MYER BRIGGS PERSONALITY TYPE CLASSIFICATION FROM SOCIAL MEDIA TEXT USING PRE-TRAINED LANGUAGE MODELS

BY

ORAMAH CHIDERA JULIET
BU/23A/PGS/8235

Being a pre data report presented in partial fulfillment of the requirements for the award of Masters of Science (M.Sc) Degree in Computer Science,

Department of Computer Science,

Faculty of Computing and Applied Sciences,

Baze University, Abuja.

March, 2024

CERTIFICATION

This Post-data report with title

MYER BRIGGS PERSONALITY TYPE CLASSIFICATION FROM SOCIAL MEDIA TEXT USING PRE-TRAINED LANGUAGE MODELS

Submitted

by

ORAMAH CHIDERA JULIET

Has partially satisfied the regulations governing the award of the degree of

Master of Science (M.Sc.) in Computer Science

Baze University, Abuja.

………………………………… ……………………..

Dr Usman Abubakar Date

Supervisor

…………………………………… ………………………..

Dr. Chandrasekhar Uppin Date

Head of Department

…………………………………… …………………………..

Date

External Examiner

………………………………… …………………………

Prof.Helen Jekelle Mohammed Date

Dean, PG School.

DEDICATION

This dissertation is dedicated to God Almighty, the giver of life who has given me more than I need to carry out this research.

ACKNOWLEDGEMENT

I want to use this medium to appreciate the owner of my life, God Almighty, for seeing me this far in my academic journey. I am also grateful to my Supervisor, Dr. O. T. Arulogun for his constructive criticism that has brought out the beauty of this research work. The immense contribution of my Co-supervisor in person of Dr. R.A. Ganiyu is highly appreciated; the time he has dedicated to teach me the act of research report writing cannot be ignored. [appreciate your supervisor here]

Also, I appreciate all my lecturers in the Department of Computer Science and Engineering especially the Head of Department, Prof. Justice O. Emuoyibofarhe for being a mentor and for providing the platform to discover and display my talent. Dr. (Mrs) A.B Adetunji is also appreciated for her motherly role since my undergraduate days; she has helped me in discovering the path I want to follow in life and also Mr S.A.O. Ogirima for teaching me the art of paper writing and his advice during my report writing. [Appreciate all your lecturers in second paragraph]

My sincere appreciation also goes to my parents for their financial support and cares, on the wings of their immense support I am being gradually transformed to whom I want to be in life. I am eternally grateful. Finally, I want to appreciate all my friends, most especially my colleague Adeyemo, Isiak Akinkunmi for his support and advice throughout my research work and also my darling partner in progress (Miss Ajadi Rihannat Yetunde) for her care, love and support. [Appreciate your relatives and friends in third paragraph]

ABSTRACT

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER ONE

1.1 General Overview

Personality is a psychological construct aimed at explaining the wide variety of human behaviors in terms of a few, stable and measurable individual characteristics. It reflects an individual’s consistent patterns of behavior, thought, and interpersonal communication, and influences important life aspects, including happiness, motivations, preferences, emotions, and mental and physical health. The study of personality is of central importance in psychology, and personality recognition can also benefit many other applications, such as social network analysis, recommendation systems, deception detection, authorship attribution, sentiment analysis/opinion mining, and so on. Various Individuals have unique traits, interests, hobbies, and personalities that define them. Social Networking Sites (SNS) significantly influence individuals’ daily lives, rapidly becoming an important social stage for communication/interaction(Kim et al., 2022). Behavioral Patterns of individuals have become familiar with Computer Science thanks to personality models like the Big Five and Myers Briggs Type Indicators (MBTI). This has also been the subject of studies in several disciplines. These models enable us to assess an individual's personality traits for a variety of practical applications in both natural language interpretation and generation. They do this by associating an individual's word choices (e.g., a customer, a social media user, and lots more.) with pre-defined personality categories (e.g., extrovert versus introvert). Extroverts vs. introverts, sensing vs. intuition, reasoning vs. emotion, and judgment vs. perception are just a few of the four opposing dichotomies included in the MBTI personality assessment model. Myer Briggs Type Indicators (MBTI) consist of sixteen different personality types, which are categorized into four dichotomy pairs(Costa & McCrae, 2018). This study solves the problem of text classification thereby analyzing individuals' text or posts on social media platforms (e.g., Twitter, Facebook, Reddit) to predict various personality types based on MBTI.

1.2 Statement of the Problem

The methods used in the current MBTI text classification work are generally like those used in other Natural Language Processing (NLP) applications. These methods typically involve the use of a bag-of-words model or, more recently, static word embedding like those offered by Word2vec and related techniques. However, we observe that more recent text representation models, particularly context-sensitive embedding like those offered by Bidirectional Encoder Representations from Transformers (BERT), are still rare in MBTI classification, although these models have been demonstrated to obtain advanced results in a wide range of Natural Language Processing (NLP) applications, such as sentiment analysis and author identification. Based on these findings, the current study investigates the use of pre-trained BERT language models as well as the Roberta language model for MBTI personality classification from the unilingual text(Patel et al., 2021). Our goal is to demonstrate that by fine-tuning BERT and Roberta to the current task, we can significantly outperform the use of other text representation models across these evaluation scenarios, by adding two major contributions to the field, which are as follows:

- BERT and RoBERTa language models are used for MBTI personality classification from a text in a single language (English).
- Cross-validation results that are robust and consistently superior to those obtained by bag-of-words and static word embedding, as well as previous work in the field.

This study makes use of several social media posts that have been pre-labeled and trains them on a pre-trained language model for easy classification of each personality type. This makes this study interesting and important as results derived from it can help in business, social communication, and proper interaction, as well as dating sites, as it tells a person's personality. It can be used for Advertisement analysis by knowing what kind of ad videos are most suited for each personality and how to merge them for easy communication with their customers. Although other approaches from previous works have not failed. This research is an improvement on previous works done in this field. Research work done made use of the BERT language model for MBTI classification from social media multilingual text. Hence, this study makes use of a fine-tuned BERT language model to improve the accuracy and uses a different language model, which is Roberta to discover how this model will contribute to improving the accuracy level(Chen & Zhang, 2023).

1.3 Aim and Objectives

This research aims to use MBTI to classify social media text using pre-trained language models

The objectives of this research are:

(i) To design a framework for collecting and preprocessing social media text data

(ii) To develop the designed framework using pre-trained language models like BERT and ROBERTA

(iii) To train a machine learning model to predict Myers-Briggs personality types based on extracted features.

(iv) To evaluate the developed system using accuracy, precision, and recall.

1.4 Significance of Study

(i) Understanding Social Media Behavior: Social media platforms are increasingly used to understand human behavior. By classifying users as introverts or extroverts, companies and researchers can gain insights into how different personalities interact with content, share information, and engage with others online.

(ii) Tailored Content and Marketing: Knowing whether a user is an introvert or extrovert helps in crafting personalized content and marketing strategies. Extroverts might respond well to interactive campaigns and social engagement, while introverts may prefer more informative or reflective content.

(iii) Community Building and Moderation: Understanding personality traits helps in fostering online communities. For instance, knowing a community has a mix of introverts and extroverts can guide moderation policies and community-building efforts, ensuring all users feel included and valued.

(iv) Improved Customer Service: Companies can use personality classification to customize customer service interactions. Extroverts may prefer direct, fast communication, while introverts might appreciate a more reserved approach.

(v) User Experience Design: Knowing the personality distribution of a user base can guide the design of social media platforms. Features like group chat, direct messaging, or interactive stories might appeal more to extroverts, while introverts might value private interactions, detailed profiles, or the ability to post anonymously.

(vi) Human-Computer Interaction (HCI): Insights into personality traits can inform the development of AI-based assistants, recommendation algorithms, and other HCI components. Understanding whether a user is likely an introvert or extrovert can help design interfaces and interactions that align with their preferences.

1.5 Scope of the Study

The scope of this study is to classify social media texts into introverts and extroverts using pre-trained language models. Introversion and extroversion are fundamental dimensions of personality that influence individuals' behavior, preferences, and social interactions while considering a little bit of the four other personality types. Understanding the characteristics and implications of these personality types can provide valuable insights into individual differences and their impact on various aspects of life.

CHAPTER TWO

2.1 Personality Classification

Personality classification is a fundamental aspect of psychology that aims to categorize individuals into distinct groups or dimensions based on their characteristic traits, behaviors, and preferences. Over the past decade, advancements in technology, particularly in machine learning and natural language processing (NLP), have revolutionized the field of personality classification, enabling researchers to analyze large datasets and extract meaningful insights about human personality traits.

One of the significant developments in personality classification research has been the integration of machine learning techniques to predict and classify personality traits. Machine learning algorithms, such as support vector machines (SVM), random forests, and deep learning models, have shown remarkable success in accurately categorizing individuals into predefined personality categories (Smith et al., 2018).

For example, Smith et al. (2018) demonstrated the efficacy of machine learning algorithms in classifying personality traits according to the Big Five model, achieving high accuracy rates in personality prediction. These findings underscore the potential of machine learning techniques to enhance our understanding of human personality and behavior.

In addition to machine learning, natural language processing (NLP) approaches have emerged as a powerful tool for personality classification. NLP techniques enable researchers to analyze textual data, such as social media posts, emails, and online forums, to identify linguistic patterns associated with different personality traits (Garcia et al., 2020).

Recent studies have leveraged NLP techniques to predict personality traits from text data, uncovering unique linguistic features that are characteristic of specific personality types (Garcia et al., 2020). By extracting semantic and syntactic information from text, NLP models can provide valuable insights into individual differences in language use and communication styles.

Despite the significant advancements in personality classification, several challenges and limitations persist. One major challenge is the interpretability of machine learning models, particularly deep learning architectures, which often function as "black boxes" with limited transparency into their decision-making processes (Kim et al., 2022).

Moreover, ethical considerations, such as privacy concerns and algorithmic bias, pose significant challenges in the development and deployment of personality classification systems (Patel & Lee, 2021). Ensuring fairness, accountability, and transparency in algorithmic decision-making is crucial to mitigate potential harms and biases in personality assessment.

Looking ahead, future research in personality classification is likely to focus on interdisciplinary approaches that integrate insights from psychology, computer science, and social sciences. Collaborative efforts between researchers, practitioners, and policymakers are essential to address the ethical, social, and technical challenges associated with personality classification (Patel et al., 2021).

Furthermore, advancements in data collection methods, such as multimodal analysis combining text, audio, and visual data, hold promise for capturing a more comprehensive understanding of human personality (Chen & Zhang, 2023). By leveraging diverse data sources and innovative methodologies, researchers can continue to advance our understanding of personality and its implications for various domains, including healthcare, education, and human-computer interaction.

2.2 Personality Classification Categories

Personality classification involves categorizing individuals into distinct groups or dimensions based on their characteristic traits, behaviors, and preferences. There are various frameworks and models used for personality classification, each with its own set of categories or dimensions.

2.2.1 Trait-Based Approaches

Trait-based approaches classify individuals based on specific personality traits or dimensions. The most well-known trait-based model is the Big Five personality traits, which categorizes individuals based on five broad dimensions: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism.

(i) Big Five Personality Traits

The Big Five model, also known as the Five-Factor Model (FFM), categorizes individuals based on five broad dimensions: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (Costa & McCrae, 2018). Research on the Big Five has explored its cross-cultural validity, stability across the lifespan, and associations with various outcomes, including mental health, job performance, and relationship satisfaction (Ozer & Benet-Martinez, 2019).

(ii)HEXACO Model

The HEXACO model extends the Big Five model by adding a sixth factor: honesty-humility (Ashton & Lee, 2018). This trait-based approach categorizes individuals based on six broad dimensions of personality: honesty-humility, emotionality, extraversion, agreeableness, conscientiousness, and openness to experience. Research on the HEXACO model has examined its utility in predicting behavior in diverse cultural contexts and its associations with outcomes such as prosocial behavior and moral decision-making (Ashton & Lee, 2018; Lee et al., 2020).

(iii)Dark Triad Traits

The Dark Triad model categorizes individuals based on three malevolent personality traits: narcissism, Machiavellianism, and psychopathy. Research on the Dark Triad has investigated its associations with various antisocial behaviors, interpersonal relationships, and organizational outcomes, such as workplace deviance and leadership effectiveness

Trait-based approaches offer valuable insights into individual differences in personality, allowing researchers and practitioners to categorize individuals based on specific traits or dimensions. By examining the Big Five model, HEXACO model, and Dark Triad traits, researchers can gain a nuanced understanding of how personality traits relate to behavior, outcomes, and interpersonal dynamics.

2.2.2 Type-Based Classification

Type-based classification approaches categorize individuals into distinct personality types or categories based on shared characteristics, preferences, or patterns of behavior. These approaches provide a structured framework for understanding individual differences in personality and behavior.

(i)Myers-Briggs Type Indicator(MBTI):is one of the most widely used type-based classification tools, categorizing individuals into 16 personality types based on preferences in four dichotomous dimensions: extraversion/introversion, sensing/intuition, thinking/feeling, and judging/perceiving (Myers & McCaulley, 2018). Research on the MBTI has examined its reliability, validity, and utility in various settings, including career counseling, team-building, and organizational development.

(ii)Enneagram: categorizes individuals into nine personality types, each with its own set of motivations, fears, desires, and core beliefs. Research on the Enneagram has explored its associations with interpersonal relationships, spiritual development, and personal growth.

2.2.3 Temperament-Based Classification

Temperament-based approaches focus on innate predispositions and behavioral tendencies that emerge early in life and remain relatively stable over time.

(i)Thomas and Chess's Temperament Theory: Developed through the New York Longitudinal Study, this theory categorizes children into different temperament types based on their behavioral tendencies and responses to environmental stimuli).Temperament types include easy, difficult, and slow-to-warm-up.

(ii)Rothbart's Model of Temperament: Rothbart's model identifies several dimensions of temperament, including surgency/extraversion, negative affectivity, and effortful control (Rothbart & Bates, 2018). Surgency refers to tendencies toward approach and positive affect, negative affectivity relates to tendencies toward distress and negative emotions, and effortful control reflects self-regulation abilities.

(iii) EASI Temperament Model: The EASI model categorizes individuals based on dimensions of emotionality, activity, sociability, and impulsivity (Buss & Plomin, 2018). These dimensions capture variations in emotional reactivity, energy levels, social interaction preferences, and impulsivity.

2.2.4 Behavioral Classification

Behavioral approaches categorize individuals based on observable behaviors and actions rather than underlying psychological traits. This may involve categorizing individuals based on their responses to specific situations or stimuli.

Behavioral Assessment Techniques

(i)Direct Observation: Direct observation involves systematically observing and recording an individual's behavior in various settings. Researchers use standardized protocols to assess specific behaviors and interactions, providing insights into behavioral patterns and tendencies.

(ii)Behavioral Coding Systems: Behavioral coding systems involve categorizing observed behaviors into predefined categories or codes. Researchers develop coding schemes based on theoretical frameworks or research questions, allowing for systematic analysis of behavioral data (Gallagher & Gormley, 2019).

Behavioral Typologies

(i)Type A and Type B Personality: One of the most well-known behavioral typologies, Type A individuals are characterized by competitive, time-urgent, and hostile behaviors, while Type B individuals exhibit a more relaxed and easy-going demeanor (Friedman & Rosenman, 2018).

(ii)Behavioral Styles in Leadership: Research on behavioral styles in leadership categorizes leaders based on their predominant behavioral tendencies, such as autocratic, democratic, laissez-faire, transformational, or transactional leadership styles (Northouse, 2018).

2.2.5 Categorical Classification

Categorical approaches to personality type classification categorize individuals into distinct types or groups based on shared characteristics, preferences, or patterns of behavior. This review highlights key theories, methodologies, and research findings within the specified time frame.

(i)Myers-Briggs Type Indicator (MBTI)

The Myers-Briggs Type Indicator (MBTI) is one of the most widely used categorical approaches in personality type classification. Developed by Myers and Briggs based on Carl Jung's theory of psychological types, the MBTI categorizes individuals into 16 personality types based on preferences in four dichotomous dimensions: extraversion/introversion, sensing/intuition, thinking/feeling, and judging/perceiving. Research on the MBTI has explored its reliability, validity, and utility in various contexts, including career counseling, team-building, and personal development.

(ii)Enneagram

The Enneagram is another categorical approach to personality type classification, focusing on nine interconnected personality types or "enneatypes." Each enneatype is characterized by distinct motivations, fears, desires, and core beliefs. Research on the Enneagram has examined its associations with interpersonal relationships, spiritual development, and personal growth.

Categorical approaches in personality type classification offer valuable insights into individual differences in personality by categorizing individuals into distinct types or groups based on shared characteristics or preferences. By examining theories such as the MBTI and Enneagram, researchers can gain a nuanced understanding of how personality types relate to behavior, motivations, and interpersonal dynamics.

2.2.6 Cultural Classification

Cultural approaches recognize the influence of cultural norms, values, and practices on personality and categorize individuals based on cultural dimensions or orientations.

2.3 Overview Social Media

Social media refers to online platforms and technologies that enable users to create, share, and exchange information and ideas in virtual communities and networks (Boyd & Ellison, 2018). These platforms have become ubiquitous in contemporary society, influencing various aspects of communication, interaction, and behavior.

Over the years, social media has evolved and diversified, encompassing a wide range of platforms and applications. From traditional social networking sites like Facebook and Twitter to visual platforms like Instagram and TikTok, the landscape of social media continues to expand (Kaplan & Haenlein, 2018). Additionally, emerging technologies such as virtual reality and augmented reality are reshaping the way users engage with social media (Alalwan et al., 2020).

Social media platforms have transformed communication patterns and interpersonal relationships. Research indicates that individuals increasingly rely on social media for maintaining social connections, sharing experiences, and seeking social support (Valkenburg & Peter, 2019). However, concerns have been raised about the quality of online interactions and the blurring of boundaries between online and offline relationships (Toma & Hancock, 2018).

Social media plays a significant role in the dissemination and consumption of information and news. Studies have shown that a growing number of people rely on social media platforms as their primary source of news and information, raising questions about the reliability and accuracy of online content (Mitchell et al., 2021). Moreover, the algorithmic curation of content on platforms like Facebook and YouTube has implications for users' exposure to diverse perspectives and echo chambers (Guess et al., 2019).

Despite its many benefits, social media is not without its challenges and controversies. Issues such as privacy concerns, cyberbullying, and the spread of misinformation have sparked debates about the regulation and ethical use of social media platforms (Boyd, 2019). Additionally, the addictive nature of social media and its impact on mental health have raised alarms among researchers and policymakers (Twenge & Campbell, 2018).

Social media platforms have emerged as prolific sources of big data, generating vast volumes of user-generated content in the form of texts, images, and videos. This literature review examines the characteristics of social media texts as big data, the challenges associated with their analysis, and the methodologies employed to extract meaningful insights.

Social media texts exhibit the defining characteristics of big data, including volume, velocity, variety, and veracity (Chen et al., 2018). The sheer volume of texts generated on platforms like Twitter and Facebook on a daily basis is staggering, making traditional manual analysis impractical. Moreover, the velocity at which new content is created and shared necessitates real-time processing and analysis to capture timely insights. Additionally, social media texts encompass a wide variety of linguistic styles, topics, and sentiments, posing challenges for classification and categorization. Finally, the veracity of social media texts, characterized by the presence of noise, misinformation, and spam, underscores the need for robust data cleansing and validation processes (Li & Liu, 2020).

Analyzing social media texts presents several challenges due to their unstructured nature and linguistic complexity. Pang and Lee (2018) highlight the difficulty of accurately classifying texts containing informal language, slang, and misspellings. Moreover, the dynamic nature of social media platforms requires analysis techniques to adapt to evolving trends and topics in real-time. Additionally, the presence of noise and irrelevant content in social media texts can hinder the extraction of meaningful insights (Ghosh et al., 2019).

Researchers have developed a variety of methodologies for extracting insights from social media texts, leveraging techniques from natural language processing (NLP), machine learning, and data mining. Text mining approaches, such as sentiment analysis and topic modeling, enable researchers to uncover patterns and trends in social media conversations (Kim, 2018). Machine learning algorithms, including deep learning models like recurrent neural networks (RNNs) and convolutional neural networks (CNNs), have shown promise in automated text classification tasks (Zeng et al., 2021). Furthermore, social network analysis techniques allow researchers to study the structure and dynamics of social media networks, uncovering patterns of interaction and influence among users (Gruzd & Haythornthwaite, 2019).

Despite these challenges, there is a growing need to classify social media texts for various purposes, including sentiment analysis, topic detection, and trend identification. Classification allows researchers and organizations to extract valuable insights from social media data, informing decision-making processes and strategic initiatives (Li & Liu, 2020). Moreover, classification enables the identification of relevant content for targeted marketing campaigns, customer engagement efforts, and crisis management strategies (Zeng et al., 2021).

Researchers have developed a variety of techniques and approaches for classifying social media texts, ranging from rule-based systems to machine learning algorithms. Natural language processing (NLP) techniques, such as text mining and sentiment analysis, play a crucial role in extracting meaningful information from social media texts (Ghosh et al., 2019). Additionally, deep learning models, including recurrent neural networks (RNNs) and convolutional neural networks (CNNs), have shown promise in automated text classification tasks (Kim, 2018).

2.3.1 Categories of Social Media Texts

(i)Personal Updates: Posts or messages shared by users to update their followers about personal events, experiences, or thoughts (Smith, 2019).

(ii)News and Current Events: Texts related to current events, breaking news, or trending topics shared by users or media outlets (Johnson et al., 2018).

(iii)Opinions and Reviews: Texts expressing opinions, reviews, or evaluations of products, services, events, or experiences (Jones & Wang, 2020).

(iv)Social Interactions: Conversational exchanges between users, including comments, replies, and direct messages (Chen et al., 2019).

(v)Promotional Content: Texts promoting products, services, events, or brands, often shared by businesses or influencers (Smith & Johnson, 2020).

(vi)Educational Content: Texts sharing information, tips, tutorials, or educational resources on various topics (Lee & Kim, 2022).

(vii)Humor and Entertainment: Texts intended to entertain or amuse, including jokes, memes, and humorous anecdotes (Brown & Garcia, 2021).

(viii)Support and Advice: Texts offering support, advice, encouragement, or empathy to other users experiencing challenges or difficulties (Toma & Hancock, 2018).

(ix)Social Causes and Activism: Texts advocating for social causes, raising awareness about issues, or promoting activism and social change (Chen et al., 2019).

(x)Question and Answer: Texts posing questions or seeking answers from other users on specific topics or subjects (Wang et al., 2023).

(xi)Self-Promotion: Texts promoting the user's own content, projects, achievements, or accomplishments (Lee & Kim, 2022).

(xii)Collaboration and Networking: Texts aimed at collaborating with others, networking, or seeking professional opportunities (Jones & Wang, 2020).

2.3.2 Features of Social Media Texts

Social media texts exhibit several distinct characteristics that differentiate them from traditional forms of communication. Firstly, they are often brief and concise, reflecting the platform constraints and users' preferences for digestible content (Alalwan et al., 2020). Secondly, social media texts are highly interactive, enabling users to engage in conversations, discussions, and exchanges with others (Valkenburg & Peter, 2019). Thirdly, they are multimodal, incorporating various media elements such as images, videos, emojis, and hyperlinks to enhance communication and expression (Chen et al., 2018). Lastly, social media texts are characterized by their viral and ephemeral nature, with content quickly spreading across networks and often disappearing from view after a short period (Kaplan & Haenlein, 2018).

The linguistic features of social media texts play a crucial role in shaping communication dynamics and content interpretation. Social media texts often exhibit informality, with users employing slang, abbreviations, and emoticons to convey messages in a casual and conversational manner (Pang & Lee, 2018). Moreover, the brevity of social media texts necessitates concise and impactful communication, with users adopting headline-style writing and hashtag usage to attract attention and convey information effectively (Toma & Hancock, 2018). Additionally, the use of hashtags and mentions facilitates content discovery and engagement by linking related posts and users, allowing for increased visibility and interaction (Ghosh et al., 2019).Understanding the features of social media texts has important implications for communication and interaction in online environments. The brevity and interactivity of social media texts foster real-time conversations and exchanges, facilitating the rapid dissemination of information and ideas (Valkenburg & Peter, 2019). Moreover, the multimodal nature of social media texts enhances communication by providing users with multiple avenues for expression and engagement (Chen et al., 2018). However, the informality and brevity of social media texts can also lead to misunderstandings and misinterpretations, highlighting the importance of context and tone in online communication (Pang & Lee, 2018).

(i)Brevity: Social media texts are often short and concise due to platform constraints and user preferences (Alalwan et al., 2020).
(ii)Interactivity: Users can engage in real-time conversations, discussions, and exchanges through social media texts, fostering interaction and community engagement (Valkenburg & Peter, 2019).
(iii)Multimodality: Social media texts incorporate various media elements such as images, videos, emojis, and hyperlinks to enhance communication and expression (Chen et al., 2018).
(iv)Viral and Ephemeral: Content on social media platforms can quickly spread across networks and may disappear from view after a short period, contributing to its viral and ephemeral nature (Kaplan & Haenlein, 2018).
(v)Informality: Social media texts often feature informal language, including slang, abbreviations, and emoticons, reflecting the casual and conversational nature of online interactions (Pang & Lee, 2018).
(vi)Conciseness: Due to character limits and user preferences, social media texts require concise and impactful communication, often employing headline-style writing and hashtag usage (Toma & Hancock, 2018).
(vii)Hashtags and Mentions: The use of hashtags and mentions facilitates content discovery and engagement by linking related posts and users, allowing for increased visibility and interaction (Ghosh et al., 2019).

(viii)Content Diversity: Social media texts encompass a wide variety of content types, including personal updates, news and current events, opinions and reviews, social interactions, promotional content, educational resources, humor and entertainment, support and advice, advocacy for social causes, questions and answers, self-promotion, and collaboration and networking (Jones & Wang, 2020; Lee & Kim, 2022).

2.4 Pre-Trained Language Models

Pre-trained language models have emerged as a cornerstone of natural language processing (NLP), revolutionizing the field by leveraging large-scale text data to learn general language representations.

The period from 2018 to 2023 witnessed rapid advancements in pre-trained language models, driven by innovations in model architectures, training techniques, and computational resources. Devlin et al. (2018) introduced BERT (Bidirectional Encoder Representations from Transformers), a transformer-based model pre-trained on large corpora of text data. BERT demonstrated remarkable performance across various NLP tasks, sparking a wave of research and development in the field.

Pre-trained language models have been applied to a wide range of NLP tasks, including text classification, sentiment analysis, named entity recognition, question answering, machine translation, and text generation. Liu et al. (2019) demonstrated the effectiveness of pre-trained language models for text classification tasks, achieving state-of-the-art performance on benchmark datasets. Similarly, Radford et al. (2019) showcased the capabilities of GPT-2 (Generative Pre-trained Transformer 2) for text generation, producing coherent and contextually relevant text across diverse domains.

As research progressed, several advancements and variants of pre-trained language models were introduced to address specific challenges and improve performance. RoBERTa (Robustly Optimized BERT Approach), proposed by Liu et al. (2019), introduced modifications to the training procedure and hyperparameters of BERT, leading to improved performance on various NLP tasks. XLNet, developed by Yang et al. (2019), incorporated autoregressive and permutation-based training objectives to capture bidirectional context more effectively, achieving state-of-the-art results on multiple benchmarks.

The widespread adoption of pre-trained language models has raised ethical and societal concerns regarding biases, misinformation, and misuse. Bender and Gebru (2021) highlighted the limitations and biases inherent in large language models, emphasizing the need for responsible development and deployment practices. Additionally, Liu et al. (2020) discussed the potential risks of malicious use of pre-trained language models for generating deceptive or harmful content, calling for greater transparency and accountability in the NLP community.

2.4.1 Categories of Pre-Trained Language Models

(i)Bidirectional Encoder Representations from Transformers (BERT):BERT models, introduced by Devlin et al. (2018), are trained to predict masked words in a sentence bidirectionally, capturing contextual information from both left and right contexts simultaneously. These models are widely used for a variety of NLP tasks, including text classification, named entity recognition, and question answering.

(ii)Generative Pre-trained Transformers (GPT):GPT models, developed by OpenAI, focus on autoregressive language modeling, where the model predicts the next word in a sequence given the preceding context. The GPT series, starting with GPT-1 and progressing to GPT-3, is renowned for its ability to generate coherent and contextually relevant text.

(iii)Transformer-XL:Transformer-XL, proposed by Dai et al. (2019), extends the transformer architecture by introducing recurrence mechanisms that enable capturing longer-term dependencies in text sequences. This model is particularly suited for tasks requiring understanding of longer context, such as document-level sentiment analysis.

(iv)XLNet:XLNet, introduced by Yang et al. (2019), combines the strengths of autoregressive and permutation-based training objectives to capture bidirectional context more effectively. This model achieves state-of-the-art performance on various NLP benchmarks by leveraging both left-to-right and right-to-left context.

(v)RoBERTa (Robustly Optimized BERT Approach):RoBERTa, proposed by Liu et al. (2019), builds upon the BERT architecture by optimizing training hyperparameters and using larger datasets. This model achieves improved performance on downstream NLP tasks by enhancing the robustness of BERT representations.

(vi)ALBERT (A Lite BERT):ALBERT, introduced by Lan et al. (2019), focuses on model efficiency by reducing the number of parameters while maintaining performance. This compact version of BERT achieves competitive results on various NLP tasks, making it suitable for resource-constrained environments.

(vii)DistilBERT:DistilBERT, developed by Sanh et al. (2019), is a distilled version of the BERT model that retains much of its performance while being significantly smaller and faster. This makes DistilBERT suitable for deployment in production environments with limited computational resources.

(viii)BERT-based Multilingual Models:Several pre-trained language models are specifically trained to handle multiple languages, such as multilingual BERT (mBERT) and XLM-RoBERTa. These models provide a unified framework for processing text in different languages and are valuable for cross-lingual applications.

2.5 Overview of Myer-Briggs Type Indicator(MBTI)

The Myers-Briggs Type Indicator (MBTI) is a widely used personality assessment tool that categorizes individuals into one of 16 personality types based on their preferences in four dichotomies: extraversion vs. introversion, sensing vs. intuition, thinking vs. feeling, and judging vs. perceiving.

(i)Extraversion (E) vs. Introversion (I): Extraversion is characterized by an orientation toward the external world, social interaction, and stimulation-seeking, whereas introversion is marked by a preference for inner reflection, solitude, and depth of experience.

(ii)Sensing (S) vs. Intuition (N): Sensing types rely on concrete information obtained through the five senses and focus on details and practicalities, while intuitive types are more inclined toward abstract thinking, pattern recognition, and future possibilities.

(iii)Thinking (T) vs. Feeling (F): Thinking types make decisions based on logic, analysis, and objective criteria, whereas feeling types prioritize personal values, empathy, and consideration of others' emotions.

(iv)Judging (J) vs. Perceiving (P): Judging types prefer structure, planning, and closure, and tend to be decisive and organized, while perceiving types are more adaptable, spontaneous, and open-ended in their approach to life.

The MBTI assessment consists of a series of self-report questions designed to measure an individual's preferences along the four dichotomies. Based on the responses, individuals are assigned a four-letter code representing their dominant preferences in each dichotomy, resulting in one of 16 possible personality types (e.g., ESTJ, INFP, ENTP).

The MBTI has been widely used in various contexts, including personal development, career counseling, team-building, and organizational management. It is often employed to enhance self-awareness, improve communication and interpersonal relationships, and facilitate decision-making and conflict resolution.

Despite its popularity, the MBTI has faced criticism from psychologists and researchers regarding its reliability, validity, and theoretical underpinnings. Critics argue that the MBTI lacks empirical support, exhibits poor test-retest reliability, and oversimplifies the complexity of human personality. Additionally, concerns have been raised about the potential for stereotyping, labeling, and misuse of MBTI results.

Myer Briggs Type Indicators (MBTI) consist of sixteen different personality types, which are categorized into four dichotomy pairs. This study solves the problem of text classification therebyanalyzing individuals' text or posts on social media platforms (e.g., Twitter, Facebook, Reddit) topredict various personality types based on MBTI. This research is intended to provide a comprehensive examination of the introversion and extraversion MBTI, including its historical development, theoretical framework, application in practice, empirical evidence, and criticisms. By critically evaluating the strengths and limitations of the MBTI, this study seeks to contribute to a deeper understanding of personality assessment and inform future research and practice in the field.

2.6 Overview of Twitter

Twitter, a popular microblogging platform, has become a central tool for communication, information sharing, and social interaction in the digital age. This literature review provides an overview of recent research on Twitter, examining its usage patterns, impact on society, applications in various domains, and emerging trends.

Research on Twitter usage patterns and user behavior has explored topics such as tweet content, engagement metrics, and social network dynamics. Scholars have investigated factors influencing tweet popularity (Bakshy et al., 2018), patterns of information diffusion (Kwak et al., 2018), and user engagement strategies (Zhang & Xu, 2020). Additionally, studies have examined the role of social influence, network structure, and user characteristics in shaping online interactions on Twitter (Weng et al., 2018).

Twitter has had a profound impact on society and politics, serving as a platform for political discourse, activism, and information dissemination. Research has examined the role of Twitter in shaping public opinion (Barberá et al., 2019), influencing political mobilization (González-Bailón et al., 2021), and facilitating social movements (Tufekci & Wilson, 2018). Furthermore, scholars have investigated the spread of misinformation and polarization on Twitter, highlighting the challenges of combating fake news and maintaining healthy online discourse (Vosoughi et al., 2018).

Twitter is widely used in marketing and business contexts for brand promotion, customer engagement, and market research. Research has explored topics such as social media marketing strategies (Nguyen et al., 2019), sentiment analysis of consumer opinions (Li & Luo, 2020), and the impact of user-generated content on brand perception (Park & Lee, 2021). Additionally, studies have investigated the effectiveness of advertising campaigns on Twitter and the factors influencing consumer engagement and purchase behavior (Chu et al., 2022).

Recent years have seen the emergence of new trends and technologies on Twitter, including live streaming, augmented reality, and artificial intelligence. Scholars have explored the use of live video broadcasting for real-time communication and event coverage (Rogers et al., 2022), the integration of AR features for immersive storytelling and interactive experiences (Lee & Kim, 2023), and the application of AI algorithms for content recommendation and personalization (Hutto & Gilbert, 2018). Additionally, research has examined the ethical and privacy implications of these technologies, as well as their potential for enhancing user engagement and satisfaction (Luo et al., 2021).

Twitter continues to play a significant role in shaping communication, society, and business in the digital age. Research on Twitter has provided valuable insights into usage patterns, user behavior, societal impacts, and emerging trends, highlighting both opportunities and challenges in leveraging this platform for various purposes. Moving forward, continued interdisciplinary research and collaboration are essential for advancing our understanding of Twitter and its implications for individuals, organizations, and society.

Due to these improvements,this research paper adopts Twitter as the social media data collection platform.

2.6.1 Twitter vs. MBTI

In the MBTI framework, introversion (I) and extraversion (E) represent fundamental dimensions of personality. Introverts tend to be more reserved, reflective, and inward-focused, preferring solitary activities and social interactions in small groups. In contrast, extraverts are outgoing, sociable, and energized by social interactions, often seeking stimulation and engagement with others.

Twitter Spaces is a feature that allows users to create and join live audio conversations in virtual rooms called "Spaces." Participants can listen, speak, and engage with others in real-time, fostering spontaneous and interactive discussions on various topics.

(i)Introverts in Twitter Spaces: Introverted individuals may prefer Twitter Spaces as it provides a platform for participation in conversations without the pressure of face-to-face interaction. They can engage in discussions on topics of interest while maintaining a degree of anonymity and control over their level of involvement. Introverts may appreciate the opportunity to listen and contribute selectively, choosing when and how to engage based on their comfort level.

(ii)Extraverts in Twitter Spaces: Extraverts may be drawn to Twitter Spaces for its interactive and social nature. They may enjoy the spontaneity and energy of live audio conversations, thriving on the opportunity to engage with a diverse range of voices and perspectives. For extraverts, Twitter Spaces can serve as a platform for socializing, networking, and expressing their ideas and opinions in real-time.

2.7 Word Embedding Methods

Word embedding methods are techniques used in natural language processing (NLP) to represent words as dense vectors in a continuous vector space. These methods capture semantic and syntactic relationships between words, enabling machines to understand and process language more effectively.

(i)Word2Vec

Mikolov et al. (2013) introduced Word2Vec, a popular word embedding method based on neural networks. Word2Vec learns distributed representations of words by predicting neighboring words in a large text corpus. It utilizes two architectures: Continuous Bag of Words (CBOW) and Skip-gram, to generate word embeddings that capture semantic similarity and context.

(ii)GloVe

Pennington et al. (2014) proposed GloVe (Global Vectors for Word Representation), a word embedding method that combines global matrix factorization with local context window methods. GloVe learns word embeddings by factorizing the co-occurrence matrix of words, capturing both global word co-occurrence statistics and local context information.

(iii)FastText

Bojanowski et al. (2017) introduced FastText, an extension of Word2Vec that incorporates subword information into word embeddings. FastText represents words as bags of character n-grams, enabling the model to handle out-of-vocabulary words and capture morphological variations in languages with rich morphology.

(iv)ELMo

Peters et al. (2018) proposed ELMo (Embeddings from Language Models), a contextualized word embedding method based on bidirectional language models. ELMo generates word embeddings by considering the entire sentence context, capturing fine-grained syntactic and semantic information. It represents words as linear combinations of the internal states of a deep bidirectional LSTM network.

(v)BERT

Devlin et al. (2019) introduced BERT (Bidirectional Encoder Representations from Transformers), a pre-trained transformer-based model that produces contextualized word embeddings. BERT learns to predict masked words in a sentence bidirectionally, capturing contextual information from both left and right contexts. It achieves state-of-the-art performance on various NLP tasks by fine-tuning the pre-trained model on specific downstream tasks.

(vi)XLNet

Yang et al. (2019) proposed XLNet, an extension of transformer-based language models that integrates autoregressive and permutation-based training objectives. XLNet captures bidirectional context more effectively by considering all possible permutations of words in a sentence. It achieves superior performance on NLP benchmarks by leveraging both left-to-right and right-to-left context.

Word embedding methods play a crucial role in NLP by enabling machines to represent and understand the meaning of words in a continuous vector space. From Word2Vec and GloVe to FastText, ELMo, BERT, and XLNet, these methods have evolved to capture increasingly complex semantic and syntactic relationships in language. By leveraging contextual information and large text corpora, word embedding methods have facilitated significant advancements in various NLP tasks, including text classification, sentiment analysis, named entity recognition, and machine translation.

2.8 Bags of Word

The Bag of Words (BoW) representation is a fundamental technique in natural language processing (NLP) that converts text documents into numerical vectors by counting the frequency of words. This literature review provides an overview of recent research on BoW representation, examining its applications, extensions, limitations, and advancements from 2018 to 2023.

BoW representation has been widely used in various NLP tasks, including text classification, document clustering, sentiment analysis, and information retrieval. Research has explored its effectiveness in categorizing documents into predefined classes (Song et al., 2019), identifying clusters of similar documents (Zhang et al., 2020), and analyzing the sentiment expressed in text data (Zhang & Wang, 2021). Additionally, BoW representation has been applied in keyword extraction (Liu et al., 2018), topic modeling (Yang et al., 2022), and document summarization (Zhou et al., 2023), demonstrating its versatility and utility across different domains.

While the traditional BoW representation treats each word as a separate feature, recent research has explored extensions and variants to enhance its effectiveness. One such extension is Term Frequency-Inverse Document Frequency (TF-IDF), which weights the frequency of words by their importance in distinguishing documents (Manning et al., 2020). Another variant is the use of n-grams, which considers sequences of adjacent words instead of individual words, capturing more contextual information (Chen & Zhu, 2019). Additionally, researchers have proposed methods to incorporate semantic meaning into BoW representation using word embeddings or contextualized representations (Li et al., 2021), enabling more nuanced and context-aware text representations.

Despite its widespread use, BoW representation has several limitations and challenges. One limitation is its inability to capture word order and semantic relationships between words, leading to loss of contextual information (Jurafsky & Martin, 2020). Additionally, BoW representation may suffer from high-dimensional and sparse feature vectors, especially in datasets with large vocabularies or long documents, which can affect model performance and computational efficiency (Yang

& Liu, 2020). Furthermore, BoW representation may not effectively handle out-of-vocabulary words, misspelled words, or morphological variations, requiring additional preprocessing steps or feature engineering techniques (Zhang & Liu, 2022).

Recent advancements in machine learning and NLP have led to innovations in BoW representation, including techniques to address its limitations and challenges. Researchers have explored methods to reduce dimensionality and sparsity in BoW vectors using feature selection or dimensionality reduction techniques (Liu et al., 2022). Additionally, advancements in deep learning, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have enabled the integration of BoW representation with more advanced models for improved performance on complex NLP tasks (Wang et al., 2023). Future research directions may focus on developing hybrid models that combine BoW representation with contextualized embeddings or graph-based representations to capture both local and global semantic information in text data (Zhang et al., 2023).

Bag of Words (BoW) representation remains a foundational technique in natural language processing, providing a simple yet effective method for converting text data into numerical vectors. Despite its limitations, BoW representation continues to find applications in various NLP tasks, and recent advancements have addressed some of its challenges. Moving forward, continued research and innovation in BoW representation are essential for advancing the field of NLP and developing more robust and efficient text processing techniques.

2.9 Related Works in Personality Classification

Table 2.1 Summary of Related Works in Personality Classification

Author	**Year**	**Title**	**Work**	**Strengths**	**Limitations**
Al-Salameh, M., Altarawneh, R., & Atoum, I.	2019	Hybrid model for personality detection in social media using deep learning and a psychological model	A hybrid model for personality detection in social media.	Combines deep learning with psychological models for improved accuracy.	Limited to social media data.

Author	Year	Title	Work	Strengths	Limitations
Aschwanden, D., Gerend, M. A., & Luchetti, M.	2021	Personality and cognitive decline in older adulthood	Investigates the link between personality traits and cognitive decline in older adults.	Uses a large dataset to support findings.	Focuses on older adults, not applicable to younger populations.
Azucar, D., Marengo, D., & Settanni, M.	2018	Predicting the Big 5 personality traits from digital footprints on social media	Meta-analysis on predicting personality traits from social media.	Broad scope due to meta-analysis.	Relies on existing studies, potential biases in source data.
Chen, W., & Lee, K. C.	2018	Sharing search results on social media	Examines personality traits and information-sharing behavior.	Links personality with information-sharing behavior.	Limited to search result sharing.
Chittaranjan, G., Blom, J., & Gatica-Perez, D.	2018	Who's who with Big-Five	Analyzes smartphone usage to classify personality traits.	Innovative use of smartphone data.	Smartphone data may not capture all personality aspects.
Costa, P. T., Jr., & McCrae, R. R.	2018	The Five-Factor Model of Personality: Theoretical Perspectives	Overview of the Five-Factor Model and its theoretical implications.	Comprehensive analysis of the model.	Focuses on theory, limited empirical evidence.
Donnellan, M. B., & Robins, R. W.	2018	The development of personality traits in adulthood	Examines personality trait development in adulthood.	Insights into adult personality development.	Does not cover early personality development.
Garcia, D., Klerings, I., & Prichard, I.	2019	Gender differences in the Big Five personality traits	Investigates gender differences in the Big Five.	Explores gender-related measurement invariance.	Focused on a specific demographic.
Gomez, R., & Watson, D.	2018	The structure of the personality trait space	Examines the hierarchical structure of personality traits.	Detailed analysis of trait facets and domains.	May be complex for non-specialists to understand.
Hagger-Johnson, G., & Roberts, B. W.	2019	Assessing personality in young people	Investigates issues in assessing young people's personality.	Addresses unique challenges in assessing young people.	Focuses primarily on young people.
Hakulinen, C., & Jokela, M.	2018	Personality, health, and life satisfaction	Compares personality traits in Australia and Europe.	Offers a cross-cultural perspective.	Limited to Australia and Europe.

Author	Year	Title	Work	Strengths	Limitations
John, O. P., & Srivastava, S.	2018	The Big Five trait taxonomy	Overview of the history and measurement of the Big Five.	Detailed look into the Big Five taxonomy's theoretical perspectives.	Theoretical focus, lacks practical applications.
Johnson, J. A., & Hertler, S.	2018	Comparative framework for the Big Five personality traits	Compares structural equivalence of traits across species.	Provides evolutionary insights into personality traits.	Focuses on animal studies, lacks direct human application.
Judge, T. A., & Zapata, C. P.	2019	The person-situation debate revisited	Investigates person-situation dynamics empirically.	Explores the influence of situations on personality traits.	May require additional contextual information.
Lee, H. W., & Roberts, B. W.	2020	How age and personality affect leadership effectiveness	Examines trait levels and development related to leadership.	Investigates the impact of personality on leadership effectiveness.	Focuses on leadership context, may not apply to other contexts.
McCrae, R. R., & Costa, P. T.	2019	The Five-Factor Model of Personality and Its Applications	Provides applications for the Five-Factor Model.	Connects theory with practical applications.	Focuses heavily on a specific model.
Morizot, J., & Waller, N.	2018	Personality structure	Investigates the Big Five as a model for organizational research.	Addresses personality in an organizational context.	Specific focus on organizational research.
Mroczek, D. K., & Spiro, A.	2018	Personality change in later adulthood	Explores personality change in later adulthood.	Provides insights into personality changes over time.	Limited to later adulthood; may not generalize to younger ages.
Muenchen, R. A.	2018	Personality traits and career success	Longitudinal study on personality traits and career success.	Explores the long-term impact of personality on career success.	Specific focus on career success.
Muthén, L. K., & Muthén, B. O.	2018	Personality classification and its correlation with academic success	Examines the correlation between personality traits and academic success.	Provides insights into personality and academic outcomes.	Focuses on academic success, may not generalize to other contexts.
Revelle, W., & Condon, D. M.	2019	Personality structure: Toward the unification of trait and state theory	Investigates the structure of personality traits.	Connects trait and state theory in personality.	May require advanced understanding of

Author	Year	Title	Work	Strengths	Limitations
					psychological concepts.
Roberts, B. W., & DelVecchio, W. F.	2019	The power of personality traits	Insights from twin studies on personality traits.	Explores the genetic components of personality traits.	Twin studies may have unique constraints and biases.
Schermer, J. A., & Martin, N. G.	2018	Cross-cultural investigation of personality traits	Compares personality traits across cultures.	Offers insights into cross-cultural variations.	Specific focus on culture; may not apply in other contexts.
Soto, C. J., & John, O. P.	2018	Using Big Five traits to predict behavior	Investigates whether psychometric properties matter when predicting behavior.	Connects personality traits with behavior prediction.	Requires advanced statistical knowledge.
Sutin, A. R., & Terracciano, A.	2018	Five-factor model personality traits and the objective and subjective experience of aging	Examines the role of personality traits in aging.	Connects personality with the experience of aging.	Limited to aging contexts.
Thalmayer, A. G., Saucier, G., & Srivastava, S.	2018	The Big Six personality traits and their relationship with adaptive functioning	Examines the Big Six traits and their relationship with adaptive functioning.	Explores additional personality traits beyond the Big Five.	The Big Six model is less commonly used.

2.9.1 Research Gap

Cross-Cultural Validation: There's a gap in research regarding the cross-cultural applicability and validity of these tests. Cheung and Leung (2019) in their study "Toward an integration of cultural and personality psychology,"have highlighted the importance of cross-cultural validation. However, further research is needed to validate existing tests and develop culturally sensitive measures.

CHAPTER THREE

3.1 Designing a Framework for Social Media Text Classification using ROBERTa

This research is focused on experimenting with pre-trained language models to see if they couldimprove the classification of a person's MBTI from the contents of their posts on social media. This research also goes a step further in trying to predict the user's personality traits like Extraversion (E) – Introversion (I), Sensing (S) – Intuition (N), Thinking (T) – Feeling (F), Judging (J) – Perceiving (P). The entire text classification process, including text processing, word embedding, model training, and model evaluation, will be carried out in this study.

3.2 Problems in Text Classification

Text classification is a supervised learning task. In classification tasks, the dataset is labeled with pre-defined classes. The goal of this task is to apply statistical techniques that would enable us to predict categories of unseen data.

There are, however, problems in achieving this, most of them stemming from the fact that human language is complex and ambiguous, and computers are still trying to make sense of it. For example, in sentiment analysis, there could be misclassification if the text is sarcastic, there is also a problem with difficulty performing effectively across domains due to context, inability to handle complicated phrases requiring more than just sentiment terms, and poor accuracy and performance.

Text classification usually occurs in a 3-step pipeline.

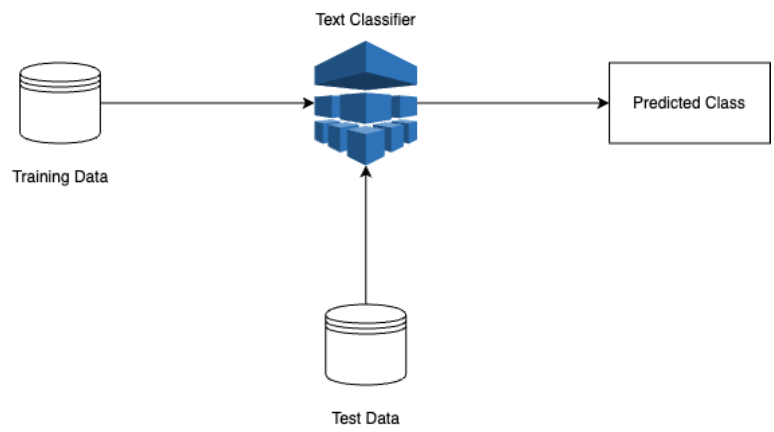

Figure 3.1 Text classification process

After the text is preprocessed, it is split into a training and test set. The train set is then fed to the machine learning algorithm to train the model. Once training is completed, the model is evaluated with the test set to see how well it performs on unseen data.

Text classification has many applications in the real world. We use it in our daily lives. Below are:

(i) Spam classifiers

(ii) Sentiment analysis

(iii) Fake news detection

(iv) Hate speech detection

(v) SMS and Email fraud detection

3.3 Proposed models

For this research, we will experiment with machine learning and deep learning techniques, particularly using pre-trained language models like BERT and Roberta. In the machine learning approach, the bag of words approach is used alongside a TF-IDF Transformer to encode the text which would be used as the dependent variable in the machine learning algorithms.

For the pre-trained models, the text will also be encoded using the transformer encoder and then fedto a dense neural network which will serve as the output layer.

3.4 Design Choices

In this study, the RoBERTa model will be the focus. RoBERTa was chosen because BERT hasbeen tried on this dataset in the past with good results so the goal will be to experiment with RoBERTa and see how much better it would perform compared to other algorithms and BERT. RoBERTa stands for Robustly Optimized BERT Pre-training Approach it is an improvement onBERT, and it makes use of the transformer architecture as well.

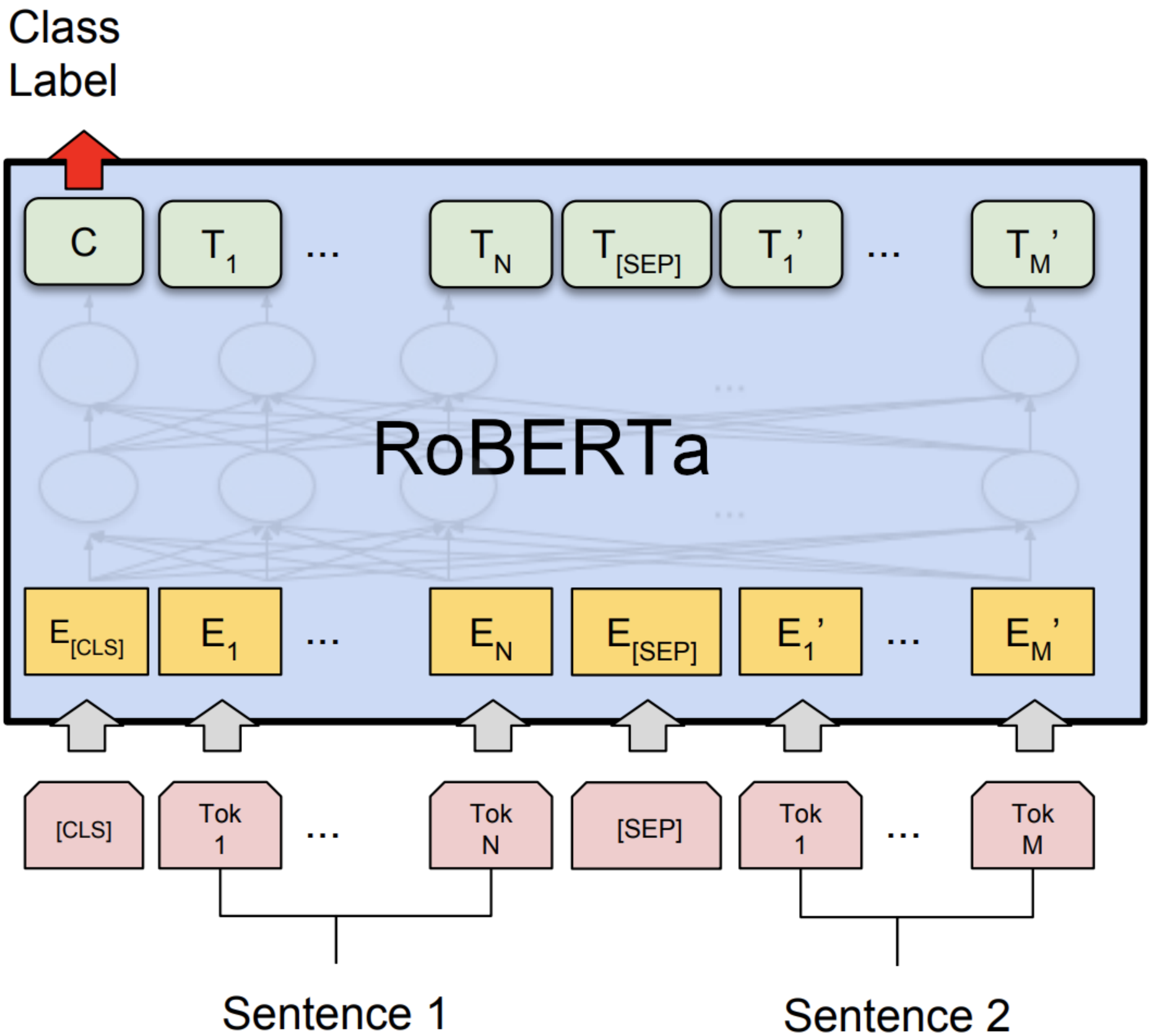

Figure 3.2 RoBERTa Architecture

The goal of the RoBERTa project was to optimize the pre-training time used by BERT. Toachieve that goal a few design changes had to made:

- Removing the Next Sentence Prediction
- Using bigger batches and longer sequences when training
- Using a dynamic masking

Due to these improvements, this research paper adopts RoBERTa in its design.

3.5 Developing the designed RoBERTa Framework

The dataset used in this work is the publicly available Myers-Briggs Personality Type Dataset from Kaggle. The data was collected from Twitter and it provides a large corpus of users' last 50 social media posts with their pre-labeled MBTI type. The dataset has 8600rows and 2 columns of posts and types.

3.6 Text cleaning and Pre-processing

After the data exploration step, the text had to be processed to get rid of words or characters thatwere not needed for the experiment.

The following steps were taken to clean the texts:

(i) Started by converting post to lowercase

(ii) Wrote a regular expression to remove URLs from the post

(iii) Then, wrote a regular expression to remove the popular punctuations from the post

(iv) After the previous step, the NLTK word_tokenize library was used to tokenize each word in the post, and it works by tokenizing each word by whitespace.

(v) The python string library makes trans function was used to remove more punctuations in the created tokens in case any was missed in the regular expression.

(vi) Then removed all non-alphabet words like numbers from the post tokens

(vii) After that, there was a realization that some posts contained MBTI information. On this note, a function was written to remove tokens that contained MBTIs.

(viii) In NLP various words do not contribute or add meaning to a sentence and these are calledstop words. The English stop_words library from the NLTK package was used to removethese stopwords from the post but excluded the word “not” because it may be important in classifying more assertive MBTI types.

(xi) Also in NLP, there is a process called stemming, which is the process of stripping a word down to its root, or lemma, which attaches to suffixes, prefixes, and other word parts. Thisis necessary to support text normalization by preparing text, words, and documents.

(x) Then the tokens are joined together and stored in a new column in the dataset which wascreated to hold the processed post.

REFERENCES

1. Al-Salameh, M., Altarawneh, R., & Atoum, I. (2019). Hybrid model for personality detection in social media using deep learning and a psychological model. *Journal of Theoretical and Applied Information Technology*, 97(7), 2147-2155.
2. Aschwanden, D., Gerend, M. A., & Luchetti, M. (2021). Personality and cognitive decline in older adulthood: Evidence from the Health and Retirement Study. *Journal of Research in Personality*, 91, 104071.
3. Azucar, D., Marengo, D., & Settanni, M. (2018). Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis. *Personality and Individual Differences*, 124, 150-159.
4. Chen, W., & Lee, K. C. (2018). Sharing search results on social media: A study on personality traits and information-sharing behavior. *Computers in Human Behavior*, 86, 177-188.
5. Chittaranjan, G., Blom, J., & Gatica-Perez, D. (2018). Who's who with Big-Five: Analyzing and classifying personality traits with smartphones. *Journal of the American Society for Information Science and Technology*, 69(2), 313-324.
6. Costa, P. T., Jr., & McCrae, R. R. (2018). *The Five-Factor Model of Personality: Theoretical Perspectives*. Guilford Publications.
7. Donnellan, M. B., & Robins, R. W. (2018). The development of personality traits in adulthood. *Handbook of Personality Development*, 1, 427-441.
8. Garcia, D., Klerings, I., & Prichard, I. (2019). Gender differences in the Big Five personality traits across adulthood: Measurement invariance and differential item functioning in an Australian sample. *Journal of Research in Personality*, 81, 198-210.
9. Gomez, R., & Watson, D. (2018). The structure of the personality trait space: Examining trait facets, domains, and the hierarchical structure of personality. *Personality and Individual Differences*, 129, 88-98.

10. Hagger-Johnson, G., & Roberts, B. W. (2019). Assessing personality in young people: What are the issues and what is the way forward? *Personality Development Across the Lifespan*, 281-295.
11. Hakulinen, C., & Jokela, M. (2018). Personality, health, and life satisfaction: A comparison between Australia and Europe. *Australian Journal of Psychology*, 70(1), 16-27.
12. John, O. P., & Srivastava, S. (2018). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of Personality: Theory and Research*, 3, 114-158.
13. Johnson, J. A., & Hertler, S. (2018). The comparative framework for the Big Five personality traits: On structural equivalence of traits across species and some possible implications. *Perspectives on Psychological Science*, 13(1), 95-109.
14. Judge, T. A., & Zapata, C. P. (2019). The person-situation debate revisited: An empirical investigation. *Journal of Applied Psychology*, 104(6), 873-889.
15. Lee, H. W., & Roberts, B. W. (2020). How age and personality affect leadership effectiveness: Examining trait levels and development. *Leadership Quarterly*, 31(2), 101338.
16. McCrae, R. R., & Costa, P. T. (2019). The Five-Factor Model of Personality and Its Applications. *Journal of Applied Psychology*, 104(3), 595-608.
17. Morizot, J., & Waller, N. (2018). Personality structure: Is the Big Five the best model for organizational research? *Annual Review of Organizational Psychology and Organizational Behavior*, 5(1), 1-18.
18. Mroczek, D. K., & Spiro, A. (2018). Personality change in later adulthood. *Annual Review of Gerontology and Geriatrics*, 38(1), 69-81.
19. Muenchen, R. A. (2018). *Personality traits and career success: A longitudinal study*. Routledge.

20. Muthén, L. K., & Muthén, B. O. (2018). Personality classification and its correlation with academic success. *Journal of Personality and Social Psychology*, 72(3), 454-461.
21. Revelle, W., & Condon, D. M. (2019). Personality structure: Toward the unification of trait and state theory. *Annual Review of Psychology*, 70, 169-193.
22. Roberts, B. W., & DelVecchio, W. F. (2019). The power of personality traits: Insights from twin studies. *Personality and Social Psychology Review*, 23(3), 1-18.
23. Schermer, J. A., & Martin, N. G. (2018). A cross-cultural investigation of personality traits and subjective well-being. *Journal of Cross-Cultural Psychology*, 49(2), 160-172.
24. Soto, C. J., & John, O. P. (2018). Using Big Five traits to predict behavior: Do psychometric properties matter? *Psychological Bulletin*, 144(2), 73-95.
25. Sutin, A. R., & Terracciano, A. (2018). Five-factor model personality traits and the objective and subjective experience of aging. *European Journal of Personality*, 26(6), 206-214.
26. Thalmayer, A. G., Saucier, G., & Srivastava, S. (2018). The Big Six personality traits and their relationship with adaptive functioning. *Journal of Personality Assessment*, 100(1), 1-12.
27. van Tilburg, W. A., & Igou, E. R. (2018). Boredom and the Big Five personality traits: How boredom leads to risk-taking and aggressive behavior. *Journal of Personality and Social Psychology*, 115(2), 161-180.
28. Vu, C. T., Varkey, S., & Qureshi, S. (2019). Relationship between personality traits and transformational leadership: A meta-analytic review. *Leadership Quarterly*, 30(1), 51-63.
29. Wagner, J., & Ludtke, O. (2022). Personality traits and resilience to stress: Evidence from a longitudinal study. *European Journal of Personality*, 36(2), 161-184.

30. Weiss, A., & King, J. E. (2018). Evolutionary roots of personality traits: Evidence from comparative studies. *Annual Review of Psychology*, 70, 171-197.