

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LA PRÉDICTION DU CHANGEMENT DIRECTIONNEL DE LA CROISSANCE DU PIB DU QUÉBEC À PLUSIEURS  
HORIZONS

MÉMOIRE  
PRÉSENTÉ  
COMME EXIGENCE PARTIELLE  
DE LA MAÎTRISE EN ÉCONOMIQUE

PAR  
AYI DEMALANVILLE BLAISE FOLLY

AVRIL 2025

UNIVERSITÉ DU QUÉBEC À MONTRÉAL  
Service des bibliothèques

Avertissement

La diffusion de ce mémoire 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.12-2023). 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

La réalisation de ce mémoire a été rendue possible grâce au concours de plusieurs personnes à qui je voudrais témoigner toute ma reconnaissance.

Tout d'abord, je voudrais exprimer ma gratitude et ma reconnaissance au Professeur Dalibor STEVANOVIĆ, pour la patience dont il a fait preuve en acceptant de diriger ce travail jusqu'à la fin. Je ne saurai jamais lui rendre toute la gratitude que je lui dois. Son enseignement et son expertise m'ont amené à découvrir le domaine intéressant de l'apprentissage automatique, et la réalisation de ce mémoire m'a permis d'entrevoir une nouvelle direction pour ma future carrière.

J'adresse aussi mes sincères remerciements à toute l'équipe pédagogique et administrative du département des sciences économiques de l'UQAM. Je souhaite particulièrement remercier Madame Martine Boisselle-Lessard pour son soutien permanent dans la réussite de ma formation.

Aux gouvernements de la République du Bénin et du Canada pour avoir financé mes études grâce au Programme Canadien de Bourses de la Francophonie (PCBF).

À ma femme Marie, mes enfants O'nel, Abygaëlle et Anaël, ma mère et ma sœur Estelle pour leur confiance et leur soutien indéfectibles malgré la distance. Trouvez dans ce travail le fruit de vos efforts, sacrifices et sollicitudes!

À mes ami(e)s à Montréal qui ont su rendre mon séjour agréable. Vous vous reconnaitrez sûrement dans ces phrases.

Merci à toutes et tous!!!

## TABLE DES MATIÈRES

LISTE DES TABLEAUX .....	vi
LISTE DES FIGURES .....	vii
RÉSUMÉ .....	viii
INTRODUCTION .....	1
CHAPITRE 1 REVUE DE LITTÉRATURE .....	8
1.1 Travaux de Merton (1981) et Henriksson et Merton (1981) .....	8
1.2 Travaux de Schnader et Stekler (1990) .....	9
1.3 Test de signe prédictif de Pesaran et Timmermann (1992) .....	10
1.4 Changement directionnel et apprentissage automatique en macroéconomie.....	11
CHAPITRE 2 APPROCHE MÉTHODOLOGIQUE .....	14
2.1 Définition du changement directionnel de la croissance du PIB .....	14
2.2 Cadre d'analyse en apprentissage supervisé de la prévision du changement directionnel .....	15
2.3 Ensemble d'informations et stratégie de prévision .....	15
2.4 Fonction de perte, erreur de prévision et critères de performance .....	16
2.4.1 Fonction de perte et erreur de prévision .....	16
2.4.2 Critères de performance du changement directionnel .....	17
2.4.3 Courbe ROC ( <i>Receiver Operator Characteristic</i> ) et critères associés .....	20
2.4.4 Choix du seuil de changement directionnel .....	21
2.5 Spécification du modèle à facteurs et des méthodes de classifications supervisée.....	23
2.5.1 Modèle à facteurs .....	23
2.5.2 Classifieur aléatoire ( <i>DummyClassifier</i> ou DC) .....	26
2.5.3 Régression logistique ( <i>Logistic Regression</i> ou LR).....	27

2.5.4	Classifieur bayésien naïf ( <i>Naive Bayes</i> ou NB) .....	27
2.5.5	Méthode des K-plus proches voisins ( <i>K-nearest neighbors</i> ou KNN) .....	28
2.5.6	Analyse linéaire discriminante ( <i>Linear Discriminant Analysis</i> ou LDA) .....	29
2.5.7	Machines à vecteurs de support ( <i>Support Vector Machines</i> ou SVM) .....	30
2.5.8	Arbres de décision ( <i>Decision Tree</i> ou DT).....	33
2.5.9	Forêt aléatoire ( <i>Random forest</i> ou RF) .....	35
2.6	Recherche des hyperparamètres optimaux et exercice de prévision .....	37
2.6.1	Recherche des hyperparamètres optimaux .....	37
2.6.2	Exercice de prévision du changement directionnel .....	39
CHAPITRE 3	DONNÉES .....	41
3.1	Présentation de la base de données .....	41
3.2	Variables d'intérêt et cible .....	42
3.3	Extraction des facteurs communs comme variables explicatives .....	45
3.3.1	Analyse des diagrammes des valeurs propres et de la variance expliquée .....	45
3.3.2	Détermination du nombre de facteurs communs.....	46
3.3.3	Facteurs communs extraits en tant que variables explicatives .....	47
CHAPITRE 4	RÉSULTATS EMPIRIQUES, DISCUSSIONS ET LIMITES MÉTHODOLOGIQUES .....	49
4.1	Résultats empiriques .....	49
4.2	Discussions .....	55
4.3	Limites méthodologiques et pistes d'amélioration .....	56
4.3.1	Limites méthodologiques .....	56
4.3.2	Pistes d'amélioration .....	58
CONCLUSION	.....	61

ANNEXE A DONNÉES ..... 63

BIBLIOGRAPHIE ..... 80

## LISTE DES TABLEAUX

Tableau 1.1	Exemple de tableau de contingence 2x2 pour l'évaluation des prévisions directionnelles.	10
Tableau 2.1	Exemple de matrice de confusion pour la prévision de la direction du PIB.....	18
Tableau 3.1	Quelques valeurs illustratives de la croissance du PIB québécois et de sa direction.....	44
Tableau 3.2	Estimation du nombre de facteurs statiques et dynamiques .....	47
Tableau 4.1	Critères de performance des huit méthodes de classification supervisée selon l'horizon	50

## LISTE DES FIGURES

Figure 2.1	Courbes ROC et qualité des prédictions .....	21
Figure 2.2	Validation croisée des séries temporelles .....	39
Figure 3.1	Évolution du niveau du PIB réel trimestriel et de son taux de croissance trimestriel au Québec du 1 <sup>er</sup> trimestre de 1981 au 1 <sup>er</sup> trimestre de 2020 .....	42
Figure 3.2	Distribution des classes du changement directionnel de la croissance du PIB au québécois	44
Figure 3.3	Scree plot et trace .....	46
Figure 3.4	Évolution des facteurs communs estimés et des séries principales à travers le temps.....	48

## RÉSUMÉ

Les acteurs privés et publics doivent comprendre l'orientation future de la croissance du PIB réel afin d'être en mesure de l'analyser et de prendre des décisions éclairées. Dans ce mémoire, nous étudions les performances hors-échantillon de huit méthodes de classification supervisée pour la prédiction du changement directionnel de la croissance du PIB réel au Québec effectuée à  $h \in \{1, 2, 4, 6, 8\}$  trimestres à l'avance. L'ensemble de données est constitué des séries macroéconomiques trimestrielles sur le Canada et ses provinces de 1981Q1 à 2020Q1. Ces séries sont issues du grand ensemble de données construit par Fortin-Gagnon *et al.* (2022), dénommé LCDMA. Compte tenu de la taille de ce jeu de données, nous avons ensuite utilisé le modèle à facteurs pour extraire des facteurs communs comme variables explicatives. Dans le cadre de la modélisation prédictive, les hyperparamètres des différentes méthodes sont optimisés à l'aide d'un algorithme génétique (GASearchCV). Par ailleurs, les prévisions sont générées à l'aide d'une approche à fenêtre glissante. Les différentes méthodes sont par la suite évaluées à l'aide de critères de performance découlant de la matrice de confusion. Selon les résultats empiriques, les performances des différentes méthodes de classification supervisée sont fortement liées à l'horizon de prévision. Aucune méthode ne s'impose de façon robuste sur tous les horizons. Ce qui souligne l'importance de l'adaptabilité desdites méthodes selon l'horizon de prévision. De plus, les faibles valeurs de l'aire sous la courbe ROC révèlent l'incapacité de certaines méthodes à discriminer les périodes d'accélération et de décélération de la croissance du PIB.

Mots clés : Méthodes de classification supervisée, changement directionnel, modèle à facteurs, données massives, algorithme génétique, fenêtre glissante, prévision hors échantillon

## INTRODUCTION

La croissance du produit intérieur brut (PIB) québécois va-t-elle s'accélérer, rester constante ou ralentir au cours des prochains trimestres ? Pour répondre à cette question, les prévisionnistes s'intéressent à comparer la croissance projetée à l'horizon  $h$  par rapport à celle réalisée à la période  $t$ , ou encore à examiner si la croissance projetée à  $h$  a effectivement évolué dans la même direction que sa valeur à  $t$ . Dans ces cas, des mesures fournissant des estimations quantitatives de la qualité attendue des prévisions sont nécessaires, ainsi que des tests de précision directionnelle qui ne portent pas sur la valeur de la prévision, mais plutôt sur la direction future de la variable projetée. Or, il ne fait aucun doute que la précision directionnelle<sup>1</sup> est une mesure importante de la performance, surtout lorsqu'on sait qu'il est difficile d'obtenir de bonnes performances en termes d'erreur quadratique moyenne (EQM) (Goulet Coulombe *et al.*, 2020). Par exemple, un modèle peut afficher une racine carrée de l'erreur quadratique moyenne plus faible qu'un autre, mais ne pas aussi bien détecter si la croissance du PIB accélère ou décélère d'un trimestre à l'autre (Sédillot et Pain, 2005). Ce qui compte finalement, ce n'est pas la valeur de la prévision, mais les décisions prises en fonction de celle-ci et des bénéfices qui en découlent (Lardic et Mpacko Priso, 1999). À titre d'exemple, un investisseur achète une action si le prix doit augmenter, ou une banque centrale relève le taux d'intérêt si les anticipations d'inflation doivent augmenter (Blaskowitz et Herwartz, 2011).

Jusqu'à présent, la littérature sur la précision directionnelle s'est principalement concentrée sur le test dit de «*market timing*» qui remonte aux travaux pionniers de Merton (1981) et de Henriksson et Merton (1981) (Meyer, 2019). Le premier auteur a proposé un cadre lié aux décisions des investisseurs pour prévoir les mouvements (hausse ou baisse) du marché, alors que les seconds en ont déduit un test statistique. L'utilisation du test de market timing dans la littérature financière a été largement étudiée (Pierdzioch et Rülke, 2015). En macroéconomie, la précision directionnelle est plutôt utilisée comme un critère pour juger de la valeur économique des prévisions macroéconomiques (Cornec et Mikol, 2011). Dans ce cas, les prévisions sont considérées comme ayant de la valeur si elles prédisent les changements directionnels, c'est-à-dire une augmentation ou une diminution, une accélération ou une décélération, mieux qu'un modèle naïf. Schnader et Stekler (1990) furent les premiers à adapter cette méthode en macroéconomie, grâce à un tableau de contingence utilisé pour tester la valeur des prévisions directionnelles du produit national brut (PNB) réel des États-Unis. Ce tableau ou sa variante a reçu une grande attention dans la littérature empirique consacrée aux changements directionnels de diverses variables macroéconomiques (voir par exemple, Öl-

---

1. On parle également de changement directionnel ou d'exactitude directionnelle

ler et Barot (2000); Pons (2001); Greer (2003); Ashiya (2006); Sinclair *et al.* (2010); Tsuchiya (2012); entre autres).

L'objectif de ce mémoire est de prédire la direction de l'évolution ou le changement directionnel du taux de croissance trimestriel du PIB québécois à plusieurs horizons  $h = 1, \dots, H$ . Il ne s'agit pas de déterminer si les futures variations du taux de croissance du PIB seront de 3%, 5% ou -1%, mais nous nous intéressons aux changements de signe, aux changements directionnels ou aux variations des prévisions macroéconomiques. Par exemple, quel sera le changement directionnel (hausse, statu quo ou baisse) du taux de croissance du PIB québécois entre le troisième et le quatrième trimestre? Kotchoni *et al.* (2019a) ont abordé cette question en étudiant la capacité à prédire la direction de plusieurs variables macroéconomiques grâce au test de signe de Pesaran et Timmermann (1992). Dans notre cas, nous considérons les prévisions du changement directionnel de la croissance du PIB québécois comme un problème de classification binaire, où le changement directionnel, encore inconnu, est classé entre deux groupes «*accélération du PIB*» et «*décélération du PIB*». Une fois cette classification effectuée, nous pouvons attribuer a priori un groupe à tout changement directionnel à venir, en s'appuyant notamment sur un ensemble d'indicateurs économiques. Ce faisant, le changement directionnel de la croissance du PIB québécois peut être estimé à l'aide de méthodes de classification supervisée.

Le domaine d'étude est pertinent pour plusieurs raisons. Premièrement, la croissance du PIB reflète la santé générale d'une économie (Liu et Gao, 2022) et la prévision de ses directions futures est une question pertinente en macroéconomie. En effet, cette direction est souvent décrite comme une succession de périodes d'accélération et de décélération. Selon Kotchoni *et al.* (2019b), «*cette dynamique d'alternance revêt une importance particulière pour la prise de décisions d'acteurs privés comme publics, entre autres parce que les fluctuations associées [· · ·] peuvent être liées à des variations substantielles du taux de chômage*». Par exemple, une croissance du PIB réel qui s'accélère, signale que l'économie fonctionne bien. Ce qui signifie que les entreprises se développent, créent davantage d'emplois et qu'il y a plus de pouvoirs d'achat. À l'inverse, lorsque la croissance ralentit, l'économie engendre moins d'emplois, moins d'embauches et les gens ont moins d'argent à dépenser. Dans ce contexte, il est essentiel d'anticiper avec précision et en temps voulu les changements directionnels de la croissance du PIB. Selon Cornec et Mikol (2011), cela n'a jamais fait l'objet d'une analyse directe, c'est-à-dire la construction d'une prévision directe sur l'«*accélération*» ou la «*décélération*» du PIB. En outre, le Québec étant l'une des provinces les plus dynamiques du Canada, une bonne anticipation des directions de l'évolution de son taux de croissance du PIB constitue un grand

atout dans la prise de décision.

Deuxièmement, la classification supervisée a rarement été mobilisée pour modéliser et prévoir le changement directionnel en macroéconomie. À notre connaissance, seuls Cornec et Mikol (2011) et Van Ruth (2014) ont exploré la question en étudiant la prédiction du changement directionnel de divers indicateurs macroéconomiques à l'aide de méthodes de classification supervisée<sup>2</sup>. Pour le Québec, les études sur les prévisions directionnelles à l'aide de la classification supervisée sont inexistantes. Goulet Coulombe *et al.* (2020) ont exploré la question, mais en appliquant le test de signe prédictif de Pesaran et Timmermann (1994) pour estimer la capacité de divers modèles d'apprentissage automatique à prévoir la direction que prendront plusieurs variables macroéconomiques du Québec et du Canada dans  $h$  trimestres.

Troisièmement, depuis quelques années, l'apprentissage automatique avec application en économie connaît un développement important et est de plus en plus mobilisée pour la prévision des grandes variables macroéconomiques (Goulet Coulombe *et al.* (2022)). Dans leur sillage, selon Charpentier *et al.* (2018), « *un ensemble de techniques souvent computationnelles, alternatives à l'économétrie classique* » a été développé et mis en œuvre, avec des résultats divers. Les contributions desdites techniques résident, soit dans les méthodologies employées, soit dans leur application unique et innovante (Gogas et Papadimitriou, 2021). De plus, les gains de prévision obtenus par rapport aux méthodes traditionnelles des moindres carrés ordinaires (MCO) sont bien documentés dans les prévisions macroéconomiques (Goulet Coulombe *et al.*, 2022). L'article de Athey et Imbens (2019) offre une description complète de certaines de ces méthodes considérées comme utiles et importantes pour les économistes.

Quatrièmement, les méthodes d'apprentissage supervisé ont été beaucoup plus orientées historiquement autour des problèmes de classification, et assez peu dans le cas où la variable d'intérêt est continue (Charpentier *et al.*, 2018). La classification sert à prédire une classe  $C$ , qui est une variable catégorielle (par exemple, un chat, un chien, un sujet sain, un groupe sanguin, etc...), en fonction de plusieurs facteurs  $x_i$  souvent mesurés :

$$C = g(x_1, \dots, x_N, \Theta) \quad (1)$$

où  $\Theta$  représente l'ensemble des paramètres de la fonction  $g$ .

---

2. Taux de croissance trimestriel du PIB français (Cornec et Mikol (2011)) et cinq variables économiques mensuelles néerlandaises, à savoir la production manufacturière, les exportations, la consommation des ménages, la confiance des producteurs et celle des consommateurs (Van Ruth (2014))

Il y a trois principaux types de classification : binaire, multiclasse et multiétiquette. Mais, le plus étudié est la classification binaire, où nous avons généralement  $C \in \{0, 1\}$ . Par exemple, on veut déterminer si un courriel arrivant dans notre boîte à lettres est un pourriel (SPAM). Il n'y a que deux éventualités possibles : oui ou non. La classification est un sujet de recherche important, applicable à de nombreux domaines tels que la finance, l'ingénierie, l'économie et la médecine. En macroéconomie, elle a été appliquée par exemple à la prévision des récessions (voir par exemple, Qi (2001); Kauppi et Saikkonen (2008); Gogas *et al.* (2015); Davig et Hall (2019); Nevasalmi (2022); Iqbal et Bowman (2018)), des points de retournement du cycle économique (Piger (2020)), à la prévision immédiate de variables macroéconomiques (Maas (2019); AbdElminaam *et al.* (2023); Soybilgen et Yazgan (2021)). Les méthodes de classification fréquemment utilisées dans ces applications, portent, entre autres, sur le classifieur bayésien naïf, la machine à vecteur de support (SVM), les arbres de décision, les forêts aléatoires, les réseaux de neurones, les K-voisins les plus proches (KNN), ainsi que sur des approches statistiques traditionnelles telles que la régression logistique, le modèle probit et l'analyse discriminante linéaire.

Cinquièmement, les grands ensembles de données sont devenus très populaires dans la recherche macroéconomique empirique. McCracken et Ng (2016) ont conçu un panel de données mensuelles, largement utilisé aux États-Unis, connu sous le nom de FRED-MD. Ils ont également proposé un panel trimestriel, FRED-QD (McCracken et Ng (2020)). Fortin-Gagnon *et al.*, (2022) ont créé un grand ensemble similaire de données dénommé « LCDMA »<sup>3</sup> qui permet de suivre, en temps réel, l'évolution de plus d'une centaine d'indicateurs économiques pour le Canada et ses provinces. Dans le cadre de ce mémoire, nous avons retenu la base de données de mars 2020 qui inclut 477 variables pour 157 périodes temporelles (1981Q1-2020Q1) après traitement. Nous nous trouvons donc confrontés à un nombre d'observations bien inférieur à celui des prédicteurs, et tous ne sont pas nécessairement pertinents au vue du problème étudié. Dans un tel contexte, les méthodes traditionnelles de classification (la régression logistique, le modèle probit, KNN, etc...) se heurtent à un obstacle bien connu qui est la malédiction de la dimensionnalité. Pour lever cet obstacle, les approches récentes en matière de prévision ont mis l'accent, entre autres, sur la nécessité de sélectionner parmi l'ensemble des variables explicatives, celles qui s'avèrent les plus informatives, afin de supprimer le bruit généré par certaines variables et de réduire le coût de calcul de la phase d'apprentissage (Wafo *et al.*, 2022). Cette nécessité a été soulignée dans la littérature récente (Borup *et al.*, 2023). Selon Yousef et Ng (2021), une estimation précise et une prévision sont rendues possibles en s'appuyant sur une certaine forme de réduction de dimension. La littérature économique propose diverses stratégies à cette

---

3. Large Canadian dataset for macroeconomic analysis

fin (Cristelli *et al.*, 2017). Mais, de manière générale, la modélisation factorielle est préférable (Goulet Coulombe *et al.*, (2020)).

Dans cette optique, nous explorons la capacité de différentes méthodes de classification supervisée à prédire le changement directionnel de la croissance du PIB québécois à plusieurs horizons  $h = 1, \dots, H$ . Nous menons une analyse détaillée des performances hors échantillon de ces prédictions, en utilisant un classifieur aléatoire comme référence. À cette fin, la base de données mentionnée précédemment sera utilisée. Compte tenu de la part importante de variation de cet ensemble de données, Fortin-Gagnon *et al.* (2022) recommandent de recourir à une structure par facteurs pour réduire la dimension. Nous adoptons donc cette approche en estimant les facteurs communs grâce à l'analyse en composantes principales. Le nouveau jeu de données, construit à partir de ces facteurs, est partitionné en deux périodes : la première sert à l'entraînement et la validation de nos différents modèles, afin de déterminer les hyperparamètres optimaux via la validation croisée et l'algorithme génétique ; la seconde est exclusivement réservée au test hors échantillon. Ensuite, nous exploitons cette période et les hyperparamètres optimaux obtenus pour entraîner nos modèles pour chaque prédiction, en adoptant un schéma de fenêtre glissante. Ce schéma facilite la prédiction des changements directionnels de la croissance du PIB québécois pour le trimestre d'horizon désigné. C'est une approche conventionnelle couramment utilisée dans la littérature de l'apprentissage automatique (voir par exemple, Goulet Coulombe *et al.*, (2020)).

Les performances de nos modèles sont obtenues à l'aide de critères issus de la matrice de confusion. Dans cet exercice de prévision, huit méthodes de classification supervisée ont été mises en œuvre : le classifieur aléatoire (DC), le classifieur bayésien naïf (NB), les K-plus proches voisins (KNN), les machines à vecteurs de support (SVM), la régression logistique (LR), les forêts aléatoires (RF), les arbres de décision (DT) et l'analyse discriminante linéaire (LDA). Notre travail est le premier à comparer les performances de prévision de ces méthodes de classification supervisée dans un contexte de données massives, appliquées au cas du Québec. Les études antérieures, pour la plupart, se sont concentrées sur un nombre restreint de variables et un seul horizon (voir, par exemple, Cornec et Mikol (2011) et Van Ruth (2014)).

Les résultats empiriques hors échantillon montrent que les performances des différentes méthodes sous-jacentes sont fortement liées à l'horizon de prévision. À court terme ( $h = 1$  et  $h = 2$ ), certaines méthodes comme LDA, SVM et LR obtiennent des performances relativement satisfaisantes, bien que leur pouvoir discriminant mesuré par l'AUROC demeure faible. À mesure que l'horizon s'allonge ( $h = 4, 6, 8$ ), les per-

formances de l'ensemble des méthodes se dégradent : le rappel diminue, l'AUROC tombe souvent sous le seuil de 50 % et aucune méthode ne parvient à maintenir un bon compromis entre détection des signaux d'accélération et réduction des faux positifs. En définitive, aucune méthode ne s'impose de façon robuste sur tous les horizons. Cette situation met en évidence deux enseignements majeurs :

- la nécessité d'adapter la méthode de classification supervisée selon l'horizon de prévision visé : aucune méthode ne domine systématiquement ;
- les faibles valeurs de l'aire sous la courbe ROC révèlent l'incapacité de certaines méthodes à discriminer les périodes d'accélération et de décélération de la croissance du PIB.

Ce mémoire apporte plusieurs contributions spécifiques à la littérature sur les données massives et l'apprentissage automatique avec applications en économie :

- il propose une application de plusieurs méthodes de classification supervisée (DC, LR, NB, KNN, SVM, RF, DT, LDA) à la prévision du changement directionnel de la croissance du PIB québécois à plusieurs horizons  $h \in \{1, 2, 4, 6, 8\}$ , un cadre peu exploré dans les travaux existants ;
- il mobilise un grand jeu de données macroéconomiques (LCDMA), et applique une réduction de dimension basée sur un modèle à facteurs, ce qui permet de traiter du problème de la malédiction de la dimensionnalité.
- il adopte un schéma de prévision hors échantillon avec fenêtre glissante, garantissant le respect de la structure chronologique des données macroéconomiques.
- l'optimisation des hyperparamètres est assurée par l'algorithme évolutif GASearchCV, en remplacement des approches classiques de recherche en grille ou de recherche aléatoire, afin de renforcer l'efficacité de l'apprentissage.

Ces choix méthodologiques s'inscrivent dans une volonté d'ancrer le travail à la fois dans les meilleures pratiques de l'apprentissage automatique et les exigences de la prévision macroéconomique. Le cadre mis en œuvre vise ainsi à évaluer la pertinence et les limites des méthodes de classification supervisée pour anticiper la direction de l'évolution de la croissance du PIB québécois, en intégrant la complexité des données disponibles.

Le mémoire est structuré comme suit : le chapitre 1 est consacré à la revue de la littérature. Il présente les principales contributions théoriques et méthodologiques relatives à la précision directionnelle. Le chapitre 2 expose l'approche méthodologique adoptée, incluant la formalisation du problème de prévision du changement directionnel de la croissance du PIB dans un cadre d'apprentissage supervisé. Il présente également les huit méthodes de classification supervisée mobilisées, ainsi que les choix méthodologiques relatifs à la

recherche des hyperparamètres et à l'exercice de prévision hors échantillon. Le chapitre 3 détaille les données utilisées et les traitements appliqués pour obtenir la variable cible et les facteurs communs servant de variables explicatives. Enfin, le chapitre 4 présente les résultats empiriques, les discute de manière critique, et identifie les principales limites méthodologiques ainsi que les perspectives d'amélioration.

## CHAPITRE 1

### REVUE DE LITTÉRATURE

L'analyse directionnelle a suscité depuis les travaux pionniers de Merton (1981) et ceux de Henriksson et Merton (1981) sur le marché boursier, l'intérêt de plusieurs chercheurs tant dans le domaine financier que macroéconomique. Notre travail s'inscrit dans le cadre de la prédiction du changement directionnel de la croissance du PIB québécois à plusieurs horizons à l'avance, ainsi que dans la mesure de la performance de ces prédictions. Cependant, il se distingue par l'approche adoptée pour prédire le changement directionnel ainsi que par le contexte étudié. À ce titre, il convient de rappeler que les travaux existants sur la thématique du changement directionnel sont largement basés sur le test de précision directionnelle de Henriksson et Merton (ci-après dénommés HM) (1981), le tableau de contingence 2x2 de Schnader et Stekler (1990)<sup>4</sup>, et le test de signe prédictif proposé par Pesaran et Timmermann (1992), que nous présenterons dans la suite. Cela nous permettra d'expliquer les considérations théoriques et méthodologiques qui sous-tendent ces différents tests.

Cette section fait également référence aux résultats de quelques travaux portant sur la prévision du changement directionnel de quelques variables macroéconomiques à l'aide des modèles d'apprentissage automatique.

#### 1.1 Travaux de Merton (1981) et Henriksson et Merton (1981)

Merton (1981) et Henriksson et Merton (1981) ont été les premiers à proposer un cadre de changement directionnel lié aux décisions des investisseurs dans la prévision des mouvements de marché à la hausse et à la baisse. Le premier a établi les conditions à respecter pour qu'une prévision d'investissement soit utile (économiquement) aux investisseurs, tandis que les seconds en ont lié un test statistique.

Dans le modèle de Merton (1981), l'investisseur exprime la direction de ces prévisions sur un marché de deux manières :

- soit, il s'attend à ce que la rentabilité du marché soit supérieur au taux sans risque (marché haussier) ;
- soit, par rapport au taux sans risque (marché en baisse), il se protège contre une performance né-

---

4. C'est une adaptation du test HM

gative du marché.

En désignant par  $\theta_t$  la prédiction de l'investisseur selon laquelle les fonds investis sur le marché  $Z_t$  seront supérieurs ou inférieurs aux actifs sans risque  $R_t$ , l'auteur postule que les probabilités conditionnelles  $p_i$  d'une prévision correcte sont données par :

$$p_1(t) = Prob(\theta_t = 1 | Z_t \leq R_t) \quad (1.1)$$

$$p_2(t) = Prob(\theta_t = 1 | R_t < Z_t) \quad (1.2)$$

avec  $\theta_t = 1$  si la prévision est correcte.

En supposant que  $p_1(t)$  et  $p_2(t)$  ne dépendent pas de l'ampleur de la différence entre  $Z_t$  et  $R_t$ , Merton (1981) montre que  $p_1(t) + p_2(t)$  est une condition nécessaire et suffisante pour l'évaluation de la capacité prédictive.

Henriksson et Merton (1981) ont développé un test statistique non paramétrique pour tester l'hypothèse :

$$H_0 : p_1(t) + p_2(t) = 1 \quad (1.3)$$

que la prévision de market-timing ( $\theta_t = 1$  ou  $\theta_t = 0$ ) n'a pas de valeur économique contre l'hypothèse alternative :

$$H_1 : p_1(t) + p_2(t) > 1 \quad (1.4)$$

Dans la littérature, le test exact de Fisher (test FE) et celui de  $\chi^2$  sont largement utilisés comme test HM. Le test original est équivalent au test FE, et celui de  $\chi^2$  est considéré comme une variante du test HM. Dans les deux tests, l'hypothèse nulle est que le changement directionnel d'une prévision et celui de la réalisation sont indépendants. Un rejet de l'hypothèse nulle implique donc qu'une prévision est un prédicteur utile de la direction réelle de la variable qui nous intéresse.

## 1.2 Travaux de Schnader et Stekler (1990)

Les résultats des travaux de Merton (1981) et Henriksson et Merton (1981) ont stimulé les recherches sur le changement directionnel des prévisions en macroéconomie. La méthode est utilisée pour fournir un critère permettant de juger de la valeur économique des prévisions macroéconomiques. Schnader et Stekler (1990) fut les premiers à l'introduire en macroéconomie au moyen d'un tableau de contingence 2x2 (voir le tableau 1.1) en testant la valeur des prévisions directionnelles du Produit National Brut (PNB) réel des États-Unis.

Tableau 1.1 – Exemple de tableau de contingence 2x2 pour l'évaluation des prévisions directionnelles.

Résultats réels	Résultats prévus		Total
	Positif (+)	Négatif (-)	
Positif (+)	$n_1$	$N_1 - n_1$	$N_1$
Négatif (-)	$N_2 - n_2$	$n_2$	$N_2$
Total	$n$	$N - n$	$N$

Note : La précision directionnelle est évaluée à l'aide d'un tableau de contingence 2x2 dont la structure est illustrée ici.  $N_1$  (respectivement  $N_2$ ) représente le nombre total d'observations dont le résultat réel est *positif* (respectivement *négatif*);  $n_1$  (respectivement  $n_2$ ) correspond au nombre de prévisions correctement anticipées *parmi les observations réellement positives* (respectivement *parmi les observations réellement négatives*);  $n$  est le nombre total de prévisions (positives) formulées, et  $N$  le nombre total d'observations; La cellule  $N_1 - n_1$  (respectivement  $N_2 - n_2$ ) indiquent les prévisions incorrectes lorsque le résultat réel est positif (respectivement le résultat réel est négatif).

Pour évaluer la valeur de la prévision directionnelle du PNB réel, Schnader et Stekler (1990) posent :

- $H_0$  : il n'y a pas d'indépendance entre les résultats prédits et réels de la variation du PNB
- $H_1$  : les résultats prédits et réels de la variation du PNB sont indépendants.

Le test le plus simple pour l'indépendance dans une table de contingence est d'utiliser le test  $\chi^2$ . La statistique de test est alors donnée par :

$$\chi^2 = \sum_i^2 \sum_j^2 \frac{(N_{ij} - N_{pij})^2}{N_{pij}} \quad (1.5)$$

où  $N_{ij}$  et  $N_{pij}$  sont les proportions des résultats réels et prévus dans la  $ij^{th}$  case.

Une valeur de  $\chi^2$  supérieure à la valeur critique indiquera un rejet de l'hypothèse  $H_0$ .

### 1.3 Test de signe prédictif de Pesaran et Timmermann (1992)

Pesaran et Timmermann (ci-après dénommés PT) (1992) présentent un test de signe non paramétrique permettant d'examiner la capacité d'un modèle à prédire le changement directionnel d'une série d'intérêt. En désignant la série d'intérêt par  $y_t$  et sa prévision par  $x_t$ , le test proposé est défini comme suit :

$$S_n = \frac{\hat{P} - \hat{P}_*}{[\hat{V}(\hat{P}) - \hat{V}(\hat{P}_*)]^{1/2}} \quad (1.6)$$

où :

–  $\hat{P} = \frac{1}{n} \sum_{t=1}^n \mathbb{1}(y_t x_t)$  est la proportion de cas positifs correctement prédits avec

$$\mathbb{1}(\cdot) = \begin{cases} 1 & \text{si } \cdot > 0 \\ 0 & \text{sinon} \end{cases}$$

–  $\hat{P}_* = \hat{P}_y \hat{P}_x + (1 - \hat{P}_y)(1 - \hat{P}_x)$  est l'estimation de la probabilité de prédire correctement les événements en supposant que les prévisions et les réalisations sont distribuées de manière indépendante

avec  $\hat{P}_y = \frac{1}{n} \sum_{t=1}^n y_t$  et  $\hat{P}_x = \frac{1}{n} \sum_{t=1}^n x_t$

–  $\hat{V}(\hat{P}) = \frac{1}{n} \hat{P}_*(1 - \hat{P}_*)$

et  $\hat{V}(\hat{P}_*) = \frac{1}{n} (2\hat{P}_y - 1)^2 \hat{P}_x (1 - \hat{P}_x) + \frac{1}{n} (2\hat{P}_x - 1)^2 \hat{P}_y (1 - \hat{P}_y) + \frac{4}{n^2} \hat{P}_y \hat{P}_x (1 - \hat{P}_x)(1 - \hat{P}_y)$  sont respectivement des estimations des variances de  $\hat{P}$  et  $\hat{P}_*$ .

Sous l'hypothèse nulle de l'incapacité de  $x_t$  à prédire  $y_t$ , le test PT suit la distribution normale standard :  $S_n \sim N(0, 1)$ . Les travaux de Pesaran et Timmermann (1994) montrent que le test du market timing est mieux interprété comme un test exact d'indépendance dans un tableau de contingence 2 x 2 où les sommes des colonnes et des lignes sont fixes (ou données a priori).

#### 1.4 Changement directionnel et apprentissage automatique en macroéconomie

Goulet Coulombe *et al.* (2020) ont récemment publié un article qui passe en revue les performances de diverses techniques d'apprentissage automatique dans la prévision de l'activité économique au Québec et au Canada. L'étude a examiné six groupes différents de modèles, notamment des modèles factoriels, des régressions pénalisées, des régressions régularisées par sous-ensembles complets, des régressions vectorielles de support, des forêts d'arbres aléatoires et des réseaux de neurones. Chacun de ces modèles propose des approches uniques pour gérer de grands ensembles de données et générer des formes fonctionnelles complexes. Les auteurs ont mené un exercice de prévision hors échantillon pour 16 variables macroéconomiques au Québec et au Canada, en utilisant de vastes ensembles de données provenant du Canada et des États-Unis. Lors de l'évaluation des performances, ils ont pris en compte deux métriques spécifiques : la première est l'erreur quadratique tandis que la seconde mesure la capacité des modèles à correctement prévoir la direction des variables cibles. C'est cette métrique qui nous intéresse et qui a été calculée à l'aide du test de signe prédictif de Pesaran et Timmermann (1992). Pour la variable cible réelle, telle que le PIB, les résultats démontrent que la plupart des méthodes d'apprentissage automatique affichent de bonnes performances. Les taux de succès sont généralement très élevés et significatifs particulièrement à court

terme au Canada et en comparaison avec le Québec. Le Random forest de type local linear forest (LLF-CSF) prévoit mieux la direction du PIB québécois à 1 et 2 trimestres, avec un taux de réussite de 85% du temps. Les modèles AR-BIC et la régression régularisée par sous-ensembles complets (T-CSR10) fonctionnent bien aux plus longs horizons au Canada comme au Québec. Dans le cas du PIB canadien, ARDI-BIC et sa variante Lasso dominant à court terme, suivi du modèle de référence à long terme. Aucune autre étude n'a démontré avec succès la possibilité d'examiner la précision directionnelle des prévisions macroéconomiques tant au Québec qu'au Canada, ce qui rend cette étude unique par ses conclusions.

Lorsque nous examinons la littérature existante en macroéconomie, il devient évident que seules quelques études ont eu recours à des méthodes de classification supervisée pour évaluer l'exactitude directionnelle des prévisions. L'une des études pionnières dans ce domaine a été menée par Cornec et Mikol (2011), dans laquelle les auteurs ont utilisé une analyse discriminante linéaire/quadratique, des arbres de décision et des machines à vecteurs de support pour prédire le changement directionnel immédiat du PIB français. La logique adoptée par ces auteurs repose sur le classement du profil  $\varepsilon_t$  de la croissance, encore inconnu, entre les deux catégories « accélération du PIB » et « décélération du PIB » :

$$\varepsilon_t = \mathbb{1}_{\{y_t \geq y_{t-1}\}}^5 \quad (1.7)$$

où  $\mathbb{1}_{[\cdot]}$  est une fonction indicatrice qui est égale 1 si l'argument est vrai et 0 sinon ;  $y_t$ , le taux de croissance au trimestre  $t$  publié dans les premiers résultats du trimestre  $t$ , c'est-à-dire à  $t+45$  jours.

En reproduisant séquentiellement le cadre dans lequel les prévisions du conjoncturiste sont réalisées chaque trimestre, et en s'appuyant notamment sur les enquêtes de conjoncture auprès des entreprises, ils ont appliqué les méthodes mentionnées précédemment pour prévoir le profil de la croissance du PIB pour le trimestre suivant. Leurs résultats ont démontré que l'analyse discriminante linéaire surpassait les autres modèles d'apprentissage automatique ainsi que les modèles de référence en termes de précision des prévisions. Un autre résultat intéressant de cette étude est le fait que les enquêtes de conjoncture contiennent des informations très pertinentes pour la prévision immédiate du profil du PIB.

Une autre étude qui suit Cornec et Mikol (2011) mais à quelques différences <sup>6</sup> près, a été menée par Van Ruth (2014). L'auteur examine le sens de variation des niveaux de cinq indicateurs économiques mensuels aux

---

5. Avec cette notation, il y a accélération au trimestre  $t$  (resp. décélération) si  $\varepsilon_t = 1$  (resp.  $\varepsilon_t = 0$ ).

6. Les différences entre les deux études se résument à :

— l'étude de Cornec et Mikol (2011) met l'accent sur le taux de croissance trimestriel du PIB français alors que celle de

Pays-Bas. Ces variables comprennent la production manufacturière, les exportations, la consommation des ménages, la confiance des producteurs et la confiance des consommateurs. L'objectif principal est de prévoir si le niveau de ces variables, noté  $y_t$  à  $t$  sera plus grand ou plus petit à  $t+1$ . Pour y parvenir elle introduit une variable binaire qui prend la valeur 0 si la variation  $\Delta$  du niveau  $y_t$  d'un mois à l'autre est négative, et la valeur 1 si elle est positive. Pour évaluer les modèles, le jeu de données a été divisé en deux parties : l'une pour l'apprentissage et l'autre pour faire des analyses de performances. Le critère d'évaluation utilisé était le pourcentage de changements correctement classés d'un mois à l'autre. Divers modèles ont été testés, tels que des modèles de régression (la régression par la méthode des moindres carrés ordinaires (MCO), le modèle linéaire généralisé - logit combiné avec le modèle dynamique à facteurs) et des modèles d'apprentissage automatique (les arbres de décision, les forêts aléatoires, les machines à vecteurs de support, l'analyse discriminante linéaire, le boosting et le bagging). Quel que soit le modèle, les résultats ont montré que les modèles prédictifs ont surpassé les références naïves. Les meilleures précisions de prévision hors échantillon se situaient entre 65,5 % et 80 % et la combinaison du modèle linéaire généralisé logit avec des estimations factorielles dynamiques s'est avérée la plus robuste et la plus performante. Cependant, il est important de noter que l'étude présentait une limite en termes de données disponibles, qui étaient relativement faibles pour obtenir une estimation robuste des techniques d'apprentissage automatique.

En conclusion, très peu d'articles traitent de l'application des méthodes d'apprentissage automatique pour la prévision du changement directionnel de la croissance du PIB. À notre connaissance, aucune étude antérieure n'a examiné les performances des méthodes de classification supervisée pour prédire le changement directionnel de la croissance du PIB québécois. Pour combler cette lacune, la présente étude est motivée par l'application de ces méthodes et la comparaison de leurs performances.

---

Van Ruth (2014) examine le sens de variation des niveaux de cinq variables économiques mensuelles aux Pays-Bas ;

- l'étude de Van Ruth (2014) teste un plus large éventail de modèles, à la fois économétriques standard (régression MCO) et le modèle linéaire généralisé-logit combiné avec le modèle dynamique à facteurs, des modèles d'apprentissage automatique (arbre de décision, forêt aléatoires, machine à vecteurs de support, analyse discriminante linéaire, boosting et bagging)

## CHAPITRE 2

### APPROCHE MÉTHODOLOGIQUE

#### 2.1 Définition du changement directionnel de la croissance du PIB

Soit  $\{y_1, \dots, y_t\}$  les valeurs observées du taux de croissance du PIB réel désaisonnalisé pour le Québec.

Dans cette étude, nous avons, entre les dates  $t$  et  $t - 1$  :

— une *accélération du PIB* si :

$$y_t > y_{t-1} \quad (2.1)$$

— une *décélération du PIB* si :

$$y_t \leq y_{t-1} \quad (2.2)$$

**Définition 1.** Le changement directionnel de la croissance du PIB entre les dates  $t - 1$  et  $t$  est une séquence de variables binaires dont les résultats dépendent de la série temporelle  $y_t$  conformément à la règle suivante :

$$d_t = \mathbb{1}_{\{y_t > y_{t-1}\}} = \begin{cases} 1 & \text{si } y_t > y_{t-1} \\ 0 & \text{si } y_t \leq y_{t-1} \end{cases} \quad (2.3)$$

où  $\mathbb{1}\{\cdot\}$  est une fonction indicatrice qui est égale 1 si l'argument est vrai et 0 sinon.

Autrement dit, il y a un *signal* de changement directionnel de la croissance du PIB, si  $y_t > y_{t-1}$ . Par contre, s'il n'y a pas de *signal* si  $y_t \leq y_{t-1}$ . Par conséquent, une accélération au cours du trimestre  $t$  (respectivement décélération) est symbolisée par  $d_t = 1$  (respectivement  $d_t = 0$ ). Le problème posé est donc celui de classification binaire similaire aux études de Cornec et Mikol (2011).

Dans les sections suivantes du document,

**Définition 2.** notre objectif est de prédire, à partir de l'information  $X_t$  disponible à la date  $t$ , les prochaines réalisations  $\{\hat{d}_{t+1|t}, \dots, \hat{d}_{t+h|t}, \dots, \hat{d}_{t+H|t}\}$  et d'évaluer la performance de ces prévisions.  $h$  désigne l'horizon de prédiction qui est le nombre de trimestres à l'avance que l'on cherche à prédire. Dans ce cas, nous devons alors prévoir le changement directionnel comme suit :

$$\hat{d}_{t+h} = \mathbb{1}_{\{\hat{y}_{t+h} > \hat{y}_{t+h-1}\}} \quad (2.4)$$

## 2.2 Cadre d'analyse en apprentissage supervisé de la prévision du changement directionnel

Soit  $\mathcal{S}_t = \left\{ (X'_t, d_t), t = 1, \dots, T \right\}$  un jeu d'observations où  $X_t = (x_{1,t}, \dots, x_{n,t})'$  représente un vecteur de  $n$  variables explicatives dans un ensemble quelconque,  $\mathcal{X}$ , et leurs étiquettes  $d_t$  qui appartiennent à un ensemble  $\mathcal{Y}$ . Pour nos besoins, nous considérons  $\mathcal{Y} = \{0, 1\}$  et  $\mathcal{X} \in \mathbb{R}^p$ . Nous supposons que les étiquettes peuvent être obtenues à partir des observations grâce à une fonction fixe et inconnue  $g$  :

$$g : \begin{cases} \mathcal{X} & \longrightarrow \mathcal{Y} \\ (X'_t, d_t) & \longmapsto d_t = g(X'_t, d_t) + \eta_{t+1} \end{cases} \quad (2.5)$$

où  $\eta_t$  est un bruit blanc (faible).

Par ailleurs, nous définissons  $\mathcal{F} \subset \mathcal{Y}^{\mathcal{X}}$  comme l'ensemble des méthodes possibles.

**Définition 3.** Du point de vue de l'apprentissage supervisée, la prévision du changement directionnel de la croissance du PIB revient donc à déterminer une « méthode » (ou différentes classes de méthodes)  $\hat{g} \in \mathcal{F}$  qui approxime « au mieux »  $g : \mathcal{X} \longrightarrow \mathcal{Y}$  telle que

$$\begin{aligned} ((X'_t, d_t), g(X'_t, d_t)) &\in \mathcal{X} \times \mathcal{Y} \\ &\text{et} \\ \hat{g}(X'_t, d_t) &\simeq g(X'_t, d_t) \end{aligned} \quad (2.6)$$

$\hat{g}$  peut prendre des valeurs dans  $\{0, 1\}$  ou utiliser une fonction intermédiaire  $g$ , qui associe à une observation un score (par exemple, la probabilité que cette observation appartienne à la classe « 0 » ou « 1 ») et déterminer un seuil permettant de classer les observations :  $\hat{g}$  est obtenue en appliquant un seuil à  $g$ , ce dernier étant appelé la *fonction de décision* ou *fonction discriminante*.

**Définition 4.** Pour une classification binaire, une fonction de décision ou fonction discriminante est définie par :

$$g : \begin{array}{ccc} \mathcal{X} & \longrightarrow & \mathbb{R} \\ (X'_t, d_t) & \longmapsto & \hat{g}(X'_t, d_t) = \begin{cases} 1 & \text{si } g(X'_t, d_t) > 0 \\ 0 & \text{si } g(X'_t, d_t) \leq 0 \end{cases} \end{array} \quad (2.7)$$

## 2.3 Ensemble d'informations et stratégie de prévision

En utilisant les notations précédentes, nous pouvons définir l'ensemble d'informations à la date  $t$  comme suit :

**Définition 5.** L'ensemble d'informations à la date  $t$  est défini par :

$$I_t = \{\underline{X}_t, \underline{d}_t\} \quad (2.8)$$

où  $\underline{X}_t = (x_t, \dots, x_1)'$  et  $\underline{d}_t = (d_t, \dots, d_1)$ .

Nous pouvons alors définir une stratégie de prévision pour un horizon donné. Par souci de simplicité, nous supposons ici que  $h = 1$  (tout en écrivant les définitions de manière générique).

**Définition 6.** Pour un modèle  $g \in \mathcal{F}_{\mathcal{M}_i}$  où  $\mathcal{F}_{\mathcal{M}_i}$  est la  $i$ ème classe/famille de méthodes choisie, une stratégie de prévision à l'horizon  $h = 1$  ( $h \geq 1$ ) sachant l'ensemble d'information à la date  $t$  est le triplet  $\Pi_{h,t}^i = (\hat{g}_\lambda, I_t, h)$  est :

$$\hat{d}_{t+h|t} := g(I_t; \hat{\lambda}) \equiv \hat{g}_\lambda(I_t) \quad (2.9)$$

où  $\hat{\lambda}$  est l'estimation du vecteur des paramètres/hyperparamètres définissant  $\mathcal{M}_i$  et  $\hat{g}_\lambda$  est la meilleure approximation/estimation de  $g \in \mathcal{F}_{\mathcal{M}_i}$ .

## 2.4 Fonction de perte, erreur de prévision et critères de performance

### 2.4.1 Fonction de perte et erreur de prévision

Pour obtenir une approximation au mieux de  $g$ , il est nécessaire de se donner un critère qui permette de mesurer sa qualité. La plupart du temps, nous utilisons une fonction de perte qui fournit une mesure de l'erreur entre la variable prédite et la variable observée.

**Définition 7.** La fonction de perte, notée  $L$  quantifie la qualité d'une prévision. Elle attribue une valeur d'autant plus élevée que la prédiction  $\hat{g}_\lambda(I_t)$  s'éloigne de la vraie valeur  $g_\lambda(I_t)$ . Plus formellement, une fonction (mesurable)  $L : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}_+$  est une fonction de perte si elle satisfait les conditions suivantes :

$$\begin{aligned} L(g_\lambda(I_t), \hat{g}_\lambda(I_t)) &= 0 & \text{pour } g_\lambda(I_t) &= \hat{g}_\lambda(I_t) \\ L(g_\lambda(I_t), \hat{g}_\lambda(I_t)) &> 0 & \text{pour } g_\lambda(I_t) &\neq \hat{g}_\lambda(I_t) \end{aligned} \quad (2.10)$$

**Définition 8.** Pour la stratégie  $(\hat{g}_\lambda, I_t, h)$ , lorsque  $\mathcal{Y} = \{0, 1\}$ , la fonction de perte  $L$  prend la forme d'une perte zéro-un, définie par :

$$L(d_{t+h}, \hat{g}_\lambda(I_t)) = \begin{cases} 0 & \text{si } \hat{g}_\lambda(I_t) = d_{t+h} \\ 1 & \text{si } \hat{g}_\lambda(I_t) \neq d_{t+h} \end{cases} \quad (2.11)$$

Elle mesure simplement l'erreur commise par la méthode  $\hat{g}_\lambda$  en prédisant  $d_{t+h}$  à partir de  $I_t$ . En d'autres termes, si la méthode prédit correctement le changement directionnel du PIB à l'horizon  $h$ , la valeur de la perte est de 0, sinon, elle est de 1.

Dans les méthodes de classification qui prédisent des probabilités, il est souvent nécessaire de convertir cette prédiction en un résultat binaire (par exemple, 0 ou 1) selon un seuil  $s$ . Dans ce contexte,

**Définition 9.** La fonction de perte  $L$  est de la forme :

$$L(d_{t+h}, \hat{g}_\lambda(I_t)) = \begin{cases} 1 & \text{si } \hat{g}_\lambda(I_t) > s \\ 0 & \text{si } \hat{g}_\lambda(I_t) \leq s \end{cases} \quad (2.12)$$

où  $0 \leq s \leq 1$  puisque notre modèle est une probabilité.

Avec cette fonction de perte, nous pouvons donc définir l'erreur de prévision associée à  $\hat{g}_\lambda(I_t)$  :

**Définition 10.** Pour toute méthode  $g$  dans l'espace des méthodes possibles, l'erreur de prévision associée à  $\hat{g}_\lambda(I_t)$  est l'espérance de la fonction de perte  $L$  :

$$\mathcal{R}(\hat{g}_\lambda) = \mathbb{E}[L(d_{t+h}, \hat{g}_\lambda(I_t))] \quad (2.13)$$

En toute rigueur,

$$\mathcal{R}(\hat{g}_\lambda) = \mathbb{E}_{X,Y \sim P}[L(d_{t+h}, \hat{g}_\lambda(I_t))] \quad (2.14)$$

où  $P$  est la loi de probabilité jointe de  $(X, Y)$ .

#### 2.4.2 Critères de performance du changement directionnel

Pour un échantillon donné,  $\mathcal{S}_t$ , il est possible de comparer les classes réelles et prédites, et de dénombrer les bonnes des mauvaises prédictions. C'est l'objet de la matrice de confusion qui est représentée sous la forme d'un tableau de contingence à double entrée qui croise les valeurs réelles avec celles prédites du changement directionnel de la croissance du PIB par un modèle. Cela nous donne une vue d'ensemble des performances de chaque méthode de classification utilisé et des types d'erreurs qu'il commet. Le tableau 2.1 présente une matrice de confusion pour deux classes ( $d_{t+h} = 1$  et  $d_{t+h} = 0$ ).

Tableau 2.1 – Exemple de matrice de confusion pour la classification binaire

		Classes réelles		Total
		$d_{t+h} = 1$	$d_{t+h} = 0$	
Classes prédites	$\hat{d}_{t+h} = 1$	Vrais positifs (VP)	Faux positifs (FP)	VP + FP
	$\hat{d}_{t+h} = 0$	Faux négatifs (FN)	Vrais négatifs (VN)	FN + VN
Total		VP + FN	FP + VN	Nb obs.

Note : Ce tableau montre les quatre catégories de résultats d'une matrice de confusion. Les **vrais positifs** (respectivement les **vrais négatifs**) représentent les cas où la méthode prédit correctement «  $\hat{d}_{t+h} = 1$  » (respectivement «  $\hat{d}_{t+h} = 0$  ») comme «  $d_{t+h} = 1$  » (respectivement «  $d_{t+h} = 0$  »); les **faux positifs**, aussi appelés « fausses alarmes » ou « erreurs de type I », représentent les cas où la méthode prédit incorrectement «  $\hat{d}_{t+h} = 0$  » comme «  $d_{t+h} = 1$  »; à l'inverse, les **faux négatifs**, également nommés « erreurs de type II », désignent les cas la méthode prédit incorrectement «  $\hat{d}_{t+h} = 1$  » comme «  $d_{t+h} = 0$  ».

Du tableau 2.1, nous pouvons dériver plusieurs métriques fréquemment utilisées pour évaluer les performances des méthodes de classification :

— *Accuracy ou exactitude*

Il représente le pourcentage de trimestres de changement directionnel, VP (vrai positif) et VN (vrai négatif), correctement prédits par rapport à toutes les prédictions. Il se calcule comme suit :

$$\text{Exactitude (\%)} = \frac{VP + VN}{VP + VN + FN + FP} * 100 \quad (2.15)$$

Dans la littérature sur l'analyse directionnelle, ce critère correspond au ratio de succès (*success ratio (SR)* en anglais) qui a été utilisé dans plusieurs études antérieures (voir, par exemple, Skabar (2013)). C'est aussi le complémentaire à un du taux d'erreur de classification<sup>7</sup> :

$$\text{Taux d'erreur de classification (\%)} = \frac{FP + FN}{\text{Nb obs.}} * 100 \quad (2.16)$$

Le terme « précision » est souvent utilisé pour traduire « Accuracy ». Cependant, la précision est un autre critère. L'exactitude est certes facile à interpréter, mais elle peut présenter des limites en cas de déséquilibre des classes ou lorsque les erreurs de type I (faux positifs) et de type II (faux négatifs) n'ont pas le même coût. Dans notre cas, où les classes sont quasi équilibrées, ce critère reste informatif, mais ne suffit pas.

7. Un taux d'erreur plus faible est généralement souhaité car il signifie que le modèle fait moins d'erreurs dans ses prédictions.

— *Rappel* (*recall* en anglais) ou sensibilité (*sensitivity* en anglais) ou taux de vrais positifs (TVP)

C'est un critère central dans la mesure où il se focalise, pour une classe donnée (par exemple « *accélération* ( $d_{t+h} = 1$ ) », sur la proportion de trimestres de changement directionnel « *accélération* ( $d_{t+h} = 1$ ) » correctement prédits :

$$\text{Rappel (\%)} = \frac{VP}{VP + FN} * 100 \quad (2.17)$$

Le rappel est une mesure appropriée si les faux positifs l'emportent sur les faux négatifs. Ce qui est le cas ici, car une mauvaise détection d'une décélération de la croissance du PIB pourrait induire des décisions retardées.

— *Précision*

La précision ou valeur positive prédictive est la proportion de trimestres de changement directionnel « *accélération* ( $d_{t+h} = 1$ ) » correctement prédits parmi tous les trimestres de changement directionnel « *accélération* » :

$$\text{Précision (\%)} = \frac{VP}{VP + FP} * 100 \quad (2.18)$$

Ce critère est essentiel si nous voulons limiter les faux positifs, par exemple, éviter d'annoncer à tort une accélération de la croissance du PIB.

— *F-mesure ou F-score*

La F-mesure ou la F-score est une mesure d'arbitrage entre la précision et le rappel à l'aide d'un paramètre  $\beta$ .

$$F_{\beta} = \frac{(1 + \beta^2) \cdot (\text{précision} \cdot \text{rappel})}{(\beta^2 \cdot \text{précision} + \text{rappel}), \beta > 0 \quad (2.19)$$

où  $\beta$  est un paramètre qui contrôle les poids relatifs de la précision et du rappel. Pour  $\beta < 1$  (resp.  $\beta > 1$ ), plus d'importance est accordée à la précision (resp. au rappel). Lorsque  $\beta = 1$ , la F-mesure ou la F1-score correspond à la moyenne harmonique de la précision et du rappel, et fournit donc une évaluation équilibrée des performances de nos méthodes, tout en tenant compte des faux positifs et des faux négatifs. Dans ce cas, nous avons :

$$\text{F1-score} = 2 * \frac{\text{Précision} * \text{Rappel}}{\text{Précision} + \text{Rappel}} \quad (2.20)$$

— Spécificité ou taux de vrais négatifs (TVN)

C'est la proportion de périodes de changement directionnel « *décélération* ( $d_{t+h} = 0$ ) » correctement prédites :

$$\text{Spécificité (\%)} = \frac{VN}{VN + FP} * 100 \quad (2.21)$$

La spécificité est utile dans des contextes où les faux positifs doivent être limités.

### 2.4.3 Courbe ROC (Receiver Operator Characteristic) et critères associés

La courbe ROC décrit l'évolution de la sensibilité en fonction du complémentaire à 1 de la spécificité lorsque le seuil varie de 0 à 1 :

$$ROC_s = (TVP_s, 1 - TVN_s) = \left( \frac{VP_s}{VP_s + FN_s}, 1 - \frac{VN_s}{VN_s + FP_s} \right) \quad \text{pour } s \in [0, 1] \quad (2.22)$$

C'est un tracé qui représente  $TVP_s$  sur l'axe vertical et  $1 - TVN_s$  sur l'axe horizontal, montrant ainsi comment le taux de vrais positifs et le taux de faux positifs varient en fonction du seuil  $s$  (voir la figure 2.1 ci-dessus). En d'autres termes, la courbe ROC nous permet de déterminer pour quel seuil nous pouvons à la fois minimiser le taux de faux positifs et maximiser le taux de vrais positifs.

A l'aide de cette courbe, nous pouvons calculer l'aire sous chaque courbe ROC : **AUROC**. Ce critère de performance indique dans quelle mesure chaque modèle est capable de faire la discrimination entre les périodes de changement directionnel « *accélération* » et « *décélération* ». Elle peut être calculée comme suit (Lahiri et Yang, 2023) :

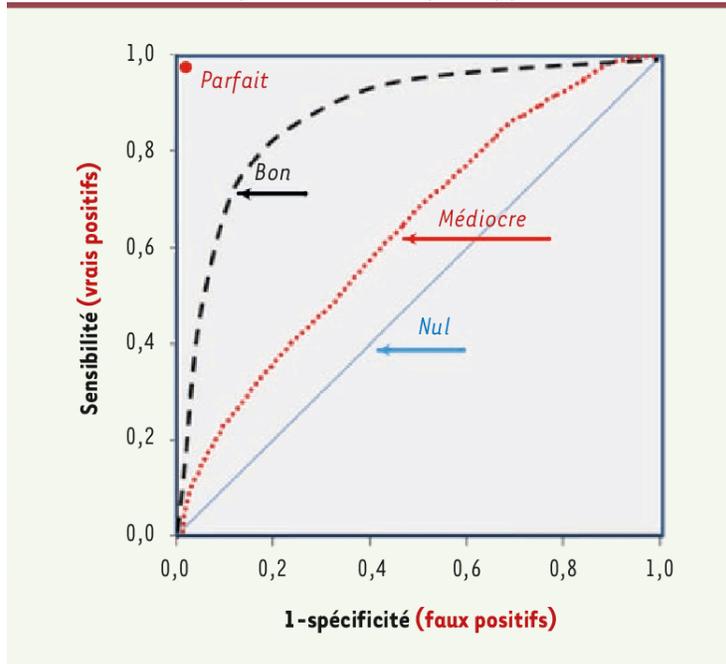
$$AUROC = \frac{1}{2} \sum_{j=1}^M (TVP(s_j) + TVP(s_{j-1})) (TFP(s_j) + TFP(s_{j-1})) \quad (2.23)$$

où  $s$  est le seuil de probabilité,  $M$  est le nombre de seuils.

La valeur de l'aire sous la courbe ROC (AUROC) varie entre 0 (cas pire) et 1 (cas meilleur). Par exemple, le point rouge dans le coin supérieur gauche de la figure 2.1 indique que la performance est parfaite. Cependant, elle peut masquer des déséquilibres dans la classification réelle si la distribution des probabilités est mal calibrée.

Figure 2.1 – Courbes ROC et qualité des prédictions

(Source: Jordan (2020) )



Note : La figure illustre une courbe ROC. Lorsqu'elle coïncide avec la droite bleue (Nul), le nombre de faux positifs est égal à celui des vrais positifs, indépendamment du seuil  $s$  sélectionné. Une méthode située sur cette ligne n'apporte aucune valeur ajoutée. En revanche, lorsque la courbe coïncide avec celle rouge en pointillés (Médiocre), cela indique une méthode qui présente une certaine capacité de discrimination, bien que moindre. Il est à noter que plus la courbe ROC s'éloigne de la droite bleue (Nul) tout en restant au-dessus (courbe en pointillés noirs), plus la méthode est performante. Le point rouge dans le coin supérieur gauche représente une méthode parfaite, où la sensibilité est égale à 1 et la spécificité à 1, conduisant à l'absence de faux positifs et à la détection de l'ensemble des vrais positifs.

#### 2.4.4 Choix du seuil de changement directionnel

Nous pouvons nous intéresser à la probabilité de prévision du changement directionnel de la croissance du PIB, soit une « *décélération* » ou une « *accélération* » en fonction d'un seuil. Concrètement, nous estimons, pour chaque observation  $X_t$ , une probabilité  $\mathbb{P}(d_{t+h} = 1 \mid X_t) = \hat{g}_\lambda(I_t)$  d'appartenance à la classe  $\hat{d}_{t+h|t} = 1$  (ou  $\hat{d}_{t+h|t} = 0$  lorsque  $\mathbb{P}(d_{t+h} = 0 \mid X_t) = \hat{g}_\lambda(I_t)$ ). Une solution envisageable consisterait à introduire une valeur seuil  $s$  fixée a priori :

$$\hat{d}_{t+h} = \mathbb{1}_{\{\hat{g}_\lambda(I_t) > s\}} = \begin{cases} 1 & \text{si } \mathbb{P}(d_{t+h} = 1 \mid X_t) = \hat{g}_\lambda(I_t) > s \\ 0 & \text{si } \mathbb{P}(d_{t+h} = 1 \mid X_t) = \hat{g}_\lambda(I_t) \leq s \end{cases} \quad (2.24)$$

Nous transformons donc les probabilités prédites en prédictions binaires basées sur ce seuil. Par exemple, un seuil de  $s = 0.50$ <sup>8</sup> permet d'interpréter une probabilité de changement directionnel de la croissance du PIB située dans l'intervalle  $[0, 0.5]$  comme est un résultat « *décélération (0)* » et une probabilité dans l'intervalle  $]0.5, 1]$  comme un résultat « *accélération (1)* ». Les résultats ainsi obtenus peuvent être résumés dans un tableau similaire à celui présenté au tableau 2.1. Le seuil  $s$  tel que  $0 \leq s \leq 1$  correspond donc au point de coupure qui détermine s'il y a une accélération ou une décélération de la croissance du PIB.

Dans le contexte de la prévision du changement directionnel de la croissance du PIB, notre objectif est de prédire avec précision si la croissance du PIB accélérera ou décélérera au cours des prochains trimestres. Or, une accélération ou une décélération de la croissance du PIB qui avait été prévue peut ou non se produire. Donc, les faux positifs et les faux négatifs doivent être évalués, comme le souligne Döpke *et al.* (2019). Les faux positifs ( $\hat{d}_{t+h} = 1$  et  $d_{t+h} = 0$ ) ou erreurs de type I consistent à prédire une accélération de la croissance du PIB alors qu'en réalité la croissance décélère. À l'inverse, les faux négatifs ( $\hat{d}_{t+h} = 0$  et  $d_{t+h} = 1$ ) ou erreurs de type II impliquent de prévoir une décélération alors que la croissance s'accélère. Ces deux types d'erreurs peuvent être coûteux, selon le cas d'utilisation spécifique. Par exemple, les tests sont typiquement biaisés vers les faux positifs au point où nous voulons contrôler la probabilité de commettre ce type d'erreur. Les faux positifs ont tendance à être plus coûteux en termes de bien-être que les faux négatifs. Ils sont susceptibles d'entraîner la mise en œuvre de mesures économiques qui peuvent aggraver une situation économique qui, en réalité, n'en était pas une. Collodel (2022) souligne que « [...] *manquer une [décélération] a un coût élevé en termes de bien-être, par exemple, des contractions importantes de la production et une augmentation du chômage* ». C'est pourquoi nous cherchons à contrôler la probabilité de commettre ce type d'erreur afin d'éviter des décisions économiques incorrectes et des ressources gaspillées.

Pour notre exercice de prévision, il est donc important de minimiser les faux positifs, même si cela implique une certaine tolérance envers les faux négatifs. Ce choix signifie que nous préférons éviter de classer à tort une décélération comme une accélération, même si cela nous amène à manquer certaines accélérations réelles. En somme, notre objectif est de déterminer un seuil qui permette de minimiser les risques liés aux prédictions erronées d'accélération (faux positifs), tout en acceptant le risque de manquer certaines accélérations (faux négatifs). Cela nous permet de prendre des décisions économiques plus prudentes et mieux alignées sur les conditions réelles de l'activité économique. Dans ce cas, nous définissons un point

---

8. Ce seuil est celui fixé par défaut dans la littérature.

spécifique sur la courbe ROC de manière à sélectionner le seuil qui nous permet d'établir un équilibre entre les faux positifs et les vrais négatifs. L'indice de Youden (Youden, 1950) est un point de la courbe où la prédiction du changement de la croissance du PIB présente la meilleure combinaison de deux critères de performances ci-après :

- sensibilité : la capacité à identifier correctement les périodes d'accélération de la croissance du PIB ;
- spécificité : la capacité à identifier correctement les périodes de décélération de la croissance du PIB.

Il est défini comme suit :

$$J = \text{Sensibilité} + \text{Spécificité} - 1 = \text{TVP} + \text{TVN} - 1 \quad (2.25)$$

## 2.5 Spécification du modèle à facteurs et des méthodes de classifications supervisée

Dans cette section, nous exposons brièvement la théorie qui sous-tend le modèle à facteurs. Ensuite, nous présentons les huit méthodes de classification supervisée pour prédire le changement directionnel de la croissance du PIB québécois. Afin de déterminer si nos méthodes choisies sont potentiellement utiles pour prédire le changement directionnel de la croissance du PIB, nous avons besoin d'une méthode de référence pour la comparaison. Le classifieur aléatoire est utilisé pour cette tâche.

Dans un souci de cohérence avec la littérature existante, nous adopterons la notation et les conventions méthodologiques utilisées par des auteurs tels que Piger (2020), Iworiso et Vrontos (2020), Vrontos *et al.* (2021a), Vrontos *et al.* (2021b).

### 2.5.1 Modèle à facteurs

La millésime de LCDMA choisie, celle de mars 2020, inclut 477 séries temporelles trimestrielles observées entre 1981Q1 et 2020Q1 et qui sont susceptibles d'influencer la prévision du changement directionnel de la croissance du PIB québécois  $d_{t+h}$ . Elle est notée  $X_t$ . Le nombre d'éléments de  $X_t$  comparés par rapport aux 157 périodes temporelles disponibles, nous interpelle sur le problème très connu dans la littérature de la *malédiction de la dimensionnalité* (Stock et Watson, 2002). Pour le contourner, l'idée consiste à réduire la dimension de  $X_t$  par les composantes principales, pour les inclure ensuite comme variables explicatives dans les différentes méthodes à appliquer. Cette réduction se traduit par un ensemble plus restreint de composantes principales non corrélées, qui capturent l'essentiel de la variation des données. C'est pourquoi

nous optons pour les étapes d'estimation des facteurs communs suivies par Fortin-Gagnon *et al.* (2022) pour réduire la dimension de notre jeu de données.

Dans cette approche, l'étape préliminaire consiste à tracer le diagramme des valeurs propres de la matrice de corrélation des données (scree plot) et celui de la proportion de variances expliquées par les  $k$  premiers facteurs (trace). Le premier graphique est une représentation courante pour déterminer le nombre de facteurs alors que le second montre la proportion de la variance de l'ensemble de données qui est expliquée de manière cumulative par chaque composante.

Nous supposons que  $X_t$  peut être représenté par une structure factorielle à deux composantes inobservables orthogonales : une composante engendrée par des facteurs communs ( $\Lambda.F_t$ ) à l'ensemble des prédicteurs et une composante idiosyncratique ( $\xi_t$ ) comme dans l'équation 2.26 :

$$X_t = \Lambda.F_t + \xi_t \text{ pour } t = 1, \dots, T \quad (2.26)$$

où  $T$  est le nombre d'observations dans les séries chronologiques,  $\Lambda$  est la matrice des pondérations de dimension  $(N \times r)$ ,  $F_t$  représente les facteurs communs de dimension  $(r \times 1)$  tels que  $F_t = (F_{1t}, \dots, F_{rt})'$  et  $r$  le nombre de facteurs communs contenus dans  $F_t$ , le vecteur  $(\xi_t) = (\xi_{1t}, \dots, \xi_{Nt})'$  de dimension  $(N \times 1)$  est composé des  $N$  composantes idiosyncratiques avec  $N$ , le nombre de prédicteurs observés et stationnaires contenus dans  $X_t$ .

L'équation 2.26 peut être estimée par la méthode d'analyse en composantes principales (ACP) qui minimise (Stock et Watson, 2002) :

$$V(r) = \min_{\Lambda, F^r} (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T (X_{it} - \Lambda_i' F_t^r)^2 \quad (2.27)$$

sous la contrainte de la normalisation  $\frac{\Lambda^{r'} \Lambda^r}{N} = I_r$  ou  $\frac{F^{r'} F^r}{T} = I_r$ .

Si nous nous concentrons  $\Lambda^k$  et utilisons la normalisation  $\frac{F^{r'} F^r}{T} = I_k$ , minimiser l'équation 2.27 équivaut à :

$$\max(\text{tr}(F^{r'}(XX')F^r)), \quad (2.28)$$

sous les contraintes :  $\frac{F^{r'} F^r}{T} = I_r$

Sous ces conditions, l'estimateur de la matrice des facteurs,  $\tilde{F}^r$  est  $T^{\frac{1}{2}}$  fois les vecteurs propres correspondant aux  $r$  plus grandes valeurs propres de la matrice  $X'X$  de taille  $T \times T$ . Étant donné  $\tilde{F}^r$ , l'estimateur

de la matrice de poids est alors :

$$\tilde{\Lambda}^{r'} = \frac{\tilde{F}^{r'} X}{\tilde{F}^{r'} \tilde{F}^r} = \frac{\tilde{F}^{r'} X}{T} \quad (2.29)$$

La solution au problème de minimisation de l'équation 2.27 ci-dessus n'est pas unique, même si la somme des résidus quadratiques  $V(r)$  est unique. Une autre solution est donnée par  $(\bar{F}^r, \bar{\Lambda}^r)$  où  $\bar{\Lambda}^r$  est construit comme  $N^{\frac{1}{2}}$  fois les vecteurs propres correspondant aux  $r$  plus grandes valeurs propres de la matrice  $X'X$  de taille  $N \times N$ . Étant donné  $\frac{\bar{\Lambda}^{r'} \bar{\Lambda}^r}{N} = I_r$ , l'estimateur de la composante principale de  $F$  est alors :

$$\bar{F}^r = \frac{X \bar{\Lambda}^r}{N} \quad (2.30)$$

Les diagrammes Scree plot et Trace sont parfois critiqués pour leur subjectivité. Par exemple, le diagramme Scree plot peut présenter de multiples "coudes" qui rendent difficile la détermination du nombre correct de facteurs ou de composantes à retenir, ce qui le rend peu fiable. C'est pourquoi de nombreuses procédures statistiques ont été proposées pour déterminer le nombre de facteurs communs  $r$  dans les modèles factoriels (voir par exemple, Fortin-Gagnon *et al.* (2022)). Dans ce travail, nous considérons le critère d'information  $IC_{p2}$  de Bai et Ng (2002) qui est le plus utilisé dans la littérature :

$$IC_{p2} = \ln(V(r, \hat{F}^r)) + r \left( \frac{N+T}{NT} \right) \ln(C_{NT}^2), \quad (2.31)$$

où  $\left( \frac{N+T}{NT} \right) \ln(C_{NT}^2) = g(N, T)$  avec  $C_{NT}^2 = \min\{\sqrt{T}, \sqrt{N}\}$  et,

$$V(r, \hat{F}^r) = (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T (X_{it} - \Lambda_i^{r'} \hat{F}_t^r)^2 \text{ avec } X_{it} \text{ l'observation } i \text{ au temps } t, \lambda_i \text{ la } i^{\text{ème}} \text{ ligne de } \lambda \text{ et } F = (F_1, F_2, \dots, F_T).$$

Pour déterminer le nombre de facteurs, il faut sélectionner la valeur de  $r$  qui minimise  $IC_{p2}$  par rapport à  $r$ . Comme l'objectif est de réduire la dimension de l'espace de données avec un nombre limité de facteurs, seules les  $r$  premières composantes principales sont choisies et utilisées comme variables explicatives dans nos différentes méthodes de classification. Pour éviter l'incertitude causée par le nombre de facteurs, nous rapportons tous les résultats à huit facteurs dans notre étude empirique <sup>9</sup>.

---

9. Fortin-Gagnon *et al.* (2022) limitent le nombre de facteurs à un maximum de  $r_{max} = 8$  dans leur analyse de 115 variables macroéconomiques canadiennes

## 2.5.2 Classifieur aléatoire (*DummyClassifier* ou DC)

Nous définissons un classifieur aléatoire comme la méthode de référence où la prévision du changement directionnel de la croissance du PIB revient à tirer aléatoirement une variable aléatoire ayant une distribution de Bernoulli de paramètre  $\frac{1}{2}$  :

$$\hat{d}_{t+h} = \begin{cases} 0 & \text{avec } \mathbb{P}(d_{t+h} = 1 \mid F_t) = \frac{1}{2} \\ 1 & \text{avec } \mathbb{P}(d_{t+h} = 1 \mid F_t) = \frac{1}{2} \end{cases} \quad (2.32)$$

Cette stratégie est notée  $(\hat{g}_\lambda, I_t, h)^a$ . Il est immédiat que l'espérance de la fonction de perte 0/1 de la stratégie aléatoire pour  $h = 1$  est  $\frac{1}{2}$ , soit  $\mathbb{E}[L_{0/1}(d_{t+1}, \hat{g}_\lambda(I_t))^a] = \frac{1}{2}$ .

En considérant une fonction de perte cumulée et normalisée de la forme :

$$L_c(d_{t+h}, \hat{g}_\lambda(I_t))^a = \frac{1}{t+h} \sum_{j=1}^{t+h} \mathbb{1}_{d_j \neq \hat{d}_j}, \quad (2.33)$$

il est alors immédiat de montrer, sous l'hypothèse d'indépendance, que  $(t+h)L_c(d_{t+h}, \hat{g}_\lambda(I_t))^a$  suit une loi binomiale de paramètre  $n = t+h$  et  $p = \frac{1}{2}$ . Dès lors, on peut évaluer la probabilité suivante pour les stratégies associées aux modèles  $\mathcal{M}_1, \dots, \mathcal{M}_k$  :

$$\mathbb{P}(L_c(d_{t+h}, \hat{g}_\lambda(I_t))^a \leq L_c(d_{t+h}, \hat{g}_\lambda(I_t))^i) \quad (2.34)$$

Une stratégie  $L_c(d_{t+h}, \hat{g}_\lambda(I_t))^i$  est statistiquement significative au seuil  $\alpha$  par rapport à une stratégie aléatoire si cette probabilité est inférieure ou égale à  $\alpha$  : la stratégie aléatoire a moins de  $\alpha\%$  de chance de conduire à un taux d'erreur plus faible que la stratégie  $i$ .

Le classifieur aléatoire est mis en œuvre dans Python à l'aide de la bibliothèque *Scikit-Learn* dédiée au machine learning, via la méthode *DummyClassifier*. Ce dernier implique plusieurs stratégies naïves :

- *most\_frequent* : prédire simplement la classe la plus fréquente ;
- *stratified* : prédire les classes en fonction de leur fréquence d'apparition ;
- *uniform* : prédire toutes les classes uniformément, quelle que soit leur fréquence d'apparition.

Dans notre cas, le *DummyClassifier* avec la stratégie *uniform* est en parfaite adéquation avec la définition d'un classifieur aléatoire. Il attribue une probabilité égale (50 %) à chacune des deux possibilités (accélération ou décélération).

### 2.5.3 Régression logistique (Logistic Regression ou LR)

La régression logistique relie la probabilité de survenue du changement directionnel  $d_{t+h}$  de la croissance du PIB  $h$  trimestres à l'avance, aux facteurs communs  $F_t$  :

$$P(d_{t+h} = 1|F_t) = \frac{e^{(\beta_0 + F_t^\top \beta)}}{1 + e^{(\beta_0 + F_t^\top \beta)}} = \pi(F_t) \quad (2.35)$$

et donc,

$$P(d_{t+h} = 0|F_t) = 1 - \pi(F_t) \quad (2.36)$$

où  $\beta_0$  une constante appelée intercept et  $\beta$  un vecteur de coefficients de régression à estimer.

La transformation logit effectuée sur  $\pi(F_t)$  dans l'équation 1.6 nous permet d'obtenir le modèle logistique linéaire simple suivant :

$$\ln(\pi(F_t)) = \ln\left[\frac{\pi_t(F_t)}{1 - \pi_t(F_t)}\right] = \beta_0 + f_t^\top \beta \quad (2.37)$$

Pour ajuster la régression logistique, les paramètres  $(\beta_0, \beta)$  sont estimés à l'aide de la méthode d'inférence statistique du maximum de vraisemblance (Talon *et al.*, 2018) :

$$\begin{aligned} (\hat{\beta}_0, \hat{\beta}) &= \operatorname{argmin}_{(\beta_0, \beta) \in \mathbb{R}^{p+1}} \left( -2\ell(\beta_0, \beta) \right) \\ &= \operatorname{argmin}_{(\beta_0, \beta) \in \mathbb{R}^{p+1}} \left( -2 \sum_{j=1}^N d_{i,t+h} \ln(P(d_{i,t+h} = 1|F_{i,t})) \right. \\ &\quad \left. + (1 + d_{i,t+h}) \ln(P(d_{i,t+h} = 0|F_{i,t})) \right) \end{aligned} \quad (2.38)$$

Une fois les paramètres estimés, il est possible d'obtenir les probabilités ajustées de changement directionnel de la croissance du PIB au Québec à  $h$  trimestre à l'avance peuvent être obtenues comme suit :

$$\hat{P}_{t+h} = \hat{\pi}(F_t) = \frac{e^{(\hat{\beta}_0 + F_t^\top \hat{\beta})}}{1 + e^{(\hat{\beta}_0 + F_t^\top \hat{\beta})}} \quad (2.39)$$

### 2.5.4 Classifieur bayésien naïf (Naive Bayes ou NB)

C'est une méthode de classification supervisée qui trouve son origine dans le théorème de Bayes qui suppose une hypothèse forte d'indépendance conditionnelle : les facteurs communs  $F_t$  sont tous conditionnellement indépendants les uns des autres, étant donné la variable à prédire  $d_{t+h}$ . La règle de Bayes associe à toute représentation  $F_t \in \mathcal{F}$  :

$$\hat{g}_{\text{Bayes}}(F_t) = \operatorname{argmax}_{d_{t+h} \in \mathcal{Y}} \Pr(F_t | d_{t+h}) \Pr(d_{t+h}) \quad (2.40)$$

La loi de probabilité jointe étant inconnue, on ne peut pas calculer  $\Pr(F_t|d_{t+h})$  et dans une certaine mesure  $\Pr(d_{t+h})$  (la loi a priori des classes). Pour surmonter cet obstacle, on suppose alors que les éléments du vecteur d'entrée  $F_t$  sont tous conditionnellement indépendants les uns des autres, étant donné la variable à prédire  $d_{t+h}$ . Par conséquent,

$$\Pr(F_t|d_{t+h}) = \prod_{i=1}^N \Pr(f_{i,t}|d_{t+h}) \quad (2.41)$$

Nous pouvons donc ré-écrire (2.40) comme suit :

$$\hat{g}_{\text{Bayes}}(F_t) = \operatorname{argmax}_{d_{t+h} \in \mathcal{Y}} \prod_{i=1}^N \Pr(f_{i,t}|d_{t+h}) \Pr(d_{t+h}) \quad (2.42)$$

En suivant Piger (2020), la probabilité a posteriori qu'un changement directionnel  $d_{t+h} = 1$  étant donné les facteurs communs ( $N = 8$ )  $F_t = (f_{1,t}, \dots, f_{N,t})'$  peut être obtenue comme suit :

$$\Pr(d_{t+h} = 1|F_t) = \hat{d}_{t+h} = \frac{\prod_{i=1}^8 \Pr(f_{i,t}|d_{t+h} = 1) \Pr(d_{t+h} = 1)}{\prod_{i=1}^8 \Pr(f_{i,t}|d_{t+h} = 1) \Pr(d_{t+h} = 1) + \prod_{i=1}^8 \Pr(f_{i,t}|d_{t+h} = 0) \Pr(d_{t+h} = 0)} \quad (2.43)$$

Dans les environnements riches en données, les performances du classifieur bayésien naïf peuvent se détériorer s'il existe un grand nombre de prédicteurs non pertinents. Dans le cadre de ce mémoire, nous avons considéré la réduction de la dimension de notre jeu de données  $X_t$  par les composantes principales comme la solution à ce problème.

### 2.5.5 Méthode des K-plus proches voisins (*K-nearest neighbors* ou KNN)

C'est un modèle non paramétrique qui fonctionne selon le principe : «dis moi qui sont tes amis, je te dirais qui tu es»<sup>10</sup>. Considérons  $N_k(F_t)$ , un voisinage de  $F_t$  contenant uniquement les  $k$  plus proches voisins de l'ensemble d'apprentissage. Pour prédire  $d_{t+h}(F_t)$ , nous déterminons les  $k$  valeurs de  $F$  les plus proches de  $F_t$  sur la base d'un vote majoritaire parmi les  $k$  plus proches voisins de  $F_t$ , c'est-à-dire que nous calculons la proportion :

$$\hat{d}_{t+h}(F_t) = \frac{1}{k} \sum_{F_i \in N_k(F_t)} I_{(d_{i+h} \in \{0,1\})} \quad (2.44)$$

où  $I(\cdot)$  est la fonction indicatrice.

Pour que la méthode KNN puisse mieux fonctionner sur un ensemble de données particulier, nous devons choisir la métrique de distance la plus appropriée. Il en existe différentes métriques de distance, mais la

10. [www.r-bloggers.com/2013/11/les-k-plus-proches-voisins-vite-il-faut-se-la-reappropriier/](http://www.r-bloggers.com/2013/11/les-k-plus-proches-voisins-vite-il-faut-se-la-reappropriier/)

plus couramment utilisée est la distance euclidienne dont la formule de calcul se présente comme suit :

$$dist(F_t, F_i) = \sqrt{\sum_{j=1}^N (F_{j,t} - F_{j,i})^2} \quad (2.45)$$

où  $F_t$  est l'échantillon d'apprentissage et  $F_i$ , l'échantillon test.

Dans cette approche, nous cherchons à identifier les observations passées ( $F_i$ ) qui sont les plus proches de  $F_t$  en termes de distance euclidienne. Une fois ces voisins trouvés, ils sont utilisés pour prédire la valeur future de  $d_{t+h}(F_t)$ . Il est important de noter que la distance est calculée uniquement sur les valeurs passées de  $F$ , et que le modèle k-NN utilise ces informations historiques pour prédire la classe future de  $d_{t+h}(F_t)$ .

Malgré sa simplicité, la méthode KNN est soumise au fléau de la dimension. En effet, les KNN sont affectés négativement lorsque le vecteur d'apprentissage contient de nombreux prédicteurs non pertinents et redondants. La précision de prédiction diminue en présence de nombreuses variables inutiles ou d'une dimension  $N$  élevée. Dans le cadre de ce mémoire, nous avons considéré la réduction de la dimension de notre jeu de données  $X_t$  par les composantes principales comme la solution à ce problème.

Avec la méthode KNN, le choix de  $k$  est important car elle détermine les performances de la méthode. Une approche classique pour avoir cette valeur est la validation croisée.

### 2.5.6 Analyse linéaire discriminante (*Linear Discriminant Analysis* ou LDA)

L'analyse linéaire discriminante est une méthode de classification linéaire basée sur l'hypothèse que chaque classe peut-être modélisée par une distribution normale et que toutes les classes partagent la même matrice de covariance. Pour être plus explicite, considérons une tâche de classification de  $k$  classes où  $k = 0$  et  $k = 1$  sont les deux classes possibles du changement directionnel de la croissance du PIB québécois. Les données de chaque classe sont générées par une distribution normale  $F_t | d_{t+h} = k \sim \mathcal{N}(\mu_k, \Sigma_k)$  dont la fonction de densité est la suivante :

$$g_k(F_t | d_{t+h} = k) = \frac{1}{(2\pi)^{d/2} |\Sigma_k|^{1/2}} e^{-\frac{1}{2} (f - \mu_k)^T \Sigma_k^{-1} (f - \mu_k)} \quad (2.46)$$

et la matrice de covariance  $\Sigma_k$  est la même pour les deux classes :

$$\Sigma_k = \Sigma \quad (2.47)$$

Pour calculer les probabilités a posteriori de changement directionnel de la croissance du PIB québécois à  $h$  trimestres à l'avance,  $\Pr(d_{t+h} = k | F_t = f_t)$  étant donné l'ensemble des prédicteurs  $F_t$ , il faut estimer les paramètres inconnus de la méthode sur les données d'apprentissage :

les probabilités a priori d'appartenance aux classes :

$$\pi_k = \Pr(d_{t+h} = k) = \frac{N_k}{N}, \text{ avec } N_k = \sum_{t=1}^N \mathbf{1}_{d_{t+h}=k} \quad (2.48)$$

les moyennes des distributions normales :

$$\mu_k = \frac{1}{N_k} \sum_{t|d_{t+h}=k} f_t \quad (2.49)$$

la matrice de covariance :

$$\Sigma = \frac{1}{N - K} \sum_{k=1}^K \sum_{t|d_{t+h}=k} (f_t - \mu_k)^T (f_t - \mu_k) \quad (2.50)$$

avec  $K$  le nombre de classes.

Une fois les paramètres sont estimés, la probabilité a posteriori de changement directionnel de la croissance du PIB québécois,  $\Pr(d_{t+h} = k | F_t = f_t)$ , étant donné l'ensemble des prédicteurs  $F_t$ , se déduit de la formule de Bayes comme suit :

$$\Pr(d_{t+h} = k | F_t = f_t) = \frac{\pi_k g_k(f_t)}{\sum_{k=1}^K \pi_k g_k(f_t)} \quad (2.51)$$

### 2.5.7 Machines à vecteurs de support (*Support Vector Machines* ou SVM)

Considérons un problème de classification binaire avec  $S_t = \left\{ \left( F_t', d_{t+h} \right), t = 1, \dots, T \right\}$  un jeu de données de  $T$  observations où  $d_{t+h} \in \{-1, +1\}$  est un vecteur cible,  $F_t = (f_{1,t}, \dots, f_{8,t})'$  représente le vecteur des facteurs communs envisagés comme variables explicatives. Dans l'espace des variables, nous supposons qu'il existe une frontière linéaire qui sépare les observations de la classe "-1" de celles de "+1". Le concept de base des machines à vecteurs de support ou séparateurs à vaste marge introduites par Cortes et Vapnik (1995), pour la classification binaire, provient du fait qu'elles reposent sur la recherche d'un hyperplan de marge optimale qui, lorsque c'est possible, classe ou sépare correctement les données tout en étant le plus éloigné possible de toutes les observations. Formellement, la recherche d'un hyperplan de marge optimale consiste à trouver un hyperplan d'équation  $g(f_t) = w \cdot F_t + b = 0$  remplissant simultanément les deux conditions suivantes :

— bonne séparation des données dans les différentes classes ( $d_{t+h} \in \{-1, +1\}$ ), au sens où :

$$g(f_t) = \begin{cases} +1 & \text{si } w \cdot F_t + b \geq 0 \\ -1 & \text{si } w \cdot F_t + b < 0 \end{cases} \quad (2.52)$$

— la marge qui est la distance des classes à l'hyperplan séparateur donnée par  $d(f_t) = \frac{|w \cdot F_t + b|}{\|w\|}$  est maximale, ce qui revient à minimiser  $\|w\|$  sous contraintes.

Mathématiquement, l'idée équivaut à résoudre le problème de minimisation <sup>11</sup> sous contraintes suivant :

$$\begin{cases} \min & \frac{1}{2} \|w\|_2^2 \\ \text{s.c.} & d_{t+h}(w \cdot F_t + b) \geq 1 \quad \forall t : 1 \leq t \leq T \end{cases} \quad (2.53)$$

On rappelle qu'un point  $(F_t, d_{t+h})$  est bien classé si et seulement si  $d_{t+h}(w \cdot F_t + b) > 0$  et nous nous imposons  $d_{t+h}(w \cdot F_t + b) \geq 1$ . De plus,  $d_{t+h} = \frac{w \cdot F_t + b}{\|w\|}$ .

Si l'on retient cette forme du problème (2.53), nous faisons face malheureusement aux problèmes classiques des méthodes d'apprentissage automatique, à savoir, le surajustement et la malédiction de la dimensionnalité. Pour résoudre ces problèmes, il est nécessaire d'introduire une forme dual de ce problème d'optimisation impliquant des multiplicateurs de Lagrange pour chaque contrainte. Dans ce cas, le problème posé (2.53) devient :

$$\begin{cases} \max_{\lambda} & \sum_{t=1}^T \lambda_t - \frac{1}{2} \sum_{t,j=1}^T \lambda_t \lambda_j d_{t+h} d_{j+h} F_t \cdot F_j \\ \text{s.c.} & \lambda_t \geq 0 \quad \forall t : 1 \leq t \leq T \\ & \sum_{t=1}^T \lambda_t d_{t+h} = 0 \end{cases} \quad (2.54)$$

L'avantage du problème (2.54) par rapport au problème (2.53) est que la fonction objectif ne dépend désormais que des multiplicateurs lagrangiens, ce qui est plus facile à résoudre analytiquement.

Soit  $\mathbf{L}(w, b, \lambda)$  le Lagrangien du problème 2.53 :

$$\mathbf{L}(w, b, \lambda) = \frac{1}{2} \|w\|_2^2 - \sum_{t=1}^T \lambda_t (d_{t+h}(w \cdot F_t + b) - 1) \quad (2.55)$$

*Calcul de  $w^*$  et  $b^*$*

La résolution du problème (2.53) passe par l'utilisation des conditions Karush, Kuhn et Tucker (KKT). Ici,  $w^*$  et  $b^*$  est solution optimale du problème (2.53) si et seulement s'il existe  $\lambda^* = (\lambda_1, \lambda_2, \dots, \lambda_T)$  où

11. Le problème 2.53 vient du fait que maximiser une marge équivaut en réalité à minimiser le carré de son inverse.

$\forall t : 1 \leq t \leq T, \lambda_t \geq 0$ , tel que  $(\mathbf{w}^*, \mathbf{b}^*)$  vérifie :

$$\forall 1 \leq t \leq T, \quad \lambda_t^*(d_{t+h}(\mathbf{w}^* \cdot F_t + \mathbf{b}^*) - 1) = 0 \quad (2.56)$$

et  $\lambda^*$  maximise  $\mathbf{L}(\mathbf{w}^*, \mathbf{b}^*, \lambda)$  sous les contraintes  $\frac{\partial L}{\partial w} = 0$  et  $\frac{\partial L}{\partial b} = 0$ , ce qui est respectivement équivalent à :

$$\mathbf{w}^* = \sum_{t=1}^N \lambda_t d_{t+h} F_t \quad (2.57)$$

où les termes non nuls dans cette somme sont ceux relatifs aux vecteurs supports. Les observations qui ne sont pas vecteurs supports ne participent pas à la définition de l'hyperplan optimal (Frédéric, 2020). Ce résultat est cohérent avec l'interprétation géométrique du problème.

Et

$$\sum_{t=1}^N \lambda_t d_{t+h} = 0 \quad (2.58)$$

Ce qui nous amène caractériser la fonction de décision  $h(f_t)$  de notre classification linéaire comme suit :

$$h(f_t) = \sum_{t=1}^N \lambda_t d_{t+h} (w \cdot F_t + b) \quad (2.59)$$

Et finalement, le classifieur SVM linéaire est donnée par :

$$g(x_t) = \text{sign}(h(f_t)) = \text{sign}\left(\sum_{t=1}^N \lambda_t d_{t+h} (w \cdot F_t + b)\right) \quad (2.60)$$

Parmi les méthodes SVM, on recense les cas linéairement séparable (qui vient d'être présenté ci-dessus) et les cas non linéairement séparable. La plupart des problèmes réels ne sont pas sujette à une séparation linéaire entre les données. Dans ce cas, il s'agira de changer l'espace des données. Nous allons donc avoir un changement de dimension. Cette nouvelle dimension est appelée « espace de re-description » (Hasan et Boris, 2006). Plus la dimension de cet espace est grande, plus la probabilité de pouvoir trouver un hyperplan séparateur entre les exemples est élevée. Nous avons donc une transformation d'un problème de séparation non linéaire dans l'espace de représentation en un problème de séparation linéaire dans un espace de re-description de grande dimension (Hasan et Boris, 2006). Dans ce cas, pour que le problème (2.54) soit réalisable, il faut que nous introduisons des variables de relaxation  $\zeta_t$  pour assouplir les contraintes du problème (2.53). Nous avons donc de nouvelles contraintes appelées contraintes de relaxation :

$$d_{t+h}(w \cdot F_t + b) \geq 1 - \zeta_t \quad \text{avec} \quad \zeta_t \geq 0 \quad \forall t : 1 \leq t \leq T \quad (2.61)$$

Les problèmes (2.53) et (2.54) deviennent respectivement :

$$\left\{ \begin{array}{l} \min \quad \frac{1}{2} \| w \|^2_2 \\ \text{s.c.} \quad d_{t+h}(w \cdot F_t + b) \geq 1 \quad \forall t : 1 \leq t \leq T \\ \zeta_t \geq 0 \quad \forall t : 1 \leq t \leq T \\ C > 0, \end{array} \right. \quad (2.62)$$

$$\left\{ \begin{array}{l} \max_{\lambda} \quad \sum_{t=1}^N \lambda_t - \frac{1}{2} \sum_{t,j=1}^N \lambda_t \lambda_j d_{t+h} d_{j+h} F_t \cdot F_j \\ \text{s.c.} \quad 0 \leq \lambda_t \leq C \quad \forall t : 1 \leq t \leq T \\ \sum_{t=1}^N \lambda_t d_{t+h} = 0 \end{array} \right. \quad (2.63)$$

où  $C$  est un paramètre qui règle le compromis entre la marge possible et le nombre de points mal classés admissibles. Le calcul de  $w^*$ ,  $b^*$ , de la fonction de décision  $h(f_t)$  et du classifieur SVM non-linéaire suivent la même démarche que pour le cas linéaire.

Il convient de signaler que la transformation non linéaire est réalisée via une fonction noyau. Nous distinguons les exemples de noyaux suivants : polynomiale, gaussien (*Radial Basis Function* ou RBF), sigmoïde et laplacien. Dans le cadre de ce mémoire, nous avons choisi d'utiliser le noyau RBF pour notre application. Ce choix se justifie par sa capacité à capturer des relations non linéaires complexes entre les variables macroéconomiques, qui sont souvent difficiles à modéliser avec des méthodes linéaires.

### 2.5.8 Arbres de décision (*Decision Tree* ou DT)

Les arbres de décision sont une méthode d'apprentissage supervisé qui prédit les valeurs de la variable à expliquer en apprenant des règles de décision dérivées des variables explicatives. Ils peuvent être utilisés en régression ou en classification. Par conséquent, ils sont appelés CART pour *Classification And Regression Tree*<sup>12</sup>.

#### *Construction d'un arbre maximal*

CART procède à une partition récursive de l'échantillon d'apprentissage sur la base d'une condition binaire de la forme  $F_t^j < \tau$  et  $F_t^j \geq \tau$ , où  $j$  et  $\tau$  peuvent être différents d'une division à l'autre (Piger, 2020). Pour

12. Ils ont été développés par Breiman *et al.* (1984)

montrer comment fonctionne cette partition récursive, un exemple à deux prédicteurs est illustré. Nous choisissons les prédicteurs  $F_t^1$  et  $F_t^2$  pour discriminer le changement directionnel  $d_{t+h}$  de la croissance du PIB québécois sous la forme d'une structure arborescente créée en fonction d'une conjonction de conditions binaires. Nous nous plaçons d'abord au nœud racine  $F_t^1$  et nous considérons que les observations sont triées par  $F_t^1$ . Celles au-dessous du seuil  $\tau_1$  sont assignées au nœud 1 ( $F_t^1 < \tau_1$ ). Celles associées aux plus grandes valeurs de  $F_{1,t}^1$  ( $F_t^1 \geq \tau_1$ ) sont ensuite triées par  $F_t^2$ . Les observations qui satisfont  $F_t^1 \geq \tau_1$  et  $F_t^2 < \tau_2$  remplissent le nœud 2, tandis que les cas  $F_t^1 \geq \tau_1$  et  $F_t^2 > \tau_2$  se retrouvent dans le nœud 3. Ainsi, à chaque étape de partitionnement récursif, un nœud « mère » est divisé en deux nœuds « fils ». La division s'arrête de deux façons : Soit les nœuds 2 et 3 ne contiennent que des observations  $d_{t+h} = 1$  ou  $d_{t+h} = 0$ , soit une règle d'arrêt s'applique, telle qu'un nombre maximal donné de divisions pour éviter un surajustement important. La totalité de ces divisions récursives partitionne l'espace d'échantillonnage en  $M$  nœuds qui ne se chevauchent pas avec :

- $A_m^*$  désignant le nœud  $m$  pour  $m = 1, \dots, M$ ;
- $T_m^*$  désignant les observations de l'échantillon d'entraînement dans chaque nœud  $m$  pour  $m = 1, \dots, M$ .

Pour tous les prédicteurs  $F_t$  qui sont dans le nœud  $A_m^*$ , la prédiction pour  $d_{t+h} = 1$ <sup>13</sup> est une constante égale à la proportion de l'échantillon de la classe 1 à l'intérieur du nœud :

$$P_{A_m^*} = \frac{1}{T_m^*} \sum_{F_t \in A_m^*} \mathbb{1}_{(d_{t+h}=1)} \quad (2.64)$$

où  $\mathbb{1}_{[\cdot]}$  est une fonction indicatrice qui est prend la valeur 1 si l'argument est vrai, et la valeur 0 sinon.

Le classifieur CART pour la classe 1 est alors donné par :

$$\hat{d}_{t+h} = \frac{1}{T_m^*} \sum_{m=1}^M P_{A_m^*} \mathbb{1}_{(d_{t+h}=1)} \quad (2.65)$$

Avec cette information, nous pouvons, selon le seuil  $s$ <sup>14</sup>, classer le changement directionnel à  $t+h$  comme une accélération ou une décélération de la croissance du PIB par rapport à l'ensemble des prédicteurs  $F_t$ .

Il convient de rappeler que pour déterminer un prédicteur de division et un point de division à chaque étape de partition récursive décrite ci-dessus, nous définissons l'impureté des nœuds fils, le plus souvent par le

---

13. Classe 1 pour l'étiquette 1 et classe 0 pour l'étiquette 0

14. Il est à noter que le seuil de probabilité  $s \in [0, 1]$ .

biais de l'indice de Gini ou d'impureté. L'indice de Gini d'un nœud  $t$  est défini comme suit :

$$G = \sum_{c=1}^C \hat{p}_t^c (1 - \hat{p}_t^c) \quad (2.66)$$

où  $\hat{p}_t^c$  est la proportion de classe  $c$  dans le nœud  $t$  et  $C$  le nombre de classes. L'indice prend des valeurs comprises entre 0 et 1 et une valeur 0 indique un nœud pur où une seule classe est présente alors que des valeurs plus élevées indiquent une plus grande diversité des classes. Les arbres de décision souffrent du problème de la variance élevée des estimateurs. La méthode souvent utilisée pour résoudre ce problème consiste, entre autres, à combiner les prévisions provenant de beaucoup d'arbres qui sont créés de façon aléatoire. D'où la forêt aléatoire.

### 2.5.9 Forêt aléatoire (*Random forest* ou RF)

Les forêts aléatoires (Breiman, 2001) sont des méthodes d'apprentissage automatique connues pour leur capacité à gérer des problèmes de prédiction bruyants, non linéaires et de grande dimension. Elles apportent des remèdes aux défauts des arbres de décision, augmentent leur robustesse et permettent de conserver tous les avantages des arbres<sup>15</sup> (Stéphane, 2015). Comme leur nom l'indique, les forêts aléatoires sont une collection d'arbres construits en tirant au hasard des arbres de classification ou de régression sur un jeu de données bootstrap. Des variables choisies au hasard font également partie du mélange. Du point de vue du compromis biais-variance, un grand ensemble d'arbres entraînés sur des échantillons bootstrap indépendants, chacun avec une variance relativement importante mais un faible biais, peut réduire considérablement la variance sans introduire de biais supplémentaire. Le pseudo-code de la méthode des forêts aléatoires est résumée dans l'algorithme 1.

---

15. Dans Stéphane (2015), il est souligné que les arbres de décision présentent de nombreux atouts : une grande facilité d'utilisation avec des réglages simples et peu de travail préliminaire sur les données, des calculs rapides, la capacité à traiter des variables à la fois qualitatives et quantitatives, la détection des interactions entre les variables et la lisibilité des résultats.

---

**Algorithme 1** Construction des forêts aléatoires pour la classification (Breiman, 2001)

---

**Entrées :**

- $S_t = \{ (F'_t, d_{t+h}) \}, t = 1, \dots, T$  : ensemble d'apprentissage avec  $F_t = (f_{1t}, f_{2t}, \dots, f_{Nt})$ .
- $B$  : nombre d'arbres dans la forêt.
- $m \in \{1, \dots, p\}$  : nombre de variables explicatives à sélectionner aléatoirement à chaque nœud.

**Algorithme :** Pour  $b = 1$  à  $B$  :

1. Générer  $B$  échantillons bootstrap de taille  $n$  à partir de  $S_t : \tilde{S}_t^1, \dots, \tilde{S}_t^B$ .
2. Pour chaque échantillon bootstrap  $\tilde{S}_t^b$ , construire un arbre en répétant de manière récursive les étapes suivantes, jusqu'à ce que la taille minimale des nœuds  $n_{\min}$  soit atteinte :
  - Choisir  $m$  variables au hasard parmi les  $p$  variables explicatives.
  - Sélectionner la meilleure coupure en ne considérant que les  $m$  variables candidates.
  - Partitionner le nœud en deux nœuds fils.
3. Noter  $\hat{h}_b(x)$  l'arbre construit.

**Sortie :** En considérant les arbres construits, la prédiction à un nouveau point  $(x, y)$  s'obtient en faisant voter ces arbres à la majorité selon l'expression suivante :

$$\hat{g}(x) = \arg \max_y \sum_{b=1}^B \mathbb{1}(\hat{h}_b(x)=y)$$

où  $\hat{h}_b(x)$  est la prédiction de la variable cible  $d_{t+h}$  en  $x$  à l'aide du  $b^{\text{ième}}$  arbre.

---

Les forêts aléatoires nécessitent la configuration de nombreuses paramètres :

- le nombre d'arbres  $b$  dans les forêts aléatoires est par défaut 500. Ce paramètre ne nécessite de réglage spécifique, mais des valeurs plus grandes conduisent à des prédictions plus stables que des valeurs plus petites.
- le calibrage du nombre de variables choisies pour la répartition des données de chaque nœud est important pour analyser avec précision les performances des forêts. Sa valeur par défaut, dans le cas de la classification est  $\sqrt{p}$ .
- lorsque le nœud est divisé, un nombre minimum d'observations doit être enregistré en dessous duquel il ne se sépare plus. Ce paramètre prend par défaut la valeur  $n_{\min} = 1$  dans le cas de la classification.
- par défaut, chaque échantillon bootstrap contient  $n$  observations tirées au hasard à partir de l'ensemble d'apprentissage.

En général, les valeurs par défaut des paramètres donnent de bons résultats (Poterie, 2018). Dans le même temps, les forêts aléatoires calculent une estimation de son erreur de généralisation : l'erreur out-of-bag

(OOB), où « out-of-bag » signifie « en dehors du bootstrap » (Genuer *et al.*, 2017). De plus, pour pallier leur manque d'interprétabilité, les forêts aléatoires intègrent des mécanismes permettant de hiérarchiser les variables explicatives en fonction de leur importance pour la réponse  $Y$ . Il s'agit principalement du coefficient de Gini et de l'importance de la permutation.

## 2.6 Recherche des hyperparamètres optimaux et exercice de prévision

Cette section décrit l'approche utilisée pour obtenir les prévisions du changement directionnel de la croissance du PIB à plusieurs horizons et évaluer les performances hors échantillon des différentes méthodes de classification supervisée. Elle s'organise en deux parties. : la première détaille la recherche des hyperparamètres optimaux pour chaque méthode et la deuxième expose l'exercice de prévision réalisé.

### 2.6.1 Recherche des hyperparamètres optimaux

Toutes nos méthodes impliquent de choisir des hyperparamètres avant de procéder à l'estimation. En effet, leur choix permet de contrôler la complexité des méthodes et, par conséquent, d'optimiser leurs performances sous-jacentes (Vrontos *et al.*, 2021b). Pour obtenir les meilleures performances d'une méthode, selon Schratz *et al.* (2019), les hyperparamètres ne peuvent pas être définis manuellement. C'est pourquoi les auteurs préconisent une optimisation automatique pour déterminer le meilleur réglage. Dans la littérature sur l'apprentissage automatique, plusieurs méthodes et configurations d'algorithmes sont proposées pour atteindre ce but. Ces méthodes incluent la recherche en grille, la recherche aléatoire, les algorithmes de recherche stochastique (par exemple, les algorithmes génétiques) et l'optimisation bayésienne.

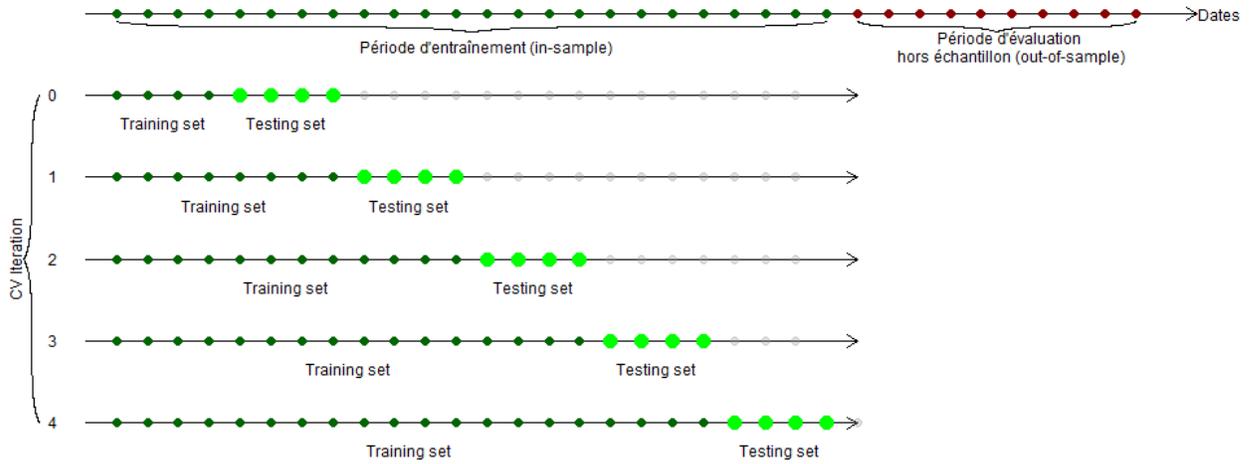
Dans le présent mémoire, le choix des hyperparamètres optimaux est réalisé à l'aide de l'algorithme génétique implémenté dans la librairie Python *GASearchCV* de *sklearn-genetic-opt*. Cette librairie se base sur les algorithmes évolutionnaires du package *DEAP (Distributed Evolutionary Algorithms in Python)*, qui imitent les principes de la génétique et de la sélection naturelle établis par Darwin (Darwin, 1964). Le fonctionnement de *GASearchCV* se distingue par la combinaison de l'algorithme génétique et de la validation croisée (*cross-validation* ou CV), ce qui permet de trouver les meilleures combinaisons d'hyperparamètres pour une méthode donnée. La validation croisée limite le surajustement en garantissant que les hyperparamètres obtenus généralisent bien sur des données non utilisées, tandis que l'algorithme génétique explore l'espace des hyperparamètres en simulant un processus évolutif (sélection, croisement et mutation) afin de maximiser une fonction d'aptitude, ici basée sur l'exactitude (*accuracy*). Pour plus

d'informations, le lecteur pourra consulter la documentation de scikit-learn à l'adresse suivante : <https://sklearn-genetic-opt.readthedocs.io/en/stable/index.html>.

Plusieurs approches de validation croisée (CV) existent, mais la plus populaire est basée sur un rééchantillonnage aléatoire (K-fold) qui paraît mal adapté pour la prévision des séries temporelles, comme dans notre cas. Selon Bergmeir *et al.* (2018), la validation croisée k-fold standard n'est pas simple à mettre en œuvre en raison de la corrélation sérielle inhérente et de la non-stationnarité des données de séries temporelles. Ceux-ci démontrent toutefois qu'elle fonctionne avec les modèles autorégressifs tant que les résidus sont non corrélés. Cela dit, il existe des situations où l'utilisation de la validation croisée K-fold peut être justifiée. Les travaux de Goulet Coulombe *et al.* (2022) ont montré que les résultats obtenus avec la validation croisée K-fold standard et celle hors-échantillon (*pseudo-out-of-sample cross-validation* ou POOS CV) sont semblables lorsque les données ne présentent pas de changements structurels significatifs au fil du temps. De plus, Basher et Sadorsky (2022) ont étudié la prédiction de la direction des prix du Bitcoin à l'aide des forêts aléatoires. Ceux-ci ont pu constater que l'utilisation de certaines caractéristiques, telles que les moyennes mobiles MA50 et MA200, qui intègrent des informations passées sur les cours des actions, permet de réduire la corrélation sérielle résiduelle, rendant l'utilisation possible de la méthode K-fold.

Pour notre approche, nous avons prêté une attention particulière à la manière d'utiliser la validation croisée qui préserve les dépendances temporelles avec *GAsearchCV*. Nous avons configuré la fonction K-fold de Python avec un argument *shuffle = False* qui garantit que la dépendance temporelle est respectée pendant la période d'entraînement. Nous avons fixé le nombre de plis à *n\_plis = 5*. Cela revient à diviser, sur une base continue, les périodes *in-sample* en 5 sous-échantillons de taille égale. Pour chacun des 5 sous-ensembles et chaque ensemble d'hyperparamètres considéré, les 4 premiers sous-ensembles (formant la période d'entraînement ou *training set* en langage d'apprentissage automatique) sont utilisés pour l'estimation. Les observations restantes, correspondant au cinquième sous-ensemble des périodes *in-sample*, servent à évaluer la performance hors échantillon avec le score d'exactitude (*accuracy*) comme critère de performance. La moyenne des scores obtenus lors des 5 opérations permet d'estimer la capacité de généralisation de la méthode. La figure 2.2 est une visualisation du déroulement de cette description qui est adaptée de Dörr (2020).

Figure 2.2 – Validation croisée des séries temporelles



Note : La figure illustre la répartition globale du jeu de données entre les périodes d'entraînement (*in-sample*) et les périodes d'évaluation hors échantillon (*out-of-sample*) dans la partie supérieure. Les observations d'entraînement sont représentées par des points verts et les observations d'évaluation par des points rouges. La répartition respecte strictement l'ordre temporel et garantit que l'ensemble d'entraînement précède l'ensemble d'évaluation. Dans la boucle CV Iteration, les données d'entraînement sont à nouveau divisées en blocs d'ensembles training (points verts) et en blocs d'ensembles testing (points verts clairs).

### 2.6.2 Exercice de prévision du changement directionnel

L'approche utilisée pour évaluer les performances des différentes méthodes de classification supervisée consiste à réaliser un exercice historique de prévision hors échantillon. Les estimations sont ainsi effectuées à partir de données strictement passées, dans le but d'évaluer la capacité desdites méthodes à anticiper le changement directionnel de la croissance du PIB. Nous avons donc divisé le jeu de données de taille  $T$  en deux parties : une période d'entraînement (*in-sample*) et une période d'évaluation hors échantillon (*out-of-sample*) tout en conservant l'ordre temporel. Pour les séries temporelles, le partitionnement aléatoire ne saurait être considéré comme une méthode appropriée, en vertu de l'exigence stipulant que les données d'entraînement doivent se produire avant celles d'évaluation (Yazdani, 2020).

La période d'entraînement, de taille  $R$  s'étend du 1<sup>er</sup> trimestre de 1981 au 4<sup>ème</sup> trimestre de l'année 1999 (76 observations trimestrielles) et sert à l'estimation des différentes méthodes ainsi qu'à l'ajustement des hyperparamètres propres à chaque méthode. La période d'évaluation hors échantillon, de taille  $P$  allant du 1<sup>er</sup> trimestre de 2000 au 1<sup>er</sup> trimestre de l'année 2020 (81 observations trimestrielles), est réservée à

l'évaluation des performances hors échantillon de nos prévisions.

En partant de l'hypothèse que les informations disponibles jusqu'au temps  $t$  sont utilisées pour entraîner nos différentes méthodes et pour produire des prévisions du changement directionnel de la croissance du PIB à  $t + h$ , lesdites prévisions sont effectuées à l'aide d'un schéma de fenêtre glissante de taille fixe, s'étendant sur une période de 73 trimestres. Cette approche implique que la période d'entraînement sur laquelle les méthodes sont estimées, se déplace au fil du temps, tandis que celle hors échantillon se rétrécit à mesure que l'exercice progresse. On parle d'estimation sur fenêtre glissante ou de « rolling forecasting » (Phung, 2023). De façon classique, pour prévoir le changement directionnel de la croissance du PIB à un trimestre  $h$  donné, par exemple pour  $h = 1$ , la première cible est  $d_{2000Q2}$  étant donné l'information disponible en  $2000Q1$ . La méthode est estimée sur la fenêtre d'entraînement allant de  $1981Q4$  à  $1999Q4$ , soit 73 trimestres et ensuite la prévision  $\hat{d}_{2000Q2|2000Q1}$  est produite. Nous enregistrons la prévision obtenue ainsi que la valeur du changement directionnel de la croissance du PIB du trimestre  $2000Q2$  que l'on souhaitait prédire. Pour la deuxième prévision, la fenêtre glisse d'un trimestre et l'ensemble d'information est mis à jour en ajoutant  $d_{2000Q2}$  dans la période d'entraînement ( $1982Q1 - 2000Q1$  (73 trimestres)), et la prévision est ensuite faite pour  $d_{2000Q3}$ . Ce processus est répété jusqu'à ce que la dernière observation de la période d'évaluation hors échantillon soit prédite. À la fin, pour chaque exercice de prévision, nous disposons d'une valeur prévue  $\hat{d}_{t+h}$  et de celle attendue  $d_{t+h}$  du changement directionnel de la croissance du PIB pour la méthode et l'horizon testés. En exploitant ces valeurs, et à partir de la matrice de confusion, il est possible de calculer les erreurs de prévision obtenues hors échantillon à chaque horizon <sup>16</sup>.

---

16. Nous étudions les horizons de prévision  $h \in \{1, 2, 4, 6, 8\}$ .

## CHAPITRE 3

### DONNÉES

#### 3.1 Présentation de la base de données

Dans le cadre de ce mémoire, nous considérons la base de données construite par Fortin-Gagnon *et al.* (2022), dénommée LCDMA et accessible au public sur le site Web de la Chaire de macroéconomie et de prévision de l'École des sciences de la gestion (ESG) de l'Université du Québec à Montréal (UQAM). Cette base de données est régulièrement mise à jour en temps réel par l'intermédiaire de Statistique Canada (StatCan). Tout comme les bases de données FRED-MD et FRED-QD de McCracken et Ng (2016), elle présente l'intérêt d'éviter aux utilisateurs d'avoir à gérer les changements de données et les révisions méthodologiques. La base de données LCDMA est utile pour capturer les points de retournement du cycle économique canadien. De plus, elle a un pouvoir prédictif substantiel lors de la prévision d'indicateurs macroéconomiques clés. En effet, elle couvre en détail plusieurs indicateurs économiques canadiens et provinciaux qui ont été transformés pour garantir leur stationnarité. Les transformations pour toutes les variables peuvent être trouvées dans Fortin-Gagnon *et al.* (2022). Elle présente toutefois l'inconvénient d'être un peu petite. En particulier, les données sont recueillies depuis 1914 mais seules celles à partir de l'année 1981 peuvent être exploitées. En ce sens, il n'est visiblement pas possible, selon Stéphane (2015), de découper ce jeu de données en trois échantillons indépendants de taille suffisante, pour l'apprentissage, la validation et le test, sachant que l'approche rigoureuse de la modélisation basée sur l'apprentissage automatique consiste à ajuster la méthode sur un échantillon d'apprentissage, à utiliser l'échantillon de validation pour choisir entre plusieurs méthodes possibles celui qui semble la mieux se généraliser, puis à utiliser l'échantillon de test pour mesurer avec le moins de biais possible la qualité de généralisation de la méthode.

Pour notre exercice de prédiction, nous avons utilisé la base de données disponible en mars 2020 qui inclut 477 variables trimestrielles pour 157 périodes temporelles (1981Q1-2020Q1) après traitement. Nous nous sommes arrêtés au premier trimestre de l'année 2020 en raison de la crise sanitaire liée au virus de la COVID, qui a généré des valeurs extrêmes dans plusieurs variables d'activités économiques.

## 3.2 Variables d'intérêt et cible

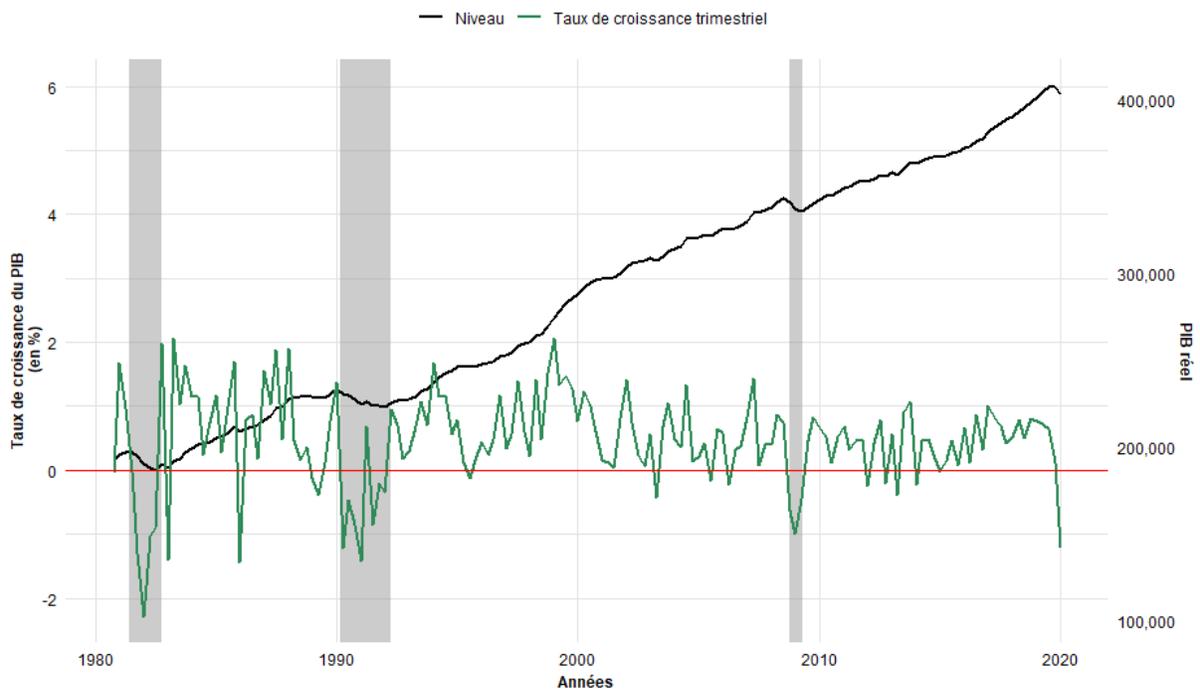
### Variable d'intérêt

La variable d'intérêt est le plus important agrégat macroéconomique : le PIB réel. Nous calculons les taux de croissance trimestriels du PIB réel à partir du niveau du PIB réel trimestriel, comme suit :

$$y_t = 100 \left( \frac{Y_t}{Y_{t-1}} - 1 \right) \quad (3.1)$$

où  $Y_t$  est le niveau du PIB réel au temps  $t$ . Le taux de croissance du PIB réel exprimé en pourcentage,  $y_t$ , constitue la variable d'intérêt qui nous servira à déterminer le changement directionnel de la croissance du PIB.

Figure 3.1 – Évolution du niveau du PIB réel trimestriel et de son taux de croissance trimestriel au Québec du 1<sup>er</sup> trimestre de 1981 au 1<sup>er</sup> trimestre de 2020



Note : L'axe de droite représente l'évolution du niveau du PIB réel (courbe noire) et celui de gauche, l'évolution de son taux de croissance (courbe verte) entre le 1<sup>er</sup> trimestre de 1981 et le 1<sup>er</sup> trimestre de 2020. Les zones ombrées représentent les périodes de récession que la province du Québec a traversées depuis le début des années 1980.

Comme illustré sur la figure 3.1, la croissance trimestrielle du PIB réel du Québec (courbe verte) a connu une évolution en dents de scie entre 1981 et 2020. Il est à noter que les plus fortes baisses de cet indicateur

ont été observées durant les récessions<sup>17</sup> et au début de la crise liée au COVID. L'activité économique a ainsi fortement diminué pendant la récession de 1981-1982 (-2,280 %). En revanche, la situation fut nettement différente durant la récession de 1990-1991, où l'économie a connu une baisse d'un peu plus de (-1,416 %) du PIB réel. La troisième récession a été la plus courte et la moins douloureuse. Celle-ci est survenue à la suite de la débâcle du marché immobilier aux États-Unis et de la crise financière. Au Québec, cette période de contraction de l'économie qui s'est étalée du 4<sup>ème</sup> trimestre de 2008 au 2<sup>ème</sup> trimestre de 2009, a été accompagnée d'une baisse de (-0,999 %) de la croissance du PIB. Quant à la courbe noire, elle illustre l'évolution trimestrielle du niveau du PIB réel au Québec sur la même période. Cette évolution, bien que ponctuée de périodes de baisse durant les récessions, affiche une tendance à long terme positive et croissante. À partir de 2010, une reprise marquée du PIB réel est observée, signalant l'amorce d'une phase de croissance soutenue. Enfin, il convient de noter que des baisses dramatiques sont observées en 2020, correspondant à la chute du taux de croissance du PIB (près de -1,218 %) et à une contraction notable du niveau du PIB réel. Cette situation est due à la crise sanitaire liée au virus de la COVID, qui a généré des valeurs extrêmes dans plusieurs variables d'activité économique.

## Cible

La cible est une variable binaire qui décrit le changement directionnel de la croissance du PIB québécois (1 = accélération, 0 = décélération) (Confère Définition 1 de la section 2.1). Elle comporte des observations dérivées sur la base de la variation du taux de croissance du PIB. Concrètement, si le signe de la variation du taux de croissance du PIB entre les dates  $t - 1$  et  $t$  est positive, l'observation prend la valeur 1; sinon, la valeur 0 est attribuée. Les séries du taux de croissance trimestriel du PIB québécois et du changement directionnel pour la période (1981Q1-2020Q1) sont illustrées dans le tableau 3.1. En outre, l'analyse de la distribution du changement directionnel, telle que présentée dans la figure 3.2 ci-dessous, ne révèle pas de disparité dans la distribution binaire des données. L'analyse des 157 points de données révèle que 79 trimestres (soit 50,32 %) sont marqués par des décélérations de la croissance du PIB selon la publication des données en temps réel. Ceci donne un ratio de distribution de classes d'environ 1:0,987, ce qui ne peut être considéré comme un cas de déséquilibre de classes.

---

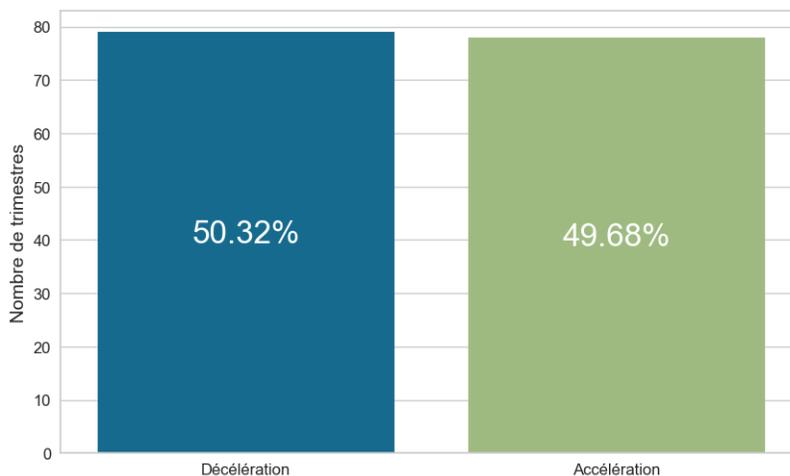
17. <https://www.cdhowe.org/turning-points-business-cycles-in-canada-since-1926/19364>

Tableau 3.1 – Quelques valeurs illustratives de la croissance du PIB québécois et de sa direction

Années	Trimestre 1		Trimestre 2		Trimestre 3		Trimestre 4	
	Croissance	Direction	Croissance	Direction	Croissance	Direction	Croissance	Direction
	du PIB		du PIB		du PIB		du PIB	
1981	0,33%	1	0,97%	0	-0,03%	0	-1,30%	0
1982	-2,31%	0	-1,05%	1	-0,891%	1	1,974%	1
1983	-1,390%	0	2,058%	1	1,033%	0	1,640%	1
...	...	...	...	...	...	...	...	...
2018	0,533%	1	0,780%	1	0,498%	0	0,816%	1
2019	0,756%	0	0,728%	0	0,651%	1	0,125%	0
2020	-1,218%	0	-	-	-	-	-	-

Note : Le tableau présente les séries du taux de croissance trimestriel et de la direction du PIB québécois pour la période (1981Q1 – 2020Q1). Par exemple, nous observons que  $d_t = 1$  au cours du 4<sup>ème</sup> trimestre de l'année 2018, indiquant que le taux de croissance du PIB au cours de ce trimestre (+0,816 %) a été supérieur au taux enregistré au 3<sup>ème</sup> trimestre de la même année (+ 0,498 %).

Figure 3.2 – Distribution des classes du changement directionnel de la croissance du PIB au québécois



Note : La figure illustre de la distribution des classes du changement directionnel de la croissance du PIB québécois. Une analyse visuelle des données révèle que 49,68 % des observations sont associées à une période de décélération, tandis que 50,32 % des observations sont liées à une période d'accélération.

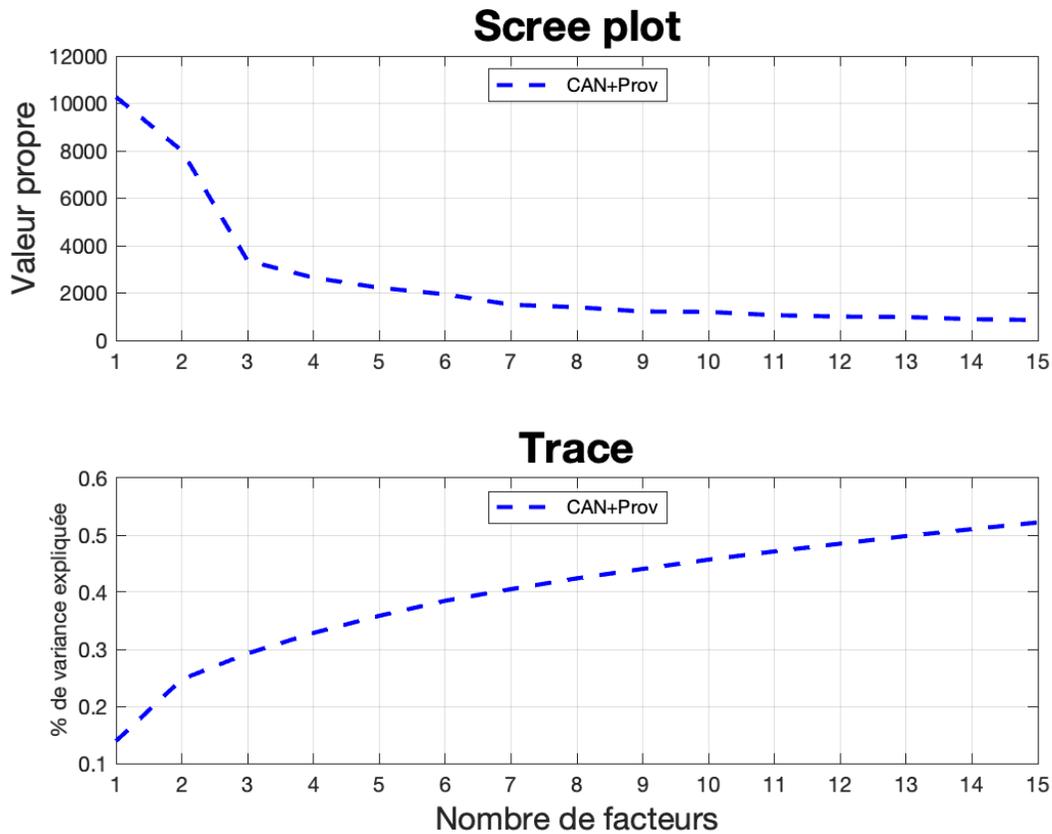
### 3.3 Extraction des facteurs communs comme variables explicatives

Dans la présente section, nous procédons à un examen détaillé du processus d'extraction des facteurs communs, ainsi que de leur utilisation en tant que variables explicatives avec les différentes méthodes de classification supervisée envisagées. Nous commençons par présenter les tests visuels : (i) "scree plot", les valeurs propres, en ordre décroissant, de la matrice de corrélation des données ; (ii) "trace" représentant le  $R^2$  de  $k$  facteurs subséquents. Ensuite, nous procédons à l'estimation du nombre de facteurs à l'aide des critères d'information et des test formels.

#### 3.3.1 Analyse des diagrammes des valeurs propres et de la variance expliquée

La figure 3.3 montre à la fois le diagramme des valeurs propres de la matrice de corrélation des données (scree plot) et celui de la proportion de variances expliquées par les  $k$  premiers facteurs (trace). Le premier diagramme (partie supérieure) montre sept changements distincts de pente (facteur  $f = 2, 3, 4, 5, 6, 7$  et  $8$ ). Au-delà du facteur  $f = 8$ , la pente de la courbe reste constante indiquant qu'il y a sept coudes et, par conséquent, sept facteurs principaux. Quant au diagramme de trace (partie inférieure), il nous donne une idée approximative de la variance du jeu de données qui est expliquée par les facteurs communs. Il faut noter qu'après avoir atteint 51,05 % de la variance expliquée autour des quatorze premiers facteurs, la courbe de la variance expliquée cumulée devient plutôt plate.

Figure 3.3 - Scree plot et trace



Note : "Scree plot" représente les valeurs propres, en ordre décroissant, de la matrice de corrélations des données. "Trace" représente le  $R^2$  de  $k$  facteurs.

### 3.3.2 Détermination du nombre de facteurs communs

Le tableau 3.2 présente le nombre de facteurs estimés en utilisant notre jeu de données et plusieurs procédures statistiques (voir Mao Takongmo et Stevanovic (2015) pour plus détails sur l'explication et la performance des différents tests et critères d'information) : Les méthodes (BN02) et (ABC) choisissent 5 et 6 facteurs statiques, tandis que le nombre de facteurs dynamiques, estimé par (ON), (AH), (HL), (BN07) et (AW), varie de 2 à 4.

Tableau 3.2 – Estimation du nombre de facteurs statiques et dynamiques

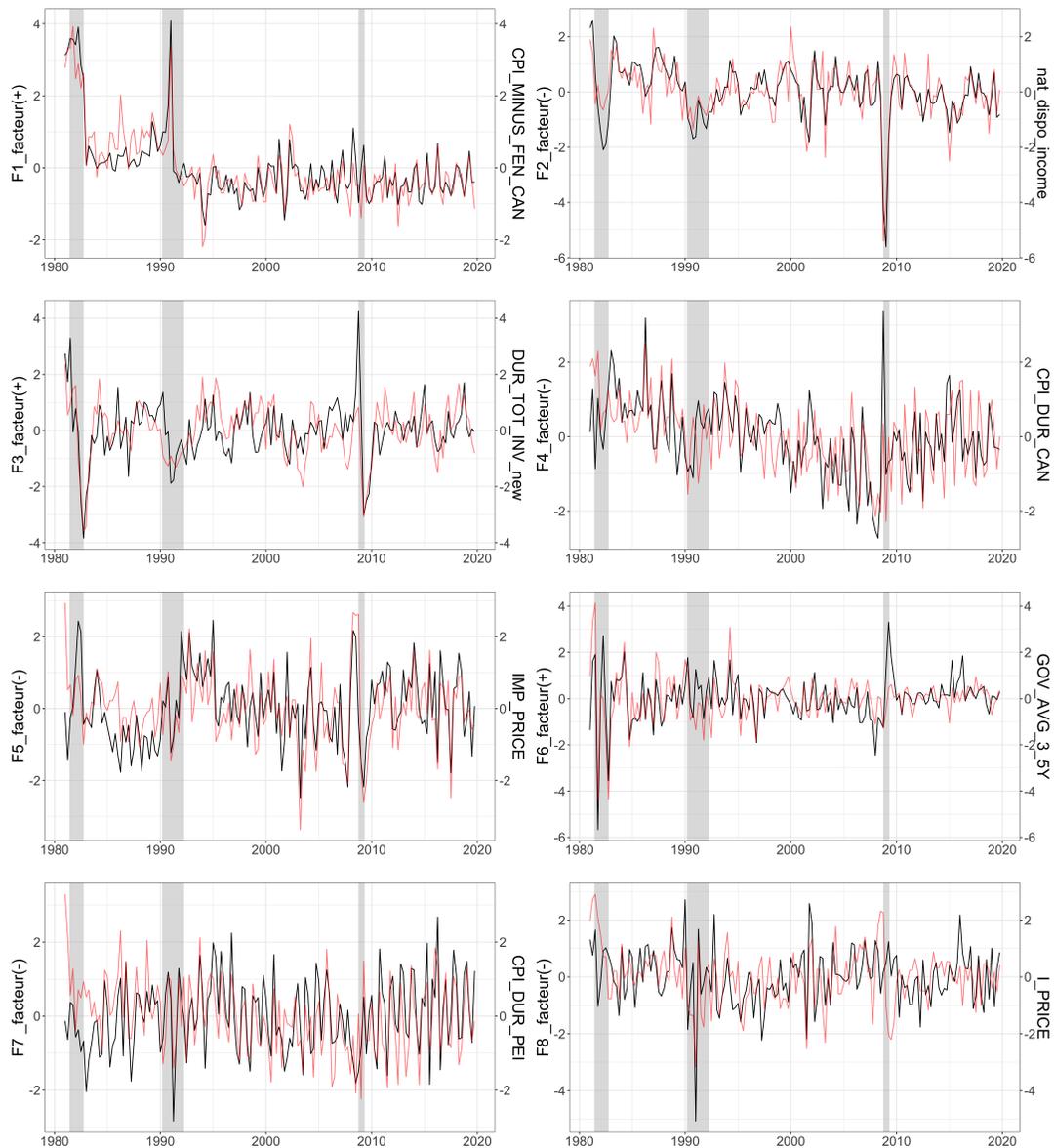
Méthodes	Nombre de facteurs
Bai et Ng (2002) : Critère d'information ICp2 (BN02)	5
Alessi <i>et al.</i> (2010) : Version modifiée de BN02 (ABC)	6
Onatski (2010) (ON)	2
Ahn et Horenstein (2013) : Test du ratio des valeurs propres (AH)	2
Hallin et Liška (2007) (HL)	3
Bai et Ng (2007) (BN07)	4
Amengual et Watson (2007) (AW)	4

Note : Le tableau représente l'estimation du nombre de facteurs statiques et dynamiques à l'aide de différents critères d'information et test.

### 3.3.3 Facteurs communs extraits en tant que variables explicatives

La figure 3.4 montre les huit premiers facteurs communs extraits à partir de notre jeu de données. Nous pouvons constater que le premier facteur suit de près les tendances sous-jacentes de l'inflation au Canada, mesurée par l'indice des prix à la consommation (IPC) excluant les aliments et l'énergie. Cela peut donner une image plus précise des tendances de l'inflation au Canada en éliminant une partie de la volatilité associée à ces deux composantes. La variable la mieux expliquée négativement par le deuxième facteur est le revenu national disponible qui fournit une estimation du revenu moyen disponible pour chaque individu après impôts et autres déductions prises en compte. Le troisième facteur est lié aux inventaires des manufactures (durables), tandis que l'indice des prix à la consommation des biens semble dominer le quatrième facteur. L'indice implicite des prix selon les importations de biens et services est la variable la plus corrélée négativement avec le cinquième facteur. Le sixième facteur est lié aux taux d'intérêt nominaux des obligations négociables de 3 à 5 ans, émises par le gouvernement du Canada et le septième à l'IPC biens durables. Enfin, l'indice implicite des prix selon la formation brute de capital fixe au Canada est à l'origine du huitième facteur.

Figure 3.4 – Évolution des facteurs communs estimés et des séries principales à travers le temps



Note : La figure montre l'évolution au fil du temps des huit facteurs communs estimés avec leurs séries principales les plus corrélées. Les bandes ombrées verticales représentent les périodes de récession au Canada, telles qu'identifiées par Cross et Bergevin (2012) du C.D. Howe Institute. Les facteurs sont représentés en noir et leurs séries principales en rouge. La rotation choisie est indiquée par (+) ou (-).

## CHAPITRE 4

### RÉSULTATS EMPIRIQUES, DISCUSSIONS ET LIMITES MÉTHODOLOGIQUES

Le présent chapitre présente les résultats empiriques hors échantillon obtenus à partir des méthodes de classification supervisée décrites à la section 2.5 et des données présentées au chapitre 3. La section 4.1 se concentre sur les performances prédictives hors échantillon et la section 4.2 propose une discussion de ces résultats. Enfin, la section 4.3 est consacrée à l'analyse des limites méthodologiques et à l'exploration de plusieurs pistes d'amélioration.

#### 4.1 Résultats empiriques

Les résultats empiriques sont présentés dans cette section. Plus précisément, le tableau 4.1 présente les performances hors échantillon des méthodes de classification supervisée sous-jacentes pour chaque horizon de prévision  $h \in \{1, 2, 4, 6, 8\}$ , sur la base des critères de performance définis dans la sous-section 2.4.2.

##### **Volet A : $h = 1$**

Pour les prévisions à  $h = 1$ , les performances des différentes méthodes démontrent une performance relativement modeste. Le classifieur aléatoire (DC), qui constitue la méthode de référence, obtient une exactitude légèrement supérieure à 50 % mais ne présente aucun pouvoir prédictif réel : son rappel, sa précision et son score F1 sont nuls. Sa spécificité de 100 % indique qu'il prédit systématiquement la classe dominante (décélération). L'aire sous la courbe (AUROC) de 50 % confirme son incapacité à discriminer entre les classes.

Les machines à vecteurs de support (SVM) et les arbres de décision (DT) ont obtenu des résultats tous aussi nuls en termes de rappel, de précision et de score F1, avec des AUROC inférieurs à 50 %. Ce qui traduit une mauvaise discrimination entre les périodes d'accélération et de décélération. Par contraste, l'analyse discriminante linéaire (LDA) affiche la meilleure exactitude (65 %), un score F1 élevé (62,16 %) et une valeur AUROC légèrement inférieure à 50 %. Cette performance suggère une certaine capacité de discrimination quant à l'identification des changements directionnels de la croissance du PIB. Les forêts aléatoires (RF) obtiennent un rappel élevé (73,68 %), mais sa faible précision (57,14 %) réduit la confiance à accorder à

ses prédictions positives (accélération). Le classifieur bayésien naïf (NB) obtient une précision remarquable (73,68 %) mais pâtit d'un rappel très bas (36,84 %).

Les performances des différentes méthodes sont globalement satisfaisantes à  $h = 1$ . Les modèles LDA, LR et SVM montrent un bon compromis entre exactitude et rappel, dépassant largement le classifieur aléatoire. Toutefois, leurs AUROC restent faibles (souvent < 60 %), ce qui limite leur pouvoir discriminant. LDA ressort comme la méthode la plus équilibrée, alliant un rappel élevé et une précision modérée.

Tableau 4.1 – Critères de performance des huit méthodes de classification supervisée selon l'horizon

Méthodes	Exactitude	Rappel	Score F1	Précision	Spécificité	AUROC
<b>Volet A : h = 1</b>						
Classifieur aléatoire (DC)	52.500	0.000	0.000	0.000	100.000	50.000
Forêts aléatoires (RF)	61.250	73.684	64.368	57.143	50.000	37.625
Machines à vecteurs de support (SVM)	52.500	0.000	0.000	0.000	100.000	29.073
Régression logistique (LR)	55.000	18.421	28.000	58.333	88.095	39.662
Arbres de décision (DT)	52.500	0.000	0.000	0.000	100.000	30.420
analyse discriminante linéaire (LDA)	65.000	60.526	62.162	63.889	69.048	46.303
K-plus proches voisins (KNN)	57.500	55.263	55.263	55.263	59.524	44.518
Classifieur bayésien naïf (NB)	63.750	36.842	49.123	73.684	88.095	40.100
<b>Volet B : h = 2</b>						
Classifieur aléatoire (DC)	53.165	0.000	0.000	0.000	100.000	50.000
Forêts aléatoires (RF)	55.696	54.054	53.333	52.632	57.143	48.938
Machines à vecteurs de support (SVM)	72.152	70.270	70.270	70.270	73.810	69.820
Régression logistique (LR)	54.430	72.973	60.000	50.943	38.095	45.560
Arbres de décision (DT)	53.165	0.000	0.000	0.000	100.000	28.636
analyse discriminante linéaire (LDA)	63.291	75.676	65.882	58.333	52.381	50.837
K-plus proches voisins (KNN)	58.228	94.595	67.961	53.030	26.190	48.198
Classifieur bayésien naïf (NB)	60.759	45.946	52.308	60.714	73.810	45.045
<b>Volet C : h = 4</b>						
Classifieur aléatoire (DC)	51.948	0.000	0.000	0.000	100.000	50.000

(suite page suivante)

(Tableau 4.1 suite) Critères de performance des huit méthodes de classification supervisée selon l'horizon

Modèles	Exactitude	Rappel	Score F1	Précision	Spécificité	AUROC
Forêts aléatoires (RF)	57.143	18.919	29.787	70.000	92.500	39.324
Machines à vecteurs de support (SVM)	55.844	43.243	48.485	55.172	67.500	30.405
Régression logistique (LR)	4.935	86.486	70.330	59.259	45.000	57.838
Arbres de décision (DT)	61.039	40.541	50.000	65.217	80.000	38.345
analyse discriminante linéaire (LDA)	62.338	100.000	71.845	56.061	27.500	44.527
K-plus proches voisins (KNN)	59.740	56.757	57.534	58.333	62.500	51.824
Classifieur bayésien naïf (NB)	59.740	54.054	56.338	58.824	65.000	46.284
<b>Volet D : h = 6</b>						
Classifieur aléatoire (DC)	50.667	0.000	0.000	0.000	100.000	50.000
Forêts aléatoires (RF)	54.667	13.514	22.727	71.429	94.737	40.363
Machines à vecteurs de support (SVM)	57.333	62.162	58.974	56.098	52.632	31.010
Régression logistique (LR)	60.000	70.270	63.415	57.778	50.000	53.129
Arbres de décision (DT)	50.667	0.000	0.000	0.000	100.000	31.366
analyse discriminante linéaire (LDA)	60.000	81.081	66.667	56.604	39.474	47.511
K-plus proches voisins (KNN)	56.000	81.081	64.516	53.571	31.579	42.354
Classifieur bayésien naïf (NB)	53.333	5.405	10.256	100.000	100.000	35.349
<b>Volet E : h = 8</b>						
Classifieur aléatoire (DC)	50.690	0.000	0.000	0.000	100.000	50.000
Forêts aléatoires (RF)	53.425	22.222	32.000	57.143	83.784	42.492
Machines à vecteurs de support (SVM)	53.425	72.222	60.465	52.000	35.135	33.559
Régression logistique (LR)	63.014	63.889	63.014	62.162	62.162	48.198
Arbres de décision (DT)	54.795	11.111	19.512	80.000	97.297	40.315
analyse discriminante linéaire (LDA)	54.795	27.778	37.736	58.824	81.081	42.568
K-plus proches voisins (KNN)	56.164	58.333	56.757	55.263	54.054	41.817
Classifieur bayésien naïf (NB)	60.274	69.444	63.291	58.140	51.351	44.670

Note : Le tableau présente les critères de performance (exactitude, rappel, score F1, précision, spécificité et AUROC) obtenus à partir du classifieur aléatoire (DC), des forêts aléatoires (RF), des machines à vecteurs de support (SVM), de la régression logistique (LR), des arbres de décision (DT), de l'analyse discriminante linéaire (LDA), des K-plus proches voisins (KNN) et du classifieur bayésien naïf (NB) à différents horizons de prévision. Les volets (A) à (E) correspondent respectivement aux prévisions  $h \in \{1, 2, 4, 6, 8\}$ .

**Volet B :  $h = 2$** 

À l'horizon  $h = 2$ , les performances s'améliorent pour certaines méthodes. Notamment, les machines à vecteurs de support (SVM) montrent une nette supériorité avec une exactitude élevée (72,15 %), un score F1 robuste (70,27 %) et surtout la meilleure AUROC (69,82 %). Ces résultats soulignent sa capacité à discriminer efficacement entre les périodes d'accélération et de décélération. L'analyse discriminante linéaire (LDA) et les K-plus proches voisins (KNN) obtiennent aussi des rappels élevés (75,68 % et 94,60 %, respectivement), mais ils présentent des précisions modérées et des valeurs AUROC proches de 50 %. Il convient de noter que la spécificité des K-plus proches voisins (KNN) tombe à 26,2 %, révélant une tendance à prédire excessivement les accélérations.

La régression logistique (LR) présente un rappel élevé (72,97 %) (capacité à détecter les accélérations réelles) alors que son faible pouvoir discriminant (AUROC = 45,56 %) limite l'intérêt opérationnel de ses prévisions. Les forêts aléatoires (RF) se révèlent globalement peu performantes sur tous les critères considérés : exactitude (55,70 %), rappel (54,05 %), précision (52,63 %) et score F1 (53,33 %). Le classifieur bayésien naïf (NB) présente une meilleure spécificité (73,81 %) mais un faible rappel (45,95 %), le rendant utile dans un cadre où éviter les faux positifs est plus important que détecter systématiquement toutes les accélérations.

Les résultats sont mitigés à cet horizon. SVM affiche des performances supérieures en exactitude, mais avec un rappel parfois faible. La méthode RF commence à se démarquer légèrement, tandis que le classifieur NB devient instable. L'écart se creuse entre la précision et le rappel, montrant la difficulté de capturer simultanément les signaux d'accélération et de décélération.

**Volet C :  $h = 4$** 

À l'horizon  $h = 4$ , la régression logistique (LR) et l'analyse discriminante linéaire (LDA) atteignent des performances élevées en termes de rappel (86,49 % et 100 %, respectivement). Toutefois, LDA présente une très faible spécificité (27,5 %) et une AUROC inférieure à 50 % (44,53 %). Ce qui limite fortement sa performance globale. LR ressort clairement du lot avec le meilleur compromis exactitude élevée (64,94 %), un score F1 élevé (70,33 %) et l'AUROC maximale (57,84 %). Ces résultats montrent une amélioration notable par rapport aux horizons précédents. Les méthodes telles que les forêts aléatoires (RF), les machines à vecteurs de support (SVM) et les arbres de décision (DT) montrent des performances faibles sur la majorité des critères, particulièrement en termes d'AUROC. Leur principale difficulté réside dans leur faible rappel (RF

(18,92 %), SVM (43,24 %) et DT (40,54 %) et leurs valeurs AUROC particulièrement basses (RF (39,32 %), SVM (30,41 %) et DT (38,35 %)), signalant des limites importantes dans leur pouvoir discriminant. Si nous souhaitons privilégier une limitation maximale des faux positifs, RF et DT montrent des spécificités élevées (RF (92,50 %) et DT (80,00 %)), mais au prix de faibles rappels (RF (18,92 %) et DT (40,54 %)).

Les performances se détériorent pour l'ensemble des méthodes. LR et LDA conservent une tendance générale stable, mais les valeurs AUROC descendent vers des niveaux critiques. Le classifieur aléatoire devient une méthode crédible, traduisant une perte de pouvoir discriminant des autres méthodes. La complexité croissante de la dynamique macroéconomique rend l'apprentissage plus difficile.

#### **Volet D : $h = 6$**

À mesure que l'horizon s'étend à  $h = 6$ , les performances des différentes méthodes baissent globalement. L'analyse discriminante linéaire (LDA) et la régression logistique (LR) obtiennent toutefois des rappels élevés (81,08 % et 70,27 %, respectivement) et des scores F1 relativement bons (66,67 % et 63,42 %, respectivement). LR obtient la meilleure AUROC (53,13 %), bien que modérée. Les méthodes les arbres de décision (DT), les forêts aléatoires (RF) et le classifieur bayésien naïf (NB) affichent des performances en termes de rappel nul ou faible (0,00 %, 13,51 % et 5,41 %, respectivement) et des exactitudes (50,67 %, 54,67 % et 53,33 %, respectivement) presque proches de celles du classifieur aléatoire (50,67 %). Ce qui les empêche d'avoir toute capacité de capturer les périodes d'accélération. Par ailleurs, les valeurs élevées de précision (80,00 %, 71,43 %, 100 %, respectivement) et spécificité (97,30 %, 94,74 %, 100 %, respectivement) sont trompeuses ici, car elles reflètent surtout une forte inertie de ces méthodes à prédire la classe négative (décélération). Enfin, les scores F1 nuls ou faibles (0,00 %, 22,73 %, 10,26 %, respectivement) ainsi que les valeurs AUROC en dessous de 50 % soulignent leur pouvoir discriminant très limité.

La détérioration est confirmée. Si LDA et DT montrent des signes de robustesse, leurs scores restent modestes. Le taux de faux positifs augmente significativement, nuisant ainsi à la précision. La robustesse des performances est fortement remise en question.

#### **Volet E : $h = 8$**

Enfin, pour l'horizon  $h = 8$ , l'érosion de la performance prédictive est encore plus notable. La régression logistique (LR) et le classifieur bayésien naïf (NB) présentent néanmoins les meilleurs résultats. LR atteint la

meilleure exactitude (63,01 %) et obtient un équilibre appréciable entre rappel et précision (autour de 63 %), avec une valeur AUROC proche de 50 % (48,20 %), ce qui reste modeste. Les autres méthodes telles que les forêts aléatoires (RF), les machines à vecteurs de support (SVM), les arbres de décision (DT), l'analyse discriminante linéaire (LDA) et les K-plus proches voisins (KNN) ne dépassent pas le seuil de performance attendu pour une prévision directionnelle robuste. Elles obtiennent des valeurs AUROC inférieures à 50 %, signe d'une discrimination des classes inférieure à celle d'un classifieur aléatoire, malgré les taux d'exactitude respectifs de 53,43 %, 53,43 %, 54,80 %, 54,80 % et 56,16 %. Par ailleurs, les méthodes RF, DT et LDA qui sont réputées pour leur capacité de détection des périodes de décélération avec des taux de spécificité élevés (83,78 %, 97,30 % et 81,08 %, respectivement), présentent simultanément une détection déficiente des périodes d'accélération, avec des rappels très bas (22,22 %, 11,11 % et 27,78 %, respectivement). En outre, les méthodes SVM et KNN qui détecte davantage de périodes d'accélération (avec un taux de rappel élevé 72,22 % et 58,33 %, respectivement), présentent une baisse de la spécificité (35,14 % et 54,05 %, respectivement), ainsi que du pouvoir discriminant faible (AUROC très basse). Dans ce contexte, la méthode KNN ressort comme la plus équilibrée, bien que sa performance reste très limitée. Les résultats sont les plus faibles. La majorité des méthodes affiche une AUROC inférieure ou égale à 50 %, ce qui équivaut à une absence totale de pouvoir discriminant. La performance est dominée par le hasard. Ce constat souligne la difficulté intrinsèque de la prévision directionnelle à cet horizon, même avec des méthodes puissantes.

Dans l'ensemble, les résultats empiriques hors échantillon révèlent une dépendance des performances des différentes méthodes à l'horizon de prévision. En effet, aux horizons ( $h = 1$  et  $h = 2$ ), certaines méthodes telles que l'analyse discriminante linéaire (LDA), les machines à vecteurs de support (SVM) et la régression logistique (LR) ont démontré des performances encourageantes, traduisant une certaine capacité à détecter les changements directionnels de la croissance du PIB. Cependant, il est à noter que, même à ces horizons, leur pouvoir discriminant (AUROC) reste faible. À mesure que l'horizon de prévision s'allonge ( $h = 4, 6, 8$ ), les performances des différentes méthodes se dégradent progressivement. Les rappels subissent une baisse, les valeurs AUROC descendent en dessous du seuil de discrimination utile, et les méthodes peinent à maintenir un équilibre satisfaisant entre la détection des accélérations et l'évitement des faux positifs. Les méthodes telles que les forêts aléatoires (RF), les arbres de décision (DT), l'analyse discriminante linéaire (LDA) et classifieur bayésien naïf (NB) deviennent inopérantes tandis que la régression logistique (LR) et les K-plus proches voisins (KNN) affichent des performances limitées mais plus équilibrées.

Aucune méthode de classification supervisée ne s'impose comme supérieure de manière robuste sur l'en-

semble des horizons et critères de performance. Ce qui met en évidence leurs limites face à la complexité de la prévision du changement directionnel de la croissance du PIB, en particulier sur plusieurs horizons. Ces observations invitent à une discussion critique sur la portée et les limites de ces méthodes, ainsi qu'à l'exploration de pistes d'amélioration méthodologiques.

## 4.2 Discussions

Les résultats obtenus révèlent des performances contrastées entre les méthodes de classification supervisée, qui sont significativement influencées par l'horizon de prévision et les critères de performance utilisés.

Trois grandes tendances se dégagent :

- À court terme ( $h = 1$  et  $h = 2$ ), les différentes méthodes obtiennent des résultats relativement satisfaisants, bien que perfectibles. À  $h = 1$ , l'analyse discriminante linéaire (LDA) se distingue par un équilibre appréciable entre l'exactitude, le rappel et le score F1, même si son AUROC reste proche de 50 %. À  $h = 2$ , les performances s'améliorent sensiblement pour certaines méthodes. Les machines à vecteurs de support (SVM) combine un bon rappel, un score F1 élevé et le meilleur AUROC, indiquant une réelle capacité à détecter efficacement les changements directionnels. La régression logistique (LR) montre également un rappel élevé, mais sa faible spécificité limite la fiabilité de la détection des signaux du changement directionnel.
- À moyen terme ( $h = 4$ ), les performances déclinent globalement. La régression logistique (LR) continue de présenter un compromis acceptable entre les critères de performance, tandis que des méthodes telles que les forêts aléatoires (RF), les arbres de décision (DT) et les machines à vecteurs de support (SVM) souffrent de rappels faibles et des valeurs AUROC bien inférieures à 50 %, ce qui révèle leur pouvoir prédictif limité à cet horizon.
- À long terme ( $h = 6$  et  $h = 8$ ), l'érosion des performances devient marquée. À  $h = 6$ , plusieurs méthodes échouent à détecter les accélérations (rappels proches de 0 %), avec des scores F1 très faibles et des AUROC sous le seuil de 50 %. À  $h = 8$ , deux groupes de méthodes émergent : d'un côté, les forêts aléatoires (RF), les arbres de décision (DT) et l'analyse discriminante linéaire (LDA) qui privilégient la prédiction de décélération (spécificité élevée, rappel faible), et de l'autre, les machines à vecteurs de support (SVM) et les K-plus proches voisins (KNN) qui identifient plus d'accélération au prix d'une forte baisse de la spécificité et de l'AUROC.

Ces résultats montrent une variabilité des performances selon les méthodes de classification et les horizons de prévision. À court terme ( $h = 1$  et  $h = 2$ ), certains modèles comme LDA, SVM et LR obtiennent de

bons résultats en termes de rappel, traduisant une capacité à détecter les accélérations de la croissance du PIB. Toutefois, leur pouvoir discriminant, mesuré par l'AUROC, reste modeste, voire inférieur à celle d'un classifieur aléatoire dans plusieurs cas. Aux horizons plus longs ( $h = 4, 6, 8$ ), la dégradation de l'ensemble des critères de performance, notamment le rappel et l'AUROC, est manifeste. Aucune méthode ne parvient à maintenir une performance stable. Les arbres de décision (DT) et forêts aléatoires (RF), pourtant robustes en général, voient leurs performances chuter significativement, suggérant une difficulté à capter la dynamique économique à long terme. Deux conclusions majeures peuvent être tirées :

- la performance des différentes méthodes varie selon l'horizon. Certaines méthodes (par exemple, LR ou KNN) s'avèrent compétitives à certains horizons mais perdent leur supériorité sur d'autres, ce qui souligne l'importance de l'adaptabilité en fonction de l'horizon de prévision ;
- les faibles valeurs de l'aire sous la courbe ROC révèlent l'incapacité de certaines méthodes à discriminer les périodes d'accélération et de décélération de la croissance du PIB.

Ces conclusions appellent à une réflexion plus large sur les limites méthodologiques de notre mémoire et les pistes d'amélioration possibles. Ces éléments sont abordés dans la section suivante.

### 4.3 Limites méthodologiques et pistes d'amélioration

Une analyse approfondie des performances hors échantillon des différentes méthodes de classification supervisée pour la prévision du changement directionnel de la croissance du PIB québécois a révélé plusieurs limites, tant sur le plan méthodologique que sur celui des résultats empiriques. Ces limites invitent à explorer des voies d'amélioration concrètes.

#### 4.3.1 Limites méthodologiques

##### **Pouvoir discriminant limité des différentes méthodes**

L'un des constats les plus marquants de cette étude réside dans les faibles valeurs observées pour l'aire sous la courbe ROC, souvent inférieures ou proches de 50 %. Cela suggère que, malgré une exactitude parfois correcte, les différentes méthodes étudiées ne parviennent pas à discriminer de manière efficace les périodes d'accélération et de décélération par rapport à un classifieur aléatoire. Cette faiblesse peut trouver son origine dans divers facteurs :

- *pertinence limitée des variables explicatives* :

Goulet Coulombe *et al.* (2020) ont observé que les modèles d'apprentissage automatique bien calibrés

peuvent améliorer les prévisions directionnelles sur des variables macroéconomiques canadiennes, mais ces gains sont fortement dépendants du type de méthode et du prétraitement (réduction de dimension par analyse en composantes principales), et de l'horizon. Dans notre étude, malgré l'utilisation d'un modèle à facteurs pour extraire des facteurs communs en tant que variables explicatives, nos résultats ne reproduisent pas entièrement ces gains, ce qui suggère que l'information pertinente sur le changement directionnel de la croissance du PIB n'a pas pu être correctement capturée par les composantes principales seules.

— *taille restreinte de l'échantillon :*

L'horizon trimestriel et la limitation du nombre d'observations disponibles réduisent le signal exploitable pour l'apprentissage. Van Ruth (2014) montrent que l'efficacité des modèles dépend fortement de la taille de l'échantillon et du type de données. Nos 157 observations trimestrielles représentent un échantillon restreint (voir chapitre 3 pour nos méthodes de classification supervisée, ce qui pourrait expliquer des résultats peu robustes en terme d'aire sous la courbe ROC.

— *définition du changement directionnel :*

La présente étude s'appuie sur une définition du changement directionnel qui s'aligne sur les approches de Cornec et Mikol (2011). Cette définition se focalise sur le taux de croissance du PIB et s'attache à identifier les périodes d'accélération ( $y_t > y_{t-1}$ ) et de décélération ( $y_t \leq y_{t-1}$ ) du PIB. Cependant, ce cadre de classification ne tient pas compte de la magnitude du changement. Ce qui peut affecter les capacités discriminantes, notamment en ce qui concerne la valeur AUROC. En effet, les méthodes de classification supervisée sont souvent en difficulté pour saisir des signaux ambigus dans les cas où les variations du taux de croissance sont marginales. Empiriquement, la croissance du PIB n'est jamais régulière : l'économie est en général tantôt en phase d'accélération, tantôt en phase de décélération, tantôt en phase de stabilité qui est souvent très courte mais ne reste jamais longtemps (Adanero-Donderis *et al.*, 2009).

### **Mauvaise calibration des hyperparamètres des différentes méthodes**

La calibration des hyperparamètres peut ne pas être optimale, malgré l'usage de la validation croisée et d'un algorithme génétique pour la recherche d'hyperparamètres. Nous pensons qu'il est possible que certaines méthodes de classification supervisée aient été sous-optimisées, notamment ceux sensibles à la configuration des paramètres (par exemple, SVM ou KNN). De plus, les grilles de recherche utilisées ont pu être trop restreintes ou mal adaptées à la complexité du problème. Cette hypothèse est d'autant plus plausible que certaines méthodes présentent une grande variabilité dans leurs performances selon les horizons. Par

ailleurs, notre manière d'utilisation de la validation croisée avec *GASearchCV*, en configurant la fonction K-fold de Python avec l'argument *shuffle = False* qui préserve les dépendances temporelles, est susceptible de provoquer une fuite d'informations au fil du temps.

### **Modification dans la sélection des variables ou des horizons de prévision**

Les résultats obtenus sont fondés sur une sélection de facteurs communs extraits de la base de données LCDMA à partir d'un modèle à facteurs, sans pour autant que des tests de robustesse n'aient été effectués (par exemple, choix exclusif d'indicateurs avancés du PIB). De même, les horizons de prévision choisis  $h \in \{1, 2, 4, 6, 8\}$  n'ont pas fait l'objet d'une analyse de sensibilité, alors qu'un changement de ces horizons pourrait influencer sur la capacité prédictive des différentes méthodes.

### **Pertinence du cadre méthodologique**

La faible valeur des critères de performance observées pourrait s'expliquer en partie par la pertinence du cadre méthodologique utilisé. Bien que l'approche de classification soit adaptée à la prédiction du changement directionnel, la stratégie de la fenêtre glissante qui permet de simuler une prévision en temps réel, pourrait réduire la quantité de données disponibles pendant la période d'entraînement (in-sample). Avec moins de données pour s'entraîner, il devient plus compliqué pour les différentes méthodes de repérer des tendances fiables, surtout si les signaux de changement directionnel dans les données sont faibles.

#### **4.3.2 Pistes d'amélioration**

Compte tenu des limites méthodologiques identifiées, plusieurs pistes d'amélioration peuvent être explorées afin de renforcer la capacité prédictive des différentes méthodes de classification supervisée utilisées et la robustesse des résultats empiriques.

### **Modification de l'ensemble d'informations**

Il est tout-à-faire possible d'améliorer le pouvoir prédictif des différentes méthodes de classification supervisée dans cette étude en modifiant l'ensemble d'informations. Une approche consisterait par exemple à utiliser des indicateurs avancés liés à l'activité économique ou des données alternatives à haute fréquence (données Google) pour capter les signaux du changement directionnel de la croissance du PIB. Une autre approche pourrait résider dans l'emploi de techniques d'apprentissage automatique, à même d'effectuer

une sélection efficace des variables et d'utiliser les variables prédictives sélectionnées comme intrants dans nos différentes méthodes de classification supervisée sous-jacentes. À titre d'exemple, la méthode LASSO pourrait être employée pour la sélection de variables. Entre autre, l'exploitation de données issues d'enquêtes de conjoncture pourrait s'avérer pertinente. Selon Hansson *et al.* (2005), les enquêtes de conjoncture présentent un caractère coïncident, en ce sens qu'elles reflètent la situation du trimestre en cours. Elles présentent également un caractère prospectif, en ce sens qu'elles anticipent également les tendances du trimestre suivant.

### **Amélioration du cadre méthodologique**

Eu égard à la taille restreinte de l'échantillon, il serait pertinent d'envisager une division de notre jeu de données qui intègre de manière périodique des données plus récentes. À cet effet, l'approche de la fenêtre expansive, qui consiste à déplacer progressivement les échantillons d'apprentissage et de validation dans le temps tout en les augmentant progressivement, apparaît comme une méthode plus pertinente. Bien que cette approche nécessite des calculs plus importants que celle de la fenêtre glissante, elle permet néanmoins de prendre en compte l'intégralité de l'historique des données. Par ailleurs, l'utilisation de méthodes plus élaborées pourrait être envisagée telles que :

- des méthodes séquentielles comme les réseaux de neurones récurrents (ou RNN pour Recurrent Neural Networks) qui permettent de modéliser les dépendances temporelles et les effets de mémoire économique ;
- des méthodes probabilistes bayésiennes ou à régimes multiples (par exemple, les modèles à changement de régime markoviens) pour capturer les effets non linéaires ou de rupture dans la dynamique économique.

Enfin, il existe d'autres approches pour aborder la question de la prévision du changement directionnel. Selon Skabar (2008), la prévision du changement directionnel peut être considérée, comme un problème de régression impliquant l'étape supplémentaire de comparaison de la valeur prévue à celle actuelle : si la valeur prévue est supérieure à la valeur actuelle de la série, on prédit une hausse ; sinon, on prédit une baisse.

### **Optimisation plus fine des hyperparamètres**

L'optimisation des hyperparamètres pourrait également être affinée à l'aide de grilles de recherche plus étendues, adaptatives ou basées sur d'autres approches de recherche des hyperparamètres optimaux (la

recherche par grille, la recherche aléatoire, l'optimisation bayésienne, etc...). Par ailleurs, il conviendrait notamment de calibrer différemment les méthodes selon l'horizon de prévision, étant donné que les configurations optimales peuvent fortement varier d'un horizon à l'autre. L'introduction de critères de performance centrés sur l'aire sous la courbe ROC ou le score F1 (au lieu de l'exactitude) pourrait aussi mieux guider la recherche de configurations efficaces.

Dans le cadre de notre mémoire, l'emploi de la validation croisée K-fold avec l'argument *shuffle=False* a été motivé par la volonté de préserver la structure temporelle des données. Toutefois, le maintien de l'ordre temporel n'est pas strictement nécessaire et peut, selon Kelly *et al.* (2023) cité par Xu et Liu (2024), conduire à une utilisation inefficace des données pour la sélection des modèles. Pour y remédier, il est possible d'utiliser le schéma de validation croisée standard (de type K-fold) qui est la meilleure pratique de sélection des hyperparamètres (Goulet Coulombe *et al.* (2020))

### **Réflexion sur la définition de la variable cible**

La définition binaire du changement directionnel fondée uniquement sur le signe de la variation du taux de croissance du PIB pourrait être enrichie par une approche tenant compte de la magnitude des variations. Une alternative serait de recourir à une approche permettant d'affiner la pertinence du signal de classification du changement directionnel de la croissance du PIB. Par exemple, pour quantifier les accélérations et les décélérations de la croissance du PIB, nous introduisons, en suivant Spelta *et al.* (2025), deux indices binaires : une décélération serait définie comme une diminution de la variation (en pourcentage) en dessous du 10<sup>ème</sup> centile de sa distribution tandis qu'une accélération serait caractérisée par une augmentation de la variation (en pourcentage) au-dessus du 90<sup>ème</sup> centile de sa distribution. Ainsi, ces définitions permettent de cibler les retournements significatifs de l'activité économique, en écartant les fluctuations mineures qui pourraient être liées à des effets transitoires ou à du bruit statistique.

### **Analyse de la sensibilité et contrôle de robustesse**

Une analyse de la sensibilité et un contrôle de robustesse des résultats aux choix de variables explicatives, d'horizons de prévision ou d'utilisation de la base de données sans passer par un modèle à facteurs préalable, permettraient de mieux cerner les configurations les plus efficaces et d'évaluer la stabilité de nos conclusions.

## CONCLUSION

L'objectif de ce mémoire était d'utiliser des méthodes de classification supervisée afin de prédire le changement directionnel de la croissance du PIB québécois sur plusieurs horizons, et d'évaluer leurs performances. Dans cet exercice, au-delà de l'utilisation du modèle à facteurs pour réduire la dimension de notre jeu de données, nous avons implémenté plusieurs méthodes populaires de classification supervisée. Il s'agit notamment, du classifieur aléatoire (DC), du classifieur bayésien naïf (NB), des K-plus proches voisins (k-NN), de la machine à vecteurs de support (SVM), du modèle de régression logistique (LR) et la forêt aléatoire (RF), les arbres de décision (DT) et l'analyse discriminante linéaire (LDA). Ces différentes méthodes ont été exécutées selon plusieurs horizons de prévision  $h \in \{1, 2, 4, 6, 8\}$ . De plus, l'algorithme génétique a été utilisé pour rechercher les hyperparamètres optimaux de ces méthodes.

Les résultats empiriques hors échantillon montrent que les performances des différentes méthodes sous-jacentes sont fortement liées à l'horizon de prévision. À court terme ( $h = 1$  et  $h = 2$ ), certaines méthodes comme LDA, SVM et LR obtiennent des performances relativement satisfaisantes, bien que leur pouvoir discriminant mesuré par l'AUROC demeure faible. À mesure que l'horizon s'allonge ( $h = 4, 6, 8$ ), les performances des différentes méthodes se dégradent : le rappel diminue, la valeur de l'aire sous la courbe ROC (AUROC) tombe souvent sous le seuil de 50 % et aucune méthode ne parvient à maintenir un bon compromis entre détection des signaux d'accélération et réduction des faux positifs. En définitive, aucune méthode ne s'impose de façon robuste sur tous les horizons. Cette variabilité des performances met en évidence deux enseignements majeurs :

- la nécessité d'adapter la méthode de classification selon l'horizon de prévision visé : aucune méthode ne domine systématiquement ;
- les faibles valeurs de l'aire sous la courbe ROC révèlent l'incapacité de certaines méthodes à discriminer les périodes d'accélération et de décélération de la croissance du PIB.

Ces enseignements mettent en lumière plusieurs limites méthodologiques : (i) la définition du changement directionnel ne prend pas en compte la magnitude des variations, (ii) la taille restreinte du jeu de données limite la capacité d'apprentissage des différentes méthodes de classification supervisée, (iii) les hyperparamètres n'ont peut-être pas été calibrés de manière optimale, et (iv) le schéma de validation utilisé (K-fold avec *shuffle=False*) avec GASearchCV est susceptible d'induire une fuite d'informations. Ces limites expliquent en partie les faibles valeurs des critères de performances et appellent à une amélioration du cadre méthodologique.

Parmi les pistes d'amélioration envisagées figurent : l'utilisation d'indicateurs avancés ou d'enquêtes de conjoncture, l'intégration de données alternatives à haute fréquence, l'emploi de méthodes sélectionnant automatiquement les variables pertinentes (par exemple, le LASSO), l'implémentation de méthodes séquentielles ou non linéaires (RNN, modèles à changement de régime), ainsi que la recherche des hyperparamètres optimaux à l'aide d'autres approches telles que la recherche en grille ou l'optimisation bayésienne. De plus, une redéfinition de la variable cible pourrait être envisagée : par exemple, en suivant Spelta *et al.* (2025), il serait possible de distinguer uniquement les phases d'accélération et de décélération, via les 10<sup>e</sup> et 90<sup>e</sup> percentiles de la distribution des variations du taux de croissance du PIB. Enfin, il existe d'autres approches pour aborder la question de la prévision du changement directionnel. Selon Skabar (2008), la prévision du changement directionnel peut être considérée, comme un simple problème de régression impliquant l'étape supplémentaire de comparaison de la valeur prévue à celle actuelle : si la valeur prévue est supérieure à la valeur actuelle de la série, on prédit une hausse ; sinon, on prédit une baisse. Ces pistes ouvrent la voie à des recherches ultérieures visant à mieux comprendre et anticiper la dynamique de la croissance du PIB.

En définitive, ce mémoire enrichit la littérature sur les données massives et l'apprentissage automatique avec applications en économie grâce à une évaluation comparative sur un cas peu étudié : celui du Québec. Il présente également des pistes concrètes pour améliorer les prévisions macroéconomiques.

## ANNEXE A

### DONNÉES

Le tableau ci-dessous présente les données utilisées. Il s'agit de 477 variables canadiennes et provinciales. La colonne "Région" correspond à la provenance de chaque variable. La colonne "code" correspond au code de transformation utilisé pour rendre stationnaires les variables : (1) pas de transformation ; (2) première différence ; (5) première différence des logarithmes ; (6) seconde différence des logarithmes.

Variable	Description	Région	Code
REAL_C	Real Final consumption expenditure, chained (2012) dollars	CAN	5
REAL_C_GOODS	Real Final consumption expenditure, Goods, chained (2012) dollars	CAN	5
REAL_C_DUR_GOODS	Real Final consumption expenditure, Durable goods, chained (2012) dollars	CAN	5
REAL_C_SERV	Real Final consumption expenditure, Services, chained (2012) dollars	CAN	5
REAL_C_G	Real General governments final consumption expenditure,, chained (2012) dollars	CAN	5
REAL_I	Real Gross fixed capital formation, chained (2012) dollars	CAN	5
REAL_I_BUS	Real Business gross fixed capital formation, chained (2012) dollars	CAN	5
REAL_I_RES	Real Gross fixed capital formation, Residential structures, chained (2012) dollars	CAN	5
REAL_I_NRES	Real Gross fixed capital formation, Non-residential structures, chained (2012) dollars	CAN	5
REAL_I_ME	Real Gross fixed capital formation, Machinery and equipment, chained (2012) dollars	CAN	5
REAL_I_G	Real Gross fixed capital formation, General governments gross fixed capital formation, chained (2012) dollars	CAN	5
INVEST_INV	Investment in inventories, chained (2012) dollars	CAN	1
REAL_EXP	Real Exports of goods and services, chained (2012) dollars	CAN	5
REAL_EXP_SERV	Real Exports of services, chained (2012) dollars	CAN	5
REAL_IMP	Real Imports of goods and services, chained (2012) dollars	CAN	5
REAL_IMP_SERV	Real Imports of services, chained (2012) dollars	CAN	5
REAL_GDP	Real Gross domestic product , chained (2012) dollars	CAN	5
C_PRICE	Implicit price index : Final consumption expenditure, 2012 = 100	CAN	5
C_PRICE_GOODS	Implicit price index : Final consumption expenditure, Goods, 2012 = 100	CAN	5

<b>Variable</b>	<b>Description</b>	<b>Région</b>	<b>Code</b>
C_PRICE_DUR_GOODS	Implicit price index : Final consumption expenditure, Durable goods, 2012 = 100	CAN	5
C_PRICE_SERV	Implicit price index : Final consumption expenditure, Services, 2012 = 100	CAN	5
G_PRICE	Implicit price index : General governments final consumption expenditure, 2012 = 100	CAN	5
I_PRICE	Implicit price index : Gross fixed capital formation, 2012 = 100	CAN	5
I_PRICE_BUS	Implicit price index : Business gross fixed capital formation, 2012 = 100	CAN	5
I_PRICE_RES	Implicit price index : Gross fixed capital formation, Residential structures, 2012 = 100	CAN	5
I_PRICE_NRES	Implicit price index : Gross fixed capital formation, Non-residential structures, 2012 = 100	CAN	5
I_PRICE_ME	Implicit price index : Gross fixed capital formation, Machinery and equipment, 2012 = 100	CAN	5
I_PRICE_G	Implicit price index : Gross fixed capital formation, General governments gross fixed capital formation, 2012 = 100	CAN	5
EXP_PRICE	Implicit price index : Exports of goods and services, 2012 = 100	CAN	5
EXP_PRICE_SERV	Implicit price index : Exports of services, 2012 = 100	CAN	5
IMP_PRICE	Implicit price index : Imports of goods and services, 2012 = 100	CAN	5
IMP_PRICE_SERV	Implicit price index : Imports of services, 2012 = 100	CAN	5
GDP_PRICE_IND	Implicit price index : Gross domestic product, 2012 = 100	CAN	5
BSI_new	GDP business	CAN	5
GPI_new	GDP goods	CAN	5
SPI_new	GDP services	CAN	5
IP_new	GDP industrial production	CAN	5
NDM_new	GDP non durable goods	CAN	5
DM_new	GDP durables	CAN	5
OILP_new	GDP mining, petrol and gas	CAN	5
CON_new	GDP construction	CAN	5
RT_new	GDP retail trade	CAN	5
WT_new	GDP wholesale trade	CAN	5
PA_new	GDP public administration	CAN	5
FIN_new	GDP finance and insurance	CAN	5
OIL_CAN_new	Crude oil production (Cubic metres)	CAN	5

Variable	Description	Région	Code
indus_capu_tot	Industrial capacity utilization rates, Total industrial	CAN	2
indus_capu_for	Industrial capacity utilization rates, Forestry and logging [113]	CAN	2
indus_capu_mine_oil	Industrial capacity utilization rates, Mining and oil and gas extraction [21]	CAN	2
indus_capu_con	Industrial capacity utilization rates ,Construction [23]	CAN	2
indus_capu_manu	Industrial capacity utilization rates, Manufacturing [31-33]	CAN	2
indus_capu_oil	Industrial capacity utilization rates, Petroleum and coal products manufacturing [324]	CAN	2
indus_capu_chem	Industrial capacity utilization rates, Chemical manufacturing [325]	CAN	2
indus_capu_mach	Industrial capacity utilization rates, Machinery manufacturing [333]	CAN	2
EMP_CAN	Employment total	CAN	5
EMP_SERV_CAN	Employment services	CAN	5
EMP_FOR_OIL_CAN	Employment forestry, fishing, mining, oil and gas	CAN	5
EMP_CONS_CAN	Employment construction	CAN	5
EMP_SALES_CAN	Employment sales (wholesale and retail trade)	CAN	5
EMP_FIN_CAN	Employment finance, insurance and real estate	CAN	5
EMP_MANU_CAN	Employment manufacturing	CAN	5
EMP_PART_CAN	Employment part time	CAN	5
UNEMP_CAN	Unemployment rate	CAN	2
UNEMP_DURA_1.4_CAN	Unemployment duration (1-4 weeks)	CAN	5
UNEMP_DURA_5.13_CAN	Unemployment duration (5-13 weeks)	CAN	5
UNEMP_DURA_14.25_CAN	Unemployment duration (14-24 weeks)	CAN	5
UNEMP_DURA_27+_CAN	Unemployment duration (27+ weeks)	CAN	5
UNEMP_DURAvg_CAN_new	Unemployment average duration	CAN	5
CLAIMS_CAN	Employment insurance initial claims, Allowed	CAN	1
TOT_HRS_CAN	Hours worked total	CAN	5
GOOD_HRS_CAN	Hours worked goods	CAN	5
JOB_BI	Business Index, 2012 = 100, Total number of jobs	CAN	5
HOURS_AV_BI	Business Index, 2012 = 100, Average hours worked	CAN	5
HOURS_TOT_BI	Business Index, 2012 = 100, Hours worked	CAN	5
LPROD_BI	Business Index, 2012 = 100, Labour productivity	CAN	5
COMP_HOUP_BI	Business Index, 2012 = 100, Total compensation per hour worked	CAN	5
U_LCOST_BI	Business Index, 2012 = 100, Unit labour cost	CAN	5
U_NLABOR_PAY_BI	Business Index, 2012 = 100, Unit non-labour payments	CAN	5

<b>Variable</b>	<b>Description</b>	<b>Région</b>	<b>Code</b>
LABOR_SH_BI	Business Index, 2012 = 100, Labour share	CAN	5
COMP_TOT_BI	Business Index, 2012 = 100, Total compensation	CAN	5
NLABOR_PAY_BI	Business Index, 2012 = 100, Non-labour payments	CAN	5
EMP_COMP	Compensation of employees - households	CAN	5
TOT_WAGE_SAL	Wages and salaries	CAN	5
nat_dispo_income	National disposable income	CAN	5
hhold_dispo_income	Households disposable income	CAN	5
corp_dispo_income	Corporations disposable income	CAN	2
nat_saving	National net saving	CAN	2
hhold_saving	Households net saving	CAN	5
corp_saving	Corporations net saving	CAN	2
corp_profit	Corporation profit before taxes	CAN	5
NHOUSE_P_CAN	New housing price index, Total (house and land)	CAN	5
NHOUSE_P_NF	New housing price index, Total (house and land)	NF	5
NHOUSE_P_PEI	New housing price index, Total (house and land)	PEI	5
NHOUSE_P_NS	New housing price index, Total (house and land)	NS	5
NHOUSE_P_NB	New housing price index, Total (house and land)	NB	5
NHOUSE_P_QC	New housing price index, Total (house and land)	QC	5
NHOUSE_P_ONT	New housing price index, Total (house and land)	ONT	5
NHOUSE_P_MAN	New housing price index, Total (house and land)	MAN	5
NHOUSE_P_SAS	New housing price index, Total (house and land)	SAS	5
NHOUSE_P_ALB	New housing price index, Total (house and land)	ALB	5
NHOUSE_P_BC	New housing price index, Total (house and land)	BC	5
hstart_CAN	House Starts (units)	CAN	5
hstart_PEI	House Starts (units)	PEI	2
hstart_NS	House Starts (units)	NS	5
hstart_NB	House Starts (units)	NB	5
hstart_QC	House Starts (units)	QC	5
hstart_ONT	House Starts (units)	ONT	5
hstart_MAN	House Starts (units)	MAN	5
hstart_SAS	House Starts (units)	SAS	5
hstart_ALB	House Starts (units)	ALB	5
hstart_BC	House Starts (units)	BC	5
hstart_NF	House Starts (units)	NF	5

<b>Variable</b>	<b>Description</b>	<b>Région</b>	<b>Code</b>
build_Total_CAN_new	Building permits (tous)	CAN	5
build_Ind_CAN_new	Building permits (industries)	CAN	5
build_Comm_CAN_new	Building permits (commerce)	CAN	5
build_Total_NF_new	Building permits (tous)	NF	5
build_Ind_NF_new	Building permits (industries)	NF	2
build_Comm_NF_new	Building permits (commerce)	NF	5
build_Total_PEI_new	Building permits (tous)	PEI	5
build_Ind_PEI_new	Building permits (industries)	PEI	2
build_Comm_PEI_new	Building permits (commerce)	PEI	5
build_Total_NS_new	Building permits (tous)	NS	5
build_Ind_NS_new	Building permits (industries)	NS	5
build_Comm_NS_new	Building permits (commerce)	NS	5
build_Total_NB_new	Building permits (tous)	NB	5
build_Ind_NB_new	Building permits (industries)	NB	2
build_Comm_NB_new	Building permits (commerce)	NB	5
build_Total_QC_new	Building permits (tous)	QC	5
build_Ind_QC_new	Building permits (industries)	QC	5
build_Comm_QC_new	Building permits (commerce)	QC	5
build_Total_ONT_new	Building permits (tous)	ONT	5
build_Ind_ONT_new	Building permits (industries)	ONT	5
build_Comm_ONT_new	Building permits (commerce)	ONT	5
build_Total_MAN_new	Building permits (tous)	MAN	5
build_Ind_MAN_new	Building permits (industries)	MAN	5
build_Comm_MAN_new	Building permits (commerce)	MAN	5
build_Total_SAS_new	Building permits (tous)	SAS	5
build_Ind_SAS_new	Building permits (industries)	SAS	5
build_Comm_SAS_new	Building permits (commerce)	SAS	5
build_Total_ALB_new	Building permits (tous)	ALB	5
build_Ind_ALB_new	Building permits (industries)	ALB	5
build_Comm_ALB_new	Building permits (commerce)	ALB	5
build_Total_BC_new	Building permits (tous)	BC	5
build_Ind_BC_new	Building permits (industries)	BC	5
build_Comm_BC_new	Building permits (commerce)	BC	5

Variable	Description	Région	Code
BUILD_IND_NRES	Building construction price indexes, 2017 = 100, 11 census metropolitan area, Non-residential buildings [2362]	CAN	5
BUILD_IND_COMM	Building construction price indexes, 2017 = 100, 11 census metropolitan area, Commercial buildings [23622]	CAN	5
BUILD_IND_WARE	Building construction price indexes, 2017 = 100, 11 census metropolitan area, Warehouse	CAN	5
BUILD_IND_INDUS	Building construction price indexes, 2017 = 100, 11 census metropolitan area, Industrial buildings [23621]	CAN	5
BUILD_IND_FAC	Building construction price indexes, 2017 = 100, 11 census metropolitan area, Factory	CAN	5
MANU_N_ORD_new	Manufacturing new orders (total)	CAN	5
MANU_UNFIL_new	Manufacturing unfilled orders (total)	CAN	5
MANU_TOT_INV_new	Manufacturing inventories (total)	CAN	5
MANU_INV_RAT_new	Manufacturing inventories to shipments ratio (total)	CAN	1
N_DUR_INV_RAT_new	Manufacturing inventories to shipments ratio (durables)	CAN	1
DUR_N_ORD_new	Manufacturing new orders (durables)	CAN	5
DUR_UNFIL_new	Manufacturing unfilled orders (durables)	CAN	5
DUR_TOT_INV_new	Manufacturing inventories (durables)	CAN	5
DUR_INV_RAT_new	Manufacturing inventories to shipments ratio (durables)	CAN	1
M3	M3 (gross)	CAN	5
M2p	M2+ (gross)	CAN	5
M_BASE1	Monetary base	CAN	5
CRED_T	Total credit	CAN	5
CRED_HOUS	Household credit	CAN	5
CRED_MORT	Mortgage credit	CAN	5
CRED_CONS	Consumption credit	CAN	5
CRE_BUS	Business credit	CAN	5
BANK_RATE_L	Bank rate	CAN	2
PC_PAPER_1M	Corporate paper rate (1 month)	CAN	2
PC_PAPER_3M	Corporate paper rate (3 months)	CAN	2
GOV_AVG_1_3Y	Governmental bonds (average rate) (1-3 years)	CAN	2
GOV_AVG_3_5Y	Governmental bonds (average rate) (3-5 years)	CAN	2
GOV_AVG_5_10Y	Governmental bonds (average rate) (5-10 years)	CAN	2
GOV_AVG_10pY	Governmental bonds (average rate) (10+ years)	CAN	2

Variable	Description	Région	Code
MORTG_1Y	Mortgage rate (1 year)	CAN	2
MORTG_5Y	Mortgage rate (5 years)	CAN	2
TBILL_3M	Treasury bills (3 months)	CAN	2
TBILL_6M	Treasury bills (6 months)	CAN	2
PC_3M.Bank_rate	Corporate paper rate (3 months) – Bank rate	CAN	1
G_AVG_1.3.Bank_rate	Government bonds (1-3 years) - Bank rate	CAN	1
G_AVG_3.5.Bank_rate	Government bonds (3-5 years) - Bank rate	CAN	1
G_AVG_5.10.Bank_rate	Government bonds (5-10 years) - Bank rate	CAN	1
TBILL_6M.Bank_rate	Treasury bond (6 months) - Bank rate	CAN	1
G_AVG_10p.TBILL_3M	Government Bonds (10+ years) - Treasury Bond (3 months)	CAN	1
RES_TOT	Total Canada's official international reserves	CAN	5
RES_USD	Canadian USD reserves	CAN	5
RES_IMF	Canadian reserve position at the IMF	CAN	5
IOIL_BP_new	Imports oil	CAN	5
EOIL_BP_new	Exports oil	CAN	5
EX_ENER_BP_new	Export energy products	CAN	5
(Sum)	Export energy products	CAN	
(Sum)	Export energy products	CAN	
(Sum)	Export energy products	CAN	
EX_MINER_BP_new	Exports non-metallic ores	CAN	5
(Sum)	Exports non-metallic ores	CAN	
(Sum)	Exports non-metallic ores	CAN	
(Sum)	Exports non-metallic ores	CAN	
(Sum)	Exports non-metallic ores	CAN	
(Sum)	Exports non-metallic ores	CAN	
EX_METAL_BP_new	Exports metal and other mineral products	CAN	5
(Sum)	Exports metal and other mineral products	CAN	
(Sum)	Exports metal and other mineral products	CAN	
(Sum)	Exports metal and other mineral products	CAN	
(Sum)	Exports metal and other mineral products	CAN	
(Sum)	Exports metal and other mineral products	CAN	
(Sum)	Exports metal and other mineral products	CAN	
(Sum)	Exports metal and other mineral products	CAN	
(Sum)	Exports metal and other mineral products	CAN	

<b>Variable</b>	<b>Description</b>	<b>Région</b>	<b>Code</b>
(Sum)	Exports metal and other mineral products	CAN	
(Sum)	Exports metal and other mineral products	CAN	
EX_IND_EQUIP_BP_new	Exports industrial machinery, pieces and equipment	CAN	5
(Sum)	Exports industrial machinery, pieces and equipment	CAN	
(Sum)	Exports industrial machinery, pieces and equipment	CAN	
EX_TRANSP_BP_new	Exports motor vehicules and parts	CAN	5
(Sum)	Exports motor vehicules and parts	CAN	
(Sum)	Exports motor vehicules and parts	CAN	
EX_CONS_BP_new	Exports consumption goods	CAN	5
(Sum)	Exports consumption goods	CAN	
(Sum)	Exports consumption goods	CAN	
IMP_METAL_BP_new	Imports metal and other mineral products	CAN	5
(Sum)	Imports metal and other mineral products	CAN	
(Sum)	Imports metal and other mineral products	CAN	
(Sum)	Imports metal and other mineral products	CAN	
(Sum)	Imports metal and other mineral products	CAN	
(Sum)	Imports metal and other mineral products	CAN	
IMP_IND_EQUIP_BP_new	Imports industrial machinery, pieces and equipment	CAN	5
(Sum)	Imports industrial machinery, pieces and equipment	CAN	
(Sum)	Imports industrial machinery, pieces and equipment	CAN	
(Sum)	Imports industrial machinery, pieces and equipment	CAN	
(Sum)	Imports industrial machinery, pieces and equipment	CAN	
(Sum)	Imports industrial machinery, pieces and equipment	CAN	
IMP_TRANSP_BP_new	Imports motor vehicules and parts	CAN	5
(Sum)	Imports motor vehicules and parts	CAN	
(Sum)	Imports motor vehicules and parts	CAN	
IMP_CONS_BP_new	Imports consumption goods	CAN	5
(Sum)	Imports consumption goods	CAN	
(Sum)	Imports consumption goods	CAN	
(Sum)	Imports consumption goods	CAN	
(Sum)	Imports consumption goods	CAN	
(Sum)	Imports consumption goods	CAN	
USDCAD_new	Exchange rate CADUSD	CAN	5
JPYCAD_new	Exchange rate CADJPY	CAN	5

<b>Variable</b>	<b>Description</b>	<b>Région</b>	<b>Code</b>
GBPCAD_new	Exchange rate CADGBP	CAN	5
CAN_EQTY_NETFLOW	Canadian equity and investment fund shares, net flows	CAN	1
CAN_SEC_NETFLOW	Canadian securities, Net flows	CAN	1
FOR_SEC_NETFLOW	Foreign securities, Net flows	CAN	1
CAN_US_SEC_NETFLOW	Canadian securities, United States, Net flows	CAN	1
CPI_ALL_CAN	Consumption price index (CPI) (all)	CAN	5
CPI_SHEL_CAN	CPI (shelter)	CAN	5
CPI_CLOT_CAN	CPI (clothing and footwear)	CAN	5
CPI_HEA_CAN	CPI (health and personal care)	CAN	5
CPI_MINUS_FOO_CAN	CPI (all minus food)	CAN	5
CPI_MINUS_FEN_CAN	CPI (all minus food and energy)	CAN	5
CPI_GOO_CAN	CPI (durable goods)	CAN	5
CPI_DUR_CAN	CPI (goods)	CAN	5
CPI_SERV_CAN	CPI (services)	CAN	5
IPPI_CAN	Industrial production price index (IPPI) (all)	CAN	5
IPPI_ENER_CAN	IPPI (energy)	CAN	5
IPPI_WOOD_CAN	IPPI (wood)	CAN	5
IPPI_METAL_CAN	IPPI (metal and construction materials)	CAN	5
IPPI_MOTOR_CAN	IPPI (motor vehicles and parts)	CAN	5
IPPI_MACH_CAN	IPPI (industrial machinery and equipment)	CAN	5
WTISPLC	Petroleum price Western Intermediate Select (WTI) (FRED)	nan	5
TSX_HI	Toronto Stock Exchange (high)	nan	5
TSX_LO	Toronto Stock Exchange (low)	nan	5
TSX_CLO	Toronto Stock Exchange (close)	nan	5
DJ_CLO	Dow Jones index (close)	nan	5
SP500	Standard and Poor's (500) index (YAHOO)	nan	5
REAL_C_QC	Real Final consumption expenditure, chained (2012) dollars	QC	5
REAL_C_GOODS_QC	Real Final consumption expenditure, Goods, chained (2012) dollars	QC	5
REAL_C_DUR_GOODS_QC	Real Final consumption expenditure, Durable goods, chained (2012) dollars	QC	5
REAL_C_SERV_QC	Real Final consumption expenditure, Services, chained (2012) dollars	QC	5
REAL_C_G_QC	Real General governments final consumption expenditure,, chained (2012) dollars	QC	5
REAL_I_QC	Real Gross fixed capital formation, chained (2012) dollars	QC	5

<b>Variable</b>	<b>Description</b>	<b>Région</b>	<b>Code</b>
REAL_I_BUS_QC	Real Business gross fixed capital formation, chained (2012) dollars	QC	5
REAL_I_RES_QC	Real Gross fixed capital formation, Residential structures, chained (2012) dollars	QC	5
REAL_I_NRES_QC	Real Gross fixed capital formation, Non-residential structures, chained (2012) dollars	QC	5
REAL_I_ME_QC	Real Gross fixed capital formation, Machinery and equipment, chained (2012) dollars	QC	5
REAL_I_G_QC	Real Gross fixed capital formation, General governments gross fixed capital formation, chained (2012) dollars	QC	5
INVEST_INV_QC	Investment in inventories, chained (2012) dollars	QC	1
REAL_EXP_QC	Real Exports of goods and services, chained (2012) dollars	QC	5
REAL_IMP_QC	Real Imports of goods and services, chained (2012) dollars	QC	5
REAL_GDP_QC	Real Gross domestic product , chained (2012) dollars	QC	5
REAL_C_ONT	Real Final consumption expenditure, chained (2012) dollars	ONT	5
REAL_C_DUR_GOODS_ONT	Real Final consumption expenditure, Durable goods, chained (2012) dollars	ONT	5
REAL_C_SERV_ONT	Real Final consumption expenditure, Services, chained (2012) dollars	ONT	5
REAL_C_G_ONT	Real General governments final consumption expenditure,, chained (2012) dollars	ONT	5
REAL_I_ONT	Real Gross fixed capital formation, chained (2012) dollars	ONT	5
REAL_I_BUS_ONT	Real Business gross fixed capital formation, chained (2012) dollars	ONT	5
REAL_I_RES_ONT	Real Gross fixed capital formation, Residential structures, chained (2012) dollars	ONT	5
REAL_I_NRES_ONT	Real Gross fixed capital formation, Non-residential structures, chained (2012) dollars	ONT	5
REAL_I_ME_ONT	Real Gross fixed capital formation, Machinery and equipment, chained (2012) dollars	ONT	5
REAL_I_G_ONT	Real Gross fixed capital formation, General governments gross fixed capital formation, chained (2012) dollars	ONT	5
INVEST_INV_ONT	Investment in inventories, chained (2012) dollars	ONT	1
REAL_EXP_ONT	Real Exports of goods and services, chained (2012) dollars	ONT	5
REAL_IMP_ONT	Real Imports of goods and services, chained (2012) dollars	ONT	5
REAL_GDP_ONT	Real Gross domestic product , chained (2012) dollars	ONT	5
OIL_ALB_new	Crude oil production (ALB) (Cubic metres)	ALB	5

<b>Variable</b>	<b>Description</b>	<b>Région</b>	<b>Code</b>
EMP_NF	Employment total	NF	5
EMP_SERV_NF	Employment services	NF	5
EMP_FOR_OIL_NF	Employment forestry, fishing, mining, oil and gas	NF	5
EMP_CONS_NF	Employment construction	NF	5
EMP_SALES_NF	Employment sales (wholesale and retail trade)	NF	5
EMP_FIN_NF	Employment finance, insurance and real estate	NF	5
EMP_MANU_NF	Employment manufacturing	NF	5
EMP_PEI	Employment total	PEI	5
EMP_SERV_PEI	Employment services	PEI	5
EMP_FOR_OIL_PEI	Employment forestry, fishing, mining, oil and gas	PEI	5
EMP_CONS_PEI	Employment construction	PEI	5
EMP_SALES_PEI	Employment sales (wholesale and retail trade)	PEI	5
EMP_FIN_PEI	Employment finance, insurance and real estate	PEI	5
EMP_MANU_PEI	Employment manufacturing	PEI	5
EMP_NS	Employment total	NS	5
EMP_SERV_NS	Employment services	NS	5
EMP_FOR_OIL_NS	Employment forestry, fishing, mining, oil and gas	NS	5
EMP_CONS_NS	Employment construction	NS	5
EMP_SALES_NS	Employment sales (wholesale and retail trade)	NS	5
EMP_FIN_NS	Employment finance, insurance and real estate	NS	5
EMP_MANU_NS	Employment manufacturing	NS	5
EMP_NB	Employment total	NB	5
EMP_SERV_NB	Employment services	NB	5
EMP_FOR_OIL_NB	Employment forestry, fishing, mining, oil and gas	NB	5
EMP_CONS_NB	Employment construction	NB	5
EMP_SALES_NB	Employment sales (wholesale and retail trade)	NB	5
EMP_FIN_NB	Employment finance, insurance and real estate	NB	5
EMP_MANU_NB	Employment manufacturing	NB	5
EMP_QC	Employment total	QC	5
EMP_SERV_QC	Employment services	QC	5
EMP_FOR_OIL_QC	Employment forestry, fishing, mining, oil and gas	QC	5
EMP_CONS_QC	Employment construction	QC	5
EMP_SALES_QC	Employment sales (wholesale and retail trade)	QC	5
EMP_FIN_QC	Employment finance, insurance and real estate	QC	5

<b>Variable</b>	<b>Description</b>	<b>Région</b>	<b>Code</b>
EMP_MANU_QC	Employment manufacturing	QC	5
EMP_ONT	Employment total	ONT	5
EMP_SERV_ONT	Employment services	ONT	5
EMP_FOR_OIL_ONT	Employment forestry, fishing, mining, oil and gas	ONT	5
EMP_CONS_ONT	Employment construction	ONT	5
EMP_SALES_ONT	Employment sales (wholesale and retail trade)	ONT	5
EMP_FIN_ONT	Employment finance, insurance and real estate	ONT	5
EMP_MANU_ONT	Employment manufacturing	ONT	5
EMP_MAN	Employment total	MAN	5
EMP_SERV_MAN	Employment services	MAN	5
EMP_FOR_OIL_MAN	Employment forestry, fishing, mining, oil and gas	MAN	5
EMP_CONS_MAN	Employment construction	MAN	5
EMP_SALES_MAN	Employment sales (wholesale and retail trade)	MAN	5
EMP_FIN_MAN	Employment finance, insurance and real estate	MAN	5
EMP_MANU_MAN	Employment manufacturing	MAN	5
EMP_SAS	Employment total	SAS	5
EMP_SERV_SAS	Employment services	SAS	5
EMP_FOR_OIL_SAS	Employment forestry, fishing, mining, oil and gas	SAS	5
EMP_CONS_SAS	Employment construction	SAS	5
EMP_SALES_SAS	Employment sales (wholesale and retail trade)	SAS	5
EMP_FIN_SAS	Employment finance, insurance and real estate	SAS	5
EMP_MANU_SAS	Employment manufacturing	SAS	5
EMP_ALB	Employment total	ALB	5
EMP_SERV_ALB	Employment services	ALB	5
EMP_FOR_OIL_ALB	Employment forestry, fishing, mining, oil and gas	ALB	5
EMP_CONS_ALB	Employment construction	ALB	5
EMP_SALES_ALB	Employment sales (wholesale and retail trade)	ALB	5
EMP_FIN_ALB	Employment finance, insurance and real estate	ALB	5
EMP_MANU_ALB	Employment manufacturing	ALB	5
EMP_BC	Employment total	BC	5
EMP_SERV_BC	Employment services	BC	5
EMP_FOR_OIL_BC	Employment forestry, fishing, mining, oil and gas	BC	5
EMP_CONS_BC	Employment construction	BC	5
EMP_SALES_BC	Employment sales (wholesale and retail trade)	BC	5

<b>Variable</b>	<b>Description</b>	<b>Région</b>	<b>Code</b>
EMP_FIN_BC	Employment finance, insurance and real estate	BC	5
EMP_MANU_BC	Employment manufacturing	BC	5
UNEMP_NF	Unemployment rate	NF	2
UNEMP_PEI	Unemployment rate	PEI	2
UNEMP_NS	Unemployment rate	NS	2
UNEMP_NB	Unemployment rate	NB	2
UNEMP_QC	Unemployment rate	QC	2
UNEMP_ONT	Unemployment rate	ONT	2
UNEMP_MAN	Unemployment rate	MAN	2
UNEMP_SAS	Unemployment rate	SAS	2
UNEMP_ALB	Unemployment rate	ALB	2
UNEMP_BC	Unemployment rate	BC	2
EMP_PART_NF	Employment part time	NF	5
EMP_PART_PEI	Employment part time	PEI	5
EMP_PART_NS	Employment part time	NS	5
EMP_PART_NB	Employment part time	NB	5
EMP_PART_QC	Employment part time	QC	5
EMP_PART_ONT	Employment part time	ONT	5
EMP_PART_MAN	Employment part time	MAN	5
EMP_PART_SAS	Employment part time	SAS	5
EMP_PART_ALB	Employment part time	ALB	5
EMP_PART_BC	Employment part time	BC	5
UNEMP_DURAvg_NF_new	Unemployment average duration	NF	5
UNEMP_DURAvg_PEI_new	Unemployment average duration	PEI	5
UNEMP_DURAvg_NS_new	Unemployment average duration	NS	5
UNEMP_DURAvg_NB_new	Unemployment average duration	NB	5
UNEMP_DURAvg_QC_new	Unemployment average duration	QC	5
UNEMP_DURAvg_ONT_new	Unemployment average duration	ONT	5
UNEMP_DURAvg_MAN_new	Unemployment average duration	MAN	5
UNEMP_DURAvg_SAS_new	Unemployment average duration	SAS	5
UNEMP_DURAvg_ALB_new	Unemployment average duration	ALB	5
UNEMP_DURAvg_BC_new	Unemployment average duration	BC	5
CLAIMS_NF	Employment insurance initial claims, Allowed	NF	1
CLAIMS_PEI	Employment insurance initial claims, Allowed	PEI	1

<b>Variable</b>	<b>Description</b>	<b>Région</b>	<b>Code</b>
CLAIMS_NS	Employment insurance initial claims, Allowed	NS	1
CLAIMS_NB	Employment insurance initial claims, Allowed	NB	1
CLAIMS_QC	Employment insurance initial claims, Allowed	QC	1
CLAIMS_ONT	Employment insurance initial claims, Allowed	ONT	1
CLAIMS_MAN	Employment insurance initial claims, Allowed	MAN	1
CLAIMS_SAS	Employment insurance initial claims, Allowed	SAS	1
CLAIMS_ALB	Employment insurance initial claims, Allowed	ALB	1
CLAIMS_BC	Employment insurance initial claims, Allowed	BC	1
MANU_NF_new	Manufacturing new orders (total)	NF	5
DUR_NF_new	Manufacturing new orders (durables)	NF	5
MANU_PEI_new	Manufacturing new orders (total)	PEI	5
DUR_PEI_new	Manufacturing new orders (durables)	PEI	5
MANU_NS_new	Manufacturing new orders (total)	NS	5
DUR_NS_new	Manufacturing new orders (durables)	NS	5
MANU_NB_new	Manufacturing new orders (total)	NB	5
DUR_NB_new	Manufacturing new orders (durables)	NB	5
MANU_QC_new	Manufacturing new orders (total)	QC	5
DUR_QC_new	Manufacturing new orders (durables)	QC	5
MANU_ONT_new	Manufacturing new orders (total)	ONT	5
DUR_ONT_new	Manufacturing new orders (durables)	ONT	5
MANU_MAN_new	Manufacturing new orders (total)	MAN	5
DUR_MAN_new	Manufacturing new orders (durables)	MAN	5
MANU_SAS_new	Manufacturing new orders (total)	SAS	5
DUR_SAS_new	Manufacturing new orders (durables)	SAS	5
MANU_ALB_new	Manufacturing new orders (total)	ALB	5
DUR_ALB_new	Manufacturing new orders (durables)	ALB	5
MANU_BC_new	Manufacturing new orders (total)	BC	5
DUR_BC_new	Manufacturing new orders (durables)	BC	5
CPI_ALL_NF	Consumption price index (CPI) (all)	NF	5
CPI_SHEL_NF	CPI (shelter)	NF	5
CPI_CLOT_NF	CPI (clothing and footwear)	NF	5
CPI_HEA_NF	CPI (health and personal care)	NF	5
CPI_MINUS_FOO_NF	CPI (all minus food)	NF	5
CPI_MINUS_FEN_NF	CPI (all minus food and energy)	NF	5

<b>Variable</b>	<b>Description</b>	<b>Région</b>	<b>Code</b>
CPI_GOO_NF	CPI (goods)	NF	5
CPI_DUR_NF	CPI (durable goods)	NF	5
CPI_SERV_NF	CPI (services)	NF	5
CPI_ALL_PEI	Consumption price index (CPI) (all)	PEI	5
CPI_SHEL_PEI	CPI (shelter)	PEI	5
CPI_CLOT_PEI	CPI (clothing and footwear)	PEI	5
CPI_HEA_PEI	CPI (health and personal care)	PEI	5
CPI_MINUS_FOO_PEI	CPI (all minus food)	PEI	5
CPI_MINUS_FEN_PEI	CPI (all minus food and energy)	PEI	5
CPI_GOO_PEI	CPI (goods)	PEI	5
CPI_DUR_PEI	CPI (durable goods)	PEI	5
CPI_SERV_PEI	CPI (services)	PEI	5
CPI_ALL_NS	Consumption price index (CPI) (all)	NS	5
CPI_SHEL_NS	CPI (shelter)	NS	5
CPI_CLOT_NS	CPI (clothing and footwear)	NS	5
CPI_HEA_NS	CPI (health and personal care)	NS	5
CPI_MINUS_FOO_NS	CPI (all minus food)	NS	5
CPI_MINUS_FEN_NS	CPI (all minus food and energy)	NS	5
CPI_GOO_NS	CPI (goods)	NS	5
CPI_DUR_NS	CPI (durable goods)	NS	5
CPI_SERV_NS	CPI (services)	NS	5
CPI_ALL_NB	Consumption price index (CPI) (all)	NB	5
CPI_SHEL_NB	CPI (shelter)	NB	5
CPI_CLOT_NB	CPI (clothing and footwear)	NB	5
CPI_HEA_NB	CPI (health and personal care)	NB	5
CPI_MINUS_FOO_NB	CPI (all minus food)	NB	5
CPI_MINUS_FEN_NB	CPI (all minus food and energy)	NB	5
CPI_GOO_NB	CPI (goods)	NB	5
CPI_DUR_NB	CPI (durable goods)	NB	5
CPI_SERV_NB	CPI (services)	NB	5
CPI_ALL_QC	Consumption price index (CPI) (all)	QC	5
CPI_SHEL_QC	CPI (shelter)	QC	5
CPI_CLOT_QC	CPI (clothing and footwear)	QC	5
CPI_HEA_QC	CPI (health and personal care)	QC	5

<b>Variable</b>	<b>Description</b>	<b>Région</b>	<b>Code</b>
CPI_MINUS_FOO_QC	CPI (all minus food)	QC	5
CPI_MINUS_FEN_QC	CPI (all minus food and energy)	QC	5
CPI_GOO_QC	CPI (goods)	QC	5
CPI_DUR_QC	CPI (durable goods)	QC	5
CPI_SERV_QC	CPI (services)	QC	5
CPI_ALL_ONT	Consumption price index (CPI) (all)	ONT	5
CPI_SHEL_ONT	CPI (shelter)	ONT	5
CPI_CLOT_ONT	CPI (clothing and footwear)	ONT	5
CPI_HEA_ONT	CPI (health and personal care)	ONT	5
CPI_MINUS_FOO_ONT	CPI (all minus food)	ONT	5
CPI_MINUS_FEN_ONT	CPI (all minus food and energy)	ONT	5
CPI_GOO_ONT	CPI (goods)	ONT	5
CPI_DUR_ONT	CPI (durable goods)	ONT	5
CPI_SERV_ONT	CPI (services)	ONT	5
CPI_ALL_MAN	Consumption price index (CPI) (all)	MAN	5
CPI_SHEL_MAN	CPI (shelter)	MAN	5
CPI_CLOT_MAN	CPI (clothing and footwear)	MAN	5
CPI_HEA_MAN	CPI (health and personal care)	MAN	5
CPI_MINUS_FOO_MAN	CPI (all minus food)	MAN	5
CPI_MINUS_FEN_MAN	CPI (all minus food and energy)	MAN	5
CPI_GOO_MAN	CPI (goods)	MAN	5
CPI_DUR_MAN	CPI (durable goods)	MAN	5
CPI_SERV_MAN	CPI (services)	MAN	5
CPI_ALL_SAS	Consumption price index (CPI) (all)	SAS	5
CPI_SHEL_SAS	CPI (shelter)	SAS	5
CPI_CLOT_SAS	CPI (clothing and footwear)	SAS	5
CPI_HEA_SAS	CPI (health and personal care)	SAS	5
CPI_MINUS_FOO_SAS	CPI (all minus food)	SAS	5
CPI_MINUS_FEN_SAS	CPI (all minus food and energy)	SAS	5
CPI_GOO_SAS	CPI (goods)	SAS	5
CPI_DUR_SAS	CPI (durable goods)	SAS	5
CPI_SERV_SAS	CPI (services)	SAS	5
CPI_ALL_ALB	Consumption price index (CPI) (all)	ALB	5
CPI_SHEL_ALB	CPI (shelter)	ALB	5

<b>Variable</b>	<b>Description</b>	<b>Région</b>	<b>Code</b>
CPI_CLOT_ALB	CPI (clothing and footwear)	ALB	5
CPI_HEA_ALB	CPI (health and personal care)	ALB	5
CPI_MINUS_FOO_ALB	CPI (all minus food)	ALB	5
CPI_MINUS_FEN_ALB	CPI (all minus food and energy)	ALB	5
CPI_GOO_ALB	CPI (goods)	ALB	5
CPI_DUR_ALB	CPI (durable goods)	ALB	5
CPI_SERV_ALB	CPI (services)	ALB	5
CPI_ALL_BC	Consumption price index (CPI) (all)	BC	5
CPI_SHEL_BC	CPI (shelter)	BC	5
CPI_CLOT_BC	CPI (clothing and footwear)	BC	5
CPI_HEA_BC	CPI (health and personal care)	BC	5
CPI_MINUS_FOO_BC	CPI (all minus food)	BC	5
CPI_MINUS_FEN_BC	CPI (all minus food and energy)	BC	5
CPI_GOO_BC	CPI (goods)	BC	5
CPI_DUR_BC	CPI (durable goods)	BC	5
CPI_SERV_BC	CPI (services)	BC	5

## BIBLIOGRAPHIE

- AbdElminaam, D. S., El-Aal, A., Abdellatif, A. et al. (2023). Nowcasting egypt gdp using machine learning algorithms. *Journal of Computing and Communication*, 2(1), 1–8.
- Adanero-Donderis, M., Darné, O. et Ferrara, L. (2009). Un indicateur probabiliste du cycle d'accélération pour l'économie française. *Économie & prévision*, 189(3), 95–114.
- Ahn, S. C. et Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica*, 81(3), 1203–1227.
- Alessi, L., Barigozzi, M. et Capasso, M. (2010). Improved penalization for determining the number of factors in approximate factor models. *Statistics & Probability Letters*, 80(23-24), 1806–1813.
- Amengual, D. et Watson, M. W. (2007). Consistent estimation of the number of dynamic factors in a large n and t panel. *Journal of Business & Economic Statistics*, 25(1), 91–96.
- Ashiya, M. (2006). Are 16-month-ahead forecasts useful? a directional analysis of japanese gdp forecasts. *Journal of Forecasting*, 25(3), 201–207.
- Athey, S. et Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11, 685–725.
- Bai, J. et Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1), 191–221.
- Bai, J. et Ng, S. (2007). Determining the number of primitive shocks in factor models. *Journal of Business & Economic Statistics*, 25(1), 52–60.
- Basher, S. A. et Sadorsky, P. (2022). Forecasting bitcoin price direction with random forests : How important are interest rates, inflation, and market volatility? *Machine Learning with Applications*, 9, 100355.
- Bergmeir, C., Hyndman, R. J. et Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics & Data Analysis*, 120, 70–83.
- Blaskowitz, O. et Herwartz, H. (2011). On economic evaluation of directional forecasts. *International journal of forecasting*, 27(4), 1058–1065.
- Borup, D., Christensen, B. J., Mühlbach, N. S. et Nielsen, M. S. (2023). Targeting predictors in random forest regression. *International Journal of Forecasting*, 39(2), 841–868.
- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5–32.
- Breiman, L., Friedman, J., Olshen, R. et Stone, C. (1984). Cart. *Classification and regression trees*.
- Charpentier, A., Flachaire, E. et Ly, A. (2018). *Econométrie et Machine Learning*. Rapport technique, arXiv. org.
- Collodel, U. (2022). *The Interplay between Macroeconomic Forecasting, Financial Crises and Complexity*. (Thèse de doctorat). Université Panthéon-Sorbonne-Paris I.
- Cornec, M. et Mikol, F. (2011). *Nowcasting GDP directional change with an application to French business survey data*. Rapport technique, Technical report.
- Cortes, C. et Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20, 273–297.

- Cristelli, M. C. A., Tacchella, A., Cader, M. Z., Roster, K. I. et Pietronero, L. (2017). On the predictability of growth. *World Bank Policy Research Working Paper*, (8117).
- Cross, P. et Bergevin, P. (2012). Turning points : Business cycles in canada since 1926. *CD Howe Institute*, 366.
- Darwin, C. (1964). *On the origin of species : A facsimile of the first edition*. Harvard University Press.
- Davig, T. et Hall, A. S. (2019). Recession forecasting using bayesian classification. *International Journal of Forecasting*, 35(3), 848–867.
- Döpke, J., Fritsche, U. et Müller, K. (2019). Has macroeconomic forecasting changed after the great recession? panel-based evidence on forecast accuracy and forecaster behavior from germany. *Journal of Macroeconomics*, 62, 103135.
- Dörr, J. O. (2020). The nonlinearity of crises.
- Fortin-Gagnon, O., Leroux, M., Stevanovic, D. et Surprenant, S. (2022). A large canadian database for macroeconomic analysis. *Canadian Journal of Economics/Revue canadienne d'économique*, 55(4), 1799–1833.
- Frédéric, S. (2020). Introduction à l'apprentissage automatique. *École des Mines de Nancy*, 2021.
- Genuer, R., Poggi, J.-M., Tuleau-Malot, C. et Villa-Vialaneix, N. (2017). Random forests for big data. *Big Data Research*, 9, 28–46.
- Gogas, P. et Papadimitriou, T. (2021). Machine Learning in Economics and Finance. *Comput Econ*, 57(1), 1–4.  
<http://dx.doi.org/10.1007/s10614-021-10094-w>. Récupéré le 2021-07-25 de  
<https://doi.org/10.1007/s10614-021-10094-w>
- Gogas, P., Papadimitriou, T., Matthaïou, M. et Chrysanthidou, E. (2015). Yield curve and recession forecasting in a machine learning framework. *Computational Economics*, 45, 635–645.
- Goulet Coulombe, P., Leroux, M., Stevanovic, D. et Surprenant, S. (2020). *Prévision de l'activité économique au Québec et au Canada à l'aide des méthodes Machine Learning*. Rapport technique, CIRANO.
- Goulet Coulombe, P., Leroux, M., Stevanovic, D. et Surprenant, S. (2022). How is machine learning useful for macroeconomic forecasting? *Journal of Applied Econometrics*, 37(5), 920–964.
- Greer, M. (2003). Directional accuracy tests of long-term interest rate forecasts. *International Journal of Forecasting*, 19(2), 291–298.
- Hallin, M. et Liška, R. (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association*, 102(478), 603–617.
- Hansson, J., Jansson, P. et Löf, M. (2005). Business survey data : Do they help in forecasting gdp growth? *International Journal of Forecasting*, 21(2), 377–389.
- Hasan, M. et Boris, F. (2006). Svm : Machines à vecteurs de support ou séparateurs à vastes marges. *Rapport technique, Versailles St Quentin, France*. Cité, 64.
- Henriksson, R. D. et Merton, R. C. (1981). On market timing and investment performance. ii. statistical procedures for evaluating forecasting skills. *Journal of business*, 513–533.
- Iqbal, A. et Bowman, K. (2018). Can machine learning improve recession prediction accuracy? 6(4), 19.

- Iworiso, J. et Vrontos, S. (2020). On the directional predictability of equity premium using machine learning techniques. *Journal of Forecasting*, 39(3), 449–469.
- Jordan, B. (2020). Scores polygéniques et risque de cancer : Chroniques génomiques. *médecine/sciences*, 36, 535–537. <http://dx.doi.org/10.1051/medsci/2020088>
- Kauppi, H. et Saikkonen, P. (2008). Predicting us recessions with dynamic binary response models. *The Review of Economics and Statistics*, 90(4), 777–791.
- Kelly, B., Xiu, D. et al. (2023). Financial machine learning. *Foundations and Trends® in Finance*, 13(3-4), 205–363.
- Kotchoni, R., Leroux, M. et Stevanovic, D. (2019a). Macroeconomic forecast accuracy in a data-rich environment. *Journal of Applied Econometrics*, 34(7), 1050–1072.
- Kotchoni, R., Stevanovic, D. et Surprenant, S. (2019b). *Identification des points de retournement du cycle économique au Canada*. Rapport technique, CIRANO.
- Lahiri, K. et Yang, C. (2023). Roc and prc approaches to evaluate recession forecasts. *Journal of Business Cycle Research*, 1–30.
- Lardic, S. et Mpacko Priso, A. (1999). Une comparaison des prévisions des experts à celles issues des modèles b var. *Économie & prévision*, 140(4), 161–180.
- Liu, N. et Gao, F. (2022). The world uncertainty index and gdp growth rate. *Finance Research Letters*, 49, 103137.
- Maas, B. (2019). Nowcasting and forecasting us recessions : Evidence from the super learner.
- Mao Takongmo, C. O. et Stevanovic, D. (2015). Selection of the number of factors in presence of structural instability : a monte carlo study. *L'Actualité économique*, 91(1), 177–233.
- McCracken, M. et Ng, S. (2020). *FRED-QD : A quarterly database for macroeconomic research*. Rapport technique, National Bureau of Economic Research.
- McCracken, M. W. et Ng, S. (2016). Fred-md : A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4), 574–589.
- Merton, R. C. (1981). On market timing and investment performance. i. an equilibrium theory of value for market forecasts. *Journal of business*, 363–406.
- Meyer, T. (2019). On the directional accuracy of united states housing starts forecasts : Evidence from survey data. *The Journal of Real Estate Finance and Economics*, 58(3), 457–488.
- Nevasalmi, L. (2022). Recession forecasting with high-dimensional data. *Journal of Forecasting*, 41(4), 752–764.
- Öller, L.-E. et Barot, B. (2000). The accuracy of european growth and inflation forecasts. *International Journal of Forecasting*, 16(3), 293–315.
- Onatski, A. (2010). Determining the number of factors from empirical distribution of eigenvalues. *The Review of Economics and Statistics*, 92(4), 1004–1016.
- Pesaran, M. H. et Timmermann, A. (1992). A simple nonparametric test of predictive performance. *Journal of Business & Economic Statistics*, 10(4), 461–465.
- Pesaran, M. H. et Timmermann, A. G. (1994). A generalization of the non-parametric henriksson-merton test of market timing. *Economics Letters*, 44(1-2), 1–7.

- Phung, L. (2023). Guide pratique des enquêtes de conjonctures & protocole de prévision en temps réel. Publié sur le site de la Direction générale du Trésor. Récupéré le 2023-10-12 de <https://www.tresor.economie.gouv.fr/Articles/2023/10/11/guide-pratique-des-enquetes-de-conjonctures-protocole-de-prevision-en-temps-reel>
- Pierdzioch, C. et Rülke, J.-C. (2015). On the directional accuracy of forecasts of emerging market exchange rates. *International Review of Economics & Finance*, 38, 369–376.
- Piger, J. (2020). *Turning points and classification*. Springer.
- Pons, J. (2001). The rationality of price forecasts : a directional analysis. *Applied Financial Economics*, 11(3), 287–290.
- Poterie, A. (2018). *Arbres de décision et forêts aléatoires pour variables groupées*. (Thèse de doctorat). INSA de Rennes.
- Qi, M. (2001). Predicting us recessions with leading indicators via neural network models. *International journal of forecasting*, 17(3), 383–401.
- Schnader, M. H. et Stekler, H. O. (1990). Evaluating predictions of change. *Journal of Business*, 99–107.
- Schratz, P., Muenchow, J., Iturritxa, E., Richter, J. et Brenning, A. (2019). Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecological Modelling*, 406, 109–120.
- Sédillot, F. et Pain, N. (2005). Modèles d'indicateurs de la croissance du pib réel dans les principales économies de l'ocde. *Revue économique de l'OCDE*, (1), 185–246.
- Sinclair, T. M., Stekler, H. O. et Kitzing, L. (2010). Directional forecasts of gdp and inflation : a joint evaluation with an application to federal reserve predictions. *Applied Economics*, 42(18), 2289–2297.
- Skabar, A. (2008). A kernel-based technique for direction-of-change financial time series forecasting. Dans *Computational Science-ICCS 2008 : 8th International Conference, Kraków, Poland, June 23-25, 2008, Proceedings, Part II 8*, 441–449. Springer.
- Skabar, A. (2013). Direction-of-change financial time series forecasting using a similarity-based classification model. *Journal of Forecasting*, 32(5), 409–422.
- Soybilgen, B. et Yazgan, E. (2021). Nowcasting us gdp using tree-based ensemble models and dynamic factors. *Computational Economics*, 57, 387–417.
- Spelta, A., Pagnotoni, P. et Pecora, N. (2025). Big brothering the economy : nowcasting and forecasting with port satellite images. *Journal of the Royal Statistical Society Series A : Statistics in Society*, p. qnae157.
- Stéphane, T. (2015). *Modélisation prédictive et apprentissage statistique avec R*. Éditions Technip.
- Stock, J. H. et Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, 97(460), 1167–1179.
- Talon, C., Dautrême, E., Remy, E., Dirat, Y. et Le Strat, C. D. (2018). Analyse de différents algorithmes de classification par apprentissage automatique sur un cas d'usage du domaine nucléaire. Dans *Congrès Lambda Mu 21 «Maîtrise des risques et transformation numérique : opportunités et menaces»*.
- Tsuchiya, Y. (2012). Is the purchasing managers' index useful for assessing the economy's strength ? a directional analysis. *Economics Bulletin*, 32(2).

- Van Ruth, F. (2014). *Analysing short term developments in key economic indicators : predicting the sign of period on period changes*. Statistics Netherlands.
- Vrontos, S. D., Galakis, J. et Vrontos, I. D. (2021a). Implied volatility directional forecasting : a machine learning approach. *Quantitative Finance*, 21(10), 1687-1706.
- Vrontos, S. D., Galakis, J. et Vrontos, I. D. (2021b). Modeling and predicting us recessions using machine learning techniques. *International Journal of Forecasting*, 37(2), 647-671.
- Wafo, T., Clyde, V., Van Oirbeek, R. et Bury, T. (2022). Comparaison des méthodes de sélection des variables appliquées à la tarification des produits d'assurance non-vie.
- Xu, X. et Liu, W.-h. (2024). Forecasting the equity premium : can machine learning beat the historical average ? *Quantitative Finance*, 24(10), 1445-1461.
- Yazdani, A. (2020). Machine learning prediction of recessions : An imbalanced classification approach. *The Journal of Financial Data Science*.
- Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, 3(1), 32-35.
- Yousuf, K. et Ng, S. (2021). Boosting high dimensional predictive regressions with time varying parameters. *Journal of Econometrics*, 224(1), 60-87.