Received: 25 October 2023 | Revised: 25 January 2024 | Accepted: 17 February 2024
DOI: 10.1111/exsy.13572

ORIGINAL ARTICLE

Expert Systems WILEY

AI unveiled personalities: Profiling optimistic and pessimistic attitudes in Hindi dataset using transformer-based models

Dipika Jain[1] | **Akshi Kumar**[2]

[1]Department of Computer Science & Engineering, Delhi Technological University, New Delhi, India

[2]Department of Computing, Goldsmiths, University of London, London, UK

Correspondence
Akshi Kumar, Department of Computing, Goldsmiths, University of London, London, United Kingdom.
Email: akshi.kumar@gold.ac.uk

Abstract

Both optimism and pessimism are intricately intertwined with an individual's inherent personality traits and people of all personality types can exhibit a wide range of attitudes and behaviours, including levels of optimism and pessimism. This paper undertakes a comprehensive analysis of optimistic and pessimistic tendencies present within Hindi textual data, employing transformer-based models. The research represents a pioneering effort to define and establish an interaction between the personality and attitude chakras within the realm of human psychology. Introducing an innovative "Chakra" system to illustrate complex interrelationships within human psychology, this work aligns the Myers-Briggs Type Indicator (MBTI) personality traits with optimistic and pessimistic attitudes, enriching our understanding of emotional projection in text. The study employs meticulously fine-tuned transformer models—specifically mBERT, XLM-RoBERTa, IndicBERT, mDeBERTa and a novel stacked mDeBERTa—trained on the novel Hindi dataset 'मनोभाव' (pronounced as Manobhav). Remarkably, the proposed Stacked mDeBERTa model outperforms others, recording an accuracy of 0.7785 along with elevated precision, recall, and F1 score values. Notably, its ROC AUC score of 0.7226 underlines its robustness in distinguishing between positive and negative emotional attitudes. The comparative analysis highlights the superiority of the Stacked mDeBERTa model in effectively capturing emotional attitudes in Hindi text.

KEYWORDS
deep learning, machine learning, MBTI, optimism, pessimism, transformers

1 | INTRODUCTION

Numerous studies have investigated personality trait detection through social media posts and online content, often employing Big Five (Mehta et al., 2020; Ong et al., 2017; Yoneda et al., 2023) and Myers-Briggs Type Indicator (MBTI) (Cerkez et al., 2021; Jain et al., 2022; Kumar et al., 2023; Ren et al., 2021; Shafi et al., 2021) theories for psychometric testing. Psycholinguistics research demonstrates that writing style and language reveal personality, aided by semantic and syntactic cues. Recent advances in NLP, driven by deep learning and Transformer-based models like GPT-3 (Brown et al., 2020), BERT (Devlin et al., 2018; Kenton & Toutanova, 2019), and RoBERTa (Liu et al., 2019), have revolutionized various tasks by capturing contextual nuances and dependencies (Table 1). Utilizing such models in personality detection enhances nuanced analyses of language patterns, uncovering intricate human behaviour and thought processes. While efforts have been made to create datasets and benchmarks for low-resource NLP tasks such as sentiment (Kumar & Albuquerque, 2021) and sarcasm detection (Jain et al., 2020), more recent

TABLE 1 Table of taxonomy.

Abbreviation	Definition
NLP	Natural Language Processing
MBTI	Myers Briggs Type Indicator
I	Introvert
E	Extrovert
T	Thinking
F	Feeling
J	Judgement
P	Perceiving
N	iNtution
S	Sensing
BERT	Bidirectional Encoder Representations from Transformers
mBERT	Multilingual BERT
GPT	Generative Pre-trained Transformer
RoBERTa	Robustly Optimized BERT Pre-training Approach
ALBERTA	A lite BERT
XLM	Cross-lingual Language Model
DeBERTa	Decoding-enhanced BERT with Disentangled Attention
mDeBERTa	Multilingual DeBERTa

initiatives have arisen to address personality detection in non-English languages. This includes languages like Arabic (Salem et al., 2019), Persian (Fatehi et al., n.d.; Anari et al., 2022), Bahasa Indonesia (Adi et al., 2018), Hindi (Khan et al., 2020; Kumar et al., 2023), and Bengali (Rudra et al., 2020), which are often characterized by limited resources. This undertaking holds particular significance for Indian languages, which are predominantly low-resource and face constraints in training NLP systems, especially those tailored for conversations. Moreover, these datasets are designed to align typically with the MBTI and Big-5 frameworks.

MBTI provides a structured framework for understanding cognitive preferences and decision-making styles. It is designed to help individuals gain insight into their cognitive preferences and tendencies in how they perceive the world and make decisions. It categorizes people into 16 personality types based on four dichotomies: extraversion (E) versus introversion (I), sensing (S) versus intuition (N), thinking (T) versus feeling (F), and judging (J) versus perceiving (P). These preferences provide a general framework for understanding how individuals process information, interact with their environment, and approach decision-making. While the MBTI provides valuable insights into an individual's cognitive preferences, it does not comprehensively capture the entirety of a person's psychological makeup. This is primarily because of the following factors:

- *Complexity of Personality*: Human personality is multi-dimensional and intricate. It encompasses a wide range of traits, behaviours, emotions, and attitudes. While the MBTI offers a structured way to describe certain cognitive tendencies, it does not account for the full complexity of human personality, which can include traits like optimism, pessimism, resilience, emotional stability, and more.
- *Variability within types*: Each of the 16 MBTI types represents a set of preferences, but there is significant variability within each type. People of the same type can have diverse attitudes, values, and behaviours. For example, two individuals of the same MBTI type might express different levels of optimism or pessimism based on their personal experiences, upbringing, and life circumstances.
- *Emotional attitudes*: Optimism and pessimism are emotional attitudes that relate to how individuals view the future and interpret events. Specific MBTI preferences are associated with tendencies that align with these attitudes (e.g., intuitive types leaning toward optimism due to their focus on possibilities), However, it is important to note that the MBTI primarily assesses cognitive preferences and doesn't explicitly measure or categorize emotional attitudes like pessimism or optimism in the same way it assesses cognitive preferences.
- *External Influences*: Attitudes like optimism and pessimism can be influenced by a variety of external factors, including cultural background, life experiences, upbringing, education, and personal beliefs. These influences can shape an individual's emotional outlook in ways that may not be fully captured by the MBTI.

In this study, our objective is to target one of existing research gaps from above by developing a method for profiling individuals' emotional attitudes which contribute to shaping their underlying personalities. Optimism is characterized by a positive mindset that anticipates favourable outcomes and perceives challenges as opportunities for personal development. On the other hand, pessimism involves a negative mindset that expects unfavourable outcomes, leading individuals to dwell on negatives, anticipate failure, and approach challenges cautiously. The interplay between one's attitude toward optimism or pessimism frequently reveals compelling associations with specific MBTI personality types. This intriguing phenomenon can be further elucidated through various correlations. For instance, there exists a statistically positive link between an individual's optimism and their extraverted tendencies. This implies that those who lean toward optimism often exhibit extroverted traits, displaying an inclination toward sociability, enthusiasm, and a proactive approach to engaging with the world. Similarly, a parallel trend emerges in the connection between pessimism and emotional stability. Here, individuals who tend toward pessimism may demonstrate a greater emphasis on emotional stability. This suggests that they could exhibit a heightened awareness of their emotions and possess coping strategies that help them navigate challenges with resilience. On the contrary, the relationships are more complex when considering the interplay between pessimism and extraversion, as well as optimism and introversion. Statistically, negative associations are discernible. This implies that those who are pessimistic might exhibit tendencies that diverge from extraversion, possibly showcasing introverted traits such as introspection, depth of thought, and a preference for solitude. Conversely, individuals characterized by optimism may align less with introversion, potentially displaying qualities associated with extraversion.

These intricate interconnections bear profound implications within the realm of information technology, particularly in the domain of dialogue prediction and analysis. By harnessing an awareness of these correlations, technology can potentially elevate its capacity to forecast and decode individual behaviour based on their conversational interactions. As a result, this heightened understanding could empower technology to anticipate responses, decipher nuanced meanings, and adapt interactions, thereby facilitating more effective and personalized communication experiences.

As pessimism and optimism detection in text can enhance decision-making, facilitate proactive interventions, and provide valuable insights across a wide range of industries and applications, this research primarily focuses to explore the accurate and efficient identification of emotional attitudes along the dimensions of pessimism and optimism using transformer-based models. The existing Indian low-resource Hindi language dataset 'विशिष चरित्र_MBTI' (Kumar et al., 2023) (pronounced as vishesh charitr) is re-annotated with pessimistic and optimistic class labels to create a new dataset 'मनोभाव' (pronounced as Manobhav) for profiling text-based emotional attitude. The re-annotation is an indirect mapping from MBTI personality trait to optimistic and pessimistic attitude categories. This two-stage annotations are based on the human psychology "Chakra" system which we put forward in this research. The proposed "Chakra" system establishes the correlation between key elements of human psychology. To summarize, the major contributions of this research are as follows:

- Introducing the human psychology "Chakra" system to demonstrate an intricate interplay within human psychology as a chakra system with four wheels, each representing a crucial facet: personality, emotion, attitude, and behaviour.
- Creating a novel dataset, 'मनोभाव' (pronounced as Manobhav) with two emotional attitude class labels {pessimistic; optimistic}
- Demonstrating the potential of four Transformer-based models, namely, mBERT, XLM-RoBERTa, IndicBERT (Dabre et al., 2021), mdeBERTa and a stacked mdeBERTa in uncovering the intricate nuances of emotional attitudes using Hindi text.
- Comparing the efficacy of these models using appropriate evaluation metrics like precision, recall, F1-score for elucidating the performance of each model's contribution to the field of psychometric natural language processing.

The rest of the paper is organized as follows. Section 2 discusses the overall human psychological make-up in terms of "Behaviour," "emotion," "attitude," and "personality" finally introducing the proposed human psychology "Chakra" system. Section 3 discusses some recent work done in the field of personality detection specifically in Indian languages. Section 4 presents the comprehensive details of the curated Hindi emotional attitude detection dataset "मनोभाव" (pronounced as Manobhav). In Section 5, the paper discusses the transformer-based model utilized in the research. The in-depth comparison and the results obtained are thoroughly analysed in Section 6. Finally, Section 7 presents concluding remarks and outlines potential avenues for future improvements in this work.

2 | OVERALL HUMAN PSYCHOLOGICAL MAKE-UP

"Behaviour," "emotion," "attitude," and "personality" are interconnected concepts that collectively contribute to understanding human psychology and individual differences. While they are distinct, they influence and shape each other in complex ways.

- *Behaviour*: Behaviour refers to the observable actions, reactions, and responses of an individual in various situations. It encompasses both verbal and nonverbal actions. Behaviour is influenced by a combination of internal factors (such as thoughts, emotions, and attitudes) and external factors (such as environmental stimuli and social context).
- *Emotion*: Emotion refers to a complex psychological state that involves feelings, physiological changes, and cognitive responses. Emotions are often triggered by internal or external stimuli and play a significant role in guiding behaviour. Emotions are subjective experiences that can range from joy and happiness to sadness, anger, fear, and more.
- *Attitude*: Attitude refers to a person's overall evaluation, belief, or opinion about a particular object, person, group, idea, or situation. Attitudes can be positive, negative, or neutral and are shaped by an individual's experiences, beliefs, values, and emotions. Attitudes influence how a person perceives and responds to the world around them.

- *Personality*: Personality is a relatively stable and enduring pattern of thoughts, feelings, behaviours, and characteristics that distinguish one individual from another. It encompasses a wide range of traits, including cognitive, emotional, and behavioural tendencies. Personality traits contribute to an individual's consistency in behaviour and emotional responses across different situations and over time.

In a cohesive fusion, these elements intertwine to craft a nuanced tapestry of human existence. Within this narrative, emotions delicately influence attitudes, attitudes intricately mould personality, and personality seamlessly directs behaviour. The relationships among these concepts can be described as follows:

- *Behaviour and Emotion*: Emotions can drive behaviour. For example, feeling excited may lead to energetic and enthusiastic behaviour, while feeling anxious might result in cautious or avoidant behaviour. Similarly, behaviours can influence emotions. Engaging in positive behaviours may lead to increased feelings of happiness.
- *Behaviour and Attitude*: Attitudes can shape behaviour. If someone has a positive attitude toward exercise, they are more likely to engage in physical activity. On the other hand, behaviours can also impact attitudes. Repeatedly engaging in a behaviour may lead to attitude change through cognitive dissonance or reinforcement.
- *Emotion and Attitude*: Emotions can influence attitudes. For example, experiencing positive emotions in a certain context may lead to the development of a positive attitude toward that context. Attitudes can also influence emotions. Having a negative attitude toward a task may result in feelings of frustration or stress when engaging in that task.
- *Personality and behaviour/Emotion/Attitude*: Personality traits play a role in shaping behaviour, emotions, and attitudes. For instance, an extraverted individual might engage in more social behaviours and experience positive emotions in social settings. Certain personality traits may also predispose individuals to specific attitudes, such as a more open-minded attitude in individuals with high openness to experience.

In summary, behaviour, emotion, attitude, and personality are intertwined elements of human psychology. They interact and influence each other in intricate ways, contributing to the complexity of individual human experiences and expressions.

2.1 | Human psychology "chakra" system

The word "chakra" means "wheel" or "disk" in Sanskrit. A chakra model is a well-established conceptual framework stemming from ancient Indian traditions, like Yoga and Hinduism, depicting the body's energy centres as spinning wheels or vortexes. It represents the body's energy centres, often depicted as spinning wheels or vortexes of energy that correspond to specific physical, mental, emotional, and spiritual aspects of human existence. Now visualizing this intricate interplay within human psychology as a chakra system with four wheels, each representing a crucial facet: personality, emotion, attitude, and behaviour reveals a captivating analogy. Like spinning wheels of energy, these harmoniously interact, contributing to the holistic balance of our inner emotional and mental landscape as shown in Figure 1.

- *Personality Wheel*: This foundational wheel embodies the essence of who we are, housing our enduring traits, tendencies, and core characteristics. It rotates with the rhythms of our unique nature, influencing how we perceive and respond to the world around us.
- *Emotion Wheel*: Enveloping the personality wheel, this vibrant sphere radiates with the hues of our inner feelings. It spins with the dance of joy, sorrow, excitement, and more, infusing our experiences with a vivid spectrum of emotional shades. Psychologist Robert Plutchik developed one of the most popular emotion wheels, known as the Plutchik wheel (Tromp & Pechenizkiy, 2014).
- *Attitude Wheel*: Nurtured by the emotion wheel, this wheel of perspective spins, shaping how we interpret and engage with life's tapestry. It revolves with optimism, pessimism, or neutrality, casting a distinct filter through which we view our surroundings.
- *Behaviour Wheel*: The outermost wheel, powered by the synergy of personality, emotion, and attitude, propels our outward actions and interactions. It rotates with precision, orchestrating our responses and movements, translating inner dynamics into tangible expressions.

In this harmonious chakra system, the personality wheel serves as the core, while the emotion, attitude, and behaviour wheels revolve around it. Just as the balanced flow of energy in a chakra system promotes well-being, the fluid interaction of these four wheels nurtures a holistic understanding of human psychology, guiding our journey through self-discovery and personal growth.

3 | RELATED WORK

Various studies collectively underscore the dynamic interaction between language, culture, and personality across a diverse range of languages and contexts. Here is a year-wise chronological summary of the personality/ attitude/emotion-based studies conducted on low-resource languages.

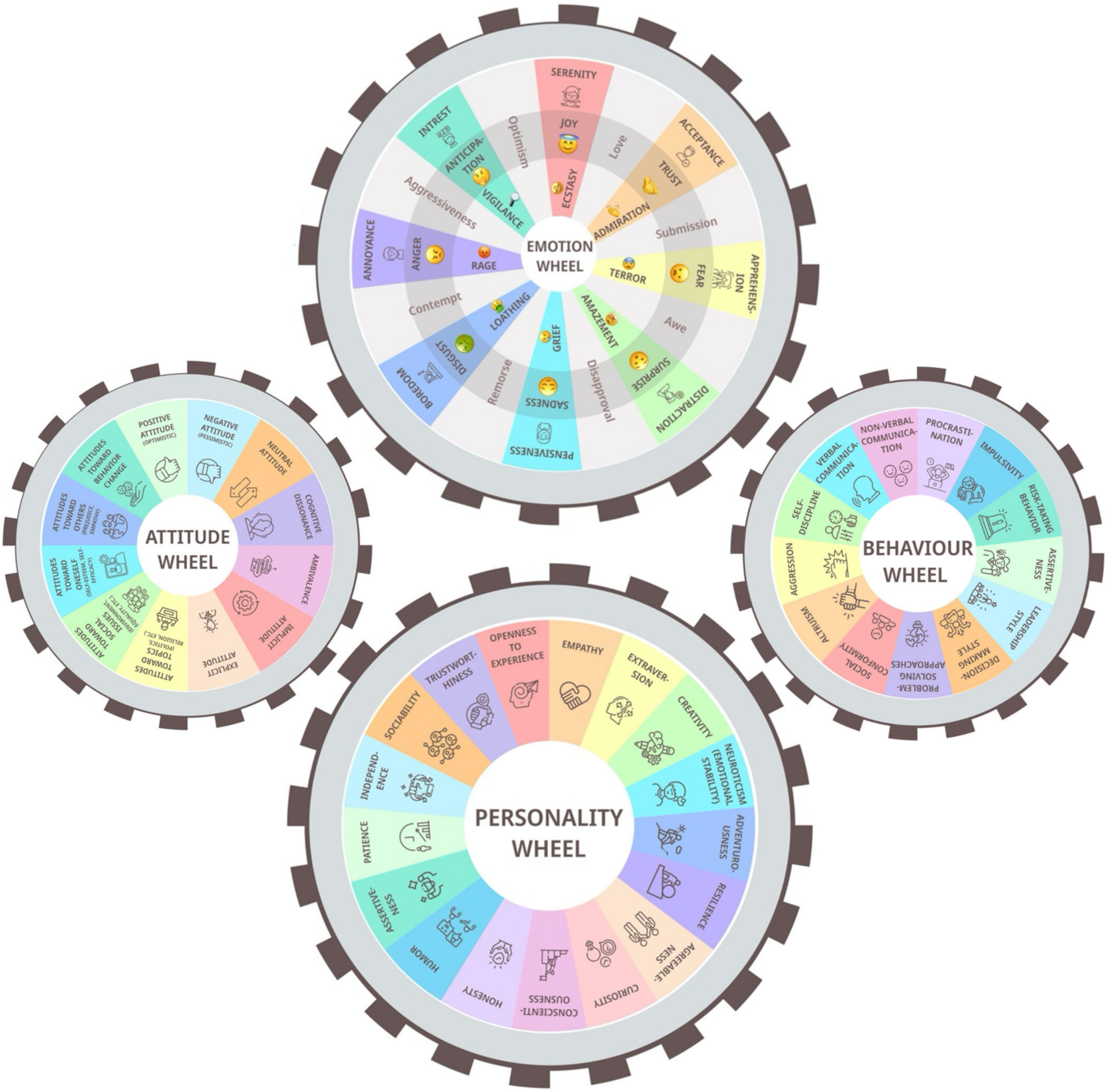

FIGURE 1 The human psychology "Chakra" system.

In 2011, Rahman and Gautam (Rahman & Gautam, 2011) conducted a study that delved into the language attitudes of Hindi-Urdu speakers in Patiali. Through an examination of preferences and attitudes toward Hindi and Urdu, the research provided valuable insights into how sociolinguistic factors influence language choices and shape cultural identity. In the same year, Koolagudi et al. (2011) curated the "IITKGP-SEHSC" corpus to refine emotion analysis and human-computer interaction through spoken Hindi, emphasizing the role of emotions in effective communication. Moving to 2013, Singh et al. (2013) carried out a seminal study focusing on personality traits within the Hindi language. Their research unveiled the intricate connection between linguistic subtleties, cultural influences, and the manifestation of personality traits unique to Hindi. This study emphasized the interplay of language, culture, and personality, enriching the empirical landscape of personality psychology. In 2015, Singh et al. (2015) introduced the validation of the Vedic Personality Inventory in Hindi. By adapting this inventory, they provided a reliable tool for assessing personality traits aligned with Vedic principles, catering to the Hindi-speaking population and enhancing the accuracy of personality evaluations. Stepping into 2019, Kumar et al. (2019) presented the "Bhaav" text corpus, a collection of textual content extracted from Hindi stories. This corpus served as a valuable resource for analysing emotions within Hindi literature, contributing to a deeper understanding of how emotions are intricately intertwined with language.

Fast forwarding to 2020, Khan et al. (2020) introduced the Vyaktitv dataset, which encompassed a diverse range of Hindi conversations along with visual content. This dataset facilitated research on personality traits and communication dynamics, showcasing the significance of linguistic nuances and paralinguistic cues in understanding the complexities of personality expression. In 2021, Junaid et al. (Asghar et al., 2021) employed advanced deep learning techniques to identify psychopathic traits within social media text. This innovative approach not only enhanced personality analysis but also opened a new avenue for AI-driven exploration of complex personality traits within the digital communication landscape. Yuan et al. (2021) explored the expression of personality traits through the Chinese language, using advanced natural language processing techniques to bridge linguistic subtleties with cross-cultural personality assessment. Moving on to 2022, Yang et al. (2022) undertook a comprehensive review of computational methods for analysing personality traits. Their study highlighted the disruptive potential of technology in the field of personality assessment, showing how various methodologies and algorithms predict traits from digital traces across different domains.

More recently, in 2023, a series of studies advanced the exploration of personality traits in various languages and contexts. Kumar et al. (2023) used a novel approach to infer MBTI personality traits in Hindi through a fusion of social psychology and computational analysis. Dandash and Asadpour (2023) employed machine learning techniques to identify personality traits in Arabic-speaking Lebanese users on social media, considering linguistic nuances and cultural context to enhance accuracy. Alsubhi et al. (2023) predicted personality traits using machine learning on Saudi Arabic tweets, revealing the intricate interplay between language, culture, and personality in social media communication.

Extensive research has been conducted to explore the use of datasets based on The MBTI for the advancement of automated personality profiling. However, a notable research gap exists in the domain of understanding emotional attitudes (optimistic or pessimistic) of individuals through an analysis of their spoken or written expressions. Remarkably, there has been a lack of prior research that explores the relationship between a person's stated emotions and their overall personality profile. The 'Manobhav' dataset, denoted as 'मनोभाव' in Hindi, emerges as a pioneering endeavour aimed at bridging this research gap by establishing a connection between the intricate layers of human psychology, particularly the interplay between personality traits and emotional attitudes.

4 | 'मनोभाव' (PRONOUNCED AS MANOBHAV) DATASET

The development of the Hindi dataset 'मनोभाव' (pronounced as Manobhav) included re-annotating an existing Indian low-resource Hindi language dataset 'वशिष चरत्रि_MBTI' (pronounced as vishesh charitr) dataset with pessimistic and optimistic class labels for profiling text-based emotional attitude. The primary dataset has a total of 969 quotes scrapped through Google and labelled for 16 different personality traits of MBTI trait theory as INTP, ISTP, INTJ, ISTJ, INFP, ISFP, INFJ, ISFJ, ENTP, ESTP, ENTJ, ESTJ, ENFP, ESFP, ENFJ and ESFJ. It is now with the help of psychologists were relabelled and mapped these traits of human personality into optimistic and pessimistic emotional attitudes that individuals hold toward various situations, events, and outcome. The 'मनोभाव' (pronounced as Manobhav) dataset has 969 quotes out of which 488 quotes are labelled as optimistic while the rest 481 are labelled as Pessimistic.

4.1 | MBTI mapping to pessimistic-optimistic attitudes

The Myers-Briggs Type Indicator (MBTI) is a widely known personality assessment tool that categorizes individuals into one of 16 personality types based on preferences in how they perceive and interact with the world. While the MBTI does not explicitly measure traits like pessimism and optimism, certain aspects of personality types might be associated with these emotional attitudes. A general overview of how certain MBTI personality types might relate with optimism and pessimism are as follows:

a. Optimistic attitude:
   - Extraverted Intuitive Types (ENFP, ENTP): These types like brainstorming and are open to new ideas. They might have an optimistic outlook due to their focus on potential and their tendency to see multiple angles.
   - Extraverted Feeling Types (ESFJ, ENFJ): Valuing harmony and relationships, these types may lean toward optimism as they prioritize positive interactions and often work to uplift others.
   - Perceiving Types (ESFP, ESTP, ENFP, ENTP): Perceiving types are more adaptable and open-ended, which can lead to an optimistic outlook since they are comfortable exploring different options and adjusting to changing circumstances.

b. Pessimistic attitude:
   - Introverted Sensing Types (ISTJ, ISFJ): These types often focus on past experiences and known information. They might exhibit cautiousness, which can sometimes lead to a more pessimistic viewpoint when dealing with uncertainty.
   - Introverted Thinking Types (ISTP, INTP): Their analytical and critical thinking can sometimes lead them to consider potential problems and challenges, which might align with a more cautious or pessimistic mindset.
   - Judging Types (ISTJ, ISFJ, INFJ, INTJ): Judging types value structure and organization, which can sometimes lead to a more realistic or cautious approach, potentially leaning toward pessimism.

The following Figure 2 visually summarizes this mapping.

It is important to remember that these attitudes or tendencies are not absolute for each personality type. Moreover, optimism and pessimism are influenced by a range of factors beyond personality type. People of any personality type can develop either of these two tendencies. The MBTI based dataset in both English and Hindi languages have proven to be quite an accurate parameter for individuals to understand their personalities. However, there has been no prior research on mapping a person's emotional attitudes (optimistic or pessimistic) based on their statements or dialogues. The 'मनोभाव' (pronounced as Manobhav) pioneer attempt in establishing an interaction between the personality and attitude chakras of human psychology.

4.2 | Data pre-processing

The pre-processing of the dataset involved a systematic sequence of procedures geared toward enhancing the dataset's quality and coherence while considering the intricacies of Hindi language specifics. Normalization and standardization of the Hindi text data assume significance, encompassing strategies to homogenize variances in spelling, accents, and diacritics. The process of tokenization, considering the distinctive linguistic structure of Hindi, plays a pivotal role in segmenting the text into meaningful units such as words or phrases, facilitating subsequent analytical operations. The extraction of textual features, including removing common stop words to enhance the data's analytical quality. Moreover, throughout the data cleaning procedure, manual validation and review play a pivotal role in rectifying intricate linguistic anomalies that automated processes might overlook. Figure 3 shows the snapshot of the pre-processed data with two columns as type and posts with 969 rows.

5 | TRANSFORMER-BASED MODELS

Transformer-based models are ideal for NLP with text due to their exceptional context understanding. Their self-attention mechanism captures both nearby and distant word relationships, vital for tasks like translation and understanding. Pre-training on large text data equips models like BERT to grasp nuanced meanings, adaptable for various tasks via fine-tuning. This versatility reduces the need for task-specific models. Transformers are computationally efficient and have revolutionized NLP by efficiently processing language data. This section of the paper discusses the transformer-based models, mBERT, mDeBERTa, XLM-RoBERTa, IndicBERT and novel stacked mDeBERTa which help in attitude prediction from the dataset as shown in Figure 4.

5.1 | mBERT (multilingual bidirectional encoder representations from transformers) model

Multilingual BERT (mBERT), developed by Google AI, is a transformative variant of the original BERT model that is designed to tackle the challenges of multilingual natural language processing (NLP). Unlike the standard BERT, which was mainly trained on English text, mBERT is pre-trained on a diverse and expansive collection of text data in multiple languages. This multilingual pre-training equips mBERT with the ability to capture language-agnostic patterns, syntax, and context that transcend individual languages. The BERT multilingual base model (uncased) model is pre-trained on 102 languages and some of them are: Georgian, German, Gujarati, Hindi, Hungarian and etc. mBERT's multilingual nature

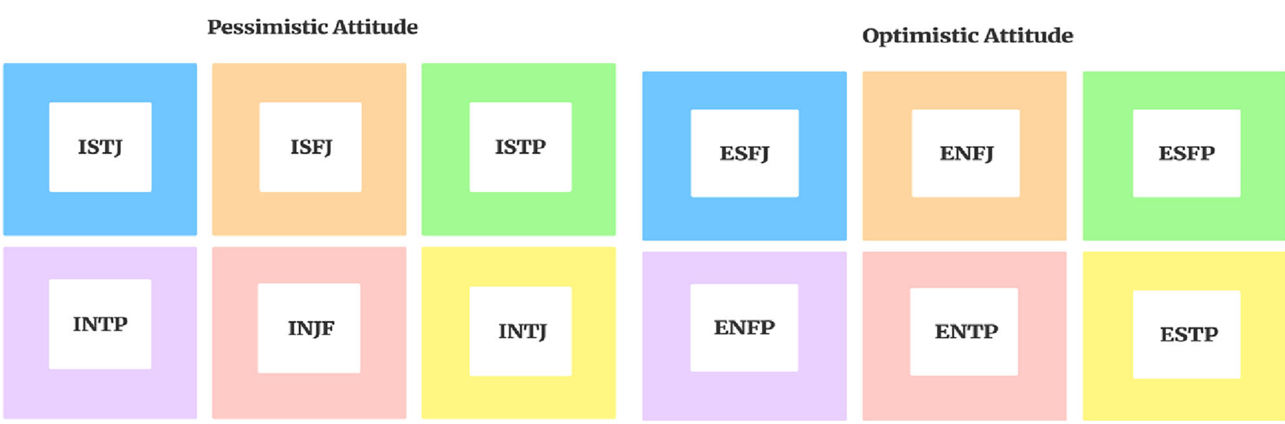

FIGURE 2 Mapping MBTI personality traits to pessimistic-optimistic attitudes.

type	posts
Optimistic	अगर एक आदमी सब कुछ नष्ट कर सकता है, तो एक लड़की इसे क्यों नहीं बदल सकती
Pessimistic	मैं अविवाहित हूं
Optimistic	आप जिस सड़क पर हैं उसे बदलने का अभी भी समय है
Pessimistic	लोग हमेशा मेरे लिए नहीं रहे हैं लेकिन संगीत हमेशा रहा है।
Optimistic	स्वागत
Pessimistic	कुतिया मैं यहीं हूँ गुस्सा पागल गुस्सा पागल
Pessimistic	कोई हम और वे नहीं हैं, यह एक भ्रम है।
Optimistic	नुकसान की गारंटी देने का सबसे अच्छा तरीका छोड़ देना है

FIGURE 3 'मनोभाव' dataset snapshot.

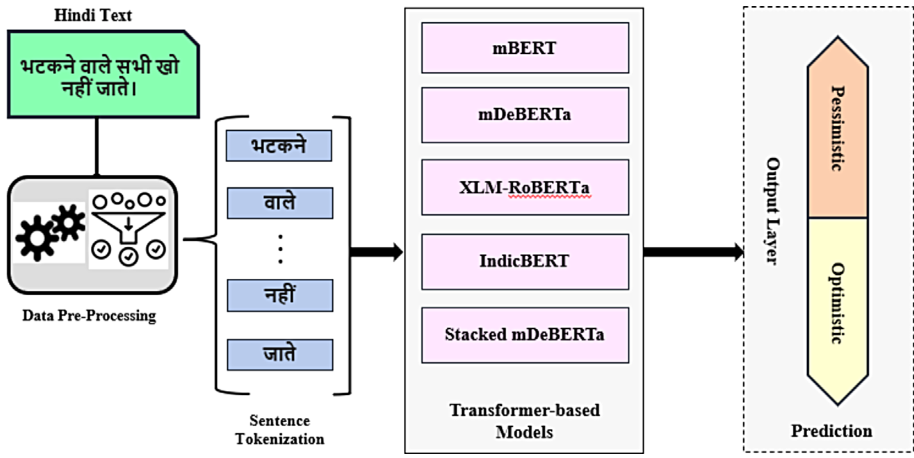

FIGURE 4 Attitude prediction using transformer-based models.

empowers it with cross-lingual transfer capabilities. It learns to recognize similarities in linguistic structures and relationships that hold true across different languages. Moreover, mBERT offers an intriguing advantage in zero-shot and few-shot learning scenarios. This means that the model can be fine-tuned with just a small amount of labelled data from a new language, even if it is not explicitly trained on that language.

5.2 | mDeBERTa (multilingual decoding-enhanced BERT with disentangled attention) model

mDeBERTa is multilingual version of DeBERTa (He et al., 2020) which use the same structure as DeBERTa and is trained with CC100 (Common Crawl 100) multilingual data. It is an advanced transformer-based language model that represents a significant evolution of the original mBERT architecture which is mainly based on BERT architecture. It is developed to address some limitations of traditional transformer models and further enhance the performance of natural language processing tasks. The mDeBERTa V3 base model have 12 layers and a hidden size of 768. It has 86 M backbone parameters with a vocabulary containing 250 K tokens which introduces 190 M parameters in the Embedding layer. This model was trained using the 2.5 T CC100 data.

5.3 | XLM-RoBERTa (multilingual RoBERTa) model

XLM-RoBERTa, developed by Facebook AI, combines BERT and RoBERTa (A Robustly Optimized BERT Pre-training Approach) models for advanced multilingual language understanding. It is trained to comprehend text across languages, using robust pre-training techniques inspired by RoBERTa. The model excels in cross-lingual tasks, learning from diverse languages. It improves contextual understanding through masked language modelling and uses byte-pair encoding for versatile tokenization. XLM-RoBERTa adapts well to specific languages via fine-tuning and employs unsupervised tasks like translation and cross-lingual modelling. Overall, it is a powerful and language-agnostic model for multilingual natural language processing.

5.4 | IndicBERT model

IndicBERT is a specialized variant of the BERT model that has been tailored and optimized for Indic languages, a group of languages spoken in the Indian subcontinent. Similar to other BERT-based models, IndicBERT aims to capture rich contextual information from text data, making it well-suited for a wide range of natural language processing (NLP) tasks. The distinct feature of IndicBERT lies in its fine-tuning and customization to better handle the linguistic nuances, structures, and characteristics specific to Indic languages. This involves training the model on large and diverse datasets from Indic languages, enabling it to learn patterns and representations that are particularly relevant to these languages.

IndicBERT is a multilingual ALBERT (Lan et al., 2019) model pre-trained exclusively on 12 major Indian languages. It is pre-trained on a novel monolingual corpus of around 9 billion tokens and subsequently evaluated on a set of diverse tasks. IndicBERT has much fewer parameters than other multilingual models (mBERT, XLM-R etc.) while it also achieves a performance on-par or better than these models.

The 12 languages covered by IndicBERT are: Assamese, Bengali, English, Gujarati, Hindi, Kannada, Malayalam, Marathi, Oriya, Punjabi, Tamil, Telugu.

5.5 | Proposed stacked mDeBERTa model

Stacking two mDeBERTa models involves employing two instances of the mDeBERTa architecture in a sequential manner. Each instance can be thought of as a separate model, and by stacking them, they form a composite architecture with specific benefits and applications. Some of the potential details and implications of stacking two mDeBERTa models are:

- Increased Representational Power: Stacking two models allows for a more complex and powerful representation of input data. The second model can capture higher-order relationships and nuances in the outputs of the first model, leading to more expressive representations.
- Feature Hierarchies: The first mDeBERTa model might capture lower-level features and structures, while the second model can build upon those features to capture more abstract and high-level patterns in the data.
- Fine-tuning and Specialization: The first model can serve as a general feature extractor, and the second model can be fine-tuned for specific tasks or domains. This approach can lead to better task-specific performance.
- Error Correction: Stacking models can help mitigate errors or inconsistencies that might occur in the predictions of a single model. The second model might learn to correct errors made by the first model, enhancing overall accuracy.
- Ensemble Effect: Stacking can be seen as a form of model ensemble, where the combined predictions of two models can lead to improved overall performance by reducing bias and variance.

It is important to note that while stacking models can offer advantages, it increases computational requirements, as each model adds to the processing load. At the same time, hyperparameters of both models need to be tuned carefully to achieve optimal performance. This includes architectural choices, learning rates, and regularization parameters.

6 | RESULTS AND DISCUSSION

This section encompasses a detailed analysis of model performance, their potential, and the hyperparameters crucial for their effectiveness. The evaluation provides valuable insights into model choice and improvement strategies, fostering a deeper understanding of the models' capabilities in text-based attitude prediction tasks. A concise summary of the limitations of the study is also presented.

6.1 | Performance of transformer-based models on 'मनोभाव' (pronounced as Manobhav) dataset

In this sub-section, we provide a comprehensive analysis of the performance of various Transformer-based models on the 'मनोभाव' dataset. A comparative evaluation based on metrics including Accuracy, Precision, Recall, F1 Score, and ROC AUC Score is presented in Table 2. The results highlight the effectiveness of each model in capturing the nuances of the dataset.

TABLE 2 Performance of all transformer-based models on 'मनोभाव' dataset.

Models	Accuracy	Precision	Recall	F1 score	ROC AUC score
mBERT	0.6712	0.6580	0.6712	0.6633	0.5911
XLM-RoBERTa	0.7474	0.7332	0.7474	0.7338	0.6618
IndicBERT	0.6851	0.6672	0.6851	0.6731	0.5979
mDeBERTa	0.7543	0.7411	0.7543	0.7411	0.6700
Stacked mDeBERTa	0.7785	0.7726	0.7785	0.7745	0.7226

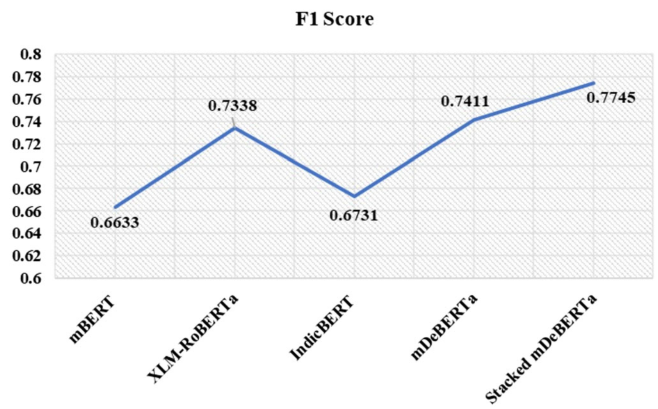

FIGURE 5 F1 comparison of transformer-based models.

Furthermore, Figure 5 visually depicts the comparison of F1 scores across all the evaluated Transformer-based models.

It is evident that the proposed Stacked mDeBERTa model outperforms the other models in terms of accuracy, precision, recall, F1 score, and ROC AUC score. It demonstrates a well-rounded performance in identifying positive cases while maintaining a good balance between avoiding false positives and capturing true positives. The high ROC AUC score (0.7226) indicates that the model is effective at ranking positive cases higher than negative cases across a range of thresholds. This implies a good ability to distinguish between the two classes. The ROC curves of all five models are shown in Figures 6–10.

6.2 | Discussion of model performance and implications

The performance of the various Transformer-based models on the 'मनोभाव' dataset yields valuable insights into their capabilities and areas for improvement. This section delves into the performance of each model, highlighting their strengths, potential enhancements, and implications for real-world applications.

- *XLM-RoBERTa and Stacked mDeBERTa Performance*:
  Among the models evaluated, both XLM-RoBERTa and Stacked mDeBERTa exhibit the strongest performance across various evaluation metrics. Notably, Stacked mDeBERTa slightly outperforms the others in terms of accuracy, precision, recall, F1 Score, and ROC AUC Score. This suggests that the ensemble technique used in Stacked mDeBERTa has effectively harnessed the strengths of multiple models, leading to enhanced predictive capabilities. The superior performance of these models underscores the significance of leveraging advanced pre-trained models for natural language processing tasks.
- *mBERT and IndicBERT Potential*:
  On the other hand, mBERT and IndicBERT exhibit moderate performance compared to XLM-RoBERTa and Stacked mDeBERTa. This indicates that while they have the potential to capture the underlying patterns in the data, there is room for improvement. Fine-tuning these models or making specific modifications could enhance their predictive capacity. Hyperparameter tuning, utilizing larger training datasets, or exploring different architectures could potentially elevate their performance.
- *Model Choice and Objectives*:
  The selection of a model ultimately hinges on the specific objectives of the task at hand and the trade-offs between precision, recall, and overall accuracy. XLM-RoBERTa provides a balanced choice that performs well across various metrics, making it suitable for tasks where a

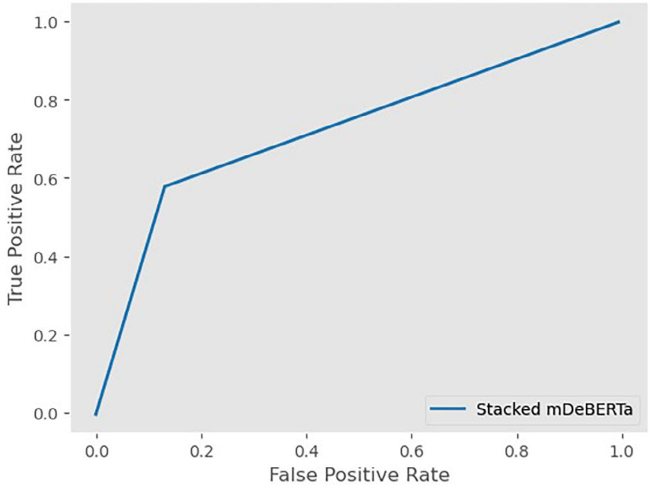

FIGURE 6 ROC curve for stacked mDeBERTa.

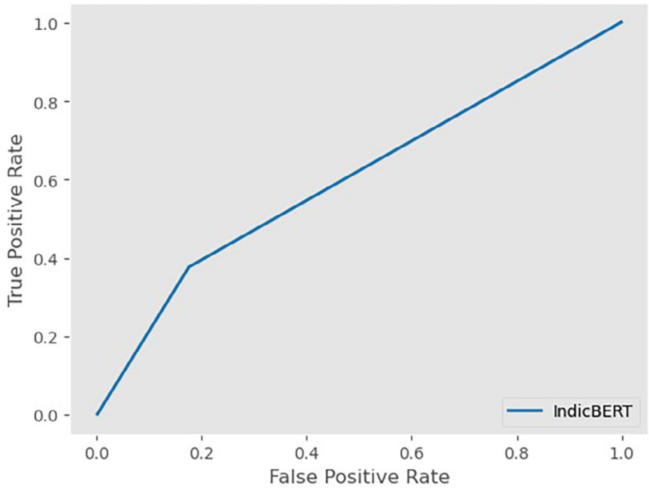

FIGURE 7 ROC curve for IndicBERT.

compromise between different performance measures is essential. Stacked mDeBERTa, with its superior performance, is particularly well-suited for tasks where high precision, recall, and accuracy are of paramount importance.

- *Potential for Improvement*:
  All models, including the top performers, show potential for enhancement, especially in terms of their ability to distinguish between classes. Ensemble techniques, like the one employed in Stacked mDeBERTa, might be further explored to boost discriminatory capabilities. Hyper-parameter tuning, adjusting model architectures, and incorporating domain-specific features could collectively contribute to better results. Continuous refinement and optimization are key to maximizing model performance.
- *Misclassified Instances and Class Imbalances*:
  A deeper analysis of misclassified instances and potential class imbalances within the dataset could provide valuable insights into areas for improvement. Understanding why certain instances are misclassified can guide refinements in the models or pre-processing steps. Addressing class imbalances, if present, through techniques like data augmentation or re-sampling could lead to more accurate predictions.

In conclusion, the evaluation results offer a clear roadmap for model selection and improvement. Leveraging the strengths of models like XLM-RoBERTa and exploring ensemble techniques such as Stacked mDeBERTa can lead to superior performance. However, the choice of model should be aligned with the task's specific goals and trade-offs between different evaluation metrics. The iterative process of refining models, tuning hyperparameters, and conducting in-depth analyses will undoubtedly contribute to achieving optimal results in natural language processing tasks.

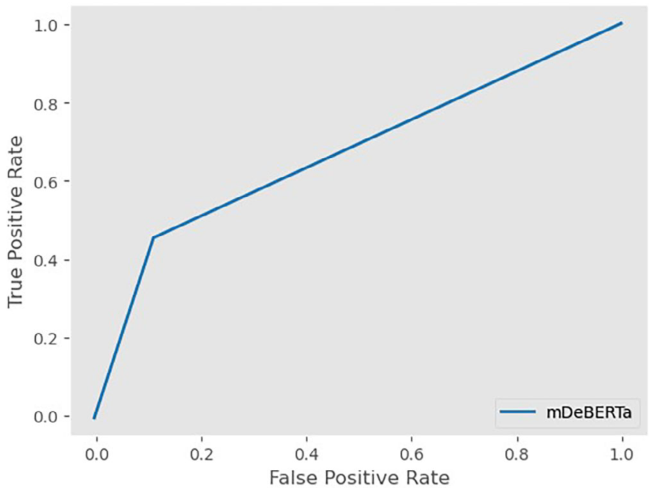

FIGURE 8 ROC curve for mDeBERTa.

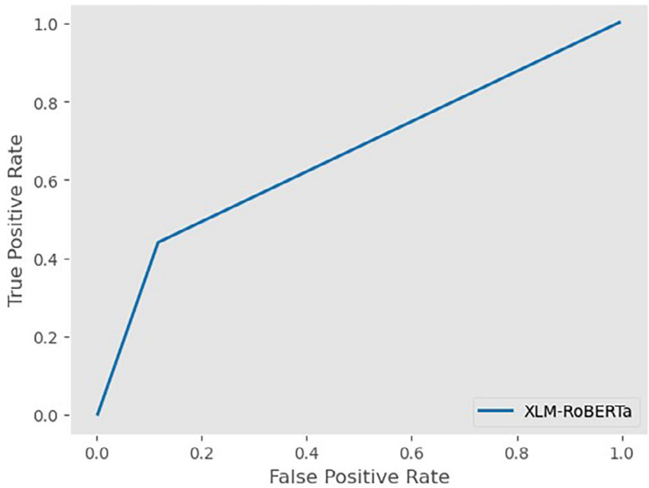

FIGURE 9 ROC curve for XLM-RoBERTa.

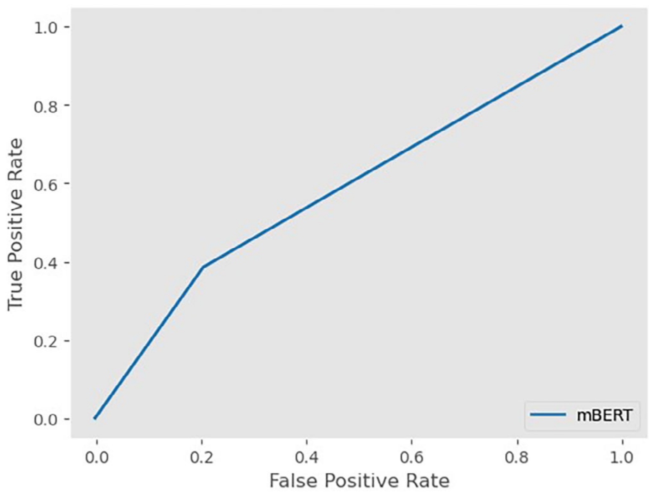

FIGURE 10 ROC curve for mBERT.

6.3 | Hyperparameters for models

Fine-tuning hyperparameters based on empirical testing is crucial to achieving optimal performance across various natural language processing tasks. In this section, we delve into the hyperparameters that were selected for each of the models in the experimental process. The chosen hyperparameters play a crucial role in influencing the performance and convergence of the models. The choices made can significantly impact the convergence speed, generalization ability, and overall effectiveness of the models. We considered the following hyperparameters:

- *Epoch*: Different models have varying numbers of epochs, indicating the number of times the entire dataset was passed through the model during training.
- *Learning rates*: Learning rates control the step size in gradient descent and play a role in determining how quickly the model converges.
- *Batch size*: Batch size influences the number of instances processed before updating the model's parameters, affecting training speed and memory usage.
- *Weight decay*: Weight decay is used to prevent overfitting by adding a regularization term to the loss function.
- *Epsilon*: The choice of epsilon (eps) influences numerical stability in optimization algorithms.

The selection of hyperparameters is a critical aspect of model training and performance optimization. The hyperparameters are tuned through experimentation to optimize the models' performance on the specific task. The variations in hyperparameters across models may account for differences in their performance. As the values chosen reflect a balance between convergence speed, generalization ability, and potential overfitting, the following are the hyperparameter values that were set for each model:

a. M-BERT: Epoch = 6, seed_value = 100, learning_rate = 4e−5, weight_decay = 0.005 eps = 1e−8, batch_size = 10
b. XLM-RoBERTa: Epoch = 10, seed_value = 100, learning_rate = 2e−5, eps = 1e−8, weight_decay = 0.01, batch_size = 16
c. mDeBERTa: Epoch = 7, seed_value = 17, learning_rate = 2e−5, eps = 1e−8, batch_size = 15
d. IndicBERT: Epoch = 10, seed_value = 25, learning_rate = 5e−5, batch_size = 10
e. Stacked DeBERTa
   - First model: mDeBERTa_1: learning_rate = 15e−6, Epochs = 6, weight decay = 0.01, batch size = 6
   - Second model: mDeBERTa_2: learning_rate = 18e−6, Epochs = 6, weight decay = 0.005, batch size = 6

6.4 | Limitations

While this study brings valuable insights into personality profiling through linguistic analysis and the application of transformer-based models, it is important to acknowledge its limitations. Firstly, the study primarily focuses on emotional attitude detection within the scope of the proposed “Chakra” system, which, while intriguing, might not encompass the entirety of emotional and personality complexity. The study's reliance on the re-annotated ‘मनोभाव’ dataset, though innovative, might still lack certain nuances or cultural variations in emotional attitudes. Additionally, the study's analysis is restricted to Hindi dialogues, limiting its generalizability to other languages and cultures. The proposed Stacked mDeBERTa model, while promising, comes with increased computational requirements and hyperparameter tuning complexities, which might hinder its practicality in resource-constrained environments. Another significant limitation is related to the dataset's imbalance in terms of optimistic and pessimistic labels. With the ratio of optimistic labels to pessimistic labels being close to 3.5:1, the dataset is skewed toward the optimistic class. Consequently, during the training phase, the model has access to a much larger number of optimistic instances compared to pessimistic instances, which affects its ability to effectively learn and predict the pessimistic class. This imbalance leads to a scarcity of pessimistic labels for training, and as a result, the model's predictive performance for pessimistic labels suffers. Specifically, the model's correct prediction rate for pessimistic labels is low. This limitation underscores the importance of balanced datasets to ensure that models can learn from sufficient instances of all classes, leading to more accurate and representative predictions.

7 | CONCLUSION

This study probes into the intricate realm of personality detection through the lens of linguistic analysis and NLP advancements. By embracing the complexities of emotional attitudes and their interplay with cognitive tendencies, the research opens doors for technology to enhance its ability to comprehend and respond to human interactions. The innovative proposition of the “Chakra” system adds depth to our understanding of human psychology, illustrating the interconnectedness of personality, emotion, attitude, and behaviour. By introducing the concept of emotional attitudes as integral components of personality, this study bridges the gap between traditional psychometric theories and the complexities of real-world human interactions. The creation of the 'मनोभाव' (Manobhav) dataset is a noteworthy contribution, enabling the study of emotional attitudes within the Indian linguistic landscape. The alignment of MBTI traits with optimistic and pessimistic attitudes introduces a new layer of analysis, further enriching our comprehension of how individuals communicate and project their emotions through text. Further, the evaluation of Transformer-based models—mBERT, XLM-RoBERTa, IndicBERT, mdeBERTa, and a stacked mdeBERTa—adds a practical dimension to the study. The proposed Stacked mDeBERTa model introduces a novel approach to enhancing the power of transformer-based architectures. By sequentially employing two instances of the mDeBERTa architecture, this composite structure offers several advantages. The meticulous comparison of the models provides insights into their efficacy in capturing the nuances of emotional attitudes from Hindi text. Among the models, XLM-RoBERTa and the proposed Stacked mDeBERTa demonstrate the highest accuracy, precision, recall, F1 score, and ROC AUC score. These findings highlight the potential of advanced pre-trained models and the efficacy of ensemble techniques in enhancing prediction capabilities. On the other hand, mBERT and IndicBERT exhibit moderate performance, suggesting room for improvement through fine-tuning, architectural adjustments, and the exploration of domain-specific features.

The study not only introduces an innovative model architecture but also provides a roadmap for selecting, refining, and optimizing transformer-based models. Through these advancements, the research contributes to the growing field of natural language processing, guiding the development of more accurate and effective models for attitude prediction tasks. Moreover, addressing the limitations about its focus on attitude detection using the "Chakra" system, the dataset's specificity, the complexities of the proposed Stacked mDeBERTa model would further enhance the robustness and applicability of the research findings. Simultaneously, addressing the data imbalance could potentially involve data augmentation techniques or resampling strategies to enhance the model's performance on the underrepresented class.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

Adi, G. Y., Tandio, M. H., Ong, V., & Suhartono, D. (2018). Optimization for automatic personality recognition on twitter in Bahasa Indonesia. *Procedia Computer Science*, *1*(135), 473–480.

Alsubhi, S., Alhothali, A., & Almansor, A. (2023). ARABIG5: The big five personality traits prediction using machine learning algorithm on Saudi Arabic tweets. *IEEE Access*, *11*, 112526–112534.

Anari, M. S., Rezaee, K., & Ahmadi, A. (2022). TraitLWNet: A novel predictor of personality trait by analyzing Persian handwriting based on lightweight deep convolutional neural network. *Multimedia Tools and Applications*, *81*(8), 10673–10693.

Asghar, J., Akbar, S., Asghar, M. Z., Ahmad, B., Al-Rakhami, M. S., & Gumaei, A. (2021). Detection and classification of psychopathic personality trait from social media text using deep learning model. *Computational and Mathematical Methods in Medicine*, *2021*, 1–10.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A. & Agarwal, S. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *33*, 1877–1901.

Cerkez, N., Vrdoljak, B., & Skansi, S. (2021). A method for MBTI classification based on impact of class components. *IEEE Access*, *9*, 146550–146567.

Dabre, R., Shrotriya, H., Kunchukuttan, A., Puduppully, R., Khapra, M. M., & Kumar, P. (2021). IndicBART: A pre-trained model for indic natural language generation. *arXiv Preprint* arXiv:2109.02903.

Dandash, M., & Asadpour, M. (2023). Recognizing personality traits using Twitter & Facebook for Arabic speaking users in Lebanon. *International Journal of Information and Communication Technology Research*, *15*(1), 45–55.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv Preprint* arXiv:1810.04805.

Fatehi, S., Anvarian, Z., Madani, Y., Mehditabar, M., & Eetemadi, S. MBTI personality prediction approach on Persian Twitter. [28_Paper.pdf (winlp.org)].

He, P., Liu, X., Gao, J., & Chen, W. (2020). Deberta: Decoding-enhanced bert with disentangled attention. *arXiv preprint* arXiv:2006.03654.

Jain, D., Kumar, A., & Beniwal, R. (2022). Personality BERT: A transformer-based model for personality detection from textual data. In *Proceedings of international conference on computing and communication networks: ICCCN 2021* (pp. 515–522). Springer Nature Singapore.

Jain, D., Kumar, A., & Garg, G. (2020). Sarcasm detection in mash-up language using soft-attention based bi-directional LSTM and feature-rich CNN. *Applied Soft Computing*, *91*, 106198.

Kenton, J. D. M. W. C., & Toutanova, L. K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of naacL-HLT* (Vol. 1, p. 2).

Khan, S. N., Leekha, M., Shukla, J., & Shah, R. R. (2020). Vyaktitv: A multimodal peer-to-peer hindi conversations-based dataset for personality assessment. In *2020 IEEE sixth international conference on multimedia big data (BigMM)* (pp. 103–111). IEEE.

Koolagudi, S. G., et al. (2011). IITKGP-SEHSC: Hindi speech corpus for emotion analysis. In *2011 international conference on devices and communications (ICDeCom)*. IEEE.

Kumar, A., & Albuquerque, V. H. C. (2021). Sentiment analysis using XLM-R transformer and zero-shot transfer learning on resource-poor Indian language. *Transactions on Asian and Low-Resource Language Information Processing*, *20*(5), 1–13.

Kumar, A., Beniwal, R., & Jain, D. (2023). Personality detection using kernel-based ensemble model for leveraging social psychology in online networks. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *22*, 1–20.

Kumar, Y., Mahata, D., Aggarwal, S., Chugh, A., Maheshwari, R., & Shah, R. R. (2019). Bhaav-a text corpus for emotion analysis from Hindi stories. *arXiv preprint* arXiv:1910.04073.

Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., & Soricut, R. (2019). Albert: A lite bert for self-supervised learning of language representations. *arXiv preprint* arXiv:1909.11942.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. & Stoyanov, V. (2019). Roberta: A robustly optimized bert pre-training approach. *arXiv Preprint* arXiv:1907.11692.

Mehta, Y., Majumder, N., Gelbukh, A., & Cambria, E. (2020). Recent trends in deep learning-based personality detection. *Artificial Intelligence Review*, *53*, 2313–2339.

Ong, V., Rahmanto, A. D. S., Williem, S. D., Nugroho, A. E., Andangsari, E. W., et al. (2017). Personality prediction based on twitter information in Bahasa Indonesia. In: *Proceedings of the 2017 federated conference on computer science and information Systems* (pp. 367–372). Prague, Czech Republic.

Rahman, M., & Gautam, A. K. (2011). Language attitude of Hindi-Urdu speakers in Patiali a perspective. *Language in India*, *11*(6), 118.

Ren, Z., Shen, Q., Diao, X., & Xu, H. (2021). A sentiment-aware deep learning approach for personality detection from text. *Information Processing & Management*, *58*(3), 102532.

Rudra, U., Chy, A. N., & Seddiqui, M. H. (2020). Personality traits detection in Bangla: A benchmark dataset with comparative performance analysis of state-of-the-art methods. In *2020 23rd international conference on computer and information technology (ICCIT)* (pp. 1–6). IEEE.

Salem, M. S., Ismail, S. S., & Aref, M. (2019). Personality traits for Egyptian twitter users dataset. In *Proceedings of the 2019 8th international conference on software and information engineering* (pp. 206–211).

Shafi, H., Sikander, A., Jamal, I. M., Ahmad, J., & Aboamer, M. A. (2021). A machine learning approach for personality type identification using MBTI framework. *Journal of Independent Studies and Research Computing*, *19*(2), 6–10.

Singh, J. K., Misra, G., & De Raad, B. (2013). Personality structure in the trait lexicon of Hindi, a major language spoken in India. *European Journal of Personality*, *27*(6), 605–620.

Singh, R., Singh, M., Rajesh, S. K., Ilavarasu, J. V., Pradhan, B., & Deshpande, S. (2015). Hindi version of Vedic personality inventory-reliability and construct validity. *International Journal of Yoga-Philosophy*, *Psychology and Parapsychology*, *3*(1), 1.

Tromp, E., & Pechenizkiy, M. (2014). Rule-based emotion detection on social media: putting tweets on Plutchik's wheel. *arXiv Preprint* arXiv:1412.4682.

Yang, L., Li, S., Luo, X., Xu, B., Geng, Y., Zeng, Z., Zhang, F., & Lin, H. (2022). Computational personality: A survey. *Soft Computing*, *26*(18), 9587–9605.

Yoneda, T., Lozinski, T., Turiano, N., Booth, T., Graham, E. K., Mroczek, D., & Terrera, G. M. (2023). The big five personality traits and allostatic load in middle to older adulthood: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *148*, 105145.

Yuan, C., Hong, Y., & Junjie, W. (2021). Personality expression and recognition in Chinese language usage. *User Modeling and User-Adapted Interaction*, *31*, 121–147.

AUTHOR BIOGRAPHIES

Dipika Jain is full-time PhD scholar within the Department of Computer Science and Engineering at Delhi Technological University (DTU) in Delhi, India. She did her Post graduation from Maharishi Dayanand University, Haryana, India. Her research lies in the realm of personality detection, where she seamlessly integrates psycholinguistics and natural language processing (NLP).

Dr. Akshi Kumar is a Senior Lecturer and Director-Post Graduate Research (PGR) in the Department of Computing at the Goldsmiths, University of London, United Kingdom. Dr. Kumar has been endorsed by the Royal Academy of Engineering, United Kingdom as an Exceptional promise in the field of Artificial Intelligence/Data Science in 2022. She has received 9 research awards for Excellence in Research from various National and International organizations. Her name has been included in the "Top 2% scientist of the world" list by Stanford University, USA in 2023, 2022 and 2021. She has published more than 100 peer-reviewed journal papers, 65+ conference papers with 5 best paper awards and 2 patents with Indian Patent Office. She has been serving as an Associate editor and guest editor in various high impact journals with reputed publishers. Her research interests are in affective computing, social network and media analytics, NLP, and AI for pervasive healthcare.

How to cite this article: Jain, D., & Kumar, A. (2024). AI unveiled personalities: Profiling optimistic and pessimistic attitudes in Hindi dataset using transformer-based models. *Expert Systems*, e13572. https://doi.org/10.1111/exsy.13572