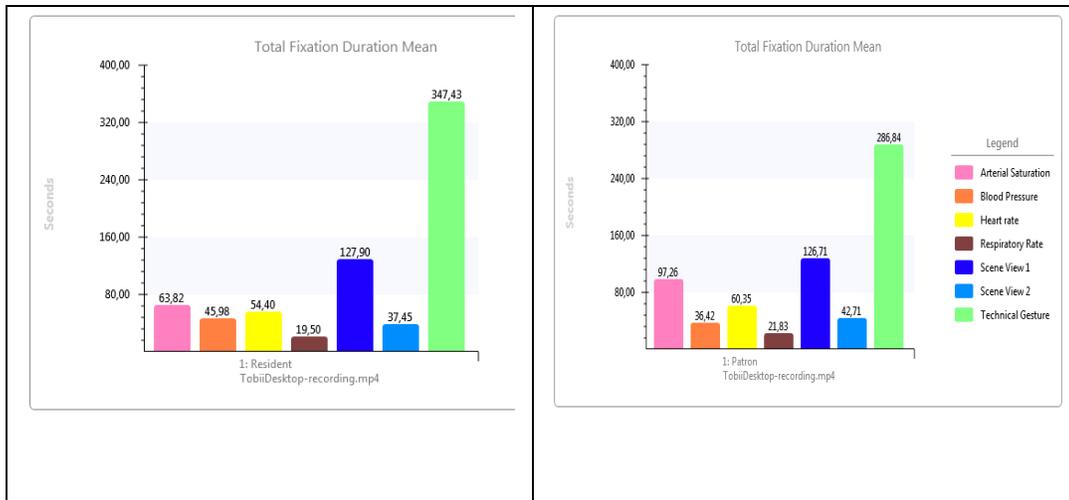


Figure 16 : Moyenne de la durée totale de fixation de Novice Vs Expert

Graphique 2 : Nombre de fixations

Cette métrique mesure le nombre fixation sur un AOI spécifique. La figure ci-dessous présente la moyenne du nombre de fixations pour les novices et les experts. Le nombre de fixations n'est pas nécessairement corrélé à la durée totale, car il est possible d'avoir un petit nombre de fixations ayant une durée élevée chacune.

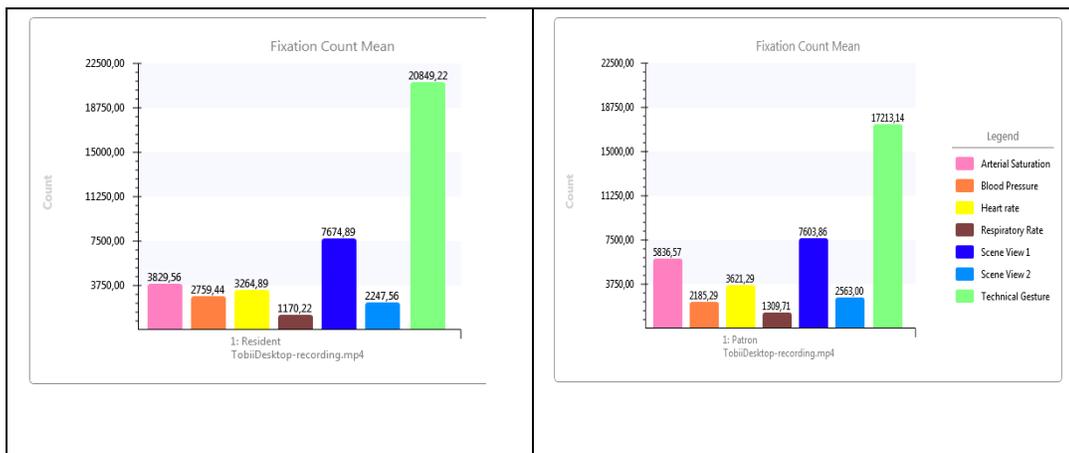


Figure 17 : Moyenne du nombre de fixations Novice Vs Expert

Graphique 3: Durée des visites

Cette métrique mesure la somme de toutes les durées de visites dans un AOI. Une visite individuelle est définie comme l'intervalle de temps entre la première fixation sur l'AOI et la fin de la dernière fixation dans le même AOI où il n'y a pas eu de fixations en dehors du AOI. La figure ci-dessous présente la moyenne de la durée totale de la visite pour les novices et les experts.

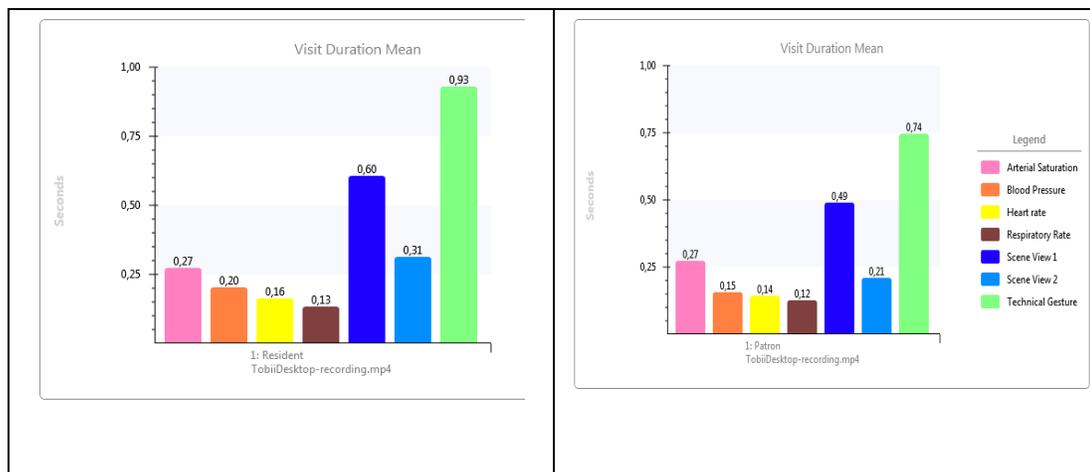


Figure 18 : Moyenne de la durée des visites Novice Vs Expert

Graphique 4 : Le heatmap (carte thermique)

Le heatmap (carte thermique) constitue une façon populaire de visualiser les données de suivi oculaire. Un heatmap utilise différentes couleurs pour montrer le nombre et la durée de fixations effectuées par les participants dans certaines zones de la simulation.

- Le rouge indique le plus grand nombre de fixations ou la plus longue durée de fixations. Ceci est suivi du jaune puis du vert.
- Une zone sans couleur sur le heatmap signifie une absence de fixations

La figure 20 montre un heatmap pour le novice; la figure 21 présente en termes de pourcentage l'information provenant du heatmap. Les figures 22 et 23 présentent les mêmes résultats pour les experts.



Figure 19 : Heatmap pour les novices

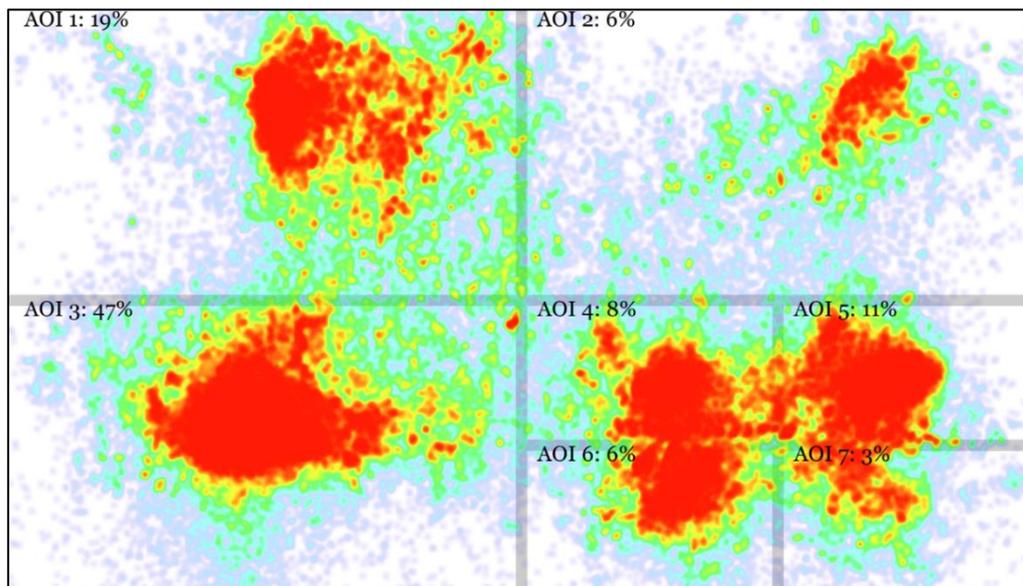


Figure 20 : Heatmap avec pourcentage pour les novices



Figure 21 : Heatmap pour les experts

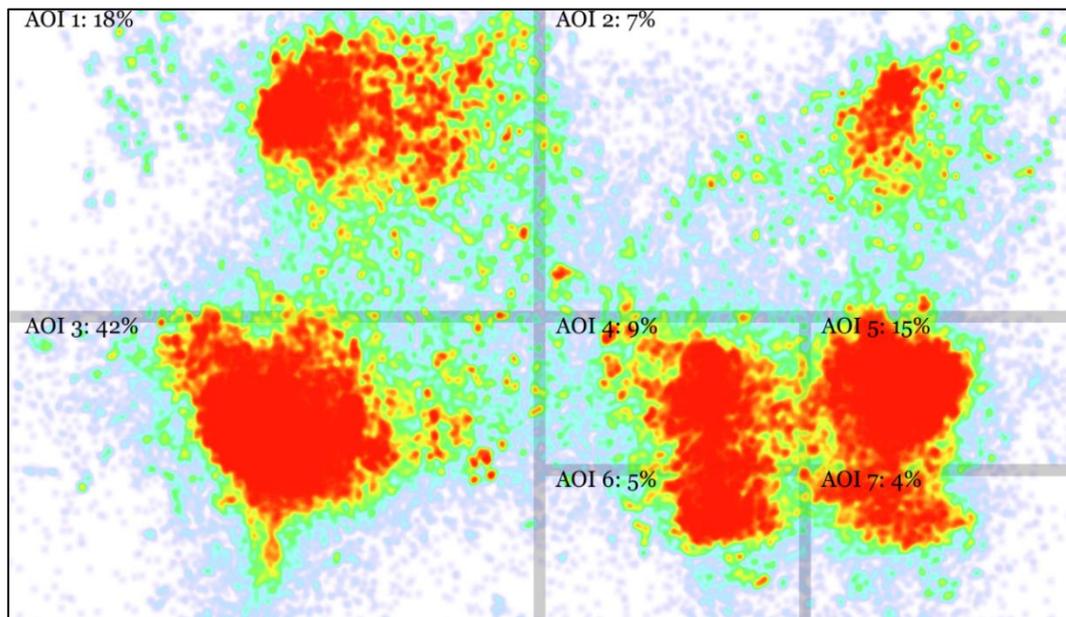


Figure 22 : Heatmap avec pourcentage pour les experts

Somme toute, l'analyse préliminaire et exploratoire a permis de dégager certaines observations présentées ci-dessous :

- **Observation 1** : En moyenne, les experts et les novices ont leur durée de fixation totale la plus élevée sur le geste technique suivi de la vue de la scène globale 1 (figure 17). Ce résultat est confirmé par le nombre de fixations (figure 18).
- **Observation 2** : En moyenne, la durée des visites pour les novices et les experts montre que même si nous avons un plus grand nombre de fixations de l'expert sur la saturation artérielle, ils finissent par avoir la même durée de visite moyenne. Donc les novices ont fait moins de fixations, mais ont passé plus de temps par fixations (voir figure 17 et figure 19).
- **Observation 3** : En moyenne, les experts ont regardé la saturation pour une durée plus longue et le geste technique pour une durée plus courte que les novices (voir figure 17 et figure 19).
- **Observation 4** : En moyenne, la zone de la vue de la scène globale 2 est été la moins regardée par les novices et les experts (figure 17, figure 18 et figure 19).
- **Observation 5** : Le heatmap montre que les novices ont passé moins de temps à regarder l'AOI de la saturation et plus de temps à regarder l'AOI geste technique que les experts.

Sur la base d'observations sur les graphiques générés avec le Tobii Studio, une analyse statistique a permis de dégager que les novices ont passé beaucoup moins de temps sur la saturation (AOI 5) que les experts ($M = 59$ vs $M = 107$ s, $p = 002$, test t apparié à deux variables). Inversement, les novices ont passé beaucoup plus de temps à la vue technique (AOI 3) que les experts ($M = 382$ contre $M = 266$ s, $p = 042$, test t apparié à deux variables). La figure 23, montre une comparaison du nombre de fixations dans les AOI pour les novices et les experts. Les barres d'erreur indiquent des intervalles de confiance à 95%. Les autres comparaisons n'ont pas été statistiquement significatives.

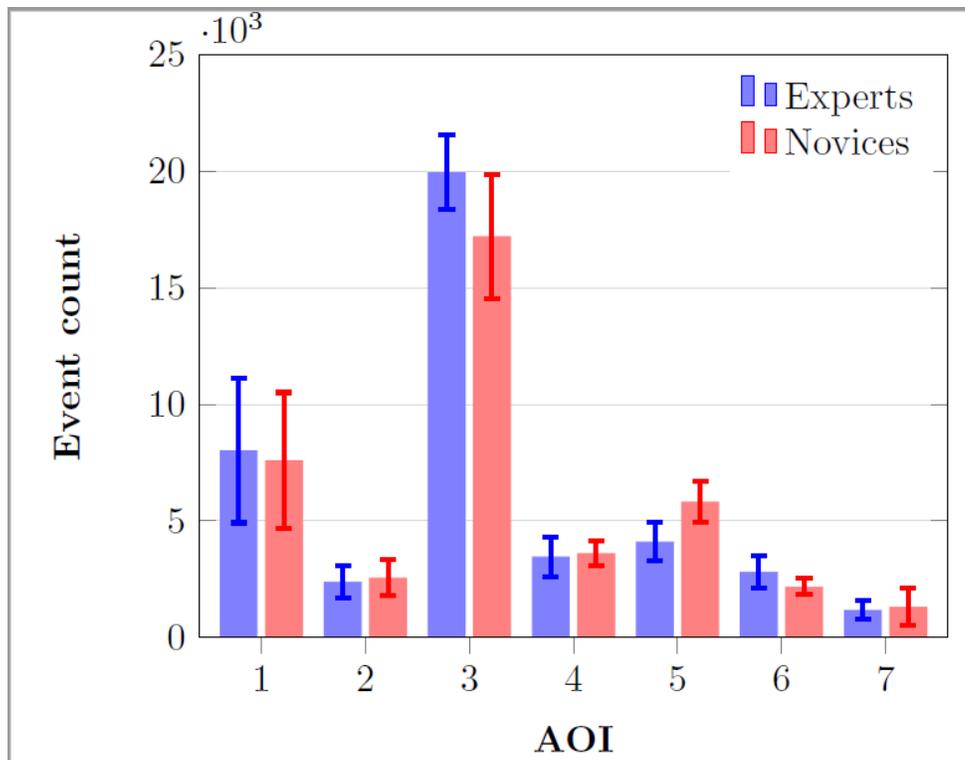


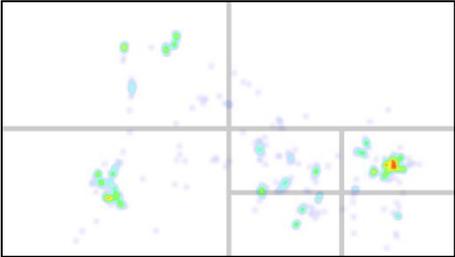
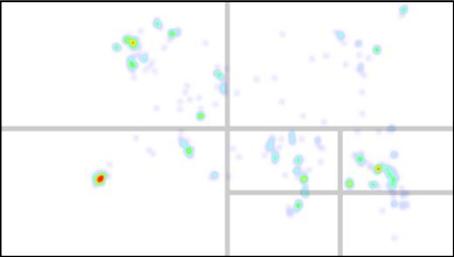
Figure 23 : Comparaison du nombre de fixations dans les différents AOI pour les novices et les experts

Dans l'ensemble, l'analyse préliminaire ne donne pas des résultats très significatifs pour différencier les caractéristiques de la perception visuelle des novices et des experts. Par conséquent, une analyse comportementale a été effectuée dont le but est de faire une analyse plus fine des mouvements oculaires autour des événements clés préalablement présentés.

6.3 Analyse comportementale des mouvements oculaires autour des évènements clés

Une revue de la littérature ainsi que les résultats de nombreuses recherches ont démontré que les perceptions visuelles intentionnelles pour supporter une décision ainsi que le temps de fixation suivi d'une réaction ou décision se produisent généralement autour de 500 ms (Hoang Duc, Bays et al. 2008, L. Orquin et Mueller Loose 2013, Krauzlis, Goffart et al. 2017). Ainsi pour l'analyse comportementale autour des évènements clés nous avons considéré les données de mouvements oculaires correspondant à une durée de 2 secondes soit 1 seconde avant et 1 seconde après chaque évènement clé. Le but est de donner une petite marge avant et après chaque évènement. La liste des évènements "clés" dans la simulation est présentée dans le tableau 3.

En utilisant ces annotations et un programme développé pour d'une part filtrer et agréger les données de mouvements oculaires autour des évènements clés et d'autre part générer des heatmap progressifs, nous avons observé des différences plus saillantes entre les novices et les experts. Nous présentons dans le tableau 5 un extrait des heatmap progressifs obtenus. Les lignes épaisses représentent les bords des zones d'intérêt. Par la suite, dans le tableau 6, nous présentons un résumé des observations sur les heatmap progressifs des mouvements oculaires autour des évènements clés.

| Temps | Novice | Expert |
|---|---|--|
| 01:37 Diminution de la saturation |  |  |

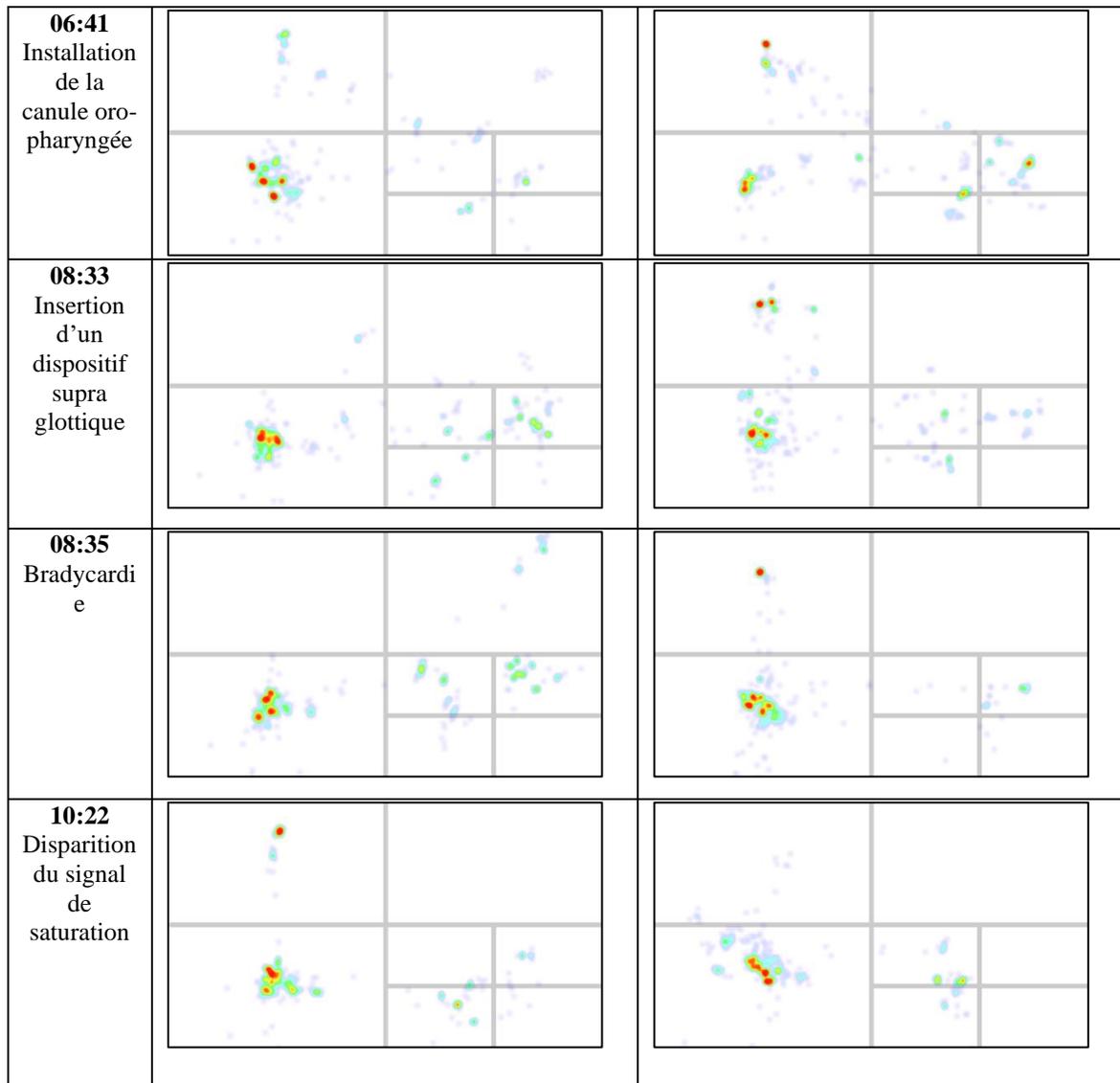


Tableau 5 : Heatmap progressif des mouvements oculaires des novices et des experts autour des évènements clés

Toutes les observations sur les heatmaps progressifs des mouvements oculaires autour des évènements clés sont résumées dans le tableau ci-dessous :

| Évènements clés | Novice | Expert |
|--|--|--|
| 01 :10 Diminution réponse verbale | Tous les AOI | Les signes vitaux (AOI 4 - AOI 7) AOI 3 : Geste technique |
| 01:25 Fermeture des yeux | Plus sur les AOI 1 (Vue globale de la scène 1) et AOI 6 (Pression artérielle) | Plus sur les AOI 4 (Fréquence et rythme cardiaque), AOI 5 (Saturation), AOI 6 (Pression artérielle) |
| 01:37 Diminution de la saturation | Plus sur AOI 5 (Saturation) | Plus sur AOI 3 (Geste technique) |
| 02:09 Hypoventilation | AOI 1 : Vue globale de la scène 1 AOI 3 : Geste technique AOI 5 : Saturation | AOI 3 : Geste technique AOI 5 : Saturation |
| 02:41 Appel à l'aide | AOI 3 : Geste technique AOI 5 : Saturation Partiel AOI 6 : Pression artérielle | AOI 3 : Geste technique AOI 5 : Saturation |
| 03:35 Ventilation au Masque | AOI 1 : Vue globale de la scène 1 AOI 3 : Geste technique AOI 5 : Saturation | AOI 1 : Vue globale de la scène 1 AOI 3 : Geste technique AOI 5 : Saturation |
| 06:41 Installation de la canule oro- pharyngée | AOI 1 : Vue globale de la scène 1 AOI 3 : Geste technique AOI 5 : Saturation | AOI 3 : Geste technique AOI 5 : Saturation |

| | | |
|--|--|--|
| 07:35 Utilisation du Glidescope | AOI 1 : Vue globale de la scène 1 AOI 3 : Geste technique AOI 5 : Saturation | AOI 1 : Vue globale de la scène 1 AOI 3 : Geste technique AOI 5 : Saturation |
| 08 :33 Insertion d'un dispositif supra glottique | AOI 3 : Geste technique AOI4 : Fréquence et rythme cardiaque AOI 5 : Saturation AOI 6 : Pression artérielle | AOI 1 : Vue globale de la scène 1 AOI 3 : Geste technique AOI 4 : Fréquence et rythme cardiaque AOI 5 : Saturation AOI 6 : Pression artérielle |
| 09:39 Déclenchement du Code Bleu | AOI 1 : Vue globale de la scène 1 AOI 3 : Geste technique | AOI 1 : Vue globale de la scène 1 AOI 3 : Geste technique AOI 4 : Fréquence et rythme cardiaque AOI 5 : Saturation |
| 10:22 Disparition du signal de saturation | AOI 3 : Geste technique AOI 1 : Vue globale de la scène 1 | AOI 3 : Geste technique |
| 10 :32 Procédure de Cricotomie | AOI 3 : Geste technique | AOI 3 : Geste technique |

Tableau 6 : Résumé des observations sur les heatmaps progressifs des mouvements oculaires autour des événements clés

Les observations sur les heatmaps progressifs suggèrent que les novices et les experts ont des mouvements oculaires différents la plupart du temps autour des événements

clés. Ces modèles de comportement suggèrent donc que les novices et les experts ont tendance à voir différentes zones pour supporter leur RC au moment des complications cliniques à reconnaître pour une bonne prise en charge clinique et les indices à surveiller pour une meilleure sensibilité situationnelle. Par exemple au moment de :

- La diminution de la saturation (**01:37**) les novices focalisent sur le signe vital problématique alors que les experts focalisent sur les gestes techniques pour résoudre le problème. Avec un patient en CICO (situation d'intubation et de ventilation impossible) ayant une diminution de la saturation de l'hémoglobine au niveau des capillaires sanguins les autres signes vitaux sont certes importants à surveiller, toutefois une prise en charge urgente (par exemple avec un masque facial) est prioritaire.
- L'hypoventilation (**02:09**) qui est un apport d'air insuffisant au niveau des poumons qui provoque une diminution de la quantité d'oxygène. Cette observation qui pourrait être une hypothèse qui explique la diminution de la saturation (**01:37**), est toutefois faite après dans l'axe du temps. Ceci pourrait induire en erreur le novice suivant une approche hypothético-déductive, car elle nécessite plutôt une induction constructive pour supporter la génération des autres phases d'hypothèse.
- L'installation de la canule oro-pharyngée (**06:41**), les experts focalisaient lors des manipulations techniques non seulement sur le patient, mais également sur l'interaction de l'équipe de prise en charge (Vue globale de la scène 1). Cette observation été corroborée à d'autres moments comme à 08h33 (Insertion d'un dispositif supra glottique).

Des observations avec les heatmaps progressifs confirment les intuitions empiriques de la part des instructeurs du CAAHC qui montrent que généralement la situation est prise en charge d'une façon globale et rapidement anticipée par les experts, alors que les novices adoptent une approche hypothético-déductive de manière analytique et

évolutive. Nous nommons les évènements clés autour desquels se produisent les différences significatives les phases de différenciation des mouvements oculaires. Dans la chapitre VIII, nous présenterons l'algorithme MDSE (Multimodal Dynamic Sequence Extraction) qui servira à extraire les motifs comportementaux des experts lors des phases de différenciation, ce qui servira à bâtir le modèle du comportement de l'expert dans le STI.

6.4 Analyse de la trajectoire globale (*scanpath*) des fixations et saccades

Notre objectif dans cette analyse est de visualiser dans une perspective globale la trajectoire des saccades des mouvements oculaires des novices et des experts. Le tableau 7 montre trois exemples (Novice1, Novice2, Novice3 Versus Expert1, Expert2, Expert3) dans chaque groupe. Pour éviter le désordre visuel, chaque figure présente un exemple type de trajectoire en couleur noire au-dessus de toutes les trajectoires. Les trajectoires des novices sont indiquées avec un fond en rouge, celles des experts avec un fond bleu.

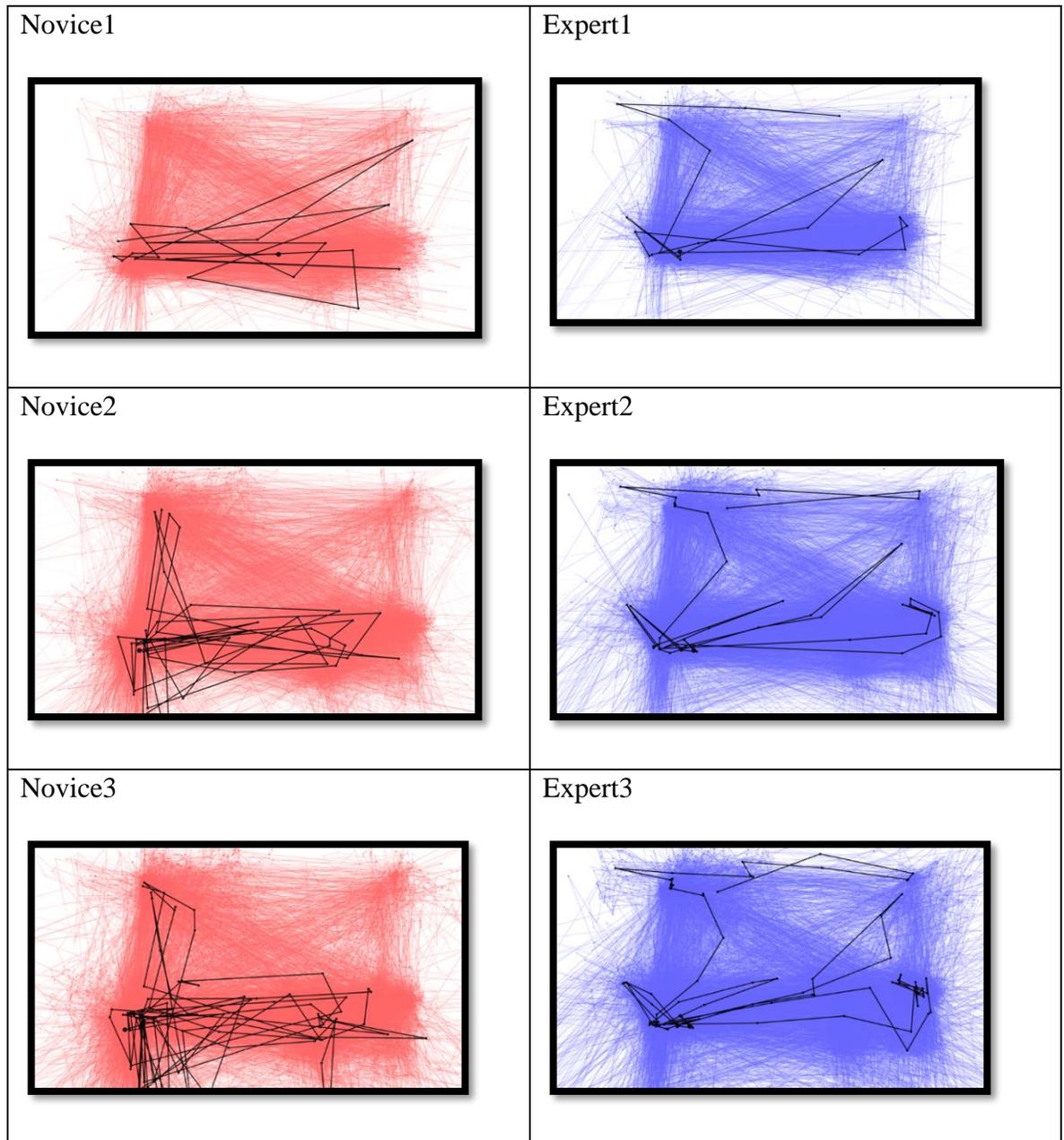


Tableau 7 : Exemple type de trajectoire. Les saccades des novices sont indiquées avec un fond en rouge et celles des experts en bleu

Bien que les figures du tableau 7 soient une représentation statique des mouvements du regard, elles suggèrent les observations suivantes :

1. Les novices ont plus de mouvements oculaires plus saccadés, c'est-à-dire qu'ils ont tendance à déplacer plus souvent leur regard (nombre de fixations, spécifiquement dans l'AOI des gestes techniques) que les experts et à se concentrer moins (temps moyen par fixation) sur chaque partie de l'écran.
2. Les experts présentent des mouvements moins erratiques, en ce sens que leurs zones de regard sont globalement plus étroites et par la suite font des saccades plus grandes.

Afin d'obtenir des métriques sur ces observations des *scanpath*, nous avons utilisé l'algorithme de Smith-Waterman (Smith et Waterman 1981). C'est un algorithme en bio-informatique pour étudier l'alignement des séquences dans une chaîne d'ADN. Les séquences biologiques sont traitées comme des chaînes de caractères et l'algorithme cherche à optimiser l'alignement de deux chaînes en prenant une chaîne comme référence (X) et l'autre pour la comparaison (Y). L'alignement se fait par insertion et substitution afin de maximiser le nombre de caractères qui coïncident entre X et Y. La recherche de l'alignement optimal entre X et Y se fait en deux phases. La première par le calcul d'une matrice M des scores d'alignement partiel et la seconde par la construction de l'alignement à partir de la matrice M.

Il existe d'autres algorithmes d'alignement de séquences, toutefois l'avantage de Smith-Waterman est qu'il garantit l'optimalité de la solution. L'implémentation de l'algorithme de Smith-Waterman se fait par programmation dynamique, ce qui fait qu'il est complexe et coûteux en temps d'exécution. Il existe des variantes meilleures en temps d'exécution comme l'algorithme Myers&Miller (Myers et Miller 1988) qui sont toutefois très spécifiques. Finalement, une autre approche possible serait de passer par la distance de Levenstein pour calculer la distance entre deux séquences oculaires et par la suite minimiser cette distance à des fins de comparaison.

L'algorithme de Smith-Waterman a été utilisé par (Khedher, Jraidi et al. 2018) pour la comparaison de données oculaires de cliniciens novices à une séquence oculaire de référence considérée comme optimale. Dans notre recherche, nous n'avons pas à priori une séquence optimale. Notre objectif dans cette étape d'analyse exploratoire et comparative a été de faire une analyse comparative croisée des mouvements oculaires entre : Expert Vs Expert et Novice Vs Novice. Ceci permettra d'obtenir une mesure objective et numérique des similarités/différences dans les mouvements oculaires.

Comme approche, nous avons fait l'alignement temporel des séquences issues de la dynamique des mouvements oculaires (fixation et saccades) en optimisant un score de similitude dérivé des caractéristiques suivantes :

- Le nombre de correspondances entre les fixations dans les AOI
- Le nombre de disparités, donc nécessite une substitution pour avoir une équivalence
- Le nombre d'insertions et de suppressions pour avoir une équivalence.

La valeur du score de similitude est plus élevée pour deux séquences presque similaires. Plusieurs étapes de prétraitements intermédiaires ont été faites et nous présentons dans l'annexe B un extrait de séquences prétraitées pour l'exécution de l'algorithme de Smith-Waterman.

Les différentes valeurs croisées obtenues via l'algorithme de Smith-Waterman ont été normalisées pour avoir des valeurs entre 0 et 1. Par la suite, pour la visualisation, les scores normalisés ont été mis dans des matrices de corrélation sous forme de carte thermique :

- Les teintes de vert indiquent de fortes similitudes
- Les teintes de jaune indiquent de faibles similitudes
- Les teintes de rouge indiquent des séquences quasi différentes

Le tableau 8 présente les scores de similitudes des scanpaths des novices obtenus via l'algorithme de Smith-Waterman et mis dans une matrice sous forme de carte thermique. Le tableau 9 présente la même matrice pour les experts. Sur la base des

différentes valeurs croisées, nous concluons qu'il y a de fortes similitudes dans le mouvement oculaire des experts. Cependant, pour les novices, les similitudes sont faibles ou les séquences quasi différentes. Avec ces résultats de scores de similitudes des séquences des experts, nous nous attendons à avoir des motifs de mouvements oculaires généralement similaires et distinctifs des experts.

| | Novice1 | Novice2 | Novice3 | Novice4 | Novice5 | Novice6 | Novice7 |
|---------|---------|---------|---------|---------|---------|---------|---------|
| Novice1 | NA | 0,01 | 0,03 | 0,04 | 0,04 | 0,06 | 0,03 |
| Novice2 | 0,1 | NA | 0,15 | 0,15 | 0,17 | 0,17 | 0,15 |
| Novice3 | 0,03 | 0,15 | NA | 0,47 | 0,60 | 0,64 | 0,60 |
| Novice4 | 0,04 | 0,15 | 0,47 | NA | 0,52 | 0,55 | 0,48 |
| Novice5 | 0,04 | 0,17 | 0,60 | 0,52 | NA | 0,21 | 0,80 |
| Novice6 | 0,06 | 0,17 | 0,64 | 0,55 | 0,21 | NA | 0,18 |
| Novice7 | 0,03 | 0,15 | 0,60 | 0,48 | 0,80 | 0,18 | NA |

Tableau 8 : Tableau croisé des scores de similitudes des scanpth des novices obtenus via l'algorithme de Smith-Waterman

| | Expert1 | Expert2 | Expert3 | Expert4 | Expert5 | Expert6 | Expert7 |
|---------|---------|---------|---------|---------|---------|---------|---------|
| Expert1 | NA | 0,81 | 0,7 | 0,81 | 0,88 | 0,7 | 0,82 |
| Expert2 | 0,81 | NA | 0,7 | 0,81 | 0,72 | 0,82 | 0,73 |
| Expert3 | 0,7 | 0,7 | NA | 0,88 | 0,90 | 0,75 | 0,90 |
| Expert4 | 0,81 | 0,81 | 0,88 | NA | 0,68 | 0,60 | 0,65 |
| Expert5 | 0,88 | 0,72 | 0,90 | 0,68 | NA | 0,65 | 0,91 |
| Expert6 | 0,7 | 0,82 | 0,75 | 0,60 | 0,65 | NA | 0,67 |
| Expert7 | 0,82 | 0,73 | 0,90 | 0,65 | 0,91 | 0,67 | NA |

Tableau 9 : Tableau croisé des scores de similitudes des scanpth des experts obtenus via l'algorithme de Smith-Waterman

6.5 Conclusion et discussion

Dans cette partie nous avons présenté le protocole expérimental, exposé le déroulement des expérimentations ainsi que la description des approches et outils pour la collecte, la validation et l'analyse des données. Par la suite, nous avons effectué des analyses exploratoires et comparatives en utilisant des techniques de visualisation de données. Les éléments visuels obtenus (graphiques, des cartes, etc.) ont permis de faire quelques observations qui fournissent des indices qui serviront de base pour des analyses plus poussées avec des techniques d'apprentissage machine.

1. Analyse préliminaire : le but est de comparer les novices aux experts en utilisant des statistiques descriptives sur les fixations dans les AOI. Il y a quelques observations intéressantes sur les similarités et différences. Toutefois, nous obtenons peu de résultats statistiquement significatifs pour différencier les caractéristiques des fixations.
2. Analyse comportementale : le but est de faire une analyse plus fine des mouvements oculaires autour d'évènements clés bien identifiés. Les observations des mouvements oculaires sur les heatmaps progressifs fournissent des indices que généralement la situation est prise en charge d'une façon globale et rapidement anticipée par les experts, alors que les novices adoptent une approche analytique et évolutive.
3. Analyse globale des trajectoires des fixations et saccades (*scanpath*): les observations sur les représentations statiques des formes caractéristiques des saccades montrent que les novices ont des mouvements oculaires plus saccadés. Toutefois les experts ont globalement des mouvements oculaires plus étroites et par la suite font des saccades plus grandes. Les scores de comparaisons de similitudes effectuées avec l'algorithme de Smith-Waterman prouvent que les experts ont des mouvements oculaires assez similaires, alors que les novices ont beaucoup de différences.

Les observations issues de l'analyse préliminaire, comportementale et des trajectoires des fixations/saccades fournissent globalement quelques pistes d'interprétations. En effet la rapidité avec laquelle les hypothèses initiales du RC sont générées est une caractéristique du comportement de l'expert (Joseph et Patel 1990). Il a tendance à réduire assez vite l'incertitude sur les hypothèses candidates alors que les novices utilisent une stratégie similaire à la recherche en profondeur. Les novices considèrent et évaluent chaque hypothèse (Arocha, Patel et al. 1993) à partir desquelles ils génèrent d'autres hypothèses en collectant activement toutes les informations pour soutenir leur RC (Patel, Arocha et al. 1994). Dans le cas précis de notre expérimentation, les sources d'informations sont les AOI et les mouvements oculaires le moyen pour collecter ces informations. Dès lors les mouvements oculaires (saccade, fixation) et les observations des graphiques reflètent ces différences comportementales. L'expert a globalement ce que nous appelons la perception visuelle axée sur la vérification. Le but est de rassembler des informations spécifiques sur l'environnement, cela impliquera typiquement des mouvements oculaires moins erratiques, mais avec des saccades plus larges. Au contraire, si la perception visuelle visant à soutenir activement les décisions, c'est-à-dire à analyser l'environnement en mobilisant des connaissances spécifiques, ceci implique généralement plus de mouvements oculaires saccadés et erratiques que nous appelons la perception visuelle guidée par la décision. C'est le cas des novices. En prenant en considération les analyses préliminaires, comportementales et des trajectoires; les observations nous suggèrent que nous pourrions construire des modèles prédictifs de classification qui prendront en compte la dynamique des mouvements oculaires au cours du temps. Étant donné une séquence particulière de mouvements oculaires, est-elle susceptible d'être un comportement de novice ou expert? Il s'agit d'une tâche de classification binaire d'extrait de séquences de perception visuelle provenant de l'enregistrement des vidéos des participants. Également, est-il possible d'extraire les séquences perceptuelles caractéristiques (similaires) et différentes entre

novices et experts. Dans la partie IV, nous répondrons à ces questions en présentant l'architecture DeepConvLSTM pour la classification des séquences et l'algorithme MDSE (Multimodal Dynamic Sequence Extraction) qui servira à extraire les motifs comportementaux qui serviront à bâtir le modèle de l'expert pour le STI.

PARTIE 3 - ANALYSE DES DONNÉES : CLASSIFICATION DES SÉQUENCES,
EXTRACTION DE MOTIFS

CHAPITRE VII

CLASSIFICATION DES SÉQUENCES - DEEPCONVLSTM VS WKM-1NN

Dans ce chapitre, nous avons exploré la problématique suivante : pouvons-nous classifier les cliniciens novices et experts en analysant seulement la séquence de leurs mouvements oculaires? Plus spécifiquement, étant donné un extrait de séquence de mouvements oculaires provenant d'un participant à l'expérimentation, est-ce qu'il s'agit d'un comportement perceptuel d'un expert ou d'un novice. Le classifieur obtenu est utilisé par le STI EyeLab pour classifier des sous-séquences de mouvements oculaires des cliniciens pour une mise à jour dynamique du modèle comportemental de l'apprenant.

7.1 Introduction

De précédentes recherches ont démontré que l'analyse de la perception visuelle peut être utilisée pour catégoriser les novices et les experts dans plusieurs domaines.

- Plus concrètement, Hermens et al. (Hermens, Flin et al. 2013) donnent un aperçu de plusieurs études comparant les mouvements oculaires chez les chirurgiens novices et experts.

- D'autres chercheurs ont démontré que les mouvements oculaires pour l'interprétation d'un artefact clinique, par exemple une image de radiographie dépendrait de l'expertise du clinicien (Cooper, Gale et al. 2009).

Cependant, à notre connaissance il n'y a pas de travaux antérieurs dans le contexte des mouvements oculaires pour supporter le raisonnement clinique en anesthésiologie.

L'analyse préliminaire et exploratoire des données a permis d'observer de subtiles, mais importantes différences dans la perception visuelle tout au long de la progression du cas clinique. Également, les heatmaps progressifs nous montrent que les novices ont parfois un comportement similaire à celui des experts. Ces observations nous confirment qu'il serait possible de bâtir un modèle prédictif qui permette de classer les cliniciens novices et experts en utilisant les données provenant des séquences de mouvements oculaires.

7.2 Nature des données

Avec une configuration de 60Hz, l'oculomètre (Tobii) collecte des points de données des mouvements oculaires toutes les 17ms. Chaque point de données est identifié par un *timestamp* (estampille temporelle) et les coordonnées : abscisse x et ordonnée y. Ces points sont traités et agrégés dans différentes catégories de mouvements oculaires à savoir les fixations dans les différents AOI et les saccades. Avec une durée de 13 minutes pour la simulation, nous obtenons un total de 46000 points de données par participant. Considérer ces 46000 points de données comme s'il s'agissait d'une séquence unique serait sans grand intérêt pour les objectifs de la recherche, nécessiterait un grand nombre de données d'entraînement et donnerait des modèles difficilement généralisables. Étant donné que l'objectif est de déterminer si des extraits de séquences oculaires durant la performance reflètent un comportement de novice ou d'expert; chaque séquence complète a été découpée en de plus petites séquences de

taille égale. Pour les expérimentations, nous avons utilisé des tranches de séquence de longueur $s = 500$. Ceci représente environ 8 secondes de séquence oculaire hormis les quelques erreurs d'indexation temporelles dans les données générées par l'oculomètre.

Chaque tranche de séquence est considérée comme une instance indépendante. Chaque séquence est une série temporelle de taille fixe composée de 500 points de données successives. Les points de données représentent l'AOI regardé à chaque 17ms. Dans chaque point de données, la fixation sur un AOI en cet instant est indiquée par la valeur de 1. Le tableau 10 montre une illustration de séquence utilisée pour la classification. Une fixation plus ou moins prolongée sur un AOI se manifeste par des valeurs consécutives de 1 dans cet AOI. Les saccades sont déterminées par les changements de valeurs (0 et 1) dans les différents AOI dans le temps. Par exemple, dans le tableau 10, de l'instant représenté par le point de données 6 à l'instant 7, il y a une saccade de l'AOI-4 à l'AOI-1.

Dans le protocole expérimental, un seuil de 75% a été fixé pour la qualité des données oculaires collectées par l'oculomètre. Ceci signifie qu'environ 25% des données pourront être manquantes ou imprécises. Les facteurs qui conditionnent la qualité des données ne sont pas prévisibles à priori. Ceci rend quasi impossible un prétraitement temps réel des séquences pour la classification. Il y a par exemple des facteurs environnementaux (par exemple la luminosité), la distance du participant par rapport à l'écran de l'oculomètre, la distraction du participant (regarder en dehors des AOI) et la sensibilité/précision dans les réglages de l'oculomètre. Dans le tableau 10, les points de données 5 et 9 représentent des données manquantes. Le point de données 6 est possiblement une donnée imprécise, car elle présente des caractéristiques de perturbations que nous avons nommées spikes.

| Participant | Point de données | AOI-1 | AOI-2 | AOI-3 | AOI-4 | AOI-5 | AOI-6 | AOI-7 |
|-------------|------------------|-------|-------|-------|-------|-------|-------|-------|
| P1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 3 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 6 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| P1 | 7 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 8 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 10 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| P1 | ... | ... | ... | ... | ... | ... | ... | ... |
| P1 | 500 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |

Tableau 10 : Illustration d'une séquence de donnée utilisée pour l'apprentissage et la classification

Somme toute, pour la tâche de classification, les données oculaires complètes ont été divisées en séquences. La modélisation des séquences permet de préserver les caractéristiques principales des séquences perceptivo-décisionnelle à savoir :

- Les caractéristiques spatiales : les fixations et saccades dans différents AOI
- Les caractéristiques séquentielles : l'ordre temporel dans laquelle les différents mouvements oculaires ont été exécutés
- Les caractéristiques de manques de précision et bruits imprévisibles

Échantillonnage

Des techniques d'échantillonnage sont utilisées pour subdiviser les données pour faire d'une part l'apprentissage et d'autre part une validation du modèle appri. Nous avons créé une partition d'apprentissage avec 80% des séquences et les 20% restantes sont utilisés pour tester chaque modèle. Avec 46000 points de données par participant et pour des séquences de longueur 500 nous obtenons 92 séquences par participant. Pour les 14 participants, nous obtenons un total de 1288 séquences. Le tableau 11 résume la structure des données pour l'apprentissage machine.

| Partition | Novices | Experts | Total |
|--------------|---------|---------|-------|
| Entraînement | 515 | 515 | 1030 |
| Test | 129 | 129 | 258 |
| Total | 644 | 644 | 1288 |

Tableau 11 : Nombre de séquences de mouvement oculaires par classe pour l'ensemble des données

7.3 Choix de modèle et classification des séquences

La classification des données séquentielles présente une certaine complexité, car il faudrait préserver l'ordre des attributs. En faisant une combinaison judicieuse d'algorithmes en apprentissage machine nous obtenons quelques approches intéressantes au problème de la classification des séquences oculaires. Nous avons comparé deux catégories d'approches dont la première est fondée sur des techniques d'apprentissage connexionniste et la seconde sur des techniques d'apprentissage statistique.

7.3.1 Approches fondées des techniques d'apprentissage connexionniste

Les données oculométriques qui sont utilisées pour la classification sont d'une part de nature spatiale avec des fixations et saccades sur les images et d'autre part séquentielle en considérant l'ordonnement temporel dans lesquelles les AOI sont regardés. Il est possible de combiner adéquatement différentes architectures de RNA pour obtenir une architecture profonde spécialisée. En prenant en considération les caractéristiques spatiales et temporelles des données oculométriques pour effectuer la classification, une architecture profonde constituée d'un réseau de neurones à convolution et d'un LSTM pourrait être un modèle adéquat.

Réseau de neurones à convolution

Les réseaux de neurones à convolution (CNN ou ConvNet : Convolutional Neural Networks) désignent une sous-catégorie de réseaux de neurones à propagation avant spécialement conçus pour la vision par ordinateur. En 2012 le CNN nommée AlexNet s'illustre avec très bonne performance au ILSVRC (ImageNet Large Scale Visual Recognition Challenge) organisé par le projet ImageNet. Un CNN est un RNA avec une architecture profonde, donc effectue l'extraction d'un ensemble attribut (*features*) pour effectuer par exemple une tâche de classification d'image. Un CNN est constitué de quatre types de couches spécialisés à savoir :

- La couche de convolution : c'est la couche d'entrée qui reçoit les images et détecte les attributs de chacune d'entre elles en utilisant une opération de convolution. Une convolution est une opération mathématique sur deux fonctions pour produire une fonction agrégée, intégrant des caractéristiques d'une fonction avec l'autre. La valeur convolutive la plus élevée est atteinte lorsque le noyau est similaire à une partie particulière de l'entrée. La sortie est une carte d'activation qui indique l'emplacement des attributs dans l'image.
- La couche de pooling : prend en entrée la carte d'activation, et a pour objectif de réduire la taille des images, tout en préservant leurs caractéristiques

importantes. Le principe consiste à faire glisser une petite fenêtre pas à pas sur toutes les parties de l'image et prendre la valeur maximum ou moyenne à chaque pas. Une taille de fenêtre de 2 ou 3 est généralement utilisée.

- La couche de correction ReLU : Cette couche est nommée en référence à la fonction d'activation ReLU (Rectified Linear Units : Unité de rectification linéaire). C'est une fonction réelle non-linéaire définie par $\text{ReLU}(x) = \max(0, x)$ qui sert de fonction d'activation. Avec l'activation ReLU, à chaque fois qu'il y a une valeur négative dans un pixel, on la remplace par un 0 et toute valeur positive par elle-même. Le principe est simple, mais offre plusieurs avantages sur un plan fondamental. C'est une fonction dérivable et permet la rétropropagation, simple à calculer et empêche les valeurs apprises de rester coincées autour de 0 ou d'exploser vers l'infini. Elle présente toutefois l'inconvénient du problème du ReLU mourant, car pour des valeurs négatives ou proches de 0, le gradient de la fonction est nul. Dans ce cas le réseau ne peut pas effectuer la rétropropagation efficacement. Plusieurs variantes de ReLU ont été proposées pour pallier ce problème dont le Leaky ReLU, le SineReLU, etc.
- La couche « entièrement connectée » qui est une couche de type perceptron qui permet de faire la classification.

Les RNA de type LSTM

Les réseaux récurrents sont convenables pour le traitement de données séquentielles ou temporelles, mais souffrent d'un problème d'explosion ou de dissipation du gradient lors du processus d'apprentissage (Bengio, Pascanu et al. 2013). Il y a également d'autres défis à considérer dans le traitement des données séquentielles comme la dépendance à long terme. Pour adresser ces problèmes, les réseaux récurrents à mémoire court et long terme (LSTM : Long short-term memory LSTM) permettent le blocage des entrées et des sorties à travers le réseau ainsi que la remise à zéro de l'état de certains neurones grâce à l'utilisation de portes d'entrée (input gates), de portes de

sortie (output gates) et de portes d'oubli (forget gates) respectivement. Le LSTM crée une connexion entre les activations de la porte d'oubli et le calcul des gradients, cette connexion crée un chemin pour le flux d'informations à travers la porte d'oubli pour les informations que le LSTM ne doit pas oublier. Ce mécanisme de fonctionnement peut être décrit à l'aide de l'ensemble d'équations présenté ci-dessous.

$$\begin{aligned}
 z^{(t)} &= g(W_z x^{(t)} + R_z \hat{y}^{(t-1)} + b_z) \\
 i^{(t)} &= \sigma(W_i x^{(t)} + R_i \hat{y}^{(t-1)} + p_i \odot c^{(t-1)} + b_i) \\
 f^{(t)} &= \sigma(W_f x^{(t)} + R_f \hat{y}^{(t-1)} + p_f \odot c^{(t-1)} + b_f) \\
 c^{(t)} &= i^{(t)} \odot z^{(t)} + f^{(t)} \odot c^{(t-1)} \\
 o^{(t)} &= \sigma(W_o x^{(t)} + R_o \hat{y}^{(t-1)} + p_o \odot c^{(t-1)} + b_o) \\
 \hat{y}^{(t)} &= o^{(t)} + \odot h(c^{(t)})
 \end{aligned}$$

Équation : Ensemble d'équations du mécanisme de fonctionnement d'un LSTM

Dans cette équation le symbole \odot note la multiplication ponctuelle de deux vecteurs. Les fonctions σ , g et h sont des fonctions non linéaires, W et R sont des matrices de poids et tous les termes b sont des termes de biais. Les points suivants sont à noter:

1. L'élément le plus important de la LSTM est l'état de la cellule désigné par $c^{(t)} = i^{(t)} \odot z^{(t)} + f^{(t)} \odot c^{(t-1)}$. L'état de la cellule est mis à jour en fonction de l'entrée de bloc $z^{(t)}$ et de l'état de la cellule précédente $c^{(t-1)}$. La porte d'entrée $i^{(t)}$ détermine quelle fraction de l'entrée du bloc se trouve dans l'état de la cellule (donc appelée porte). La porte d'oubli $f^{(t)}$ détermine la quantité de l'état de la cellule précédente à conserver.
2. La sortie $\hat{y}^{(t)}$ est déterminée en fonction de l'état de la cellule $c^{(t)}$ et de la porte de sortie $o^{(t)}$, qui détermine dans quelle mesure l'état de la cellule affecte la sortie.

3. Le terme $z^{(t)}$ est appelé entrée de bloc et produit une valeur basée sur l'entrée courante et la sortie précédente.
4. Le terme $i^{(t)}$ est appelé porte d'entrée. Il détermine la quantité d'entrées à conserver dans l'état de la cellule $c^{(t)}$.
5. Tous les p termes sont des connexions par judas, ce qui permet à une fraction de l'état de la cellule de prendre en compte dans le calcul du terme en question.
6. Le calcul de l'état de la cellule $c^{(t)}$ ne rencontre pas le problème de dissipation du gradient (c'est ce qu'on appelle le carrousel d'erreurs constant). Cependant, les LSTMs peuvent être affectés par l'explosion du gradient, l'écrêtage de gradient est donc utilisé pendant l'entraînement.

Il existe plusieurs variantes du modèle de base comme :

- Encoder-Decoder LSTM : un réseau LSTM code les séquences d'entrée et un autre réseau LSTM décode l'encodage en une séquence de sortie
- Bidirectional LSTM : les séquences d'entrée sont présentées et apprises à la fois en avant et en arrière
- Generative LSTM : le LSTM apprend si bien la relation de structure dans les séquences d'entrée qu'il peut générer de nouvelles séquences plausibles.

Nous adoptons le LSTM pour les séquences, car il a surpassé de nombreuses autres approches dans une variété de tâches impliquant le traitement de données séquentielles. Nous pouvons citer entre autres l'écriture manuscrite (Graves, Fernandez et al. 2007) et la reconnaissance vocale (Graves, Mohamed et al. 2013).

Architecture profonde DeepConvLSTM

Les paramètres intrinsèques aux données, les objectifs du projet et l'état de l'art dans les architectures profondes ont été considérés pour faire ce choix de combiner dans une architecture profonde un CNN et un LSTM (Ordóñez et Roggen 2016). Nous avons développé une architecture de réseau de neurones profond constitué d'un bloc convolutif suivi d'un bloc récurrent comme présenté à la figure 25. L'entrée du réseau est comme dans le modèle de donnée c'est-à-dire une série temporelle (une séquence de vecteurs 2D avec 500 points), alors que la sortie du réseau est soit novice, soit expert.

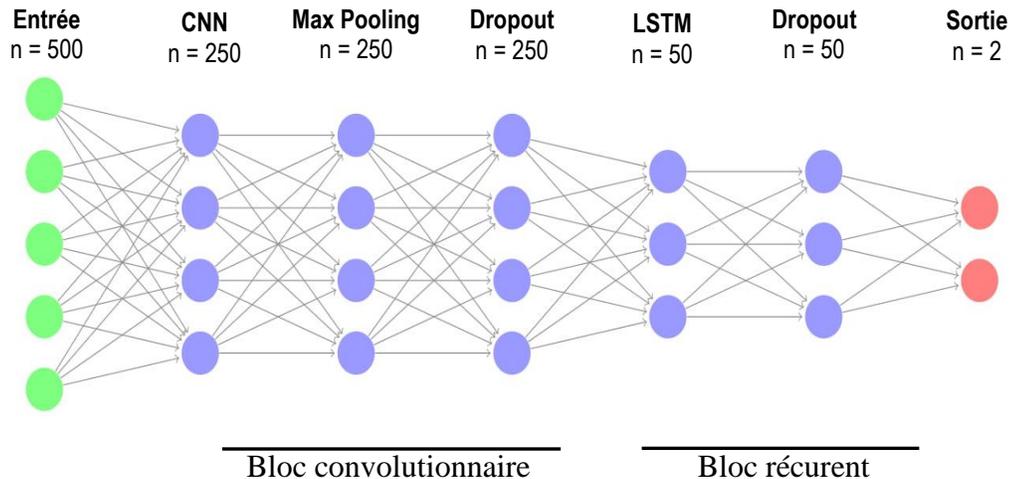


Figure 24 : Topologie de l'architecture DeepConvLSTM

Le dessin est simplifié pour illustrer la topologie. Chaque dimension de couche (n) est indiquée sous son titre.

Inspiré par un travail récent qui a appliqué avec succès des CNN pour la modélisation des séquences (Bradbury, Merity et al. 2016), nous ajoutons une couche de convolution temporelle unidimensionnelle à l'entrée du réseau, suivie d'une couche de Max Pooling. Les attributs consolidés à la fin de ce processus servent à alimenter le LSTM. En d'autres termes, la couche CNN apprend des caractéristiques spatiales qui sont ensuite

appries en tant que séquences par la couche LSTM. De cette façon, nous combinons les propriétés d'apprentissage de la structure spatiale des CNN avec l'apprentissage séquentiel des LSTM.

La couche de Max Pooling est en soi un processus de discrétisation visant trois objectifs :

1. Réduire la dimensionnalité d'entrée, en filtrant la représentation de données initiale;
2. Éviter le surapprentissage, en fournissant une forme abstraite de la représentation des données;
3. Réduire le coût de calcul, en réduisant le nombre de paramètres à apprendre.

Les deux couches CNN et Max Pooling ont une taille de noyau de 2. La couche LSTM est entièrement connectée à 50 cellules. Les couches de dropout ont une probabilité de 0,2 puisque c'est la valeur recommandée pour la plupart des cas en apprentissage machine (Srivastava 2013). Ces couches ont pour effet de réduire le surapprentissage et d'améliorer les performances du modèle. Nous avons entraîné le réseau de neurones avec 60 époques et une taille de lot de 56 (mini-batch) sur un processeur i7 à 3,30 GHz avec 16 Go de RAM. Après chaque époque, le modèle est évalué par rapport à la partition de test, pour avoir une idée de la performance du modèle pendant l'entraînement, après quoi les données sont sauvegardées dans des fichiers de log pour l'époque suivante. Le modèle a été optimisé en utilisant l'algorithme ADAM (Kingma et Ba 2014) avec une entropie croisée binaire comme fonction de perte.

Résultats

Le tableau 12 résume les résultats en termes de précision de classification, d'intervalles de confiance calculés selon la méthode de Wilson pour les distributions binomiales (Wilson 1927) et l'aire sous la courbe ROC qui est une mesure normalisée de la performance d'un classificateur (AUC : Area under the ROC Curve). Plus précisément,

l'AUC est la probabilité qu'un classificateur classe une instance positive quelconque mieux qu'une instance négative choisie au hasard. Plus l'AUC est proche de 1, meilleur est le classement.

| Précision | Int de Conf. 95% | AUC |
|-----------|------------------|------|
| 90,2 | 91,4 | 0,92 |

Tableau 12 : Résumé des résultats de l'expérimentation

Comme observé sur la figure 27, la performance obtenue par le modèle est remarquable. La forme en coude apparaît dans les tâches de classification binaire; c'est-à-dire lorsque les données et les modèles de prédictions sont tous deux binaires.

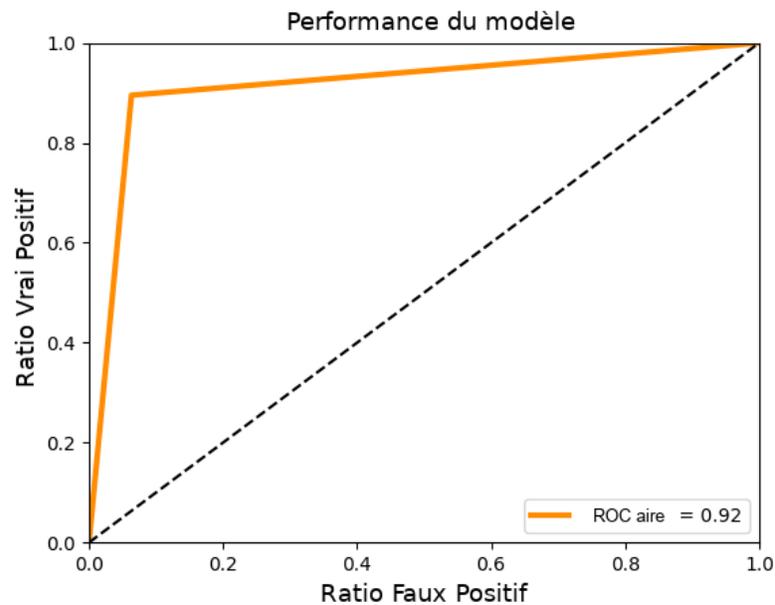


Figure 25 : Performance obtenue par l'architecture DeepConvLSTM

D'autre part, sur la figure 27, nous pouvons observer que le modèle a commencé à converger après 50 époques

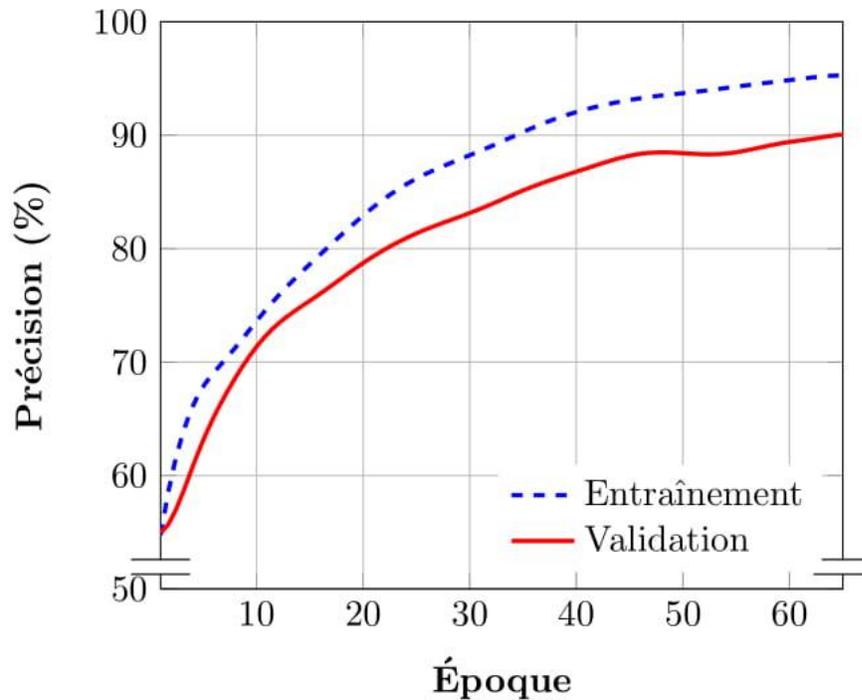


Figure 26 : Convergence obtenue par l'architecture DeepConvLSTM

7.3.2 Approches fondées des techniques d'apprentissage statistique

Nous discuterons brièvement ici de deux principales catégories d'approches de classification que nous nommons une classification basée sur les vecteurs d'attributs et une classification basée sur un modèle des données.

1. Classification basée sur les vecteurs d'attributs

Dans la classification basée sur les vecteurs d'attributs, une séquence de données est transformée en un vecteur d'attributs en utilisant une procédure de sélection d'attributs (*feature selection*) descriptifs des caractéristiques des séquences. Pour appliquer ces techniques de sélection d'attributs sur des données de nature séquentielle, il faudrait d'une part une discrétisation des données et d'autre part des métriques de distance pour mesurer la similitude entre les séquences.

Le choix des métriques de distance est essentiel à la performance du classificateur qui sera utilisé. Pour la classification simple des séries chronologiques, la distance euclidienne est une option largement adoptée (Xing, Pei et al. 2010). Toutefois la distance euclidienne est sensible aux distorsions dans la dimension temporelle. La déformation temporelle dynamique (DTW : *Dynamic Time Warping*) est proposée pour surmonter ce problème de la distance euclidienne. De plus, le DTW ne nécessite pas deux séries chronologiques de même longueur (Keogh et Pazzani 2000). L'implémentation de DTW nécessite une programmation dynamique et présente une complexité algorithmique quadratique. Par conséquent, il est coûteux en calcul sur un grand ensemble de données. Des variantes plus performantes comme le DSW (Dynamic State Warping) ont été proposées par (Gong et Chen 2018).

Après avoir obtenu une représentation vectorielle des données séquentielles, la classification pourrait se faire avec de nombreuses approches bien établies. Nous avons entre autres les machines à vecteurs de support (SVM : Support Vector Machines), plusieurs variantes des arbres de décision, la régression logistique et un perceptron multicouche :

- Les SVM (Vapnik 1995) sont sans doute un des classificateurs les plus puissants pour une tâche de classification binaire, car elles permettent de travailler sur des données de grandes dimensions, nécessitent un faible nombre d'hyperparamètres et ont un fondement théorique solide qui garantit un résultat optimal. Le SVM fournit un hyperplan de séparation maximal dans un espace vectoriel comme frontière de décision entre les 2 classes. Pour des données non-

linéairement séparables, l'astuce du noyau (kernel trick) permet une redescription des données.

- Les arbres de décision permettent d'extraire des règles logiques, de cause à effet, un ensemble de choix sous la forme graphique pour la classification. L'avantage réside donc dans le fait d'obtenir un résultat interprétable plus facilement pour supporter les prises de décision.
- La régression logistique ou modèle logit qui est une régression binomiale qui permet de prédire une sortie catégorielle à partir de variables prédictives (variables explicatives) contrairement à la régression linéaire.

2. Classification basée sur le modèle de données

Cette catégorie suppose que les données appartenant à une classe sont générées par un modèle sous-jacent M . Le modèle M définit donc la distribution de probabilité des séquences dans la classe. Habituellement, un modèle est défini sur la base de certaines hypothèses et les distributions de probabilité sont décrites par un ensemble de paramètres. Dans l'étape d'apprentissage, les paramètres de M sont appris. Dans l'étape de classification, une nouvelle séquence est affectée à la classe avec la plus forte vraisemblance.

Le classifieur Naive Bayes est un exemple simple de cette catégorie. Il fait l'hypothèse que les attributs représentatifs des données séquentielles sont indépendants les uns des autres. Les probabilités conditionnelles des attributs d'une classe sont apprises lors de la phase d'entraînement. Cependant, l'hypothèse d'indépendance requise par Naive Bayes n'est pas souvent garantie dans la pratique.

Un modèle de Markov caché (HMM : Hidden Markov Model) peut modéliser la dépendance entre les éléments dans les séquences (Baum et Petrie 1966, Rabiner et Juang 1986). Les HMM supposent toutefois que le système modélisé est un processus markovien avec des états non observés, où l'état est décrit par une seule variable aléatoire discrète. Dans l'étape de classification, une séquence inconnue est alignée

avec le profil HMM dans chaque classe par programmation dynamique. Une nouvelle séquence sera donc classée dans la classe qui a le score d'alignement le plus élevé. Les HMM fournissent donc des structures flexibles qui peuvent modéliser des sources complexes de données séquentielles. Cependant, les HMM supposent une série d'hypothèses (par exemple l'indépendance conditionnelle), ce que les RNA n'imposent pas. De plus, les HMM ne peuvent gérer qu'un nombre limité de dépendances d'étape, tandis que les LSTM peuvent gérer des dépendances à long terme.

Architecture WKM-1NN

L'architecture d'apprentissage utilisée est constituée de l'algorithme WKM et d'un classifieur k-NN. WKM (warped K-means) est une variante de K-means pour catégoriser des données séquentielles basées sur les distances. WKM est un algorithme rapide et combiné à un classifieur; il a surpassé les approches comparables dans des tâches de la classification des séquences (Leiva et Vidal 2013). En plus de fournir une représentation compacte des séquences de données, WKM les rend robustes au bruit ou aux distorsions dans les données. L'algorithme WKM capitalise sur la séquentialité des données et commence par une partition initiale appropriée, en utilisant une interpolation linéaire par morceaux, ce qui conduit à une partition initiale des données non répartie linéairement. Ensuite, WKM itère sur les points de données en utilisant une procédure d'optimisation.

L'architecture d'apprentissage utilisée WKM comprime les données séquentielles de longueur 500 en c segments homogènes disjoints (ou unités élémentaires) avec $1 < c \ll n$. Ensuite, le centroïde de chaque segment est utilisé comme entrée d'un classifieur 1-NN. Le nombre de blocs c doit être fourni en entrée. Par conséquent, parce que le c optimal pour la classification est inconnu à l'avance, nous avons testé différentes valeurs de c entre 1 à 100.

Résultats

Le tableau 13 résume les résultats en termes de précision, intervalle de confiance et aire sous la courbe ROC.

| Précision | Int de Conf. 95% | AUC |
|-----------|------------------|------|
| 72,6 | 71,1 | 0:68 |

Tableau 13 : Résumé des performances obtenues par le modèle WKM+1NN

Étant donné que le modèle non supervisé a été testé avec différentes valeurs de segmentation c , nous obtenons le meilleur résultat de précision de classification avec $c = 4$ segments. Fait intéressant, la vidéo simulation montrée aux participants a été découpée en 7 AOI, dont quatre sont regroupées dans le coin inférieur droit de l'écran correspondant aux signes vitaux. Il y a donc au fait quatre grandes zones principales et selon la précision de classification observée, le modèle semble bien corrélé avec cette observation.

7.4 Discussion et positionnement

Dans ce chapitre, nous avons exploré la problématique suivante : pouvons-nous classifier les cliniciens novices et experts en analysant seulement la séquence de leurs mouvements oculaires.

Le classifieur obtenu est utilisé par le STI EyeLab pour déterminer si des extraits de mouvements oculaires de l'apprenant reflètent un comportement de novice ou d'expert pour supporter des décisions du module de tutorat. Il s'agit d'une tâche de classification binaire de données séquentielles dans un cadre d'exploitation qui limite les possibilités de prétraitement. Deux architectures dont l'une est fondée sur des techniques

d'apprentissage connexionniste (**DeepConvLSTM**) et l'autre sur des techniques d'apprentissage statistique (**WKM-1NN**) ont été utilisées. Ces deux architectures d'apprentissage machine ont pu effectuer la tâche avec différentes performances et une bonne précision. L'approche DeepConvLSTM donne de meilleurs résultats en termes de précision de classification (90,2), d'intervalle de confiance (91,4%) et d'aire sous la courbe ROC (0,92).

Pour la classification des séquences, des tranches de point de données de longueur $s = 500$ ont été utilisées. Ce qui représente environ 8 secondes des séquences oculaires. Certaines études affirment que le temps pour une perception visuelle suivie des décisions est de quelques centaines de millisecondes (Cohen, Elger et al. 2008). Il serait donc intéressant d'étudier l'impact de la longueur de la séquence oculaire sur la précision du modèle de classification. Toutefois pour déterminer la durée de séquence convenable pour la classification, il faudrait aller au-delà de la durée physiologique minimale pour une perception et décision, mais considérer dans une perspective pédagogique la durée convenable pour obtenir des indices comportementaux visuels suffisants pour soutenir un diagnostic cognitif. Ce diagnostic cognitif va mettre à jour le modèle de l'apprenant qui à son tour va déclencher les décisions du module pédagogique. Par conséquent, l'analyse de cet impact possible de la longueur de séquence sur la précision est une autre avenue intéressante pour des travaux futurs, qui à son tour ouvre de nombreuses questions de recherche : existe-t-il un intervalle du temps à partir duquel nous pouvons obtenir des données oculaires exploitables sur un plan pédagogique dans un STI ? Est-ce qu'un chevauchement du contexte des segments pourrait-il conduire à une meilleure généralisation du modèle?

CHAPITRE VIII

EXTRACTION DE MOTIFS COMPORTEMENTAUX

Dans ce chapitre nous allons présenter la fouille de motifs séquentiels dans les données de mouvement oculaires. Contrairement à la classification des séquences du chapitre VII, le traitement de fouille de motifs séquentiels se fera en mode *offline*. L'objectif est d'extraire des séquences caractéristiques qui permettront de mettre à jour le modèle de l'expert. Dans la fouille de motifs séquentiels, les modèles de séquence sont souvent mélangés avec du bruit et d'autres ensembles de données irréguliers qui rendent difficile la recherche des motifs fréquents. La méthodologie adoptée vise à faire des prétraitements pour supprimer le bruit et autres valeurs aberrantes ou irrégulières de l'ensemble de données, par la suite, extraire les motifs fréquents qui doivent être exploitables et contextualisés.

8.1 Introduction

De précédentes recherches ont démontré l'intérêt et les avantages de bâtir un modèle des connaissances du domaine en utilisant des fouilles de données séquentielles sur les

traces de résolutions de problèmes (Fournier-Viger, Nkambou et al. 2013). Plusieurs objectifs sont visés et dans le cas spécifique de cette recherche, nous visons entre autres à :

1. Extraire les motifs de mouvement oculaires fréquents spécifiques à chaque participant (novices et experts)
2. Extraire les motifs de mouvement oculaires communs aux experts
3. Extraire les motifs de mouvement oculaires communs aux novices
4. Comparer les motifs : similitudes, différences, etc.
5. Extraire des séquences oculaires contextualisées autour des évènements clés

8.2 Nature des données

Avec une configuration de 60Hz, l'oculomètre (Tobii) collecte des points de données des mouvements oculaires tous les 17ms. Pour une durée de 13 minutes, nous obtenons un total de 46000 points de données séquentielles par participant. Le tableau 14 montre une illustration de séquences utilisées pour la classification, après le prétraitement initial. Dans le protocole expérimental, un seuil de 75% a été fixé pour qualité des données oculaires collectées par l'oculomètre. Ceci signifie qu'environ 25% des données pourront être manquantes ou imprécises. Trois étapes de prétraitements ont été effectuées pour obtenir les données qui seront utilisées par l'algorithme d'extraction de séquence.

| Parti- cipant | Point de données | AOI-1 | AOI-2 | AOI-3 | AOI-4 | AOI-5 | AOI-6 | AOI-7 |
|------------------|---------------------|-------|-------|-------|-------|-------|-------|-------|
| P1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 3 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| P1 | 4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 6 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| P1 | 7 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| P1 | 8 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| P1 | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 10 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |

Tableau 14 : Illustration d'une séquence de données utilisée pour des fouilles de données séquentielles avec les données manquantes et imprécises

Étapes de prétraitement

Les prétraitements visent à identifier ou supprimer les données manquantes ou imprécises. Une étude approfondie des données et des techniques de visualisation de données ont permis d'observer que ces données se composent de trois types de bruits.

Étape 1 : Des données manquantes.

C'est lorsqu'il n'y a aucune fixation dans un AOI pour un point de données. Dans le tableau 14, ceci est illustré par les points de données 5 et 9.

Étape 2 : Des données imprécises

C'est lorsqu'il y a par exemple deux fixations dans différents AOI pour un point de données. Dans le tableau 14, ceci est illustré par les points de données 3.

Étape 3 : Des perturbations (spikes)

C'est lorsqu'il y a par exemple un très court changement d'état (saccade) à l'intérieur d'une séquence continue dans un AOI. La figure 29 est un graphique à barres qui montre la dynamique de changement pour un ensemble de données brutes sur 1000 points de données. La figure indique la présence de pointes qui constituent la présence de spikes. C'est le cas dans l'AOI-1 au point de données 200 et l'AOI-6 au point de donnée 800.

Le spike génère une fixation dans un AOI. Toutefois, il faudrait une durée de fixation supérieure à un seuil pour être considéré comme valide. Dans les recherches, par exemple (Jraidi, Khedher et al. 2019) utilise un seuil de fixation valide de 250 ms. De plus en nous basant sur les recherches qui prennent en considération la durée physiologique pour une perception visuelle pouvant supporter une décision, la valeur par défaut du seuil de fixation valide a été mise à 250 ms. Ce seuil de durée de fixation valide est paramétrable dans le programme de prétraitement.

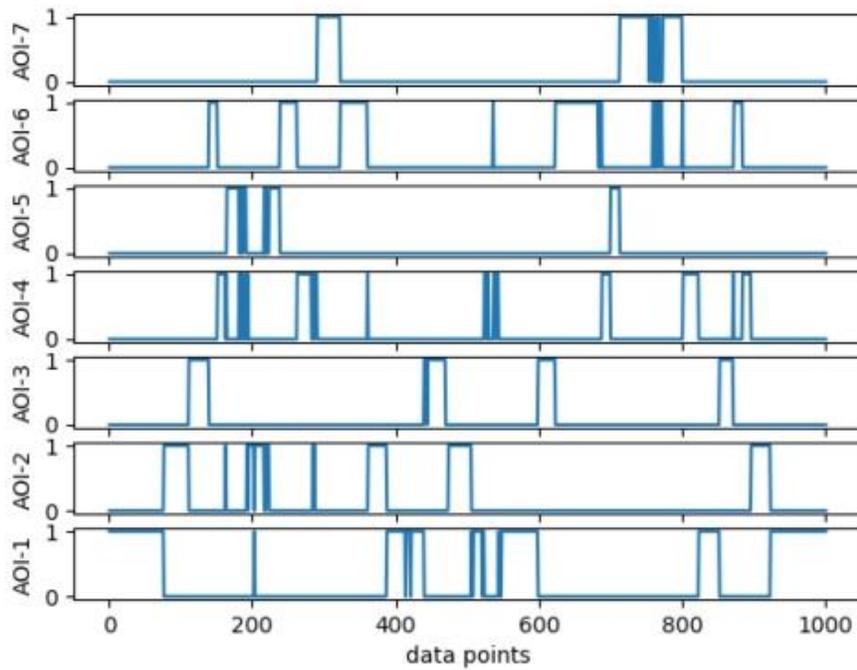


Figure 27 : Graphique à barres illustrant la dynamique de changement dans les AOI pour un ensemble de 1000 points de données avant la suppression des spikes

La figure 29 présente les étapes de prétraitement.

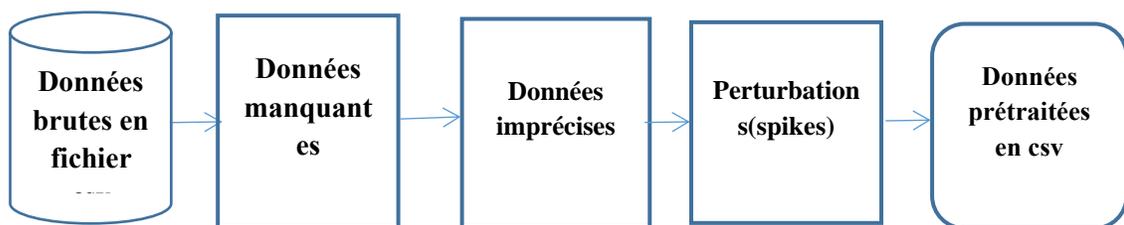


Figure 28 : Illustration des étapes de prétraitement

Après le prétraitement sur les données du tableau 14, nous obtenons les données illustrées au tableau 15, qui ne contiennent pas les données manquantes et les données imprécises. Au total, 11.72% des données ont été supprimées à l'étape 1 et 2.6% à l'étape 2.

| Participant | Point de données | AOI-1 | AOI-2 | AOI-3 | AOI-4 | AOI-5 | AOI-6 | AOI-7 |
|-------------|------------------|-------|-------|-------|-------|-------|-------|-------|
| P1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 3 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| P1 | 4 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| P1 | 5 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| P1 | 6 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| P1 | 7 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |

Tableau 15: Illustration d'une séquence de donnée après prétraitement

La figure 30 présente les graphiques à barres qui montrent la dynamique de changement dans les AOI après la suppression des spikes.

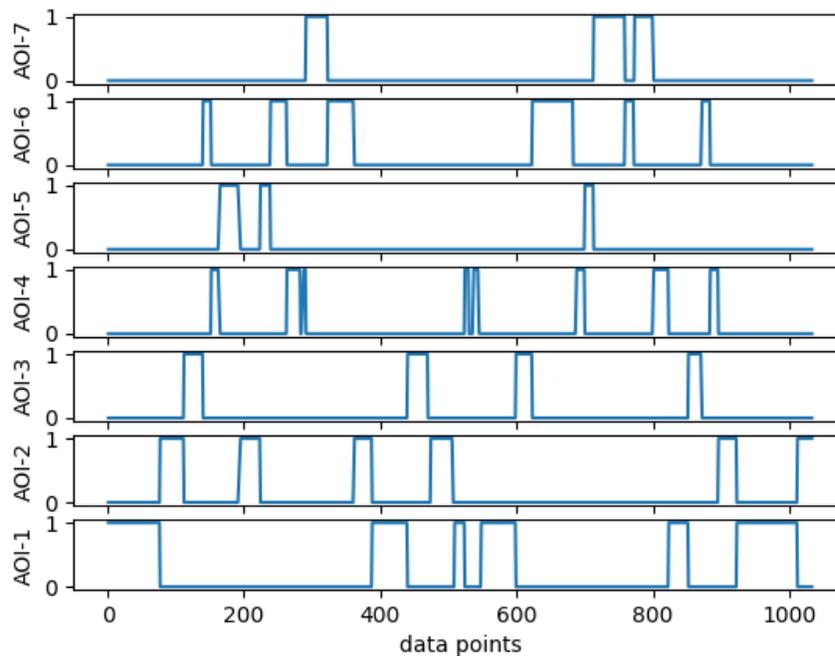


Figure 29 : Graphique à barres illustrant la dynamique de changement dans les AOI pour un ensemble de 1000 points de données après la suppression des spikes

8.3 Spécificités et contraintes sur les données de mouvements oculaires

Les caractéristiques ci-dessous décrivent les spécificités des séquences de mouvements oculaires à considérer pour le choix du modèle d'extraction des motifs fréquents :

- Les séquences oculaires sont indépendantes et de longueur variable.
- Les séquences oculaires sont constituées de points de données séquentielles selon un ordre chronologique. Ce sont donc des séries chronologiques et l'extraction des motifs doit conserver l'ordre temporel.
- Les séquences oculaires sont des séries chronologiques. Toutefois de manière classique, dans les fouilles de motifs dans les données chronologiques il n'y a pas de contrainte quant à la durée minimale pour un motif. Pour les données

oculaires, nous considérons une fixation valide pour une durée de 250 millisecondes. Ceci signifie un minimum de 15 points de données à l'intérieur d'un AOI pour être considéré valide et rajouté au motif en cours d'extraction.

- Les motifs à extraire des séquences de mouvements oculaires pour chaque participant doivent être mutuellement exclusifs. Les motifs sont extraits transaction par transaction de manière indépendante. Chaque transaction représente les mouvements oculaires complets d'un clinicien. Par la suite, il faudrait déterminer les motifs similaires entre transactions.
- Les séquences oculaires sont multidimensionnelles par transaction. Il y a exactement sept dimensions (AOI), toutefois l'algorithme doit permettre un nombre variable de dimensions.

Le choix d'algorithme pour l'extraction des motifs devrait se faire en prenant en considération les spécificités et contraintes des séquences de mouvements oculaires.

8.4 Étude et comparaisons de quelques algorithmes d'extraction de motifs

Le processus générique adopté par la plupart des algorithmes d'extraction de motifs qui serviront aux règles d'association est composé de deux sous-étapes :

1. La recherche d'itemsets fréquents
2. La génération des règles d'association. Les règles d'association extraites n'impliquent pas nécessairement un lien de causalité, mais seulement la présence simultanée de deux itemsets dans des enregistrements.

Différentes métriques comme le support et la confiance sont utilisées par la suite pour sélectionner les motifs convenables. Nous illustrons dans le tableau 16 un exemple de base de transactions pour une extraction de motifs. Chaque client constitue une seule transaction avec un identifiant unique associé.

| TID | Items | | | |
|----------|-------|--------|---------|---------|
| Client 1 | Pain | Lait | | |
| Client 2 | Pain | Beurre | Fromage | Œufs |
| Client 3 | Lait | Beurre | Fromage | Coke |
| Client 4 | Pain | Lait | Beurre | Fromage |
| Client 5 | Lait | Pain | Beurre | Coke |

Tableau 16 : Exemple de transactions pour une extraction de motifs

8.4.1 Algorithme APriori

L'algorithme APriori (Agrawal et Srikant 1994) détermine les règles d'association présentes dans un ensemble de transactions, pour un seuil de support et un seuil de confiance fixés. A-Priori fonctionne en deux phases :

- Phase 1 : Rechercher les ensembles d'items fréquents (EIF)
- Phase 2 : Utiliser ces EIF pour déterminer les règles d'association dont la confiance est supérieure au seuil préalablement défini.

Le processus de découverte des ensembles d'items fréquents de la phase 1 est itératif et incrémental. Il débute par la construction de EIF avec un seul item, ensuite on réitère pour construire des EIF avec deux items, ainsi de suite.

A-Priori offre l'avantage de la simplicité à comprendre à un niveau conceptuel et donne des résultats faciles à comprendre. De plus, tout sous-ensemble d'un itemset fréquent est fréquent. Il y a toutefois deux inconvénients majeurs à cet algorithme. La première est la performance due au nombre considérable d'accès à la base de transaction D pour générer les candidats. Le second inconvénient est la qualité en termes d'exploitabilité

versus la quantité selon les valeurs de support/confiance. Étant l'un des premiers algorithmes d'extractions de motifs fréquents, plusieurs variantes ont été proposées pour pallier les limites et améliorer la performance. L'algorithme Apriori-TID (Agrawal & Srikant, 1994) est une variante qui permet de stocker à chaque itération les identificateurs des transactions contenant les sous-ensembles fréquents dans l'ensemble. D'autres solutions ont été également proposées aux défis de génération des candidats comme :

- Direct Hashing and Pruning (DHP) (Park et al., 1997) : utilise des tables de hachage afin de réduire le nombre de candidats générés.
- Partition (Savasere et al., 1995) : propose un partitionnement de la base de données afin d'optimiser la consommation de mémoire.
- Dynamic Itemset Counting (DIC) (Brin, Motwani, Ullman & Tsur, 1997) : quand un k -itemset est considéré comme fréquent, tous les $(k + 1)$ -itemsets candidats que le k -itemset peut produire sont générés.
- Tree Projection (Agrawal et al., 2001) : utilise l'arbre d'un ensemble d'énumérations pour générer des candidats.

Mis à part les algorithmes qui sont des variantes optimisées de Apriori, d'autres algorithmes sont basés sur des hyperstructures de données compactes (FP-tree).

8.4.2 Algorithme FP-Growth

L'algorithme Frequent Pattern growth permet la découverte des itemsets fréquents, mais contrairement à Apriori il n'y a pas de génération des itemsets candidats (Han, Pei et al. 2000). Le processus se déroule en deux phases :

- Phase 1 : Construire des arbres FP-tree
- Phase 2 : Extraire les itemsets fréquents directement de ces arbres FP-tree.

La construction de l'arbre FP-tree s'effectue suivant les étapes ci-dessous. Le processus de construction de l'arbre Fp-tree s'arrête lorsque toute la base de transactions est parcourue.

1. Calculer le support minimal.
2. Calculer chacune des occurrences d'un item constituant la base de transactions.
3. Établir un critère de priorité pour ces items.
4. Faire le tri des items en fonction de leur priorité.
5. Établir le nœud racine.
6. À partir de chaque nœud père insérer les enfants en partant du nœud racine
7. Valider la structure de l'arbre FP-Growth.

Le principal avantage de FP-Growth est la performance. L'algorithme ne fait que deux balayages de la base de transactions. Le premier balayage pour trouver les k-itemsets et pour construire la liste des items fréquents. Le second pour construire la structure de l'arbre FP-tree. Le mécanisme de fonctionnement de l'algorithme FP-Growth via la construction de l'arbre FP-tree lui confère l'avantage de performance, mais est à la base de l'inconvénient principal de l'algorithme. La construction de l'arbre se fait en mémoire qui est plus performante en lecture/écriture que le disque pour accéder par exemple à la base de transaction. Dans des bases de transaction volumineuses, les calculs nécessaires à la construction de l'arbre et le stockage en mémoire peuvent donc prendre beaucoup de ressources au niveau du processeur et de la mémoire.

8.4.3 Algorithme ECLaT

L'algorithme ECLaT (Equivalence CLass Transformation) est un algorithme de recherche en profondeur basé sur l'intersection d'ensemble introduit par (Zaki 2000). Cet algorithme fait l'extraction de motifs fréquents en passant par une représentation intermédiaire verticale des données de la base de transactions. Dans cette méthode de

représentation des données, toutes les transactions qui contiennent un ensemble d'éléments particulier sont regroupées dans le même enregistrement. L'algorithme ECLaT transforme les transactions du format horizontal en format vertical en scannant la base de données une fois. Par la suite, de manière itérative un croisement des $k+1$ itemset fréquents est effectué. Étant donné que la représentation intermédiaire verticale ainsi que les traitements de l'algorithme ECLaT se passe en mémoire, l'inconvénient principal provient des ressources au niveau du processeur et de la mémoire.

8.4.4 Algorithme N-gram

Cet algorithme n'est pas à la base un algorithme d'extraction de motifs séquentiels, mais plutôt utilisé pour les tâches de prédiction de séquence en linguistique computationnelle et en bio-informatique. Les séquences prédites peuvent toutefois servir de base pour une classification.

Un n-gram est une sous-séquence de n items construits à partir d'une séquence donnée. À partir d'une séquence d'item il est possible d'obtenir la fonction de vraisemblance de l'apparition de l'item suivant. Il faudrait donc construire une distribution de probabilité pour le prochain item à partir d'un corpus d'apprentissage. Le modèle obtenu est en fait un modèle de Markov d'ordre n , dans lequel les n derniers items sont utilisés pour la prédiction de l'item suivant. Le modèle pour les calculs de probabilités suppose une indépendance locale d'un item à l'autre. Pour l'exploitation du modèle, les distributions de probabilité sont lissées en attribuant des probabilités non nulles aux items qui n'étaient pas présents dans la phase d'apprentissage.

Il existe d'autres algorithmes récents comme l'algorithme P-Mine (Baralis, Cerquitelli et al. 2013) optimisé pour un traitement parallèle, l'algorithme LP-Growth (Pyun, Yun et al. 2014), l'algorithme EXTRACT (Feddaoui, Fa et al. 2016). Finalement, même s'il

existe de nombreux algorithmes, (Chee, Jaafar et al. 2019) propose une classification en trois catégories principales selon le mécanisme de fonctionnement.

Mis à part ces algorithmes génériques d'extraction de séquence, il y a des algorithmes conçus spécifiquement pour les données éducationnelles comme CMRules et PHARules déjà présentés.

8.5 Discussions et limites des algorithmes disponibles pour l'extraction de séquence

Les motifs extraits par exemple par les algorithmes comme Apriori et FP-Growth sont utilisés pour déterminer des règles d'association qui mettent en évidence les tendances séquentielles générales dans les bases de transaction ou les co-occurrences d'itemsets. Cela s'applique bien à des applications dans une catégorie de domaines que nous nommons l'analyse du panier d'épicerie. Même si leurs principes de fonctionnement diffèrent, les algorithmes A-Priori, FP-Growth et ECLaT sont dans cette catégorie. Ces algorithmes exigent plusieurs transactions réalisées par différents acteurs. Par la suite, il s'agit de trouver les co-occurrences communes à plusieurs transactions. Ces co-occurrences servent à extraire les motifs fréquents et les règles d'associations. Également, l'ordre temporel d'exécution n'est pas considéré. Par exemple, un achat de lait, céréale et bière est équivalent à un achat de céréale, lait et bière. Comme présenté dans les contraintes et caractéristiques dans les données de mouvements oculaires et par rapport aux objectifs du projet, les algorithmes de cette catégorie d'analyse du panier d'épicerie ne sont pas convenables pour plusieurs raisons. En effet, notre objectif est d'extraire les motifs séquentiels dans une seule transaction à la fois et par la suite, de trouver les motifs communs ou différents pour plusieurs transactions. Les motifs devront donc être extraits initialement transaction par transaction de

manière indépendante. Également, l'extraction des motifs de mouvements oculaires doit conserver l'ordre temporel.

L'algorithme ECLaT ne convient pas pour une autre raison, car les données oculaires sont mutuellement exclusives dans chaque transaction. Eclat construit un arbre en combinant les identifiants et en choisissant des colonnes communes et il n'y a pas de colonnes communes dans les données oculaires.

Finalement, l'algorithme N-gram sert à l'extraction de motifs dans une seule transaction, il pourrait être un choix convenable. Par exemple, les co-occurrences d'une suite de mots dans un paragraphe. Le paragraphe est considéré comme la séquence de laquelle les motifs seront extraits et représente une seule transaction d'autant plus qu'il ne s'agit pas d'extraire les co-occurrences de mots provenant de paragraphes rédigés par différentes personnes. L'inconvénient principal avec N-gram provient du fait que l'extraction fournit des motifs de longueur fixe et prédéterminée. Par exemple, pour un motif de longueur 2, nous parlerons de bi-gram. De plus, N-gram suppose une indépendance locale d'item à item dans la séquence. Cette condition simplifie le problème et ne peut pas être garantie dans le cas des données de mouvements oculaires.

8.6 L'algorithme MDSE (Multidimensional Dynamic Sequence Extraction)

L'algorithme MDSE (Multidimensional Dynamic Sequence Extraction) est conçu en prenant trois critères en considération :

1. Les spécificités et contraintes sur les données de mouvements oculaires.
2. Les limites des algorithmes existants pour l'extraction de motifs.
3. Les objectifs spécifiques du projet de recherche.

L'algorithme MDSE est conçu en se basant sur les apports de trois algorithmes existants notamment : FP-Growth, N-gram et PHARules. Le principal avantage de FP-Growth est la performance. Les phases de construction des arbres FP-tree et d'extraction des motifs fréquents ont été modifiées pour procéder transaction par transaction. N-gram fournit justement l'avantage de détecter les co-occurrences dans une séquence (transaction). Finalement, PHARules offre la capacité à prendre en compte l'importance des phases dans les traces de résolution. Les phases sont identifiées par les événements clés. L'implémentation de l'algorithme MDSE permet de faire des appels itératifs en prenant en paramètre les événements clés pour retourner les motifs extraits durant ces phases.

8.6.1 Entrée de l'algorithme MDSE

Aplatissement dimensionnel des séquences (Dimensional Sequence Flattening)

Il s'agit d'une étape de redescription de l'espace multidimensionnel des séquences en une nouvelle dimension contenant tous les items. L'objectif est d'avoir une compression des séquences en une seule dimension tout en gardant l'expressivité séquentielle et temporelle. Sur un plan technique, il s'agit d'une projection de toutes les fixations sur un nouvel axe temporel. Le tableau 17 présente l'opération Dimensional Sequence Flattening (DSF) et le résultat obtenu est : (AOI-1, AOI-1, AOI-1, AOI-4, AOI-3, AOI-3, AOI-5). Nous présentons dans tableau 18 un extrait de séquence des mouvements oculaires pour l'expert 5 après l'opération DSF.

| Point de données | AOI-1 | AOI-2 | AOI-3 | AOI-4 | AOI-5 | AOI-6 | AOI-7 | DSF |
|------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | AOI-1 |
| 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | AOI-1 |
| 3 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | AOI-1 |
| 4 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | AOI-4 |
| 5 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | AOI-3 |
| 6 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | AOI-3 |
| 7 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | AOI-5 |

Tableau 17 : Exemple du Dimensional Sequence Flattening

AOI-1,AOI-7, AOI-6,AOI-4, AOI-6,AOI-1, AOI-3, AOI-1,AOI-2, AOI-3, AOI-1, AOI-3,
 AOI-1,AOI-7, AOI-3,AOI-7, AOI-6, AOI-7,AOI-5, AOI-7,AOI-6, AOI-4,AOI-6, AOI-3,
 AOI-6, AOI-3, AOI-4,AOI-1, AOI-3, AOI-1,AOI-7, AOI-6, AOI-5,AOI-3, AOI-1, AOI-
 3,AOI-5, AOI-7,AOI-3, AOI-6, AOI-4, AOI-3, AOI-4, AOI-5,AOI-7, AOI-6, AOI-4,AOI-5,
 AOI-3,AOI-6, AOI-7, AOI-5,AOI-3, AOI-6,AOI-3, AOI-1, AOI-3,AOI-2, AOI-4,AOI-1,
 AOI-2,AOI-1, AOI-3,AOI-4, AOI-5,AOI-1, AOI-3, AOI-1,AOI-4, AOI-3, AOI-1, AOI-
 3,AOI-4, AOI-6, AOI-4, AOI-5,AOI-4, AOI-6,AOI-1, AOI-3,AOI-1, AOI-3, AOI-1,AOI-5,
 AOI-6,AOI-2, AOI-1, AOI-3,AOI-5, AOI-6, AOI-5, AOI-7, AOI-5,AOI-1, AOI-3, AOI-
 4,AOI-1, AOI-3, AOI-4,AOI-3, AOI-4,AOI-5, AOI-3,AOI-5, AOI-7,AOI-4, AOI-5,AOI-6,
 AOI-4, AOI-6,AOI-3, AOI-4,AOI-3, AOI-1, AOI-3,AOI-3, AOI-4,AOI-5, AOI-4,AOI-3,
 AOI-1, AOI-3,AOI-1, AOI-3, AOI-1, AOI-3,AOI-1, AOI-3, AOI-6,AOI-7, AOI-5,AOI-6,
 AOI-5, AOI-6,AOI-1, AOI-4,AOI-3, AOI-6,AOI-3, AOI-1, AOI-3, AOI-6, AOI-3, AOI-
 6,AOI-6, AOI-3, AOI-1, AOI-3,AOI-6, AOI-3, AOI-4, AOI-3,AOI-6, AOI-7, AOI-3, AOI-6,
 AOI-5,AOI-3, AOI-4, AOI-6, AOI-4, AOI-5, AOI-4,AOI-3, AOI-6, AOI-3,AOI-1, AOI-3,

AOI-1,AOI-3, AOI-1, AOI-3,AOI-1, AOI-3, AOI-2,AOI-6, AOI-4,AOI-1, AOI-3, AOI-1,AOI-2, AOI-3, AOI-1, AOI-2,AOI-3, AOI-1, AOI-3, AOI-1, AOI-3,AOI-2, AOI-1, AOI-3,AOI-1, AOI-3, AOI-1,AOI-2, AOI-1, AOI-3, AOI-1,AOI-4, AOI-3,AOI-2, AOI-1, AOI-3, AOI-1,AOI-2, AOI-5, AOI-2,AOI-3, AOI-1,AOI-6, AOI-3, AOI-1, AOI-3,AOI-4, AOI-1,AOI-5, AOI-2, AOI-5, AOI-2,AOI-3, AOI-1,AOI-6, AOI-4,AOI-6, AOI-3,AOI-7, AOI-5,AOI-3, AOI-1, AOI-2,AOI-6, AOI-1,AOI-2, AOI-5,AOI-1, AOI-3,AOI-2, AOI-1, AOI-3,AOI-5, AOI-2,AOI-3, AOI-1, AOI-3,AOI-5, AOI-2,AOI-7, AOI-3,AOI-5, AOI-6,AOI-4, AOI-3, AOI-6,AOI-3, AOI-1,AOI-2, AOI-5,AOI-3, AOI-1, AOI-3, AOI-1, AOI-2, AOI-1,AOI-2, AOI-1, AOI-3,AOI-1, AOI-3, AOI-1, AOI-3, AOI-1, AOI-3,AOI-2, AOI-3

Tableau 18 : Un extrait de séquence des mouvements oculaires pour l'expert 5 après l'opération DSF

8.6.2 Sortie de l'algorithme MDSE

Le tableau 19 présente la nomenclature adoptée pour la représentation des motifs extraits par l'implémentation de l'algorithme MDSE. Ci-dessous les explications de chaque élément du tableau :

- `sequence_pattern` : Il s'agit du motif extrait : ['AOI-2', 'AOI-1', 'AOI-6'] par l'algorithme MDSE
- `ind_AOI_length` [96, 43, 134] : indique le nombre d'occurrences de chaque élément du motif pris individuellement. Nous avons dans ce cas spécifique 96 fixations dans l'AOI-2 suivi de 43 fixations dans l'AOI-1 et 134 fixations dans l'AOI-6.
- `start_timestamp` : correspond à l'instant de début du motif. Cette valeur est en milliseconde, ce qui signifie que 575000 millisecondes correspondent à 9:35 depuis le début de la simulation.
- `sequence_length_sum` : c'est la longueur totale du motif extrait. C'est le nombre total de fixations contenu dans le motif. Donc la valeur 273 correspond

à $96+43+134$. Étant donné qu'il y a 17 millisecondes entre fixations, la durée de ce motif extrait de: $273*17 = 4641$ millisecondes (4.6 secondes).

- `end_timestamp` : correspond à l'instant de fin du motif. Cette valeur est en milliseconde, ce qui signifie que 585000 correspond à 9:45 depuis le début de la simulation. Théoriquement, on pourrait s'attendre à avoir une valeur de `end_timestamp = start_timestamp + 17*sequence_length_sum`. Ce qui dans le cas présenté donnerait 579641. Toutefois, dans la pratique pour les motifs extraits, la valeur du `end_timestamp` est supérieure, car seules les fixations valides sont considérées pour l'extraction des motifs.

Cette nomenclature des motifs nous donne une représentation contextualisée et exploitable pour comparer la dynamique des mouvements oculaires. Dans l'exemple du tableau 19 avec le début de séquence à 9:35 et la fin à 9:45; nous obtenons un motif représentatif de la dynamique des mouvements oculaires d'un clinicien autour de l'évènement clé : lancement du code bleu. Précisément, il s'agit de 4 secondes avant l'évènement et 6 secondes après l'évènement.

| Nom | Valeur | Description |
|----------------------------------|-----------------------------|--------------------------------------|
| <code>sequence_pattern</code> | ['AOI-2', 'AOI-1', 'AOI-6'] | Motif extrait |
| <code>ind_AOI_length</code> | [96, 43, 134] | Nombre d'éléments par items |
| <code>start_timestamp</code> | 575000 | Début de la séquence en milliseconde |
| <code>sequence_length_sum</code> | 273 | Longueur du motif extrait |
| <code>end_timestamp</code> | 585000 | Fin de la séquence en milliseconde |

Tableau 19 : Nomenclature de la représentation des motifs extraits

8.6.3 Dérroulement de l'algorithme MDSE

Nous présentons dans le tableau 20, le pseudocode de l'algorithme MDSE. Ceci sera suivi d'explications de certains points importants de l'algorithme.

Algorithm : MDSE-Growth

Input: preprocessed and flattened eye tracking data.

Output: Frequent patterns with counts and contextualized information.

Method: call MDSE

Procedure MDSE-growth (Tree, a) {

(01) if Tree contains a single prefix path then { //

(02) let P be the single prefix-path part of Tree

(03) let L represent the length of sequence

(04) let X be the long sequence ($L > \text{minValidFixation}$)

(05) let Y be the short sequence ($L < \text{minValidFixation}$)

(06) Concat (join) X_1, X_2, \dots, X_n sequence till Y breaks sequence

(07) Let ST be start time stamp, note the time stamp at start of sequence

(08) Let ET be end of time stamp, note the time stamp at the end of sequence

(09) Remove sequence (X) with length = 1

(10) for each combination of pattern (denoted as β) of the nodes in the path P do

(11) Repeat the steps (06) to find more patterns ($\beta_1, \beta_2, \beta_3, \dots, \beta_n$)

(12) Note the length of pattern (add the individual sequence length)

(13) generate pattern $\beta \cup a$ with support = minimum support length of nodes in β

(14) let freq pattern set(Q) be the set of patterns so generated;

(15) Group duplicates patterns (β) assign multiple freq (Q) counts

(16) return (freq pattern set(β) \times freq pattern set(Q), ST, ET and L

Tableau 20 : Déroulement de l'algorithme MDSE

Explications du déroulement de l'algorithme MDSE

L'algorithme MDSE débute comme FP-Growth. En effet FP-Growth identifie d'abord un ensemble d'item fréquents, puis les trie et construit des arbres en fonction de leur fréquence, finalement attribue des nœuds et des liens parents (ce qui devient un motif). Cependant, pour notre recherche nous avons besoin non seulement des items qui vont constituer le motif, mais également leur fréquence pour identifier des motifs similaires et différents. L'étape d'assignation des nœuds est donc différente dans MDSE. Dans FP-Growth, les nœuds et les liaisons avec les parents sont formés pour séparer les motifs en différentes longueurs afin d'identifier les sous-arbres et les nœuds. Dans notre recherche, les étapes de prétraitement proposées identifient implicitement des séparations, car il faudrait considérer le seuil pour avoir des fixations valides pouvant être ajoutées aux motifs. Finalement, les items dans le motif extrait doivent conserver leurs ordonnancement temporel. Dans l'algorithme MDSE :

- L'arbre est construit dans un ordre temporel croissant.
- Pour un motif en cours de construction ayant une nouvelle sous-séquence de fixation dans un AOI supérieur à `minValidFixation` (seuil pour avoir des fixations valides), la sous-séquence est ajoutée au motif. Si la sous-séquence est de longueur inférieure à `minValidFixation`, elle n'est pas rajoutée au motif.
- Les motifs avec un seul item sont supprimés
- Le *timestamp* de début, le *timestamp* de la fin et le nombre des items des motifs sont conservés.

À un niveau procédural d'implémentation, les étapes d'extractions sont présentées ci-dessous :

1. Let start = 1 and end = length of dataset
2. Let L be the data points (View time length) = 30
3. For repeat the procedure from start to end
4. Start scan the dataset according with order of time stamps
5. If seq(1) < L then Pattern = null
6. If seq(1) > L, If seq(2) < L then Pattern = null
7. If seq(1) > L, If seq(2) > L then Pattern = seq(1), seq(2)
8. If pattern seq(1), seq(2), If seq(3) < L then Pattern = seq(1), seq(2)
9. If pattern seq(1), seq(2), If seq(3) > L then Pattern = seq(1), seq(2), seq(3)
10. Repeat the procedure till seq(n) < L then new pattern
11. Repeat the procedure till seq(n) > minValidFixation then continue pattern joining seq(n)
12. Once end is reached, stop
13. Assign unique tokens to patterns, repeated tokens count as frequency

Au niveau de la performance, MDSE parcourt la base de transaction deux fois comme FP-growth pour extraire des motifs de longueurs variables. Il ne génère pas d'items candidats tout comme FP-growth. Toutefois MDSE hérite des inconvénients précédents mentionnés de FP-growth. Comparé à N-gram, il n'y a aucun apriori par rapport à l'indépendance des items et la taille des motifs est variable.

8.6.4 Résultats obtenus

Nous présentons dans le tableau 21 un exemple de motifs extrait par l'algorithme MDSE pour l'expert identifié par le numéro 5.

| | sequence_pattern | Ind_AOI_length | sequence_ length_sum | start_ timestamp | end_ timestamp |
|----|--|----------------------------|-------------------------|---------------------|-------------------|
| 0 | ['AOI-3', 'AOI-1', 'AOI-3'] | [32, 200, 77] | 309 | 10710 | 15326 |
| 1 | ['AOI-1', 'AOI-3', 'AOI-2'] | [62, 163, 140] | 365 | 21425 | 26475 |
| 2 | ['AOI-6', 'AOI-4'] | [79, 51] | 130 | 32874 | 33724 |
| 3 | ['AOI-1', 'AOI-3', 'AOI-1'] | [34, 124, 84] | 242 | 36623 | 40090 |
| 4 | ['AOI-1', 'AOI-3', 'AOI-1', 'AOI-2'] | [59, 189, 86, 41] | 375 | 42639 | 47905 |
| 5 | ['AOI-3', 'AOI-1', 'AOI-3', 'AOI-1', 'AOI-3'] | [87, 269, 68, 100, 130] | 654 | 49872 | 59320 |
| 6 | ['AOI-2', 'AOI-1', 'AOI-3'] | [53, 165, 33] | 251 | 60487 | 63786 |
| 7 | ['AOI-1', 'AOI-3', 'AOI-1'] | [37, 39, 276] | 352 | 76851 | 82101 |
| 8 | ['AOI-2', 'AOI-1', 'AOI-3', 'AOI-1'] | [78, 144, 109, 212] | 543 | 83950 | 91699 |
| 9 | ['AOI-1', 'AOI-3'] | [88, 96] | 184 | 100548 | 102148 |
| 10 | ['AOI-2', 'AOI-1', 'AOI-3', 'AOI-1'] | [31, 32, 51, 158] | 272 | 103448 | 107464 |
| 11 | ['AOI-2', 'AOI-5', 'AOI-2'] | [262, 72, 59] | 393 | 112980 | 115163 |
| 12 | ['AOI-3', 'AOI-1'] | [328, 159] | 487 | 126811 | 129461 |
| 13 | ['AOI-1', 'AOI-3', 'AOI-1', 'AOI-3'] | [49, 146, 82, 315] | 592 | 136943 | 145992 |
| 14 | ['AOI-3', 'AOI-1'] | [41, 129] | 170 | 150441 | 152591 |

Tableau 21 : Exemple de motifs extraits par l'algorithme MDSE pour l'expert 5

Sur la base des motifs, nous avons procédé aux extractions et comparaisons nécessaires pour la recherche. Il s'agit notamment de :

1. Extraire les motifs de mouvement oculaires fréquents spécifiques à chaque participant (novices et experts)
2. Extraire les motifs de mouvement oculaires communs aux experts
3. Extraire les motifs de mouvement oculaires communs aux novices
4. Comparer les motifs : similitudes, différences, etc.
5. Extraire les séquences oculaires contextualisées autour des événements clés

Le tableau 22 présente un exemple de motifs similaires pour les experts. Le tableau présente uniquement un sommaire des motifs similaires, la fréquence des motifs et la longueur.

| No | Patterns | Count | Length |
|----|-----------------------------|-------|--------|
| 1 | ['AOI-3', 'AOI-1', 'AOI-3'] | 25 | 297 |
| 2 | ['AOI-1', 'AOI-3', 'AOI-1'] | 16 | 254 |
| 3 | ['AOI-2', 'AOI-1', 'AOI-3'] | 11 | 309 |
| 4 | ['AOI-1', 'AOI-3'] | 78 | 169 |
| 5 | ['AOI-3', 'AOI-1'] | 56 | 167 |
| 6 | ['AOI-2', 'AOI-1'] | 22 | 153 |
| 7 | ['AOI-6', 'AOI-3'] | 9 | 168 |
| 8 | ['AOI-3', 'AOI-1', 'AOI-2'] | 8 | 222 |
| 9 | ['AOI-2', 'AOI-3'] | 10 | 110 |
| 10 | ['AOI-3', 'AOI-4'] | 14 | 199 |
| 11 | ['AOI-4', 'AOI-3'] | 16 | 167 |

| | | | |
|----|--------------------|----|-----|
| 12 | ['AOI-3', 'AOI-6'] | 6 | 151 |
| 13 | ['AOI-5', 'AOI-4'] | 10 | 96 |
| 14 | ['AOI-1', 'AOI-2'] | 14 | 110 |

Tableau 22 : Exemple de motifs similaires pour les experts

Les motifs fréquents des experts sont identifiés comme motifs caractéristiques des mouvements oculaires des experts et serviront pour le modèle de l'expert du STI.

8.7 Discussion et positionnement

Dans ce chapitre nous avons présenté la fouille de motifs séquentiels dans les données de mouvement oculaires. Après une comparaison des algorithmes courants pour l'extraction de motifs au regard des contraintes des données de mouvement oculaires ainsi qu'aux objectifs du projet, nous avons proposé l'algorithme MDSE. L'algorithme MDSE est conçu en se basant sur les apports de trois algorithmes existants notamment : FP-Growth, N-gram et PHARules. L'implémentation de l'algorithme offre également l'avantage de fournir en sortie un modèle de représentation de motifs qui sous-tendent des connaissances perceptivo-décisionnelles exploitables dans un STI. Le traitement de fouille de motifs séquentiels se fera en mode *offline*. L'objectif est d'extraire des motifs caractéristiques qui permettront de mettre à jour le modèle de l'expert.

L'algorithme MDSE est optimisé pour les contraintes et objectifs du projet de recherche. Toutefois à un niveau fondamental MDSE serait un algorithme de choix pour l'extraction de motifs dans des données séquentielles dans le temps présentant les

caractéristiques des données du projet. MDSE permettra de bénéficier des caractéristiques suivantes : des motifs de longueur variable par séquences et non par co-occurrence dans des transactions, l'ordre chronologique, un extrait orienté phase et contextualisé. Nous présenterons dans la généralité des solutions des domaines dans lesquels nous pouvons avoir des séquences de données auxquelles pourraient s'appliquer l'algorithme MDSE.

PARTIE 4 – SPÉCIFICATIONS, CONCEPTION ET IMPLÉMENTATION DU
SYSTÈME TUTORIEL INTELLIGENT

CHAPITRE IX

SPÉCIFICATIONS ET FONCTIONNEMENT DU STI EYELAB

Dans ce chapitre nous présentons l'utilisation de l'oculométrie cognitive dans les interactions humain-machine dans différents domaines. En se basant sur les résultats concluants de l'utilisation de l'oculométrie cognitive dans le domaine des STI ainsi que dans le domaine du RC nous avons défini les principes pédagogiques soutenant les services tutoriels du fonctionnement du STI EyeLab. Le STI EyeLab est une plateforme web qui intègre des fonctionnalités de suivis oculaires et des fonctionnalités d'un STI.

9.1 Utilisation de l'oculométrie cognitive

L'oculométrie cognitive vise l'étude des comportements sous-jacents à la perception visuelle et l'interprétation des opérations cognitives associées. Les techniques d'oculométrie cognitive ont été utilisées dans de nombreux domaines comme le marketing web (Nielsen et Pernice 2009) , l'éducation (Homof, Cavender et al. 2004),

la médecine (Hutchinson, White et al. 1989), l'armée, les environnements d'apprentissage en ligne (Conati et Merten 2007).

9.1.1 Marketing web

Les travaux de (Nielsen et Pernice 2009) ont permis de démontrer que les utilisateurs du web parcourent généralement le contenu d'une page sous forme de la lettre « F ». La majorité des utilisateurs commencent par lire la bande horizontale du haut, ensuite une bande horizontale du milieu et enfin ils parcourent la zone verticale gauche. L'étude du comportement des lecteurs vis-à-vis des pages web a également démontré que les utilisateurs ne lisent pas les pages mot à mot, mais sautent d'un endroit à l'autre. Ces études ont contribué à l'amélioration dans la conception des interfaces de site web qui exploite l'optimisation dans la présentation du contenu en vue d'influencer indirectement le processus décisionnel de l'internaute.

9.1.2 Médical

Le système ERICA (Eye-gaze-Response Interface Computer Aid) de (Hutchinson, White Jr, Martin, Reichert, & Frey, 1989) permet aux personnes handicapées d'appeler des infirmières à partir du mouvement des yeux.

L'oculométrie cognitive a été aussi utilisée dans des recherches et développements d'outils dans le domaine médical. Par exemple, les travaux de (Charlier et al. 1992)⁹ ont contribué à mettre au point un microscope guidé par le regard. Ceci est utile à un chirurgien dont les deux mains sont occupées par la manipulation d'instruments lors d'une opération. Dans une telle situation, le déplacement du champ visuel du microscope ne peut pas être facilement délégué à une autre personne vu la lenteur et

⁹ <http://www.google.fr/patents/WO1996013743A1?cl=fr>

les risques d'erreurs élevées qui y sont associés. La commande par les pieds risque d'augmenter la charge cognitive du clinicien, car les mains sont aussi en mouvement et la commande par la voix, serait inefficace, car la voix est discontinue et inadéquate pour des réglages fins. La commande par le regard reste donc la meilleure solution.

9.1.3 Domaine militaire

L'oculométrie est beaucoup utilisée dans le domaine militaire aussi bien pour l'entraînement que des applications concrètes. SBG Labs, un fournisseur de la défense américaine installé dans la Silicon Valley a développé un système de ciblage contrôlé par le regard pour l'US Air Force. En effet, à l'heure actuelle, pour effectuer une telle manœuvre le pilote devrait tourner et incliner la tête et des capteurs intégrés au casque assurent une interprétation de ses mouvements en orientation pour le ciblage. Tous ces mouvements deviennent difficiles lorsque la vitesse atteint plusieurs Mach dus à l'effet d'apesanteur. Polhemus, une entreprise basée dans le Vermont a développé des dispositifs de eye-tracking pour suivre les mouvements oculaires des étudiants pilote dans des simulateurs de vol permettant de détecter si l'apprenant a balayé les jauges dans le bon ordre, etc. Eye-Com une entreprise basée au Nevada, a conçu un masque équipé d'oculomètre pour les Navy SEALs permettant de détecter chez le soldat d'éventuelles détresses physiologiques pouvant impacter sa performance.

9.1.4 Environnement d'apprentissage en ligne

L'oculométrie cognitive est utilisée pour alimenter le modèle de l'apprenant dans le but d'évaluer le comportement métacognitif durant ses interactions avec un environnement d'apprentissage (Conati et Merten 2007), la métacognition étant généralement reliée à une meilleure compréhension de ce qui est observé, donc un meilleur apprentissage. Les travaux de (Beatty 1982) ont montré une corrélation entre la charge cognitive et la dilatation de la pupille et (Barrier 2008) a observé une corrélation positive entre la dilatation de la pupille et la difficulté de la tâche.

L'oculométrie cognitive est également utilisée pour analyser la capacité de l'apprenant à réaliser une autorégulation de son apprentissage. Certains chercheurs orientent plutôt leurs études sur l'utilisation des mouvements oculaires comme indices des capacités cognitives de l'apprenant (Steichen et al. 2013). L'objectif est d'exploiter des données oculométriques pour déduire trois types de capacités cognitives de l'apprenant :

1. La vitesse perceptuelle (une mesure de la vitesse de réalisation de tâches perceptuelles simples).
2. La mémoire visuelle de travail (une mesure de la capacité de rétention et manipulation d'information visuelle et spatiale).
3. La mémoire verbale de travail (une mesure de la capacité de rétention et de manipulation d'information verbale).

Dans les STI, (Kardan et Conati, 2012) ont utilisé les données de suivi oculaire pour améliorer le modèle de l'apprenant et fournir un soutien adaptatif. Ils ont également fait des investigations sur les comportements oculaires des apprenants qui conduisent à une réussite ou un échec lors d'un apprentissage dans un STI. Finalement (Bull, Cooke, & Mabbott, 2007 et Mathews et al., 2012) exploitent le suivi oculaire pour stimuler la métacognition dans un modèle ouvert de l'apprenant

9.2 Oculométrie cognitive et transfert d'expertise

Certaines recherches ont utilisé le suivi oculaire pour améliorer l'apprentissage à partir d'exemples élaborés en ciblant principalement la façon de guider l'attention de l'utilisateur vers des mouvements oculaires caractéristiques des experts. Spécifiquement dans le domaine médical, (Litchfield, Ball et al. 2010) ont mené une série de trois expérimentations pour observer comment guider l'attention d'un clinicien en utilisant les mouvements oculaires d'autres cliniciens dans le domaine de la radiologie. Il s'agissait de valider l'amélioration des performances des radiologues lors de la lecture des radiographies thoraciques.

- Dans l'expérience 1, ils n'ont trouvé aucune différence significative entre la performance des membres dans les deux groupes : novice et expert.
- Dans l'expérience 2, seuls les novices s'améliorent lorsqu'ils reçoivent les indices sur les mouvements oculaires des experts.
- Dans l'expérience 3, ils ont réexaminé la contribution de l'image, de la tâche et du niveau d'expertise des modèles afin de valider le gain en apprentissage.

Les conclusions de (Litchfield, Ball et al. 2010) suggèrent que guider l'attention visuelle des radiologues novices via les mouvements oculaires des radiologues experts donne un gain en performance comportementale de courte durée pour les novices quant à leurs mouvements oculaires pour rechercher adéquatement les indices décisionnels dans les images de radiographies.

9.3 Fondement du cadre pédagogique dans le STI EyeLab

En se basant sur les résultats de recherche concluants de l'utilisation de l'oculométrie cognitive dans le domaine des STI ainsi que dans le transfert d'expertise du RC nous avons défini les principes pédagogiques qui soutiennent le fonctionnement du STI EyeLab.

De nombreuses études ont prouvé qu'aux premiers stades de l'apprentissage, les apprenants tirent davantage profit à avoir des exemples élaborés (c'est-à-dire des problèmes avec des solutions ou traces de résolution) que de la tentative de résoudre les problèmes sans aide (Rourke & Sweller, 2009; Schwonke et al., 2009; Sweller, Ayres et Kalyuga, 2011). Ces études démontrent un gain en apprentissage à utiliser des exemples élaborés et Sweller (2006) explique ces résultats en se basant sur la théorie de la charge cognitive. En effet, lorsque les exemples fournissent des explications étape par étape et les connaissances nécessaires pour résoudre les problèmes, il y a une diminution de la charge cognitive dans la mémoire de travail des apprenants. Ceci

s'applique surtout dans le cas de l'apprenant novice qui n'a pas les connaissances du domaine préalable pour une auto-explication des solutions. L'expertise perceptivo-décisionnelle dans le domaine médical est empirique et acquise par la pratique dans des situations réelles ou dans un environnement de simulation. De ce fait, le novice ne dispose pas à priori de connaissances du domaine pour définir adéquatement ses mouvements oculaires.

Toutefois, les exemples ne doivent pas être fournis à l'apprenant sous forme d'exercices d'apprentissage passifs (c'est-à-dire simplement du matériel à lire ou à voir). Au contraire, ils doivent être interactifs en demandant à l'apprenant d'explorer l'exemple et fournir des explications pas à pas (Brusilovsky et Yudelso, 2008; McLaren et Isotani, 2011).

Dans le chapitre VII, avec l'architecture DeepConvLSTM nous avons fait la classification d'extraits de séquences oculaires. Par la suite, dans le chapitre VIII, avec l'algorithme MDSE nous avons fait l'extraction de motifs oculaires contextualisés dans le but de dégager des régularités caractéristiques du novice et de l'expert. Les extraits de ces deux approches sont utilisés dans l'implémentation du STI EyeLab. Les motifs oculaires extraits sont représentés dans une perspective pédagogique exploitable par un STI. Avec une grande similarité comportementale dans les mouvements oculaires des experts, le modèle de l'expert est basé principalement sur les motifs extraits des mouvements oculaires des experts. L'approche globale dans EyeLab est une revisualisation de l'enregistrement vidéo de la performance du clinicien novice avec l'intervention d'un agent tuteur.

9.4 Cadre d'utilisation et fonctionnalités du STI EyeLab

Le STI EyeLab est utilisé lors du débriefing post-simulation par les cliniciens novices sous l'accompagnement des instructeurs du CAAHC et permet entre autres :

- Une revisualisation de la performance du clinicien novice avec les fixations et saccades ainsi que l'intervention d'un agent tuteur artificiel. L'agent tuteur utilise les analyses fournies par un module nommé *Sankofa*, pour afficher de manière proactive des rétroactions personnalisées en relation avec le parcours oculaire de l'apprenant. Sankofa est un symbole Adinkra commun à plusieurs peuples d'Afrique de l'Ouest. Il symbolise le fait de capitaliser sur ce qui a été bien fait dans le passé et de l'exploiter dans le présent afin de progresser de manière positive grâce à une utilisation bienveillante des connaissances acquises.
- L'analyse du processus décisionnel du clinicien novice en fonction des données issues de la cartographie de ses mouvements oculaires.
- L'extraction des règles comportementales permettant un enrichissement du modèle de l'expert.

Nous présentons ci-dessous la liste de certaines fonctionnalités par profil d'utilisateur.

Profil administrateur

- Gestion des apprenants (ajouter, modifier, supprimer, activer/désactiver)
- Gestion des instructeurs (ajouter, modifier, supprimer, attribuer des apprenants).
- Téléverser une vidéo, extrait de simulation ou image qui sera affiché aux apprenants
- Modifier la configuration (par exemple le nombre d'images par seconde pour capter les points de données des fixations et saccades des apprenants)

- Définir la durée minimale pour une fixation valide
- Définir manuellement les AOI dans les vidéos ou images téléversées
- Visualiser la page de résultats des données oculométriques des apprenants
- Exécuter l'algorithme MDSE pour l'extraction des motifs et mise à jour de la base des connaissances du modèle de l'expert
- Télécharger en format csv les données oculaires pour chaque apprenant.
- Visualiser la matrice de comparaison des *scanpath* (séquences de fixation et saccades)
- Visualiser le fichier de log des interventions du tuteur et modifier les messages à afficher

Profil instructeur

- Se connecter au STI EyeLab
- Visualiser la page de résultats des données oculométriques de ses apprenants (enregistrements des fixations et saccades, graphiques par exemple carte thermique)
- Supprimer l'expérimentation d'un apprenant
- Télécharger en format csv les données oculaires de ses apprenants
- Visualiser le modèle de ses apprenants

Profil Apprenant

- Se connecter au STI EyeLab
- Faire la calibration des mouvements oculaires
- Visualiser les vidéos et images qui lui sont attribuées
- Revisualiser sa performance en mode tutorat avec l'intervention du tuteur

9.5 Guidage des mouvements oculaires dans le STI EyeLab

Selon (Conati, Jaques et al. 2013) les recherches utilisant le guidage oculaire et les STI résument sur deux axes principaux. Le premier est la recherche dans les données oculaires des sources d'information pour la modélisation des apprenants en vue de formuler des rétroactions personnalisées. La seconde exploite les données de mouvements oculaires pour tenter de comprendre les comportements oculaires pertinents des apprenants.

Pour guider les novices à développer graduellement un modèle de mouvement oculaire des experts, il sera utile de leur fournir des indices personnalisés sur la façon dont les experts allouent leur attention visuelle pour soutenir leur RC. Cela est fait en utilisant des indices visuels pour guider l'attention des apprenants sur l'AOI pertinent afin d'effectuer les fixations suggérées et les saccades adéquates tout au long de l'évolution du cas clinique présenté dans la vidéo. Un modèle ouvert de l'apprenant est utilisé, car (Mathews, Mitrovic et al. 2012) présentent les avantages de ce modèle ouvert pour susciter la métacognition et l'autoévaluation de la part de l'apprenant.

Le STI EyeLab permet à l'apprenant de rejouer la vidéo de sa performance avec les indices visuels pour stimuler sa métacognition et sa réflexion. Le principe est basé sur la théorie du transfert de l'apprentissage (Li, Cohen et al. 2011). De ce fait pour le modèle du tuteur, la technique d'indices incrémentaux utilisés dans le traçage de modèle des Cognitive Tutors (Anderson, Corbett et al. 1995) a été utilisé. Les indices incrémentaux prennent la forme de:

- Messages de suggestion : si l'apprenant fait des fixations et saccades (*scanpath*) non conformes aux motifs de l'expert au moment des événements clés. Un score normalisé issu de l'algorithme de Smith-Waterman est utilisé pour la comparaison du *scanpath* du novice avec un motif de mouvement

oculaire de l'expert autour de l'évènement clé. Un seuil de similarité de 0,7 est défini. En dessous de ce seuil, les messages de suggestions sont affichés.

- Demande de validation ou avertissement : si plusieurs extraits de séquence de fixations dans les AOI sont classifiés de comportement de novice malgré les messages de suggestion du tuteur.
- Suggestions globales à la fin de la revisualisation.

Tous les indices intègrent des dimensions perceptuelles qui indiquent les AOI impliquées ainsi que l'ordre chronologique. Les indices incrémentaux sont affichés au-dessus de la zone d'affichage de la vidéo. Également, en mode tutorat les contours des AOI sont affichés à l'apprenant. Étant donné que les vidéos sont d'une durée relativement courte, il y a une pause automatique de la vidéo pour garder l'attention de l'apprenant durant l'affichage des indices.

Certaines captures d'écran du STI EyeLab sont disponibles dans les appendices:

APPENDICE A : Interface d'administration.

APPENDICE B : Interface de détection des pupilles et du visage. Cette première étape de détection du visage est exécutée par le clinicien novice avant d'entrer dans la session de tutorat. Elle génère 68 points de données pour le contour de la face, des yeux du nez et de la bouche. Ces points serviront à bâtir le modèle d'apprentissage machine pour la détection des mouvements oculaires.

APPENDICE C : Interface de détection et calibration de mouvements oculaires. Cette seconde étape est exécutée par le clinicien novice avant d'entrer dans la session de tutorat. Elle rajoute à la détection des yeux la détection du mouvement des pupilles.

APPENDICE D : Interface de EyeLab en mode tutorat. Elle présente :

1. le cas d'étude qui est spécifiquement un CICO (Can not Intubate / Can not Oxygenate);

2. des consignes à l'apprenant : Revisualiser la vidéo de votre performance en faisant attention aux indices de l'agent tuteur;
3. l'intervention de l'agent tuteur symbolisé par Sankofa. Les indices présentés contiennent les évènements clés et dans ce cas spécifique une suggestion.
4. La vidéo de la performance initiale de l'apprenant.
5. La trajectoire affichée correspond au mouvement oculaire suggéré provenant des séquences contextualisées extraites et enregistrées dans le modèle de l'expert.

APPENDICE E : Interface de comparaison des séquences de mouvements oculaires. Cette page permet à l'instructeur de visualiser des scores de similarité entre :

1. les séquences de mouvements oculaires des apprenants comparées à celle disponible dans le modèle de l'expert;
2. les séquences de mouvements oculaires des apprenants de manière croisée.

9.6 Discussion et positionnement

Les mécanismes de fonctionnement et les services tutoriels implémentés dans le STI EyeLab sont basés sur des résultats de recherche et des approches éprouvées dans le domaine des STI. Ces choix sont motivés par le fait que ces approches et services tutoriels éprouvés permettraient d'avoir des gains en apprentissage en utilisant le STI EyeLab. Il faudrait toutefois des expérimentations avec les cliniciens novices pour valider le gain en apprentissage escompté.

L'expérimentation du STI EyeLab sera une étude longitudinale in situ qui ne sera pas menée dans le cadre de cette thèse compte tenu de son ampleur. Le protocole devrait permettre de recruter suffisamment de cliniciens pour effectuer des préexpérimentations et post-expérimentations avec le STI EyeLab en vue d'évaluer les gains en apprentissage comportemental par rapport aux mouvements oculaires des

novices. D'abord il serait important de faire évaluer EyeLab auprès d'instructeurs du CAAHC. En effet ceux-ci auraient éventuellement des suggestions par rapport à l'interface proposée et aux stratégies pédagogiques choisies. Ils pourraient également expérimenter l'interface instructeur qui serait utilisée par des instructeurs pour appliquer le système à d'autres scénarios typiques d'intervention dans un but de formation. Cette évaluation par les instructeurs permettrait de valider de façon préliminaire le fonctionnement du STI.

Enfin STI EyeLab devrait être évalué en tant qu'outil pour l'enseignement. Cette évaluation pourrait être réalisée de la façon suivante : Deux groupes d'utilisateurs Novices seraient utilisés. Des deux groupes un aurait une interaction avec la simulation en utilisant EyeLab qui remplace le Tobii par une interaction sur le web, en suivant le protocole utilisé dans la thèse. Cependant le deuxième groupe aurait suite à cette interaction l'occasion d'expérimenter le STI. La trace du suivi oculaire de son interaction avec EyeLab lui serait présentée avec des commentaires et des conseils proposés par le tuteur artificiel *Sankofa*, qui permet d'afficher de manière proactive des rétroactions personnalisées. Suite à cette interaction avec le STI l'étudiant reverrait la simulation et le suivi oculaire serait évalué et comparé à sa performance lors du premier essai avec la simulation afin d'évaluer si son comportement se rapproche davantage du comportement d'un expert.

CHAPITRE X

CONCEPTION ET IMPLÉMENTATION DU STI EYELAB

Nous présentons dans ce chapitre certains détails de l'implémentation du STI EyeLab ainsi que les technologies utilisées. Étant un projet réalisé dans un cadre de recherches universitaires, les choix technologiques ont porté sur des plateformes, des langages et librairies qui sont *open source* avec une architecture modulaire. Notre objectif est de développer le STI EyeLab comme une application web moderne, réactive et dynamique. Plusieurs défis technologiques ont été relevés dans la conception et l'implémentation du STI EyeLab. Nous allons présenter certains défis, les choix d'architecture et les approches techniques pour l'implémentation.

10.1 Quelques défis de l'implémentation de EyeLab

EyeLab est une application web qui vise à capturer les mouvements oculaires des utilisateurs via la webcam. Ces mouvements oculaires seront captés au niveau du client et envoyés à des modules de traitement au niveau serveur. Le défi est de capter via le navigateur client les mouvements oculaires et de les transférer de manière optimale au serveur. Comme solution, nous avons opté pour une approche consistant à capturer des

images de l'utilisateur via la webcam connectée à l'ordinateur. Les images offrent l'avantage d'être plus légères et plus facile à compresser qu'une vidéo. Au niveau du serveur, les images successives seront traitées par un module spécial pour reconstituer une vidéo. Une vidéo se compose d'une série d'images individuelles, également connues sous le nom de trames. La fréquence des images donne l'impression d'un phénomène continu au niveau de la perception visuelle et détermine la fluidité de la vidéo. Ainsi entre 15 et 30 images par seconde les mouvements paraissent fluides et continus. La solution adoptée est constituée de l'utilisation de WebRTC pour capter via webcam les images de l'utilisateur; compresser et encoder ces images en Base64; enregistré temporairement ces images encodées dans le *Web Storage* du navigateur; transférer ces images au serveur en utilisant une communication via SocketIO.

- WebRTC (Web Real-Time Communication : communication en temps réel pour le Web) est un API open source développé en JavaScript/C++ qui permet des communications en temps réel sans plugin entre navigateurs. Ceci facilite le développement des applications qui utilisent le navigateur pour capter et transférer des données comme la voix, des images, des vidéos et des fichiers.
- Base64 : C'est un codage de l'information utilisant 64 caractères qui offre l'avantage d'une représentation de données binaires (par exemple, image) sous forme de chaîne de caractères compressée. La représentation en Base64 facilite le transfert de données binaires du client au serveur.
- Web Storage : Web Storage est une solution adaptée aux besoins de stockage de données plus ou moins volumineuses dans le navigateur. Web Storage dispose de deux interfaces : `sessionStorage` et `localStorage` dont la différence porte sur la durée et la portée des données persistées.
- La communication via SocketIO : de manière typique la communication entre le navigateur client et le serveur se passe via le protocole http. SocketIO est une alternative intéressante pour transférer de gros volumes de données entre client

et serveur avec une meilleure vitesse de transfert et temps de réponse des requêtes client/serveur.

Il existe des bibliothèques JavaScript pouvant être utilisées pour capter les mouvements oculaires via la webcam. Une des plus populaires est : WebGazer.js¹⁰. La bibliothèque WebGazer présente certaines limites par rapport aux exigences du projet. Elle ne transmet pas de données, d'images ou de vidéos du navigateur client vers un serveur. Ceci pose un problème pour créer la vidéo de revisualisation de la performance du clinicien novice avec les fixations et saccades. Également, les données générées ne sont pas suffisantes pour les exigences de projet.

Un autre défi est le choix d'une infrastructure de développement côté serveur. Mis à part les exigences classiques comme la maintenabilité, l'évolutivité, la productivité; le choix d'infrastructure doit tenir compte d'autres exigences. Il doit permettre de faire des calculs scientifiques sur de gros volumes de données, de faire du traitement d'images, et d'intégrer des algorithmes d'apprentissage machine. En tenant compte de ces exigences, l'infrastructure de développement web Django utilisant le langage Python a été choisie. Avec ce choix, nous disposons de plusieurs bibliothèques spécialisées dont :

- Python-SocketIO : qui est une bibliothèque de Python pour les communications via SocketIO.
- OpenCV (Open Computer Vision) : qui dispose de nombreuses fonctionnalités de traitement d'images, de traitement de vidéos et de reconnaissance de forme.

¹⁰ <https://webgazer.cs.brown.edu/>

- Dlib : qui est une librairie de traitement d'image développée en C++ pouvant être intégrée dans un projet Python.
- Numpy : qui est une librairie de Python pour faciliter la manipulation et les calculs mathématiques sur des données sous forme matricielle ou tableaux multidimensionnels.
- Pandas : qui est une librairie de Python pour lire et manipuler des données stockées dans des formats comme le csv.
- SciPy : qui est une librairie de Python contenant des modules pour les calculs scientifiques.
- Matplotlib : qui est une librairie de Python permettant de visualiser des données sous forme de graphiques, de tracer des fonctions mathématiques.
- Seaborn : qui est une librairie de Python permettant de générer des visualisations plus poussées que Matplotlib (par exemple les heatmaps).

10.2 Le choix d'architecture

Nous avons opté pour une architecture en micro-services pour l'agilité du développement, la diversité des technologies utilisées et la mise à l'échelle. Tout d'abord du point de vue de l'agilité, étant donné que les micro-services sont déployés indépendamment, il est plus facile de gérer les corrections de bogues et les versions de fonctionnalités. Nous pouvons ainsi mettre à jour un service sans redéployer l'application entière. De plus, de nouvelles fonctionnalités peuvent être retardées en attendant qu'une correction de bogue soit intégrée, testée et publiée. Du point de vue des technologies, notre application utilise un nombre important de bibliothèques externes et d'API open-source. Une approche micro-service permet de choisir la pile de technologie la plus appropriée pour chaque service. Du point de vue de la performance, les services peuvent être mis à l'échelle indépendamment, ce qui permet de mettre à

l'échelle les services qui nécessitent le plus de ressources, sans faire évoluer l'ensemble de l'application. À noter que, pour le moment, les données sont centralisées et accessibles par tous les micro-services, comme dans le cas des applications monolithiques. Nous sommes conscient que cette configuration contredit au principe de l'isolation des données dans les architectures en micro-services. Dans des travaux futurs, nous allons distribuer les données afin de faciliter la modification du schéma et l'évolution du volume des données. Le diagramme des micro-services est présenté dans la figure 31.

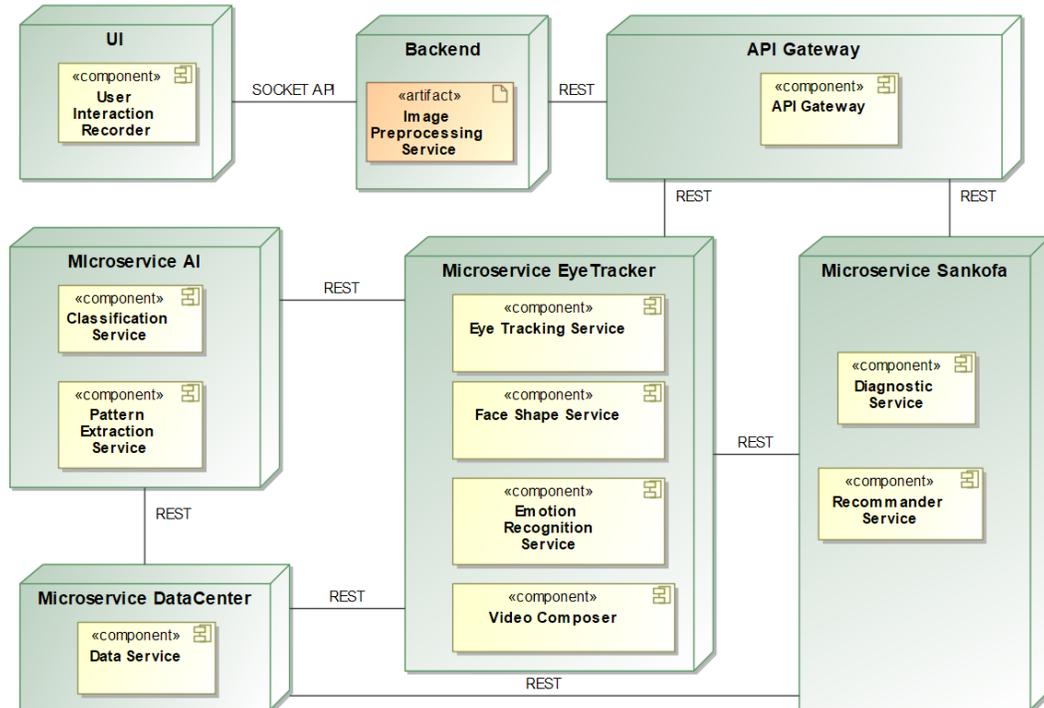


Figure 30 : Diagramme des micro-services dans EyeLab

10.3 Quelques étapes du mécanisme de fonctionnement avec les composants associés

Nous présentons de manière sommaire quelques étapes du mécanisme de fonctionnement en prenant comme exemple le profil d'un clinicien novice. Les composants impliqués sont également présentés à la figure 32.

Le clinicien novice se connecte à la plateforme EyeLab en utilisant des identifiants attribués au moment de la création du compte. Il lance une vidéo de la simulation qui lui a été attribuée par un instructeur. La vidéo contient des AOI statiques préalablement définis par l'instructeur, les coordonnées qui délimitent les AOI sont stockées dans le centre de données. Les éléments du composant UI de EyeLab font le chargement de la vidéo dans le navigateur et déclenchent automatiquement la webcam. La webcam est utilisée pour prendre des images du visage du clinicien novice à chaque frame de la vidéo. Ces images sont utilisées pour extraire les points de l'écran visualisés par le clinicien novice à chaque frame de la vidéo de simulation.

Les images sont encodées et compressées par les composants UI, puis sont envoyées au serveur à l'aide de socketIO. Plusieurs traitements sont effectués par des modules spécialisés du composant Tracker. Ces modules vont traiter les images reçues, extraire les points de données oculaires, générer une nouvelle vidéo en fonction des points de données oculaires et des frames de la vidéo. Pour un meilleur résultat de superposition, les images de la caméra sont converties selon la taille et la résolution d'écran de l'utilisateur. Les traitements du composant Tracker sont effectués sur de nouveaux threads afin qu'ils n'affectent pas la performance du site web principal. Certaines étapes de ces traitements sont décrites dans la prochaine section (Processus d'extraction des données oculaires).

Les coordonnées (x,y) des points regardés à l'écran sont converties en fixation dans les AOI avec le timestamp associé; ce qui donne des points de données de mouvements oculaires. Ces données sont envoyées au module de centre de données. Les détails des éléments du centre de données sont présentés dans la section : Les choix technologiques pour la persistance de données.

Chaque suite de 500 points de données de mouvements oculaires du clinicien novice est envoyée à DeepConvLSTM pour classifier le comportement comme novice ou expert. Le résultat de la classification est mis dans le modèle de l'apprenant. L'extraction des motifs de mouvements oculaires via l'algorithme MDSE est exécutée manuellement sur une base périodique. Les motifs extraits servent à l'enrichissement du modèle de l'expert.

Le composant Sankofa exploite les traces de résolutions (séquences et motifs de mouvement oculaires) disponibles dans le centre de données pour effectuer un diagnostic et recommandation exploitable par le module de tutorat.

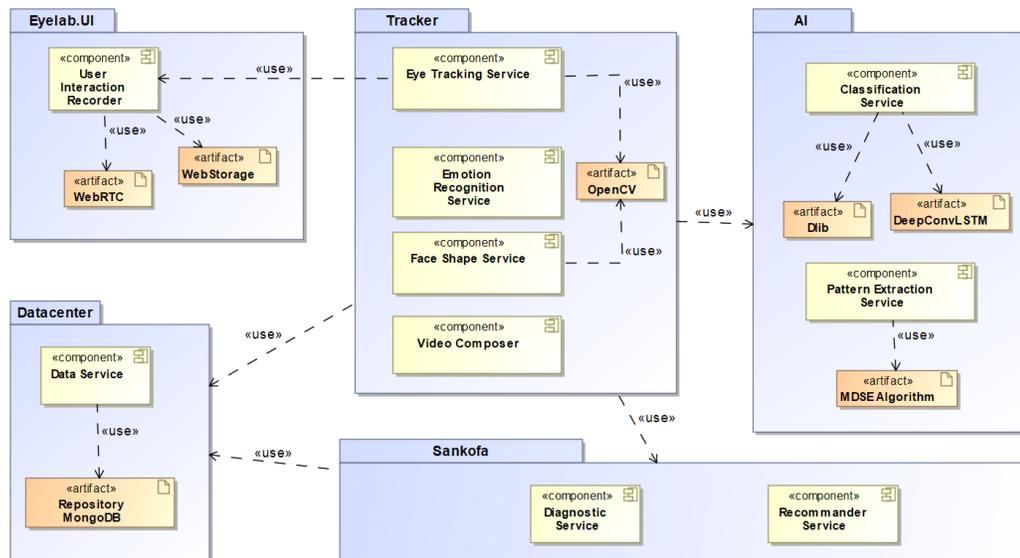


Figure 31 : Diagramme des composants de EyeLab

10.4 Processus d'extraction des données oculaires

Plusieurs étapes sont impliquées dans ce processus qui inclut la détection de la forme du visage et la détection du mouvement oculaire. La détection de l'émotion via l'expression faciale est disponible de manière optionnelle.

Détection de la forme du visage

Pour la détection de la forme du visage, le modèle de points de contour de visage de dlib 68 a été utilisé. Dans EyeLAB, la première image de l'utilisateur captée par la webcam est traitée pour obtenir le résultat de la forme du visage. Par la suite, cette forme est stockée en mémoire et la même forme de visage est conservée tout au long de l'expérimentation avec des réorientations au besoin.

Détection des mouvements oculaires

Il n'y a pas de librairie Python pour la détection du mouvement oculaire. Ce module a été développé spécifiquement pour le projet de recherche et contient les étapes ci-dessous :

Étape 1 : Pour détecter les mouvements oculaires, il faudrait commencer par trouver le visage de la personne ainsi que les éléments du visage principalement les yeux. Pour trouver les contours des éléments du visage, l'approche Haar Cascades qui déjà implémentée dans OpenCV est utilisée. Elle est très compacte et retourne 68 points de données pour le contour de la face, des yeux, du nez et de la bouche (figure 33).

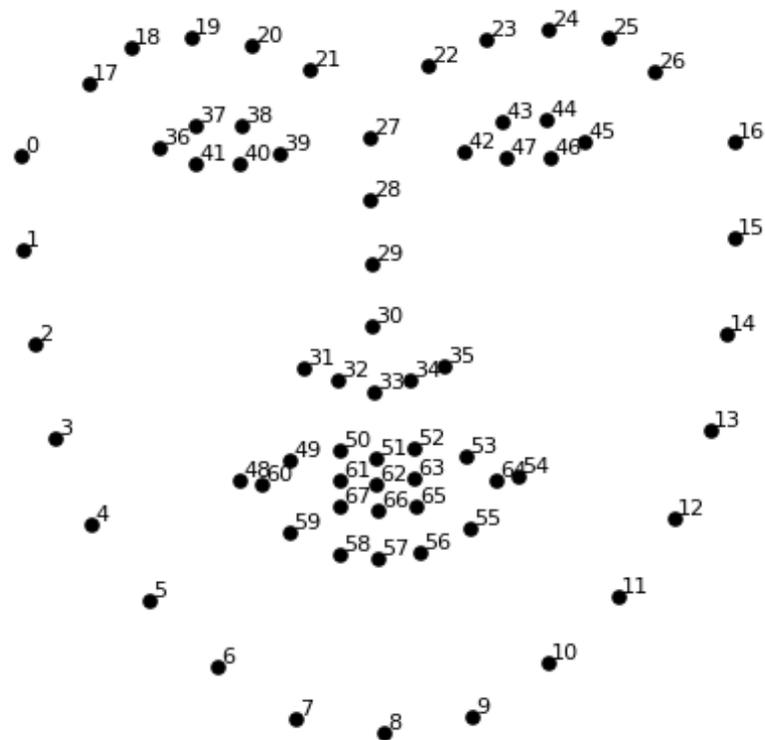


Figure 32 : L'approche Haar Cascades avec 68 points de données pour le contour de la face, des yeux du nez et de la bouche

Étape 2 : Après avoir trouvé la forme et les éléments visage, il faudrait obtenir la position des yeux à partir des 68 points de données descriptifs. Comme sur la figure 33, les points 36-41 contiennent l'œil gauche et les points 42-47 contiennent l'œil droit.

Étape 3 : Pour détecter les mouvements oculaires, il faudrait trouver la position des yeux. Par la suite, la caractéristique la plus importante est le mouvement des pupilles. Il faudrait donc détecter les pupilles ainsi que les mouvements des pupilles même si la position du visage ne change pas.

Étape 4 : Pour détecter les pupilles et ses mouvements, nous avons utilisé des techniques de traitement d'image combinées avec des modèles mathématiques développés sur mesure. Le principe est de trouver un contour plus foncé (différentes couleurs de pupilles : noir, gris, vert, bleu) dans un fond blanc selon un ratio ajustable. Le contour plus foncé constitue les pupilles et les changements de coordonnées du contour plus foncé définissent les mouvements des pupilles.

Étape 5 : Après avoir trouvé la forme du visage, la position des yeux et les mouvements de pupilles, il faudrait maintenant obtenir le point de l'écran où la personne regarde. Pour cela l'angle entre les yeux et la caméra par rapport à la taille de l'élément présenté sur l'écran sont utilisés pour bâtir un modèle mathématique. Ce modèle permet de calculer les coordonnées des points observés. Une visualisation par carte thermique (Heatmap) a été utilisée pour une meilleure compréhension dans les changements de valeurs des coordonnées à l'écran.

NOTE : Les étapes 3 à 5 correspondent à une extraction d'attributs (*features extraction*) pour un projet en IA, car nous ne disposons pas de jeu de données pour bâtir un modèle de détection de mouvements oculaires. Ultimement, un objectif après avoir obtenu beaucoup de données dans la plateforme EyeLab, serait de bâtir un modèle d'apprentissage machine pour la détection des mouvements oculaires.

Les étapes du processus d'extraction des données oculaires sont illustrées dans le diagramme de processus de la figure 34.

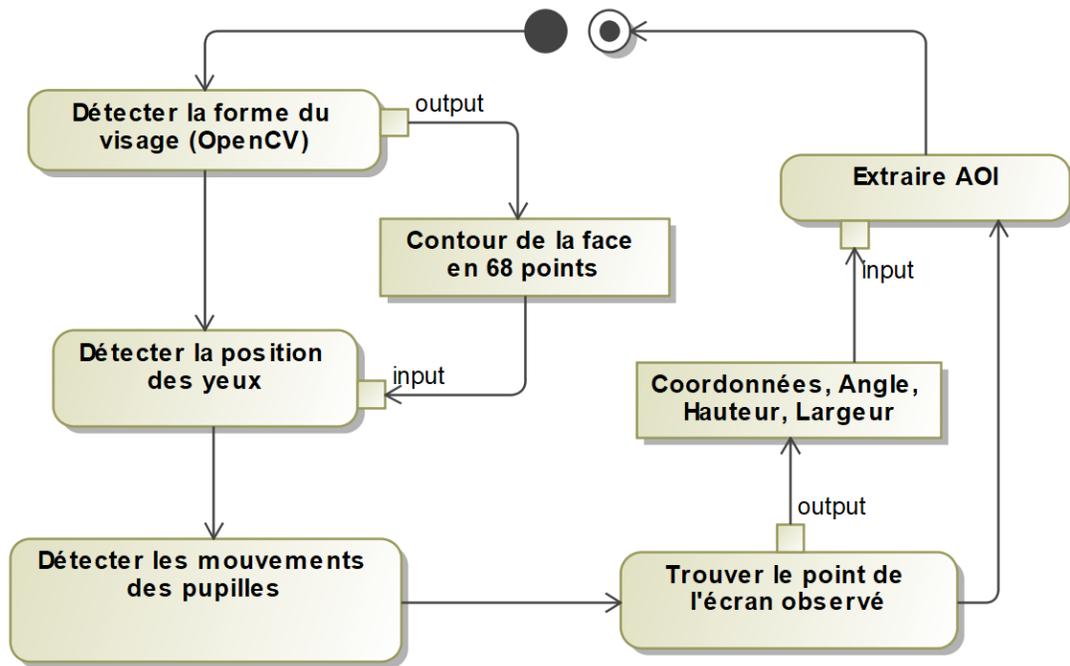


Figure 33 : Processus d'extraction des données oculaires

10.5 Les choix technologiques pour la persistance de données

Les choix technologiques pour la persistance de données doivent être faits en tenant compte des spécificités liées aux données du projet. Il s'agit notamment des trois V descriptifs des données de Big Data à savoir le : Volume, Vitesse et Variété. À ces 3V nous rajoutons la périodicité P.

Volume : désigne la quantité de données générée. Lorsqu'un utilisateur regarde la vidéo, les données sont générées à chaque frame. En supposant que la vidéo est constituée de 30 frames par seconde; pour une vidéo de 13min, il y a $13 \times 60 \times 30 = 23400$ données générées par un utilisateur durant la session d'expérimentation. EyeLab est

déployé comme application web afin d'être utilisé simultanément par plusieurs utilisateurs durant les sessions d'expérimentations. Ce volume de données serait donc multiplié par le nombre d'utilisateurs simultanés.

Vitesse : désigne la fréquence à laquelle les données sont générées. Cette vitesse génère le volume dans les cas des données du projet. À titre indicatif en prenant l'exemple des données pour le volume, un seul participant génère 1800 données par minute. Ce nombre est multiplié par le nombre d'utilisateurs simultanés.

Variété : désigne la diversité dans le format et type des données. Les données des mouvements oculaires contiennent les AOI regardés par un clinicien et le *timestamps* associés. La granularité et le format décrivant chaque point de données peuvent changer. Par exemple, si un instructeur décide d'ajouter des AOI à une vidéo existante. Également pour des fins d'analyse, d'autres informations contextuelles peuvent être rajoutées aux points de données comme la forme du visage, l'expression faciale, etc.

Périodicité : les expérimentations auront lieu de manière ponctuelle étant donné qu'il y a des démarches administratives d'approbation pouvant prendre quelques mois avant chaque expérimentation. Les 3V sont respectés durant les périodes d'expérimentations. Entre les expérimentations, la plateforme est utilisée de manière ponctuelle par les instructeurs.

Avec les contraintes 3V-P, le choix pour la persistance des données a porté sur une base NoSQL, spécifiquement MongoDB. Le paradigme NoSQL nous offre beaucoup des solutions au regard des contraintes présentées. Nous pouvons citer entre autres :

- L'enregistrement de données non structuré typiquement en format JSON

- La rapidité d'insertion de grand volume de données : cette performance d'insertion est due à la simplicité de la structure et l'absence de contraintes d'intégrité à respecter avant les insertions de données.

Il y a toutefois certains inconvénients comme la lecture et l'agrégation complexe des données. Ceci peut nécessiter l'utilisation de pipeline de transformation et agrégation de données. Toutefois, cet inconvénient ne représente pas un problème pour le projet d'autant plus qu'il y aura seulement des lectures occasionnelles des données.

CONCLUSION : CONTRIBUTIONS DE LA THÈSE, PERSPECTIVES ET TRAVAUX FUTURS

Nous faisons ici un sommaire des recherches effectuées, des résultats obtenus ainsi que des conclusions pertinentes. Ceci met en évidence le caractère bidisciplinaire informatique et sciences cognitives des contributions. Nous y décrivons également les limites des solutions proposées ainsi que les perspectives envisageables pour la suite de la recherche. Les travaux futurs y seront présentés dans une optique de validation longitudinale du STI, et d'évaluation de la généralité des solutions proposées par leurs applications dans d'autres domaines.

Contribution de la thèse

Ce projet doctoral adresse la problématique du transfert d'expertise. Il s'agit ici d'une expertise comportementale spécifiquement les mouvements oculaires dans le domaine de l'anesthésiologie pour supporter le raisonnement clinique. Nous l'avons nommée l'expertise perceptivo-décisionnelle. C'est l'aptitude du clinicien à réguler adéquatement ses mouvements oculaires pour capter les informations nécessaires à la compréhension du cas clinique pour supporter son raisonnement et ses décisions cliniques. Trois problématiques ont été soulevées notamment :

1. Existe-t-il une différence entre la perception visuelle du novice et de l'expert ?

2. Existe-t-il une interdépendance entre la perception visuelle et le raisonnement clinique dans le cadre de l'expertise perceptivo-décisionnelle ?
3. Comment peut-on acquérir et modéliser les connaissances du domaine pour l'expertise perceptivo-décisionnelle ?

Ces problématiques ont permis de formuler certaines hypothèses qui ont servi à définir nos objectifs de recherche.

Notre première hypothèse est que l'analyse des données issues de la perception visuelle pourrait permettre la détection de régularités spécifiques aux novices et aux experts. Cette hypothèse a été vérifiée en mettant en place un protocole expérimental qui a permis de collecter et d'analyser les données oculaires de cliniciens novices et experts dans une tâche authentique qui nécessite la perception visuelle et le raisonnement clinique. Une grande quantité de données a été collectée ce qui a nécessité le développement d'outils spécialisés de visualisation de données pour appréhender la dynamique de changement dans les données. Par la suite, des analyses statistiques ont été effectuées ainsi qu'un algorithme pour l'extraction de motifs de mouvements oculaires. Les résultats obtenus confirment l'hypothèse. En effet, l'analyse des mouvements oculaires autour d'événements clés ont permis d'observer des différences dans les mouvements oculaires. Par la suite, l'algorithme MDSE a permis d'extraire des motifs séquentiels caractéristiques des mouvements oculaires des novices et des experts.

Notre deuxième hypothèse est que certaines connaissances de l'expertise perceptivo-décisionnelle peuvent être extraites et modélisées. Il s'agit ici d'une modélisation pédagogique, c'est-à-dire exploitable pour supporter un tutorat. Cette hypothèse a été vérifiée avec le développement et l'implémentation de l'algorithme MDSE qui

fournit en sortie des motifs exploitables caractéristiques du comportement oculaire des experts. Il s'agit notamment du début du comportement, l'ordre temporel dans lequel les mouvements oculaires ont été exécutés, la durée dans chaque AOI observés ainsi que la durée globale. Les motifs extraits par l'algorithme MDSE sont donc exploitables par le module de tutorat d'un STI via une approche de modèle tracing pour effectuer des interventions tutorielles.

Notre troisième hypothèse est qu'un STI pourrait être utilisé pour transférer certaines connaissances au novice pour accélérer son cheminement vers l'expertise. Il s'agit ici de connaissances comportementales, spécifiquement les mouvements oculaires. Cette hypothèse a été partiellement vérifiée d'autant plus que le STI EyeLab est développé. Toutefois, l'expérimentation pour évaluer le gain en apprentissage ne fait pas partie de la thèse.

Somme toute avec les hypothèses validées, les travaux dans cette thèse ont permis d'atteindre les objectifs ci-dessous :

Objectif 1 : Étude exploratoire et comparative des mouvements oculaires de cliniciens novices et experts. Cet objectif a été atteint en mettant en place un protocole expérimental authentique suivi de trois analyses sur les données recueillies. La première est une analyse préliminaire qui a permis de comparer les novices aux experts en utilisant des statistiques descriptives sur les fixations dans les AOI. La deuxième est une analyse comportementale plus fine qui a permis de comparer les mouvements oculaires des novices et experts autour des événements clés sur le plan de la sensibilité situationnelle. La dernière analyse a porté sur les trajectoires globales des fixations des saccades. Elle nous a permis de faire des comparaisons croisées des *scanpaths* via un algorithme de comparaison de séquences.

Objectif 2 : classification des séquences des mouvements oculaires pour alimenter le modèle de l'apprenant et extraction de motifs oculaires comportementaux. Les données des mouvements oculaires présentent des caractéristiques complexes à considérer dans les analyses. Elles sont en effet très bruitées; spatiales avec la dynamique de changement des fixations/saccades et séquentielles avec un ordre temporel à conserver. Deux architectures ont été développées et comparées pour faire la classification. La première est une architecture profonde de réseau de neurones artificiel nommée DeepConvLSTM qui combine un réseau de neurones à convolution pour l'apprentissage spatial et un réseau LSTM pour l'apprentissage séquentiel. La seconde architecture nommée WKM-1NN combine des techniques d'apprentissage statistique de compression dynamique temps et de classification. DeepConvLSTM a été la plus performante basée sur des métriques de comparaison de classifieurs en apprentissage machine.

Pour l'extraction de motifs oculaires comportementaux, nous avons fait un état de l'art sur les algorithmes d'extraction de motifs dans les données transactionnelles ainsi que les données éducationnelles. Les limites de ces algorithmes ont justifié le développement d'un nouvel algorithme d'extraction de motifs séquentiels en considérant les spécificités des données oculaires comme l'ordre temporel, la durée des fixations valides et une sortie exploitable dans un STI. L'algorithme nommé MDSE (Multidimensional Dynamic Sequence Extraction) a permis d'extraire les motifs caractéristiques des mouvements oculaires des experts, des novices pour des fins de comparaisons.

Objectif 3 : Spécifications et conception d'un STI. Le STI nommé EyeLab est développé comme une application web moderne, réactive et dynamique. Plusieurs défis technologiques ont été relevés dans la conception et l'implémentation du STI EyeLab.

Nous avons présenté certains défis, les choix d'architecture et les approches techniques pour l'implémentation.

Les objectifs démontrent une perspective intégratrice qui nécessite l'analyse de mouvement oculaire; l'utilisation de l'IA principalement l'apprentissage machine avec les réseaux de neurones et les fouilles de données séquentielles, le développement d'un STI. Sur la base de ces trois objectifs qui ont été atteints durant cette thèse, nous dégageons les contributions ci-dessous :

En sciences cognitives, la contribution est sur un plan méthodologique dans l'étude des mouvements oculaires dans la réalisation d'une tâche cognitive. Cette méthodologie inclut l'approche expérimentale, ainsi que les techniques d'analyse utilisées.

En informatique, la contribution se décline en deux axes :

1. L'algorithme MDSE d'extraction de motifs séquentiels dédié aux contraintes et spécificités des données séquentielles des mouvements oculaires. L'implémentation de l'algorithme offre également l'avantage de fournir en sortie un modèle de représentation de motifs qui sous-tendent des connaissances perceptivo-décisionnelles exploitables dans un STI
2. La plateforme web EyeLab dédiée aux expérimentations, à la collecte et à l'analyse des mouvements oculaires dans un processus cognitif de résolution de problème. EyeLab inclut un STI développé dans une architecture modulaire.

Les travaux réalisés ont été acceptés et présentés à plusieurs conférences ainsi qu'une publication dans une revue scientifique. Il s'agit notamment :

- **Komi Sodoké**, Roger Nkambou, Aude Dufresne, Issam Tanoubi « Toward a deep convolutional LSTM for eye gaze spatiotemporal data sequence classification ». Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020). 672-676

- Mathieu Tourangeau, Issam Tanoubi, Roger Perron, Marie-Ève Bélanger, **Komi Sodoké**, Pierre Drolet, Arnaud Robitaille , L Mihai Georgescu « Comparing the Visual Perception of Novice and Expert Clinicians: An Exploratory Simulation Based Study ». *Anaesthesia Critical Care & Pain Medicine*. 2020;39(2):247-249.
- Issam Tanoubi : « Visual perception exploration using eye-tracking technology in a Cannot intubate / Cannot oxygenate simulation. Présentation orale d’abrégé dans session communications libres encadrées “Simulation en Anesthésie ». 22/9/2017 ;
- Issam Tanoubi, **Komi Sodoké**, Roger Perron et co. « Utilisation du Eye Tracking pour l’exploration du raisonnement clinique et de l’expertise : Quelle méthodologie de recherche ? Comment interpréter les données? »

Quelques limites dans les travaux

Augmentation du nombre de participants

La principale difficulté rencontrée au cours des travaux de cette thèse a été celle de la collecte de données. Il y a relativement peu de cliniciens en anesthésiologie comparés à d’autres domaines de la médecine. De plus, les cliniciens devraient se déplacer au laboratoire Cité à l’Université de Montréal. Théoriquement, il aurait été possible d’avoir plus de cliniciens, mais des contraintes personnelles, des urgences médicales, etc. ont poussé certains à reporter à plusieurs reprises le rendez-vous d’expérimentation ou l’annuler tout simplement.

Une solution à ce problème de recrutement aurait été d'étendre le domaine d'étude à des disciplines connexes ou complémentaires comme l'inhalothérapie, etc. Également avec le déploiement de la plateforme web EyeLab, les cliniciens n'ont plus besoin de faire le déplacement au laboratoire Cité pour faire les expérimentations. Tout ordinateur connecté sur Internet, avec un navigateur web récent équipé d'une webcam, microphone et haut-parleur serait suffisant pour conduire les expérimentations. Les expérimentations peuvent donc se faire directement dans les hôpitaux durant les heures de pause.

Soumettre les motifs extraits à une meilleure contextualisation et validation des experts

Les motifs extraits sont contextualisés autour des événements clés qui constituent les phases de différenciation des mouvements oculaires des novices et des experts. Ceci ne garantit toutefois pas leur pertinence au regard de l'expertise du domaine. Par exemple, nous avons obtenu un total de 152 motifs avec des caractéristiques (début, durée, ordre temporel dans les AOI, fin) plus ou moins différentes. Même si ces motifs ont été extraits automatiquement à partir des données oculaires par apprentissage machine, il aurait été intéressant d'avoir une validation des experts ou des instructeurs du CAAHC pour sélectionner les plus pertinentes à intégrer dans le module de l'expert du STI EyeLab.

Cadre formel et automatisé autour des rétroactions dans le STI EyeLab

Un des objectifs d'un STI est de fournir des rétroactions hautement personnalisées à chaque apprenant. Dans le STI EyeLab, ces messages prennent la forme d'indices incrémentaux en intégrant les dimensions perceptuelles dans les AOI pertinents. Les messages des indices ont été rédigés au cas par cas. Ceci signifie que dans un contexte clinique donné, des cliniciens novices ayant des mouvements oculaires non conformes auront le même message. Le STI EyeLab dispose toutefois d'un module d'analyse

d'expression faciale afin de détecter l'état affectif, ainsi que d'un historique comportemental (Novice Versus Expert retourné par l'architecture DeepConvLSTM). Ces informations sont stockées dans le modèle de l'apprenant et peuvent être exploitées dans un cadre formel pour la génération automatique de messages très personnalisés.

Travaux futurs, expérimentations du STI et évaluation de la généricité

Les travaux futurs comprennent principalement une expérimentation du STI et une évaluation de la généricité des approches proposées et algorithmes développés.

Le STI EyeLab a été conçu et développé, mais il devrait être expérimenté pour en évaluer l'interface d'interaction, le gain en apprentissage et l'efficacité pour la formation des professionnels. Un protocole pour l'expérimentation du STI EyeLab dans le cadre d'une étude in situ a été présenté.

L'évaluation de la généricité permettra de valider si la méthodologie, les approches d'analyses, les architectures d'apprentissage machine, les algorithmes et la plateforme EyeLab pourraient être exploités de manière efficiente dans d'autres domaines présentant des besoins similaires. À titre d'exemple, nous pouvons citer le pilotage d'avion, car le balayage visuel des instruments du tableau de bord est lié aux connaissances relatives aux contrôles de l'état de l'appareil à différentes phases de vol. Nous avons également de nombreux jeux vidéos qui nécessitent une bonne coordination de perception visuelle à l'écran suivi de décisions de la part du joueur. Finalement, nous pouvons citer les interfaces humain-machine, la navigation sur des sites de e-commerce, etc. Dans cadre spécifique de futures collaborations, le CAAHC envisage l'utilisation du STI EyeLab pour des recherches portant sur les mouvements oculaires des infirmières sur les formulaires et ordonnances médicaux. L'évaluation de

la généricité va englober aussi bien le cadre fondamental que les aspects d'application de l'extrait de notre recherche.

ANNEXE A

QUESTIONNAIRE POST EXPÉRIMENTATION

1. Vous êtes ?
 - Homme
 - Femme

2. Quel âge avez-vous?
 - 28 ans

3. Vous êtes Résident (quel niveau) ou Patron (combien d'années de pratique)
 - Résident 4

4. Je suis à l'aise à prendre en charge une situation CICO (intubation et ventilation impossible)
 - Fortement en désaccord
 - En désaccord
 - Neutre
 - En accord
 - Fortement en accord

5. Avez-vous déjà pris en charge une situation CICO chez un vrai patient?
 - OUI
 - NON

6. Cette situation était réaliste?
 - Fortement en désaccord
 - En désaccord
 - Neutre
 - En accord
 - Fortement en accord

7. Je me suis senti à l'aise pendant la vidéo
 - Fortement en désaccord
 - En désaccord
 - Neutre
 - En accord
 - Fortement en accord

8. Selon vous, qu'avez-vous regardé le plus
 - Pouls
 - Saturation
 - Fréquence cardiaque
 - Rythme respiratoire
 - Cou/Collet
 - Médecin
 - Assistants

9. Commentaires généraux (expérimentations, prise en charge, évolution du cas, etc.)
 - Le vidéo était réaliste, je n'étais simplement incertain par rapport aux commentaires que je devais faire ou non.
 - Il était intéressant de visualiser une telle situation d'urgence/crise en tant que spectateur. La prise en charge a été adéquate en termes de gestes techniques et il n'y a pas eu d'erreur de fixation dans les gestes techniques; c'est-à-dire que lorsqu'une manœuvre était effectuée, il y avait une évaluation rapide et si elle était infructueuse, il passait à l'étape suivante. L'urgence de la situation a été reconnue et prise en charge rapidement.

- Je crois qu'il aurait dû avoir un meilleur échange d'information entre l'infirmière et Dr Tanoubi quant au contexte de la désaturation, le type de chirurgie, les antécédants du patient, comment la chirurgie s'est déroulée, le médicament injecté juste avant la désaturation.
- il y aurait pu avoir une demande à l'aide légèrement plus rapidement dont pour le chirurgien ORL, et l'examen sous le collier cervical plus rapidement pour voir s'il y avait une raison cachée par ce dernier qui aurait pu expliquer la désaturation
- L'évolution du cas m'a semblé réaliste puisque le patient a été réanimé selon la cause de son ACR, soit l'hypoxie.

Nous vous remercions de votre participation à notre projet de recherche. Votre implication et votre temps sont grandement appréciés.

ANNEXE B

EXTRAIT DE SÉQUENCES PRÉTRAITÉES POUR L'EXÉCUTION DE L'ALGORITHME DE SMITH-WATERMAN

AOI-3AOI-1AOI-3AOI-1AOI-3AOI-2AOI-6AOI-4AOI-1AOI-3AOI-1AOI-1AOI-
3AOI-1AOI-2AOI-3AOI-1AOI-3AOI-1AOI-3AOI-2AOI-1AOI-3AOI-1AOI-3AOI-
1AOI-2AOI-1AOI-3AOI-1AOI-1AOI-3AOI-2AOI-1AOI-3AOI-1AOI-2AOI-5AOI-
2AOI-3AOI-1AOI-1AOI-3AOI-1AOI-3AOI-3AOI-1AOI-5AOI-2AOI-5AOI-2AOI-
2AOI-1AOI-6AOI-4AOI-6AOI-3AOI-7AOI-5AOI-3AOI-1AOI-2AOI-2AOI-1AOI-
2AOI-5AOI-1AOI-3AOI-2AOI-1AOI-3AOI-5AOI-2AOI-3AOI-1AOI-3AOI-5AOI-
2AOI-2AOI-3AOI-5AOI-6AOI-4AOI-3AOI-6AOI-3AOI-1AOI-2AOI-5AOI-3AOI-
1AOI-3AOI-1AOI-2AOI-1AOI-2AOI-1AOI-3AOI-1AOI-3AOI-1AOI-3AOI-1AOI-
3AOI-2AOI-3AOI-1AOI-1AOI-3AOI-1AOI-1AOI-3AOI-1AOI-3AOI-2AOI-4AOI-
3AOI-3AOI-1AOI-2AOI-1AOI-3AOI-1AOI-3AOI-2AOI-1AOI-3AOI-1AOI-3AOI-
3AOI-4AOI-3AOI-1AOI-3AOI-1AOI-3AOI-6AOI-4AOI-3AOI-1AOI-3AOI-2AOI-
3AOI-1AOI-1AOI-3AOI-1AOI-2AOI-3AOI-1AOI-1AOI-3AOI-1AOI-3AOI-1AOI-
3AOI-1AOI-3AOI-3AOI-1AOI-3AOI-4AOI-5AOI-6AOI-4AOI-5AOI-4AOI-5AOI-
5AOI-2AOI-4AOI-3AOI-4AOI-3AOI-3AOI-4AOI-3AOI-4AOI-3AOI-1AOI-3AOI-
4AOI-3AOI-1AOI-3AOI-4AOI-3AOI-4AOI-3AOI-1AOI-3AOI-1AOI-3AOI-4AOI-
3AOI-3AOI-4AOI-5AOI-4AOI-6AOI-3AOI-6AOI-3AOI-3AOI-6AOI-6AOI-3AOI-
3AOI-1AOI-3AOI-4AOI-3AOI-5AOI-4AOI-6AOI-3AOI-3AOI-4AOI-6AOI-3AOI-
6AOI-3AOI-4AOI-3AOI-4AOI-6AOI-7AOI-6AOI-3AOI-4AOI-6AOI-3AOI-6AOI

ANNEXE C

EXEMPLE DE MOTIFS EXTRAITS PAR L'ALGORITHME MDSE POUR LE
NOVICE 5

| | sequence_pattern | Ind_AOI_length | sequence_ length_sum | start_ timestamp | end_ timestamp |
|----|--|---------------------------|-------------------------|---------------------|-------------------|
| 0 | ['AOI-2', 'AOI-1', 'AOI-6', 'AOI-3'] | [96, 43, 134, 39] | 312 | 24498 | 28098 |
| 1 | ['AOI-4', 'AOI-3'] | [35, 261] | 296 | 29547 | 33897 |
| 2 | ['AOI-3', 'AOI-4', 'AOI-3'] | [34, 42, 588] | 664 | 36213 | 46712 |
| 3 | ['AOI-4', 'AOI-5', 'AOI-4', 'AOI-3', 'AOI-1', 'AOI-2'] | [50, 39, 58, 133, 71, 82] | 433 | 48228 | 54611 |
| 4 | ['AOI-2', 'AOI-6', 'AOI-7'] | [87, 58, 39] | 184 | 70175 | 71792 |
| 5 | ['AOI-1', 'AOI-3', 'AOI-1'] | [103, 89, 138] | 330 | 75824 | 79607 |
| 6 | ['AOI-3', 'AOI-6'] | [99, 112] | 211 | 82174 | 84040 |
| 7 | ['AOI-4', 'AOI-5', 'AOI-1'] | [43, 59, 47] | 149 | 85906 | 87673 |
| 8 | ['AOI-1', 'AOI-3'] | [62, 39] | 101 | 92222 | 92872 |
| 9 | ['AOI-3', 'AOI-1'] | [37, 51] | 88 | 107753 | 108603 |
| 10 | ['AOI-4', 'AOI-3'] | [38, 158] | 196 | 122385 | 125018 |
| 11 | ['AOI-6', 'AOI-3'] | [44, 74] | 118 | 126218 | 127451 |
| 12 | ['AOI-4', 'AOI-3'] | [36, 53] | 89 | 130884 | 131767 |
| 13 | ['AOI-3', 'AOI-6'] | [72, 82] | 154 | 140466 | 141832 |
| 14 | ['AOI-3', 'AOI-2'] | [88, 120] | 208 | 148515 | 150514 |
| 15 | ['AOI-6', 'AOI-3'] | [36, 260] | 296 | 180610 | 184943 |

| | | | | | |
|----|--------------------------------------|-------------------|-----|--------|--------|
| 16 | ['AOI-2', 'AOI-5', 'AOI-4', 'AOI-3'] | [65, 59, 33, 445] | 602 | 191875 | 200824 |
|----|--------------------------------------|-------------------|-----|--------|--------|

APPENDICE A

INTERFACE D'ADMINISTRATION

Interface administration



- Ajouter un apprenant
- Liste des apprenants
- Voir les performances
- Récupérer des données de mouvement oculaires



- Ajouter un instructeur
- Liste des instructeur
- Attribuer des apprenants



- Téléverser une vidéo ou une image
- Liste des contenus



- Définir la durée minimale pour une fixation valide
- Définir la durée des frames
- Définir les AOI
- Visualiser la page de résultats
- Extraire les motifs

APPENDICE B

INTERFACE DE DÉTECTION DES PUPILLES ET DU VISAGE

EyeLab
Mode tutorat

Apprenant : Maxime Landry

Interface de détection des pupilles et du visage

Consigne : Regarder à l'intérieur de la zone grise. EyeLab va détecter vos pupilles par la suite les contours de chaque élément de votre visage

Bonne calibration : 90%



Powered By: EyeLab

APPENDICE C

INTERFACE DE DÉTECTION ET CALIBRATION DE MOUVEMENTS OCULAIRES

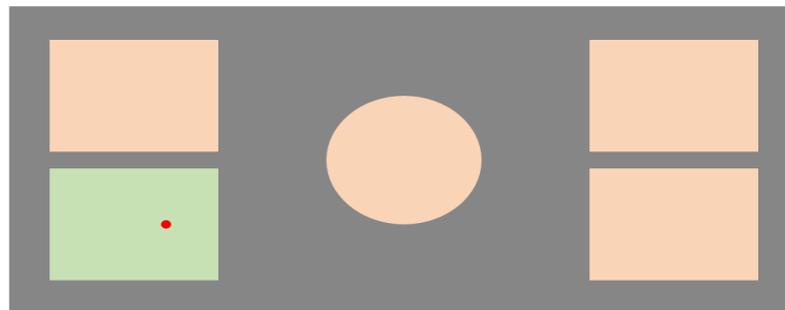
EyeLab
Mode tutorat

Apprenant : Maxime Landry

Interface de détection et calibration de mouvements oculaires

Consigne : Regarder à l'intérieur de chacune des formes ci-dessous. EyeLab vous confirme s'il détecte bien vos mouvements oculaires

Bonne calibration : 90%



Powered By: EyeLab

APPENDICE D

INTERFACE DE EYELAB EN MODE TUTORAT

EyeLab
Mode tutorat

Apprenant : Maxime Landry

Cas d'étude
Can not Intubate / Can not Oxygenate

Consigne
Revisualiser la vidéo de votre performance en faisant attention aux indices de l'agent tuteur

08:35 - Bradycardie

Cette situation fait suite à l'insertion du dispositif supra glottique et pourrait provoquer une diminution de la saturation
Indices : Surveiller les changements de valeur dans la **ZONE 4**, par la suite **ZONE 6** suivi de votre prise en charge **ZONE 3**



The screenshot displays a video feed of a medical procedure with overlaid zones and vital signs. The zones are labeled as follows:

- ZONE 1: Top left corner of the video feed.
- ZONE 2: Top right corner of the video feed.
- ZONE 3: Bottom left corner of the video feed.
- ZONE 4: Center of the video feed, containing a large green '97' representing SpO2.
- ZONE 5: Top right of the vital signs panel, containing 'SpO2 100/90'.
- ZONE 6: Bottom left of the vital signs panel, containing 'NBP Auto 1 min Sys. 139/80 90 ZONE 6 (99)'.
- ZONE 7: Bottom right of the vital signs panel, containing 'awRR 16'.

Powered By: EyeLab

APPENDICE E

INTERFACE DE COMPARAISON DES SÉQUENCES DE MOUVEMENTS OCULAIRES

Interface de comparaison des séquences de mouvements oculaires

- [Apprenant Vs Modèle de l'expert]
- Apprenant Vs Apprenant

| | Modèle de l'expert | |
|---------|--------------------|--------------|
| Novice1 | 0,04 | Voir Détails |
| Novice2 | 0,17 | Voir Détails |
| Novice3 | 0,60 | Voir Détails |
| Novice4 | 0,52 | Voir Détails |
| Novice5 | 0,19 | Voir Détails |
| Novice6 | 0,21 | Voir Détails |
| Novice7 | 0,82 | Voir Détails |
| Novice8 | 0,09 | Voir Détails |
| Novice9 | 0,24 | Voir Détails |

RÉFÉRENCES

BIBLIOGRAPHIE

Agrawal, R. and R. Srikant (1994). Fast Algorithms for Mining Association Rules in Large Databases. Proceedings of the 20th International Conference on Very Large Data Bases, Morgan Kaufmann Publishers Inc.: 487-499.

Aguirre, A., et al. (2014). "A Novel Approach to Diagnosing Motor Skills." Learning Technologies, IEEE Transactions on **7**(4): 304-318.

Allen Newell, S. H. (1972). Human Problem Solving, Prentice-Hall.

Anderson, J., et al. (1995). "Cognitive Tutors: Lessons Learned." Journal of the Learning Sciences **4**: 167-207.

Anderson, J. R. (1987). "Skill acquisition: Compilation of weak-method problem situations." Psychological Review **94**(2): 192-210.

Anderson, J. R. (1993). Rules of the Mind, Hillsdale, NJ: Erlbaum.

Arocha, J. F., et al. (1993). "Hypothesis Generation and the Coordination of Theory and Evidence in Novice Diagnostic Reasoning." Medical decision making : an international journal of the Society for Medical Decision Making **13**: 198-211.

Arocha, J. F., et al. (2005). "Identifying reasoning strategies in medical decision making: A methodological guide." Journal of Biomedical Informatics **38**(2): 154-171.

Balacheff, N. (2013). cK ϕ , a model to reason on learners' conceptions.

Baralis, E., et al. (2013). "P-Mine: Parallel itemset mining on large datasets." 2013 IEEE 29th International Conference on Data Engineering Workshops (ICDEW): 266-271.

Barrier, G. (2008). La communication non verbale : comprendre les gestes : perception et signification. France, Issy-les-Moulineaux.

Baum, L. E. and T. Petrie (1966). "Statistical inference for probabilistic functions of finite state Markov chains " Annals of Mathematical Statistics **37**: 1554-1563.

Beatty, J. (1982). "Task-evoked pupillary responses, processing load, and the structure of processing resources." Psychological Bulletin **2**: 276-292.

Bengio, Y., et al. (2009). Curriculum learning. Proceedings of the 26th Annual International Conference on Machine Learning. Montreal, Quebec, Canada, ACM: 41-48.

Bengio, Y., et al. (2013). On the difficulty of training recurrent neural networks. Proceedings of the 30th International Conference on International Conference on Machine Learning - Volume 28. Atlanta, GA, USA, JMLR.org: III-1310-III-1318.

Benner, P. (1982). "From Novice to Expert." The American Journal of Nursing **82**: 402-407.

Bradbury, J., et al. (2016). "Quasi-Recurrent Neural Networks." CoRR abs/1611.01576.

Brousseau, G. (1998). Théorie des situations didactiques. Grenoble.

Bull, S. and J. Kay (2010). Open Learner Models. **308**: 301-322.

Burton, R. R. and J. S. Brown (1986). Toward a natural-language capability for computer-assisted instruction. Readings in natural language processing, Morgan Kaufmann Publishers Inc.: 605-625.

Carr, B. and I. Goldstein (2004). "Overlays: A Theory of Modelling for Computer Aided Instruction."

Cen, H., et al. (2006). Learning Factors Analysis – A General Method for Cognitive Model Evaluation and Improvement, Berlin, Heidelberg, Springer Berlin Heidelberg.

Chase, W. G. and H. A. Simon (1973). "Perception in chess." Cognitive Psychology **4**: 55-81.

Chee, C.-H., et al. (2019). "Algorithms for frequent itemset mining: a literature review." Artificial Intelligence Review **52**(4): 2603-2621.

Chieu, V. M., et al. (2010). "Student Modeling in Orthopedic Surgery Training: Exploiting Symbiosis between Temporal Bayesian Networks and Fine-grained Didactic Analysis." Int. J. Artif. Intell. Ed. **20**(3): 269-301.

CNRTL (2012). "Perception visuelle." 2015, from <http://www.cnrtl.fr/definition/perception>.

Cohen, M. X., et al. (2008). "Oscillatory Activity and Phase–Amplitude Coupling in the Human Medial Frontal Cortex during Decision Making." Journal of Cognitive Neuroscience **21**: 390-402.

Conati, C., et al. (2013). "Understanding Attention to Adaptive Hints in Educational Games: An Eye-Tracking Study." International Journal of Artificial Intelligence in Education **23**.

Conati, C. and C. Merten (2007). "Eye-tracking for user modeling in exploratory learning environments: An empirical evaluation." Know.-Based Syst. **20**(6): 557-574.

Cooper, L., et al. (2009). Radiology image perception and observer performance: how does expertise and clinical information alter interpretation? Stroke detection explored through eye-tracking. SPIE, Sahiner, B. and Manning, D.J. (eds.). Medical Imaging 2009: Image Perception, Observer Performance, and Technology Assessment.

D'mello, S. and A. Graesser (2007). Mind and Body: Dialogue and Posture for Affect Detection in Learning Environments. Proceedings of the 2007 conference on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That Work, IOS Press: 161-168.

Dewey, J. (1933). "How we think: A restatement of the relation of reflective thinking to the educative process." Boston: Heath & Company.

Endsley, M. R. (2015). "Situation Awareness Misconceptions and Misunderstandings." Journal of Cognitive Engineering and Decision Making **9**(1): 4-32.

Ericsson, A., et al. (2006). The Cambridge Handbook of Expertise and Expert Performance, Cambridge University Press.

Evens, M. W., et al. (1997). CIRCSIM-Tutor: an intelligent tutoring system using natural language dialogue. Proceedings of the fifth conference on Applied natural language processing: Descriptions of system demonstrations and videos. Washington, DC, Association for Computational Linguistics: 13-14.

Feddaoui, I., et al. (2016). EXTRACT: new extraction algorithm of association rules from frequent itemsets. Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. Davis, California, IEEE Press: 752-756.

Feltovich, P. J., et al. (2006). Studies of expertise from psychological perspectives. The Cambridge handbook of expertise and expert performance. C. Ericsson KA, P. J., Feltovich, R. R and Hoffman. New York, Cambridge University Press.

Flavell, J. (1976). "Metacognitive aspects of problem-solving." The nature of intelligence: 231-236.

Fournier-Viger, P., et al. (2012). "CMRules: Mining sequential rules common to several sequences." Know.-Based Syst. **25**(1): 63-76.

Fournier-Viger, P., et al. (2013). "A Multiparadigm Intelligent Tutoring System for Robotic Arm Training." Learning Technologies, IEEE Transactions on **6**(4): 364-377.

Fournier-Vigier, P. (2010). Un modèle hybride pour le support à l'apprentissage dans les domaines procéduraux et mal définis. Informatique Cognitive, Université du Québec à Montréal. **PhD**: 229.

Francisco Varela, E. T., Eléonor Rosch (1993). L'inscription corporelle de l'Esprit, Sciences cognitives et expérience humaine, Seuil.

Franklin, A., et al. (2011). "Opportunistic decision making and complexity in emergency care." Journal of Biomedical Informatics **44**: 469-476.

Frasson, C. and P. Chalfoun (2010). Managing Learner's Affective States in Intelligent Tutoring Systems Advances in Intelligent Tutoring Systems. Studies in Computational Intelligence. **308**: 339-358.

Gombert, J.-E. (1990). Le développement métalinguistique. Paris, Presses Universitaires de France.

Gong, Z. and H. Chen (2018). "Sequential data classification by dynamic state warping." Knowl. Inf. Syst. **57**(3): 545-570.

Graves, A., et al. (2007). Unconstrained online handwriting recognition with recurrent neural networks. Proceedings of the 20th International Conference on Neural Information Processing Systems. Vancouver, British Columbia, Canada, Curran Associates Inc.: 577-584.

Graves, A., et al. (2013). Speech Recognition with Deep Recurrent Neural Networks.

Gruppen, L. and A. Frohna (2000). "Clinical Reasoning." International handbook of research in medical education. Kluwer Academic: 205-230.

Hagedorn, R. (1996). "Clinical Decision Making in Familiar Cases: A Model of the Process and Implications for Practice." British Journal of Occupational Therapy **59**(5): 217-222.

Han, J., et al. (2000). Mining frequent patterns without candidate generation. Proceedings of the 2000 ACM SIGMOD international conference on Management of data. Dallas, Texas, USA, ACM: 1-12.

Hatano, G. and K. Inagaki (1986). Two courses of expertise. Child Development and Education in Japan. H. Stevenson, H. Azuma and K. Hakuta. New York, Freeman.

Hermens, F., et al. (2013). "Eye movements in surgery : A literature review." Journal of Eye Movement Research **6**(4): 1-11.

Hoang Duc, A., et al. (2008). "Eye movements as a probe of attention." Progress in brain research **171**: 403-411.

Homof, A., et al. (2004). EyeDraw: a system for drawing pictures with the eyes. CHI '04 extended abstracts on Human factors in computing systems (CHI EA '04). New York, NY: 1251-1254.

Hubel, D. H. and T. N. Wiesel (1959). "Receptive fields of single neurones in the cat's striate cortex." The Journal of physiology **148**(3): 574-591.

Hubert Dreyfus, S. D. (1986). Mind over Machine. New York, McMillan.

Hutchinson, T. E., et al. (1989). "Human-computer interaction using eye-gaze input." Systems, Man and Cybernetics, IEEE Transactions on **19**(6): 1527-1534.

Joseph, G. M. and V. Patel (1990). "Domain Knowledge and Hypothesis Generation in Diagnostic Reasoning." Medical Decision Making - MED DECIS MAKING **10**: 31-44.

Jraidi, I., et al. (2019). Assessing Students' Clinical Reasoning Using Gaze and EEG Features, Cham, Springer International Publishing.

Just, M. A. and P. A. Carpenter (1980). "A theory of reading: from eye fixations to comprehension." Psychol Rev **87**(4): 329-354.

Keogh, E. J. and M. J. Pazzani (2000). Scaling up dynamic time warping for datamining applications. Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining. Boston, Massachusetts, USA, ACM: 285-289.

Khedher, A., et al. (2018). "Local Sequence Alignment for Scan Path Similarity Assessment." International Journal of Information and Education Technology **8**: 482-490.

Kiesewetter, J., et al. (2016). "Knowledge is not enough to solve the problems – The role of diagnostic knowledge in clinical reasoning activities." BMC Medical Education **16**(1): 303.

Kingma, D. P. and J. Ba (2014). "Adam: A Method for Stochastic Optimization." CoRR **abs/1412.6980**.

Klawonn, F. and F. Höppner (2006). "Equi-sized, Homogeneous Partitioning." Knowledge-Based Intelligent Information and Engineering Systems **4252**: 70-77.

Koedinger, K. R., et al. (2006). The cognitive tutor authoring tools (CTAT): preliminary evaluation of efficiency gains. Proceedings of the 8th international conference on Intelligent Tutoring Systems. Jhongli, Taiwan, Springer-Verlag: 61-70.

Krauzlis, R., et al. (2017). "Neuronal control of fixation and fixational eye movements." Philosophical Transactions of the Royal Society B: Biological Sciences **372**: 20160205.

Kumar, K. M. V. M., et al. (2012). "Sequential Pattern Mining With Multiple Minimum Supports by MS-SPADE." International Journal of Computer Science Issues **9**(5): 285-292.

L. Orquin, J. and S. Mueller Loose (2013). "Attention and choice: A review on eye movements in decision making." Acta Psychologica **144**: 190-206.

Lallé, S. and V. Luengo (2011). Intégration de données haptiques brutes dans des systèmes experts de diagnostic des connaissances. Con-férence Internationale Francophone sur l'Extraction et la Gestion des Connaissances: 599-610.

Le, N.-T. and W. Menzel (2008). "Using Constraint-Based Modelling to Describe the Solution Space of Ill-Defined Problems for Logic Programming." Lecture Notes in Computer Science **4823**: 367-379.

Leiva, L. A. and E. Vidal (2013). "Warped K-Means: An algorithm to cluster sequentially-distributed data." Information Sciences **237**: 196-210.

Li, N., et al. (2011). A Machine Learning Approach for Automatic Student Model Discovery.

Litchfield, D., et al. (2010). "Viewing another person's eye movements improves identification of pulmonary nodules in chest x-ray inspection." Journal of experimental psychology. Applied **16**: 251-262.

Liu, B., et al. (1999). Mining association rules with multiple minimum supports. Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining. San Diego, California, USA, ACM: 337-341.

Lo, D., et al. (2009). "Non-redundant sequential rules-Theory and algorithm." Inf. Syst. **34**(4-5): 438-453.

Lynch, C., et al. (2006). "Defining "Ill-Defined Domains"; A literature survey."

Marcus, S. and J. McDermott (1989). "SALT: a knowledge acquisition language for propose-and-revise systems." Artif. Intell. **39**(1): 1-37.

Mathews, M., et al. (2012). Do Your Eyes Give It Away? Using Eye Tracking Data to Understand Students' Attitudes towards Open Student Model Representations.

May, S., et al. (2010). "Limited clinical reasoning skills used by novice physiotherapists when involved in the assessment and management of patients with shoulder problems: a qualitative study." The Journal of Manual & Manipulative Therapy **18**(2): 84-88.

Mitrovic, A., et al. (2003). A comparative analysis of cognitive tutoring and constraint-based modeling. Proceedings of the 9th international conference on User modeling. Johnstown, PA, USA, Springer-Verlag: 313-322.

Mitrovic, A. and A. Weerasinghe (2009). Revisiting Ill-Definedness and the Consequences for ITSs. **200**: 375-382.

Mizoguchi, R. and J. Bourdeau (2015). "Using Ontological Engineering to Overcome AIED Problems: Contribution, Impact and

Perspectives." International Journal of Artificial Intelligence in Education.

Monteiro, S. and G. Norman (2013). "Diagnostic Reasoning: Where We've Been, Where We're Going." Teaching and learning in medicine: 26-32.

Moritz, S. and G. Blank (2008). Generating and Evaluating Object-Oriented Designs for Instructors and Novice Students. ITS for Ill-Defined Domains Workshop, ITS 2008. Montreal, Canada.

Myers, E. W. and W. Miller (1988). "Optimal alignments in linear space." Bioinformatics 4(1): 11-17.

Nathan, M. J., et al. (2001). Expert Blind Spot: When Content Knowledge Eclipses Pedagogical Content Knowledge. Proceeding of the Third International Conference on Cognitive Science: 644-648.

Neistadt, M. E. (1996). "Teaching strategies for the development of clinical reasoning." American Journal of Occupational Therapy 50(8): 676-684.

Nendaz, M., et al. (2005). "Le raisonnement clinique: données issues de la recherche et implications pour l'enseignement." Revue internationale francophone d'éducation médicale 6: 235-254.

Nielsen, J. and K. Pernice (2009). Eyetracking Web Usability. Indianapolis, IN, New Riders Pub.

Noël, B. (1997). La métacognition. Paris, Bruxelles, De Boeck Univesité.

Nwana, H. (1990). "Intelligent tutoring systems: an overview." Artificial Intelligence Review 4(4): 251-277.

Ogan, A., et al. (2006). The challenges in adapting traditional techniques for modeling student behavior in ill-defined domains Intelligent Tutoring Systems for Ill-Defined Domains Workshop.

Ohlsson, S. (1996). "Learning from Performance Errors." Psychological Review **3**(2): 241-262.

Ordóñez, F. and D. Roggen (2016). "Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition." Sensors **16**: 115.

Patel, V., et al. (1994). Diagnostic Reasoning and Medical Expertise. **31**: 187-252.

Pei, J., et al. (2004). "Mining Sequential Patterns by Pattern-Growth: The PrefixSpan Approach." IEEE Trans. on Knowl. and Data Eng. **16**(11): 1424-1440.

Pyun, G., et al. (2014). "Efficient frequent pattern mining based on Linear Prefix tree." Know.-Based Syst. **55**: 125-139.

Rabiner, L. R. and B. H. Juang (1986). "An introduction to hidden Markov models." IEEE ASSP Magazine **3**: 4-16.

Rayner, K. (1998). "Eye movements in reading and information processing: 20 years of research." Psychol Bull **124**(3): 372-422.

Remolina, E., et al. (2004). Intelligent Simulation-Based Tutor for Flight Training. Interservice/Industry Training, Simulation, and Education: 1-13.

Riccucci, S., et al. (2007). Knowledge acquisition in Intelligent Tutoring System: a data mining approach. 6th Mexican Conference on Artificial Intelligence.

Romero, C. and S. Ventura (2010). "Educational data mining: a review of the state of the art." Trans. Sys. Man Cyber Part C **40**(6): 601-618.

Sandeep, M., et al. (1995). An intelligent tutoring system for simulator-based helicopter flight training. Flight Simulation Technologies Conference, American Institute of Aeronautics and Astronautics.

Smith, T. F. and M. S. Waterman (1981). "Identification of common molecular subsequences." Journal of molecular biology **147** 1: 195-197.

Srivastava, N. (2013). Improving neural networks with dropout. Computer Science, University of Toronto. **PhD**.

Szathmary, L., et al. (2007). Towards Rare Itemset Mining. Tools with Artificial Intelligence, 2007. ICTAI 2007. 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI '07). **1**: 305-312.

Tchétagni, J., et al. (2006). "A Framework to Specify a Cognitive Diagnosis Component in ILEs." Jl. of Interactive Learning Research: 269-293.

Toussaint, B.-m. (2015). Apprentissage automatique à partir de traces multi-sources hétérogènes pour la modélisation de connaissances perceptivo-gestuelles, Université de Grenoble. **PhD**.

Vapnik, V. (1995). The Nature of Statistical Learning Theory. New York, Springer.

Vergnaud, G. (1990). La théorie des champs conceptuels. Recherches en Didactique des Mathématiques: 133–170.

Vleuten, V. D., et al. (2008). Assessing clinical reasoning. Clinical Reasoning in the Health Professions. S. L. Joy Higgs, Nicole Christensen, Elsevier Health Sciences: 413-419.

Webb, G. I. (1995). "OPUS: an efficient admissible algorithm for unordered search." J. Artif. Int. Res. **3**(1): 431-465.

Wenger, É. (1987). Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge. Los Altos, USA, Morgan Kaufmann.

Wilson, E. B. (1927). "Probable Inference, the Law of Succession, and Statistical Inference." Journal of the American Statistical Association **22**(158): 209-212.

Xing, Z., et al. (2010). "A brief survey on sequence classification." SIGKDD Explor. Newsl. **12**(1): 40-48.

Zaki, M. J. (2000). "Scalable Algorithms for Association Mining." IEEE Trans. on Knowl. and Data Eng. **12**(3): 372-390.

Zouaq, A. and R. Nkambou (2008). "Building Domain Ontologies from Text for Educational Purposes." Learning Technologies, IEEE Transactions on **1**(1): 49-62.