UNIVERSITY OF QUEBEC IN MONTREAL

Using Sentiment Analysis Techniques to Determine the Social Identity of Persons from their writing on social media

PRESENTED
AS PARTIAL REQUIREMENT
FOR THE MASTER'S IN COMPUTER SCIENCE

By

ZOHREH SOLEYMANI

April 2025

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

UTILISATION DES TECHNIQUES D'ANALYSE DE SENTIMENT POUR DÉTERMINER L'IDENTITÉ SOCIALE DES PERSONNES À PARTIR DE LEURS ÉCRITS SUR LES RÉSEAUX SOCIAUX

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DU PROGRAMME DE MAÎTRISE EN INFORMATIQUE

Par

ZOHREH SOLEYMANI

Avril 2025

ACKNOWLEDGMENTS

First, I express my deepest gratitude to my professor and research supervisor, Professor Hafedh Mili. His valuable advice, insightful feedback, and continuous support have been essential throughout the development of this thesis. His encouragement and expertise guided me at every step, and I am truly grateful for the opportunity to learn under his supervision.

I would also like to sincerely thank my family, especially my husband, for their unconditional love, endless patience, and unwavering support. Even when we were separated by distance, their belief in me gave me the strength and determination to persevere through the challenges. Their encouragement was a constant source of motivation that helped me stay focused and committed to my goals.

My heartfelt thanks also go to my friends, who stood by me during the most difficult times. Their encouragement, attentive listening, and practical help made a significant difference and reminded me that I was never alone on this journey.

Finally, I am deeply grateful to all who contributed, directly or indirectly, to the realization of this work. Your support has been invaluable, and I sincerely appreciate each and every one of you.

ABSTRACT

The marketing literature has shown that people's *social identity* influences their purchasing decisions. To the extent that is true, it becomes important to be able to recognize people's social identity from their on-line writing.

The question of *identity* has been studied in the social sciences literature. Stets et al.'s *identity theory* identified three types of identities (Stets et Serpe, 2016): 1) personal identities, 2) role identities, and 3) group identities. We argue that only *personal identities* and *role identities* are likely to: 1) influence consumer behavior, and 2) manifest themselves in on-line writing. Detecting identities in online writing is difficult because: 1) people may activate several identities simultaneously, and 2) there are no available datasets that are tagged by people's identities.

In this thesis, we propose strategies for uncovering *personal identities* and *role identities* that enable us to overcome these difficulties. For the case of *personal identities*, we propose to focus on those identities that involve taking a *polar stance* (*for* or *against*) with regard to some position. By doing so, *identity detection/-confirmation* can be framed as a *sentiment analysis* problem: determining the *polarity* of posts with regard to the identity in question. We applied this strategy to *veganism*, and obtained 88.7% with CNN+LSTM. Regarding *role identities*, we framed the problem as one of distinguishing between two (multiples) roles, determined in advance, and unlikely to be held by the same person. We tested the approach on the *Parent* and *child* identities, and obtained a 96% for Accuracy, Precision, Recall, and F1-score, with BERT. However, the techniques that we propose can only, 1) confirm identities, known in advance, and 2) works well for identities not likely to be held by the same person.

Overall, the thesis showed that NLP, machine learning–and sentiment analysis techniques in particular–can be used to detect people's social identities from their on-line writing. More work is needed to generalize the approach to *discover identities*, and to apply to more socially useful applications, including e-health.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

UQAM Université du Québec à Montréal.

CEM Customer Experience Management.

Résumé

Cette thèse examine l’application des techniques d’analyse de sentiment pour détecter les identités sociales dans les publications sur les réseaux sociaux — notamment les Tweets — en se concentrant particulièrement sur l'identité sociale végane. Dans un marché compétitif où l’autonomisation des consommateurs a transformé la gestion de l’expérience client, comprendre l’interaction complexe des facteurs de comportement des consommateurs, notamment le désir d’objectif et l'identité sociale, est devenu essentiel. S’appuyant sur le modèle de Bagozzi, cette recherche explore comment le désir d’objectif — l'un des éléments clés du processus d'achat — peut révéler les motivations des consommateurs et orienter les stratégies commerciales. Elle analyse également le rôle de l'identité sociale dans la formation des objectifs, en particulier au sein de la communauté végane, où l’appartenance à un groupe influence les actions et décisions individuelles.

La recherche étudie les défis liés à la détection de l'identité sociale, en abordant les biais, la désinformation, ainsi que les dynamiques culturelles et politiques complexes qui façonnent l'identité sur les plateformes numériques. Grâce à l'utilisation de modèles d'apprentissage automatique et d'apprentissage profond — notamment les K plus proches voisins (KNN), les arbres de décision, le perceptron multicouche (MLP) et l'architecture CNN+LSTM avec des vectorisations BERT et TF-IDF —, l’étude parvient à classifier les identités sociales à partir de données Twitter, atteignant une précision optimale de 88,7 % avec la vectorisation BERT. Les résultats montrent une variation minimale de performance entre les modèles traditionnels d'apprentissage automatique et les modèles avancés d'apprentissage profond, mettant en lumière l'efficacité nuancée des techniques d’analyse de sentiment dans les applications liées à l’expérience client, en particulier lorsque le volume de données est limité.
Cette recherche contribue à l’avancement de la gestion de l’expérience client en démontrant comment l’analyse de sentiment peut être utilisée comme outil d’interprétation de l’identité sociale, permettant ainsi aux entreprises de communiquer de manière plus significative avec des segments de consommateurs guidés par leur identité.
Enfin, elle traite des limites critiques des pratiques actuelles d’analyse de sentiment et propose des méthodes améliorées pour mieux gérer les biais et la désinformation dans la détection des identités sociales.

CHAPTER 1
INTRODUCTION

1.1 Context

In today's competitive and dynamic market, where consumers are more informed and empowered, effective customer experience management has become essential (Figure 2.2). The dynamics shaping consumer behavior and the factors influencing customer decisions are complex (Figure 2.1). *Goal desire*, which reflects consumers' needs and desires, has been identified by the marketing literature as the first step in the purchasing decision process (Bagozzi *et al.*, 2007). In turn, what consumers desire to acquire has been shown to depend, in part, on the consumers' *social identity* (Bagozzi *et al.*, 2007). Thus, we can assume that social identity plays an important role in consumer behavior.

Consider a lady named Rose who is a manager–a so-called *role identity* (Burke et Stets, 2023)–a fan of the *Montreal Canadiens* hockey team–*a group identity* (Burke et Stets, 2023)–and an active member of an animal protection group (another group identity). She also identifies as a moral person–a *personal identity* (Burke et Stets, 2023). While searching on-line for a new pair of boots, Rose sees an advertisement from her favorite brand. As she begins browsing the company's website to find the best option, she comes across a "Just for You" section featuring boots that are consistent with her values. These boots are made without animal leather, reflecting her commitment to animal protection. Additionally, she finds Chelsea boots, which are ideal for her role as a woman in a managerial position: these boots are comfortable without high heels, which is important considering how much she walks in her job. This discovery excites Rose, prompting her to add several pairs to her virtual shopping cart. The website further enhances her shopping experience by recommending personalized options, such as engraving her favorite hockey team's logo (Montreal Canadiens) on the boots, acknowledging her fandom. The site also suggests environmentally friendly artificial leather polish, recognizing Rose's commitment to moral choices. Finally, it offers handmade shoelaces purchased directly from the producer and advertised as Fair Trade, again catering to Rose's moral values.

The above example illustrated how different social identities played a role in Rose's experience, influencing different parts of her *customer journey*. This raises the question of how an online shopping application would know about a potential customer (Rose)'s social identities. Research has shown that identifying the social identity (or identities) of a potential customer is challenging, in general, because of biases, self-categorization, group commitment, and misinformation (Abrams et Hogg, 1990; Ellemers *et al.*, 1999; Park *et al.*, 2023). This is even more challenging if we do not have an elaborate consumer profile from which we can identify or guess a consumer's social identity(ies).

This thesis explores the extent to which a potential consumer's social identities can be determined by analyzing their writing on common social media. In particular, we explore the use of natural language processing (NLP) and AI techniques to determine the social identity of social media users based on their writing.

1.2 Research Objectives

The primary objective of this research is to explore the extent to which natural language processing (NLP) and AI techniques can be used to detect the social identities of a consumer that: 1) are evident in their online writing, and 2) are known to influence purchasing behavior.

First, as the previous example showed, consumers may have several identities which may influence a consumer's behavior to varying degrees. Second, it is not clear that all those identities are detectable in the consumer's social media writings. Third, it is not clear that traditional *sentiment analysis techniques*, which combine NLP and AI, can detect those identities from consumers' on-line writing. Sentiment analysis techniques have been very effective at detecting *opinion polarity*, e. g. determining whether a product review (e. g. a drill or a computer), or a service review (phone company, or restaurant) is positive. Even so, the review of *cultural/hedonistic products* was shown to be problematic (Abbasi, 2007). However, it is not clear that sentiments analysis techniques can detect social identity markers–whatever those are–in online writing.

1.3 Methodology

As the introductory example showed, a person can have several *identities*, each one of which can potentially influence the *desirability* of various products or services–the *goal desire* step of the 'customer journey', according to the marketing literature (Bagozzi *et al.*, 2007). Thus, our first goal would be to:

- (O_1): study the social sciences literature to identify the different kinds of social identities.

Having identified the different types of identities, we need to figure out, among those identities:

- (O_2): which kinds of identities are more likely to influence consumers' desire for certain products or services, but also
- (O_3): which kinds of identities are more likely to *manifest themselves in a detectable way* in consumers' online social media presence.

Once we have identified the types of identities that satisfy both of the preceding criteria, for each candidate identity type, we have to:

- (O_4): design a strategy to detect identities of that type
- (O_5): apply the corresponding strategy, and evaluate/test its effectiveness.

As it turned out, the *strategy* depends heavily on the kind of on-line writing data that we have. In other words, the strategy will be very much data-driven. Given a dataset, the next question is to adopt or devise a technique that would enable us to detect the target identity from that data set. Thus, we can think of a strategy as a combination of $<$identity, data set, discovery technique$>$.

Our research has identified two types of identities that, a) are likely to influence consumer behavior (O_2), and b) are likely to manifest themselves in online writing:

- *Personal identities*, including a person's moral values. Using our introductory example, Rose's self-perceived *morality* is a personal identity
- *Role identities*. Using our introductory example, Rose's *manager* role is a role identity

For personal identities, we explored the extent to which *sentiment analysis* techniques can be used to detect a personal identity. We framed the problem as one of detecting how a person feels about a specific *value*. Positive sentiments suggest that the writer adheres to that value; negative sentiments suggest otherwise. We tested our hypothesis on *veganism*, as a personal identity, and the results showed a high level of accuracy (see Chapter 4).

With *role identities*, we proposed a different approach: *roles* influence, first and foremost, *what* we talk about, and not so much how we feel about it. Thus, we propose a supervised learning approach where we train a supervised machine learning model with two data sets, labeled by two different roles $Role_A$ and $Role_B$, with the hope of recognizing unlabeled text as one written by a person playing either $Role_A$ or $Role_B$.

1.4 Contributions

The important contributions of this thesis include demonstrating the effectiveness of sentiment analysis for detecting social identity. By leveraging machine learning and deep learning techniques, the study shows how businesses can better understand consumer motivations and develop targeted strategies that consider identity-based factors. This research advances sentiment analysis applications in customer experience management by providing insights into how goal desire and social identity influence consumer behavior.

1.5 Contents of the thesis

This thesis consists of five chapters:

1. *Chapter 1: Literature review*: The literature review examines relevant research areas, starting with customer experience management and the factors influencing consumer behavior through goal desire. It then delves into social identity theory, which reveals the existence of different kinds of identities, illustrated in the introductory example. *Sentiment analysis and opinion mining* is the third area covered in the literature review.

2. *Chapter 2: Methodology*. This chapter presents the methodological framework, detailing the methodology (research objectives O_1 through O_5) sketched in Section 1.3.

3. *Chapter 3: discovering personal identities* deals with the discovery of *personal identities* by using the example of *veganism*. In particular, we framed the question of checking whether someone is *vegan* by assessing how they feel about *veganism*, and used sentiment analysis techniques to that end, leading to good results.

4. *Chapter 4: discovering role identities*, which deals withe discovery of role identities. Here, we propose a supervised classification approach that aims at confirming one identity among a predeter-

mined set, based on online writing.

5. *Chapter 5: Conclusion*. The concluding chapter summarizes the key findings and contributions of the thesis. It also discusses the limitations of the study and suggests directions for future research.

Details about the data sets and the various programs are presented in the appendix.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This review of the literature focuses on key research areas that support the thesis's objective of detecting social identity through analysis of social media content. Several fields intersect in this research. The first area, customer experience management, provides the initial motivation for this work. In particular, we will present Bagozzi framework (Bagozzi *et al.*, 2007) which identified *social identity* as an influence factor for a key step in *customer journeys*, namely, the desirability of specific products or services.

The second area explores *social identities*, with the hope of identifying how social identities can manifest themselves, in general, and in (online) writing.

Lastly, we look at *sentiment analysis and opinion mining*, which combines natural language processing (NLP) techniques and machine learning to extract different pieces of information from natural language text.

2.2 Customer experience management

The following discussion is largely derived from the work of (Mili *et al.*, 2016; Benzarti *et al.*, 2022).

2.2.1 Understanding the purchasing process

As shown in figure 2.1 according to the comprehensive model of consumer behavior proposed by Bagozzi, there are several steps a person goes through to make a purchase, and various factors influence the consumer during each step(Mili *et al.*, 2016).

2.2.1.1 Goal desire

Understanding what the customer wants and needs. Several factors influence the setting of a goal.

- **Goal feasibility:** I might dream about travelling to Mars, but I wouldn't realistically consider it a goal because it's not something I can achieve.

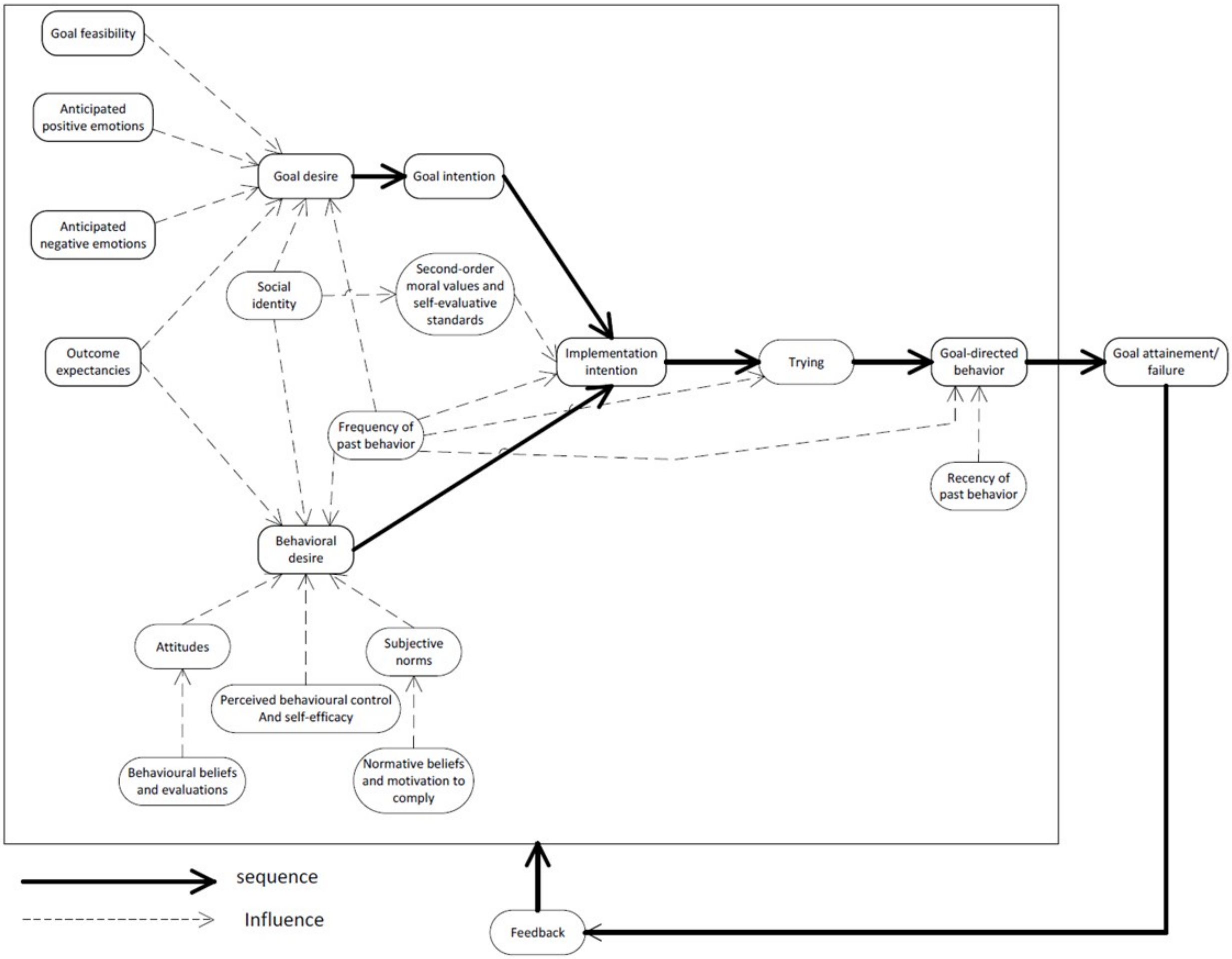


Figure 2.1 Perception of (Mili *et al.*, 2016) from the model presented by Bagozzi

- **Anticipated positive emotions:**This involves how good it will feel to achieve the goal, based on the positive emotions I expect to experience if I succeed, and how likely I believe it is that I will succeed.
- **Anticipated negative emotions:**This considers how bad I'll feel if I don't attain the goal, including the negative emotions I expect to experience if I fail and how likely I think failure is.
- **Outcome expectancies:**This relates to how much money you expect to gain if you pursue the goal, considering the potential outcomes of success versus failure.
- **Social identity** This refers to how individuals feel about being part of a group, whether it makes them feel good or bad to be a member of a specific group.
- **Frequency of past behavior:** Although Bagozzi's framework didn't originally include this, it is an im-

portant aspect of the theory of trying, which underpins this framework. The concept is that customers might go through the process of goal setting a few times, but eventually, they may take a shortcut:"I've thought this through many times and found it worthwhile to consider or pursue this goal" (Mili *et al.*, 2016).

2.2.1.2 Goal intention

Goal Desire refers to how strongly we want to achieve a goal, but it doesn't involve the decision to pursue it. This step is about making up our minds.

2.2.1.3 Behavioral desire

Researchers have identified six key influences on behavioral desire: 1) social identity, 2) outcome expectancies, 3) frequency of past behavior, 4) attitudes, 5) subjective norms, and 6) perceived behavioral control and self-efficacy. The first three were discussed in the previous section. Here, we will discuss the remaining three:

- **Attitudes:** The Theory of Reasoned Action (TRA), introduced by (Ajzen et Fishbein, 1975), highlights how attitudes affect whether someone will act. A positive attitude toward an action encourages a consumer to act, while a negative attitude discourages them. Attitudes toward an action are a combination of two things: beliefs about the outcome of the action and the evaluation of that outcome. For example, if someone decides to buy jeans, they might believe that the jeans will make them look fashionable (the outcome). If they value being fashionable, this positive evaluation encourages them to make the purchase.

- **Subjective norms:** Defined by (Vallerand *et al.*, 1992) , subjective norms represent an individual's perception of social pressure to either perform or not perform an action. They are a combination of two elements: 1) The belief that their colleagues, spouse, or peer group expects them to perform the action (a probability). 2) The perceived reward or penalty for respecting or not respecting that expectation. For instance, a high school student might believe with 90% certainty that their fellow students expect them to wear a coat. By not wearing a coat, they may think that they won't be invited to eat lunch with the group. This combination of expectation and consequence influences their decision.

- **Perceived behavioral control and self-efficacy:** Introduced by the Theory of Planned Behavior, these concepts play an important role in determining whether someone will engage in a particular behavior. "Perceived behavioral control" is defined as an individual's perception of how easy or difficult it is to perform the behavior(Ajzen et Madden, 1986). It involves a combination of two factors: 1) The belief in possessing the necessary resources or abilities to perform the behavior. 2) The perceived importance of those resources or abilities. For instance, if someone wants to quit alcohol and believes that strong willpower is essential, but feels their flaw is low willpower, they may perceive quitting as too difficult. In this research, we consider behavioral control and self-efficacy as synonymous concepts; both representing an individual's belief in their ability to perform a behavior.

2.2.1.4 Implementation intention

In the planning phase, once I've selected a goal or a goal-directed behavior, I will begin making plans to achieve the goal. The key here is that my consumption is aimed at reaching a desired end state, such as achieving a specific goal. During this phase, implementation intention involves identifying the steps or individual actions necessary to reach my goal. These steps can themselves become goals if their execution presents challenges. The planning phase is influenced by the frequency of past behavior. If I've successfully achieved the goal multiple times before, I can reuse the same solution or implementation plan. However, if I have never achieved the goal, I will need to build a new implementation plan. Another influencing factor is second-order moral values and self-evaluative standards. Values are the criteria or frames of reference by which people justify actions and evaluate both actions and individuals. Self-evaluative standards represent how consumers perceive themselves or what they aspire to become. For example, if our shopper values social responsibility, the grocery store might promote the latest politically correct products to appeal to this value. Moral values and self-evaluative standards are influenced by the consumer's social identity.

2.2.1.5 Trying

According to the original theory of trying, the act of trying was defined as a subjective state that reflects how much a person believes they have or will try to act. Later, this definition was expanded to include the execution of a previously established plan, with each action potentially presenting problems that could hinder progress. Trying involves monitoring the progress toward the objective and making adaptations to the plan if necessary. For example, consider the goal intention of giving a friend a graduation gift (behavioral desire). This involves planning, or implementation intention, which includes steps like going to a mall, selecting specific stores to visit, and choosing a gift that aligns with a budget. Executing this plan involves carrying out these steps going to the mall, visiting the stores, etc. However, the consumer may encounter issues at any step, such as transportation difficulties, not finding suitable stores, or finding options that are not within their budget.The trying phase is influenced by both the frequency of past behavior and the recency of past behavior. The recency of past behavior refers to how recently a behavior was performed. In the theory of trying, recency can affect current behavior similarly to the frequency of past behavior, as the behavior is still fresh in the consumer's mind.

2.2.1.6 Goal-directed behavior

Refers to the 'final act' in the consumption process, such as making a purchase.

2.2.1.7 Goal attainment/failure

This is the step where the consumer evaluates the extent to which they have achieved their goal.

2.2.1.8 Feedback

Based on the previous assessment, the consumer can adjust any of the choices or actions made during the 'consumption process' including the selection of goals to pursue.

This 'comprehensive model' of consumer behavior, proposed by (Bagozzi, 2006) with very minor adaptations of our own, is the result of 'merging' consumer behavior models related to different types of consumption. In some cases, certain steps or influencing factors may be irrelevant. Our framework for Customer Experience Management (CEM) should provide guidance or tools for customizing this process for specific consumption or CEM scenarios.

2.2.2 The Influencing Factors of the Various Steps in Bagozzi's Model and How to Capitalize on Them to Enhance the Customer's Experience

CEM involves managing the interactions between a business and its clientele. Whether existing or potential, customers interact with a company to satisfy certain needs through the products or services offered by the company. The company, in turn, interacts with its clientele, driven by its primary purpose of meeting customer needs, which ultimately leads to compensation for the company by selling its products or services at a price exceeding production cost. The same principle applies to public service entities, which derive their value from fulfilling public needs. Customers, each with their unique desires for survival and happiness, have specific needs that are met by engaging with a company's offerings as they pursue their goals. Customer experience encompasses the dynamics between customers and companies during the process of fulfilling these needs, whether through a product or service. The family lifecycle theory verifies that people's needs evolve through different life stages, influenced by various factors such as underlying processes (e.g., child-rearing) and the resources available at each stage.

2.2.3 Understanding the Purchasing Process: A Cognitive Approach

Marketers and social psychologists have extensively studied consumer behavior to understand its key mechanisms. They adopt a broad definition of 'consumer,' encompassing products/services (like food, clothing, cell phone plans), behaviors (such as dieting, exercising), and beliefs (like political affiliations or social values). At its core, consumption is a conscious act aimed at satisfying a need or desire.

Various psychological models have been proposed, including the theory of reasoned action (Ajzen et Fishbein, 1975), the theory of planned behavior (Ajzen, 1985), the MODE model (Fazio, 1986), the theory of trying (Bagozzi et Warshaw, 1990), the theory of self-regulation(Bagozzi, 1992), and others. The theory of reasoned action (TRA) views all actions as purposeful, beginning with the formation of an intention to act, followed by the action itself(Ajzen et Fishbein, 1975). TRA highlights that intentions are shaped by two factors: attitudes toward the action and subjective norms. Attitudes toward the action refer to the perceived likelihood of a certain outcome, while subjective norms involve the individual's belief about others' expectations regarding their behavior.

For instance, someone may want to buy Nike shoes because they believe it will enhance their appearance (personal outcome) and because they feel pressure from their peers to wear this type of sneakers (social influence). However, researchers have identified some limitations in this theory, leading to further refinement. The theory of planned behavior (Ajzen, 1985) introduces the concept of perceived behavioral control, which is considered an individual's confidence in executing the action. The MODE model (Ajzen et Madden, 1986), which considers motivation and opportunity as determining factors, addresses impulsive actions.

Further refinements, such as the theory of trying (Bagozzi et Warshaw, 1990), recognize that actions themselves can become goals and that the complexity of consumption actions needs to be considered in the pursuit of consumption goals. These models illustrate ongoing efforts to understand the multifaceted nature of consumer behavior and to adapt theories that better explain and predict consumer actions.

For this thesis, we'll focus on the synthesis presented by (Bagozzi *et al.*, 2007), which consolidates the previously discussed factors. The model, depicted in Figure2.2, encompasses various steps and influencing factors. We'll provide an overview of these elements. For the following reasons, it should be considered that not all steps are always done at the time of purchase.

Firstly, depending on the type and complexity of consumption decisions, there are a variety of paths in this process. For instance, routine or low-risk consumption activities (like picking up a bottle of milk) might not involve much planning.

Secondly, social psychologists distinguish between desirable goals and desirable behaviors, often with the former preceding the latter. For instance, someone may aim to lead a healthier lifestyle as a desirable end state, which could involve various behaviors like dieting or exercising. However, for simplicity, we'll consider these as separate "process paths" without emphasizing the relationship between goals and behaviors.

Thirdly, according to the model, there is a difference between desires, intentions, plans, and actual actions. A consumer may have several desires (like buying both a motorcycle and a car), but they may only intend to pursue one (e.g. buying a car). This difference applies to behaviors as well, such as New Year's resolutions. Once a consumer has decided to pursue a goal or behavior, they need a plan of action to achieve it. For instance, if the goal is to buy a car, the consumer must create a plan on how to accomplish this, which is referred to as an implementation intention in the model. After planning, the consumer then executes it by carrying out the actions outlined in the plan. In the model, this is represented by the steps of trying and goal-directed behavior.

2.2.4 Customer experiences management pattern

Figure 2.2 illustrates the concept. Consumers are depicted as "active systems" (living organisms) whose internal processes (such as survival and the pursuit of happiness) necessitate physical resources (like transportation and food) or abstract conditions (such as happiness or fulfillment). These internal needs trigger consumption processes, either to replenish depleted resources ("we are out of cereal") or to achieve desired conditions ("I need to dine out"). These consumption processes consist of several steps. Based on psycho-sociological studies of consumers, the decisions made in these steps are influenced by various factors. Some of these factors involve interactions between the enterprise and the consumer, which can be initiated by the consumer, such as visiting the enterprise's website, or by the enterprise, through communication or advertising. Given that many steps in the consumption process involve choices influenced by psycho-sociological factors, it is beneficial for the enterprise to understand the value of these factors. Some of these influencing factors, such as anticipated positive emotions, are changeable, and the enterprise may benefit from adjusting them to its advantage. For instance, the enterprise could highlight how purchasing its products can improve life quality or emphasize the ease of obtaining credit through the purchase (perceived behavioral control). These factors are referred to as "read-write" influence factors in figure 2.2, in contrast to "read-only"(RO) influence factors that enterprises cannot change or act upon. Importantly, while all read-write influence factors have potential actionable impact from a CEM standpoint, not all may be worth pursuing.

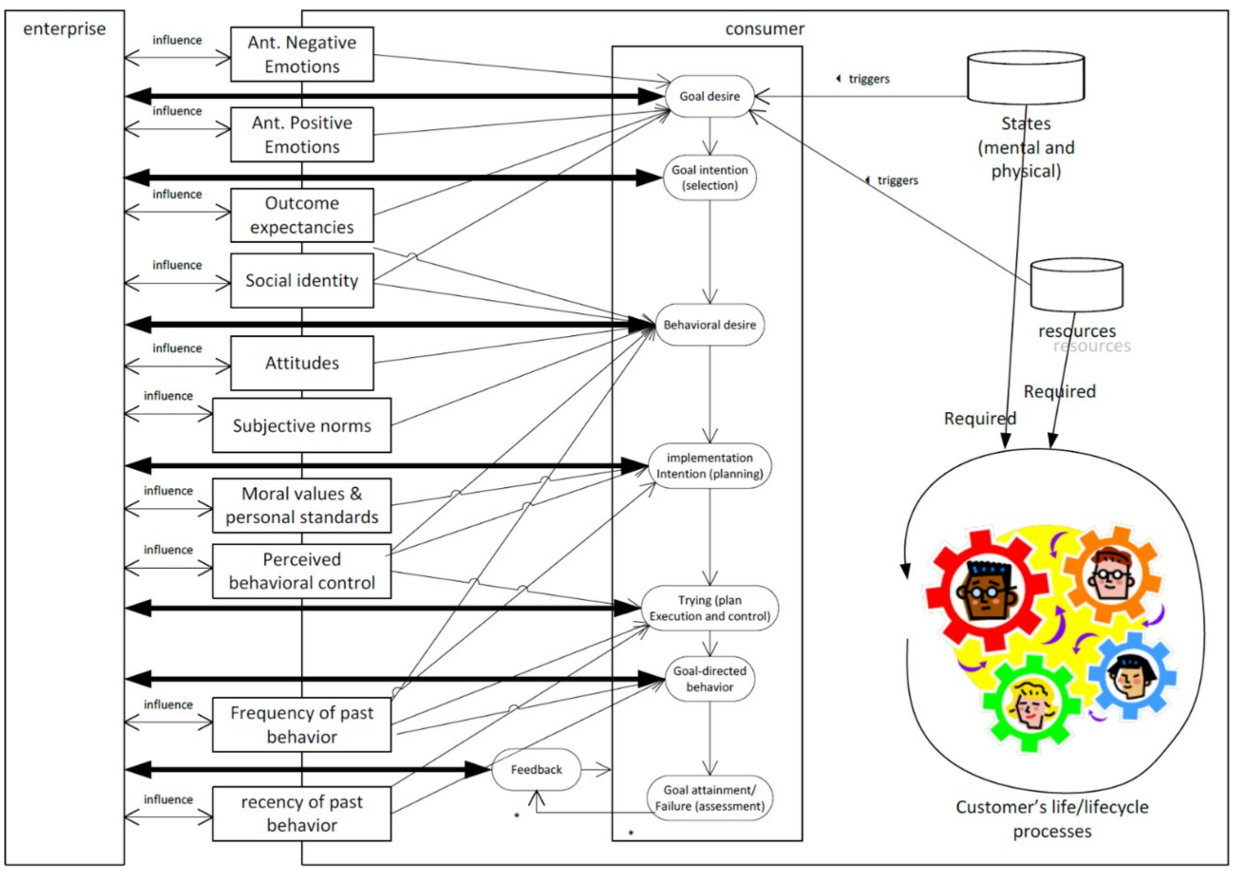


Figure 2.2 influence factors that enterprises cannot change or act upon

- **Seller Influence on Goal Desire:**

  Sellers (particularly in industries like tobacco, beer, and fashion) influence anticipated emotions and the social identity associated with their products. However, this influence often occurs outside direct one-to-one interactions between the customer and the seller's e-commerce system.

- **Different Channels Accompany Different Process Subsets:**

  Various channels accompany different subsets of the purchasing process, often corresponding to distinct entry points. For instance, entering an e-commerce site via a web search for a product may bypass several early steps of the process that would typically occur when physically walking into a store.

- **Consumer Interaction with Multiple Sellers**

  Consumers may interact with multiple sellers during certain process steps, especially early on before committing to a specific product or seller. For example, location-based services in a shopping mall may present advertisements from various restaurants or stores, highlighting competition among sellers for the customer's attention.

- **Complexity of Trying and Goal-Directed Behavior:**

  Consumers may interact with multiple sellers during certain process steps, especially early on before committing to a specific product or seller.

Considering these points, the subsequent section will explore the various steps of the purchasing process to determine how the pattern depicted in figure 2.2 can be effectively applied.

2.3 Social Identity

Social identity plays a crucial role in shaping goal desire during the purchasing process. This influences decision-making in the following ways:

- Decision-making is a product of the role-taking/role-making process, where various components, including social networks, social status, role expectations, choice behavior, and personal attributes, can impact a person's choices (Stets et Serpe, 2016).

- Even though human behavior is inherently unpredictable because our behavior in a given situation is influenced by the outcomes of previous behaviors in similar–unbeknownst–situations (Stets et Serpe, 2016), the increasing use of internet in online shopping does keep track of past experiences.

Next, we will discuss what social identity is and explore different types of social identities.

2.3.1 Social identity and Categorizing

The concept of social identity originates from the idea of a social group, which involves more than two people who share a common sense of belonging ("us" in contrast to "them"). This is distinct from personal identity, which pertains to your unique personality traits that you do not share with others ("me" and "you") (Hogg, 2000). There is ongoing debate over whether social identity and personal identity should be considered the same or different concepts.

For the purposes of this research, we will assume they are the same (Hogg, 2000). People create their own self-definitions, but these definitions are often shaped by the groups to which they belong. Social structures influence social interactions by either facilitating or constraining entry into and exit from social relationships. This doesn’t mean that people are incapable of creativity and social change; rather, there is societal pressure to conform and maintain the social order (Stets et Serpe, 2013).

Large social groups, such as those defined by race, class, gender, and socioeconomic status, play significant roles in society. These groups act as boundaries that impact an individual’s life chances, including their ability to form social relationships (Stets et Serpe, 2013).

Social structures like neighborhoods, clubs, and teams are referred to as intermediate groups. These groups also influence the likelihood of certain types of social relationships forming. Proximate structures, such as families or school clubs, are those closest to interpersonal interactions (Stets et Serpe, 2013). Social identity greatly influences how individuals perceive themselves and their role in the world. In fact, social identities provide people with a place in the world, shaping their sense of meaning and functioning (Dingle *et al.*, 2013; Cruwys *et al.*, 2014). Social identity is strengthened when social structural features of the world allow people to feel that their ingroup identity is positive, distinct, and enduring (Turner *et al.*, 1979; Cruwys *et al.*, 2014). When an identity is activated in a situation, a perceptual control model loop occurs. This loop has five parts:

- *Identity Standard or self meaning*: This refers to the meanings and associations individuals hold about their identity.
- *Perceptual Input*: This is how you perceive and feel about yourself in each situation, including your self-concept and the feedback you receive from others.
- *Comparator*: A mental mechanism that compares your current perceptions with your identity standard. This involves comparing your perceptions of your identity with the meanings associated with your identity standard.
- *Comparator*: This represents how you feel about the comparison between your current identity perceptions and your identity standard. You may feel good or bad depending on whether they align.
- *Output*: This is your behavior in response to this process (your actions or reactions to the environment), which carries meaning.

The idea of multiple identities stems from (James, 1890) notion that a person has multiple selves, with each self corresponding to how different people perceive the person in various contexts. Today, we speak of multiple identities rather than multiple selves, but the underlying idea is the same. Throughout our lives, we adopt different identities, and in any given situation, multiple identities can be activated. Research suggests that multiple identities are more likely to be activated simultaneously when they share similar meanings (Deaux, 1992; Stets, 1995). For example, a feminine gender identity and a mother identity might be activated together because both carry meanings of caring and nurturing. Therefore, for someone with a maternal social identity, expressing their feminine side might feel more natural and comfortable (Stets et Serpe, tion). There are three types of social identity:

- *Role Identity*: Defined by the social roles one occupies, such as being a teacher.
- *Group Identity*: Defined by membership in specific groups, like a book club.
- *Personal Identity*: Defined by specific characteristics of a person, such as being athletic (Stets et Serpe, tion)

An identity comprises meanings related to a person's different roles in society (role identities), the groups they belong to (group identities), and the unique way they see themselves (personal identities)(Burke et

Stets, 2009). Role identities are shaped by the expectations associated with social positions, such as being a teacher, student, or parent.

Here is another categorization of social identity:

- **Group Identities:**

  These are the meanings associated with interacting with specific sets of people, such as your family or work colleagues. It involves trying to fulfill the expected behaviors of the group.

- **Person Identities:**

  This identity is based on the unique characteristics that make you an individual, distinct from just being a role or group member. It's how you see yourself as unique. For example, seeing yourself as a moral person, which guides your behavior.

These identities often overlap and cannot be easily separated. In group settings, people can assume various roles, and each role brings its own unique personal identities into play. If we focus too heavily on one identity, we might overlook important aspects related to other identities.

For example, consider a student attending school. When the school identity is activated, such as during a pep rally, students might perform school chants. Some students may take the lead, embodying their class role, while others might be rowdier due to their personal identity of being more aggressive. In this situation, behaviors simultaneously reflect group, role, and personal identities. Therefore, it is essential to recognize that these different identity bases are often activated simultaneously during interactions (Stets, Jan E., and Richard T. Serpe). Additionally, the sense-of-place scale, presented by (Jorgensen et Stedman, 2001; Jorgensen et Stedman, 2006) and later validated by (Scannell et Gifford, 2013) , consists of three dimensions:

- **Place Identity:**

  Cognitive indicators, such as feeling that "Everything about my lake property is a reflection of me."

- **Place Attachment:**

  Affective indicators, like "I really miss my lake property when I'm away from it for too long."

- **Place Dependence:**

  Conative evaluative indicators, such as "My lake property is the best place for doing the things that I enjoy most" (Belanche *et al.*, tion).

In the context of intergroup behavior, people often rely on their group as a guide for their thoughts and actions, acting more as group members than as individuals (Brown, tion) As either in-group members or out-group members, our behavior is shaped by norms that are closely tied to our social identity(Sherif, 1936; Tankard et Paluck, 2016; Brown, tion)

According to (Turner, 1982), recognizing oneself as part of a group plays a vital role in various group behaviors. For instance, when individuals strongly identify as university students, this identity influences their thoughts and motivations. This affiliation with being a student serves as the foundation for engaging in group-related activities, such as attending lectures, expressing emotions related to group events (like celebrating a sports victory or feeling upset about cuts to student services), and collaborating with the group toward common objectives, such as supporting the college team or protesting budget reductions. Conversely, these activities and emotions are less likely to influence individuals who identify more strongly with other aspects of their identity, such as being an employee or a vegetarian. This example demonstrates that when group membership forms a significant part of our self-definition, it shapes our thoughts, emotions, and behaviors in various situations. This influence occurs because, firstly, social identity enables individuals to share a collective understanding of the social world, and secondly, it establishes a foundation for mutual social influence, facilitating the coordination of perceptions and actions (Turner, 1991; Haslam et Authors, tion).When groups compare themselves, they seek to demonstrate distinctiveness from other groups. Since social identity not only describes but also evaluates you, inter-group comparisons often aim to achieve a distinctive and positive evaluation of one's own group (Hogg et Authors, tion). Conformity refers to the tendency of individuals to change their behavior to align with the self-defining prototype of their group (Hogg et Authors, tion). Small groups are more commonly found in collectivist societies than in individualist societies (Hogg et Authors, tion).

2.3.2 Intergroup behavior and changing identity

This section focuses on how social identity can change over time. This will include the importance of considering these changes as customers progress through different life stages, such as getting married or attempting to extend their social identities to enhance their happiness. Subordinate or stigmatized groups, in particular, may seek to change or add new social identities (Brown, Rupert). Life transitions can be more positive and less harmful if individuals have access to social identities from multiple group memberships, allowing them to gain and maintain positive group connections and resources (Haslam et Authors, tion). When you stereotype someone by placing them into a category, you perceive them based on a generalized idea. This can also happen to yourself, known as self-stereotyping. Because prototypes not only describe but also prescribe appropriate ways to feel and behave, self-stereotyping can lead to normative behavior.

Normative behavior refers to actions and attitudes that conform to social or group norms. Expected or "normal" ways of feeling, thinking, and behaving in a particular group or culture.(Hogg et Reid, 2006) As a result, self-categorization fosters conformity within a group, as well as patterns of in-group liking, trust, and solidarity.

This process makes people seem less like individuals and more like representations of their groups. It leads to conformity within groups and fosters feelings of liking, trust, and solidarity among group members(Hogg et Authors, tion). Social identity is guided by two basic motivations: self-enhancement and the reduction of uncertainty. According to the intergroup social comparison theory, groups strive to be superior and distinct(Hogg et Authors, tion).
People continuously change their identity during social interactions, and the identities that are effective at one time may not be as effective at another (Stets et Serpe, tion). However, while the journey of life changes is always personal, it can be made not only personally meaningful but also more enjoyable rather than unpleasant (Haslam et Authors, tion). The individualistic explanations for intergroup behavior often focus on factors like family upbringing or personal frustration, which can overlook the importance of societal positions and group interactions. This oversight can similarly affect the understanding of behavioral changes over time (Brown, tion).

During identity changes, there can be compromises related to the effects of support, control, belonging, and inclusion during periods of transition. As we look into the future(Jetten et Authors, 2009), an increase in social psychological resources can sometimes lead to suffering. However, when life changes alter your group affiliations, the psychological benefits from your previous groups can still provide access to resources that support adaptation to change (Haslam et Authors, tion). Overall, life changes can be more fruitful and less harmful if individuals connect to positive social identities and derive benefits from them (Haslam et Authors, tion). Social Identity Theory primarily focuses on the motivations driving subordinate or stigmatized groups to change their group affiliations. However, it also sheds light on how changes in identity can influence our lives (Brown, tion). In contemporary societies, identities are multifaceted, particularly for minority groups. For example, many immigrants choose to maintain both their heritage culture and their new culture. When factors like religion and gender are added to this mix, the intersections of identity become numerous and intricate (Berry, 1997; Nguyen et Benet-Martínez, 2013; Schwartz *et al.*, 2014; Crenshaw, 1994; Deaux et Verkuyten, 2014). Recent research suggests that the source of meaning in our lives is often social, particularly within the groups we maintain and adapt to during periods of change. For instance, in cases of immigration or recovery from depression, maintaining multiple group memberships can be effective in preventing relapse(Haslam et Authors, tion).

2.3.3 Effects on life performance

According to identity theory, your identity can influence your role performance or behavior, emotions, physical and mental health (such as stress, anxiety, and depression), as well as aspects of your self-concept, including self-esteem, self-efficacy, and self-authenticity (Stets et Serpe, tion).

Social identity contributes to health and well-being by fostering social connections with in-group members, which provide trust, support, self-esteem, a sense of control, and agency (Haslam et Authors, tion). According to this concept, the more important group memberships an individual has, the better their psychological state will be due to the support provided by these memberships(Haslam et Authors, tion). Similarly, life changes can be more positive and less harmful if individuals have access to social identities from multiple group memberships, allowing them to gain and maintain positive group connections and resources(Haslam et Authors, tion).

Social identity can enhance self-esteem, so the more identities we have, the more social identity we can possess. This can contribute to well-being or reputation. However, there is a fact that as one identity becomes more salient, others may become less significant. To maintain multiple social identities derived from different groups, it is important that these groups have minimal conflict and reinforce each other, which can lead to greater satisfaction (Hogg et Authors, tion).

Research has shown that people tend to protect and promote their group status against other groups because group power is perceived as self-power, and the collection of different identities contributes to self-enhancement and self-esteem (Hogg et Authors, tion). In situations involving positive emotions, individuals feel they have exceeded their standards, while negative emotions arise when they fall short of these standards. In non-verified situations that lead to negative emotions, individuals may intensify their behavior to convince others to change their perception of this identity and to value it more highly (Hogg et Authors, tion).

2.4 An overview of sentiment analysis and opinion mining

This section is based on Bing Liu's book (Liu, 2020a). Liu has been doing research on sentiment analysis and opinion mining for over 20 years, and has published a highly cited book presenting the state of the art in 2012 (Liu, 2012b). This literature review is based on the second edition, published in 2020 (Liu, 2020a). The introduction (next) is based on Chapter 1 (Liu, 2020a); Section 2.4.2 is based on chapter 2. The document sentiment analysis section (Section 2.4.3 is based on Chapter 3 (Liu, 2020b), and the aspect-based sentiment analysis section (Section 2.4.4 is based on Chapter 5 (Liu, 2020c).

2.4.1 Introduction

Sentiment Analysis (SA), originated in computer science, studies opinions and sentiments expressed in textual data. Although the term opinion and sentiments are usually used interchangeably in the field, they have a subtle distinction. Opinions represent a concrete viewpoint while emotions resemble an expression. The rise of social media led to the availability of large-scale opinionated data causing the sentiment analysis tools to gain popularity in other fields of science such as management sciences, social sciences, and economics. They provide valuable insights for businesses, governments and individuals to gauge public opinion and customers to make more informed decisions.

2.4.1.1 Sentiment analysis applications

Sentiment analysis provides multi-faced benefits for businesses, governments and individuals. Businesses require the customers' feedback to maintain the quality of their products and services and maximize their profit. Furthermore, public opinions and sentiments help governments to make timely and more informed decisions regarding the policies and social concerns. Finally, the insights of sentiment analysis benefits individuals to make better political decisions and purchases. Sentiment Analysis has beneficial applications in various domains such as healthcare, tourism, finance, and political elections. For instance, they can be used to predict election outcomes or stock market trends. Sentiment Analysis has become a vital tool for decision-making in modern society. Big technology companies such as Google, Microsoft, Amazon, and platforms like Weibo and Yelp monitor public opinion and trends using sentiment analysis tools from the huge opinionated content made available.

2.4.1.2 Sentiment analysis research

The availability of large-scale online data has enabled sentiment analysis research to flourish to support various applications since 2000. It has attracted research work in other fields such as natural language processing (NLP), data mining, and management sciences. For example, in management sciences, the primary focus is the study of consumer opinions' impact on business, while in NLP the main focus is on developing more precise algorithms to extract sentiments from text and summarize them.

2.4.1.2.1 Different levels of analysis

There are three levels of granularity for sentiment analysis: document level, sentence level, and aspect based level. At the document level we focus on identifying the overall sentiment of the entire document (positive, neutral or negative). In sentence level, focus is on whether a single sentence is positive, negative or neutral and it is the most fine grained of the three. Aspect-based sentiment analysis explores an individual's sentiments about different aspects or features of the same object (product, service, political party, etc.). For example, the sentence "This camera has a high resolution However, the battery life could be better" has the positive sentiment about the quality of the picture and negative sentiment about the battery life. This level of analysis can give us more detail that the other levels can not provide.

2.4.1.2.2 Sentiment Lexicon and Its Issues

A simple–but not very effective–way to assess the sentiment expressed in a document, is to look for words that connote a positive sentiment, or conversely, a negative one. For example words such as "good" or "amazing" are positive while words such as "awful" or "terrible" are negative.

A *sentiment lexicon* is a dictionary of words or phrases containing positive or negative words. A sentiment lexicon can be used to analyze the sentiment of a sentence. However, in real applications, it might face challenges due to the complexities of language. One challenge is words implying both positive and negative depending on the context. The word 'sucks' is positive, if we are talking about a vacuum cleaner, and negative about anything else. Another serious challenge is sarcasm apparent in more political views which may complicate the sentiment identification. Furthermore, a sentence might contain sentiment words but not actually have a sentiment.

2.4.1.2.3 Natural Language Processing Issues

Sentiment Analysis is a subfield of the NLP field. Sentiment analysis is a restricted problem of NLP, which aims to identify the *polarity* of the sentiment of a piece of text (positive or negative), without attempting to provide a full understanding of the text. Proper NLP-based understanding involves many more complex tasks such as co-reference resolution (detection of different words referring to the same thing), negation handling (study of how negation influences meaning), and word sense disambiguation (understanding the meaning of a word depending on the context). Although sentiment analysis is a narrow domain in NLP, it could help in solving more broader problems of NLP.

2.4.1.2.4 Opinion Spam Detection

Social media has provided a platform where views can be presented anonymously. This provides the possibility that opinions or posts are generated by fake sources (opinion spams) and further propagated in the social media environment to manipulate public opinion (Harmon, 2004; Kost, 2012; Streitfeld, 2012a; Streitfeld, 2012b). *Opinion spam detection* is a major issue in social media, and it leverages NLP and behavioral data analysis to identify fake content. Platforms such as Yelp and Dianping use spam detection systems to ensure the trustworthiness of their online social medium (Liu, 2012a)

2.4.2 Sentiment Analysis: Research Issues and Problem Definitions

We start by defining the *sentiment analysis problem* (Subsection 2.4.2.1). In Subsection 2.4.2.2, we identify the different types of opinions.

2.4.2.1 Problem Definitions

Sentiment analysis can be defined as a structured methodology made for converting unstructured natural language text into meaningful and actionable insight (Liu, 2012b). It has a primary goal of analyzing subjective data to identify sentiment opinion targets and the relationship between them. The process involves categorizing opinion into interrelated sub-problems for facilitating the development of the robust and accurate solution. Sentiment is inherently subjective, unlike factual information, and is influenced by personal experience, different viewpoints and ideological stances.

2.4.2.1.1 What is an Opinion?

Liu (Liu, 2020a) defines an opinion as a quadruple (e, a, s, h, t) where:

1. e is the entity that is the target of the *opinion* or *sentiment*;
2. a is the aspect of the entity e that is the target of the opinion or sentiment;
3. s is the sentiment itself;
4. h is the *opinion holder*, in this case, "I" (the narrator); and
5. t is the time at which the opinion was expressed.

For example, in the sentence, "I like the quality of the pictures taken by my iPhone", e is my iPhone, a is the quality of the picture, s is positive, I am the opinion holder (h), and the time (t) is now.

For the purposes of our work, we will emphasize the importance of every single of these components for meaningful and accurate analysis. Indeed, the lack of any of these elements can lead to incomplete interpretation.

Finally, note that sentiments may be explicitly stated or inferred indirectly that require nuanced understanding. Distinguishing between these modes of expression is critical for an accurate extraction and classification.

2.4.2.1.2 Sentiment Analysis Tasks

Roughly speaking, the main tasks of sentiment analysis consist of identifying the five components of a sentiment that we defined previously: 1) identifying the entity, 2) identifying the aspect about which the sentiment is expression, 3) identifying (classifying) the sentiment itself, 4) identifying the opinion holder, and 5) identifying the time or temporal trends of the opinion.

Each one of these steps presents its own challenges, and has been addressed in the literature using a mix of NLP and machine learning techniques.

2.4.2.2 Types of Opinions

Liu proposed different classifications of opinions (Liu, 2020a), along different dimensions. One such classification contrasts *regular opinions* with *comparative opinions*. Regular opinions express sentiments directly about an entity or its aspects, while comparative opinions evaluate similarities or differences between entities. For example, "I like the quality of the pictures taken by my iPhone" is a regular opinion whereas "The quality of the picture on the iPhone 14 is much better than the quality of the picture on the iPhone 10".

Similarly, opinions can be subjective, showing personal views, or so called 'fact-implied', deriving sentiment from factual statements. For example "I liked the second song better than the first one" is subjective whereas "the iPhone 14 is more expensive than the Samsung A7" is factual. Liu proposed other classifications, each of which raises its own research challenges in reaching accurate sentiment analysis (Liu, 2020a).

In the following paragraphs, we examine *regular* versus *comparative* opinions, *explicit* versus *implicit* opinions, and *emotional* versus *rational* opinions.

2.4.2.2.1 Regular and Comparative Opinions

Regular opinions focus on sentiments about entities or aspects (explicitly or implicitly), and often is about highlighting benefits or addressing issues. Comparative opinions include evaluating multiple entities, often utilizing superlative or comparative language. For accurate sentiment classification and comparative evaluations, understanding these concepts is very important.

2.4.2.2.2 Explicit and Implicit Opinions

While explicit opinions clearly show sentiment and its target, implicit opinions require conclusion, because the sentiment or target is not directly stated. For managing implicit opinions as it relies on contextual understanding, background knowledge, and commonsense reasoning, we should handle challenges. Developing effective methods to tackle implicit opinions significantly improve the accuracy and depth of sentiment analysis.

2.4.2.2.3 Subjectivity and Emotion

In this paragraph, we explain emotional sentiment, distinguishing it from rational opinions. *Emotions* are defined as specific, intense, and short-lived feelings. Emotions are to be distinguished from *moods*, which are more diffused and non temporary. Sentiment analysis often prioritizes emotions according to their certain targets, but analyzing moods provides insights into sustained psychological states. Emotions are also categorized into primary, secondary, and tertiary levels that enable fine-grained sentiment analysis to specific applications (Liu, 2020a).

2.4.3 Document Sentiment Classification

2.4.3.1 Introduction

The main objective of document sentiment classification is to detect the general sentiment (positive or negative) of a document, for example a product review. This task is less fine-grained than aspect based-sentiment analysis, as it considers the document as one cohesive entity rather than multiple entities and aspects. This task is particularly useful in product review analysis where the sentiment of an opinion regarding a specific product needs to be identified. However, this model is not beneficial for longer and more complex documents such as blogs, social media posts, or forums where there are multiple opinion holders, expressing themselves about a variety of topics.

2.4.3.2 Sentiment Classification Using Supervised Learning

The dominant method in document sentiment classification is supervised learning approaches, because we often have labeled data sets. Indeed, most product reviews include a combination of text and star ranking. Thus, we can consider the star ranking as a labeling for the textual review.

Traditional machine learning models such as Naïve Bayes (NB) and Support Vector Machines (SVMs), when combined with good *feature engineering*, produced good results (Liu, 2020b). Example features include individual words, n-grams, words in a document with specific part-of-speech tags (e.g. adjectives), *sentiment shifters* (Liu, 2020b). However, these algorithms can be easily derailed by simple linguistic variabilities. Deep learning models are able to handle such problems by taking into account context and word order. Indeed, word embeddings, Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, are better able to handle the complexity of text in documents (Liu, 2020b).

2.4.3.3 Sentiment Classification Using Unsupervised Learning

Unsupervised methods rely on sentiment words and phrases that are expressive of the general sentiment of a text, when labels (e.g. stars) are not available.

One of the approaches is to leverage fixed syntactic patterns that are likely to suggest the sentiment expressed. This method extracts important word pairs such as adjectives followed by a noun in a piece of text. Then using the search engine data pattern, it calculates a measure called Sentiment Orientation (SO) for each phrase, and averages over all the phrases in the text to get a general value. A positive SO represents a positive sentiment, and vice versa. SO represents how more likely that a phrase often co-occurs with a positive reference word compared to how likely that phrase often co-occurs with a negative reference word. For example, ‘beautiful sounds’ are more likely to be visited near the positive reference word ‘excellent’ in a corpus of text rather than visited near the negative reference word ‘poor’.

Another approach uses sentiment lexicons. Sentiment lexicons are dictionaries of words that researchers in different domains have created and assigned positive or negative scores to words. This method sums up the sentiment score of each word in a piece of text to derive a general score. A positive score represents a positive sentiment and vice versa. Although this method is useful when there are no labels, there are a few challenges. One is having a bias toward positive sentiments as there are generally more positive words. Further, it requires adjusting for specific domains as a positive word in one domain might be negative in another domain. For example, ’dark’ is a negative sentiment word in general, although it might be a positive sentiment word for a movie review. Also, handling negations and intensity might be difficult. For example, a negative combination may represent a slightly positive sentiment. For example, the negative phrase ‘not bad’ connotes a (mildly) positive sentiment.

2.4.3.4 Sentiment Rating Prediction

Sentiment rating *prediction* refers to predicting the *rating* of a sentiment, as opposed to its *polarity*; for example, trying the predict the number of stars assigned, from 1 to 5. This is a more challenging problem than sentiment classification where only positive or negative labels need to be predicted. One approach considers the problem as a regression task: the relationship between features in the text and the rating must be learned to provide an estimate of the rating. Another approach considers it as a multi-class classification problem where there are multiple categories of intensities. However, the ratings are ordered whereas in a multi-class classification, there is no ordering between the classes.

A third method is *graph-based semi-supervised learning*. This method creates a graph where reviews are nodes, and two reviews are linked if they are similar. The strength of similarity is assigned as link weights, and determined via a similar measure such as number of co-occurrence words, cosine similarity, semantic similarity, distance-based (Chong *et al.*, 2020). The unlabeled reviews are assigned a rating using rating propagation techniques. More recent and advanced approaches like deep learning architectures with hierarchical attention mechanisms improved the prediction task accuracy. However, this method is more effective in longer reviews as it needs to build context by identifying key words and sentences to determine the sentiment.

2.4.4 Aspect-Based Sentiment Analysis

This section is a summary of (Liu, 2020c).

Aspect-based sentiment analysis (ABSA) advances sentiment classification from a general to a specific level by focusing on specific aspects of entities. ABSA corresponds to the quintuple formulation shown in Section 2.4.1.2, namely defining sentiment analysis as elucidating the quintuple (e, a, s, h, t), where e refers to the entity, a refers to the aspect, s refers to the sentiment polarity, h identifies the opinion holder, and t shows the time of the opinion.

Aspect-based sentiment analysis enables us to better understand "mixed reviews" by distinguishing overall positive impressions from nuanced criticisms of particular features. This ability is useful for applications like product reviews, when opinions about specific features of a product (e.g. “voice quality” or “battery life”) are important. In the remainder of this section, we present some of the research issues raised by ABSA, and describe some of the approaches used to solve them.

2.4.4.1 Aspect Sentiment Classification

Aspect sentiment classification identifies the polarity (positive, negative, neutral) of sentiments that are related to specific aspects. There are two main approaches: 1) supervised learning, and 2) lexicon-based approaches. Supervised learning uses advanced machine learning models like SVM and LSTM that integrate syntactic dependencies and semantic relationships to define textual features from the text. New methods like attention mechanisms and memory networks further enhance these models by considering complicated relations between aspects and contextual words. Lexicon-based approaches use sentiment lexicons and composition rules to assess polarity.

Supervised approaches require the availability of extensive labeled data. On the other hand lexicon-based methods are domain-independent and still need labor-intensive rule creation.

2.4.4.2 Rules of Opinions and Compositional Semantics

Linguistic termed "rules of sentiment composition" modulate sentiment polarity. These constructs include negation, modality and conjunctions like "but," which significantly impact sentiment interpretation. For example negations "not" invert sentiment polarity, in “not bad.” Modal verbs such as "should" and "might" make sentiment intensity and direction to expressing obligations or possibilities. Compositional semantics shows how discrete sentiment-bearing elements merge into comprehensive sentiment expressions, underscoring the complexity inherent in analyzing nuanced linguistic structures. These rules exemplify the necessity for sophisticated NLP methodologies. revise

2.4.4.3 Aspect Extraction

Aspect extraction deals with identifying the specific aspects referring to their entities discussed in a piece of text. This is particularly useful for business owners to know customers’ sentiment toward specific features of their products. Aspect extraction is a key step in aspect-based sentiment analysis to provide a more fine-grained analysis. Consider this example: “This phone’s camera has an amazing resolution, and the battery life is decent”. The aspects are “camera’s resolution” and “battery life”. Aspect extract is not always explicit as in the previous sentence. For example; “This laptop is slow”, implicitly states that the aspect is “speed” or “performance”. There are multiple techniques for aspect extraction such as rule-based methods, statistical methods, supervised methods, deep learning approaches, and unsupervised methods. Rule-based methods are based on linguistic patterns such as noun phrases (‘This laptop is light-weight”) or adjective-noun pairs (“light-weight laptop”). Although they are simple to use; however, they require clear linguistic patterns and are domain-specific. Statistical methods utilized the concept of frequency and co-occurrence of terms. One simple approach is to find words which frequently appear close to opinion words (“good”, “excellent”, “poor”, “expensive”, “durable”). The terms which appear often near the opinion words are likely to be the aspects within a domain. However, this approach has difficulty in finding implicit or indirect aspects, and sometimes random non-aspect words might appear close to opinion. Another statistical method is utilizing the Pointwise Mutual Information (PMI) to measure the strength of relation between two terms (the aspect and the opinion word). It starts by estimating PMI for opinion words; for each opinion word the terms with highest PMI's are selected as aspects. This approach requires a large dataset to calculate accurate probabilities and to provide less noisy aspects selection. Aspect extraction using supervised methods requires (1) a labeled dataset where each piece of text is annotated with a set of aspects and non-aspects existing in the training corpus, and (2) an engineered features set. The most popular method is sequential learning (sequential labeling) such as Hidden Markov Models (HMM) and Conditional Random Fields (CRF). HMM considers a hidden state (aspect and non-aspect) for each observation (token) in a piece of text. The goal is to predict the series of hidden states given a series of observations, which is based on two main assumptions. Firstly, each hidden state (aspect or non-aspect) at the current step is only dependent on the previous hidden state which is called markov property. Additionally, (2) the observation (token) is only dependent on its current hidden state. An observable state could be a pair of a token and its POS. The model tries to learn the parameters that maximize the observation probability to predict a series of hidden states for a given new observation. CRF tries to address the limiting assumption in HMM, that is observations on multiple steps can be dependent on hidden state at a specific step. HMM uses a feature function to model the conditional probability of the hidden states given the observation sequence. The feature function for each step is dependent on current and previous hidden state and observations from any step. The feature function would consider the tokens, POS, Short dependency path (syntactic relations among words), Word distance (distance between aspect and opinion word). Deep learning methods are highly helpful for aspect extraction as they learn and extract excellent features than other automated feature engineering approaches. The relationship among aspects and its context can be well represented from the learned feature representations. Nomours deep learning models have been proposed to extract aspects, and a few examples are given in the following. One approach is to combine a CRF feature extraction model with deep learning to use continuous word embeddings instead of discrete features. These models can jointly extract aspects and their corresponding sentiments. Another approach is to jointly model aspects and their sentiment relation as a bi-directional LSTM. An example of an unsupervised method is constructing aspect embeddings using neural word embeddings of selected aspects. These aspects have been filtered out using an attention mechanism to represent words which are likely to be aspects. Furthermore, an autoencoder identifies aspects by reconstructing input word embeddings.

2.4.4.4 Grouping Aspects into Categories

After aspects have been extracted, they need to be grouped into categories. This is critical to provide a summary of opinion analysis. Different words or phrases might resemble the same aspect. For example, “sound quality” and “voice quality” refer to the same aspect of a device. A simple approach is to use a lexical dictionary such as WordNet to find synonyms of the same aspect. However, this approach is a bit tricky due to several reasons. Synonyms might be highly domain specific. Also, antonyms might be used to describe the same aspect. Furthermore, grouping aspects together cannot be performed in an unsupervised manner as they are highly subjective. A set of aspects might be categorized in different groups depending on the user’s preference. There are a few approaches such as Dictionary-Based Approaches, semi-supervised learning, soft constraints, multilevel latent semantic association, Constrained-LDA. For example, a semi-supervised learning approach requires some seed labels (labeled data) as guidance. It uses Expectation-Maximization (EM) Algorithm to help in classifying unlabeled data, which has two steps (1) Expectation (E) and Maximization (M). This algorithm firstly initializes the model parameters by calculating the probability of assignment of labeled seed aspects. Then, the E-step calculates the probability of unlabeled aspects assignments based on the current parameters with the help of some prior knowledge. The prior knowledge could be aspects who share common words, semantic similarity or co-occurrence patterns. Then, the M step, iteratively updates the model parameters such that the likelihood of observing data is maximized.

2.4.4.5 Entity, Opinion Holder, and Time Extraction

Entity extraction is a similar task to the named entity recognition (NER) problem. Opinion target extraction method could be applicable to the entity extraction problem because opinion target might serve as entity as well. NER has been studied in many disciplines such as information retrieval, text mining, data mining, machine learning, and NLP all which lie into the information extraction problem. The main problem is to find all mentions or manifestations of an entity in a given corpus (called Entity Extraction Task). There are several unique entities; however each unique entity might have multiple different mentions and manifestations, and they have to be grouped together (called Entity Resolution Task). The main focus here is the task of Entity Extraction. There are two main approaches (1) rule-based, and more recently (2) statistical-based. Statical-based methods use HMM and CRF models which rely heavily on labeled data which is expensive and limited. The main semi-supervised approaches are PU learning (learning from Positive and Unlabeled examples) and Bayesian Sets. Both of these methods do not require labeled data, however they operate on seed labeled entity data. For example, the PU model is a classification problem where the goal is to classify entities extracted in an unlabeled set (extraction has been done using rule-based methods) either as positive (belonging to an entity category) or negative (irrelevant). The features are based on frequencies of the surrounding terms for each entity. The learning is performed using a Expectation-Maximization with Naïve Bayes Classifier which iteratively tries to update model parameters and re-estimated the probability of class assignment based on the given data. After the classification, the identified entities are ranked based on the probability of the class assignment to the seed set. Opinion holder (person's name) and time extraction is also similar to the task of NER. However, it is not useful to extract the name of authors and the posting time in the context of social media data; as they are visible and can be easily scraped. It is more useful in the case of extracting names and time from, for example newspapers posted online. One approach is to use a CRF model with features such as surrounding words, POS of surrounding words, grammatical roles, and sentiment words to identify a person's name and organizations. Others have utilized maximum entropy (ME) model, SVM, convolution kernels, dependency parser, automatic semantic role labeling (ASRL).

2.4.4.6 Coreference Resolution and Word Sense Disambiguation

Word sense disambiguation (WSD) and coreference resolution are two popular tasks in NLP in which their issues are being discussed in the context of sentiment analysis. Word Sense Disambiguation (WSD) is the problem of identifying the specific contextual sense (meaning) of a term used in a piece of text; as many terms might have different meanings in different contexts. For example; "cheap" can mean "low quality" or "affordable" in different contexts. Another idea is that terms can have varying senses depending on the grammar tag they appear in (subjective versus objective). This is called Subjectivity Word Sense Disambiguation (SWSD). For example, the word "bold" can be both used in a subjective sense, as in "She has a bold mind", and in an objective sense, as in "The font is bold". However, the meaning is different; confident versus design. Although there are lexicons which provide information about the subjectivity versus objectivity and sense of words; however, many words can be used in both senses. The correct sense helps in reducing sentiment prediction error. Initial approaches have leveraged contextual features (surrounding words) with supervised classification techniques to identify the correct sense using labeled data in a subjectivity lexicon to solve both WSD and SWSD. The results of SWSD were better than WSD which suggests that subjectivity helps in understanding the senses better.

Coreference resolution is a task to identify all the references to the same entity in a piece of text. Consider this sentence, “I bought an iPhone two days ago. It looks very nice. I made many calls in the past two days. They were great”. “It” refers to “iPhone” and “they” refers to “Calls”. This task is of high importance in the context of sentiment analysis, because people often express the sentiment not in the same sentence they have mentioned the exact reference entity (previous sentence). Furthermore, in comparative sentences; this task becomes critical to understanding the correct sentiment for its entity. Consider this example; “The Nokia phone is better than this Motorola phone. It is cheap, too”. In this example, “it” refers to “Nokia”, and in the second part, it expresses a positive sentiment for it. However, consider this sentence “The Nokia phone is better than this Motorola phone. It is cheap, too”. The pronoun “it” now refers to “Motorola” which in the second part, expresses a negative sentiment about it. Initial approaches to solve this problem have utilized supervised clustering techniques, and supervised learning with semantic features. However, deep sequential learning with attention mechanisms can provide excellent results in solving word sense disambiguation and co-reference resolution problems; as they are designed to pay attention to the context of the words in longer pieces of text.

2.5 Conclusion

This literature review has explored key areas relevant to this research. We first studied the marketing literature to see how *social identity* influenced consumer behavior. In particular, according to Bagozzi’s framework, *social identity* influences the first step of the customer journey: the desirability of various products and services (*goal desire* (Bagozzi, 2006)).

Next, we looked at the *social sciences* literature to understand the concept of *social identity*. *Identity theory* identifies three types of identities: *personal identities*, *role identities*, and *group identities* (see e. g. (Stets, 1995; Stets et Serpe, 2013; Stets et Serpe, 2016)). We learned that each type of identity embodies a set of values, and influences the holder’s behavior. We also learned that several identities can be brought to bear to a situation, and influence a person’s behavior in different ways. The question remains as to: 1) which types of identities can influence consumer behavior, and 2) which types of identities manifest themselves in writing.

To the extent that we need to analyze on-line writing to "guess" writers’ identities, we presented an overview of *sentiment analysis and opinion mining* techniques, which combine natural language processing and machine learning techniques to extract different pieces of information from natural language text.

We learned about the main challenges and research directions in sentiment analysis, and more specifically, *aspect-based sentiment analysis*, where we try to figure out how the author feels about a specific *aspect* of an entity, as opposed to the entity as a whole.

In light of the objectives of our research, and the research challenges and opportunities identified in the literature review, we present the methodology in the next chapter.

CHAPTER 3
METHODOLOGY

3.1 Introduction

This chapter outlines the methodology. Recall from chapter 2 that customer experience management relies, in part, on understanding the factors that influence consumers' choices (e.g. (Bagozzi *et al.*, 2007)). We saw in Section 2.2.1 that the purchase process starts with *goal desire*, which is the step where a consumer desires to acquire a product or service, and that product and service desirability is influenced by, among other factors, by what the marketing literature calls *social identity*. Thus, to the extent that a retailer or service provider can guess the social identity of a consumer, they can tailor their product offering or the information they communicate to the consumer in a way that is consistent with that social identity. Hence our interest in exploring the extent to which sentiment analysis techniques, and more generally, natural language processing techniques, can be used to detect the social identity or identities of a consumer, from their on-line writing.

Stets et al. identified *three* types of identities: *personal*, *role*, and *group* identities (Stets et Serpe, 2013; Stets et Serpe, 2016), all of which can potentially influence consumer choices. Thus, we need to first revise our research objectives based on this distinction. This is done in the next section (Section 3.2, where we explore, among other things, the relationships between Stets et al. *identity theory*, and the *social identity* referred to in the marketing literature. We will also speculate on which of Stets et al. *identities* are likely to have the greatest influence on consumer behavior. We identity *personal identity* as one, and we propose a general methodology for identifying personal identities in Section 3.3. We also identify *role identity*, as the other, and discuss the identification of role identities in Section 3.4. We conclude in section 3.5.

3.2 Different identities

The literature survey identified three types of identities:

- *Personal identities*, which are based on having a personal characteristic, like being athletic, a moral person. Personal identities appear to include what the marketing literature identifies as *secondary moral values* (Bagozzi *et al.*, 2007; Schwartz, 1992; Schwartz *et al.*, 2014).

- *Role identities*, which are based on the role that a consumer plays in life. That could be a teacher, a parent, a doctor, a student, or a manager.
- *Group identities*, which refers to the group or groups that a person identifies with, and *with which it undertakes group activities* (Stets et Serpe, 2013; Stets et Serpe, 2016). It can be a sports club, an alumni association, a neighborhood association, or a work-related social or sports club.

According to Stets et al. (Stets et Serpe, 2013; Stets et Serpe, 2016), each identity embodies a set of values and corresponding behaviors which align with those values. They argue that different identities or combinations of identities, may be activated depending on the situation and on the level of commitment of the individual to her or his identities (Stets et Serpe, 2013). Further, several identities can be activated at the same time (Stets et Serpe, 2013):

> Thus, taking the lead in the school chants and doing so more aggressively than others simultaneously verifies the group (school) identity, the role (leader) identity, and the person (aggressive) identity. Thus, we need to be mindful of the simultaneous activation of the different bases of identities in interaction.

Interestingly, none of these three types corresponds to our intuitive understanding of *social identity*, which corresponds more to what Stets et al. refer to as *social categories* (Stets et Serpe, 2013), for example based on income, religion, profession, etc. Stets et al. distinguish between *social identity* and *group identity*, the latter implying that the group organizes activities in which the members participate, and through which they seek confirmation of their belonging by adopting behaviors aligned with the values of the group. There are no such group activities in *social identities*, which can be thought of as statistical aggregations of socioeconomic characteristics.

For the purposes of this thesis, we will focus on Stets et al.'s types of identities (personal, role, group), for the following reasons:

- Stets et al.'s identity theory explicitly ascribes *values* to the various identities, and *behaviors* that are aligned with those values;

- Identity theory stipulates that people who "activate" specific identities in specific situations, *seek confirmation* of their identities from other people. This is particularly true for *group identities*;
- Identity theory also stipulates that people are unhappy when their projected identities are not confirmed by others (Stets et Serpe, 2013). This accounts for our intuitive understanding of "peer pressure", which compels us to behave like members of our group–including purchasing behavior;

Further, there are easier ways to detect *social identity*, as defined by sociodemographic characteristics, than to analyze online writing. For example, there a number of ways to guess a consumer's income bracket, including their exact postal address, their zip code, the type of articles they purchase, their purchasing history, their mode of payment, etc. Similar things can be said about gender, age bracket, lifestyle, etc.

The next questions we have to ask are:

- Which of the three types of identities (personal, role, group) are more likely to influence consumption choices; and
- Which of the three types of identities (personal, role, group) are more likely to manifest themselves in online writing

Recall that a *person identity* corresponds to how a person sees themselves, in a way that makes them unique. A person can be *health conscious*, and they will buy organic foods, and stay away from processed foods and foods with high content of sugar, salt, or fat. They can be *environmentally conscious*, and buy mostly bulk food and refrain from buying bottled water. They can be proponents of *fair trade*, and will buy products that are certified fair trade. They can be *athletic*, and they may buy protein bars and sports drinks.

A person plays different *roles* in their social and professional lives, and some of these roles may influence the person's consumption habits. In fact, the travel industry was one of the first to rely on the concept of *lifecycle stages* to target different travel products to adults at different life stages: couples with no children, couples with toddlers, couples with adolescents, empty nesters, retired, etc.; couples with toddlers are easy to recognize at the cash register: baby diapers and baby formula galore.

It is less clear how *professional roles* influence a person's consumption patterns:

- *Personal consumption* will be influenced by *personal role* and income level (*social category*). If you have toddlers, you will buy baby formula, regardless of your position at work!
- *professional consumption* is typically handled by the *purchasing department* of your employer.

Professional consumption is mostly visible for *tradespeople*, who buy their own work gear and parts related to their trade.

It is less clear how *group identities* influence consumer behavior. As explained above, what distinguishes a *group identity* from a traditional *social identity* is the members of the group involvement in group activities. We have thus to think of ways such group activities influence the consumption behaviors of members of the group. Members of a tennis club will be buying tennis gear. Members of a book club will be regular book buyers. Members of an organic food cooperative will be buying organic food. From these examples, it appears that the group identity (tennis club/book club) is linked to a person identity (tennis player/book reader), and thus, it appears that a group identity will influence consumption behavior only to the extent that the underlying person identity does.

Thus, for the purposes of this thesis, we will focus on using sentiment analysis and natural language processing (NPL) techniques to detect *person* and *role* identities.

Whether these identifies manifest themselves in online writing, and how they can be identified, will be discussed in the next two sections. Section 3.3 talks about uncovering personal identities; we talk about *role identities* in Section 3.4.

3.3 Uncovering personal identities

The question that we try to answer here is to devise a strategy to uncover personal identities. Like we mentioned in the introduction (see Section 1.3), our strategy will necessarily be data-driven, and will thus be heavily influenced by the kind of dataset we have, i.e. which kind of on-line writing we have, and whether the data set is labeled or not.

Addressing the problem in the general case is difficult, because personal identities are, themselves, of different kinds, and they can manifest themselves in different ways.

The marketing literature argues that *social identity* influences *second order moral values* and *self-evaluative standards*, which in turn influence consumer behavior (Bagozzi, 2006),(Bagozzi *et al.*, 2007) (see Figure 2.1). Thus, we will:

- Focus on personal identities that involve a *polar stance* (adhering or rejecting) relative to a social value or a marked lifestyle,
- Use *sentiment analysis and opinion mining* techniques to detect the polarity of that stance.

For the purposes of this thesis, we will use the example of *veganism* as a personal identity, for two reasons. First, people are known to have strong opinions about veganism, because of the moral undertones with issues related to ethical treatment of animals and environmental consciousness, on one hand, and political correctness and "wokeness", on the other, etc. The second reason is that this has been studied thoroughly in the literature, as shown in Chapter 4.

3.4 Uncovering role identities

Recall that role identities have to do with identities related to the roles that people play in their life, both personal and professional. Recall, from Chapter 2 (Section 2.3) that identities embody values, but also influence behavior in a way that concurs with those values. But what role-related behaviors can be visible through on-line writing?

It is our hypothesis that role identities manifest themselves in on-line writing in the way that people talk about what they do. What do managers talk about? they talk about hiring, setting goals for their teams, scheduling meetings, reviewing their subordinates, etc. Doctors talk about some of the symptoms and challenges of their patients. Parents of young children talk about (breast)feeding, baby formula, diapers, colics.

If we knew which vocabulary is associated with any given role, we could use an approach similar to the lexicon-based approach to sentiment analysis: the preponderance (to be defined) of a vocabulary associated with a role, within a person's online writing, will indicate that the person plays that role in real-life. However, we anticipate challenges with such an approach. First, like we said in Section 2.3, people will typically play different *role identities*. A manager who is a parent of a toddler will probably talk about the challenges of managing a team, and caring for a toddler. Then, there is the issue of the *vocabulary*: how do we know, *empirically*, which vocabulary is a associated with a given role. Finally, there is the issue of *audience*. A manager with a toddler will talk mostly about managing on a group chat related to managing, and may mention in passing the challenges of combining the demands of management with those of caring for toddler, and vice versa. In a work - family balance group chat, a person might talk equally about managing and raising a toddler.

Topic modeling is an unsupervised learning techniques that consists of identifying *topics* as sets of words that appear often together within a document collection (Papadimitriou *et al.*, 1998), (Blei, 2012). To the extent that people who play a particular role speak to the appropriate audience, the word clusters may well embody the set of concerns shared by people playing that role, i.e. the *role vocabulary* we mentioned above. However, topic models are "unlabeled", and human interpretation is needed to give a unifying theme to these word clusters: the *role* that the person plays.

Accordingly, we prefer to: 1) use supervised learning to "learn" role-specific vocabularies, and 2) to the extent that a person can play several roles, and those multiple roles can manifest themselves simultaneously, frame the problem as a many-class classification problem. Further details about the experimental approach, and the validation protocol will be presented in chapter 5.

3.5 Conclusion

In this chapter, we presented the methodology used for recognizing people's identities from their on-line writing, with a view towards customizing their consumer journeys. The social sciences literature distinguishes between different types of identities: 1) personal identities, 2) role identities, and 3) group identities. We argued that personal identities and role identities are more likely to influence consumption behavior. They are also likely to manifest themselves in people's social media writing (Section 3.2). Thus, we proposed strategies for uncovering personal identities (Section 3.3) and role identities (Section 3.4). Chapter 4 will present the strategy for personal identities; chapter 5 will present the strategy for role identities.

CHAPTER 4

UNCOVERING PERSONAL IDENTITIES - CASE OF VEGANISM

4.1 Introduction

Our goal is to devise a strategy to uncover personal identities. Like we mentioned in the methodology chapter (see Section 3.3), addressing the problem in the general case is difficult because there are different kinds of personal identities that can manifest themselves in different ways or simultanesouly.

As a "feasibility study" of personal identity discovery through the analysis of on-line writing, we choose to focus on personal identities that embody a value-system that people can either espouse or reject, and that involves some *polarity* that can find expression in on-line writing. For the purposes of this thesis, we will use the example of *veganism* as a personal identity, for two reasons. First, people are known to have strong opinions about veganism, because of the moral undertones with issues related to ethical treatment of animals and environmental consciousness, on one hand, and political correctness and "wokeness", on the other, etc. Thus, the people who are positive about veganism tend to be vegan, themselves, and the people who 'hate' veganism as a social trend, tend to be non-vegan. Also, it so happens that veganism has been studied thoroughly in the literature (see e.g. (Shamoi et Authors, tion), (Cooper et Authors, tion), (Teh et Yap, 2023), (Steger et Kashdan, 2009), etc).

We first present veganism, and give a survey of the literature on attitudes towards veganism (Section 4.2). Next, we provide an overview of the methodology (Section 4.3). In section 4.4, we present the data preparation and collection step of the methodology. Text processing is presented in Section 4.5. The presentation and evaluation of results is presented in Section 4.6.

4.2 Personal identity case study: Veganism

In a plant-based diet, individuals refrain from consuming animal products such as meat, fish, dairy, eggs, and cheese. Instead, they focus on foods like fruits, vegetables, beans, whole grains, legumes, and nuts. The primary motivations for adopting a plant-based diet include environmental protection, animal rights, and health benefits (Teh et Yap, 2023).

Because of the moral undertones of the first two motivations, "veganism" leaves no one indifferent. Thus, sentiment analysis techniques have been used in the past to evaluate people’s feelings about veganism. The study by (Shamoi et Authors, tion) used Twitter posts to assess public opinion about a vegan diet. This research showed that as the vegan diet is becoming more common, there is a noticeable shift towards a positive viewpoint among people compared to previous years, suggesting an increase in the number of people adopting the diet. This article provides evidence that the vegan trend is growing.

A study by (Cooper et Authors, tion) sought to determine the primary reasons for choosing vegetarian food among three main factors: ethics, personal health, and environmental concerns. The results showed that few adopters were motivated by environmental or sustainability concerns in choosing vegetarian food. Instead, most tweets focus on the health benefits and taste of the food.

The study by (Olavsrud, 2023), discovered that topic modeling can be effectively used to explore and analyze large datasets, providing meaningful insights into the content. In this study, 52,000 texts related to food, specifically vegetarian and vegan topics, were collected using a specific search query. The Latent Dirichlet Allocation (LDA) method was then used to create and model these topics. Another study by (Dalayya et Authors, tion) aimed to analyze social media data to understand human perception and sentiment regarding plant-based diets and cancer using the TF-IDF method. The results revealed that many people were hesitant to adopt a completely plant-based diet. It was suggested that this was due to a lack of apps dedicated to personalized diet planning and widespread myths and misinformation about a link between cancer and plant-based diets.

It appears, from this partial research survey, that diet related issues have been thoroughly studied in the literature. It further shows that natural language processing and machine learning techniques have been widely used to extract different pieces of information from the text, including how people feel about different kinds of diets (sentiment analysis), and the primary motivations for doing so (topic modeling).

4.3 An overview of the methodology

The methodology for this study involves a series of processing steps to prepare, train, and evaluate the models used in sentiment analysis. Below is a general overview of these steps:

1. Data Collection and Preprocessing: Tweets about the vegan diet were collected from Twitter and underwent preprocessing to remove noise, such as punctuation, stop words, and irrelevant characters. This step is essential to ensure that the data fed into the models is clean and optimized for feature extraction.

2. Data Splitting: The data was divided into training, validation, and test sets. For the machine learning models, an 80-20 train-test split was used, while the deep learning models used a 64-16-20 split for training, validation, and testing, respectively.

3. Text Vectorization: The text data was converted into numerical vectors. For machine learning models, TF-IDF was used, which assigns weights based on the importance of each word in the document relative to the entire dataset (Manning et al., 2008). For deep learning models, embeddings such as Random Weights Embedding, FastText, and BERT-large-uncased were utilized, which capture word meanings based on contextual relationships (Devlin et al., 2018; Bojanowski et al., 2017).

4. Model Training: The machine learning models (KNN, MLP, Decision Tree) and deep learning models (CNN, LSTM) were trained on the respective training sets. Hyperparameter tuning was conducted on the deep learning models using the validation set to optimize performance.

5. Evaluation: The models were evaluated using accuracy, precision, recall, and F1-score metrics. These metrics provided insight into each model's performance in classifying tweet sentiments, allowing for comparisons across different model types.

6. Result Analysis: The outputs from each model were analyzed to determine their effectiveness in sentiment classification. This involved comparing the performance metrics and identifying any patterns or insights related to sentiment trends in the tweets.

4.4 Data Preparation and Collection

To collect the data, the "Sentiment Dataset with 1 Million Tweets" was used. This dataset contains tweets labeled with four different categories: positive, negative, uncertainty, and litigious, which can be used for sentiment analysis. The dataset's columns include "Text," "Language," and "Label." The total number of entries in this dataset is 929,544. Table 4.1 shows the percentage distribution of each label category in the dataset.

Label	Percentage of Total Data
Positive	28%
Negative	28%
Uncertainty / Litigious	44%
Total	929544

Table 4.1 Percentage of Total Data

4.5 Text preprocessing

To begin, data are read from the dataset, and preprocessing is performed on them. Initially, tweets that contain at least one of the terms in the CV list are selected from the available tweets.

cv = ["vegan", "plant based", "healthy", "vegetarian", "veggie", "veganism", "cruelty free", "plant milk", "beyond meat", "vegan person", "vegan diet", "Organic", "Non-GMO", "Sustainable", "Eco-friendly", "Meat alternatives", "Whole foods", "Nutrient-dense", "Ethical eating", "Farm-to-table", "Raw food"]

Using the code shown in the below, rows from the dataset that contain at least one of the terms in the CV list are selected.

```
def check_word(txt):
    txt = txt.lower()
    for k in cv:
        if k in txt:
            return True
    return False
```

After filtering the tweets, 2,829 tweets related to the vegan diet were selected for sentiment analysis, and further preprocessing was conducted on this data. During the preprocessing stage, the following elements were removed:

- HTML tags
- Links and emails within the text
- Punctuation marks

- Extra spaces
- NULL values (with index resetting performed due to the removal of NULL values)
- Any words not formed from the 26 letters of the English alphabet or numeric characters
- Words shorter than three characters

Punctuation marks are symbols used in writing to clarify and convey the meaning of written language. These marks help indicate the structure, tone, and emphasis of a sentence. Table 4.2 shows these symbols. Each mark has a specific purpose, such as indicating exclamations, pauses, questions, sentence endings, quotes, and more in written communication. Removing punctuation can simplify the text, making analysis and content processing easier.

Row	Symbol	Name
1	.	Period
2	!	Exclamation mark
3	:	Colon
4	,	Comma
5	?	Question mark
6	""	Quotation marks
7	'	Apostrophe
8	()	Parentheses
9	-	Hyphen
10	...	Ellipsis
11	/	Slash

Table 4.2 Punctuation Marks

Furthermore, NULL values, sometimes referred to as missing values, indicate the absence of data in a specific field or column in a dataset. These values suggest that the data for that observation is either unknown or unavailable. Managing NULL values is crucial in data analysis and machine learning tasks, as they can affect the accuracy and reliability of results. Common approaches to handling NULL values include deletion (removing rows or columns with missing values), imputation (replacing missing values with estimated values), or treating NULL values as a separate category. In this implementation, NULL values were removed. Additionally, the following steps were taken in the preprocessing stage to better prepare the data:

- Converting uppercase letters to lowercase: This process was applied only to words where the first letter was uppercase and the rest were lowercase.
- Lemmatization: This involves converting words to their root or base form, representing the canonical form of a word in the dictionary. For example, the word "running" is converted to "run."
- Expanding abbreviations: Abbreviations were converted to their full form (e.g., "I'm" to "I am").
- Handling malformed characters: Malformed characters were corrected and converted to their proper form.
- Stop-word removal: The stop-word list was derived from two libraries: sklearn and NLTK.

The union of the lists generated by these two libraries was considered the final stop-word list. Stop words are common words that are usually filtered out during natural language processing tasks, like text analysis, to improve performance and efficiency. These words have little or no meaning in the context of analysis. Examples of stop words in English include "is," "the," "and," "of," "in," "a," etc. Since words generated after lemmatization may become stop words and negatively affect sentiment analysis, stop words were eliminated in two stages: before and after lemmatization.

After applying these changes, the data were stored in a column named "clean text." The final preprocessed dataset contains columns specifying the text and its corresponding sentiment label. The "Label" column can indicate negative, positive, or uncertain sentiments, reflecting the user's sentiment about veganism. The litigious label was removed from the dataset and was not used in the implementation. Figure 4.1 shows the final preprocessed dataset. Table 4.3 shows the count of data points for each label in the preprocessed dataset. The total number of data points after preprocessing was 2,829, which reduced to 2,565 after removing the litigious label.

4.6 Presentation and evaluation of results

The results of the used models are shown in Table 4.4 According to this table, the lowest accuracy is for to the KNN model, with an accuracy of 57.5% and the best result between machine learning models is for

Label	Count
Positive	(33.33%) 943
Negative	(24.25%) 686
Uncertainty	(33.1%) 936
Litigious	(0.093%) 264
Total	2565

Table 4.3 Count of data points for each label in the preprocessed dataset

	text	clean-text	label
0	Powerful call for #OceanAction by @UN SG @anto...	powerful oceanaction antonioguterres UNOC open...	positive
1	@VeganRecovering Do it, see if he notices! Pro...	veganrecovering notice probably best thing tasted	uncertainty
2	@StevenMV22 he's definitely a more realistic c...	stevenmv definitely realistic contract mcginn ...	litigious
3	@JonathanTurley Also Turley: "According to tha...	jonathanturley turley according account chance...	litigious
4	#GetFit #FatLoss #Fit See this extremely good ...	getfit fatloss extremely good webpage experien...	positive
5	@DragMeFromHell That's how I deciphered it but...	dragmefromhell deciphered stuff came boring ma...	uncertainty
6	These delicious, healthy Chipotle Black Bean V...	delicious healthy chipotle black bean veggie b...	positive
7	@unhealthytruth That wasn't no accident...	unhealthytruth accident	negative
8	Interactive elements like 'Yoga with Modi Ji' ...	interactive element like yoga modi healthy yog...	positive
9	The vegan beyond beef jerky isn't bad	vegan beef jerky	negative

Figure 4.1 Final preprocessed dataset

the decision tree model, reaching an accuracy of 88.5%. The multilayer perceptron model has the highest accuracy of 82.7% comes next. Figure 4.2, 4.3 and 4.4 provide classification reports for the machine learning models.

Among the deep learning methods, the best result of 88.7% accuracy is obtained when using the BERT model. The accuracy achieved from this model is very close to the decision tree model. The BERT model is known for its advanced ability in capturing complex patterns in textual data. However, decision tree model's efficiency and simplicity enabled it to attain accuracy comparable to the BERT model.
Following BERT, models using Random Weights Embedding, TF-IDF, and FastText achieved accuracies of 84.6%, 83.8%, and 83.2%, respectively. FastText has the lowest accuracy compared to other models. Fast Text's performance is such that this model embeds words based on their character n-grams. For this im-

Model	Precision	Recall	Accuracy	F1
KNN	0.654	0.575	0.575	0.568
Decision Trees	0.884	0.885	0.885	0.884
MLP	0.827	0.827	0.827	0.826
CNN+LSTM (Random Weights Embedding)	0.846	0.846	0.846	0.844
CNN+LSTM (FastText)	0.833	0.832	0.832	0.832
CNN+LSTM (bert-large-uncased)	0.888	0.887	0.887	0.886
CNN+LSTM (TF-IDF)	0.837	0.838	0.838	0.837

Table 4.4 Model Implementation Results

plementation, FastText was used as uni-gram and bi-gram. However, FastText performs worse, compared to other models. The reason for this poor performance is due to the breaking words into single and double characters does not create distinctive features which resulting in lower accuracy compared to others. However, this method achieved higher accuracy, compared to the multilayer perceptron method.
This higher accuracy is due to the use of deep learning models (CNN+LSTM). When using the TF-IDF method and deep learning models for tweet classification, higher accuracy is obtained compared to using the FastText method with deep learning methods. In other hands, TF-IDF method can give better accuracy than using the FastText method. Figure 4.5,4.6 ,4.7, and 4.8 show classification reports for the deep learning methods.

Figures 4.9, 4.10, and 4.11 show the confusion matrix related to machine learning methods.

4.7 Conclusion

Recall from Section 2.2.1 that the marketing literature has shown that people's *social identity* influences their purchasing decisions. To the extent that is true, it becomes important to be able to recognize people's social identity from their on-line writing.

We saw in Section 2.3 that there are several *identities*, including *personal identities*. We also saw in Section 3.2 that *personal identities* are likely to: 1) influence purchasing behavior, and 2) manifest themselves in on-line writing.

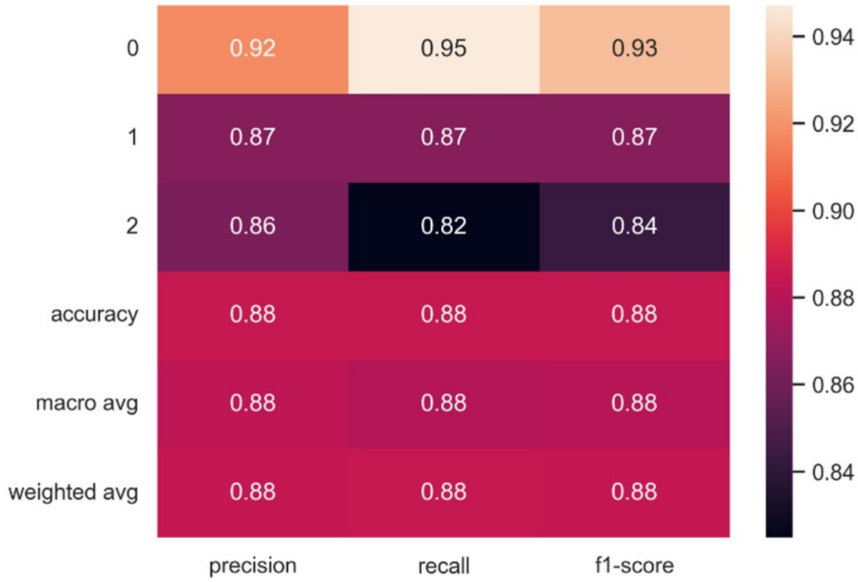


Figure 4.2 Classification Report for Decision Tree Model

Because there are many personal identities, and because individuals may 'activate' many identities at any given in time, we argued in Section 3.3 that detecting personal identities in difficult, in general; however, by focusing on personal identities that involve taking a *polar stance* (*for* or *against*) with regard to some position, we could use *sentiment analysis techniques* to determine somebody's member to the corresponding personal identity.

In this chapter, we applied this approach to the *vegan* identity, by trying to determine how they feel about *veganism*. To this end, we use the Kaggle dataset "Sentiment Dataset with 1 Million Tweets" (see `https://www.kaggle.com/datasets/tariqsays/sentiment-dataset-with-1-million-tweets` and used KNN, MLP, and different CNN+LSTM combinations (random weights embedding, FastText, BERT-large-uncased, and TF-IDF). We achieved an accuracy of 88.7% with CNN+LSTM (bert-language-uncased), and a close 88.5% with decision trees. This proves the soundness of the approach. However, it applies only to personal identities that are strongly correlated to polar positions people might have about 'something'.

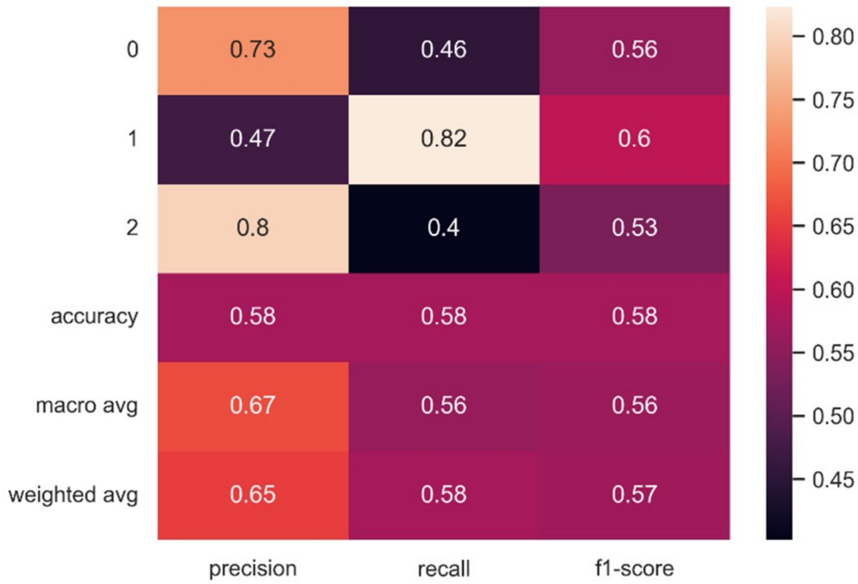


Figure 4.3 Classification Report for K-nearest Neighbors Model

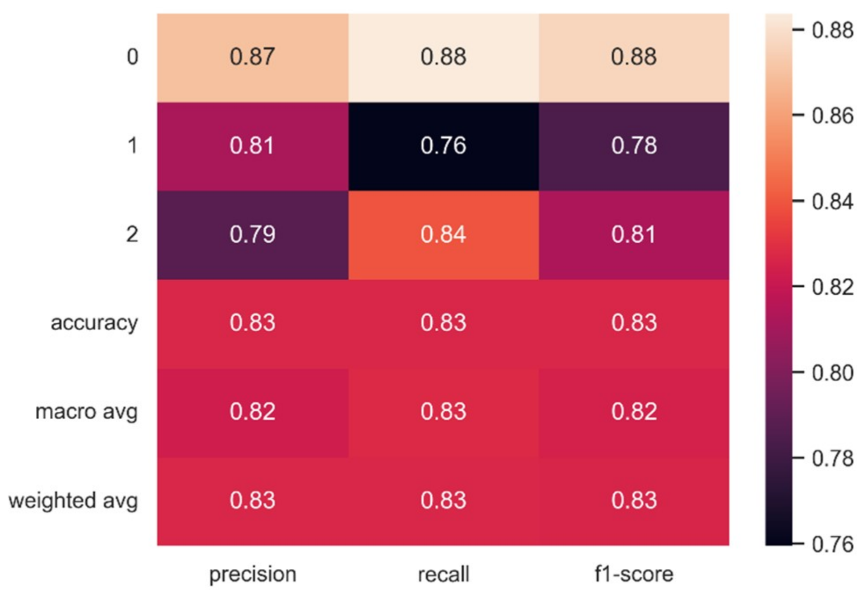


Figure 4.4 Classification Report for Multilayer Perceptron Model

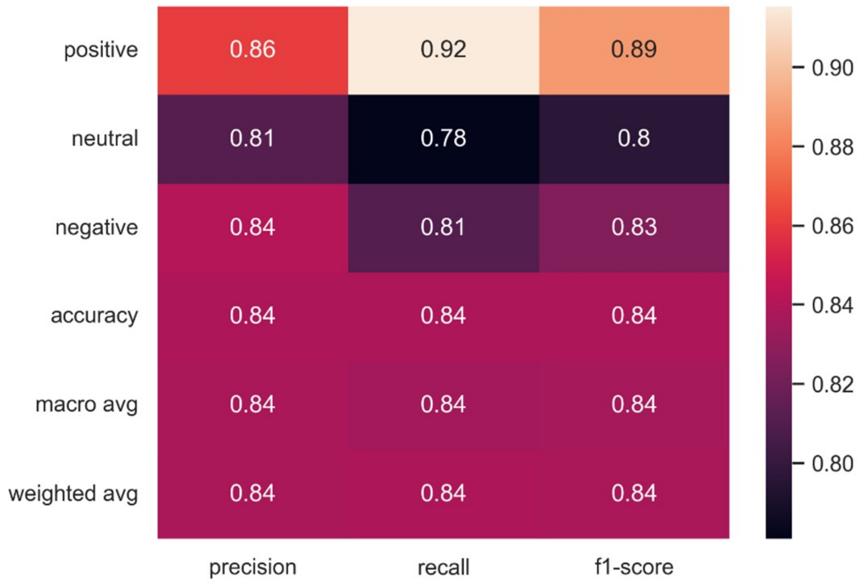


Figure 4.5 Classification Report for CNN+LSTM (TF-IDF) Model

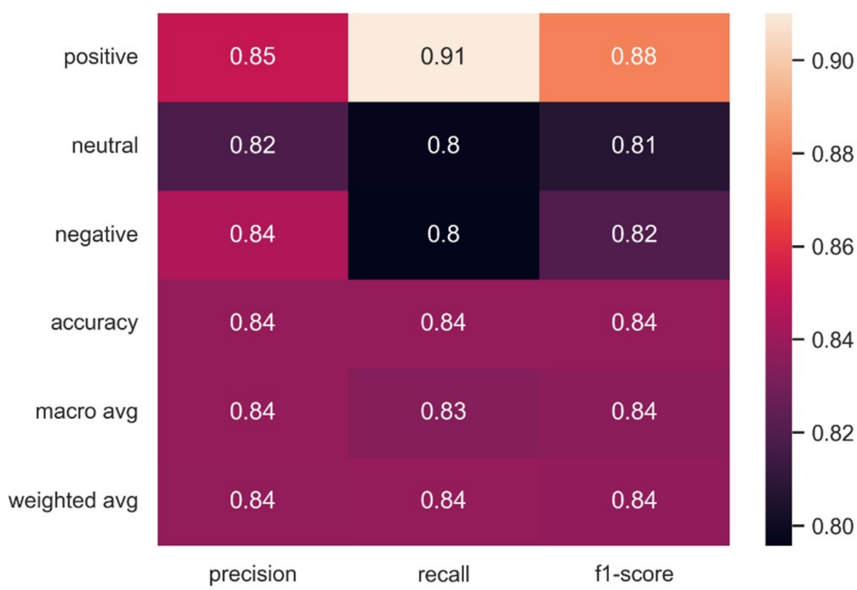


Figure 4.6 Classification Report for CNN+LSTM (Random Weights Embedding) Model

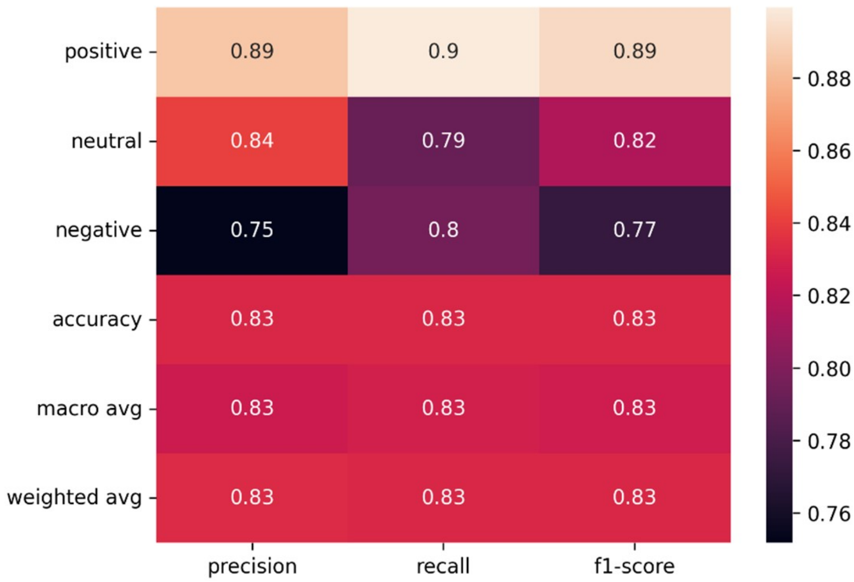


Figure 4.7 Classification Report for CNN+LSTM (FastText) Model

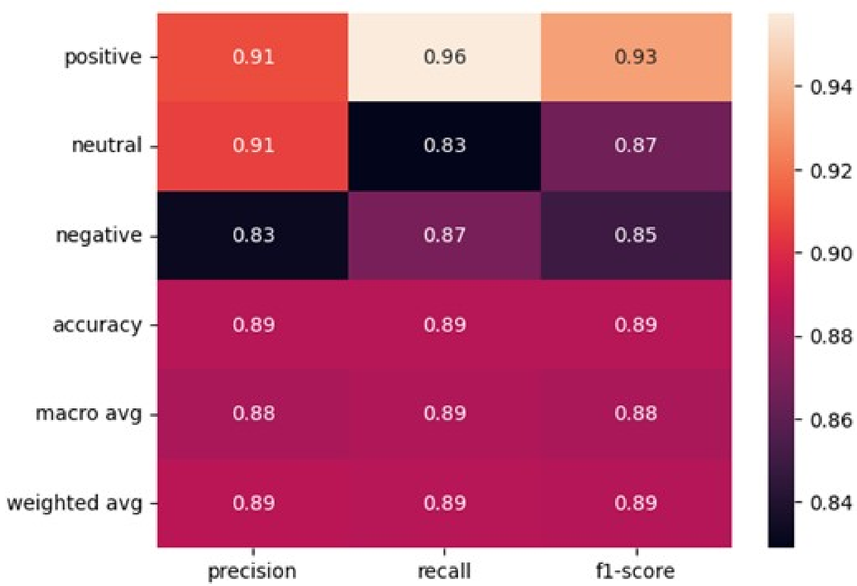


Figure 4.8 Classification Report for BERT Model

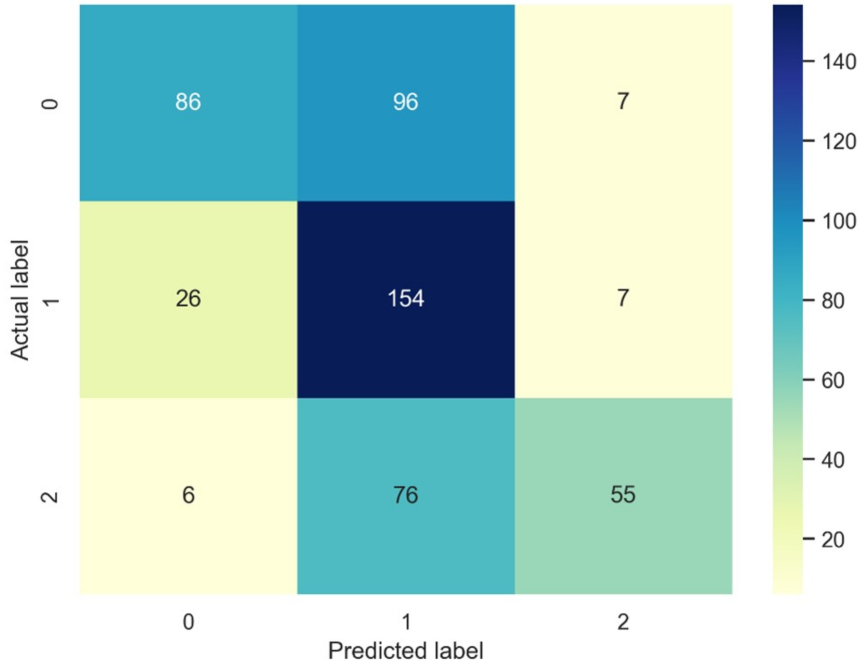


Figure 4.9 Confusion Matrix for KNN Model

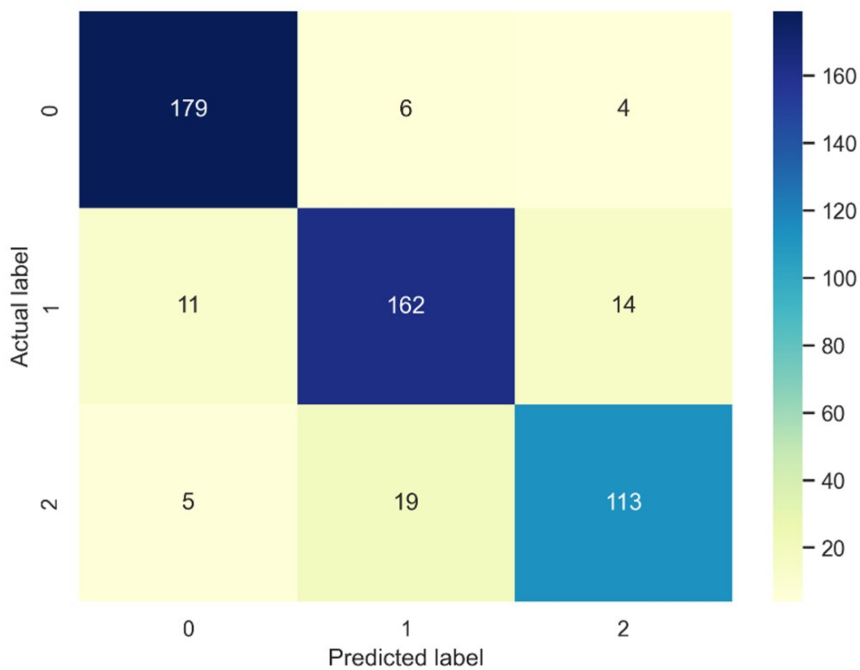


Figure 4.10 Confusion Matrix for Decision Tree Model

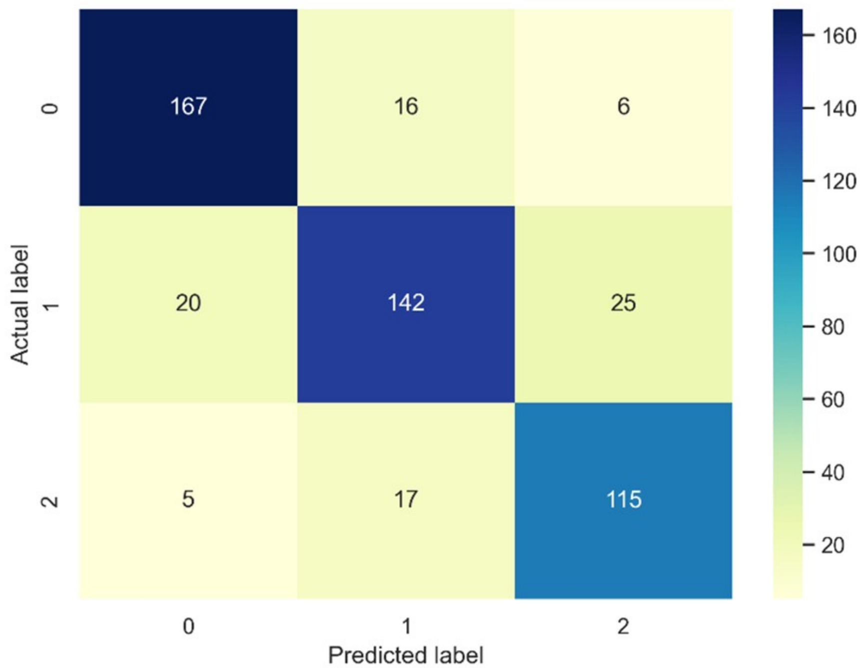


Figure 4.11 Confusion Matrix for Multilayer Perceptron Model

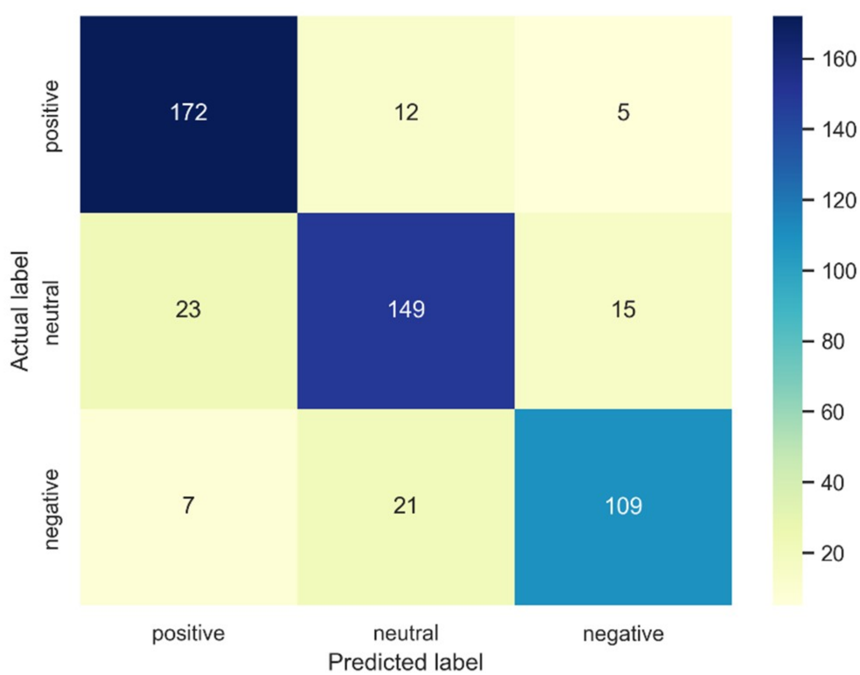


Figure 4.12 Confusion Matrix for CNN+LSTM (Random Weights Embedding) Model

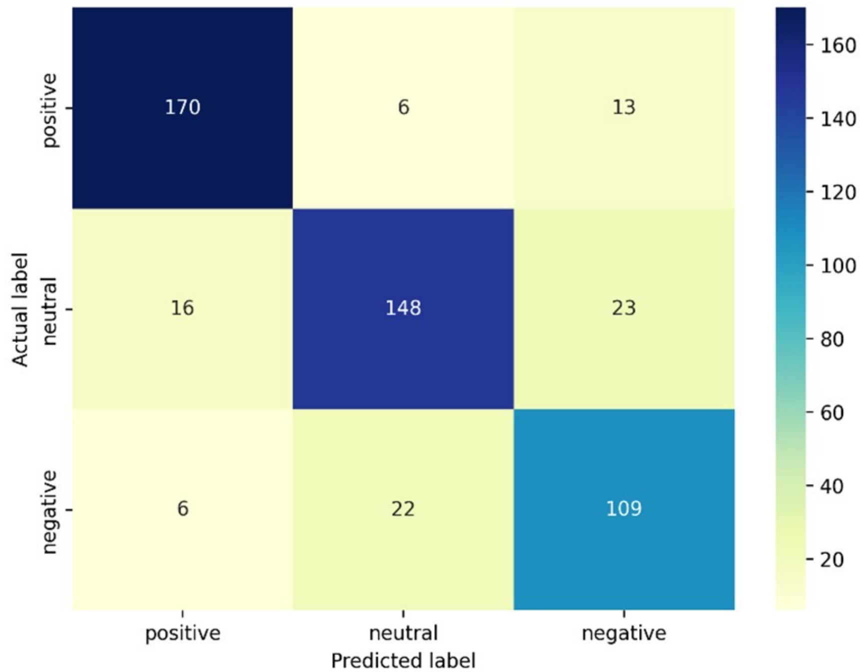


Figure 4.13 Confusion Matrix for CNN+LSTM (FastText) Model

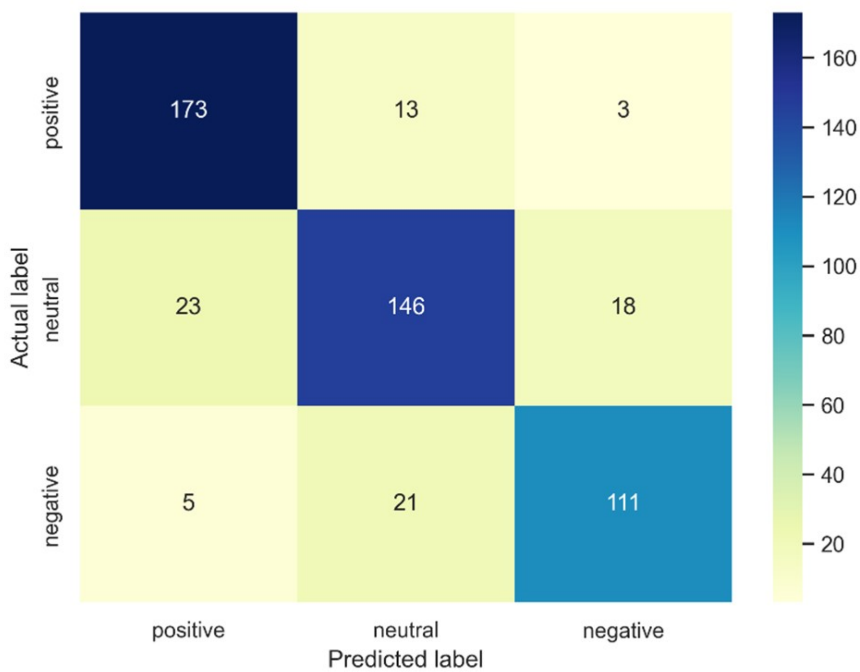


Figure 4.14 Confusion Matrix for CNN+LSTM (TF-IDF) Model

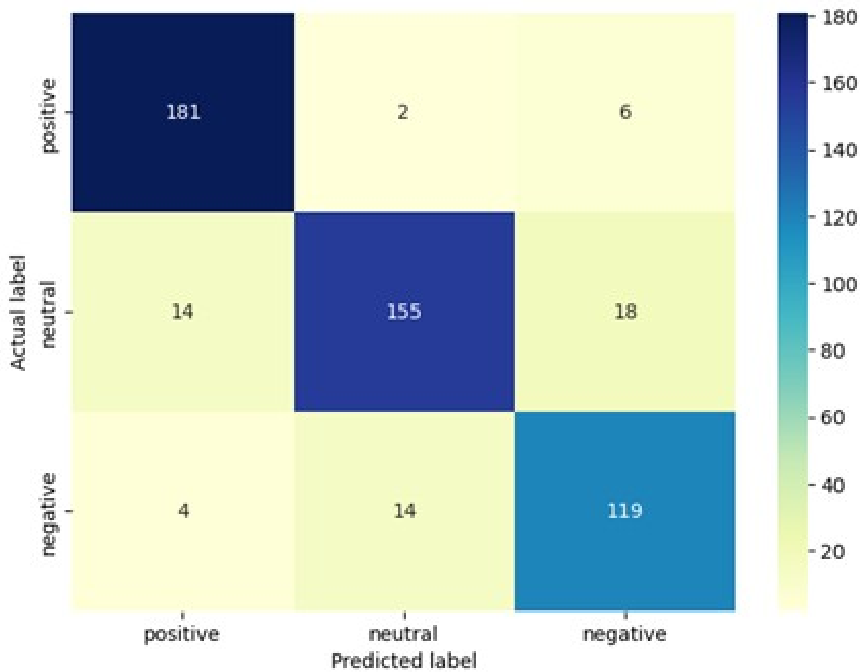


Figure 4.15 Confusion Matrix for CNN+LSTM (BERT) Model

CHAPTER 5

IDENTIFYING ROLE IDENTITIES

5.1 Introduction

Recall from Section 3.4 that our hypothesis was that role identities manifest themselves in on-line writing in the way that people talk about what they do. Each role comes with a specific vocabulary that embodies the concerns and issues that people playing those roles deal with on a daily basis. However, to the extent that people play different roles *simultaneously*, multiple roles might manifest themselves in people's online writing, depending in part on the audience (Section 3.4). We proposed in Section 3.4 to: 1) use supervised learning to "learn" role-specific vocabularies, and 2) frame the problem as a many-class classification problem. In this Chapter, we present our implementation of the strategy outlined in Section 3.4.

Section 5.2 (next) presents the methodology, where we describe our way of operationalizing the strategy outlined in Section 3.4. Section 5.3 presents the implementation, where we discuss dataset collection (Section 5.3.1) and data preparation (Section 5.3.2) and also Presentation and evaluation of results in section 5.4 present the results. We conclude in Section 5.5.

5.2 Methodology

As explained in Section 3.4 and summarized above (Section 5.1), we will use a similar strategy to the one we used for personal identities. Indeed, our approach will consist of validating the "presence" of some predetermined role identities within a person's on-line writing, as opposed to having those roles identities "emerge" through some unsupervised learning technique. For this reason, we framed our problem as one of multi-class supervised classification problem.

The first question that we need to address is related to the training data sets: how to obtain properly labeled data sets. Assume that we want to validate the presence of roles "manager" and "parent of a toddler". We need posts by people talking about their roles as managers, labeled as "manager"; and posts by people talking about their roles as parents of toddlers, labeled as "parent of a toddler".

One strategy for obtaining such a data set is to:

1. Get a hold of a dataset of posts by people
2. Label the posts by their topic, as "concerns of a manager" or "concerns of a parent of toddler"
3. *Independently* ascertain the roles played by the people who authored those posts.
4. Consider the posts by managers (parents of toddlers) that were labeled as "concerns of a manager" (concerns of a parent of toddler) as positive instances, and consider all the others as negative.

Step (2) may involve varying degrees of manual labor, depending on whether we use few shot learning methods or not, and step (3) is next to impossible: how can we *independently* (from their writing) determine the roles internet users play in real life? Add to that the potential difficulty of considering posts that talk about manager concerns but that are not written by managers, as negative instances. I could be a non manager, relating the experiences of my manager spouse, using exactly their terms.

Instead, we choose a different approach. Find datasets of posts by:

- people who are *known* or *assumed* to be managers,
- people who are *known* or *assumed* to be talking about management concerns.

and use such data sets for supervised training.

How do we find such datasets? In fact, there are lots of interest-specific *social networks* mediated with social media where people with similar interests or concerns exchange stories and tips. Reddit, for example, has a number of so-called *communities* that focus on specific topics, to which people interested in those topics can subscribe and exchange about issues and concerns of interest. Figure 5.1 shows the description of the "parenting" Reddit community with 8.2 million members, 564 of which were online at the time this picture was taken.

Figure 5.2 shows a sample of business-related communities that cater to people interested in, or involved in running businesses, with community sizes ranging from 2.1 thousands to 21 thousands members; orders of magnitude smaller than the parenting community. It is important that we pick a community with a relatively narrow focus. The r/parenting community is too broad for our purposes. Interestingly, the home page for

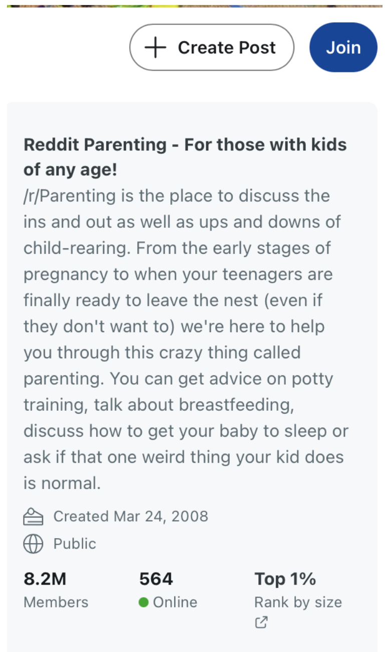


Figure 5.1 A description of the reddit parenting "community"

that community lists four specialized Wikis, including `Wiki - Early Years`, `Wiki - Older Kids`, and `Wiki - Crisis`.

Once we have resolved the labeled data set issue, we have to keep in mind that a parent with an early years child may post about the challenges of starting-up a business in the `r/parenting` community, and vice-versa: the owner of a start-up business may talk about the challenging of starting up a business, with a toddler in-tow. This means that posts from the `r/parenting` community, may be legitimately labeled with both "Parent of a toddler" and "Start-up business owner". The same is true for posts from the "Start-up business owner". Thus, our problem is one of *multi-label classification* problem.

One way to handle this is to train different binary classifiers, one for each label, and to apply them to an online user post, in no particular order. This is called the *binary relevance* method (Read *et al.*, 2011). This raises the issue of choosing the appropriate dataset as negative instances for the various binary classifiers. To use the example of Reddit communities, if the choice is between n role identities $RI_1, ..., RI_n$ corresponding to n communities C_1,..., C_n, we could use posts to community C_j as positive instances for role identity RI_j, and posts to the other communities as negative instances. However, such a naive approach goes counter to the multi-label approach. Instead, we should use, as negative instances, posts in communities that are not likely to intersect with the positive instances. For example, raising a child and building a

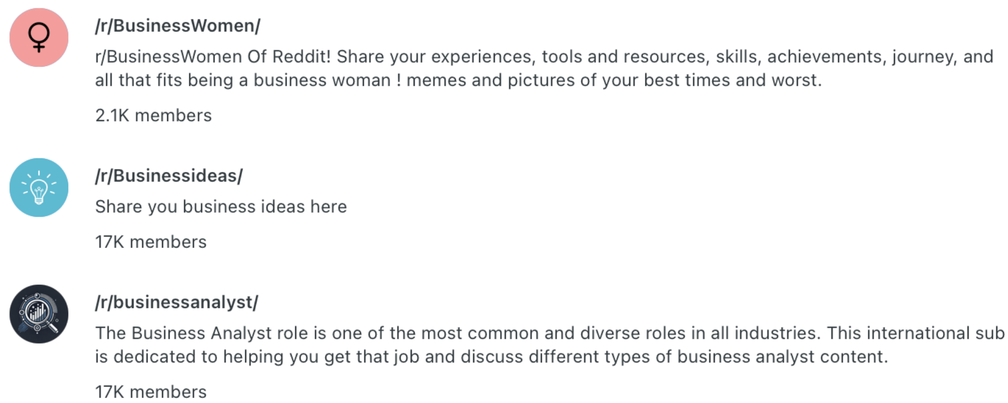


Figure 5.2 Some Reddit business-related groups

business are two important activities that compete for a finite resource: the undertaker's time. This makes it plausible that people who happen to play both roles will talk about both sets of issues in their posts. However, this is *less likely* between `r/NintendoPS4` and `r/BusinessWomen`, say.

5.3 Implementation

Per the methodology discussed previously, we choose to test our approach on two communities: 1) one community where *parents* talk about their children, and 2) one community where *children* talk about their parents. The likelihood of the same person talking their children and parents is very slim, and if they do, they will probably talk about their (probably elderly) parents in different terms from every other *child*.

We first discuss data collection (Subsection 5.3.1), then data preparation (Subsection 5.3.2). We present the training and testing of the two classifiers in Section

5.3.1 Dataset Collection

As mentioned above, we used two datesets for these project:

1. The Dataset derived from the research study "Accuracy of online discussion forums on common childhood" ailments
2. Subreddit called toxicparents from Reddit website

Reddit hosts a number of *communities*. Users on Reddit participate in topic-specific forums called subreddits (for example, r/science or r/funny). Users publish posts, share opinions, and vote on content in these subreddits.

Reddit has a free API for developers, making it possible to access detailed information about posts and comments. In this project, the API was used to collect data related to family members' perspectives toward each other, including parents' views of their children, children's opinions of their parents and grandparents. These comments were directly collected from Reddit.

ID	subreddit	post_title	comment	score	created_utc	comment_author	permalink
0	r/teenager	[mod] REMINDER ABOUT WITCH HUNTING POSTS.	Reddit mods (the ones for the whole app not yo...	17	2025-05-04 14:38:29	nermalnormal	https://reddit.com/r/
1	r/teenager	[mod] REMINDER ABOUT WITCH HUNTING POSTS.	alhamdulilah thank you these have been so anno...	23	2025-05-03 18:07:50	AwesomeeeeeeeeeAcc	https://reddit.com/r/
2	r/teenager	[mod] REMINDER ABOUT WITCH HUNTING POSTS.	Oh shoot the mods came hopefully they didn't s...	4	2025-05-04 07:02:59	PureProfessional144	https://reddit.com/r/

Table 5.1 Reddit comments from r/teenager

Parent Dataset

This dataset is derived from the research study "Accuracy of online discussion forums on common childhood ailments" (Farrell, 2018). The objective of the original study was to evaluate the accuracy of health advice provided in online discussion forums aimed at parents of young children, and to determine whether these forums serve as potential sources of misinformation.

The data were collected from publicly available parenting forums, specifically BabyCenter.com and WhatToExpect.com, both of which host large communities of parents discussing common childhood ailments. Using Google searches for five common childhood illnesses, the original researchers identified forums that appeared frequently across search results and met inclusion criteria. Posts from a six-month period were manually copied into spreadsheets and saved as CSV files for analysis.

Child dataset

The child Dataset was collected using the free Reddit API. The data were gathered from the r/toxicparents subreddit, which features discussions of people struggling with toxic parents. This subreddit was selected because it is one of the active Reddit communities on the topic, providing a sufficient number of user comments for analysis. In this subreddit, we can find posts and comments that reflect children's opinions and perceptions of their parents as discussed within the community. The collected text represents online expressions of children's perspectives on parental behavior, relationships, and emotional experiences.

5.3.2 Data Preparation

After extraction our 2 datsets that was gathered separately we combined them to create dedicated datasets.

A new dataframe had two main columns:

- comment: the text of the user's comment.
- role: the family role of the commenter, categorized as Parent and Child

From this point, preprocessing began to prepare the data for modeling. First, we performed basic taxtuel preprocessing consisting of:

- Removing URLs
- Removing special symbols

- Removing numbers
- Stripping extra spaces

Next, we performed additional filtering: included:

- We removed duplicates: we used the post title field to identify duplicates.
- We removed entries that had empty comments field.
- We deleted the entries where the post title and comments fields had fewer than 8 and 10 words, respectively.

ID	role	comment
0	child	Reddit mods the ones for the whole app not you...
1	child	Oh shoot the mods came hopefully they didn't s...
2	Child	Bro can you please ban the teenagers posting t...

Table 5.2 Sample of role-based comments from child dataset.

5.3.3 Analysis of frequent words in each class

5.3.3.1 Parent class

The WordCloud analysis shows that the most common words are baby, mom, know, will, time, need, want and help. These represent that parents' focus is on children, caregiving, responsibilities, and the need for support. We can see that words related to time (month, week and hour) and every activity verbs (sleep, go , put, try) are frequently being used by parents. This notes that parents often share their daily experiences.

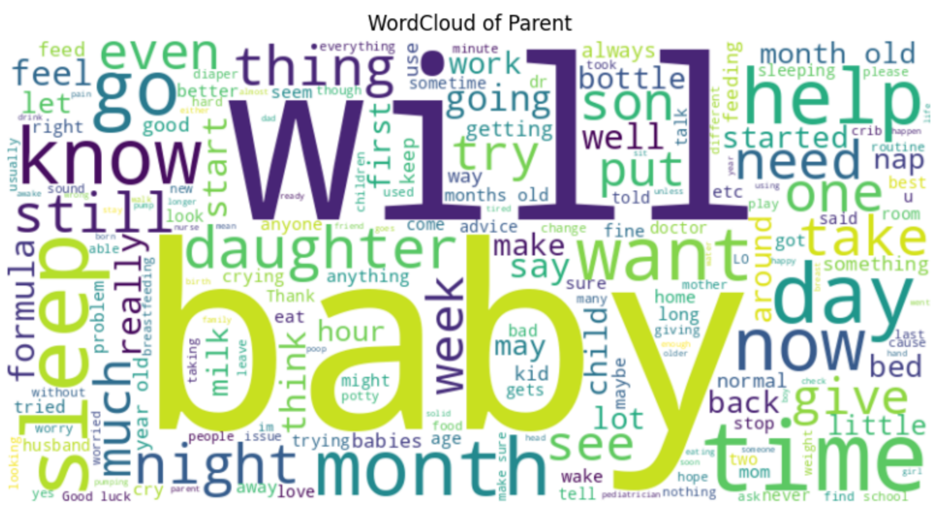


Figure 5.3 WordCloud of parents

5.3.3.2 Child class

The WordCloud analysis for children is given in the Figure below. We can see that the most frequent words were mom, dad, want, know, need, time, help, sleep, and go. The high occurrence of parent, mom, mother, dad shows that the conversations of the children category are mostly about parents. Children are using words such as want, need, help and try frequently which represents that they require independence, support or problem-solving. Many posts convey personal struggles and emotional challenges indicated by the high occurrence of emotional words (feel, say, think, cry).

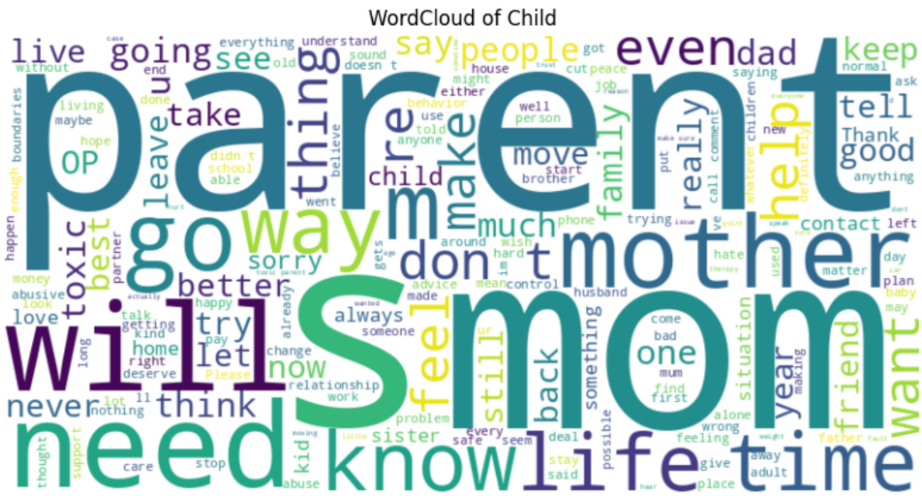


Figure 5.4 WordCloud of Child

5.3.4 Class distribution

Each family class (Parent, Child) included exactly 1,749 samples which represents a balanced dataset. The class balance is of great importance as imbalanced datasets could influence the model's prediction to favor one class over the others.

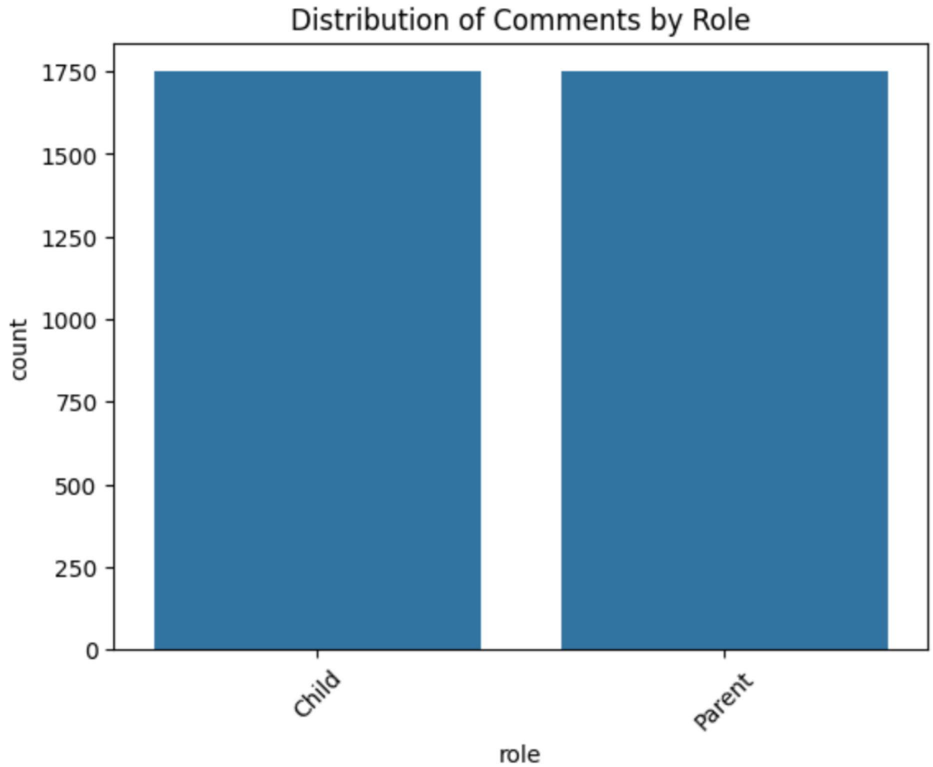


Figure 5.5 Distribution of Comments by Role

5.3.5 Modeling for Classification

The BERT model classifies the identity category of the commenters. It was implemented using the PyTorch and Hugging Face's transformer library. More specifically, the BertTokenizer tokenized the data, and the BertForSequenceClassification model did classify the identities. DataLoader efficiently manages the batches during training. The loss function is CrossEntropyLoss alongside the AdamW optimizer. The weights of the model were frozen during evaluation, while a softmax layer has been used to compute the probabilities of each class membership. Both the training and the evaluation stages have been implemented using Pytorch tools.

5.3.6 Results

The model was trained considering the following parameters:

- Batch size: 16
- Learning rate: 0.00002
- Weight decay: 0.01
- Training epochs: 3

In total, 525 steps were completed. The average training loss was 0.1804, with a total runtime of about 776 seconds ($\approx$ 10.8 samples/second).

Figure 5.3.6 shows the performance per class:

```
TrainOutput(global_step=525, training_loss=0.18046800340924946, metrics={'train_runtime': 776.2854, 'train_samples_per_second': 10.813, 'train_steps_per_second': 0.676, 'total_flos':
2208554198691840.0, 'train_loss': 0.18046800340924946, 'epoch': 3.0})
```

Performance per class:
- Child → Accuracy: 96%, Recall: 96%, F1-Score: 96%
- Parent → Accuracy: 96%, Recall: 95%, F1-Score: 96%

In general the model reached a 96% across Accuracy, Precision, Recall, and F1-score. Only 29 samples out of 700 were classified incorrectly representing a strong generalization without a bias toward any class. The ROC curve with AUC of 0.99 validates the strong separation ability of the model between the positive and negative samples. The PR curve also achieved AUC-PR of 0.99 which represents high precision and recall even in potentially imbalance scenarios.

5.4 Presentation and evaluation of results

Based on the evaluation metrics (Accuracy, Precision, Recall, F1-score), the model proved highly capable of distinguishing between parents' and children's comments. The Confusion Matrix also confirmed balanced classification, with no bias toward either class.

Together, the ROC and PR curves verified the robustness of the approach, confirming that deep learning–based models like BERT are effective for analyzing text and classifying family roles from user-generated content.

	precision	recall	f1-score	support
Child	0.96	0.96	0.96	375
Parent	0.96	0.95	0.96	325
accuracy			0.96	700
macro avg	0.96	0.96	0.96	700
weighted avg	0.96	0.96	0.96	700

Figure 5.6 Effectiveness of model

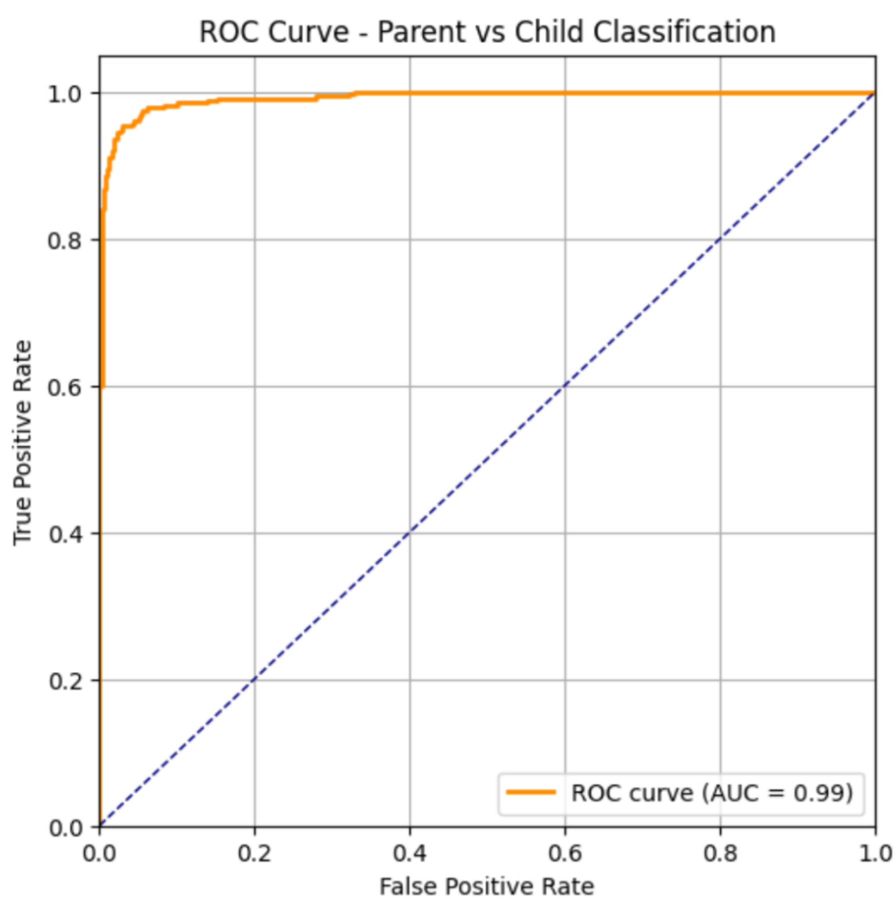


Figure 5.7 ROC Curve - Parent VS Child Classification

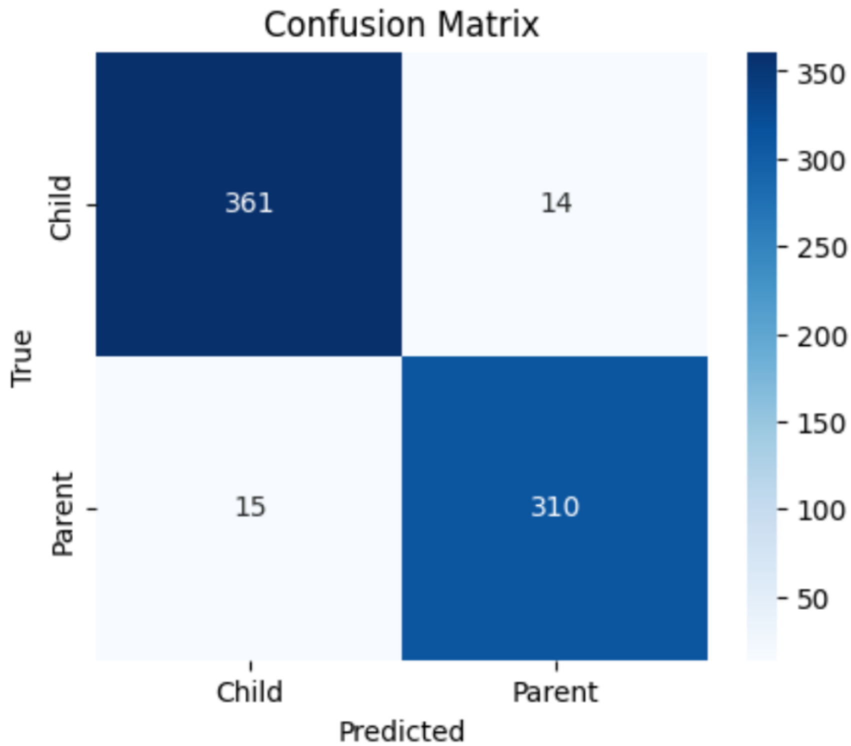


Figure 5.8 Confusion Matrix

5.5 Conclusion

We saw in Section 2.3 that there are several *identities*, including *role identities*. We also saw in Section 3.2 that *role identities* are likely to: 1) influence purchasing behavior, and 2) manifest themselves in online writing. Role identities may be reflected in the things people talk about it. A manager will talk about employees, hiring, evaluating employees, scheduling meetings, progress reports, and the like. But if the person just had a baby, they might talk about the challenges of raising a baby, and doing the managing. And depending on the audience, they will talk more about managing, or more about raising a baby.

If we had a predetermined lexicon associated with a role, we could use simple NLP techniques that look for words from that lexicon in people's writing to guess their social identity. However, there is no such vocabulary, and given the number of potential role identities, it is impractical to build or manage lexicons, one per role.

If we had a dataset of posts by people, that were tagged by the *roles* that people talk about in those posts, we could use a supervised learning approach to train a classifier, and then use that classifier to identify the roles 'activated' by people in their writing. There is no such dataset either.

The approach we have taken, explained in Section 5.2, consists of:

1. taking posts by people:1) *known to play some roles* X, 2) thought to be talking exclusively, or mostly, about them talking about those roles
2. tag those posts with the role identity class, obtaining a tagged dataset set
3. use these tagged datasets to train a multiclass classifier.

We also argued that care must be taken to choose role identities that are somewhat mutually exclusive of each other, to minimize the chances of one person activating two role identities in the same post.

To this end, we applied this approach to the *Parent* and *child* identities, by using the dataset of the research study "Accuracy of online discussion forums on common childhood ailments" (Farrell, 2018) for parent identity, and 'toxic parent' subreddit from Reddit, for child identity.(https://www.reddit.com/r/toxicparents/)

For classification, we used the BERT model. It achieved a 96% for Accuracy, Precision, Recall, and F1-score, indicating a strong generalization without a bias toward any class. The ROC curve with AUC of 0.99 validates the strong separation ability of the model between the positive and negative samples. The PR curve also achieved AUC-PR of 0.99 which represents high precision and recall even in potentially imbalanced scenarios. This proves the soundness of the approach.

CHAPTER 6

CONCLUSION

6.1 Summary of the thesis

The marketing literature has shown that people's *social identity* influences their purchasing decisions (Section 2.2.1). To the extent that is true, it becomes important to be able to recognize people's social identity from their on-line writing.

The question of *identity* has been studied in the social sciences literature (see Section 2.3). In particular, we saw that Stets et al.'s *identity theory* identified three types of identities (Stets et Serpe, 2016): 1) personal identities, 2) role identities, and 3) group identities. We argued that only *personal identities* and *role identities* are likely to: 1) influence consumer behavior, and 2) manifest themselves in on-line writing. We proposed different strategies for uncovering *personal identities* (Section 3.3 and Chapter 4) and *role identities* (Section 3.4 and Chapter 5).

Regarding personal identities, because individuals may 'activate' many identities at any given in time, we argued in Section 3.3 that detecting personal identities in difficult, in general; however, by focusing on personal identities that involve taking a *polar stance* (*for* or *against*) with regard to some position, we could use *sentiment analysis techniques* to determine a person's membership to the corresponding personal identity. We applied this strategy in Chapter 4 with regard to *veganism*, and used the Kaggle dataset "Sentiment Dataset with 1 Million Tweets" (see `https://www.kaggle.com/datasets/tariqsays/sentiment-dataset-with-1-million-tweets`. We tried different algorithms, namely KNN, MLP, and different CNN+LSTM combinations, and achieved an accuracy of 88.7% with CNN+LSTM. This proves the soundness of the approach. However, it applies only to personal identities that are strongly correlated to polar positions people might have about 'something'.

With regard to *role identities*, we had a similar problem: 1) people can 'activate' multiple roles (e.g. manager and parent) in the same post, and 2) there are no tagged datasets. The approach we have taken (see Section 5.2), consisted of: 1) getting datasets of posts *known* to be about specific roles, and 2) choosing datasets that correspond to roles that are unlikely to be activated simultaneously. We tested the approach on the *Parent* and *child* identities, by using the dataset of the research study "Accuracy of online discussion forums on common childhood ailments" (Farrell, 2018) for parent identity, and 'toxic parent' subreddit from Reddit, for child identity (https://www.reddit.com/r/toxicparents/). With the BERT model, we achieved a 96% for Accuracy, Precision, Recall, and F1-score, indicating a strong generalization without a bias toward any class.

6.2 contributions of the thesis

This thesis made three contributions:

1. It is a case study in interdisciplinary research that combines literature from marketing, social studies, and computer science, leading to a methodology that achieved good results. More specifically:
2. NLP, machine learning techniques, and more specifically, sentiment analysis techniques can be used to detect various identities in people's writing;
3. The approach and techniques used enable us to *confirm* specific identities, identified in advance. This entails careful selection of datasets for training.

This thesis provides a foundation for future research to refine identity detection models.

6.3 Directions for future research

As we said in the previous section, the methodology developed in this thesis enables us to *confirm* specific identities, identified in advance. This is probably sufficient for *customer experience management* applications.

It would be interesting–and more socially useful–to explore applications in populational health, as in the targeting of behavioral or health interventions.

On the theoretical level, it would be interesting to explore techniques that *discover* identities–as opposed to confirm them.

.1 Appendix

Part1(Group Identity) Get-Data

[1]:
```
!pip install kaggle openai
```

```
Requirement already satisfied: kaggle in /usr/local/lib/python3.10/dist-packages
(1.5.16)
Collecting openai
  Downloading openai-1.11.1-py3-none-any.whl (226 kB)
     ---------------------------------------- 226.1/226.1

kB 4.5 MB/s eta 0:00:00
Requirement already satisfied: six>=1.10 in
/usr/local/lib/python3.10/dist-packages (from kaggle) (1.16.0)
Requirement already satisfied: certifi in /usr/local/lib/python3.10/dist-
packages (from kaggle) (2023.11.17)
Requirement already satisfied: python-dateutil in
/usr/local/lib/python3.10/dist-packages (from kaggle) (2.8.2)
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-
packages (from kaggle) (2.31.0)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages
(from kaggle) (4.66.1)
Requirement already satisfied: python-slugify in /usr/local/lib/python3.10/dist-
packages (from kaggle) (8.0.3)
Requirement already satisfied: urllib3 in /usr/local/lib/python3.10/dist-
packages (from kaggle) (2.0.7)
Requirement already satisfied: bleach in /usr/local/lib/python3.10/dist-packages
(from kaggle) (6.1.0)
Requirement already satisfied: anyio<5,>=3.5.0 in
/usr/local/lib/python3.10/dist-packages (from openai) (3.7.1)
Requirement already satisfied: distro<2,>=1.7.0 in /usr/lib/python3/dist-
packages (from openai) (1.7.0)
```

```
Collecting httpx<1,>=0.23.0 (from openai)
  Downloading httpx-0.26.0-py3-none-any.whl (75 kB)
     ---------------------------------------- 75.9/75.9 kB
5.3 MB/s eta 0:00:00
Requirement already satisfied: pydantic<3,>=1.9.0 in
/usr/local/lib/python3.10/dist-packages (from openai) (1.10.14)
Requirement already satisfied: sniffio in /usr/local/lib/python3.10/dist-
packages (from openai) (1.3.0)
Collecting typing-extensions<5,>=4.7 (from openai)
  Downloading typing_extensions-4.9.0-py3-none-any.whl (32 kB)
Requirement already satisfied: idna>=2.8 in /usr/local/lib/python3.10/dist-
packages (from anyio<5,>=3.5.0->openai) (3.6)
Requirement already satisfied: exceptiongroup in /usr/local/lib/python3.10/dist-
packages (from anyio<5,>=3.5.0->openai) (1.2.0)
Collecting httpcore==1.* (from httpx<1,>=0.23.0->openai)
  Downloading httpcore-1.0.2-py3-none-any.whl (76 kB)
     --------------------------------------- 76.9/76.9 kB
6.2 MB/s eta 0:00:00
Collecting h11<0.15,>=0.13 (from httpcore==1.*->httpx<1,>=0.23.0->openai)
  Downloading h11-0.14.0-py3-none-any.whl (58 kB)
     ---------------------------------------- 58.3/58.3 kB
5.4 MB/s eta 0:00:00
Requirement already satisfied: webencodings in
/usr/local/lib/python3.10/dist-packages (from bleach->kaggle) (0.5.1)
Requirement already satisfied: text-unidecode>=1.3 in
/usr/local/lib/python3.10/dist-packages (from python-slugify->kaggle) (1.3)
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.10/dist-packages (from requests->kaggle) (3.3.2)
Installing collected packages: typing-extensions, h11, httpcore, httpx, openai
  Attempting uninstall: typing-extensions
    Found existing installation: typing_extensions 4.5.0
    Uninstalling typing_extensions-4.5.0:
```

```
    Successfully uninstalled typing_extensions-4.5.0
ERROR: pip's dependency resolver does not currently take into account all

the packages that are installed. This behaviour is the source of the following

dependency conflicts.

llmx 0.0.15a0 requires cohere, which is not installed.

llmx 0.0.15a0 requires tiktoken, which is not installed.

tensorflow-probability 0.22.0 requires typing-extensions<4.6.0, but you have

typing-extensions 4.9.0 which is incompatible.

Successfully installed h11-0.14.0 httpcore-1.0.2 httpx-0.26.0 openai-1.11.1
typing-extensions-4.9.0
```

[39]:
```python
import pandas as pd
from tqdm import tqdm
import glob

tqdm.pandas(desc="my bar!")
```

[2]:
```python
# cv = [
#     "Covid",
#   "Covid 19",
#   "Covid19",
#   "corona virus",
#   "corona",
#   "virus",
#   "Covid_19",
#   "corona_virus",
```

```
#   "pandemic",
#   "pandemy",
#   "quarantine",
#   "social distancing",
#   "COVID",
#   "SARS",
#    "lockdown",
# #   "",
# #    "",
# #   "",
# #    "",
# #   "",
# #    "",
# #   "",
# #    "",
# #   "",
#   ]

# cv = [c.lower() for c in cv]
# cv = set(cv)

cv = [
      "vegan",
      "plantbased",
      "healthy",
      "vegetarian",
      "veggie",
      "veganism",
      "cruelty free",
      "plant milk",
      "beyond meat",
```

```
      "vegan person",
      "vegan diet",
      'Organic',
      'Non-GMO',
      # 'Sustainable',
      # 'Eco-friendly',
      'Meat alternatives',
      'Whole foods',
      'Nutrient-dense',
      'Ethical eating',
      'Farm-to-table',
      'Raw food'
 ]

cv = [c.lower() for c in cv]
cv = set(cv)
```

[20]:
```
# def check_word(txt, kw):
#   txt = txt.lower()
#   for k in kw:
#     if(k in txt):
#       return True

def check_word(txt):
  txt = txt.lower()
  for k in cv:
    if(k in txt):
      return True
```

[5]:
```
!mkdir dataset
```

[6]:
```
# from google.colab import files

# uploaded = files.upload()

# for fn in uploaded.keys():
#   print('User uploaded file "{name}" with length {length} bytes'.format(
#       name=fn, length=len(uploaded[fn])))

# Then move kaggle.json into the folder where the API expects to find it.
!mkdir -p ~/.kaggle/ && mv kaggle.json ~/.kaggle/ && chmod 600 ~/.kaggle/kaggle.
↪json
```

[7]:
```
!kaggle datasets download "kazanova/sentiment140"
```

```
Downloading sentiment140.zip to /content
 90% 73.0M/80.9M [00:00<00:00, 195MB/s]
100% 80.9M/80.9M [00:00<00:00, 175MB/s]
```

[8]:
```
!unzip sentiment140.zip
```

```
Archive:  sentiment140.zip
  inflating: training.1600000.processed.noemoticon.csv
```

[18]:
```
df = pd.read_csv('training.1600000.processed.noemoticon.csv',␣
↪encoding="latin-1", header=None)
df.columns = ['target', 'ids', 'data', 'flag', 'user', 'text']
df.head()
```

[18]:
```
   target         ids                          data      flag  \
0       0  1467810369  Mon Apr 06 22:19:45 PDT 2009  NO_QUERY
1       0  1467810672  Mon Apr 06 22:19:49 PDT 2009  NO_QUERY
2       0  1467810917  Mon Apr 06 22:19:53 PDT 2009  NO_QUERY
```

```
3       0  1467811184  Mon Apr 06 22:19:57 PDT 2009  NO_QUERY
4       0  1467811193  Mon Apr 06 22:19:57 PDT 2009  NO_QUERY

              user                                              text
0  _TheSpecialOne_  @switchfoot http://twitpic.com/2y1zl - Awww, t...
1    scotthamilton  is upset that he can't update his Facebook by ...
2         mattycus  @Kenichan I dived many times for the ball. Man...
3          ElleCTF    my whole body feels itchy and like its on fire
4           Karoli  @nationwideclass no, it's not behaving at all...
```

[19]:
```
df.shape
```

[19]: (1600000, 6)

Content

It contains the following 6 fields:

target: the polarity of the tweet (0 = negative, 2 = neutral, 4 = positive)

ids: The id of the tweet (2087)

date: the date of the tweet (Sat May 16 23:58:44 UTC 2009)

flag: The query (lyx). If there is no query, then this value is NO_QUERY.

user: the user that tweeted (robotickilldozr)

text: the text of the tweet (Lyx is cool)

[ ]:
```
# df['date_column'] = pd.to_datetime(df['Job Posting Date'])
```

[21]:
```
df['vegan'] = df.text.apply(check_word)
```

[25]:
```
df[df.vegan==True].shape
```

[25]: (2506, 7)

[26]:
```
df[df.vegan==True].to_csv('dataset/dataset1.csv', index=False)
```

[27]:
```
!kaggle datasets download "tariqsays/sentiment-dataset-with-1-million-tweets"
```

```
Downloading sentiment-dataset-with-1-million-tweets.zip to /content
 86% 65.0M/75.4M [00:02<00:00, 30.3MB/s]
100% 75.4M/75.4M [00:02<00:00, 35.1MB/s]
```

[28]:
```
!unzip sentiment-dataset-with-1-million-tweets.zip
```

```
Archive:  sentiment-dataset-with-1-million-tweets.zip
  inflating: dataset.csv
```

[30]:
```
df2 = pd.read_csv('dataset.csv')
df2.columns = ['text', 'lang', 'label']
df2.head()
```

[30]:

	text	lang	label
0	@Charlie_Corley @Kristine1G @amyklobuchar @Sty...	en	litigious
1	#BadBunny: Como dos gotas de agua: Joven se di...	es	negative
2	https://t.co/YJNiO0p1JV Flagstar Bank disclose...	en	litigious
3	Rwanda is set to host the headquarters of Unit...	en	positive
4	OOPS. I typed her name incorrectly (today's br...	en	litigious

[31]:
```
df2.shape
```

[31]: (937854, 3)

[32]:
```
df2 = df2[df2.lang=='en']
df2.shape
```

[32]: (871310, 3)

[33]:
```
df2['vegan'] = df2.text.apply(check_word)
```

[35]:
```
df2[df2.vegan==True].shape
```

[35]: (2829, 4)

[36]:
```
df2[df2.vegan==True].to_csv('dataset/dataset2.csv', index=False)
```

[49]:
```
!zip -r ./data.zip /content/dataset
```

```
adding: content/dataset/ (stored 0%)
adding: content/dataset/dataset2.csv (deflated 58%)
adding: content/dataset/dataset1.csv (deflated 60%)
```

[ ]:
```

```

[ ]:
```

```

Pre-processing

[1]:
```
import pandas as pd
import numpy as np
from tqdm import tqdm
import os

import matplotlib.pyplot as plt
import seaborn as sns

from textblob import TextBlob

import nltk
# nltk.download('stopwords')
# nltk.download('punkt')
```

```
# nltk.download('wordnet')

from nltk import pos_tag
from nltk.corpus import wordnet as wn, stopwords
from nltk.stem.wordnet import WordNetLemmatizer

import re
import string
from unidecode import unidecode
import contractions
from bs4 import BeautifulSoup
from sklearn.feature_extraction import _stop_words
```

```
[nltk_data] Downloading package stopwords to
[nltk_data]     C:\Users\AI\AppData\Roaming\nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\AI\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package wordnet to
[nltk_data]     C:\Users\AI\AppData\Roaming\nltk_data...
[nltk_data]   Package wordnet is already up-to-date!
```

[2]:
```
add = 'dataset/ds1-tags.csv'
input_col = 'text'
col = 'clean-text'
```

[3]:
```
def remove_accented_chars(text):
    text = unidecode(text)
    return text
```

[4]:
```
def expand_contractions(text):
    text = contractions.fix(text)
    return text
```

[5]:
```
def strip_html_tags(text):
    soup = BeautifulSoup(text, "html.parser")
    stripped_text = soup.get_text(separator=" ")
    return stripped_text
```

[6]:
```
def remove_link_email(txt):
    txt = txt.replace("...", "")
    txt = re.sub(r"http\S+", "", txt)
    txt = txt.replace('\S*@\S*\s?', "")
    txt = re.sub(r'[^\w\s]', '', txt)
    return txt
```

[7]:
```
def remove_punctuation(text):
    punctuationfree = "".join([i for i in text if i not in string.punctuation])
    return punctuationfree
```

[8]:
```
def lower_case(input_string):
    words = input_string.split(' ')
    for i in range(len(words)):
        if not words[i].isupper():
            words[i] = words[i].lower()
    return ' '.join(words)
```

[9]:
```
is_noun = lambda pos: pos[:2] == 'NN'
```

[10]:
```
stop_words = set(stopwords.words('english')).union(set(_stop_words.
 ↪ENGLISH_STOP_WORDS))
```

[11]:
```
l_word = 3 #min length of word
ngram_range = (1, 2)
```

[12]:
```python
df = pd.read_csv(add)
```

[13]:
```python
df.shape
```

[13]:
```
(2506, 9)
```

[14]:
```python
df.head()
```

[14]:
```
   target                          data      flag             user  \
0       0  Mon Apr 06 23:11:06 PDT 2009  NO_QUERY        amitgupta
1       0  Tue Apr 07 00:07:07 PDT 2009  NO_QUERY  TheGrottoTweets
2       0  Tue Apr 07 02:42:01 PDT 2009  NO_QUERY          ipodque
3       0  Tue Apr 07 02:49:46 PDT 2009  NO_QUERY      moximillian
4       0  Tue Apr 07 03:12:37 PDT 2009  NO_QUERY        Jayme1988

                                                text  negative  neutral  \
0  Have an invite for &quot;Healthy Dining&quot; ...      0.56     0.39
1  @TexasVegetarian Oh, God, ow.  That must have ...      0.91     0.08
2  Creepy Outdoor on the speedway: Alli is now in...      0.91     0.08
3  The nowhere land - not 100% sick, but definate...      0.83     0.16
4  @Seamonkey86 I am on a healthy eating kick! I ...      0.01     0.12

   positive     final
0      0.05  negative
1      0.01  negative
2      0.01  negative
3      0.02  negative
4      0.87  positive
```

[15]:
```python
df[col] = df[input_col].apply(lower_case)
```

[16]:
```python
df.dropna(subset=[col], inplace=True)
```

[17]:
```python
df['len'] = df[col].str.count(' ')
# df = df[df.len  >= l_word]
```

```
[18]: df.reset_index(drop=True, inplace=True)
```

```
[19]: df[col] = df[col].apply(lambda x:re.sub(r'\W+', ' ', x))
      df[col] = df[col].apply(remove_accented_chars)
      df[col] = df[col].apply(expand_contractions)
      df[col] = df[col].apply(strip_html_tags)
      df[col] = df[col].apply(remove_link_email)
      df[col] = df[col].apply(lambda x:remove_punctuation(x))
      df[col] = df[col].apply(nltk.word_tokenize)
      # df[col] = train[col].apply(lambda x: [word for (word, pos) in nltk.pos_tag(x)␣
       ↪if is_noun(pos)] )
      df[col] = df[col].apply(lambda x: [item for item in x if item not in stop_words])
      df[col] = df[col].apply(lambda x: [re.sub('[^a-zA-Z]+', '', item) for item in x])
      df[col] = df[col].apply(lambda x : ' '.join([w for w in x if len(w.
       ↪strip())>l_word]))
      df[col] = df[col].apply(lambda x: " ". join([w.lemmatize() for w in TextBlob(x).
       ↪words]))
      df[col] = df[col].apply(nltk.word_tokenize)
      df[col] = df[col].apply(lambda x: [item for item in x if item not in stop_words])
      df[col] = df[col].apply(lambda x: " ". join(x))
```

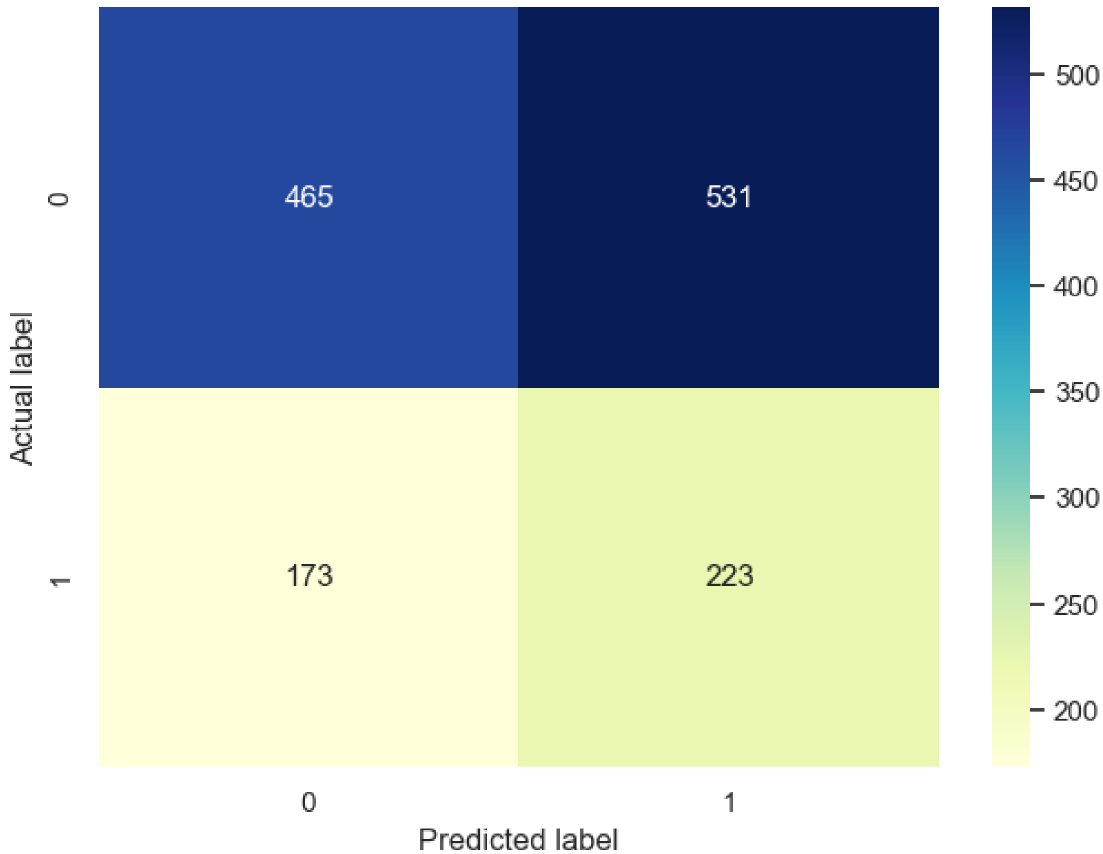
```
C:\Users\AI\AppData\Local\Temp\ipykernel_11116\497320988.py:2:
MarkupResemblesLocatorWarning: The input looks more like a filename than markup.
You may want to open this file and pass the filehandle into Beautiful Soup.
  soup = BeautifulSoup(text, "html.parser")
```

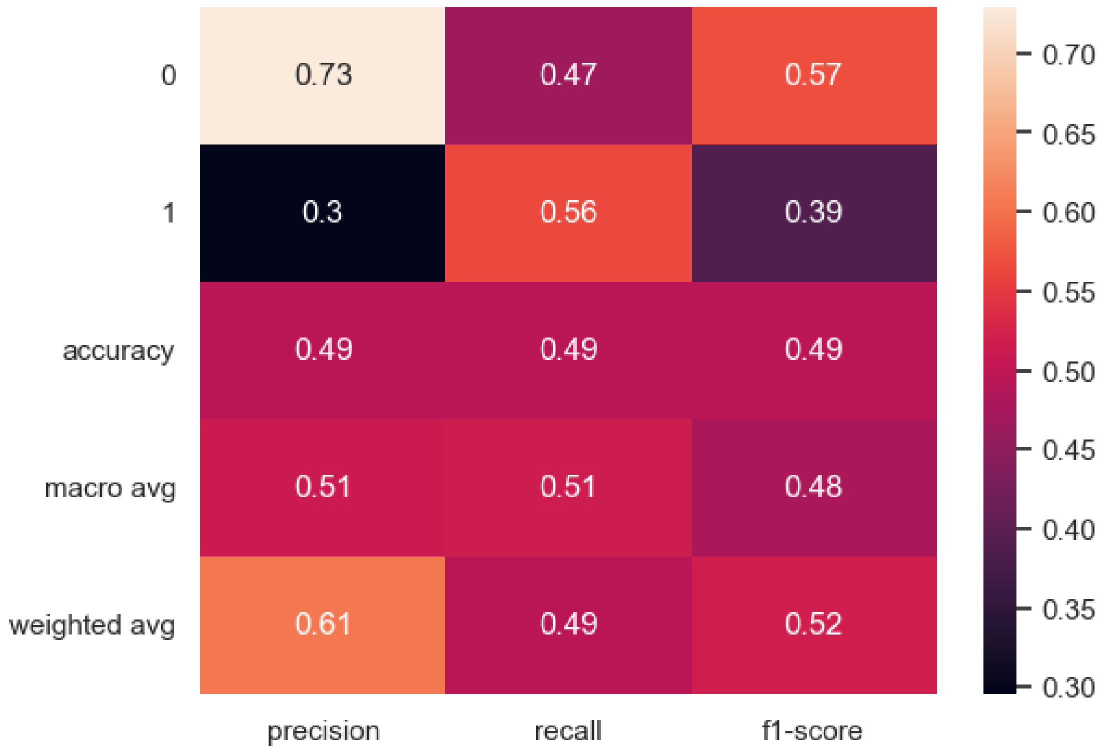
```
[20]: df.reset_index(drop=True, inplace=True)
```

```
[21]: df.to_csv(add[:-4]+'-preprocessed.csv', index=False)
```

Machine learning and LLM

e: 0.021974299999783398


465
531
173
223
0
1
Actual label
0
1
Predicted label
500
450
400
350
300
250
200



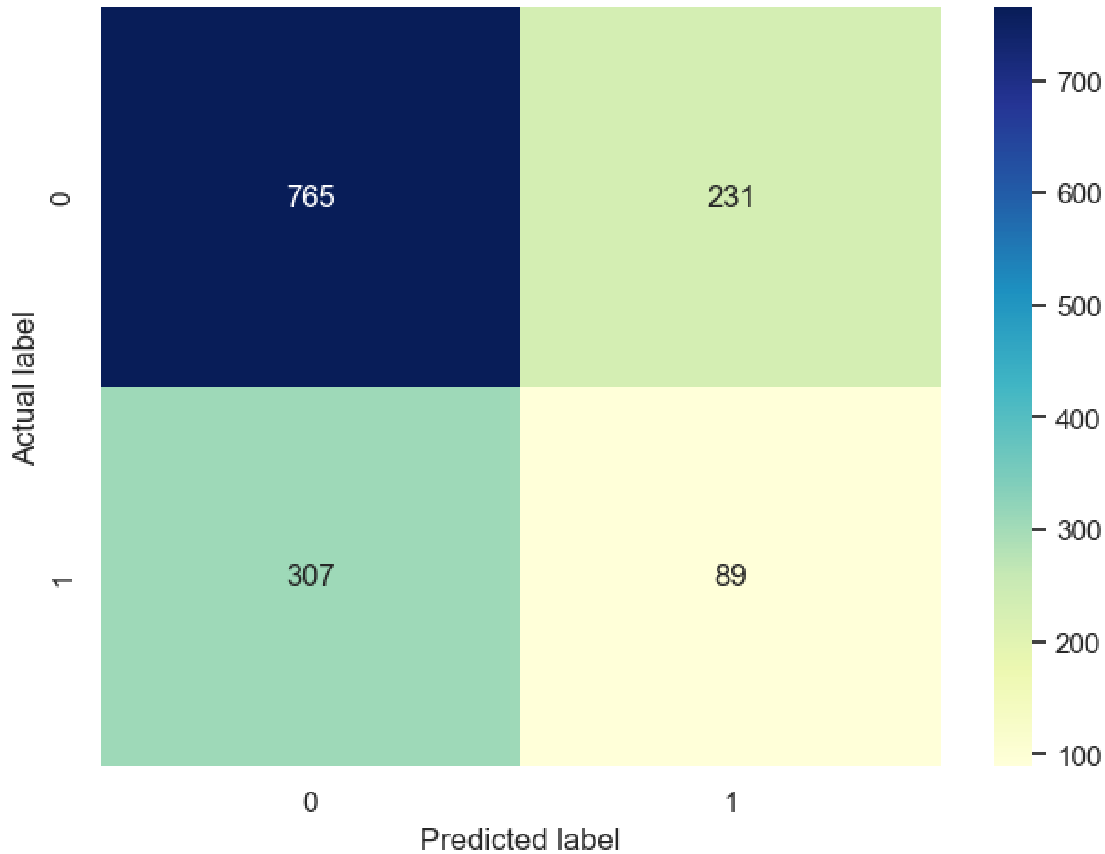
[26]:

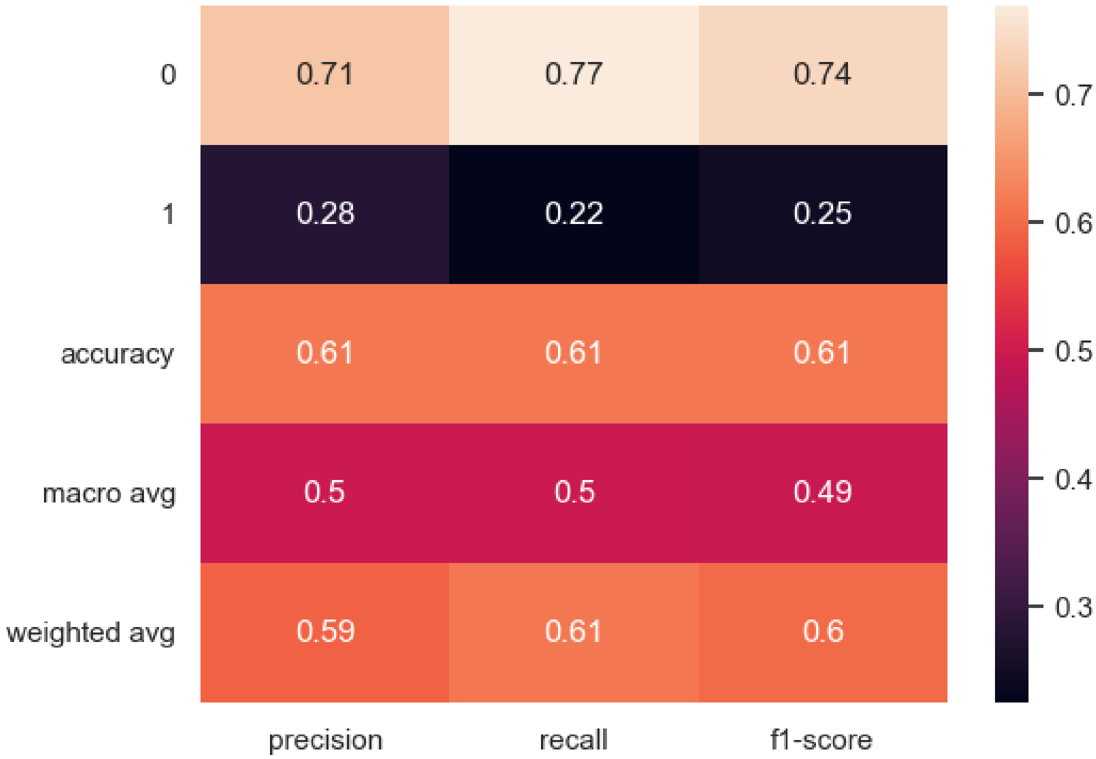
```
model_name = 'MLP'
mlp_model = MLPClassifier(random_state=1, max_iter=300)

mlp_model, conf, y_pred2 = train_model(mlp_model, x_train, y_train, x_test,␣
 ↪y_test)
plot_diagram(conf, model_name)
plot_diagram2(y_test, y_pred2, model_name)
calc_roc(conf, model_name, mlp_model, x_train, y_train, x_test, y_test)
```

```
C:\Users\AI\AppData\Local\Programs\Python\Python310\lib\site-
packages\sklearn\neural_network\_multilayer_perceptron.py:691:
ConvergenceWarning: Stochastic Optimizer: Maximum iterations (300) reached and
the optimization hasn't converged yet.
```

```
  warnings.warn(
```

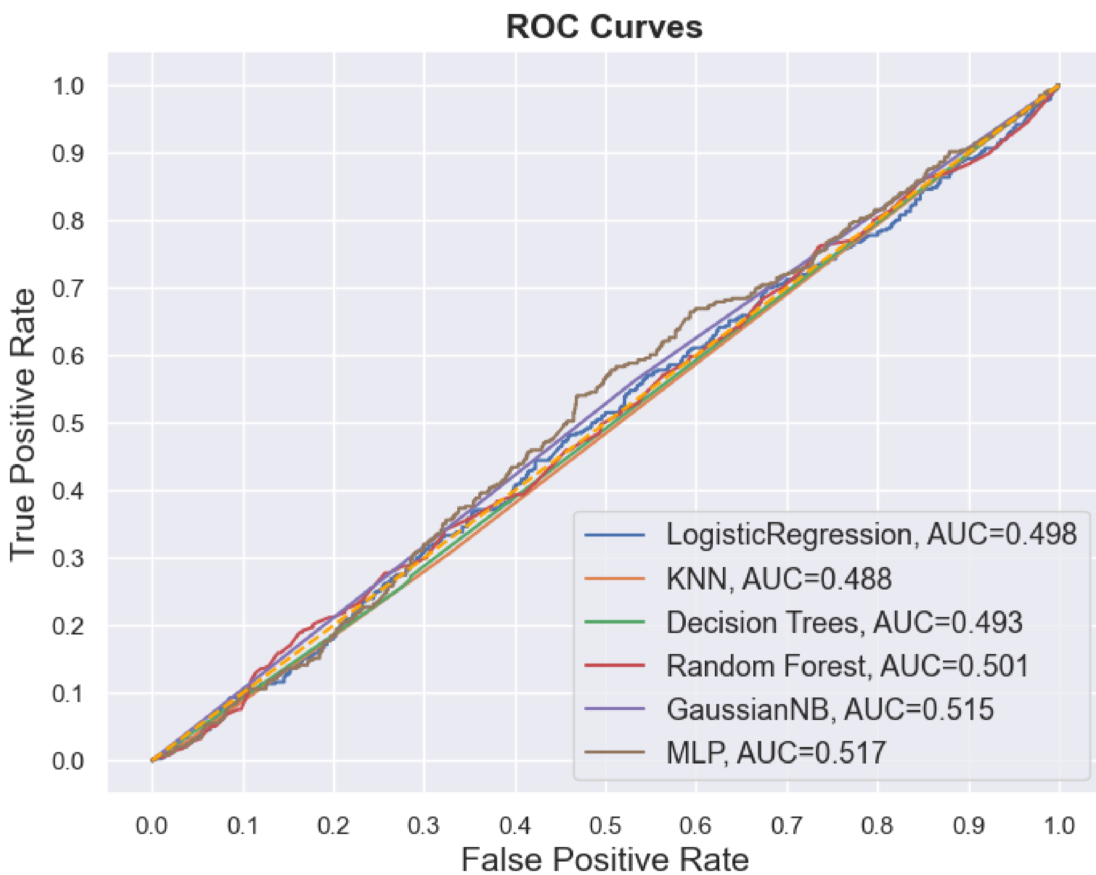
```
Train Time:  3.4679350999999522
Test Time:  0.009138000001257751
```





[27]:
```
fig = plt.figure(figsize=(8, 6))

for res in results:
    plt.plot(res[0], res[1], label=res[2])

plt.plot([0, 1], [0, 1], color='orange', linestyle='--')

plt.xticks(np.arange(0.0, 1.1, step=0.1))
plt.xlabel("False Positive Rate", fontsize=15)

plt.yticks(np.arange(0.0, 1.1, step=0.1))
plt.ylabel("True Positive Rate", fontsize=15)
```

```
plt.title('ROC Curves', fontweight='bold', fontsize=15)
plt.legend(prop={'size':13}, loc='lower right')


plt.savefig('result/roc_curves.png', dpi=500)
plt.show()
```



[28]:
```
plt.figure(figsize=(14,8))
xx = range(len(result))
index_name = result.index
plt.bar(xx, result['accuracy'], color=(0.2, 0.4, 0.6, 0.6), alpha=0.5)
```

```
plt.ylim(0.6, 1)
plt.title('Model Comparison on Test Set', fontsize=18)
plt.ylabel('Test Accuracy (%)', fontsize=16)
plt.xticks(xx, index_name, rotation=90, fontsize=16)
sns.despine()
plt.savefig('result/Comparison.png', dpi=300, bbox_inches='tight')
```

[29]:

```
result
```

[29]:

	name	accuracy	recall	precison	f1
0	LogisticRegression	0.715	0.715	0.607	0.598
1	KNN	0.715	0.715	0.512	0.597
2	Decision Trees	0.592	0.592	0.585	0.588
3	Random Forest	0.689	0.689	0.565	0.597
4	GaussianNB	0.494	0.494	0.606	0.518

```
5                MLP    0.614   0.614     0.59    0.6
```

[30]:
```python
result.to_csv('res.csv', index=False)
```

[ ]:

[ ]:

Part2 - Role Identity

Get data

[ ]:
```
pip install praw
```

[ ]:
```
import praw
import pandas as pd
import datetime
```

[ ]:
```
reddit = praw.Reddit(
    client_id="",
    client_secret="",
    user_agent=""
)
```

[ ]:
```
subreddit = reddit.subreddit("NewParents")
print(subreddit.title)
```

```
WARNING:praw:It appears that you are using PRAW in an asynchronous environment.
It is strongly recommended to use Async PRAW: https://asyncpraw.readthedocs.io.
See https://praw.readthedocs.io/en/latest/getting_started/multiple_instances.htm
l#discord-bots-and-asynchronous-environments for more info.
```

```
New Parents!
```

[ ]:
```
data = []

for post in subreddit.hot(limit=2000):
    post.comments.replace_more(limit=0)
    for comment in post.comments.list():
        data.append({
            "subreddit": subreddit.title,
```

```
            "post_title": post.title,
            "comment": comment.body,
            "score": comment.score,
            "created_utc": datetime.datetime.fromtimestamp(comment.created_utc),
            "comment_author": str(comment.author),
            "permalink": f"https://reddit.com{comment.permalink}"
        })

df = pd.DataFrame(data)
df.to_csv("child_comments.csv", index=False)
```

[ ]:
```
df
```

[ ]:
```
          subreddit                                    post_title  \
0      New Parents!                Weekly Discussion - Relationships
1      New Parents!                Weekly Discussion - Relationships
2      New Parents!                Weekly Discussion - Relationships
3      New Parents!                Weekly Discussion - Relationships
4      New Parents!                Weekly Discussion - Relationships
...             ...                                              ...
14689  New Parents!  Do toddlers need to have breakfast before drop...
14690  New Parents!  Do toddlers need to have breakfast before drop...
14691  New Parents!  Do toddlers need to have breakfast before drop...
14692  New Parents!  Do toddlers need to have breakfast before drop...
14693  New Parents!  Do toddlers need to have breakfast before drop...

                                                 comment  score  \
0      I just need a moment to rant and express my hu...      5
1      We have an 8 week old daughter who honestly is...      3
2      Bit of a vent, and my wife knows I say it to h...      2
```

```
3      Looking for advice\n\nI have a 6 week old girl...      2
4      We have been living with my in laws as we wait...      2
...                                                  ...    ...
14689  He's 1.5 years old. He doesn't go to daycare y...      1
14690     Ahh that's a good point. I'll ask the daycare!      1
14691  Around that age, we premade smoothies in reusa...      1
14692  oh i see, but what time will you need to wake ...      3
14693  That makes sense. We haven't thought of his wa...      1

              created_utc        comment_author  \
0     2025-08-26 14:15:08          DramaticAriel
1     2025-08-27 14:35:51  Brilliant-Parsley-40
2     2025-08-26 08:42:13            Kamen-Ramen
3     2025-08-29 08:53:47  NatureParticular8960
4     2025-08-29 15:51:30      feisty-tiny-okra
...                   ...                   ...
14689 2025-08-25 15:57:55               Fenix512
14690 2025-08-25 15:59:00               Fenix512
14691 2025-08-25 20:21:25            proteins911
14692 2025-08-25 15:59:47              Few_Paces
14693 2025-08-25 16:12:08               Fenix512

                                               permalink
0      https://reddit.com/r/NewParents/comments/1n0eu...
1      https://reddit.com/r/NewParents/comments/1n0eu...
2      https://reddit.com/r/NewParents/comments/1n0eu...
3      https://reddit.com/r/NewParents/comments/1n0eu...
4      https://reddit.com/r/NewParents/comments/1n0eu...
...                                                  ...
14689  https://reddit.com/r/NewParents/comments/1mztx...
14690  https://reddit.com/r/NewParents/comments/1mztx...
```

```
14691  https://reddit.com/r/NewParents/comments/1mztx...
14692  https://reddit.com/r/NewParents/comments/1mztx...
14693  https://reddit.com/r/NewParents/comments/1mztx...

[14694 rows x 7 columns]
```

codes for 4 categories classification

[1]:
```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from transformers import BertTokenizer
import torch
from transformers import BertForSequenceClassification, Trainer,␣
↪TrainingArguments
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics import (
    roc_auc_score, roc_curve, precision_recall_curve, auc,
    confusion_matrix, classification_report
)
import torch.nn.functional as F
from sklearn.metrics import classification_report, accuracy_score,␣
↪precision_score, recall_score, f1_score
import re
```

[2]:
```
df_parents=pd.read_csv("/content/parents.csv")
df1_child=pd.read_csv("/content/child_comments.csv")
```

[3]:
```
def clean_useless_comments(text):
    text = str(text)
    text = re.sub(r"http\S+|www\S+|https\S+", "", text)
    text = re.sub(r'[^\w\s.,?!\'\"]', '', text)
    text = re.sub(r'\b\d+\b', '', text)
    return text.strip()
```

[4]:
```
df2_teen=pd.read_csv("teenagers_comments.csv")
df2_teen
```

[4]:

```
          subreddit                                post_title  \
0      r/teenagers  [mod] REMINDER ABOUT WITCH HUNTING POSTS.
1      r/teenagers  [mod] REMINDER ABOUT WITCH HUNTING POSTS.
2      r/teenagers  [mod] REMINDER ABOUT WITCH HUNTING POSTS.
3      r/teenagers  [mod] REMINDER ABOUT WITCH HUNTING POSTS.
4      r/teenagers  [mod] REMINDER ABOUT WITCH HUNTING POSTS.
...            ...                                       ...
13909  r/teenagers              Bro I need songs to listen to
13910  r/teenagers              Bro I need songs to listen to
13911  r/teenagers              Bro I need songs to listen to
13912  r/teenagers              Bro I need songs to listen to
13913  r/teenagers                                  my bfs ex

                                                 comment  score  \
0      Reddit mods (the ones for the whole app not yo...     17
1      alhamdulilah thank you these have been so anno...     23
2      Oh shoot the mods came hopefully they didn't s...      4
3                   THANK GOD IM TIRED OF THESE POSTS       7
4      Bro can you please ban the teenagers posting t...      6
...                                                  ...    ...
13909                                     listen to styx      1
13910                           "Paranoid" Black Sabbath      1
13911                           "Paranoid" Black Sabbath      1
13912                                              sombr      1
13913  That's actually so annoying, she sounds incred...      1

               created_utc        comment_author  \
0      2025-05-04 14:38:29          nermalnormal
1      2025-05-03 18:07:50      AwesomeeeeeeeAcc
2      2025-05-04 07:02:59  Pure-Professional144
3      2025-05-04 00:31:50         AussieGoofball
```

```
4      2025-05-05 21:21:59        bipolar-femboy
...                    ...                   ...
13909  2025-05-24 17:25:34      Pristine-Donut22
13910  2025-05-24 17:30:45  Alternative-Way-1760
13911  2025-05-24 17:30:51  Alternative-Way-1760
13912  2025-05-24 17:31:22       curious_foxx211
13913  2025-05-24 13:41:39      FallenMeteorite5

                                               permalink
0      https://reddit.com/r/teenagers/comments/1kdz4t...
1      https://reddit.com/r/teenagers/comments/1kdz4t...
2      https://reddit.com/r/teenagers/comments/1kdz4t...
3      https://reddit.com/r/teenagers/comments/1kdz4t...
4      https://reddit.com/r/teenagers/comments/1kdz4t...
...                                                  ...
13909  https://reddit.com/r/teenagers/comments/1kugy8...
13910  https://reddit.com/r/teenagers/comments/1kugy8...
13911  https://reddit.com/r/teenagers/comments/1kugy8...
13912  https://reddit.com/r/teenagers/comments/1kugy8...
13913  https://reddit.com/r/teenagers/comments/1kubto...

[13914 rows x 7 columns]
```

[5]:
```
df2_teen['comment'] = df2_teen['comment'].apply(clean_useless_comments)
df2_teen = df2_teen[df2_teen['comment'].notna()]
```

[6]:
```
comment_df=list(df2_teen['comment'])
```

[7]:
```
post_teen=list(dict.fromkeys(df2_teen[["comment", "post_title"]]["post_title"].
↪iloc[1:]))
```

[8]:
```
df2_teen_comment = df2_teen[df2_teen['comment'].apply(lambda x: len(str(x).
 ↪split()) >= 10)]
df2_teen_comment_list=list(df2_teen_comment.comment)


df2_teen_post = df2_teen[df2_teen['post_title'].apply(lambda x: len(str(x).
 ↪split()) >= 10)]
df2_teen_post_list=list(df2_teen_post.post_title)
```

[9]:
```
c = df2_teen_comment_list + df2_teen_post_list
```

[10]:
```
role_teen_df=["teenagers" for i in range (len(c))]
df2_teen=pd.DataFrame({
    "role":role_teen_df,
    "comment":c
})
```

[11]:
```
df2_teen
```

[11]:

	role	comment
0	teenagers	Reddit mods the ones for the whole app not you...
1	teenagers	Oh shoot the mods came hopefully they didn't s...
2	teenagers	Bro can you please ban the teenagers posting t...
3	teenagers	Exactly! My post with censored usernames got d...
4	teenagers	Hi, I have a question. I have screenshotted qu...
...	...	...
10223	teenagers	If you read this you have to tell your crush y...
10224	teenagers	If you read this you have to tell your crush y...
10225	teenagers	If you read this you have to tell your crush y...
10226	teenagers	If you read this you have to tell your crush y...
10227	teenagers	If you read this you have to tell your crush y...

[10228 rows x 2 columns]

[12]:
```
df3_grandparents=pd.read_csv("/content/grandparents_comments.csv")
df4_grandpa=pd.read_csv("/content/grandpa_comments.csv")
```

[13]:
```
df4_grandpa['comment'] = df4_grandpa['comment'].apply(clean_useless_comments)
df4_grandpa = df4_grandpa[df4_grandpa['comment'].notna()]

df4_grandpa['post_title'] = df4_grandpa['post_title'].
↪apply(clean_useless_comments)
df4_grandpa = df4_grandpa[df4_grandpa['post_title'].notna()]

df3_grandparents['comment'] = df3_grandparents['comment'].
↪apply(clean_useless_comments)
df3_grandparents = df3_grandparents[df3_grandparents['comment'].notna()]

df3_grandparents['post_title'] = df3_grandparents['post_title'].
↪apply(clean_useless_comments)
df3_grandparents = df3_grandparents[df3_grandparents['post_title'].notna()]
```

[14]:
```
df4_grandpa_comment = df4_grandpa[df4_grandpa['comment'].apply(lambda x:␣
↪len(str(x).split()) >= 8)]
df4_grandpa_comment_list=list(df4_grandpa_comment.comment)

df3_grandparents_comment = df3_grandparents[df3_grandparents['comment'].
↪apply(lambda x: len(str(x).split()) >= 8)]
df3_grandparents_comment_list=list(df3_grandparents_comment.comment)

df3_grandparents_post = df3_grandparents[df3_grandparents['post_title'].
↪apply(lambda x: len(str(x).split()) >= 10)]
df3_grandparents_post_list=list(df3_grandparents_post.post_title)

df4_grandpa_post = df4_grandpa[df4_grandpa['post_title'].apply(lambda x:␣
↪len(str(x).split()) >= 10)]
```

```
df4_grandpa_post_list=list(df4_grandpa_post.post_title)
```

[15]:
```
df3_grandparents
```

[15]:
```
        subreddit                                          post_title  \
0    Grandparents         Anyone know why my Grandma acts like this?
1    Grandparents         Anyone know why my Grandma acts like this?
2    Grandparents         Anyone know why my Grandma acts like this?
3    Grandparents         Anyone know why my Grandma acts like this?
4    Grandparents         Anyone know why my Grandma acts like this?
..            ...                                                 ...
827  Grandparents                           Soon to be new Paw Paw
828  Grandparents                          Miss my Grandma so much
829  Grandparents  My grandparents just sent me this for my 33rd ...
830  Grandparents  BBC News  My Nazi grandfather, Amon Goeth, wou...
831  Grandparents  How do I get my grandparents to see The Hobbit...

                                               comment  score  \
0    Get her to a doctor. She needs a full check up...     16
1    Have her hearing checked.  The kids may be cre...     10
2    Sometime I like to look at it as backwards agi...      6
3    You sound like a stressed but very caring gran...      5
4    I think what you describe is very frustrating...      5
..                                                 ...    ...
827  Thanks! He was born this morning at  my time. ...      1
828   I'm sure she was a very kind woman. Stay strong!      1
829  Priceless. \n\nI have to remember to write a f...      1
830              You sadly really can't choose family.      1
831                                            Donuts?      1

             created_utc       comment_author  \
0    2022-03-28 05:12:26              629mrsn
```

```
1    2022-03-28 00:45:18  JaneAustinAstronaut
2    2022-04-29 20:57:13           Rememberose
3    2022-04-26 19:14:01                   NaN
4    2022-05-19 22:00:29  ILoveMyGrandsAuthor
..                   ...                   ...
827  2016-06-04 04:36:33         mrmidnight273
828  2015-10-25 17:10:35           lbfeboston
829  2015-08-28 03:16:50              ccbbb23
830  2013-10-07 14:37:45           mickeynero
831  2013-03-25 02:33:43             revmeeks

                                              permalink
0    https://reddit.com/r/grandparents/comments/tpw...
1    https://reddit.com/r/grandparents/comments/tpw...
2    https://reddit.com/r/grandparents/comments/tpw...
3    https://reddit.com/r/grandparents/comments/tpw...
4    https://reddit.com/r/grandparents/comments/tpw...
..                                                 ...
827  https://reddit.com/r/grandparents/comments/4ma...
828  https://reddit.com/r/grandparents/comments/3q4...
829  https://reddit.com/r/grandparents/comments/3in...
830  https://reddit.com/r/grandparents/comments/1nw...
831  https://reddit.com/r/grandparents/comments/15e...

[832 rows x 7 columns]
```

[16]:
```
post_grandparents=list(dict.fromkeys(df3_grandparents[["comment",
 ↪"post_title"]]["post_title"].iloc[1:]))
post_grandpa=list(dict.fromkeys(df4_grandpa[["comment",
 ↪"post_title"]]["post_title"].iloc[1:]))
```

[17]:
```python
comment_df3_grandparents=list(df3_grandparents['comment'])
comment_df4_grandpa=list(df4_grandpa['comment'])
```

[18]:
```python
a = df4_grandpa_post_list + df4_grandpa_comment_list
b = df3_grandparents_comment_list + df4_grandpa_post_list
```

[19]:
```python
role_grandparents_df=["grandparents" for i in range (len(b))]
df3_grandparents=pd.DataFrame({
    "role":role_grandparents_df,
    "comment":b
})

role_grandpa_df=["grandparents" for i in range (len(a))]
df4_grandpa=pd.DataFrame({
    "role":role_grandpa_df,
    "comment":a
})
```

[20]:
```python
df3_grandparent=pd.concat([df4_grandpa , df3_grandparents],ignore_index=True)
```

[21]:
```python
df3_grandparent
```

[21]:

	role	comment
0	grandparents	I sure do like to jerk off looking for a long ...
1	grandparents	. No pill or cock ring needed. Just sit on pap...
2	grandparents	. No pill or cock ring needed. Just sit on pap...
3	grandparents	. No pill or cock ring needed. Just sit on pap...
4	grandparents	I want to bend you over, push your pretty fac...
...	...	...
3140	grandparents	Grandpa having a chill out day want to join me
3141	grandparents	Grandpa having a chill out day want to join me
3142	grandparents	Grandpa having a chill out day want to join me
3143	grandparents	Grandpa having a chill out day want to join me
3144	grandparents	Grandpa having a chill out day want to join me

```
[3145 rows x 2 columns]
```

[22]:
```python
df1_child=df1_child.iloc[:1749]
df3_grandparent=df3_grandparent.iloc[:1749]
df2_teen=df2_teen.iloc[:1749]
```

[23]:
```python
df3_grandparent
```

[23]:
```
             role                                            comment
0     grandparents  I sure do like to jerk off looking for a long ...
1     grandparents  . No pill or cock ring needed. Just sit on pap...
2     grandparents  . No pill or cock ring needed. Just sit on pap...
3     grandparents  . No pill or cock ring needed. Just sit on pap...
4     grandparents  I  want to bend you over, push your pretty fac...
...            ...                                                ...
1744  grandparents  Yes that is an excellent point. As I said in t...
1745  grandparents  Yeah unfortunately my sister's family just nee...
1746  grandparents  I can understand this point of view. Call me a...
1747  grandparents  From this elaboration, I also now wonder if it...
1748  grandparents  Sounds like you might have a rmildlynomil. \n\...

[1749 rows x 2 columns]
```

```
No charts were generated by quickchart
```

[ ]:
```python
#df3_grandparent.to_csv("grandparentssss_comments.csv",index=False)
```

[ ]:
```python
df4_grandpa
```

[ ]:
```
            role                                           comment
0    grandparents                                [61] Sexy Sunday.
1    grandparents                                59 year old Grandad
```

```
2     grandparents                    Are you watching or joining in!
3     grandparents   Bi-curious, fit, healthy, clean, youthful [68]...
4     grandparents              I can be a handful some days   [60]
..              ...                                                ...
686   grandparents                                        Good morning
687   grandparents        (54) just sat here waiting for you. DMs open
688   grandparents                          Gramps needs a play mate (67)
689   grandparents                                      (56) married dad
690   grandparents   64 Grandpa having a chill out day want to join me

[691 rows x 2 columns]
```

[24]:
```python
df_complete=pd.concat([df_parents , df1_child , df3_grandparent, df2_teen␣
↪],ignore_index=True)
```

[25]:
```python
df_complete.to_csv("compelte_dataset.csv", index=False)
```

[26]:
```python
df_complete
```

[26]:
```
           role                                            comment
0        Parent  Sounds like she's getting some constipation. M...
1        Parent  As mothers I think we always have a right to g...
2        Parent  I would think she would be more concerned. Yes...
3        Parent  I would think she would be more concerned. Yes...
4        Parent  It's called the baby exchange where we meet in...
...         ...                                                ...
6991  teenagers  Bro I have a twink physique, u got exactly wha...
6992  teenagers  Remember to take care of yourself now. The mor...
6993  teenagers  wth bro where are these type of dudes?? in my ...
6994  teenagers  Thank you!  \n years of intensive training and...
6995  teenagers  Thank you so much!   \nI hope at least good ma...

[6996 rows x 2 columns]
```

```
<google.colab._quickchart_helpers.SectionTitle at 0x7a3cf3dfb190>
```

```
from matplotlib import pyplot as plt
df_complete['lable'].plot(kind='hist', bins=20, title='lable')
plt.gca().spines[['top', 'right',]].set_visible(False)
```

```
<google.colab._quickchart_helpers.SectionTitle at 0x7a3d04da7950>
```

```
from matplotlib import pyplot as plt
import seaborn as sns
df_complete.groupby('role').size().plot(kind='barh', color=sns.palettes.
 ↪mpl_palette('Dark2'))
plt.gca().spines[['top', 'right',]].set_visible(False)
```

```
<google.colab._quickchart_helpers.SectionTitle at 0x7a3cf3d06790>
```

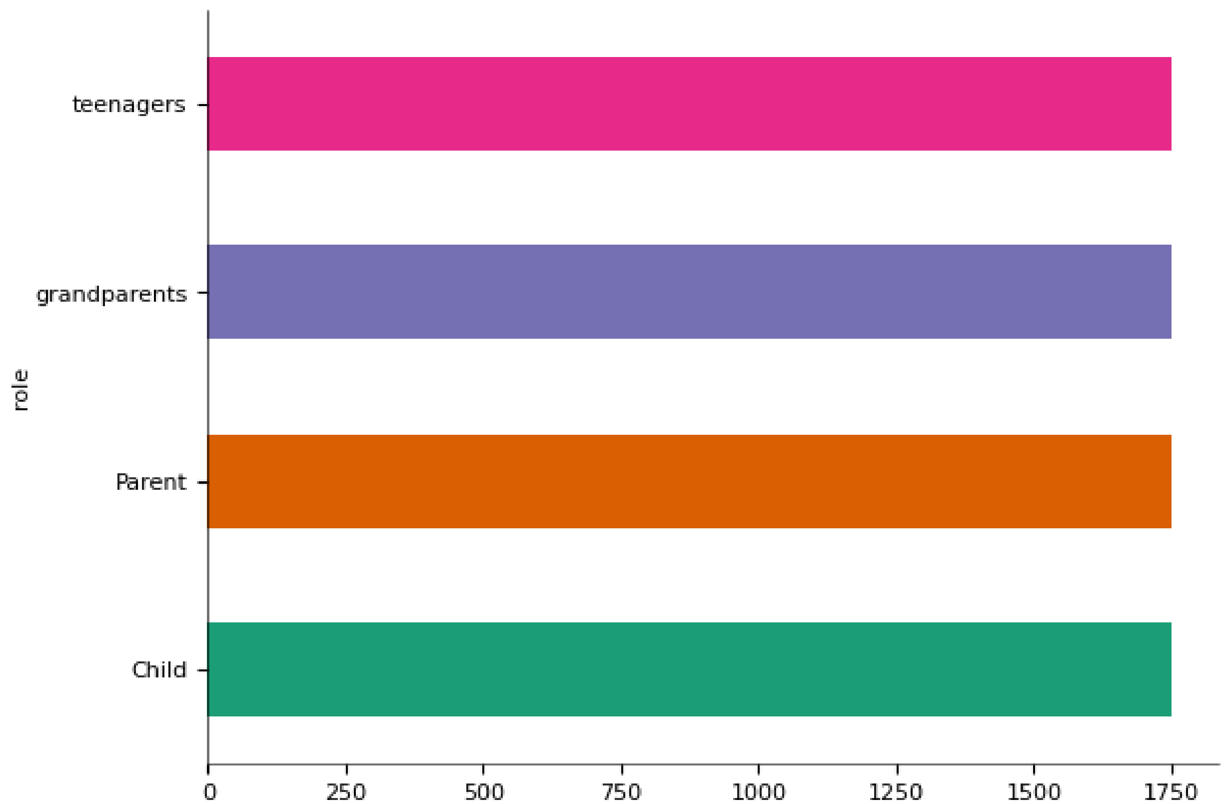
```
from matplotlib import pyplot as plt
df_complete['lable'].plot(kind='line', figsize=(8, 4), title='lable')
plt.gca().spines[['top', 'right']].set_visible(False)
```

```
<google.colab._quickchart_helpers.SectionTitle at 0x7a3cf3d06e10>
```

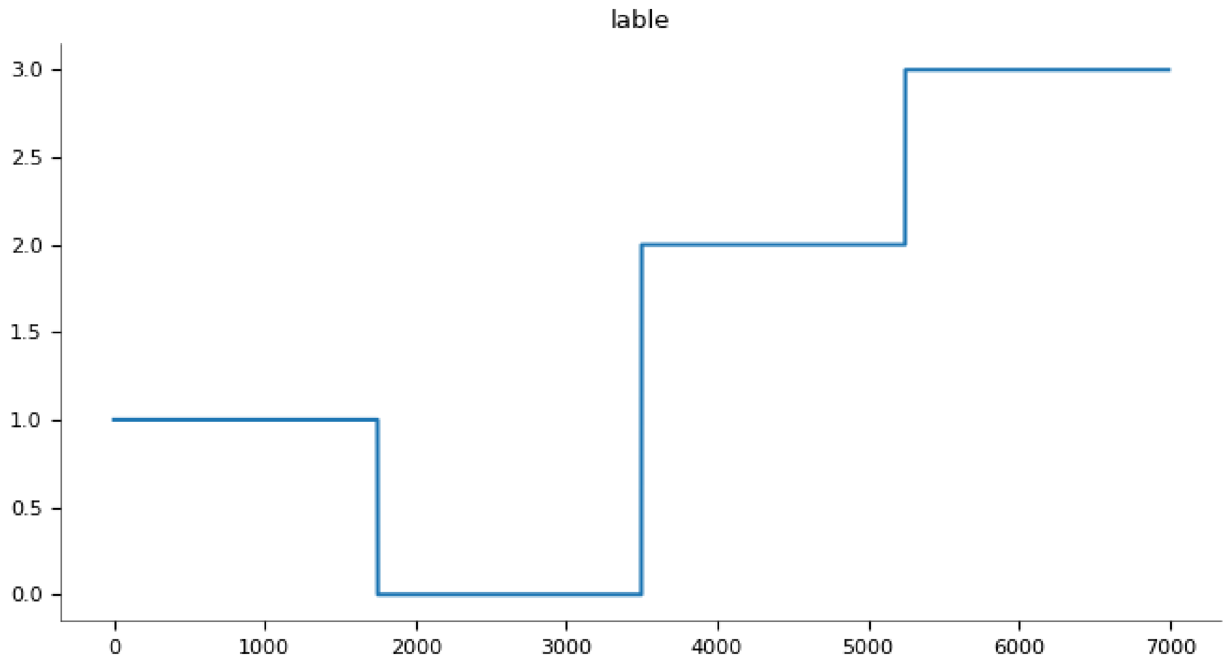
```
<string>:5: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in
v0.14.0. Assign the `y` variable to `hue` and set `legend=False` for the same
effect.
```

```
from matplotlib import pyplot as plt
import seaborn as sns
figsize = (12, 1.2 * len(df_complete['role'].unique()))
plt.figure(figsize=figsize)
sns.violinplot(df_complete, x='lable', y='role', inner='box', palette='Dark2')
sns.despine(top=True, right=True, bottom=True, left=True)
```

Categorical distributions

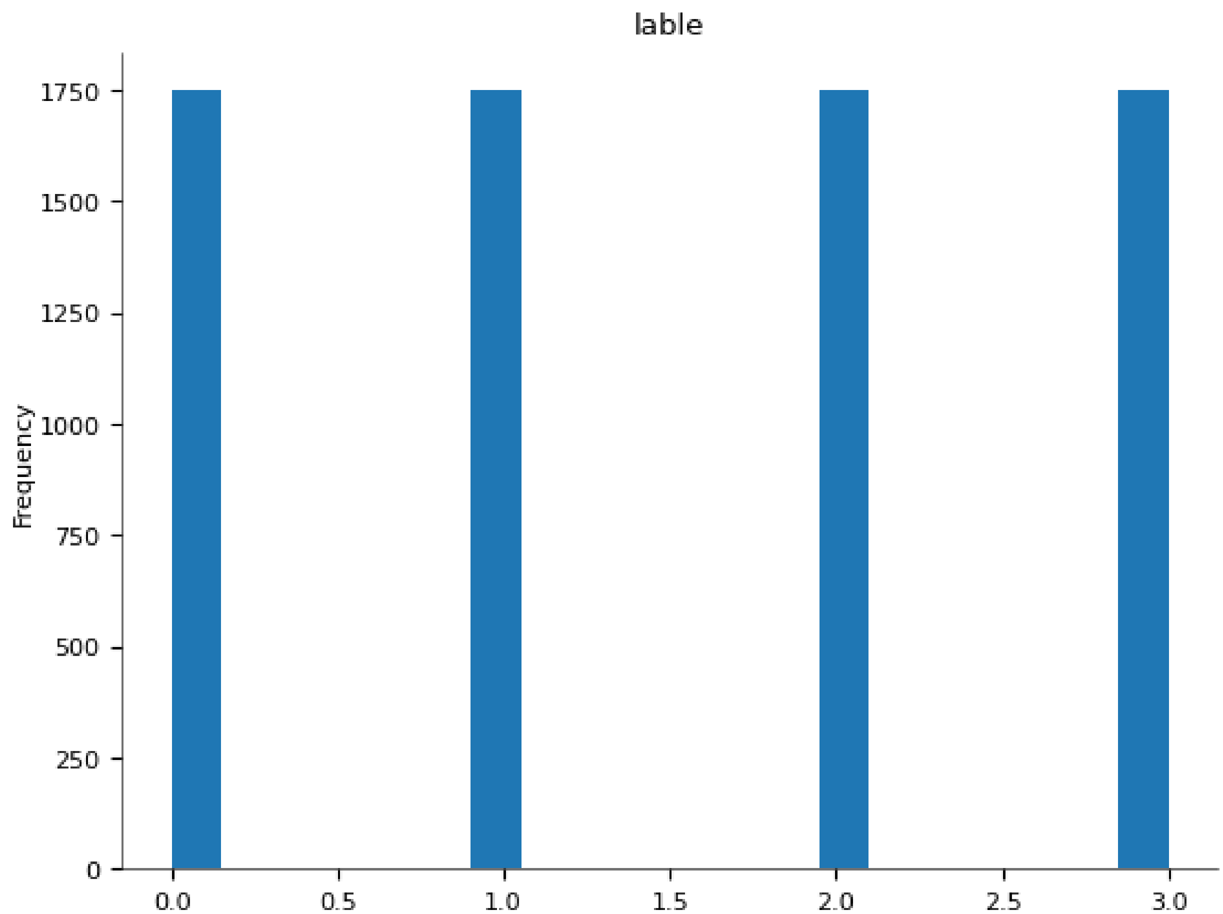


Values



teenagers).

distributions



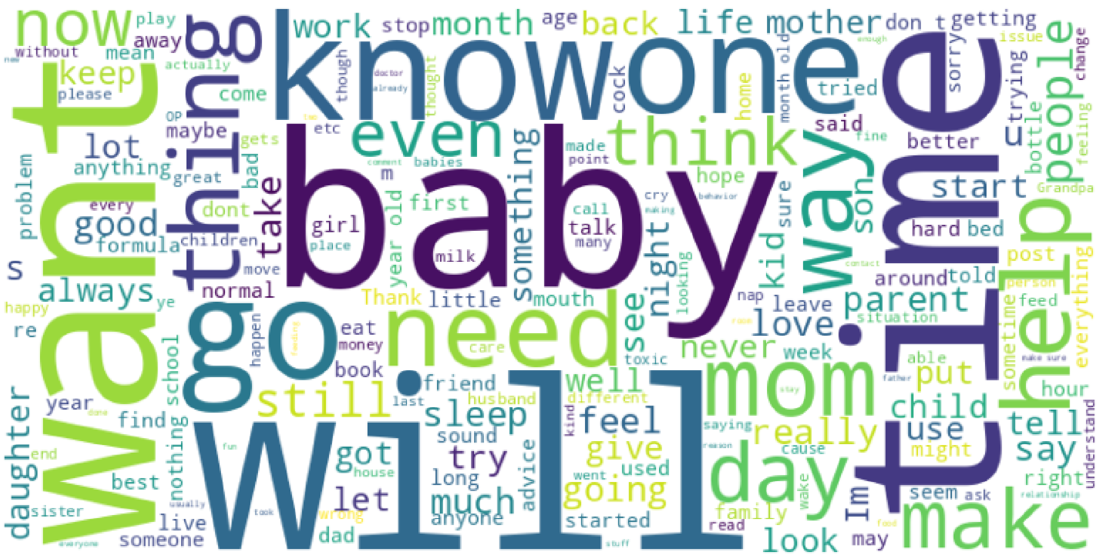
#EDA

[ ]:
```
from collections import Counter
from wordcloud import WordCloud
```

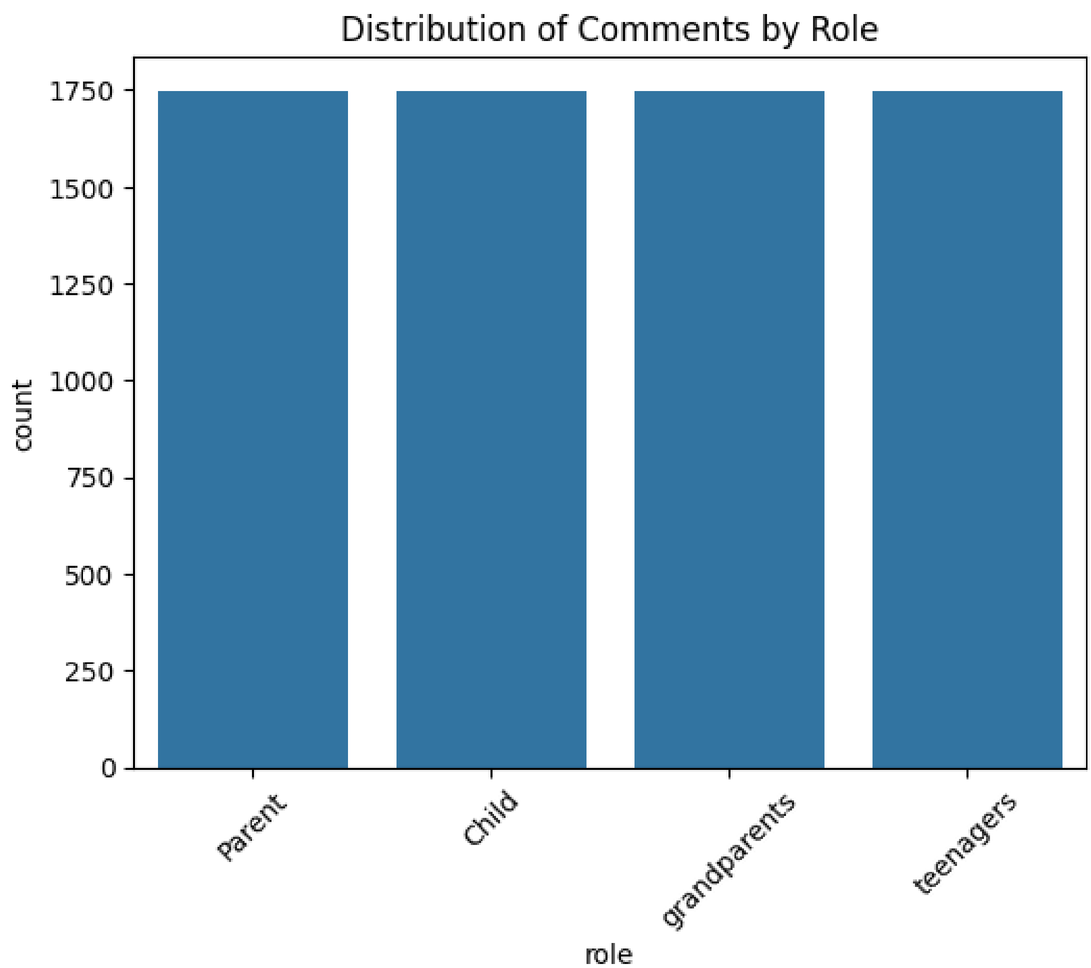
[ ]:
```
all_words="".join(df_complete.comment.tolist())
wordcloud=WordCloud(width=800,height=400,background_color="white").
 ↪generate(all_words)
plt.figure(figsize=(10 , 5))
plt.imshow(wordcloud,interpolation="bilinear")
plt.axis("off")
```

```
plt.title("WordCloud of All Comments")
plt.show()
```

WordCloud of All Comments



```
[ ]: sns.countplot(data=df_complete,x="role")
plt.title("Distribution of Comments by Role")
plt.xticks(rotation=45)
plt.show()
```

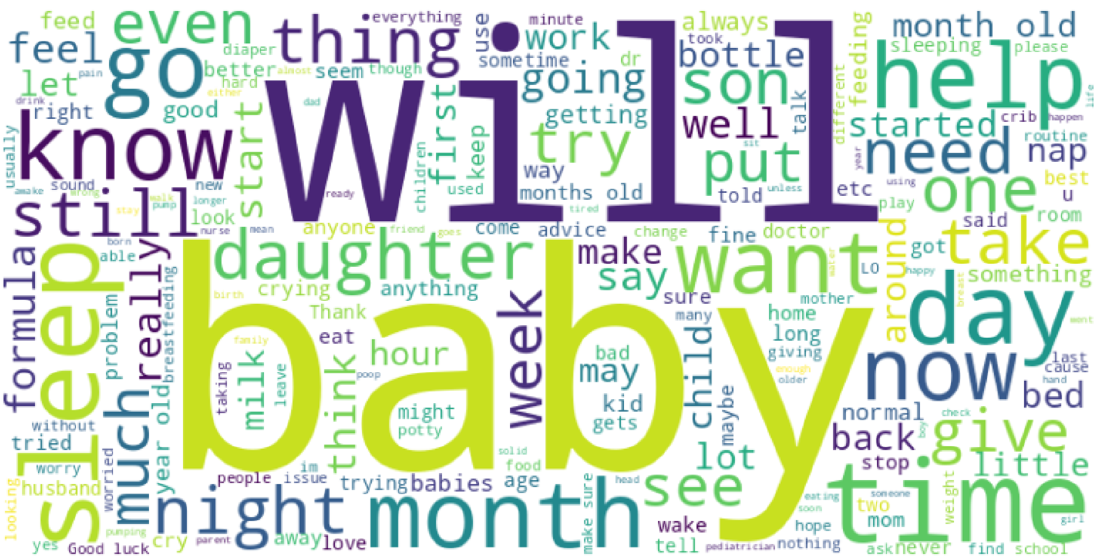


[ ]:
```
def findwords(col):
  parent_comments=df_complete[df_complete["role"]==col]["comment"]
  all_words="".join(parent_comments)
  wordcloud=WordCloud(width=800,height=400,background_color="white").
↪generate(all_words)
  plt.figure(figsize=(10 , 5))
  plt.imshow(wordcloud,interpolation="bilinear")
  plt.axis("off")
  plt.title(f"WordCloud of {col}")
```

```
plt.show()
```

[ ]:
```
findwords("Parent")
```

WordCloud of Parent



[ ]:
```
findwords("teenagers")
```

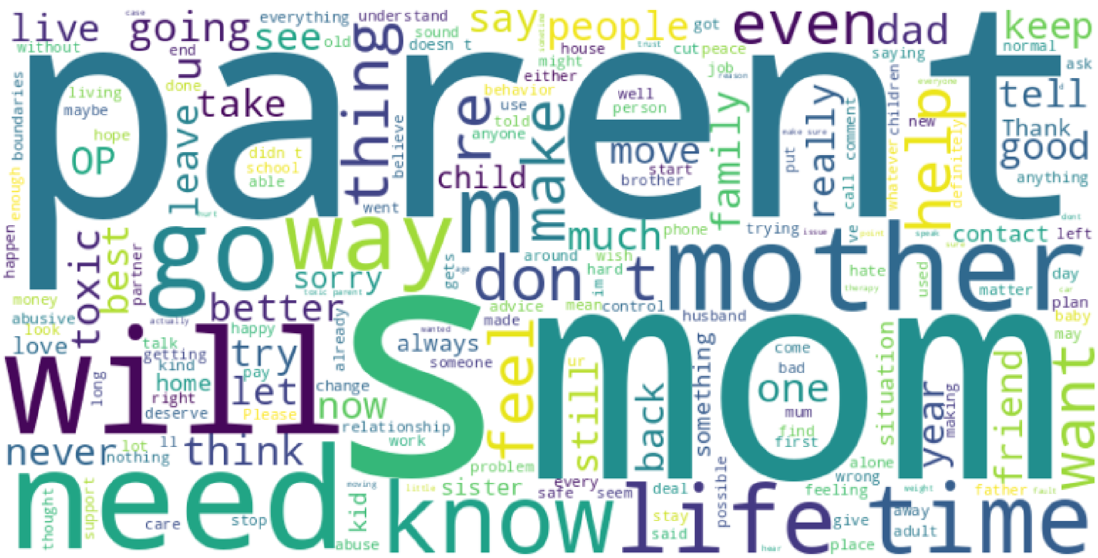
WordCloud of teenagers


```
[ ]: findwords("grandparents")
```

WordCloud of grandparents

```
[ ]: findwords("Child")
```

WordCloud of Child



#CREAT MODEL

```
[27]: le=LabelEncoder()

df_complete["lable"]=le.fit_transform(df_complete["role"])
```

```
[58]: train_texts,test_texts,train_lables,test_lables=train_test_split(df_complete["comment"],df_comp
 ↪1,random_state=42)
```

```
[59]: tokenizer=BertTokenizer.from_pretrained("bert-base-uncased")

train_encodings=tokenizer(list(train_texts),truncation=True, padding=True)

test_encodings=tokenizer(list(test_texts),truncation=True, padding=True)
```

```
[60]: class RedditDataset(torch.utils.data.Dataset):

  def __init__ (self,encodings, labels):

    self.encodings=encodings
```

```
    self.labels=labels
  def __len__(self):
    return len(self.labels)
  def __getitem__(self,idx):
    item={key:torch.tensor(val[idx]) for key,val in self.encodings.items()}
    item["labels"]=torch.tensor(self.labels[idx])
    return item
```

[61]:
```
train_dataset=RedditDataset(train_encodings,train_lables.tolist())
test_dataset=RedditDataset(test_encodings,test_lables.tolist())
```

[53]:
```
model=BertForSequenceClassification.
 ↪from_pretrained("bert-base-uncased",num_labels=len(le.classes_))
```

```
Some weights of BertForSequenceClassification were not initialized from the
model checkpoint at bert-base-uncased and are newly initialized:
['classifier.bias', 'classifier.weight']
You should probably TRAIN this model on a down-stream task to be able to use it
for predictions and inference.
```

[33]:
```
import os
os.environ["WANDB_DISABLED"]="true"
```

[67]:
```
training_args = TrainingArguments(
    output_dir="./results",
    per_device_train_batch_size=8,
    per_device_eval_batch_size=8,
    warmup_steps=200,
    weight_decay=0.01,
    logging_dir="./logs",
    logging_steps=25
)
```

```
trainer = Trainer(
    model=model,
    args=training_args,
    train_dataset=train_dataset,
    eval_dataset=test_dataset
)

trainer.train()
```

```
Using the `WANDB_DISABLED` environment variable is deprecated and will be
removed in v5. Use the --report_to flag to control the integrations used for
logging result (for instance --report_to none).
```

```
<IPython.core.display.HTML object>
```

[67]:
```
TrainOutput(global_step=2361, training_loss=0.03904784380312148,
metrics={'train_runtime': 1799.8702, 'train_samples_per_second': 10.494,
'train_steps_per_second': 1.312, 'total_flos': 4969730854453248.0, 'train_loss':
0.03904784380312148, 'epoch': 3.0})
```

[68]:
```
predictions = trainer.predict(test_dataset)
preds = torch.argmax(torch.tensor(predictions.predictions), axis=1)
print(classification_report(test_lables, preds, target_names=le.classes_))
```

```
<IPython.core.display.HTML object>
```

```
              precision    recall  f1-score   support

       Child       0.92      0.88      0.90       180
```

```
      Parent       0.93      0.97      0.95       180
grandparents       0.96      0.93      0.95       178
   teenagers       0.89      0.92      0.91       162

    accuracy                           0.93       700
   macro avg       0.93      0.93      0.93       700
weighted avg       0.93      0.93      0.93       700
```

[69]:
```python
accuracy = accuracy_score(test_lables, preds)
weighted_precision = precision_score(test_lables, preds, average='weighted')
weighted_recall = recall_score(test_lables, preds, average='weighted')
weighted_f1 = f1_score(test_lables, preds, average='weighted')

print(f"Accuracy: {accuracy}")
print("Weighted Precision:", weighted_precision)
print("Weighted Recall:", weighted_recall)
print("Weighted F1 Score:", weighted_f1)
```

```
Accuracy: 0.9257142857142857
Weighted Precision: 0.9260810957779484
Weighted Recall: 0.9257142857142857
Weighted F1 Score: 0.9256205829843341
```

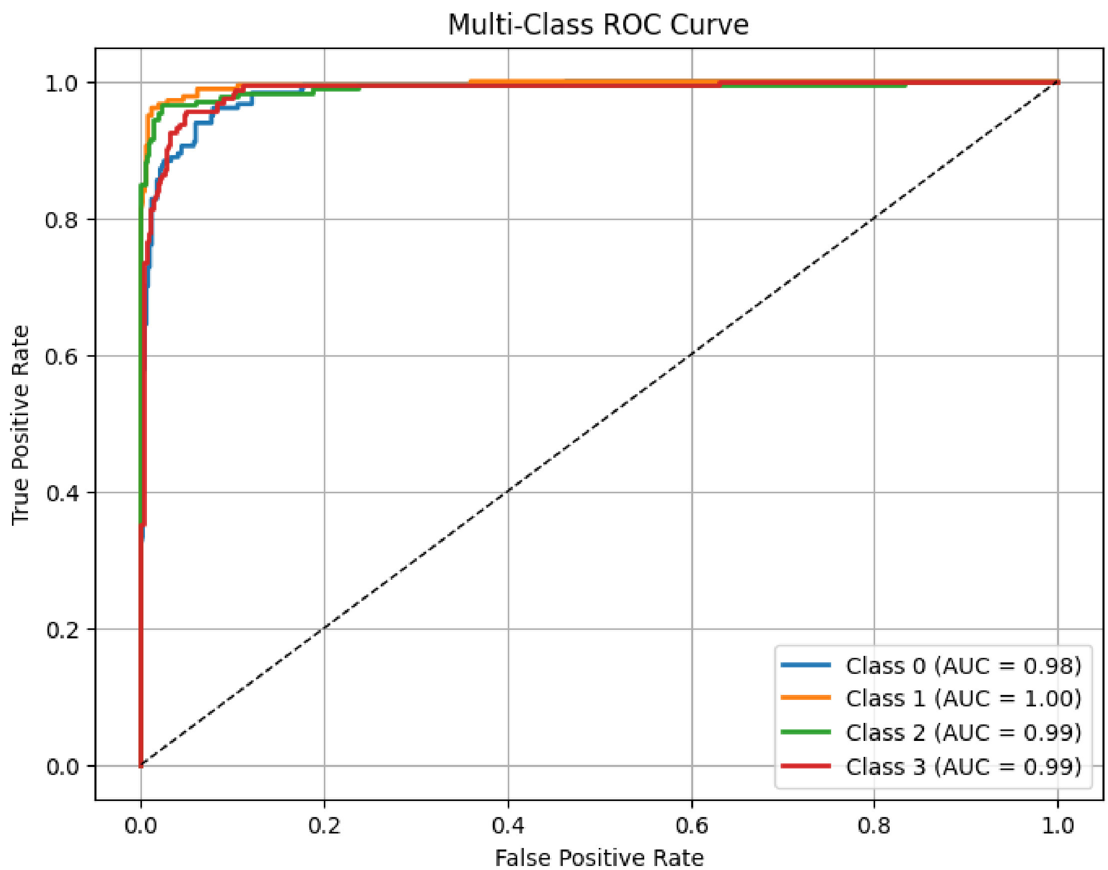
[70]:
```python
logits = torch.tensor(predictions.predictions)
probs = F.softmax(logits, dim=1)
child_probs = probs[:, 1].numpy()
```

[74]:
```python
y_true = le.fit_transform(test_lables)
```

[76]:
```python
n_classes = len(le.classes_)
y_true_bin = label_binarize(y_true, classes=range(n_classes))
```

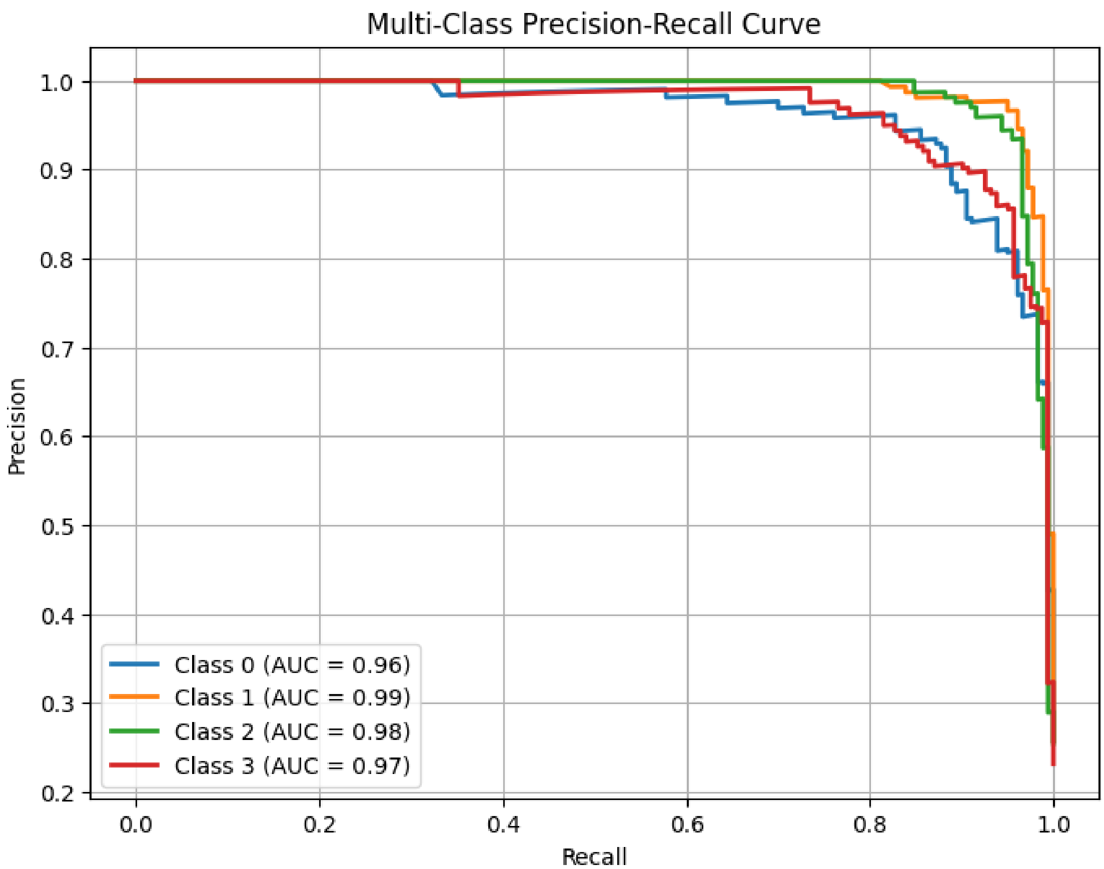
[77]:

```
plt.figure(figsize=(8, 6))

for i in range(n_classes):
    fpr, tpr, _ = roc_curve(y_true_bin[:, i], probs[:, i])
    roc_auc = auc(fpr, tpr)
    plt.plot(fpr, tpr, lw=2, label=f'Class {le.classes_[i]} (AUC = {roc_auc:.
 ↪2f})')

plt.plot([0, 1], [0, 1], 'k--', lw=1)
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.title('Multi-Class ROC Curve')
plt.legend(loc='lower right')
plt.grid()
plt.show()
```



[78]:
```
plt.figure(figsize=(8, 6))

for i in range(n_classes):
    precision, recall, _ = precision_recall_curve(y_true_bin[:, i], probs[:, i])
    pr_auc = average_precision_score(y_true_bin[:, i], probs[:, i])
    plt.plot(recall, precision, lw=2, label=f'Class {le.classes_[i]} (AUC =␣
↪{pr_auc:.2f})')

plt.xlabel('Recall')
plt.ylabel('Precision')
plt.title('Multi-Class Precision-Recall Curve')
```

```
plt.legend(loc='lower left')
plt.grid()
plt.show()
```



Codes for 2 categories classification

```
[ ]: import pandas as pd
     import numpy as np
     from sklearn.model_selection import train_test_split
     from sklearn.preprocessing import LabelEncoder
     from transformers import BertTokenizer
     import torch
     from sklearn.preprocessing import label_binarize
     from transformers import BertForSequenceClassification, Trainer,␣
      ↪TrainingArguments
     import matplotlib.pyplot as plt
     import seaborn as sns
     from sklearn.metrics import (
         roc_auc_score, roc_curve, precision_recall_curve, auc,
         confusion_matrix, classification_report , average_precision_score
     )
     import torch.nn.functional as F
     from sklearn.metrics import classification_report, accuracy_score,␣
      ↪precision_score, recall_score, f1_score
```

```
[ ]: df1 = pd.read_csv('/content/child_comments.csv')
     df2 = pd.read_csv('/content/parents.csv')
```

```
[ ]: df2=df2.iloc[:1749]
     df1=df1.iloc[:1749]
```

```
[ ]: df_com = pd.concat([df1,df2],ignore_index=True)
```

```
[ ]: df_com
```

```
[ ]:       role                                          comment
     0     Child        My dad wouldnt let me remove my wisdom teeth
     1     Child      I don't want my toxic father at my graduation
```

```
2       Child                                        Mom Situation
3       Child  I want to move out from my toxic mother, but I...
4       Child  My parents won't let me sleep at my boyfriend'...
...       ...                                                ...
3493   Parent             Why has my baby stopped self settling?
3494   Parent                     Baby at water park? Life vest?
3495   Parent  Is anyone looking for a family to adopt their ...
3496   Parent     What is the best pram to buy for a new mother?
3497   Parent                      Looking for adoptive parents.

[3498 rows x 2 columns]
```

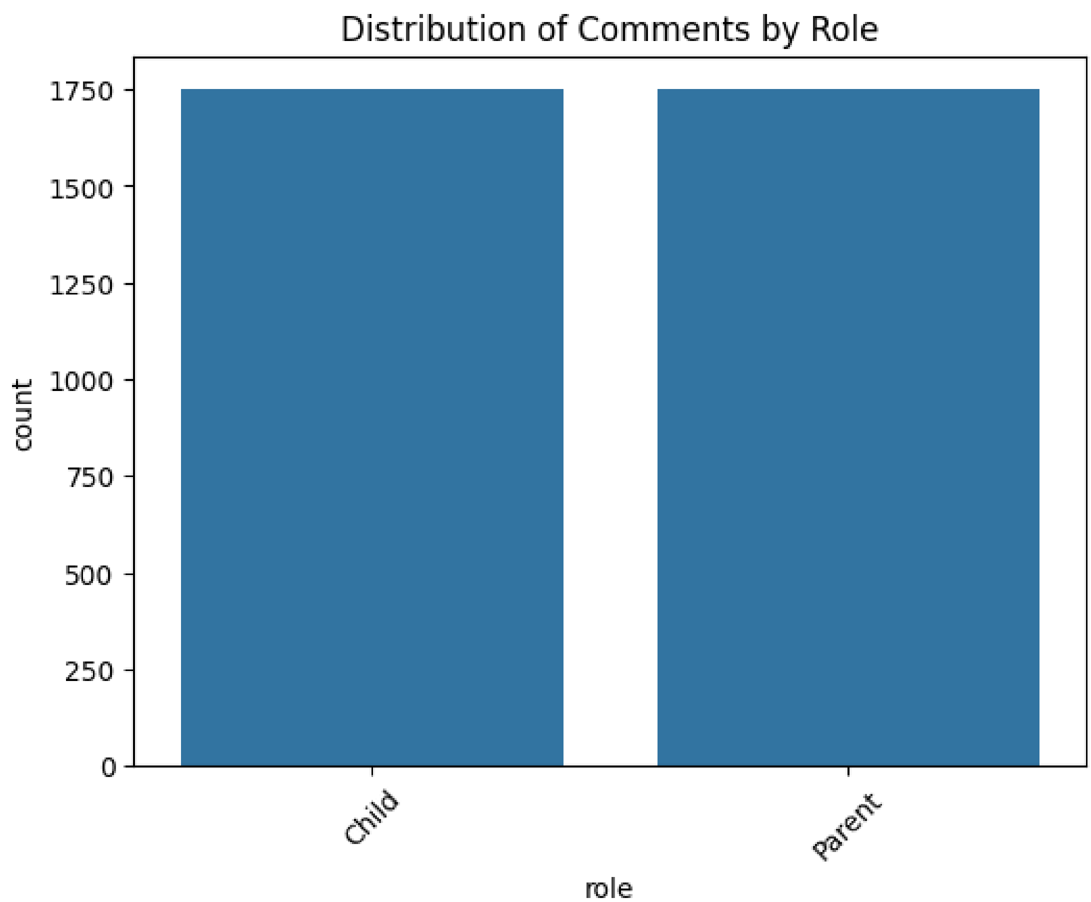
#EDA

[ ]:
```
from collections import Counter
from wordcloud import WordCloud
```

[ ]:
```
all_words="".join(df_com.comment.tolist())
wordcloud=WordCloud(width=800,height=400,background_color="white").
 ↪generate(all_words)
plt.figure(figsize=(10 , 5))
plt.imshow(wordcloud,interpolation="bilinear")
plt.axis("off")
plt.title("WordCloud of All Comments")
plt.show()
```

[ ]:

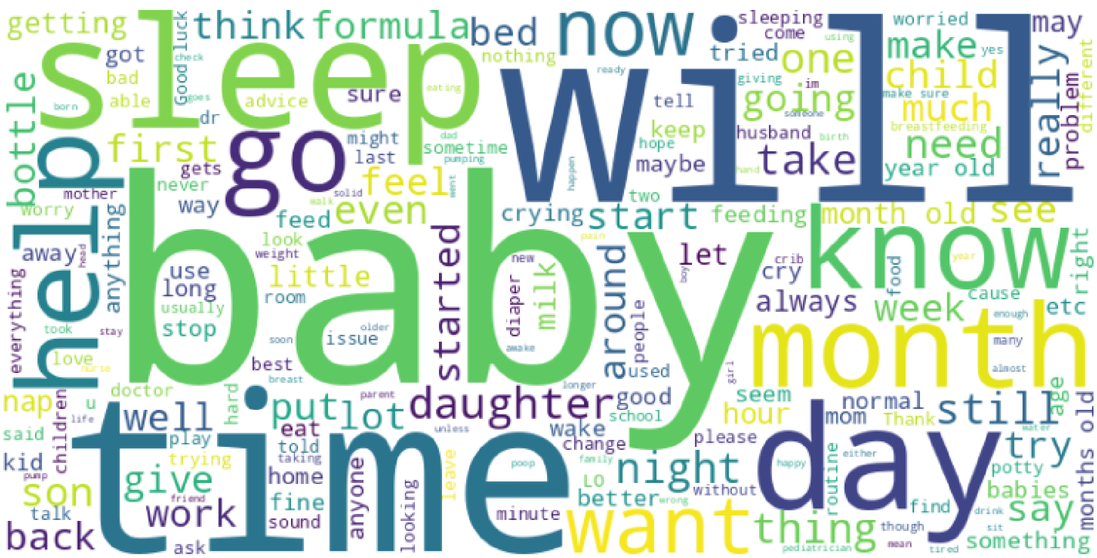
```
sns.countplot(data=df_com,x="role")
plt.title("Distribution of Comments by Role")
plt.xticks(rotation=45)
plt.show()
```




```
[ ]: def findwords(col):
       parent_comments=df_com[df_com["role"]==col]["comment"]
       all_words="".join(parent_comments)
       wordcloud=WordCloud(width=800,height=400,background_color="white").
     ↪generate(all_words)
       plt.figure(figsize=(10 , 5))
       plt.imshow(wordcloud,interpolation="bilinear")
       plt.axis("off")
       plt.title(f"WordCloud of {col}")
       plt.show()
```

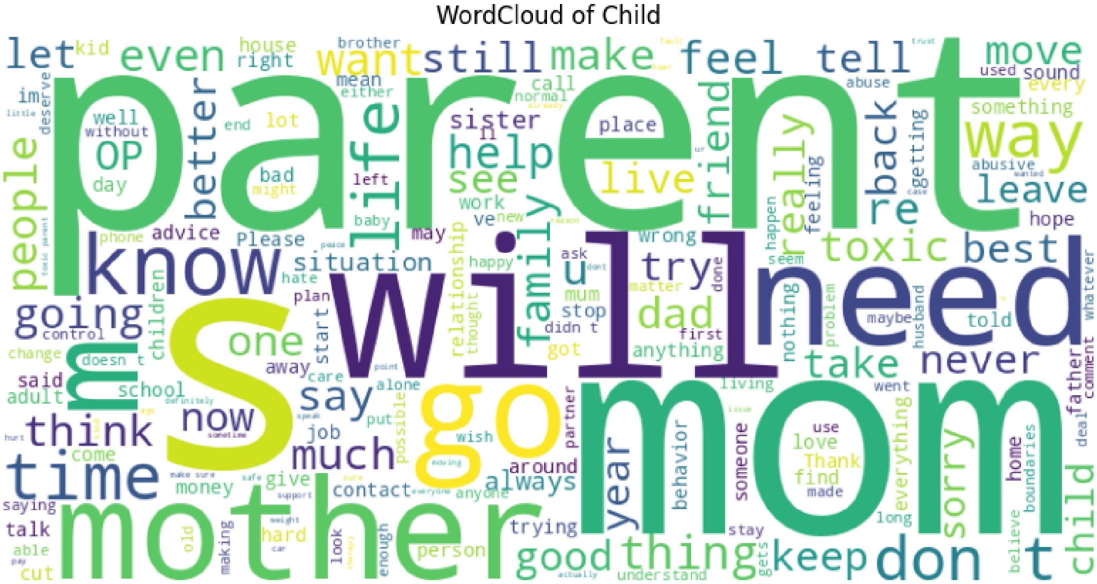
```
[ ]: findwords("Parent")
```

WordCloud of Parent



```
[ ]: findwords("Child")
```

WordCloud of Child



#creat model

```
[ ]: le = LabelEncoder()
df_com['label'] = le.fit_transform(df_com['role'])

train_texts, val_texts, train_labels, val_labels =␣
 ↪train_test_split(df_com['comment'], df_com['label'], test_size=0.2,␣
 ↪random_state=42)
```

```
[ ]: tokenizer = BertTokenizer.from_pretrained("bert-base-uncased")

train_encodings = tokenizer(list(train_texts), truncation=True, padding=True)
val_encodings = tokenizer(list(val_texts), truncation=True, padding=True)
```

```
/usr/local/lib/python3.11/dist-packages/huggingface_hub/utils/_auth.py:94:
UserWarning:
The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab
(https://huggingface.co/settings/tokens), set it as secret in your Google Colab
and restart your session.
You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access
public models or datasets.
  warnings.warn(
```

```
tokenizer_config.json:   0%|          | 0.00/48.0 [00:00<?, ?B/s]
```

```
vocab.txt:   0%|          | 0.00/232k [00:00<?, ?B/s]
```

```
tokenizer.json:   0%|          | 0.00/466k [00:00<?, ?B/s]
```

```
config.json:   0%|          | 0.00/570 [00:00<?, ?B/s]
```

[ ]:
```
class FamilyDataset(torch.utils.data.Dataset):
    def __init__(self, encodings, labels):
        self.encodings = encodings
        self.labels = labels
    def __len__(self):
        return len(self.labels)
    def __getitem__(self, idx):
        item = {key: torch.tensor(val[idx]) for key, val in self.encodings.
 ↪items()}
        item['labels'] = torch.tensor(self.labels[idx])
        return item

train_dataset = FamilyDataset(train_encodings, train_labels.tolist())
val_dataset = FamilyDataset(val_encodings, val_labels.tolist())
```

[ ]:
```
model = BertForSequenceClassification.from_pretrained("bert-base-uncased",␣
 ↪num_labels=len(le.classes_))
```

```
Xet Storage is enabled for this repo, but the 'hf_xet' package is not installed.
Falling back to regular HTTP download. For better performance, install the
package with: `pip install huggingface_hub[hf_xet]` or `pip install hf_xet`
WARNING:huggingface_hub.file_download:Xet Storage is enabled for this repo, but
the 'hf_xet' package is not installed. Falling back to regular HTTP download.
For better performance, install the package with: `pip install
huggingface_hub[hf_xet]` or `pip install hf_xet`
```

```
model.safetensors:   0%|          | 0.00/440M [00:00<?, ?B/s]
```

```
Some weights of BertForSequenceClassification were not initialized from the
model checkpoint at bert-base-uncased and are newly initialized:
['classifier.bias', 'classifier.weight']
You should probably TRAIN this model on a down-stream task to be able to use it
for predictions and inference.
```

[ ]:
```python
import os
os.environ["WANDB_DISABLED"] = "true"

training_args = TrainingArguments(
    output_dir="./results",
    per_device_train_batch_size=16,
    per_device_eval_batch_size=16,
    warmup_steps=200,
    weight_decay=0.01,
    logging_dir="./logs",
    logging_steps=50
)

trainer = Trainer(
    model=model,
    args=training_args,
    train_dataset=train_dataset,
    eval_dataset=val_dataset
)

trainer.train()
```

```
Using the `WANDB_DISABLED` environment variable is deprecated and will be
removed in v5. Use the --report_to flag to control the integrations used for
logging result (for instance --report_to none).
```

<IPython.core.display.HTML object>

[ ]: TrainOutput(global_step=525, training_loss=0.18593088536035446, metrics={'train_runtime': 756.1611, 'train_samples_per_second': 11.101, 'train_steps_per_second': 0.694, 'total_flos': 2208554198691840.0, 'train_loss': 0.18593088536035446, 'epoch': 3.0})

[ ]:
```
predictions = trainer.predict(val_dataset)
preds = torch.argmax(torch.tensor(predictions.predictions), axis=1)
print(classification_report(val_labels, preds, target_names=le.classes_))
```

<IPython.core.display.HTML object>

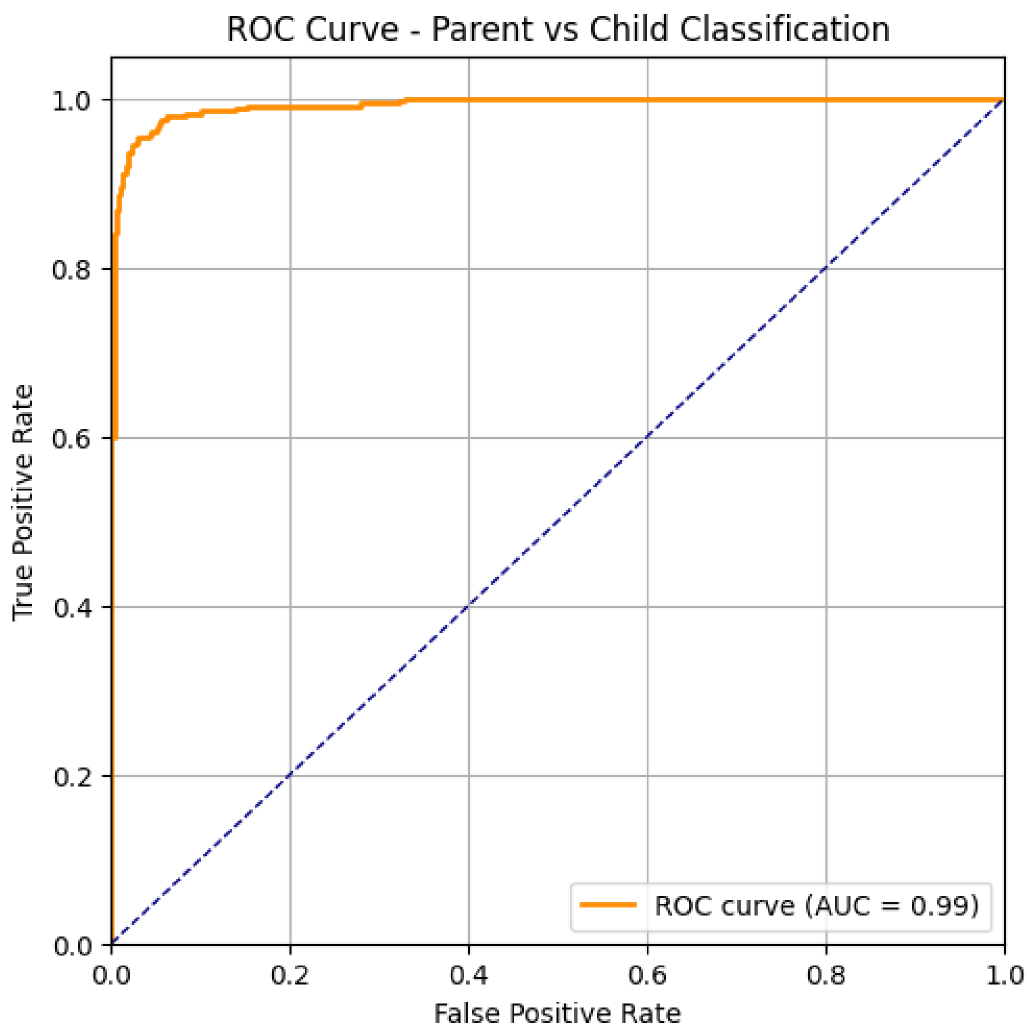
```
              precision    recall  f1-score   support

       Child       0.96      0.97      0.96       375
      Parent       0.97      0.95      0.96       325

    accuracy                           0.96       700
   macro avg       0.96      0.96      0.96       700
weighted avg       0.96      0.96      0.96       700
```

[ ]:
```
accuracy = accuracy_score(val_labels, preds)
weighted_precision = precision_score(val_labels, preds, average='weighted')
weighted_recall = recall_score(val_labels, preds, average='weighted')
weighted_f1 = f1_score(val_labels, preds, average='weighted')

print(f"Accuracy: {accuracy}")
print("Weighted Precision:", weighted_precision)
print("Weighted Recall:", weighted_recall)
print("Weighted F1 Score:", weighted_f1)
```

```
Accuracy: 0.9614285714285714
Weighted Precision: 0.961559132197777
Weighted Recall: 0.9614285714285714
Weighted F1 Score: 0.9613969536198768
```

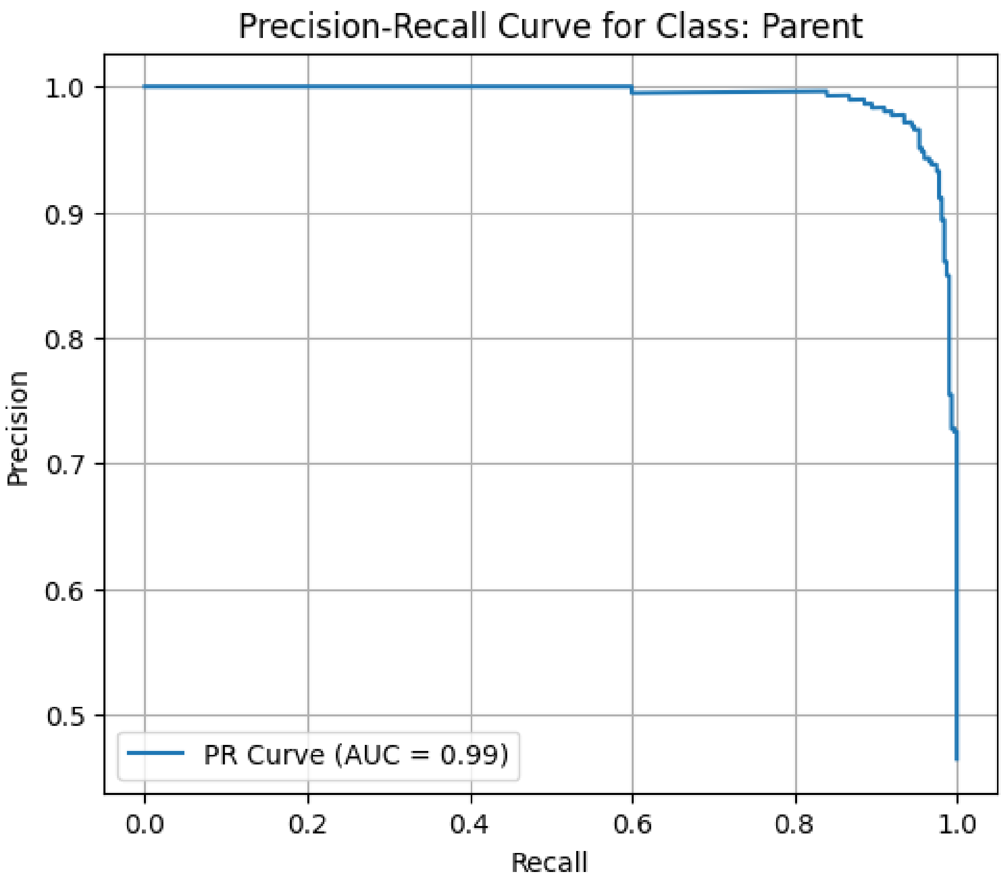
[ ]:
```
logits = torch.tensor(predictions.predictions)
probs = F.softmax(logits, dim=1)
child_probs = probs[:, 1].numpy()
parent_probs = probs[:, 0].numpy()
```

[ ]:
```
fpr, tpr, thresholds = roc_curve(val_labels, child_probs)
roc_auc = auc(fpr, tpr)

plt.figure(figsize=(6, 6))
plt.plot(fpr, tpr, color='darkorange', lw=2, label=f'ROC curve (AUC = {roc_auc:.
 ↪2f})')
plt.plot([0, 1], [0, 1], color='navy', lw=1, linestyle='--')
plt.xlim([0.0, 1.0])
plt.ylim([0.0, 1.05])
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.title('ROC Curve - Parent vs Child Classification')
plt.legend(loc="lower right")
plt.grid()
plt.show()
```



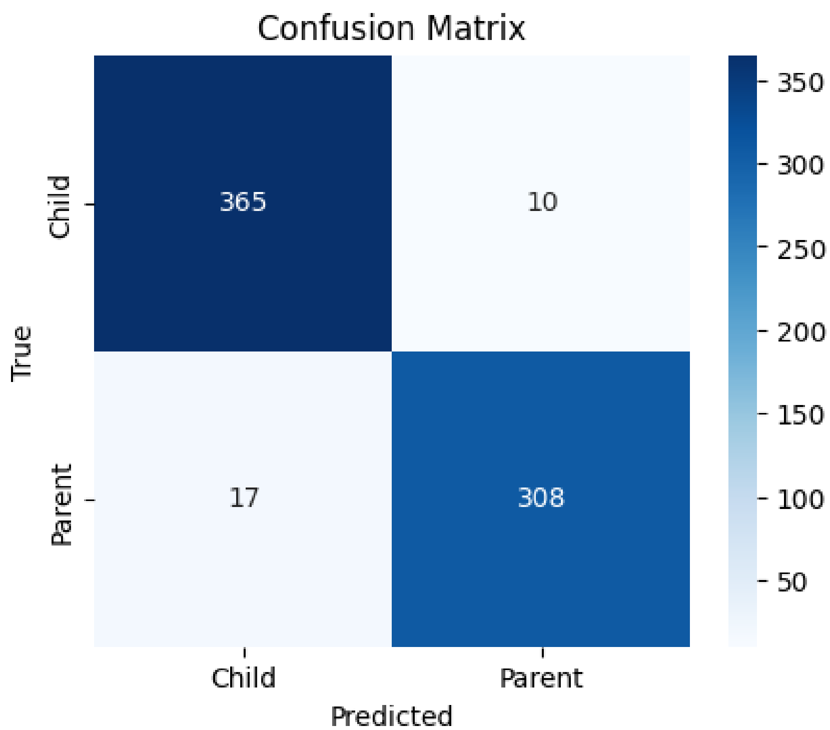
[ ]:
```
precision, recall, _ = precision_recall_curve(val_labels, child_probs,␣
  ↪pos_label=1)
pr_auc = auc(recall, precision)

plt.figure(figsize=(6, 5))
plt.plot(recall, precision, label=f"PR Curve (AUC = {pr_auc:.2f})")
plt.xlabel("Recall")
plt.ylabel("Precision")
```

```
plt.title("Precision-Recall Curve for Class: Parent")
plt.legend()
plt.grid()
plt.show()
```



[ ]:
```
preds = torch.argmax(logits, axis=1)


cm = confusion_matrix(val_labels, preds)
labels = ['Child', 'Parent']
```

```
plt.figure(figsize=(5, 4))
sns.heatmap(cm, annot=True, fmt='d', cmap="Blues", xticklabels=labels,␣
 ↪yticklabels=labels)
plt.xlabel("Predicted")
plt.ylabel("True")
plt.title("Confusion Matrix")
plt.show()
```




```
[ ]:
```

BIBLIOGRAPHY

Abbasi, A. (2007). *Affect Intensity Analysis of Dark Web Forums*. (Thèse de doctorat). University of Maryland. Available at: https://www.cs.cornell.edu/home/llee/omsa/omsa.pdf.

Abrams, D. et Hogg, M. A. (1990). *An introduction to the social identity approach*.

Ajzen, I. (1985). *Action Control From Cognition to Behavior*.

Ajzen, I. et Fishbein, M. (1975). A bayesian analysis of attribution processes. *Psychological Bulletin*, *82*. http://dx.doi.org/10.1037/h0076477

Ajzen, I. et Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology*, *22*. http://dx.doi.org/10.1016/0022-1031(86)90045-4

Bagozzi, R. P. (1992). The self-regulation of attitudes, intentions, and behavior. *Social psychology quarterly*, 178–204.

Bagozzi, R. P. (2006). Explaining consumer behavior and consumer action: From fragmentation to unit. *Seoul Journal of Business*, *12*(2), 111–143.

Bagozzi, R. P., Dholakia, U. M. et Pearo, L. R. (2007). Antecedents and consequences of online social interactions. http://dx.doi.org/10.1080/15213260709336804

Bagozzi, R. P. et Warshaw, P. R. (1990). Trying to consume. *Journal of consumer research*, *17*(2), 127–140.

Belanche, D., Casaló, L. V. et Flavián, C. (Year of Publication). Title of the article. *Journal Name*, *Volume Number*(Issue Number), Page Range. http://dx.doi.org/DOINumber

Benzarti, I., Mili, H., de Carvalho, R. M. et Leshob, A. (2022). Domain engineering for customer experience management. *Innovations in Systems and Software Engineering*, *18*. http://dx.doi.org/10.1007/s11334-021-00426-2

Berry, J. W. (1997). *Immigration, Acculturation, and Adaptation*. Cambridge: Applied Psychology: An International Review.

Blei, D. M. (2012). Probabilistic topic models. *Commun. ACM*, *55*(4), 77–84. http://dx.doi.org/10.1145/2133806.2133826. Récupéré de https://doi.org/10.1145/2133806.2133826

Brown, R. (Year of Publication). *Title of the Book*. Publisher Location: Publisher Name.

Burke, P. J. et Stets, J. E. (2009). *Identity Theory*. New York: Oxford University Press.

Burke, P. J. et Stets, J. E. (2023). *Identity Theory: Second Edition, Revised and Expanded*. http://dx.doi.org/10.1093/oso/9780197617182.001.0001

Chong, Y., Ding, Y., Yan, Q. et Pan, S. (2020). Graph-based semi-supervised learning: A review.

Neurocomputing, *408*, 216–230. http://dx.doi.org/https://doi.org/10.1016/j.neucom.2019.12.130. Récupéré de https://www.sciencedirect.com/science/article/pii/S0925231220304938

Cooper, K. et Authors, O. (Year of Publication). Understanding consumers' perceptions of the value of vegetarian food: An analysis of over 120,000 tweets related to vegetarianism. *Journal Name*, *Volume Number*(Issue Number), Page Range. http://dx.doi.org/DOINumber

Crenshaw, K. (1994). Mapping the margins: Intersectionality, identity politics, and violence against women of color. In M. A. Fineman et R. Mykitiuk (dir.), *The Public Nature of Private Violence* 93–118. New York: Routledge.

Cruwys, T., Haslam, S. A., Dingle, G. A., Haslam, C. et Jetten, J. (2014). Depression and social identity: An integrative review. *Personality and Social Psychology Review*, *18*. http://dx.doi.org/10.1177/1088868314523839

Dalayya, S. et Authors, O. (Year of Publication). Analyzing social media data to understand human perception and sentiment regarding plant-based diets and cancer using the tf-idf method. *Journal Name*, *Volume Number*(Issue Number), Page Range. http://dx.doi.org/DOINumber

Deaux, K. (1992). *Social Identity*. New York: Springer.

Deaux, K. et Verkuyten, M. (2014). The social psychology of multiculturalism: Identity and intergroup relations. *European Journal of Social Psychology*, *44*(1), 1–13. http://dx.doi.org/10.1002/ejsp.2014

Dingle, G. A., Brander, C., Ballantyne, J. et Baker, F. A. (2013). 'to be heard': The social and mental health benefits of choir singing for disadvantaged adults. *Psychology of Music*, *41*. http://dx.doi.org/10.1177/0305735611430081

Ellemers, N., Kortekaas, P. et Ouwerkerk, J. W. (1999). Self-categorisation, commitment to the group and group self-esteem as related but distinct aspects of social identity. *European Journal of Social Psychology*, *29*. http://dx.doi.org/10.1002/(sici)1099-0992(199903/05)29:2/3<371::aid-ejsp932>3.0.co;2-u

Farrell, A. (2018). Accuracy of online discussion forums on common childhood ailments. *Journal of the Medical Library Association (JMLA)*, *106*(4), 480–486. http://dx.doi.org/10.5195/jmla.2018.355. Récupéré de https://jmla.pitt.edu/ojs/jmla/article/view/355

Harmon, D. (2004). Title of the article or work. *Journal or Source*.

Haslam, C. et Authors, O. (Year of Publication). Title of the article or chapter. *Journal Name*, *Volume Number*(Issue Number), Page Range. http://dx.doi.org/DOINumber

Hogg, M. A. (2000). Subjective uncertainty reduction through self-categorization: A motivational theory of social identity processes. *European review of social psychology*, *11*(1), 223–255.

Hogg, M. A. et Authors, O. (Year of Publication). Title of the article. *Journal Name*, *Volume Number*(Issue

Number), Page Range. `http://dx.doi.org/DOINumber`

Hogg, M. A. et Reid, S. A. (2006). Social identity, self-categorization, and the communication of group norms. *Communication Theory*, *16*(1), 7–30.

James, W. (1890). The consciousness of self.

Jetten, J. et Authors, O. (2009). Title of the article. *Journal Name*, *Volume Number*(Issue Number), Page Range. `http://dx.doi.org/DOINumber`

Jorgensen, B. S. et Stedman, R. C. (2001). Sense of place as an attitude: Lakeshore owners attitudes toward their properties. *Journal of Environmental Psychology*, *21*, 233–248. `http://dx.doi.org/10.1006/jevp.2001.0227`

Jorgensen, B. S. et Stedman, R. C. (2006). A comparative analysis of the relationship between sense of place and individual well-being in urban, suburban, and rural settings. *Journal of Environmental Psychology*, *26*, 36–44. `http://dx.doi.org/10.1016/j.jenvp.2005.09.001`

Kost, T. (2012). Title of the article or work. *Journal or Source*.

Liu, B. (2012a). Opinion spam detection. In *Sentiment Analysis and Opinion Mining*, volume 5 de *Synthesis Lectures on Human Language Technologies* 113–125. Morgan & Claypool Publishers

Liu, B. (2012b). *Sentiment Analysis and Opinion Mining* (1 éd.). Synthesis Lectures on Human Language Technologies. Springer Cham. Copyright ©Springer Nature Switzerland AG 2012. Topics: Artificial Intelligence, Natural Language Processing (NLP), Computational Linguistics, `http://dx.doi.org/10.1007/978-3-031-02145-9`. Récupéré de `https://doi.org/10.1007/978-3-031-02145-9`

Liu, B. (2020a). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions* (2nd éd.). Cambridge University Press.

Liu, B. (2020b). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press, (2nd éd.).

Liu, B. (2020c). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press, (2nd éd.).

Mili, H., Benzarti, I., Meurs, M. J., Obaid, A., Gonzalez-Huerta, J., Haj-Salem, N. et Boubaker, A. (2016). *Context aware customer experience management: A development framework based on ontologies and computational intelligence*, volume 639

Nguyen, A.-M. D. et Benet-Martínez, V. (2013). Biculturalism and adjustment: A meta-analysis. *Journal of Cross-Cultural Psychology*, *44*(1), 122–159. `http://dx.doi.org/10.1177/0022022111435097`

Olavsrud, M. A. (2023). Topic modeling as a tool for exploring and analyzing large datasets. *Journal Name*, *Volume Number*(Issue Number), Page Range. `http://dx.doi.org/DOINumber`

Papadimitriou, C. H., Tamaki, H., Raghavan, P. et Vempala, S. (1998). Latent semantic indexing: a probabilistic analysis. Dans *Proceedings of the Seventeenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, PODS '98, p. 159–168., New York, NY, USA. Association for Computing Machinery. `http://dx.doi.org/10.1145/275487.275505`. Récupéré de `https://doi.org/10.1145/275487.275505`

Park, S., Strover, S., Choi, J. et Schnell, M. (2023). Mind games: A temporal sentiment analysis of the political messages of the internet research agency on facebook and twitter. *New Media and Society*, *25*. `http://dx.doi.org/10.1177/14614448211014355`

Read, J., Pfahringer, B., Holmes, G. et Frank, E. (2011). Classifier chains for multi-label classification. *Machine Learning*, *85*(3), 333–359. `http://dx.doi.org/10.1007/s10994-011-5256-5`. Récupéré de `https://doi.org/10.1007/s10994-011-5256-5`

Scannell, L. et Gifford, R. (2013). Personally relevant climate change: The role of place attachment and local versus global message framing in engagement. *Environment and Behavior*, *45*(1), 60–85. `http://dx.doi.org/10.1177/0013916511421196`

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. volume 25 de *Advances in Experimental Social Psychology* 1–65. Academic Press

Schwartz, S. J., Luyckx, K. et Vignoles, V. L. (2014). Handbook of identity theory and research. *Journal of Adolescence*, *37*(6), 769–777. `http://dx.doi.org/10.1016/j.adolescence.2014.04.010`

Shamoi, E. et Authors, O. (Year of Publication). Using twitter posts to gain insights into public opinion about a vegan (plant-based) diet. *Journal Name*, *Volume Number*(Issue Number), Page Range. `http://dx.doi.org/DOINumber`

Sherif, M. (1936). *The Psychology of Social Norms*. New York: Harper.

Steger, M. F. et Kashdan, T. B. (2009). Title of the article. *Journal Name*, *Volume Number*(Issue Number), Page Range. `http://dx.doi.org/DOINumber`

Stets, J. et Serpe, R. (2013). Identity theory. *Handbook of Social Psychology*, 31–60. `http://dx.doi.org/10.1007/978-94-007-6772-0_2`

Stets, J. E. (1995). Role identity theory. *Annual Review of Sociology*, *21*, 103–128.

Stets, J. E. et Serpe, R. T. (2016). *New directions in identity theory and research*. Oxford University Press.

Stets, J. E. et Serpe, R. T. (Year of Publication). Title of the article. *Name of the Journal*, *Volume Number*(Issue Number), Page Range. `http://dx.doi.org/DOINumber`

Streitfeld, D. (2012a). Title of the article or work. *Journal or Source*.

Streitfeld, D. (2012b). Title of the article or work. *Journal or Source*.

Tankard, M. E. et Paluck, E. L. (2016). Norm perception as a vehicle for social change. *Social Issues and Policy Review*, *10*(1), 181–211. `http://dx.doi.org/10.1111/sipr.12022`

Teh, P. L. et Yap, W. L. (2023). Exploring the beneficial effects of clustering in discussions related to veganism. *Journal Name*, *Volume Number*(Issue Number), Page Range. http://dx.doi.org/DOINumber

Turner, J. C. (1982). *Towards a Cognitive Redefinition of the Social Group*. Cambridge: Cambridge University Press.

Turner, J. C. (1991). *Social Influence*. Pacific Grove, CA: Thomson Brooks/Cole.

Turner, J. C., Brown, R. J. et Tajfel, H. (1979). Social comparison and group interest in ingroup favouritism. *European Journal of Social Psychology*, *9*. http://dx.doi.org/10.1002/ejsp.2420090207

Vallerand, R. J., Deshaies, P., Cuerrier, J.-P., Pelletier, L. G. et Mongeau, C. (1992). Ajzen and fishbein's theory of reasoned action as applied to moral behavior: A confirmatory analysis. *Journal of personality and social psychology*, *62*(1), 98.