International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-14 Issue-5, January 2026

Enhancing Accuracy of MBTI Personality Prediction Using Deep Ensemble Models and Data Augmentation Techniques

Devraj Patel, Sunita V Dhavale, Bhushan B Mhetre

Abstract: ***Personality traits prediction from text has broad applications in various fields such as recruitment, job performance analysis, adaptive learning and personalised systems. Although traditional psychological assessments are widely used today, they may be subjective and impractical for large-scale deployment because they require the physical presence of a psychologist. This study presents an automated personality prediction model utilising text data. To address class imbalance, a significant factor that degrades model performance on the personality text dataset, a two-tier oversampling strategy has been implemented. The primary contribution of this study is to systematically evaluate the efficacy of various Deep Learning Architectures, including Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and Bidirectional LSTMs, for MBTI prediction. Additionally, we have explored various ensemble learning approaches by combining separable CNNs, LeNet-5, and LSTM and BiLSTM models, thereby further improving prediction accuracy and generalisation. The experimental results show that integrating the proposed oversampling technique ensemble with the ensemble learning framework achieves higher accuracy, exceeding 87%, and outperforms previous models based solely on a single architecture or machine learning methods. The proposed method enables large-scale personality assessments to be deployed anywhere, at any time, reducing the need for the physical presence of psychologists.***

Keywords: ***Deep Learning Ensemble, MBTI Classification, Oversampling Strategy, Personality Prediction, Text-based Assessment.***

Nomenclature:
LSTM: Long Short-Term Memory
CNN: Convolutional Neural Networks
MBTI: Myers–Briggs Type Indicator
DL: Deep Learning
BiLSTM: Bidirectional Long Short-Term Memory
BERT: Bidirectional Encoder Representations from Transformers
TF-IDF: Term Frequency-Inverse Document Frequency
E–I: Extraversion–Introversion
N–S: Intuition–Sensing
F–T: Feeling–Thinking
J–P: Judging–Perceiving
GPT-2: Generative Pretrained Transformer 2

Manuscript received on 30 October 2025 | First Revised Manuscript received on 14 November 2025 | Second Revised Manuscript received on 17 December 2025 | Manuscript Accepted on 15 January 2026 | Manuscript published on 30 January 2026.
*Correspondence Author(s)

Devraj Patel*, Department of Computer Science & Engineering, Defence Institute of Advanced Technology (DU), Girinagar, Pune (MH.), India. Email ID: devraj.patel23@gmail.com, ORCID ID: 0000-0003-4874-4124.

Dr. Sunita Vikrant Dhavale, Associate Professor, Department of Computer Science & Engineering, Defence Institute of Advanced Technology (DU), Girinagar, Pune (M.H.), India. Email ID: sunitadhavale@gmail.com, ORCID ID: 0000-0003-0306-6198.

Dr. Bhushan B. Mhetre, Department of Psychiatry, Smt. Kashibai Navale Medical College and General Hospital, Narhe, Pune (MH.), India. Email ID: bhushanbmhetre@gmail.com, ORCID ID: 0009-0008-3479-1054.

I. INTRODUCTION

The assessment of personality traits plays a crucial role across various domains, including security operations, job recruitment, performance evaluation and analysis, leadership, education, personalised learning, and teamwork. Currently, many prominent and widely used traditional approaches require the presence of experienced psychologists and lengthy questionnaires, which often suffer from subjectivity and bias because they rely entirely on human judgment. Furthermore, there is a risk of candidate manipulation, as the questionnaire is fixed and the questions are known to everyone. Sometimes, the unavailability of experts prevents correct analysis. Additionally, the new normal, post-COVID-19 pandemic era, has witnessed the rapid growth of online communication using natural language data, particularly from social media platforms and digital interactions, which have become a rich resource for predicting individual personality traits. Automatic personality prediction models use such datasets and provide a scalable, efficient alternative to traditional personality assessment procedures. These models can be deployed anywhere to enhance the effectiveness of teamwork activities and strengthen the success of operations that require aligning individual personality traits with organisational goals.

In addition to job selection and recruitment, automatic personality models contribute to personalised training programs that match individual differences, understanding, and take-off levels. The teaching-learning process becomes more effective when automatic personality models are used, as personal learning and interests vary from person to person. Analysis of general communication data in natural language would further enhance dynamic monitoring of personality changes over time and inform corrective measures for leadership development, career planning, and stress management.

There are various personality models in use, but the Myers–Briggs Type Indicator (MBTI) has been widely accepted within organisations due to its ability to provide valuable insights and broad categorisation that helps individuals understand their interpersonal engagement, social preferences, and decision-making approaches. Briggs et al. [1] developed the MBTI system, which classifies human behaviour into four dimensions: interpersonal engagement, information gathering capabilities, decision-making approaches, and planning and execution skills, as presented in Table I.

Table I: Myers-Briggs Type Indicator Dimensions

Dimension I: Interpersonal Engagement	
Extraversion (E)	Focuses on the external world and interactions, tends to recharge by being around people.
Introversion (I)	Focuses on inner thoughts and feelings, tends to recharge by spending time alone.
Dimension II: Information Gathering	
Intuition (N)	Focuses on patterns, possibilities, and future implications.
Sensing (S)	Focuses on concrete information gathered through the five senses and present realities.
Dimension III: Decision-Making Style	
Feeling (F)	Makes decisions based on personal values and considerations for others.
Thinking (T)	Makes decisions based on logic, objectivity, and rational analysis.
Dimension IV: Planning and Execution	
Judging (J)	Prefers structure, organisation, closure, and decision-making.
Perceiving (P)	Prefers to keep options open, be adaptable, and be spontaneous.

Other personality frameworks also provide valuable perspectives. The Big Five Model, also known as OCEAN [2], categorises personality traits into five traits: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. It is one of the widely accepted models that provides a comprehensive view of an individual's thoughts, behaviour, and emotional status. The Dark Triad [3], used mainly by psychiatrists, focused on socially aversive personality traits defined as Machiavellianism (manipulativeness), Narcissism (excessive self-focus), and Psychopathy (impulsivity and lack of empathy). This personality framework becomes vital in contexts such as risk assessment, organisational security, and leadership evaluation. Further, the DISC Model [4] categorises an individual's personality into Dominance, Influence, Steadiness, and Conscientiousness. Like the MBTI, the DISC model is also frequently applied in the workplace and leadership training, offering practical insights into understanding communication styles and team dynamics. The HEXACO Model [5] is an extension of the Big Five, introducing Honesty–Humility as a sixth dimension alongside Emotionality, eXtraversion, Agreeableness, Conscientiousness, and Openness. This model provides a comparatively more in-depth understanding of behavioural traits such as sincerity, fairness, and modesty.

The MBTI is a self-report questionnaire that combines four facets of personality into 16 distinct personality types, based on psychological theories. Each personality type provides detailed insights into self-understanding, communication styles, team dynamics, and professional suitability, as presented in Table II.

Table II: MBTI Types and Context in the Work Environment

MBTI Type	Context in the Work Environment
ISTJ	Reliable and detail-oriented, excels in logistics management.
ISFJ	Supportive and compassionate, thrives in personnel management roles.
INFJ	Insightful and visionary, excels in strategic leadership roles.
INTJ	Analytical and innovative, excels in developing military strategies.
ISTP	Practical and adaptable, thrives in hands-on roles such as engineering.
ISFP	Creative and flexible, excels in roles involving artistic expression.
INFP	Idealistic and empathetic, excels in counselling roles.
INTP	Logical and innovative, excels in problem-solving roles.
ESTP	Energetic and resourceful, excels in dynamic roles like special operations.
ESFP	Outgoing and sociable; excels in roles that involve relationship-building.
ENFP	Inspirational and innovative, excels in creative communication roles.
ENTP	Charismatic and strategic, excels in innovation and problem-solving roles.
ESTJ	Organized and assertive, excels in leadership roles requiring structure and decision-making.
ESFJ	Supportive and sociable, excels in team-building and personnel management roles.
ENFJ	Charismatic and empathetic, excels in leadership roles involving mentorship.
ENTJ	Strategic and decisive, excels in leadership roles requiring vision and planning.

In this study, we propose a novel approach to predict 16 personality types from publicly available text data automatically. This research aims to provide automatic support for psychologists to conduct quick personality assessments of individuals and to enhance its application in job recruitment and in the analysis of personalised training requirements. The publicly available text dataset exhibits class imbalance, which affects model training and evaluation. To address imbalanced datasets and improve prediction accuracy, we proposed an integrated framework that combines effective dataset augmentation, deep learning architectures, and ensemble learning strategies. The main contributions of this study are as follows:

A. Proposed an integrated framework combining data augmentation techniques and ensemble learning using CNN and LSTM architecture that effectively captures both local textual patterns and long-range sequential dependencies.
B. Demonstrated the effectiveness of data augmentation using transformer-based LLM in addressing the class imbalance problem on MBTI text datasets.
C. Presented the comprehensive performance analysis of the proposed model, including dimension-wise evaluation and comparison with the existing state-of-the-art methods.
D. Validated the proposed models on real-world data,

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

which shows their robustness and applicability for practical use. use in domains such as recruitment, training, personalised learning and social media analysis.

E. Listed the future scope of research in automatic personality recognition using hybrid architectures and interpretable models for quick adoption across corporate and government organisations.

The subsequent sections are organised as follows: Section II reviews related work on personality recognition using deep learning approaches; Section III describes the proposed methodology and experimental setup; Section IV presents the results and analysis of the proposed ensemble techniques; and Section V concludes the study and outlines potential directions for future research.

II. RELATED WORK

Automated personality trait prediction from text has advanced significantly over the past decade, evolving from feature-based and rule-based methods to deep learning and transformer-based approaches. Early work relied on linguistic markers (word use, syntax, semantics) to examine frameworks such as MBTI, Big Five, Dark Triad, and HEXACO, laying the foundation for understanding how personality is reflected in language across marketing, mental health, and user behaviour.

Machine learning classifiers such as Support Vector Machines, Naïve Bayes, and Random Forests have been used to predict personality traits (e.g., MBTI, Big Five) from social media texts, forum posts, and other sources.[1],[6]. Though useful, these often cannot capture sequential/contextual dependencies inherent in natural language.

Deep learning (DL) models, such as CNNs, LSTMs, and BiLSTMs, have been adopted to capture these dependencies better. For MBTI prediction, Ontoum et al. [7] used BiLSTM to obtain 83.55% accuracy. Lin [8] applied BiLSTM with data augmentation and pretrained embeddings, reaching nearly 78.75% accuracy. Pradnyana et al. [9] combined BiLSTM + GloVe + QER and achieved 85.96%.

More recently, transformer and attention-based models have been applied to MBTI prediction. Kumar et al. [10] employ BERT embeddings and classification, combined with TF-IDF and other models, and report an average accuracy of approximately 87%. Later, through a customised Emoji Dataset (EmoMBTI-Net), an emoji-enhanced dataset is introduced, and fine-tuned transformer models (BART, RoBERTa, DeBERTa) are presented [11]. The fine-tuned BART model achieves an accuracy of 0.875 in MBTI prediction. Shahnazari et al. utilise RoBERTa to infer MBTI traits and gender from Telegram conversational data, achieving ~86.16% accuracy in MBTI type detection [12]. Bama et al. present a transformer variant with hierarchical layers and label-attention mechanisms tailored to MBTI classification [13] and report higher accuracy than many state-of-the-art methods. A comparison of recent transformer- and deep-learning-based prediction studies is presented in Table III.

Table III: Comparison of Recent Transformer and Deep Learning MBTI Prediction

Study	Dataset / Source	Model / Method	Accurac y / F1	Key Findings
Ontoum et al. [7]	Social network profile text	BiLSTM	83.55%	No heavy augmentation; class imbalance not fully addressed
Lin et al. [8]	Kaggle MBTI	BiLSTM + pretrained embeddings	78.75% / 78.52 F1	Moderate performance; fewer ensembles methods
Pradnyana et al. [9]	Kaggle MBTI	BiLSTM + GloVe + QER	85.96%	Limited exploration of transformer models
A Kumar et al. [10]	MBTI type dataset	BERT + TF-IDF	87%	Less on ensemble; fewer augmentation strategies
Akshi et al. [11]	Reddit	Transformers (BART, RoBERTa, DeBERTa) with emoji	87.5%	Requires emoji data; domain specificity
Shahnazari et al. [12]	Telegram conversatio ns	RoBERTa fine-tuned	86.16%	Mostly conversational data; may not generalize to other text genres

Despite these advances, issues remain: class imbalance in MBTI datasets, generalisation across different text sources, and performance variation among MBTI classes. Some studies also lack strong external validation (e.g., annotated by psychologists) or fail to combine augmentation and ensemble modelling.

III. METHODOLOGY

A. Overview

A detailed overview of the proposed methodology is presented in Fig. 1. The proposed approach suggests a novel MBTI personality prediction model using customised ensemble techniques, including CNNs, long short-term memory (LSTM), and BiLSTM. The proposed approach also involves a data augmentation technique using transformers (GPTs) to enhance the number of imbalanced data points. This study conducted a detailed analysis of the impact of various parameters on neural network architecture. It assessed the prediction accuracy of real-time data samples cross-annotated by psychologists.

In the context of the MBTI personality recognition task, each online post in dataset X was linked with four multilabel binary categories extraversion/introversion (E/I), sensing/intuition (S/N), thinking/feeling (T/F), and judging/perceiving (J/P). Using deep-learning techniques, the objective is to assign one category from each label dimension to each post x sub I. This constitutes a multi-labelled

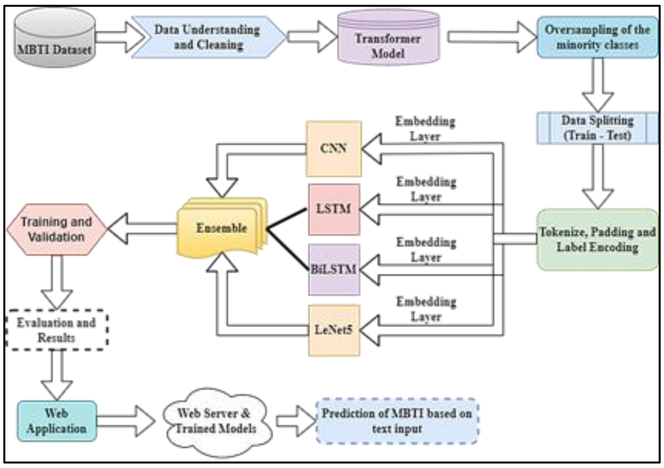
Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

supervised binary classification task within a DL framework. To achieve this, each MBTI trait was broken down into four binary dimensions, denoted as $y_{\left(\frac{E}{I}\right)}, y_{\left(\frac{N}{S}\right)}, y_{\left(\frac{F}{T}\right)}, y_{\left(\frac{J}{P}\right)}$. The model was trained independently for each dimension and predictios were made separately, resulting in four pediction functions: $f(E/I)(X), f_{\left(\frac{N}{S}\right)}(X), f_{\left(\frac{F}{T}\right)}(X), f_{\left(\frac{J}{P}\right)}(X)$. The ultimate prediction for any MBTI type x_i is obtained by merging the projections of each dimension into a multilabel vector, Y_{MBTI}.

For example, a binary vector [1, 0, 1, 0] suggests that the individual associated with the post x_i falls into the 'ESFP' category, signifying traits such as outgoing, enthusiastic, spontaneous, friendly, and sociable. Automated personality prediction helps understand individuals' social and emotional preferences, thereby enhancing team-building capabilities, particularly in domains such as security and manufacturing, where teamwork and cohesion are paramount.

B. Dataset

For our human personality prediction experiments, we selected the Kaggle platform and utilised the MBTI dataset [14] to train the deep learning models. The dataset comprises two columns, namely, Type and Post. It consists of records for 8675 individuals, each containing a snippet of their most recent 40-50 posts and a corresponding four-letter MBTI personality trait description (e.g., INTJ, ENFP, ISTP, etc.). The dataset was primarily collected from the social media platform 'personalitycafe.com', where users must complete a personality questionnaire before posting comments in the forum.



[Fig.1: Overview Diagram of the Proposed Process]

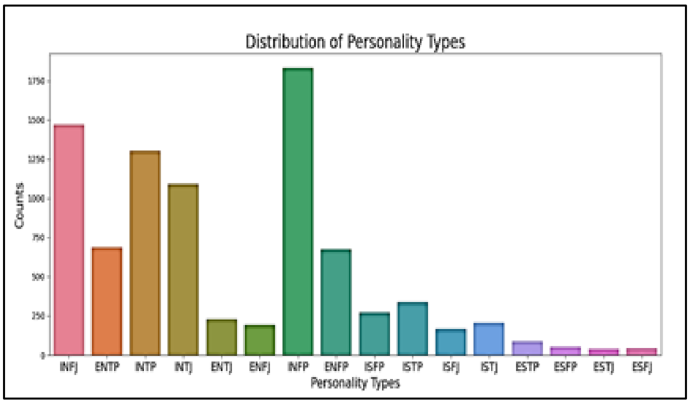
i. Data Understanding and Cleaning

Following an in-depth investigation, we discovered that the dataset has a sparse distribution of personality traits, as shown in Fig. 2. The data for a few traits, such as INFP, INFJ, INTP, and INTJ, were extensive compared to the data for other traits, primarily ESTP, ESFP, ESTJ, and ESFJ [Fig. 3] Upon further exploration, it was observed that specific personality traits, such as "E-Extroversion" and "S-Sensing", were predominantly imbalanced in the dataset. We address the data imbalance problem in the Kaggle MBTI personality types dataset by augmenting it with a pretrained transformer model and oversampling.

Additionally, several features and patterns in the text may introduce noise, compromising the accuracy of future studies. The dataset contains several issues because it was obtained from an internet forum. These problems involve variances in contractions, inconsistent letter case, noise elements (e.g., user mentions, hashtags, and URLs), and the potential for adding unrelated information. Complex additions include unlemmatized words, explicit mentions of personality types, and specific text patterns. By carefully addressing these problems, the dataset was made suitable for personality prediction. The steps taken to prepare the textual data within a specified column in the dataset are as follows:

- To ensure uniformity, the first step was to standardise the text case. Lowercasing is used to ensure that related words are consistently recognised, even when their representations in capital and lowercase differ.
- A specific function is then used to expand the contractions inside the dataset. When text is entered, this function expands contractions to their complete forms. For instance, "don't" become "do not," while "can't" becomes "cannot".
- To exclude pointless hyperlinks, mentions (such as @ and #), and URLs from the text, a regular expression was employed.
- To understand the role of language cues in expressing personality traits, the dataset was filtered so that only alphabetic characters remained. To maintain the basic language fundamental terms, multiple spaces and non-alphabetic characters are substituted.
- Certain personality type words, such as ENFJ and INTP, in both cases (upper and lower), were eliminated from the posts to prevent direct assistance in efficient training.
- Finally, the terms were reduced to their base form using lemmatisation. This approach creates sets of synonyms that improve semantic consistency.

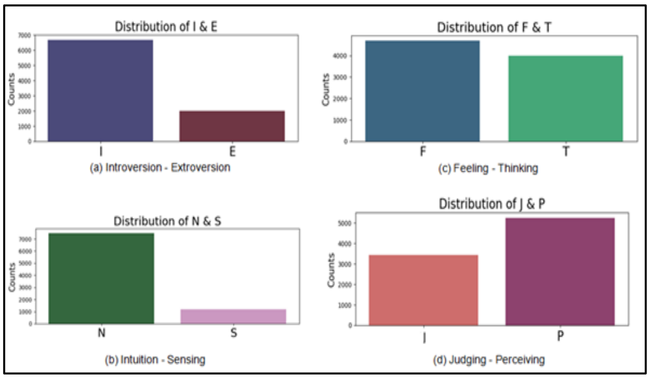
To address the need to enhance data quality, stop words that contribute little to contextual understanding, such as special characters, punctuation, and numerical data, were excluded from the dataset. Single and short character words, such as I, me, and you, were retained for relevance to personality features.



[Fig.2: Personality Distribution on the MBTI Dataset]

Retrieval Number: 100.1/ijrte.D831514041125
DOI: 10.35940/ijrte.D8315.14050126
Journal Website: www.ijrte.org

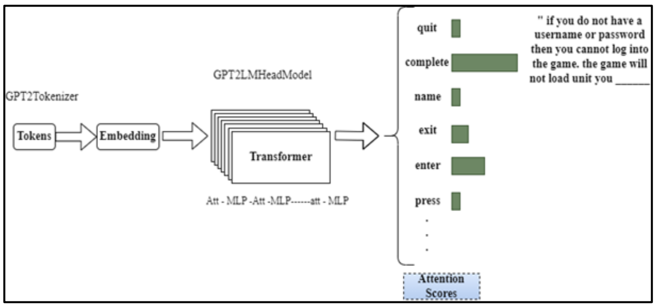
Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)



[Fig.3: Data Imbalance]

ii. Data Augmentation using Transformers

We utilised the transformer architecture, Generative Pretrained Transformer 2 (GPT-2) [15], to generate an additional dataset from the original dataset. GPT-2 is a language model trained on 1.5 billion parameters that uses the transformer architecture to generate text. The transformer architecture utilises self-attention mechanisms to assess the relative significance of words within a sentence. GPT-2 is built using transformer decoder blocks, which produce one token at a time and can process up to 1024 tokens. Each token passes through all decoder blocks along its own path, creating a vector along that path. This vector can then be scored against the model's vocabulary.



[Fig.4: Transformer Model]

In the GPT-2 model, as illustrated in Fig. 4, the randomness of sampling was controlled through the temperature parameter. A higher temperature typically produces more diverse outputs at the cost of coherence, whereas a lower temperature yields safer but repetitive text. We choose a moderate value of 0.7 for the temperature parameter. It helps balance creativity with contextual text generation, further ensuring that the created MBTI text dataset is diverse and relevant within the given context. Further, we considered top_k (k=50) and top_p (p=0.95) sampling. 'top_k' limits the generated vocabulary to the top 50 likely tokens in each step, whereas 'top_p' sampling considers only those tokens whose probability reaches 95%. Both values in the text generation allow for controlled randomness without affecting the text's fluency. These values make the text generation more dynamic in the context and also avoid over-constraint in the model.

Other parameters include a context window size of 150 tokens. It was considered to provide sufficient context while maintaining computational efficiency during generation. A sample size of 5 was selected to balance computational cost by generating multiple candidate sequences from which high-quality outputs can be selected. The 'no-repeat n-gram' size was set to 2 to avoid repetitive phrasing, which further prevents the model from generating identical bi-grams in the generated text [15],[16].

In this study, we employed a unique data-augmentation method. First, we generated five datasets by splitting the original text data into chunks of varying lengths, each containing 1 to 5 sentences. In the original dataset, the ||| symbol was used as a delimiter to separate the individual sentences within each post. Each dataset is then augmented using the transformers library. Thus, each dataset contained generated versions of the original posts, with each chunk generated independently to maintain coherence and contextual similarity within the dataset.

Each generated dataset underwent another round of data cleaning and was then merged with the original text within the post. The inclusion of generated text in the original text also served as additional noise, further improving the model training and its robustness. It was found that the final augmented dataset has nearly 14% more tokens than the original dataset, as shown in Table IV.

Table IV: Token Counts for MBTI Types

MBTI Types	Number of Tokens in the Original Dataset	Number of Tokens in the Augmented Dataset	% Increase
E	1,182,960	1,357,874	14.78
I	3,951,834	4,481,054	13.39
N	4,454,368	4,990,319	12.03
S	680,426	848,609	24.71
F	2,819,270	3,195,072	13.32
T	2,315,524	2,643,856	14.17
J	2,050,403	2,335,636	13.91
P	3,084,391	3,503,292	13.58

Although the data size has increased and more data points are obtained for executing deep learning models, there are some prominent issues and constraints in generating data using the transformer architecture, particularly for models such as GPT-2:

- Transformer models have a finite context window, which potentially leads to coherence issues in longer text passages.
- The generated text may inadvertently reflect biases or contain offensive language present in the training data.
- Balancing coherence and diversity in generated text is challenging and requires careful parameter tuning.
- Generating data relies heavily on pretrained models, which may not always generalise well without additional customisation.

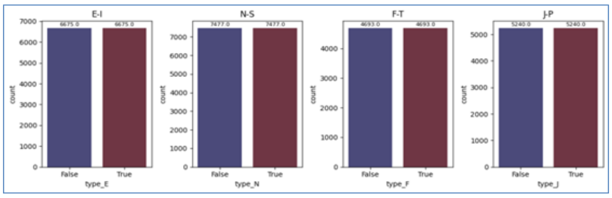
Google's transformer-based, pretrained bidirectional encoder representations from transformers (BERT) language model was used to check the quality of the generated text [17]. The similarity between the generated and reference (original) texts was measured using the BERT score. It is computed from the contextual embeddings of the generated and referenced texts, yielding a more nuanced evaluation of the generated text's quality.

In this method, we used the "score" function in the "bert_score" package to calculate the F1 score, precision, and recall for both the GPT-generated text and the original text obtained from Kaggle. The average F1 score in the GPT-generated dataset for 4917 posts was 0.9286. This value demonstrates strong contextual mapping and association between the original and generated text, making it suitable for data augmentation.

iii. Oversampling the Minority Classes in the Dataset

Because text generation with GPT has certain limitations, we employed oversampling to address class imbalance in the dataset. In this study, the Random Over-Sampling technique, implemented with the Python package 'RandomOverSampler', was used to oversample minority classes. This method works by randomly duplicating instances from the minority class until a balanced distribution is reached. A threshold value can be set to oversample the minority class and undersample the majority class, thereby obtaining a balanced distribution without compromising text quality. The balanced representation of the MBTI traits in the augmented dataset is shown in Fig.5.

As the number of prominent instances in the minority class increases, the random oversampling method mitigates the risk that the model will exhibit a bias toward the majority class during training. Consequently, the model becomes more proficient at accurately classifying instances from both the majority and minority classes, thereby further improving the model's generalisation.



[Fig.5: Over-Sampled Dataset Representation]

iv. Data splitting: Training, Validation and Testing

Only the relevant columns in the dataset, such as "cleaned_post," "E-I," "N–S," "F-T," and "J-P," were used to train the proposed method. One-hot encoding was used to change categorical variables such as "E-I," "N–S," "F-T," and "J-P" into dummy variables, as shown in Fig. 6. Each dimension is represented by a binary column (True/False), which prepares the dataset for training the deep learning models for binary classification tasks. We used stratified splitting in the dataset to preserve the uniformity in the class distribution. The dataset was divided into training (80%) and testing (20%) sets for each personality dimension.

	cleaned_post	E-I	N-S	F-T	J-P	type_E	type_I	type_N	type_S	type_F	type_T	type_J	type_P
8669	always think cat fi doms reason website become...	I	S	F	P	False	True	False	True	True	False	False	True
8670	thread already exists someplace else heck dele...	E	N	F	P	True	False	True	False	True	False	False	True
8671	many question thing would take purple pill pic...	I	N	T	P	False	True	True	False	False	True	False	True
8672	conflicted right come wanting child honestly m...	I	N	F	P	False	True	True	False	True	False	False	True
8673	long since personalitycafe although seem chang...	I	N	F	P	False	True	True	False	True	False	False	True

[Fig.6: Data Splitting and Label Encoding]

v. Word Embedding

The word embedding process plays a crucial role in language processing tasks, particularly in personality recognition from text data, as it enables the encoding of semantic meaning in text, reduces feature dimensionality, and further improves the dataset's generalisation capabilities. A few popular techniques include Word to Vector (word2vec) [18], Term Frequency-Inverse Document Frequency (TF-IDF) [19], and Global Vectors for Word Representation (GloVe) [21].

In our study, we trained the models with customised word embeddings using the Keras embedding layer. This approach enabled the capture of the dataset's unique language patterns and semantics. Using customised embedding also helped tailor the representation space to the dataset's characteristics and further optimised it for personality-related modelling tasks. Additionally, compared to pretrained embeddings such as Word2Vec or GloVe, this approach helped incorporate domain expertise, primarily psychological knowledge, potentially yielding more effective semantic representations and more interpretable results.

In our experiment, we configured the embedding layer with three key parameters: input_dim, the vocabulary size; output_dim, the vector representation length; and input_length, the maximum sequence length per iteration. We considered input dimension (input_dim) as the maximum words, i.e 5000, output dimension (out_dim) as 32 and the input length as the maximum length of vocabulary

C. Ensemble Deep Learning Models

We explored the design of ensemble deep learning (DL) models to improve prediction accuracy for MBTI personality types from textual data. Ensemble approaches generally combine the complementary architectures of individual models to leverage their strengths, thereby enhancing the generalisation and robustness of the overall model in personality prediction tasks. In this study, we primarily used CNN, LSTM, and LeNet5 architectures.

i. Ensemble of Convolutional Neural Network (CNN) with the LeNet-5 Architecture

In this study, one-dimensional convolutional neural networks (Conv1D) [22] have been used to capture local patterns in sequential text data. In a Conv1D layer, a convolutional filter (or kernel) slides along the temporal dimension of an input sequence and performs element-wise multiplication and summation to generate output features. The network learns these filter weights during training to detect meaningful patterns in the sequence.

Let $X \in R^{T \times D}$ denote an input sequence, where T represents the sequence length (number of tokens or time steps), and D indicates the feature dimensionality (e.g., embedding size). Assume a Conv1D layer has K filters, each of size F (filter length). The output feature map $Z \in R^{(T-F+1)\times K}$ is computed as:

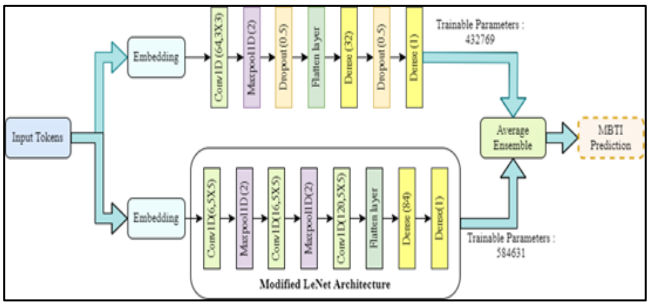
$$Z_{(i,k)} = \sigma\left(\sum_{m=1}^{F} W_{(m,k)} \, . X_{(i+m-1)} + b_k\right) \quad \ldots \quad (1)$$

where:

- $Z_{(i,k)}$ is the activation of the k-th filter at position i in the output feature map,
- $W_{(m,k)}$ represents the m-th weight of the k-th filter,
- $X_{(i+m-1)}$ is the input vector at position (i + m − 1),
- b_k is the bias term for the k-th filter,
- $\sigma(.)$ denotes the activation function, typically ReLU.

We employed the LeNet-5 architecture [23] to ensemble the Conv1D results. The LeNet-5 architecture consists of multiple convolutional and pooling layers, followed by a fully connected layer. It was initially developed for image classification, but when combined with a CNN, it can effectively extract hierarchical features from personality-related text data.

In this study, we propose an ensemble approach that integrates Conv1D with LeNet-5 [23] for predicting MBTI traits. Conv1D captures sequential dependencies while LeNet-5 extracts hierarchical feature representations. The combination enables the model to leverage both temporal and structural patterns present in MBTI text sequences, resulting in improved prediction performance, as shown in Fig. 7.



[Fig.7: Proposed Conv1D-LeNet Ensemble Model for MBTI Personality Prediction]

ii. *Ensemble of Long Short-Term Memory (LSTM) with a Convolution Network*

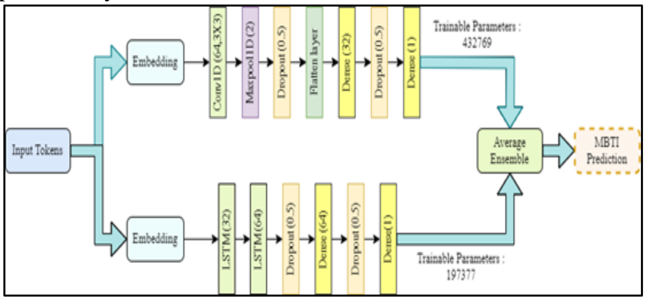
In this ensemble approach, we employed a Long Short-Term Memory (LSTM) network [25] in conjunction with a 1D Convolutional Neural Network (Conv1D) [26]. This approach leverages the complementary strengths of each architecture: LSTMs are well-suited for capturing long-range dependencies and temporal patterns in sequential data. At the same time, Conv1D efficiently extracts local feature patterns and reduces the computational complexity of text data.

Mathematically, let X R to the power TXD as in the above para...X RTxD denote the input sequence of data with length T and feature dimension D (i.e. word embeddings). The LSTM layer computes a sequence of hidden states {h1, h2, hT} that further encodes temporal dependencies using gated mechanisms (input, forget, and output gates). On the other hand, the Conv1D layer applies K filters of size F along the temporal dimension, generating feature maps that capture local contextual patterns within the LSTM outputs. The combined output is passed through fully connected layers for the final prediction of MBTI personality traits.

The LSTM-CNN ensemble design provides several advantages in the model training and evaluation: -

- *Accuracy:* As the LSTM captures long-term dependencies, it becomes crucial for understanding personality-related context in text.
- *Efficiency:* The use of Conv1D reduces model parameters and computational cost by summarising local patterns before passing to the fully connected layers.
- *Robustness:* The LSTM-CNN ensemble mitigates the limitations of individual models, hence improving the generalisation across diverse text samples.

The proposed ensemble architecture of LSTM-CNN is presented in Fig. 8, which illustrates the flow from input sequences through LSTM and CNN layers to the final classification layer. The hyperparameters were accordingly adjusted to obtain the optimised results in predicting personality traits.



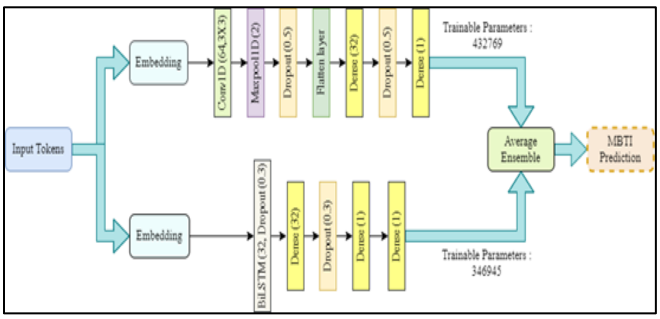
[Fig.8: Proposed Conv1D-LSTM Ensemble Model for MBTI Personality Prediction]

iii. *Ensemble of Bidirectional Long Short-Term Memory (BiLSTM) with a Convolution Network*

In this ensemble model for MBTI text prediction, we propose integrating a Bidirectional Long Short-Term Memory (BiLSTM) network [9] and a 1D separable convolutional network (CNN) [24]. The BiLSTM processes input sequences in both forward and backward directions, thereby enabling the model to capture contextual dependencies from past and future tokens simultaneously, which is crucial for understanding nuanced patterns in personality-related text.

Let X R to the power TXD denote the input sequence of length T and embedding dimension D. The BiLSTM computes a forward hidden state ht and a backward hidden state ht← for each time step t, producing a concatenated hidden representation h to the base t. This representation encodes bidirectional context across the sequence.

The BiLSTM output is then merged with results from a 1D separable convolutional layer, which applies channel-wise convolutions to each feature map independently. This ensemble method reduces computational overhead while effectively capturing the local patterns within the BiLSTM representations. The overall output is subsequently aggregated using an average ensemble method, which combines predictions across multiple filters and time steps to improve the model's robustness and prediction accuracy.



[Fig.9: Proposed Conv1D-BiLSTM Ensemble Model for MBTI Personality Prediction]

Fig. 9 illustrates the proposed ensemble architecture of the BiLSTM–CNN model, showing the flow from input sequences through the BiLSTM and CNN layers to the final prediction of MBTI traits.

The combination of

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

BiLSTM and Convolutional Network captures dependencies from both past and future tokens using BiLSTM, which enhances sequence modelling, and Separable convolutions reduce model parameters and computational cost without compromising feature quality. The average ensemble aggregation further stabilises the predictions and improves generalisation across diverse MBTI text input sequences.

iv. Evaluation Metrics

We employed both 10-fold cross-validation and standard split dataset techniques to evaluate the performance of our proposed ensemble model. Additionally, the final performance of the models was assessed using the following key metrics: -

- **Accuracy:** This evaluation metric, primarily used for binary classification tasks, provides an overall assessment of the model but may not be sufficient for imbalanced datasets.

$$Accuracy = \frac{(Number\ of\ Correct\ Predictions)}{(Total\ Number\ of\ Predictions)} \quad \ldots \quad (2)$$

- **Precision:** It is also called the positive predictive value. Precision measures the fraction of relevant predictions among all predictions.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad \ldots \quad (3)$$

- **Recall:** It is also referred to as the actual positive rate or sensitivity. It is the proportion of relevant instances that the model correctly retrieves.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad \ldots \quad (4)$$

- **F1-Score:** A balanced metric for assessing false positives and false negatives. It is the harmonic mean of precision and recall, offering a higher value for the optimal balance.

$$F1 - Score = \frac{2 * (Precision * Recall)}{(Precision + Recall)} \quad \ldots \quad (5)$$

To optimise the learning process, the models were trained using the Adam optimiser [26], which updates the model weights according to the equation below. It is designed to handle deep hierarchical attention from textual data, making it suitable for classifying personality-related features.

IV. RESULTS AND DISCUSSIONS

In this study, we used the above performance metrics to analyse the performance of various DL-based methods.

A. Performance Comparison of the Original vs Augmented Datasets

Table V summarizes the dimension-wise accuracy of the individual models. Table VI presents the dimension-wise accuracy for the four MBTI personality dimensions—Extraversion–Introversion (E–I), Intuition–Sensing (N–S), Feeling–Thinking (F–T), and Judging–Perceiving (J–P)—across the three individual lightweight neural network models.

Table V: Performance of Models on the MBTI Dataset

Models	Trainable Parameters	Average Accuracy on Original Dataset	Average Accuracy on Augmented Dataset
CNN1D	432769	0.731	0.787
LSTM	197377	0.736	0.755
BiLSTM	346945	0.717	0.742

Among the individual models, CNN1D achieved the highest overall performance, with an average accuracy of 78.7%, outperforming both LSTM (75.5%) and BiLSTM (74.2%). Primarily, the CNN1D model performed better in the N–S and E–I dimensions, demonstrating its ability to capture discriminative local patterns within the text data. While all three models performed reasonably well on the F–T dimension, LSTM and BiLSTM lagged on the J–P dimension, highlighting their potential limitations in modelling specific sequential dependencies.

Table VI: Dimension-Wise Performance of Individual Models

Models	E-I	N-S	F-T	J-P	Avg Acc.
CNN1D	0.87	0.93	0.7	0.65	0.787
LSTM	0.84	0.92	0.69	0.57	0.755
BiLSTM	0.78	0.87	0.69	0.63	0.742

Based on these results, CNN1D was selected as the primary model to form an ensemble with LSTM and BiLSTM. Combining these models leverages the complementary strengths of CNNs and LSTMs to enhance the overall accuracy and robustness of MBTI personality prediction.

B. Dimension-wise Performance (Accuracy in %) of Proposed Ensemble Models

All ensemble models demonstrated a significant improvement over the individual lightweight models, indicating that the combination of complementary architectures enhances both local feature extraction and sequence modelling. Table VII presents the dimension-wise accuracy of the proposed ensemble models for the four MBTI personality dimensions.

Among the ensembles, CNN1D + LSTM achieved the best performance with an average accuracy of 87.7%. Notably, this ensemble model achieved the highest performance across the N–S, F–T, and J–P dimensions, highlighting its effectiveness in integrating sequential learning from an LSTM with convolutional feature extraction from a CNN1D. The CNN1D + BiLSTM and CNN1D + LeNet5 ensembles also demonstrated strong performance, with average accuracies of 85.2% and 82.5% respectively. The experimental results clearly highlighted the contributions of bidirectional context and hierarchical feature learning.

Retrieval Number: 100.1/ijrte.D831514041125
DOI: 10.35940/ijrte.D8315.14050126
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

Table VI: Performance Comparison of Ensemble Deep Learning Models

Ensemble Models	Trainable Parameters	E-I	N-S	F-T	J-P	Avg Accuracy
CNN1D + LeNet5	1017400	0.91	0.90	0.73	0.76	0.825
CNN1D + LSTM	**630146**	**0.90**	**0.97**	**0.81**	**0.83**	**0.877**
CNN1D + BiLSTM	779714	0.89	0.93	0.78	0.81	0.852

Since the CNN-LSTM performed better than the other ensemble models, we selected it for further comparison with state-of-the-art methods in the literature.

C. Performance Comparison of the Proposed Ensemble Model and Existing Works

The comparative performance of the proposed CNN1D + LSTM ensemble with representative state-of-the-art MBTI prediction models from the literature is presented in Table VIII. The proposed ensemble achieves an average accuracy of 87.7%, outperforming all previously reported models.

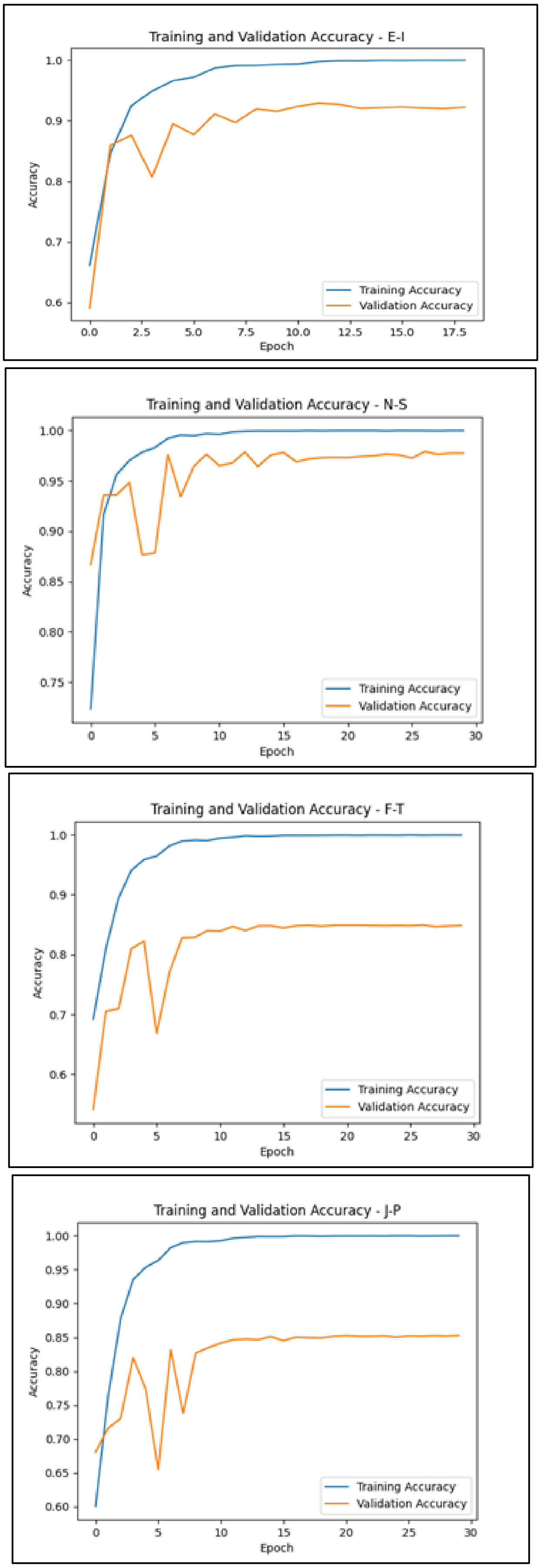
Table VII Comparison of Proposed Ensemble Model with Existing Deep Learning and Transformer-based MBTI Prediction Studies

Authors / Study	Models / Methods	Dataset / Source	Average Accuracy
Sakdipat Ontoum et al. [7]	BiLSTM	Social network profile text	83.55%
Hao Lin et al. [8]	DLP-BiLSTM + Pretrained embeddings	Kaggle MBTI	78.52%
Aditra Pradnyana et al. [9]	Hybrid BiLSTM with GloVe + QER	Kaggle MBTI	85.96%
A Kumar et al (2023) [10]	BERT + TF-IDF	MBTI dataset	85.0%
Akshi et al. (2024) [11]	Transformers (BART, RoBERTa, DeBERTa) with emoji context	Reddit	87.5%
Shahnazari et al. [12]	RoBERTa fine-tuned	Telegram conversatio ns	86.16%
Proposed Model (CNN1D + LSTM)	**CNN1D + LSTM Ensemble**	**Augmente d MBTI**	**87.70%**

The improvement over previous models underscore the effectiveness of combining a 1D CNN with an LSTM for MBTI prediction. While earlier approaches relied primarily on sequential modelling or hybrid embeddings, the proposed ensemble leverages both local feature extraction (via CNN1D) and long-range sequence modelling (via LSTM), resulting in superior generalisation across all four MBTI dimensions. This demonstrates the potential of ensemble architectures for achieving more accurate, robust predictions of personality traits from text data.

D. Training and Validation Accuracy Plots

The training and validation accuracies for each trait during training with the best-performing ensemble model, consisting of Conv1D with an LSTM, are shown in Fig. 10. Slight overfitting was observed in the two dimensions (F-T and J-P). In contrast, it aligns perfectly with the other two dimensions (E-I and N-S), providing a strong fit for the personality prediction model.



[Fig.10: Training and Validation Accuracy of Each Dimension]

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

V. CONCLUSIONS

In this study, we propose a novel ensemble deep learning framework for predicting MBTI personality types from textual data that combines CNNs and LSTMs to leverage both local feature extraction and long-range sequential dependencies. Our approach also demonstrates that leveraging multiple models can more effectively capture complex patterns in text data than a single model. The combination of CNNs and LSTMs enables us to leverage the CNN's ability to capture local dependencies and patterns, as well as the LSTMs' strength in understanding long-term dependencies in sequential data. The experimental results further demonstrated that integrating the oversampling technique with transformers into ensemble deep learning methods is lightweight, robust, and outperforms both individual models and other state-of-the-art methods. The superior performance of data augmentation with transformers also highlights its effectiveness and suitability for handling data imbalance in the MBTI text dataset.

The hybrid architecture that integrates attention mechanisms or transformer encodings with ensemble models may be explored to boost accuracy further and improve generalisation. The Explainable AI (XAI) tools can also be explored in the future to visualise the learned features and identify the decision-making features in the feature set that could be crucial for practical deployment, particularly in organisations across recruitment, selection and assessment, training, and personalised learning domains.

Overall, the proposed approach demonstrated a scalable, accurate, and reliable methodology for automated MBTI personality prediction using an imbalanced dataset, with potential for broad applicability across corporate, governmental, and social research contexts.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to DIAT, Pune, for providing access to the laboratory facilities essential for conducting this research. We also extend our appreciation to Dr Rajendra S. Deodhar, Sc-G, ARDE, DRDO, Pune, for his valuable insights and constructive feedback on the manuscript, which significantly improved the quality of this work.

DECLARATION STATEMENT

This section provides detailed information on funding, ethical considerations, and the availability of data and code for this research. All authors confirm their consent for publication, declare no competing interests, and outline the contributions and responsibilities related to this study.

As the article's author, I must verify the accuracy of the following information after aggregating input from all authors.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** This article has not been funded by any organizations or agencies. This independence ensures that the research is conducted objectively and without external influence.
- **Ethical Approval and Consent to Participate:** The content of this article does not necessitate ethical approval or consent to participate with supporting documentation.
- **Data Access Statement and Material Availability:** The adequate resources of this article are publicly accessible.
- **Author's Contributions:** The authorship of this article is contributed equally to all participating individuals.

REFERENCES

1. Ryan, G., Katarina, P., Suhartono, D.: MBTI personality prediction using machine learning and SMOTE for balancing data based on statement sentences. Information 14(4) (2023), DOI: https://doi.org//10.3390/info14040217.
2. Soto, C. J. (2018). Big Five personality traits. In M. H. Bornstein, M. E. Arterberry, K. L.Fingerman, & J. E. Lansford (Eds.), The SAGE encyclopedia of lifespan human development (pp. 240-241). Thousand Oaks, CA: Sage. URL: https://www.researchgate.net/publication/324115204_Big_Five_personality_traits.
3. Kus Hanna Rahmi (2024). The Dark Triad Personality: The Impact and How to Manage at Work. International Journal of Research and Innovation in Social Science (IJRISS), 8(02), 2074-2082. DOI: https://doi.org//10.47772/IJRISS.2024.802147.
4. Utami, E., Hartanto, A.D., Adi, S., Oyong, I., Raharjo, S.: Profiling analysis of disc personality traits based on Twitter posts in Bahasa Indonesia. Journal of King Saud University - Computer and Information Sciences 34(2), 264–269 (2022) DOI: https://doi.org/10.1016/j.jksuci.2019.10.008.
5. Jia, R., Bahoo, R., Cai, Z., Jahan, M.: The hexaco personality traits of higher achievers at the university level. Frontiers in Psychology Volume 13 - 2022 (2022) DOI: https://doi.org/10.3389/fpsyg.2022.881491.
6. Majima, Seiyu & Markov, Konstantin. (2022). Personality Prediction from Social Media Posts using Text Embedding and Statistical Features. 235-240. DOI: https://doi.org/10.15439/2022F133.
7. Ontoum, S., Chan, J.H.: Personality type based on Myers-Briggs Type Indicator with text posting style by using traditional and deep learning. arXiv preprint arXiv:2201.08717 (2022) DOI: https://doi.org/10.48550/arXiv.2201.08717.
8. Lin, H.: Dlp-personality detection: a text-based personality detection framework with psycholinguistic features and pre-trained features. Multimedia Tools and Applications 83, 1–20 (2023) DOI: https://doi.org/10.1007/s11042-023-17015-z.
9. Pradnyana, G.A., Anggraeni, W., Yuniarno, E.M., Purnomo, M.H.: Enhancing MBTI personality trait prediction from imbalanced social media data using hybrid query expansion ranking and GloVe-BiLSTM. In: 2023 IEEE International Conference on Fuzzy Systems (FUZZ), pp. 1–6 (2023). DOI: https://doi.org/10.1109/FUZZ52849.2023.10309718.
10. Kumar, Akshi & Beniwal, Rohit & Jain, Dipika. (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. DOI: https://doi.org/10.1145/3571584.
11. Kumar, A., Jain, D.: Emombti-net: Introducing and leveraging a novel emoji dataset for personality profiling with large language models. Lecture Notes in Networks and Systems Proceedings of International Conference on Computing and Communication Networks (2024) DOI: https://doi.org/10.21203/rs.3.rs-4768237/v1
12. Shahnazari, K., & Ayyoubzadeh, S.M. (2025). Who Are You Behind the Screen? Implicit MBTI and Gender Detection Using Artificial Intelligence. ArXiv, https://arxiv.org/abs/2503.09853.
13. Bama S, Hema M S, Esakkirajan S, Nageswara Guptha M, A hierarchical transformer network with label attention for personality prediction by MBTI classification, Applied Soft Computing, Volume 178,113267 (2025). DOI: https://doi.org/10.1016/j.asoc.2025.113267.
14. Bronchal, L.: MBTI Dataset. Kaggle (2018). Accessed: Oct. 21, 2025. URL: https://www.kaggle.com/datasets/datasnaek/mbti-type.
15. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I.: Language models are unsupervised multitask learners. In: Computer Science, Linguistics (2019). https://api.semanticscholar.org/CorpusID:160025533.
16. Holtzman, A., Buys, J., Du, L., Forbes, M., Choi, Y.: The curious

case of neural text degeneration. arXiv preprint arXiv:1904.09751 (2019). https://arxiv.org/abs/1904.09751.
17. Zhang, T., Kishore, V., Wu, F., Weinberger, K.Q., Artzi, Y.: Bertscore: Evaluating text generation with Bert. arXiv (2019). https://arxiv.org/abs/1904.09675.
18. C. Yang and C. Ding, "Learning word embedding with better distance weighting and window size scheduling," 2024, https://arxiv.org/abs/2404.14631.
19. Qaiser, S., Ali, R.: Text mining: Use of tf-idf to examine the relevance of words to documents. International Journal of Computer Applications 181 (2018) DOI: https://doi.org/10.5120/ijca2018917395.
20. Singgalen, Yerik. (2024). Implementation of Global Vectors for Word Representation (GloVe) Model and Social Network Analysis through Wonderland Indonesia Content Reviews. Jurnal Sistem Komputer dan Informatika (JSON). 5. 559−569. DOI: https://doi.org/10.30865/json.v5i3.7569
21. Park, J.S., Kim, J.: MBTI personality type prediction model using WWT analysis based on the CNN ensemble and GAN. Human-Centric Computing and Information Sciences 13-14 (2023) DOI: https://doi.org/10.22967/HCIS.2023.13.014.
22. Zhang, Jingsi & Yu, Xiaosheng & Lei, Xiaoliang & Wu, Chengdong. (2022). A novel deep LeNet-5 convolutional neural network model for image recognition. Computer Science and Information Systems. 19. 36-36. DOI: https://doi.org/10.2298/CSIS220120036Z,
23. Guo, Y., Li, Y., Wang, L., Rosing, T.: Depthwise convolution is all you need for learning multiple visual domains. In: Proceedings of the AAAI Conference on Artificial Intelligence. AAAI'19/IAAI'19/EAAI'19. AAAI Press, (2019). DOI: https://doi.org/10.1609/aaai.v33i01.33018368.
24. Kosan, M.A., Karacan, H., Urgen, B.A.: Predicting personality traits with semantic structures and LSTM-based neural networks. Alexandria Engineering Journal 61(10),8007–8025 (2022) DOI: https://doi.org/10.1016/j.aej.2022.01.050.
25. Kaiser, L., Gomez, A., Chollet, F.: Depthwise separable convolutions for neural machine translation. ArXiv abs/1706.03059 (2017) DOI: https://doi.org/10.48550/arXiv.1706.03059.
26. Ngartera, L. and Diallo, C. (2024) A Comparative Study of Optimisation Techniques on the Rosenbrock Function. Open Journal of Optimization, 13, 51-63. DOI: https://doi.org/10.4236/ojop.2024.133004.

AUTHOR'S PROFILE

Devraj Patel received his MTech in Computer Science and Engineering (CSE) from the National Institute of Technology, Surathkal, Karnataka, in 2018. He is a research scholar in the Department of Computer Science & Engineering at the DIAT, Pune. His areas of interest include automatic personality recognition, machine learning, and natural language processing. He conceptualised the proposed work in this study and implemented the technique in Python libraries.

Dr. Sunita Vikrant Dhavale received her M.E. (CS) from VIT, Pune, in 2009 and her PhD in Computer Science from the Defence Institute of Advanced Technology, Pune, in 2015. She is working as an Associate Professor at the Defence Institute of Advanced Technology, Pune. Her areas of interest are computer vision, deep learning, and cybersecurity. In this study, her guidance facilitated the integration of novelty into the proposed system, making a significant contribution to this paper. We implemented this work under her valuable direction.

Dr Bhushan B. Mhetre received his MD (Psychiatry) from the Regional Institute of Medical Sciences, Imphal, Manipur. He works as an Associate Professor and Post Graduate guide at Smt. Kashibai Navale Medical College, Narhe, Pune. His areas of interest are Addiction Psychiatry, Personality Disorders, and First Episode Psychosis. He has annotated and validated the collection of real-time data for personality prediction.