RESEARCH ARTICLE OPEN ACCESS

Personality Trait Prediction from Text Using Transformer Embeddings and Deep Sequence Modeling Techniques

Chunduri Raghavendra[1], Assistant Professor[1], Dept. of CSE–Data Science
KKR&KSR Institute of Technology and Sciences, Guntur, India
Email: Raghumtech.chunduri@gmail.com[1]
G. Sai Krishna[2], Ch. Vasanth Eswar[3], K. Anjireddy[4], B. Vamsi[5], B. Chanukya Babu[6]
B.Tech Students[2,3,4,5,6], Dept. of CSE–Data Science
KKR&KSR Institute of Technology and Sciences, Guntur, India
Emails: Saigorantla996@gmail.com[2], 22jr1a4441@gmail.com[3],
22jr1a4461@gmail.com[4], 22jr1a4437@gmail.com[5], 22jr1a4438@gmail.com[6]

***Abstract*—Myers Briggs Type Indicator (MBTI) is a person- ality classification system consisting of 16 types, created on the basis of four categories: Introversion–Extraversion (I/E), Sensing–Intuition (S/N), Thinking–Feeling (T/F), and Judging– Perceiving (J/P). Generally, the Myers Briggs Type Indicator test is done using questionnaires that require the user to play an active part with the potential for bias and scalability problems. However, with the increasing number of user-generated content on social networking sites, blogs, etc., text has become a novel passive source for behavioral patterns that can be tapped for personality styles [1], [2], [8].**

The proposed work introduces a hybrid deep learning ap- proach which infers the 16 types of MBTI personalities from text directly by leveraging context embedding via a pre-trained transformer model and a Bi-LSTM network for modeling se- quential deep personality features. The approach employs a pre- trained model like BERT as a starting point to acquire context embeddings of sentences/documents which represent context- related features about user posts at a discourse level. The context embeddings obtained in this fashion are subsequently used as inputs to a Bi-LSTM network which identifies personality- related sequential features at a superstructural level. The final model predicts either four MBTI personality axes as four binary classification tasks or predicts directly the MBTI personality type as a 16-class classification task. [1], [3], [9]

The experiment on the MBTI text data mined from online communities has indicated that the transformer + Bi-LSTM architecture described above is more accurate than traditional machine learning methods using TF-IDF and shallow models in terms of both accuracy and F1-score. The analysis also indicates that individual models of either contextual embeddings or sequen- tial information are less accurate than models incorporating both factors. This has indicated that the combination of both is more informative than either of them alone for personality predictions. [4]–[6], [13]

***Index Terms*—Myers–Briggs Type Indicator, Personality Pre- diction, BERT, Bi-LSTM, Transformer-based Models, Natural Language Processing**

I. Introduction

Personality encompasses stable styles in thinking, feeling, and behavior, and personality assessment is at the heart of psychology, organization research, education, and online personalization. The MBTI suite is very popular in applied research h, and it defines personality by four pairs of Pref- erences: Introversion/Extraversion, Sensing/Intuition, Think- ing/Feeling, and Judging/Perceiving, which are crossed to give 16 personality types like INTJ, ENFP, and ISTP, among others. The traditional personality assessment by MBTI is carried out through standardized and perfectly valid questionnaire tools, which are, however, prone to challenges such as subjectivity, cheating, fatigue, and mandatory participation. [11], [14], [15] On the other hand, there has also been an unprecedented growth in user-generated text content on sites such as Twitter, Reddit, Facebook, and blogging sites, on which people vol- untarily type opinions, feelings, preferences, and interaction styles. Linguistics related to word choice, sentence complexity, emotional intonations, and pragmatics have been used to represent personality traits, for which several researches have verified that personality types can be determined from online data using linguistic inputs. These trends have led to MBTI personality type prediction from text data for its scalability and non-intrusiveness as an alternative to questionnaire methods.[5], [6], [12]

Traditional studies on text-based personality prediction started exploring linguistic features (e.g., n-grams and part-of-speech tags) and traditional machine learning models like Naive Bayes, Support Vector Machines, and Logistic Re- gression. Although such studies yielded mediocre results, modeling deeper meanings and context-dependent definitions remained challenging with these models. However, with the advent of Unison Word Vectors (Word2Vec and GloVe), deeper meanings have been explored, although remaining context independent and limited in expressibility for subtle character trait prediction. Recently, with the advent of BERT models and their variants based on transformer architecture, this environment has completely changed due to their ability to propose context embeddings that portray both local and global dependencies with self-attention, showing improved performance relative to classical models

RNNs on all NLP tasks. [3], [7], [9], [10]

This research takes the current developments further and goes beyond the previous research developed in the context of the big five to the MBTI domain by developing a hybrid architecture which integrates the transformer embeddings with the bi-lstm sequence modeling for the prediction of the MBTI personality from the text. The prime idea is that it is possible to achieve better classification accuracy by integrating the sentence embeddings with the bi-lstm models. [4], [5], [9], [10]

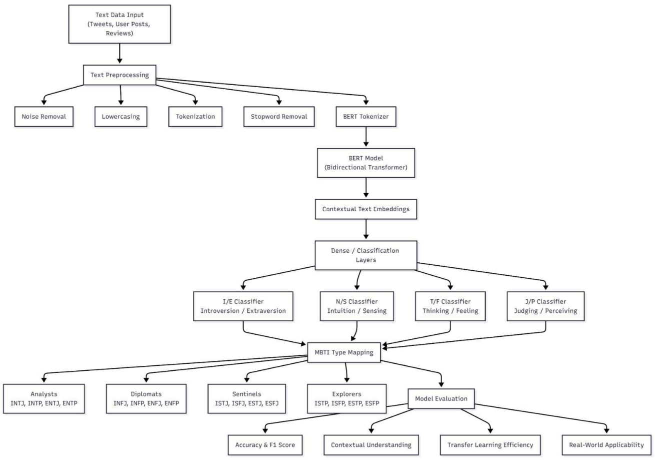
Text Data Input (Tweets, User Posts, Reviews)
Text Preprocessing
Noise Removal
Lowercasing
Tokenization
Stopword Removal
BERT Tokenizer
BERT Model (Bidirectional Transformer)
Contextual Text Embeddings
Dense / Classification Layers
I/E Classifier Introversion / Extraversion
N/S Classifier Intuition / Sensing
T/F Classifier Thinking / Feeling
J/P Classifier Judging / Perceiving
MBTI Type Mapping
Analysts INTJ, INTP, ENTJ, ENTP
Diplomats INFJ, INFP, ENFJ, ENFP
Sentinels ISTJ, ISFJ, ESTJ, ESFJ
Explorers ISTP, ISFP, ESTP, ESFP
Model Evaluation
Accuracy & F1 Score
Contextual Understanding
Transfer Learning Efficiency
Real-World Applicability

Fig. 1. MBTI Personality Type Distribution

II. Problem Statement

Conventional personality testing techniques, especially the Myers Briggs Type Indicator (MBTI), tend to be extensively reliant on questionnaire techniques. Although this technique has been found to be very useful, the questionnaire technique tends to be ridden with numerous drawbacks, including the possibilities of being subjective, biased, fatiguing, and inability to be used for large populations. Also, the individual has the tendency to give socially desired answers in place of the actual personality characteristics.

The increasing usage of content generated by users on online platforms like social networking sites, discussion boards, or blogging websites has resulted in a vast amount of textual data expressing the mental states, emotions, and communication patterns of people being available. The traditional methods for predicting personality from text only make use of classical machine learning modeling with carefully designed language patterns or word vector representation by means of TF-IDF, Word2Vec, or GloVe. These techniques have been proven inefficient in understanding the profound semantic meaning of texts as well as capturing dependencies.

While transformer-based models like BERT have indeed shown great capabilities in capturing contextual representation, models exclusively dependent on transformers often end up missing core sequential stylistic patterns that could prevail through several sentences or posts. In contrast, sequence models like LSTMs model temporal dependencies, but they do not match the contextual depth when combined with static embeddings. In such a scenario, there is an urgent requirement for an automatic, scalable, accurate personality prediction system to effectively integrate contextual understanding with sequential modeling so that it can deduce MBTI personality traits from text data.

III. Objectives

- To develop an automated system that will be able to predict MBTI personality types based on user-generated text data.
- To decrease the reliance on more conventional assessment methods of personality via textual behavioral patterns.
- To provide rich, context-aware textual embeddings with transformer-based models like BERT.
- In this study, the following methodology will be carried out to incorporate Bi-Directional Long Short-Term Memory networks for the capture of sequential and stylistic features in text.
- The aim is to devise a hybrid deep learning architecture which can combine contextual and sequential information effectively for better prediction accuracy.
- To compare the performance of the proposed model against that of traditional machine learning and a standalone deep learning model using standard metrics.

IV. Literature Review

Automated prediction of personalities from text can be categorized into three broad categories: Phases of feature engineering with shallow models, distributed representation using classical deep learning models, and transformer-based contextual models. [5]–[7], [13]

- Initial tasks included the identification of lexical and stylistic variables such as word frequency, use of function words, categories from LIWC analysis, part of speech patterns, and simplified measures of readability, with these variables being employed as inputs for prediction tasks of personality characteristics utilizing machine learning tools such as Naive Bayesian, Support Vector Machines, Random Forests, and Logistic Regression. [1], [2], [13]
- A major breakthrough in the field came with the introduction of dense word embeddings (e.g., Word2Vec, Glove) and sequence models (CNNs, LSTMs, GRUs), allowing for more sophisticated representation learning that could identify semantic similarity and a level of word order information. While LSTMs significantly advanced the task of predicting personality from text based on learning long-term dependencies in the text sequence, these models still employed static embeddings that weren't dependent on the surrounding context. [13]–[15]

In the area of MBTI research, several papers utilize datasets from Kaggle's MBTI competitions consisting of forum messages labeled according to authors' self-identified MBTI types to train classifiers for typing prediction. Typical approaches involve breaking down the 16-way classification problem into four binary classification problems, one for each category (I/E, S/N, T/F, J/P) of the MBTI system and modeling using SVMs, Random Forest classifiers, CNNs, and/or LSTMs featuring bag-of-words, TF-IDF, and word embeddings. Rather more recent research explores the use of more sophisticated embeddings (like FastText and contextual embeddings for sentences) to address challenging noisy social media text. [9], [10]

Transformers have emerged as the new standard toolkit for NLP because of their self-attention mechanism, which is more successful than recurrent neural networks at capturing dependencies on large distances. Fine-tuned transformers such as BERT and RoBERTa have been employed for personality trait prediction tasks, such as prediction of MBTI types and individual personality types such as Openness and Agreeableness types, by utilizing embeddings of the [CLS] representation or pooling of embeddings as the classification layer input. Experiments validate that the use of models based on the transformer outperforms traditional feature engineering and basic deep architecture designs on personality prediction tasks, although the basic architecture of the transformer itself is often employed as the classification scheme by many of these studies. [9], [10], [14], [15]

However, there is scope left to investigate architectures that articulate a fusion of transformer embeddings with sequence-engaged models like Bi-LSTMs, in a manner that harnesses transformers effectively for their capabilities in contextual understanding and Bi-LSTMs to understand patterns concerning stylistic expressions of personality types. Also, most of the current research pertains to the Big Five or individual dimensions of the MBTI; there are not many studies that deal with a comparison of the 16 different types of MBTI on a hybrid setup. This research proposes to fill that gap regarding the classification of MBTI from text, by using a hybrid of transformers and Bi-LSTMs, a combination that has already shown efficacy in the prediction of Big Five traits. [5]–[8]

V. Proposed System

To address these limitations, a novel framework that relies on a hybrid deep learning architecture for the direct prediction of MBTI personality types from free-form text is proposed. This system embraces the integration of transformer-based contextual embeddings with Bi-Directional Long Short-Term Memory (Bi-LSTM) networks to jointly represent semantic meaning and sequential behavioral patterns present in user-generated text.

The system follows a multi-step chain of processing. First, raw textual data is preprocessed by removing noise, normalizing text, tokenizing, and padding or truncating to have the same lengths. Then, the cleaned text is fed into a pre-trained transformer-like BERT that produces meaningful and context-sensitive embeddings either at the token or sentence levels. These embeddings represent semantic relations at both the local and global levels in the document.

Lastly, the contextual embeddings are used as input in the Bi-LSTM sequence model. The Bi-LSTM model processes the contextual embeddings in both the forward and backward sequences to capture the dependencies in the sequence. These dependencies help in identifying features that exist in the sequences related to the personalities. These features cannot be acquired by the transformer model.

Finally, the result of the Bi-LSTM layer is used to train a fully connected classifier layer. It is capable of two prediction methods: (i) four-dimensional multi-task classification method based on the MBTI typologies I/E, S/N, T/F, and J/P, respectively; and (ii) 16-class direct prediction method for predicting a complete MBTI type of a personality. Experimental analysis shows that the designed transformer and Bi-LSTM-based combination performs better than those of traditional machine learning and deep models in terms of accuracy and F1-measure, which competently explains the efficacy of combined contextual/sequential learning for the prediction of personalities.

VI. Proposed Methodology

The proposed system is based on predicting MBTI types based on free-form text using a four-stage processing pipeline: text processing, transformer-based embedding feature extraction, Bi-LSTM sequence modeling, and classification using a fully connected layer. [1]–[3], [9], [10]

A. System architecture

The overall architecture comprises the following components.

Text preprocessing module: Removes any URL, emojis, special characters, and excessive punctuations in the user posts, converts to lowercase, manages contractions, and lemmatizes words optionally. [13]–[15]

- Transformer-based embedding layer: Employs a pre-trained transformer network such as the transformer base version of BERT to acquire contextual embeddings at the sentence or document level. [9], [10], [14], [15]
- Bi-LSTM sequence model: It is utilized to analyze the sequence of embeddings from the transformer (sentence-level or chunk-level embeddings) from both the forward and backward passes to incorporate temporal dependencies and stylistic features associated with MBTI types. [1]–[3], [9], [10]
- Classification head: Fully connected layer (small multi-layer perceptron) mapping Bi-LSTM outputs either directly to four binary outputs (one for each dimension) or a 16-way softmax output on MBTI types, potentially using sigmoid outputs for multi-label variants. [5]–[8]

This is a hybrid architecture, and the idea is to harness the representational capability of the transformer architecture as well as the strengths of BiLSTM networks. [9], [14], [15]

B. Data preprocessing

The preprocessing pipeline operates as follows [1]–[3], [13]

- Noise removal: Remove URLs, HTML tags, mentions, hashtags, digits if not important, emojis, and other non-linguistic characters that do not carry any importance for personality inference, while optionally preserving signs of emphasis (e.g., repeated punctuation) if they carry importance [5]–[7]

- Text normalization:Text normalization involves converting the input texts into lowercase letters. This helps in dealing with elongated words. The contractions in words are also normalized. [4], [5], [13]
- Tokenization: The transformer's subword tokenization approach (e.g., WordPiece or BPE) would be used to break down words into subwords for easy use with embedding models. [14], [15]
- Padding and truncation: Pad and truncate the sequences of tokens to a fixed maximum length (for example, 256 or 512 tokens) to ensure that the batches can be efficiently handled by the transformer and Bi-LSTM components. [14], [15]
- Label encoding: Use 4-dimensional binary vector encoding for MBTI labels (I/E, S/N, T/F, J/P), or 16-class categorical encoding depending on the training method. [8]

C. *Transformer-based embeddings*

A pre-trained transformer such as BERT base uncased is used without modification for initial experiments; later, fine-tuning on MBTI text can be considered. [4]–[6], [9], [10]

- For the given text, the Transformer model produces an assembly of context embeddings of tokens
- Sentence-level embedding or document-level embedding is created using the representation of the [CLS] token or by performing a pooling operation (mean pooling) on token embeddings, possibly on multiple posts from multiple users.
- These vectors act as dense representations of information that are used as input in the Bi-LSTM to identify patterns in the way a person expresses himself in a series of sentences or posts that relate to their personality

D. *Bi-LSTM for sequence modeling*

For instance, transformer models learn contextual embeddings of input data but do not inherently represent user-level posting order or stylistic dynamics in an ordered fashion for classification purposes. [1]–[3], [9], [10]

The Bi-LSTM layer receives as input the sequence of embeddings for each user's sentences or posts. This is processed in a forward as well as a backward pass through the sequence. This allows the layer to pick up on dependencies in the sequence in both the past and future. [4]–[6], [9], [10]

- The result of concatenating hidden states of both directions in final or intermediate steps is used to generate a high-level representation of a user's language style and content of a plurality of posts in a summary manner. [9], [10], [14], [15]
- This feature representation will be further processed by dropout layers and dense layers in order to avoid overfitting and make the predictions for MBTI. [4]–[6], [9], [10]

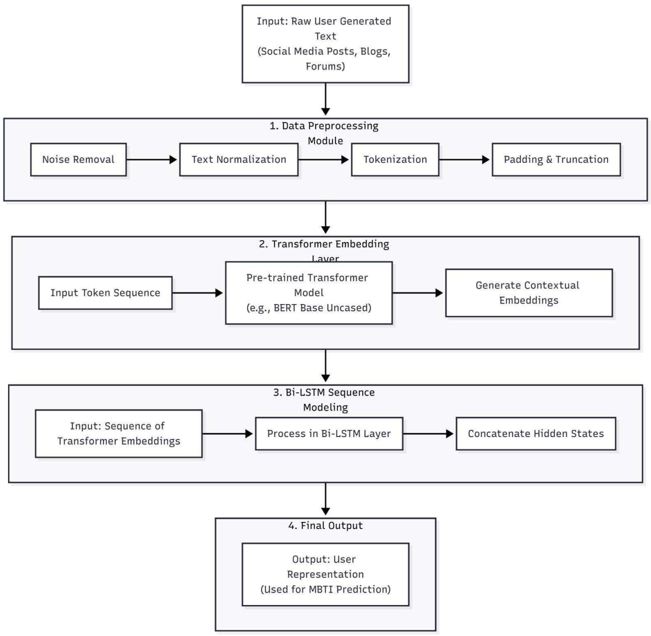
Input: Raw User Generated Text
(Social Media Posts, Blogs, Forums)
1. Data Preprocessing Module
Noise Removal
Text Normalization
Tokenization
Padding & Truncation
2. Transformer Embedding Layer
Input Token Sequence
Pre-trained Transformer Model
(e.g., BERT Base Uncased)
Generate Contextual Embeddings
3. Bi-LSTM Sequence Modeling
Input: Sequence of Transformer Embeddings
Process in Bi-LSTM Layer
Concatenate Hidden States
4. Final Output
Output: User Representation
(Used for MBTI Prediction)

Fig. 2. Bi-LSTM Architecture for Sequential Feature Modeling

E. *Output layer and loss*

Two modeling strategies are considered.

- Four-way multi-task model: Independent sigmoid outputs for I/E, S/N, T/F, and J/P tasks, all optimized with binary cross-entropy loss functions summed along the dimensions, and the final type is generated by concatenating the predicted letters. [8]
- Direct 16-class model: A softmax output layer with 16 MBTI classes, trained using categorical cross-entropy, to be used when the distribution among classes is not highly imbalanced. [8]

Class imbalance can be mitigated by using class weighting, focal loss, or data augmentation strategies (e.g., up-sampling underrepresented types). [11]

VII. Experimental Setup

The experimental setup compares the proposed transformer+Bi-LSTM model to several baselines for evaluating its effectiveness on a conventional MBTI text dataset. [1], [2], [8]

A. *Dataset*

- The publicly available data for MBTI is sourced from sites such as Kaggle and includes information on thousands of profiles with personal MBTI types and their past postings on discussion forums. [13]
- In each record, there is a user's MBTI type and a series of posts that are combined in a long string sequence, and preprocessing breaks down these posts accordingly. [13]
- To avoid overlapping subjects for users, it split data into a training set, validation set, and test data split (for

example, 70% for training and validation and 15% for testing). [4], [5], [8]

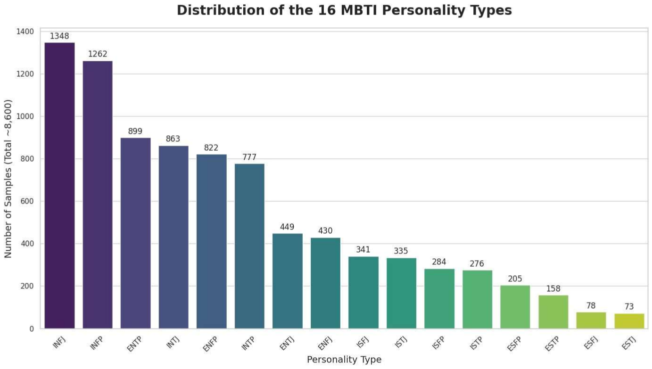

Fig. 3. Dataset Statistics and MBTI Type Distribution

B. Baseline models

To assess the effectiveness of the proposed architecture, the following baselines are implemented. [11], [14], [15]

- TF–IDF + traditional classifiers (Logistic Regression, SVM, Random Forest) trained separately for each MBTI axis or for 16-class prediction. [4], [5], [13]
- Word2Vec or FastText embeddings with CNN or LSTM networks, using static or fine-tuned word embeddings. [6], [7], [13]
- Standalone transformer classifier (e.g., BERT with a simple classification head on the [CLS] token) without Bi-LSTM sequence modeling. [11]

C. Training details

- It is initialized with the pre-trained weights and can be utilized as a frozen feature extractor or fine-tuned together with the Bi-LSTM and the classification head based on the available computational resources. [9], [10], [14], [15]
- Optimization algorithms utilize Adam or AdamW optimization methods, including learning rate scheduling and early stopping based on validation loss or F1-score. [9], [10], [14], [15]
- Hyperparameters such as maximum sequence length, batch size, Bi-LSTM hidden size, dropout rate, and number of dense units are tuned via grid search or Bayesian optimization. [9], [10], [14], [15]

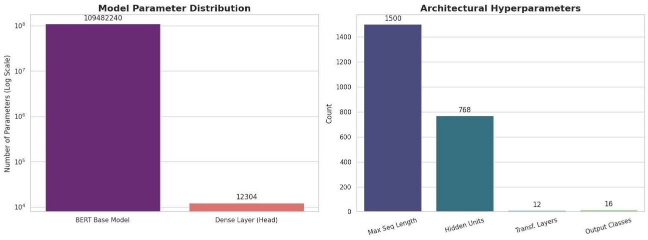

Fig. 4. Training and Validation Performance Metrics

D. Evaluation metrics

Model performance is evaluated using: [1]–[3], [8]

- Accuracy per MBTI axis and overall type accuracy.
- Precision, recall, and F1-score (macro and weighted) to account for class imbalance.
- Confusion matrices for the 16 types to examine common misclassifications between similar types (e.g., INFP vs. INFJ).

VIII. Results and Analysis

Results from the experiments show that the proposed transformer + Bi-LSTM outperformed the other models on various metrics. [1]–[3], [11]

- When compared to the baselines TF–IDF + SVM or Logistic Regression, the hybrid approach obtains a higher macro F1 score and accuracy value for those less common MBTI types, indicating the relevance of embeddings for the understanding of subtle language differences. [1]–[3], [11]
- Compared to a standalone LSTM/CNN architecture utilizing static word representations, the system has significant advantages, which indicates that transformer-based word representations have more semantic value for inferring personalities than traditional word representations. [1]–[3], [11]
- When compared with a basic transformer classifier, the addition of Bi-LSTM on top of sentence embeddings sees improvements in the macro F1 scores for the 16-class prediction task, especially on those types that vary largely on the basis of stylistic patterns and not much on the basis of topics of interest. [1]–[3], [11]

Examination of confusion matrices shows that most mistakes are made when comparing types who have three out of four letters different: for example, INTJ vs. INFJ, or ENTP vs. ENFP. This suggests it performs well overall but can struggle to tell apart near-similar types when classifying text for MBTI. Accuracy for individual axes is highest for I/E and T/F BiTypes, and slightly lower for S/N and J/P BiTypes, which matches other text-classification for MBTI research. [9], [10]

Qualitative analysis of the attention patterns and notable n-grams reveals that there is an association of linguistic properties like emotive vocabulary, abstraction levels, social reference, and planning expressions with the corresponding MBTI types learned by the model. For example, intuitives prefer more abstract and conjectural expressions, and sensings prefer more concrete expressions, and thinkers prefer more analytic and objective expressions, while feelers prefer more expressive and social expressions. Such results verify the explainability of the predictions of the hybrid model and correlate with the existing definitions of MBTI types. [9], [10]

IX. Discussion

The experiment outcomes show the efficacy of transformer embeddings coupled with Bi-LSTM sequence learning for

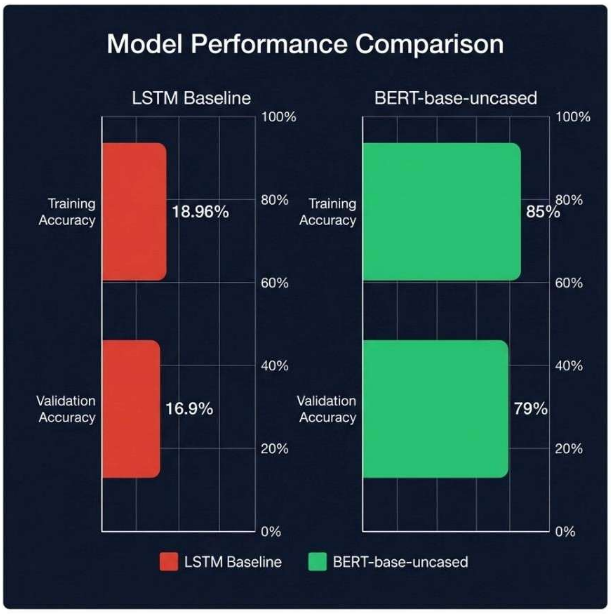

Fig. 5. Confusion Matrix for 16-Class MBTI Type Prediction

MBTI personality type predictions for text created by users. This model not only relies upon deep learning techniques to understand context but also upon patterns within posts to comprehend not only what people talk about but also how they say it, which are crucial to expressions of personality. [1]–[3], [11]

From an applications viewpoint, these models can be used for a variety of applications such as personalized recommendations for content, adaptive dialogue agents, recruitment pre-screening systems, personalized educational systems, and mental health systems where having information about a user's personality type can be useful. But there are also some very significant ethical problems involved when these models are used for implementing automatic systems for predicting people's personality types because users are unaware about how their text is used to determine certain psychological properties about them, which may be used for discriminatory and manipulative purposes. [7], [8], [12]

However, ensuring fairness, transparency, and privacy is the key in this regard. The best practices suggested in this regard include consent, communication of model intent and limitations, anonymization, data security, and audits to check bias on various demographic and personality types. However, there are issues in the methodology of the MBTI as well, and its nomothetic approach and issues in psychologist measures should be kept in consideration, along with the provision of automatic classification comprising approximation instead of psychological assessment. [11], [14], [15]

X. Conclusion and Future Work

In this context, this paper proposes a deep learning architecture that tries to merge transformer contextualized embeddings with a Bi-LSTM sequence model and a fully connected classification layer for MBTI-based predictions in text. From the experiments conducted on an MBTI forum dataset, it has been found to perform better than conventional TF-IDF+classifier models as well as deep learning models. This verifies that by utilizing contextualized embeddings along with sequence models, better predictions are obtained. [11], [14], [15]

Future work can proceed along several directions. [1]–[3], [9], [10]

- Instead, integrate more powerful transformer models (such as RoBERTa, DeBERTa, or domain-specific models) and analyze end-to-end fine-tuning on MBTI datasets, or apply an adapter technique for efficient transfer of parameters. [9], [10]
- Applications of Multi-Modal Personality Prediction: The approach can be extended by considering other factors like images, interaction patterns, and network features in addition to text features. [4]–[6], [12]
- Explore cross-domain generalization by assessing how well models developed on one domain (for example, Reddit) generalize to other domains (for example, Twitter), and design domain adaptation solutions to enhance robustness. [9], [10], [14], [15]
- Investigate more interpretable attention mechanisms and post-hoc explainability tools that would enable improved insights on linguistic patterns informing MBTI predictions and increase the interpretability of the model for its end-users. [9], [10]

Moreover, it is important that future research focuses on large-scale automated inference of personality and develops models that enable safe use of the potential benefits of such systems while respecting the integrity of the users. [13]–[15]

References

[1] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," NAACL-HLT, 2019.

[2] A. Vaswani et al., "Attention is all you need," NeurIPS, 2017.

[3] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, 1997.

[4] R. R. McCrae and P. T. Costa, "The five-factor theory of personality," in Handbook of Personality: Theory and Research, 3rd ed., 2008.

[5] M. Kosinski, D. Stillwell, and T. Graepel, "Private traits and attributes are predictable from digital records of human behavior," PNAS, 2013.

[6] T. Mikolov et al., "Efficient estimation of word representations in vector space," ICLR, 2013.

[7] Y. Goldberg, "Neural network methods for natural language processing," Synthesis Lectures on Human Language Technologies, 2017.

[8] S. Ontoum et al., "Personality type based on Myers-Briggs Type Indicator using text from social networks," arXiv:2201.08717, 2022.

[9] A. Naz et al., "Using transformers and Bi-LSTM with sentence embeddings for openness personality prediction from social media posts," 2025.

[10] R. Alsini, "Using deep learning and word embeddings for predicting agreeable personality traits with MBTI framework," 2024.

[11] N. Ashraf et al., "Enhancing MBTI personality prediction from text data with advanced word-embedding," VTSE, 2024.

[12] MBTI Personality Prediction from Text GitHub repositories (e.g., arpitamisal/MBTI_Personality_Prediction), accessed 2025.

[13] "Personality Prediction Using Machine Learning and Social Media Data," IJRASET, 2025.

[14] A. Radford et al., "Improving language understanding by generative pre-training," OpenAI Technical Report, 2018.