P-ISSN: 2709-6254	Journal of Development and Social Sciences	July-September 2025, Vol. 6, No. 3
O-ISSN:2709-6262	https://doi.org/10.47205/jdss.2025(6-III)57	[664-676]

Journal of Development and Social Sciences
www.jdss.org.pk

RESEARCH PAPER

Text-Based Personality Recognition Based on User Content

[1]Muhammad Asfand E Yar, [2]Mujahid Ahmad and [3]Qadeer Hashir

1. Associate Professor, Head Enabling Technology, Centre of Excellence Artificial Intelligence, Dept. of Computer Science, Bahria University, Islamabad, Pakistan
2. Student, Department of Computer Science, Bahria University, Islamabad, Pakistan.
3. Student, Centre of Excellence Artificial Intelligence, Dept of Computer Science, Bahria University, Islamabad, Pakistan

Corresponding Author: asfandeyar.buic@bahria.edu.pk

ABSTRACT

Personality recognition from textual data, a topic of growing interest, has gained significant importance in the fields of psychology, marketing, and human-computer interaction. This research explores the domain of text-based personality recognition, focusing on user-generated content to uncover the complex aspects of an individual's personality traits. It leverages state-of-the-art transformer-based models, including BigBird, Albert, and DistilBERT, enhanced with NLP statistical features. The primary objective is to evaluate and compare these cutting-edge models' performance comprehensively, concatenated with NLP statistical features, against conventional methods for personality trait recognition across diverse textual datasets, including Facebook and essay datasets. This research employs two classifiers, BiGRU and BiLSTM, to classify the five personality traits of the Big Five personality trait model using Facebook and essay datasets. BigBird, when combined with NLP statistical features and using BiLSTM as a classifier, achieves impressive F1 scores for traits EXT, NEU, AGR, CON, and OPN, demonstrating accuracies that underline the effectiveness of this approach. The findings show that pre-trained models in combination with NLP statistical features have improved the performance of the Personality recognition model in terms of accuracy and F1-score across the myPersonality datasets.

KEYWORDS BigBird, DistilBERT, Albert, BiGRU, BiLSTM, Big Five, Personality Recognition

Introduction

Personality includes different patterns of thoughts, emotions, and behaviour. Distinguishing individual's personality and structuring their relationships in society is difficult. Personalities have far-reaching consequences for decision-making, relationships, and overall well-being. In today's interconnected world, social media and the internet made possible effective cross-cultural communication, so by exploring individual communication it is possible to predict an individual's personality. Personality recognition is important in a variety of contexts, like commercial interaction, individual interactions, and recruitment processes.

Automatic personality recognition is an emerging field in the domain of Natural Language Processing (NLP). The approach of Personality trait identification from various text sources, such as tweets, Facebook postings, blogs, and essays, has gained the interest of researchers. Deep learning and machine learning techniques have been used to improve system performance in personality prediction. Notably, algorithms like XGBoost have improved the performance of the Big Five personality prediction models (Tadesse, M. M., et al., 2018), while pre-trained transformers like BERT, XLNet, and Roberta have shown promise in personality detection (Christian, H. et al., 2021). Deep learning models have also contributed to improved personality prediction accuracy (Yuan, C., et al., 2018), (Tandera, T., et al. 2017), (Zheng, H., et al. 2019).

We use the power of pre-trained transformers, notably BigBird, Albert, and DistilBERT, in conjunction with NLP features such as TF-IGM, Sentiment Analysis, Emotion-

Based Features (NRC Lexicon), and Linguistic and Textual Attributes in this study. The extraction of features is our key focus. Models such as BiLSTM and BiGRU are used for classification to aid with personality recognition. The overarching goal is to improve the Big Five personality prediction model's performance by evaluating it against prior state-of-the-art approaches.

Literature Review

BERT-based emotion recognition studies assess research goals, methods, and results, and provide insights into textual emotion recognition by Transformer models (Acheampong et al., 2021). Al.. Studies on personality prediction from social media posts classify the methods as linguistic, interaction-based, and network-based strategies and offer indications for improving the prediction accuracy (Aung, Z. M. M., 2019). Supervised machine learning algorithms were used to predict users' personality traits based on the entered text (Das, K. A. H., et al., 2022), (William, P., , et al. 2022), (Bruno, A., 2022). XGBoost was used to categorize personality traits from user text. The methodology included data collection, resampling, pre-processing, feature selection, and classification using the MBTI model. A performance comparison was performed between XGBoost and other classifiers, supported by various evaluation metrics (Cherukuru, R. K., et al., 2022). Research into personality recognition through deep learning techniques has led to various methodologies and approaches and contributed to the advancement of the field. A personality recognition model was developed using deep learning techniques (Yu, J., et al., 2017), (Yuan, C., et al., 2018), (Xue, X., et al. 2021). A personality profile was created using Computational Psychology by assessment on the MBTI scale. The pre-processing involved deleting hyperlinks, numbers, and punctuations as well as using derivation tools (WordNet Lemmatizer and Lancaster Stemmer). The feature vector was formed by the combination of TF-IDF, EmoSenticNet (10 emotions), LIWC, and ConceptNet (300 floating point numbers) features. These three classifiers trained and tested on the MBTI dataset using a 70:30 split ratio are the SVM, Neural Networks and Naive Bayes. The results showed that the highest accuracy was 86.27 percent with SVM (Bharadwaj, S., 2018). In order to implicate personality depending on the existence of text and emoji information, researchers integrate emojis in personality recognition frameworks with the assistance of bidirectional long-term and short-term memory (BiLSTM) and focus (Zhou, L., et al. 2022) A deep learning hybrid model that tries to categorize text by certain personality features. The text was subjected to tokenization, stop word stripping and lowercase. The architecture was a combination of the DNN-CNN+LSTM model, a word representation-processing layer through the embedding layer, a feature-processing layer through CNN, a long-term information learning layer through LSTM and a classification layer through SoftMax (Ahmad, H et al., 2021). To investigate the possibility of using the input of filters of different lengths to identify personality, a distinctive method that uses CNN and the AdaBoost methodology was considered. The YouTube personalities and stream-of-consciousness studies were used as datasets. Local features were extracted using word embedding according to the Skip-Gram model. The use of AdaBoost to scale the classifier using varying n-grams reflects the importance of the pooling and the dropping strategies (Sadr, H., et al. 2016). Multi-model deep learning is a set of architecture that combines NLP functionality with pre-trained transformers, including BERT, RoBERTa, and XLNet. Preprocessing includes removing URLs, symbols, and emoticons, followed by English translation, lowercase, contraction expansion, stop word removal, and derivation. Feature extraction uses techniques such as word piece tokenization, token embedding, segment embedding, and position embedding, using CLS and SEP tokens to enhance the contextual meaning of the. This approach produces the best results for all major personality traits, including openness (70,85). %), conscientiousness (88.85%).49%), extraversion (81.17%), agreeableness (69.33%), and neuroticism (75.08%) (Christian, H., et al., 2021). The use of state-of-the-art DL-based NLP models meets the challenge of identifying and categorizing personality types using different fonts and text styles. Two datasets are indexed: MyPersonality (Facebook) and Essays (Penne and King). The research article proposes data-level and classifier-level fusion strategies to improve personality prediction performance. Pre-trained language models (Elmo, ULMFiT, BERT) are adopted, and combining the Essays and MyPersonality datasets further improves the proposed model (El-Demerdash, K., et al., 2022). It is not novel to predict personality traits using data from Facebook and Twitter. For example, (Tandera, T., et al. 2017)used an open-source Facebook personality dataset called MyPersonality, which contains 250 users' status data and attributes and maps to the huge five-personality model. The main feature extraction method is Linguistic Inquiry and Word Count (LIWC), which is a linguistic analytical tool that aids in the analysis of quantitative texts and provides a calculation number of words that have the meaning of categories based on a psychological dictionary.

Material and Methods

We turn to two different datasets: The essay dataset (Pennebaker, J. W., et al., 2007) and the myPersonality dataset (Stillwell, D. J., et al. 2015). A pre-processing step is used in each dataset to provide data quality and consistency. The two datasets undergo several pre-processing steps to ensure that the textual contents become available to further analysis. The pre-processing stage helps in cleaning, standardizing, and normalizing the text data, therefore, something to build on.

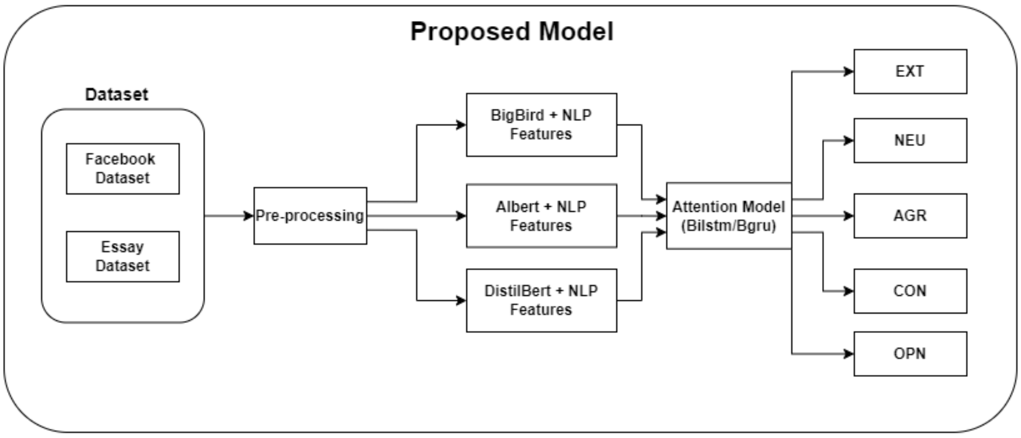

Fig. 1: Framework Architecture of Text-based Personality Recognition

We use the advantages of higher-order language model, i.e., BigBird, ALBERT and DistilBERT to extract detailed semantic elements of text (Zaheer, M., et al., 2020) (Lan, Z., et al. 2020), (Sanh, V., 2019), and extract elements with the help of NLP statistics. The two-step feature extraction approach summarizes the natural linguistic complexities in the data and provides a full-scale image to be further analyzed. Each dataset contains a designated set of featured data which is fed into the model to perform classification tasks. This step includes the implementation of the BiLSTM models, which were also used in the study (Zhou, L., et al., 2022) and the BiGRU. These models have been carefully designed to capture complex dependencies and relationships within textual data.

Dataset

Two different datasets were used to perform the analysis presented in this article, each providing unique information about the relationship between textual data and personality traits. The first dataset, dubbed the “myPersonality dataset”, includes a total of 250 with 9917 individual posts Facebook users. This dataset comes from the myPersonality Project (Stillwell, D. J., et al. 2015), the second dataset used in this study is the Essay dataset, which serves as the established benchmark in this field (Pennebaker, J. W., et al., 1999). Curated by Pennebaker and Laura King, this data set consists of a large corpus of text written by 2,467 people between 1997 and 2004.

Preprocessing

This stage involves a series of steps to prepare the textual data for further analysis. The figure 2 shows illustration of each step in the pre-processing. In the text contractions were extended, hyperlinks were removed (for the essay dataset this step was not included), numerical values were removed, and the text changed to lowercase. The next steps included the elimination of symbols and special characters, tokenization into individual words, word normalization using stemming, and the omission of frequent stopwords while keeping personal pronouns for context. The words were then put back together into intelligible text strings after processing. The homogeneity, readability, and relevancy of the textual content were all improved by these extensive pre-processing efforts, creating a refined framework for in-depth analytical investigation inside both datasets.

Features Extraction

NLP Statistical Feature

Analysis of personality traits from textual data involves the extraction and use of various statistical characteristics. The table. 1 represents NLP statistical features used in the study. These features provide valuable information about the various linguistic, emotional, and psychological dimensions inherent in the textual content.

In this study, we used TF-IGM measurements instead of TF-IDF which is also used in studies (Christian, H., et al. 2021) for Personality traits classification on the Big Five model, resulting in higher performance. TF-IGM is an alternative to TF-IDF for text classification tasks as it gives high performance compared to the traditional method TF-IDF (Chin, K., 2016). We are taking the top 60 words score of TF-IGM as features. Five features of sentiment analysis are used, as they provide information about a person's emotional state, opinions, and perspectives expressed through texts (Pak, A., et al., 2010). These characteristics have assisted in exploring the relationships existing between emotions and other personality traits. The lexicon of NRC gives us the facts concerning the psychological composition of a person and his/her emotional inclinations (Mohammad, S. M., et al, 2013). The NRC Lexicon contains 8 features that reflect various emotion in the text. The characteristics illustrate the multidimensionality of the interrelationship between emotions and personality traits and introduce a new dimension of linguistic expression and emotional disposition. In addition to the above key characteristics, some other statistical characteristics, which elaborate on the personality test even more, provide more context and meaning to the analysis: The complexity and readability of text can be quantified using metrics such as the Flesch Reading Ease and Gunning Fog Score (Flesch, R., et al. 1948). This highlights the simplicity of understanding the contents by different audience. The frequency of personal pronouns (I, you, he, she, we and they) is a measure of self-referential tendencies and interpersonal styles of communication. The richness and complexity of the vocabulary and evidence of linguistic diversity and expression sophistication are presented in the attributes word variety and average word length. The social behavior count, a count of words related to social interactions and relationships, reveals a person's social behavior and provides information about possible associations with personality traits. Various count-based metrics, such as the number of capital letters, capital letters, repeated words, and the occurrence of proper nouns (PROPNAME), provide insight into different writing styles and patterns.

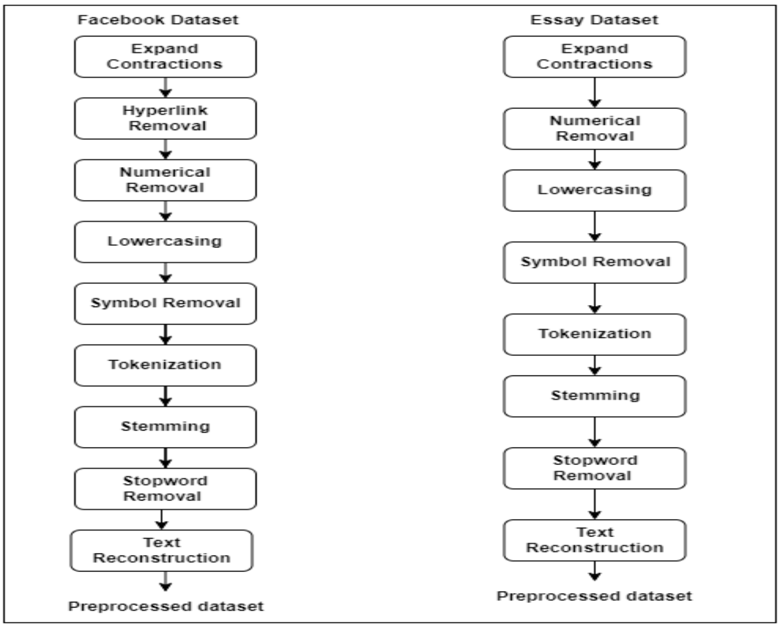

Figure 2. Pre-Processing of data

Table 1
NLP Statistical Features

Feature Name	Description	Feature Count
TF-IGM	Statistical method to find how important a word is in a document influenced by the class label of a document. This method is used based on the research performance comparison between TF-IDF and TF-IGM in text classification (Chin, K., 2016).	60
Sentiment Analysis	The sentiment analysis features include sentiment polarity, sentiment subjectivity, positive percentage, negative percentage, and neutral percentage. The researcher used a polarity sentiment analysis approach (Pak, A., et al., 2010) to extract these features.	5
Emotion-Based Features (NRC Lexicon)	Contains 14,000 sets of words in English and the relation of each word with eight common emotions: anger, fear, anticipation, trust, surprise, sadness, joy, and disgust (Mohammad, S. M., , et al 2013)	8
Linguistic and Textual Attributes	This category encompasses various linguistic and textual attributes such as readability scores, pronoun usage (first-person, second-person, third-person pronouns), word diversity, average word length, word count, character count, and counts related to social behavior, capitalization, repeated words, and occurrences of proper nouns (PROPNAME)	17
Total Statistical Features		90

Pre-trained Feature Extraction

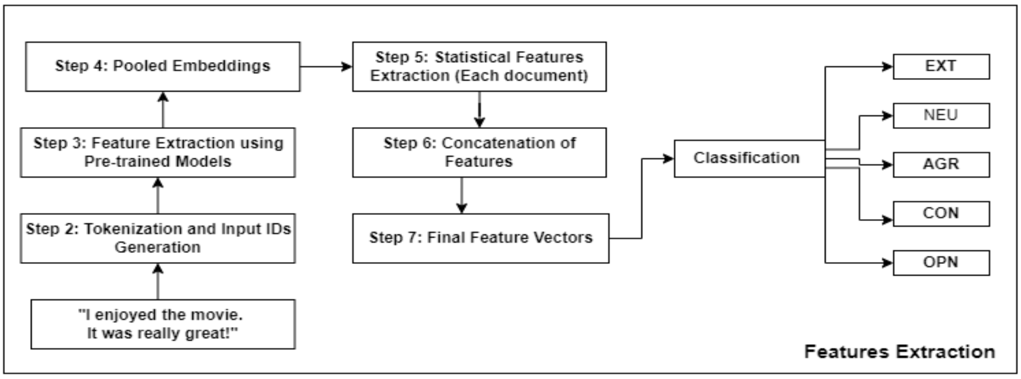

Figure 3: Training model for feature Extraction.

This research uses three pre-trained models namely BigBird, ALBERT, and DistilBERT (Zaheer, M., et al. 2020),(Lan, Z., et al. 2020), (Sanh, V., et al., 2019) to address the shortcomings of conventional models like BERT, RoBERTa, and XLNet by utilizing the distinct characteristics of each model (Christian, H., et al., 2021). Figure 3 represents the feature extraction process from the pre-trained model and the concatenation of statistical features with the pooled embedding. These models, each with a unique architecture that enables them to handle various linguistic patterns, were trained on enormous text collections. They can effectively catch details of language expressions due to their built-in systems. The method of feature extraction entails converting the text into numerical representations that capture both the linguistic characteristics and the context of the words. This process includes tokenizing WordPiece tokenization, input IDs generation, contextual embedding generation through self-attention processes, pooling embedding generation, and classification of features as a form of feature vectors. The text is tokenized using WordPiece tokenization which divides the text into subwords or tokens. These tokens are translated into input IDs to generate the input serving as a tensor. Then the contextual embeddings (also called self-attention embeddings) are extracted using self-attention mechanisms. These techniques form the contextual embedding matrix and enable each token to consider its relations with all other tokens. Contextual embeddings matrix is averaged to generate pooled embeddings that provide a concise description of the meaning of the entire text. We obtain the embeddings by combining the pooled embedding with the statistically extracted features after extracting pooled embedding. This combined vector is the feature vector of each classification model of the Big Five personality traits that comprise the language and contextual understanding characteristics.

Model Prediction

The second task after the extraction of feature vectors was to categorize those feature vectors into the predefined categories of personality traits of the Big Five personality model (EXT, NEU, AGR, CON, OPN). To do this, two different recurrent neural network (RNN) architectures were adopted BiGRU and BiLSTM. GRU architecture of the BiGRU paradigm is two directional. It performs based on the forward and reverse examination of feature vectors. This two-way process enhances the contextual understanding of the model on the relationships that exist in the feature vectors. The BiLSTM model consists of Long Short-Term Memory (LSTM) units. These units are the best at capturing long-term dependencies in sequential data, and are the units that are best used when complex time connections are required. Similar to the BiGRU model, the BiLSTM model combines both forward and backward LSTM predictions using a linear layer to obtain the prediction. In training, the computation of cross-entropy is done to calculate the difference between the predicted and actual value (ground truth). The parameter tuning process will be implemented to get the best out of the model that is generated. A grid search method will be used to perform a series of searches in order to identify optimal parameters that will produce the most desirable prediction performance. Some of the parameters that can be varied include the batch size, epoch, number of hidden layers, number of layers and learning rate.

Evaluation Matrix

An extensive assessment technique was conducted to test comprehensively the performance of our categorization models. Accuracy and F1-score are some of the important criteria that we used to evaluate the capacity of each model to predict personality traits. These actions gave an overall view of how the models were capable of classifying data with high reliability. A categorization report was provided to each model and personality trait category which was an important part of the evaluation process. This report was very insightful on the performance of the models across different features, and we were able to identify areas that could be improved.

Experiment

We not only scraped NLP statistical features but also applied three transformer-based feature extraction models (BigBird, ALBERT, and DistilBERT). Large-scale text datasets were used to train each model on word and sentence representations in context. Two recurrent neural networks (RNN) architectures BiGRU and BiLSTM were used to test personality traits classification. These models were chosen because they perform well in sequential data analysis tasks. To identify the optimal configuration for each model, hyperparameters were tuned using grid search. We tested the model on three different hidden sizes: 64, 128, and 256, three different numbers of layers: 2, 3, and 4, two batch sizes: 16 and 32, and two learning rates: 0.001 and 0.0001, Also used test the model on different epochs 5, 10 and 15.

Model weights were set up according to best practices for each architecture. DataLoader instances were used to load the training and validation sets. We used the Cross-Entropy Loss, which is appropriate for multi-class classification jobs. The model was trained over numerous epochs for each combination of hyperparameters. The model with the highest validation accuracy was chosen as the best. To evaluate model performance, we used the evaluation measures (Accuracy and F1-score).

Result and Discussion

In this section, we discuss the results of our study on the problem of personality traits classification based on different transformer-based models: BigBird, Albert, and DistilBERT. Another thing that we examine is whether the addition of new NLP statistical variables in them has any impact.

myPersonality Dataset

Table 2 results of the BiLSTM model give some idea of how different transformer model predicts personality characteristics. The Extraversion (EXT) aspect of the BigBird model had an accuracy of 56.69 percent and an F1-score of 64.33 percent. Applying the Albert model gives the accuracy and F1 score of 70.67 and 71.82 respectively. The maximum accuracy and F1-score of DistilBERT were 78.62 and 79.07, respectively. When they were complemented with new NLP statistical variables, these models proved to be significantly more accurate predictors, and the BigBird+NLP and Albert+NLP models achieved 85.16% and 84.50 accuracy, respectively. The findings indicate that transformer models are able to learn complex linguistic regularities about extraversion and that their predictions can be further maximized when other variables are taken into account. Performance of the trait of Neuroticism (NEU) was a stabilized pattern. The model accuracy of BigBird was 70.42, the F1-score was 37.74, and the model accuracy of Albert and DistilBert was 76.04 and 87.39, respectively, and the F1 score was 60.22 and 76.44. The further improvement of NLP statistical characteristics increased the performance, just as the extraversion characteristic, with accuracies of 85.97% and 86.78% in the BigBird+NLP and Albert+NLP models, respectively. This level of accuracy was obtained with the DistilBERT without providing any NLP statistical features. In the case of Agreeableness (AGR) attribute. BigBird, Albert, and DistilBERT got accuracy of 84.04, 83.28 and 86.98 respectively with F1-scores of 18.18, 12.23 and 47.01 respectively. NLP statistical features were added again and the BigBird+NLP model achieved 92.35% accuracy and an F1-score of 74.62. The models had been doing a great job in the board of the Conscientiousness (CON) trait. The BigBird, Albert and DistilBERT models had high accuracies of 95.39, 98.02 and 97.72 with F1-score of 26.02, 79.79 and 77.39 respectively. The BigBird+NLP and Albert+NLP models were able to reach 98.48% and 97.52% accuracies with the NLP statistical features added. At least, the score Openness to Experience (OPN) was also good at prediction. The BigBird model was 96.40% accurate and its F1-score was 32.38%.

the Albert and DistilBERT models achieved accuracies of 97.92 and 98.23 respectively and an F1-score of 69.63 and 81.68. The introduction of NLP statistical features did not significantly affect the model performance with BigBird + NLP model achieving 98.33 % and Albert + NLP model reaching 97.21 %.

Table 2
Facebook Dataset (BiLstm)

Traits	**Metric**	**Bigbird**	**Albert**	**Distilbert**	**Bigbird + NLP statistical features**	**Albert + NLP statistical features**	**Distilbert + NLP statistical features**
EXT	Accuracy	0.5669%	0.7067%	0.7862%	**0.8516%**	0.8450%	0.6084%
	F1-Score	0.6433	0.7182	0.7907	**0.8277**	0.8271	0.6529
NEU	Accuracy	0.7042%	0.7604%	**0.8739%**	0.8597%	0.8678%	0.7351%
	F1-Score	0.3774	0.6022	**0.7644**	0.7834	0.7599	0.5052
AGR	Accuracy	0.8404%	0.8328%	0.8698%	**0.9235%**	0.9063%	0.8323%
	F1-Score	0.1818	0.1223	0.4701	**0.7462**	0.7141	0.0461
CON	Accuracy	0.9539%	0.9802%	0.9772%	**0.9848%**	0.9752%	0.9463%
	F1-Score	0.2602	0.7979	0.7739	**0.8454**	0.8032	0.3977
OPN	Accuracy	0.9640%	0.9792%	0.9823%	**0.9833%**	0.9721%	0.9630%
	F1-Score	0.3238	0.6963	0.8168	**0.8156**	0.7179	0.4823

The BiGRU model generated results presented in Table 3 that were not similar to the BiLSTM results. Extraversion (EXT) trait was 62.41% and 62.83% accuracy and F1-score respectively in the BigBird model. In turn, the Albert and DistilBERT models were more accurate (77.96% and 68.59, respectively). The inclusion of NLP statistical features did not have any visible impact on the performance of the models, and the BigBird+NLP model obtained an accuracy of 75.08%. The same thing was with the Neuroticism (NEU) trait. The accuracy of the BigBird model was 70.97 per cent and F1-score 51.40 per cent, whereas the accuracies of the Albert and DistilBERT models were 76.85 per cent and 76.90 per cent, with F1-scores of 69.51 per cent and 51.90 per cent, respectively. Moderate performance was obtained with the addition of NLP statistical features; the accuracy of the BigBird+NLP model was 76.19 and that of Albert+NLP was 76.85. Once again, the performance of the models regarding the Agreeableness (AGR) trait was constant. The BigBird, Albert, and DistilBERT models had the accuracy of 82.42%, 89.11%, and 52.40%, respectively, and the F1-score of 20.59%, 64.23%, and 52.40%, respectively. Inclusion of NLP statistical features led to insignificant change in performance, and the BigBird + NLP model attained an accuracy of 87.18%.

The BiGRU has done very well with Conscientiousness (CON). The BigBird, Albert, and DistilBERT models have high accuracies of 94.98, 96.91 and 97.32 with F1 of 43.43, 65.14 and 68.26, respectively. NLP statistical features also marginally improved the performance, the BigBird+NLP and Albert+NLP models achieved maximum accuracy of 97.37% and 96.91, respectively. The trait Openness to Experience (OPN) has large values of accuracy and F1-score. The F1-score and the accuracy of the BigBird model were 54.66% and 96.30, respectively, and the F1-score and the accuracy of the Albert and DistilBERT models were 76.47% and 56.06 respectively. Adding NLP statistical features to model performance did not significantly affect the performance, and the BigBird+NLP and Albert+NLP models achieved the accuracy of 98.23% and 97.97, respectively.

Table 3
Facebook Dataset (BiGRU)

Traits	Metric	Bigbird	Albert	Distilbert	Bigbird +NLP statistical features	Albert +NLP statistical features	Distilbert +NLP statistical features
EXT	Accuracy	0.6241%	**0.7796%**	0.6859%	0.7508%	0.7796%	0.6859%
	F1-Score	0.6283	**0.7106**	0.6846	0.7185	0.7106	0.6846
NEU	Accuracy	0.7097%	**0.7685%**	0.7690%	0.7619%	0.7685%	0.7690%

	F1-Score	0.5140	**0.6951**	0.5190	0.6466	0.6951	0.5190
AGR	Accuracy	0.8242%	**0.8911%**	0.8546%	0.8718%	0.8911%	0.8546%
	F1-Score	0.2059	**0.6423**	0.5240	0.5125	0.6423	0.5240
CON	Accuracy	0.9498%	0.9691%	0.9732%	**0.9737%**	0.9691%	0.9732%
	F1-Score	0.4343	0.6514	0.6826	**0.7451**	0.6514	0.6826
OPN	Accuracy	0.9630%	0.9797%	0.9412%	**0.9823%**	0.9797%	0.9412%
	F1-Score	0.5466	0.7647	0.5606	**0.7619**	0.7647	0.5606

Essay dataset

The findings obtained from the essay dataset utilizing the BiLSTM and BiGRU models, along with BigBird, Albert, and DistilBERT embeddings, show a significant difference from the performance found in the myPersonality Facebook dataset. In this case, the prediction of the models was noticeably limited across all features. The Classification Reports show that most personality qualities have low accuracy and F1 scores. When assessing the performance of the BiLSTM model with the integration of BigBird and NLP statistical characteristics, the findings show difficulties in discriminating qualities. While the accuracy measures do not show substantial accuracy values, the models struggle to identify each personality feature reliably. Similar trends may be seen in the BiGRU model findings with different embeddings.

The poor performance can be attributed to a variety of variables, such as the unique nature of the essay dataset, potential noise or unpredictability in the data, and differences in writing styles and content compared to the myPersonality Facebook dataset. Furthermore, the addition of NLP statistical features did not result in significant gains, implying that the essay dataset's linguistic and structural qualities may not correspond well with the features used.

Comparison

Table 4 provides a complete assessment of previous research endeavors' personality characteristic outcomes using different machine learning and deep learning models. Table 4 focuses solely on the myPersonality dataset, which includes several approaches such as deep learning, machine learning, and model averaging.

To assess the efficacy of their models, researchers used a mix of performance metrics, including f1-score and Accuracy. When the comparison of our model with other proposed model are examined, our proposed model has improved performance in term of accuracy and f1-measure. Furthermore, the table reveals that models that combine NLP statistical features and pre-trained models outperform those depending only on individual pre-trained model features or using only NLP Statistical features. This supports the idea that integrating NLP features leads to a significant improvement in model performance when forecasting personality traits.

Table 4
Comparison of Model Performance

Research	EXT	NEU	AGR	CON	OPN
Tandera et al. 2017	78.95% on MLP	79.49% on MLP	67.39% on CNN ID	62.00% on GRU	79.31% on MLP and CCN ID
Tadesse, M. M., et al., 2018	78.6% SNA+ XGB	68.0%SNA+ XGB	65.3% SNA+ XGB	69.8% SNA+ XGB	73.3% SNA+XGB, also onAlso on LWIC+ XGB
Yuan et al., 2018	57.0% On CNN	60.0% On CNN	57.0% on CNN	58.0% on CNN	76.0% On CNN

Christian, H., et al, 2021	76.92% On Model Averaging	78.21% On Model Averaging	72.33% On XLNet + NLP Features	70.85% On Model Averaging	86.17% On Model Averaging
Our model	**85.16% on Big bird + NLP features + Bilstm**	**87.39% On Distilbert + Bilstm**	**92.35% On Big bird + NLP features + Bilstm**	**98.48% on Big bird + NLP features + Bilstm**	**98.33% on Big bird + NLP features + Bilstm**
Results based on F1-Score					
Zheng and Wu, et al. 2019	0.71 On PMC +LIWC + unigram	0.70 On PMC +LIWC + unigram	0.68 On PMC+ LIWC + unigram	0.64 On PMC +LIWC	0.65 On PMC+LIWC With oror without “unigram”
Christian, H., et al, 2021	0.748 On Model Averaging	0.709 On XLNet + NLP Features	0.701 On XLNet + NLP Features	0.652 On Model Averaging	0.912 On Model Averaging
Our model	**0.82 on Big bird + NLP features + BiLstmBilstm**	**0.76 on Distilbert +BiLstm**	**0.74 on Big bird + NLP features + BiLstm**	**0.84 on Big bird + NLP features +BiLstm**	**0.81 on Big bird + NLP features + Bilstm**

Discussion

This research shed some light on many significant issues of personality prediction using text information. The difference in performance between the myPersonality Facebook data and the essay data indicates that the nature of the data, and the certain linguistic characteristics of the data, exert a significant influence on the model performance. In the case of Facebook dataset, a combination of NLP statistical features and pre-trained transformer models came with incredibly accurate prediction of personality traits. It shows how effective these transformer models are in capturing language and context nuances that may be of value in certain applications such as the classification of personality traits. The prediction capabilities of the models were also improved with the NLP statistical components, thus resulting in more accurate and higher F1 scores of several personality traits. These results prove that the transformer models of NLP statistical characteristics may be viewed as a perspective technique to complement the performance of personality trait prediction in some situations and the information related to social media. On the other hand, the data of the essay provided dissimilar issues. The models need assistance in order to generalize even with the same models and feature engineering methods. Such a gap shows how the dataset-specific characteristics, the writing style, and the change of content affect the model performance. Moreover, the fact that there was a limited improvement with the addition of NLP statistical features within the essay dataset indicates that such features are not perfectly suited to the linguistic and structural dimensions of essay-based data. Further studies are needed to better ensure prediction model generality of personality traits by designing feature engineering techniques tailored to essay dataset characteristics.

The comparatively low performance recorded with some personality traits in Table 2 and 3 can be explained by a number of factors. First, these traits are making it very difficult in textual data. Characteristics such as EXT and NEU have no clear linguistic expressions and therefore cannot be identified by models using a text-only method. Personality traits may be inherently context-dependent making their prediction even more difficult. In such instances, the training data does not include a wide variety of linguistic manifestations of these characteristics, and the models do not generalize well. Second, the performance difference is also influenced by the selection of model architecture. Transformers of different models do not all work equally well to detect fine linguistic nuances that correlate to various personality features. Although certain personalities can be quite compatible with the strengths of a given model, some will not, leading to poor performance. Examples of this include models such as Bigbird and Albert, which may not perform well on properties that have more complex linguistic structures, or properties whose properties are less directly marked in text. When that happens, the structure of the model and its inherent biases prevent it to perform well in particular traits. When seeking to enhance the performance of difficult personality traits, the architecture of the model must be considered with regard to the nature of the data.

Conclusion

We have discussed one of the curious directions in this work the ability to predict personality traits using a text with the help of a high-quality deep learning engine such as BigBird, ALBERT, or DistilBERT. We have explored two different data sets, the myPersonality Facebook dataset and an essay dataset, and both of them are subject to certain obstacles and challenges. In our models, we were able to infer the characteristics of personalities using the rich Facebook posts in the myPersonality Facebook sample. Combining context-laden embeddings with NLP statistical features, we improved the accuracy and the F1 score by a significant margin. This leads to the potential to use high-order deep-learning models to obtain dynamic information about personality assuming the dynamism of social interactions on the web. But this was a harder puzzle with the essay data set. Even when we trained our models to run in the myPersonality environment, they could not predict personality attributes using the broad and fine information in the essays. In this case, the accuracy or the F1 scores were also lower, and it is worth noting that special procedures that reflect specific linguistic features of a given textual source are needed. Further study should be conducted on how to optimize transformer-based models, feature engineering, multimodal, cross-cultural study, and ethical problems. These methods can produce new findings and extend the limitations of text-based personality prediction.

Recommendations

For future work, it is recommended to design dataset-specific feature engineering techniques that align with the linguistic and structural properties of diverse textual sources. Incorporating multimodal approaches that integrate textual data with additional modalities, such as images, social interactions, or behavioural signals, may enhance the robustness of personality recognition models. Furthermore, conducting cross-cultural validation studies can ensure generalizability across different populations, while addressing ethical and privacy-preserving mechanisms (e.g., federated learning or differential privacy) will support the responsible deployment of personality prediction systems in real-world applications.

References

Acheampong, F., Nunoo-Mensah, H., & Chen, W. (2021). Transformer models for text-based emotion detection: A review of BERT-based approaches. *Artificial Intelligence Review, 54*(1), 1–41. https://doi.org/10.1007/s10462-021-09958-2

Ahmad, H., Asghar, M. U., Asghar, M. Z., Khan, A., & Mosavi, A. H. (2021). A hybrid deep learning technique for personality trait classification from text. *IEEE Access, 9*, 146214–146232. https://doi.org/10.1109/ACCESS.2021.3123456

Aung, Z. M. M., & Myint, P. H. (2019). Personality prediction based on content of Facebook users: A literature review. *2019 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, 34–38. https://doi.org/10.1109/SNPD.2019.8935692

Bharadwaj, S., Sridhar, S., Choudhary, R., & Srinath, R. (2018). Persona traits identification based on Myers-Briggs type indicator (MBTI): A text classification approach. *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 1076–1082. https://doi.org/10.1109/ICACCI.2018.8554828

Bruno, A., & Singh, G. (2022). Personality traits prediction from text via machine learning. *2022 IEEE World Conference on Applied Intelligence and Computing (AIC)*, 588–594. https://doi.org/10.1109/AIC55036.2022.9848937

Chin, K., Zhang, Z., Long, J., & Zhang, H. (2016). Turning from TF-IDF to TF-IGM for term weighting in text classification. *Expert Systems with Applications, 66*, 1–12. https://doi.org/10.1016/j.eswa.2016.09.009

Christian, H., Chowanda, A., Suhartono, D., et al. (2021). Text-based personality prediction from multiple social media data sources using pre-trained language model and model averaging. *Journal of Big Data, 8*(1), 1–20. https://doi.org/10.1186/s40537-021-00459-1

Cherukuru, R. K., Kumar, A., Srivastava, S., & Verma, V. K. (2022). Prediction of personality traits using machine learning on online texts. *2022 International Conference for Advancement in Technology (ICONAT)*, 1–8. https://doi.org/10.1109/ICONAT53423.2022.9725910

Das, K. A. H., & Prajapatd, H. (2022). Personality identification based on MBTI dimensions using natural language processing. *International Journal of Creative Research Thoughts, 8*(6), 1653–1657. https://ijcrt.org/papers/IJCRT2006219.pdf

El-Demerdash, K., El-Khoribi, R. A., Shoman, M. A. I., & Abdou, S. (2022). Deep learning-based fusion strategies for personality prediction. *Egyptian Informatics Journal, 23*(1), 47–53. https://doi.org/10.1016/j.eij.2021.05.004

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*(3), 221–233. https://doi.org/10.1037/h0057532

Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., & Soricut, R. (2020). ALBERT: A lite BERT for self-supervised learning of language representations. *arXiv*. https://arxiv.org/abs/1909.11942

Mohammad, S. M., & Turney, P. D. (2013). NRC emotion lexicon. *National Research Council Canada*.

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. *Proceedings of LREC*, 1320–1326.

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 77*(6), 1296–1312. https://doi.org/10.1037/0022-3514.77.6.1296

Pennebaker, J. W., Chung, C., Ireland, M., Gonzales, A., & Booth, R. (2007). The development and psychometric properties of LIWC2007. *LIWC.net*.

Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter. *arXiv*. https://arxiv.org/abs/1910.01108

Sadr, H., Tarkhan, M., & Mohades Deilami, F. (2022). Contextualized multidimensional personality recognition using a combination of deep neural network and ensemble learning. *Journal/Conference TBD*.

Stillwell, D. J., & Kosinski, M. (2015). myPersonality project website. *Cambridge Psychometrics Centre*.

Tadesse, M. M., Lin, H., Xu, B., & Yang, L. (2018). Personality predictions based on user behavior on the Facebook social media platform. *IEEE Access, 6*, 61959–61969. https://doi.org/10.1109/ACCESS.2018.2876502

Tandera, T., Hendro, Suhartono, D., Wongso, R., & Prasetio, Y. L. (2017). Personality prediction system from Facebook users. *Procedia Computer Science, 116*, 604–611. https://doi.org/10.1016/j.procs.2017.10.016

William, P., Badholia, A., Patel, B., & Nigam, M. (2022). Hybrid machine learning technique for personality classification from online text using HEXACO model. *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, 253–259. https://doi.org/10.1109/ICSCDS53736.2022.9760970

Xue, X., Feng, J., & Sun, X. (2021). Semantic-enhanced sequential modeling for personality trait recognition from texts. *Applied Intelligence, 51*(11), 7705–7717. https://doi.org/10.1007/s10489-021-02277-7

Yuan, C., Wu, J., Li, H., & Wang, L. (2018). Personality recognition based on user-generated content. *2018 15th International Conference on Service Systems and Service Management (ICSSSM)*, 1–6. https://doi.org/10.1109/ICSSSM.2018.8465006

Yu, J., & Markov, K. (2017). Deep learning-based personality recognition from Facebook status updates. *2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST)*, 383–387. https://doi.org/10.1109/ICAwST.2017.8256484

Zaheer, M., Guruganesh, G., Dubey, A., Ainslie, J., Alberti, C., Ontañón, S., & Ahmed, Z. (2020). Big bird: Transformers for longer sequences. *arXiv*. https://arxiv.org/abs/2007.14062

Zheng, H., & Wu, C. (2019). Predicting personality using Facebook status based on semi-supervised learning. *Proceedings of the 2019 11th International Conference on Machine Learning and Computing (ICMLC '19)*, 59–64. https://doi.org/10.1145/3318299.3318363

Zhou, L., Zhang, Z., Zhao, L., & Yang, P. (2022). Attention-based BiLSTM models for personality recognition from user-generated content. *Information Sciences, 596*, 460–471. https://doi.org/10.1016/j.ins.2022.03.038