International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 17 (2025) pp. 169–188
©Cerebration Science Publishing

Article

Unveiling Human Essence: Deep Learning in Personality Traits Detection

Nouha Arfaoui [1,*], Mariem Mkhini [2], Aboubacar Sidiki Sidibe [2], Bertille Baron [2] and Serawork Wallelign [2]

[1] Research Team in Intelligent Machines, National Engineering School, Gabes, 6029, Tunisia
[2] Kaisens Data, 16 Pl. de l'Iris, 92400 Courbevoie, France, mmkhinini@kaisensdata.fr (M.M.); asidiki@kaisensdata.fr (A.S.S.); bbaron@kaisensdata.fr (B.B.); swallelign@kaisensdata.fr (S.W.)
* Correspondence author: arfaoui.nouha@yahoo.fr

Received date: 1 May 2024; Accepted date: 19 July 2024; Published online: 6 January 2025

Abstract: Natural Language Processing (NLP) and Deep Learning (DL) are two branches of Artificial Intelligence (AI). They offer several methods, models and algorithms to facilitate the user work. The first one satisfies the requirement for natural language interaction between a user and the machine. The second one is based on multi-layered neural networks to simulate the human brain function in order to recognize complex patterns in structured and unstructured data. In this work, we propose exploiting the performances of NLP and DL to automate the detection of the personality traits from text. The personality is considered as the combination of various factors including behavior, emotion, thoughts, etc. The detection of personality traits has a crucial impact on our daily life improving, hence, our personal growth, social interactions, professional environments, customer experiences, and more. Many works in the literature proposed solutions to study the different personality traits from text. Because of the huge amount of data extracted from emails, social medias, etc., automating this task becomes crucial hence the use of DL and NLP. To achieve our goal, we implement three different approaches. We evaluate them and we compare them to extract the best one to our case. We used, for this purpose, Big Five dataset which represents the five traits of a person which are Openness, Conscientiousness, Extroversion, Agreeableness and Neuroticism. We used, also, NRC Emotion Lexicon which is a list of English words used to facilitate the extraction of emotional nuances from text data.

Keywords: Artificial Intelligence; Deep Learning; Personality Traits Detection; CNN; LSTM; BERT; XLNET; Leaky ReLU; Tanh; Sigmoid

1. Introduction

Deep learning (DL) is an Artificial Intelligence (AI) technique that uses multi-layered neural networks to simulate the human brain function. It can recognize complex patterns in pictures, text, sounds, and other data. It can, also, automate several tasks that were limited to human because of their intellectual capacity such as describing images or transcribing audio files to text.

Natural Language Processing (NLP) is an branch of artificial intelligence that encompasses a wide range of complex, advanced, and demanding language-related activities, including summarization, machine translation, and question answering [1]. In this context, models, methods, and algorithms are designed and putted into practice to address real-world issues related to language comprehension [1]. Its goal is to enable computers to comprehend words or sentences written in human languages. It was created to make user work easier and to fulfill the need for natural language communication between the user and the machine [2].

The machine learning community is becoming increasingly interested in NLP applications as a result of the remarkable performance improvement brought about by recent advancements in DL.

In this context, we propose combining the techniques of NLP with DL to automate the detection of the personality traits. Deep learning-based neural network models are very good at identifying personality traits because of their ability to learn and NLP offers powerful tools to understand the context in which words are used.

The personality is defined as the characteristic set of behaviors, cognitions, and emotional patterns [3]. It can manifest in different ways such writing, speech, decision, etc. [4]. With the continuous development of the internet, then, the social media, the emails, etc., people find more space to express themselves using words. The quantity of data that can be used for studying the people personality is huge, hence the importance of the automation of this task which requires finding a way to measure this personality to be able to classify it for a good comparison. To achieve this task, it is crucial to go back to psychology [4]. Different models had been proposed and studied. We mention, as example, Cattell's 16 Factor Model traits [5], Eysenck Personality Questionnaire (EPQ) [6], Myers–Briggs Type Indicator (MBTI) [7], and Big Five [8]. The latter is used by the majority of the research for classification [9]. It represents the five traits of a person: Openness, Conscientiousness, Extroversion, Agreeableness and Neuroticism. They are called OCEAN in abbreviate.

The purpose of this work is to ensure an automatic detection of a personality based on text data type. Automated personality recognition systems is the subject of various industrial applications. For example, using data extracted from the conversations between players, Yang and Huang [10] studied their personality, then they created a system for recommending games. Yin et al., [11], based on the hobbies and the personality of a client, developed a model concerning automobile purchase intentions. There is also the Word polarity detection. It can be used to solve the problem of word polarity or to distinguish the sarcastic and non-sarcastic content [12]. Personality recognition is very similar to NLP applications. Both of them focus on mining users' attributes from texts [13]. Hence, we are using powerful text modeling techniques to improve the performance of personality recognition [14,15], and we combine it with Deep Neural Network (DNN) that demonstrated remarkable performance [16,17].

In this work, we propose implementing and evaluating several Deep Learning algorithms: BERT-CNN, BERT-LSTM, Word2Vec-CNN, Word2Vec-LSTM, and two Ensemble models. For the two latter, we combine BERT-CNN and BERT-LSTM with logistic regression in the first case and Word2Vec-CNN and Word2Vec-LSTM with logistic regression in the second case. We calculate the accuracy value in the different cases to check the best algorithm having the highest value.

The rest of paper is organized as following. In Section 2, we present some of the existing works using different deep learning algorithm for personality detection. In Section 3, we describe the used datasets. In Section 4, we present the different techniques used to ensure the data augmentation. Section 5 focuses on the techniques of preprocessing to get clean data from the used dataset.In Section 6, we define the used metric of evaluation which is in our case the accuracy. Next, implement and evaluate three different approaches using the accuracy. At the end of this section, we discus the obtained results. In Section 8, we deploy the solution using Django and we give the different implemented interfaces. Finally, we conclude this work and we give some perspectives.

2. State of the Art

In this section, we summarize some of the existing works that studied the personality traits extraction using deep learning techniques.

Through this work [18], the authors are looking for the best personality traits extraction framework. They used exclusive five special CNN networks having the same architecture. Each architecture plays the role of a binary classifier for the different traits to be positive or negative. With the deep learning algorithm, the authors tested different activation functions: sigmoid, tanh, and leaky ReLU. Based on the experiments, tanh and leaky ReLU performed better over sigmoid activation function.

In [19], the authors proposed a model named 2CLSTM which is the result of the concatenation of Long Short Term Memory networks(LSTM) and Convolution Neural Network (CNN). The model detects the personality traits using the structure of the text that can be an important feature, also it uses Latent Sentence Group (LSG) to capture the abstract feature combination based on closely connected sentences. To validate their model, the authors used two datasets which are: Stream-of-consciousness essays and YouTube personality dataset to well explain the versatility of their models.

The authors, in [20], proposed Deep Learning-based Bagged-SVM in order to automate the personality detection from the text using BERT to extract contextualized word embeddings from textual data. The proposed solution can be easily used by a large number of people without access to huge computation resources. Compared to the state of the art, the proposed model outperforms by 1.04%

In [21], the authors proposed exploiting the rich semantic information that can be extracted from users' generated texts since they can be the most directed way to express one's feeling and opinion. For this purpose, they implemented a combination of deep learning model CNN and AdaBoost. CNN is efficient with natural language processing and AdaBoost can combine various classifiers with respective filter size. The solution has been validated using Essay dataset.

The authors proposed, in [22], Transformer-MD (Multi-Document Transformer model). The latter can access to information in the other posts of the user using Transformer-XL's memory tokens. It uses dimension attention mechanism on the top to obtain trait-specific representations for multi-trait personality detection.

In [23], the authors proposed attention-based BILSTM for psychopath's detection to classify input text into psychopath and non-psychopath traits. This approach has the advantage of using increased data set size for an efficient classification. It stores the past information using Backward LSTM and future information using Forward LSTM.

Table 1 describes for each cited work the used Word Embedding, the Datatset, the DL algorithms, the used metric(s) for the evaluation, and its (their) the value.

Table 1. Comparative Table.

Ref.	Embedding	DataSet	Model	Evaluation Metric	Value
[18]	Googles Pre-trained Word and Phrase Vectors	Stream of Consciousness essays	CNN	average F1score	Sigmoid = 0.3311 Tanh = 0.4725 Leaky ReLU=0. 4907
[19]	GloVe	Stream of Consciousness essays and Youtube personality dataset	CNN and RNN	Precision	stream-of-consciousness: AGR=0.5887 ; EXT=0.5564 CON =0.5352; NEU =0.5352 OPN=0.5419 YouTube AGR= 0.5802; EXT= 0.6769 CON= 0.6117; NEU= 0.6128 OPN=0.5546
[20]	BERT	Stream of Consciousness essays	BB-SVM	average accuracy	0.5903
[21]	-	Stream of Consciousness essays	Combination of CNN and AdaBoost	Accuracy	**CNN-AdaBoost-Rand** EXT =0.6005 ; NEU=0.6091 AGR=0.5811 ; OPN=0.5934 CON=0.6394 **CNN-AdaBoost-Static** EXT=0.6012 ; NEU=0.6105 AGR=0.5835 ; OPN=0.5968 CON=0.6405 **CNN-AdaBoost-Non-Static** EXT=0.6045 ; NEU=0.6171 AGR=0.5861 ; OPN=0.6001 CON=0.6418 **CNN-AdaBoost-2channel** EXT=0.6125 ; NEU=0.6193 AGR=0.5902 ; OPN=0.6016 CON=0.6463
[22]	XLNet	Kaggle and Pandora MBTI personality datasets	Transformer-MD	Average macro-F1	Transformer-MDDA= 0.6092
[23]	Keras embedding	Dataset from different social media sites like Facebook and Twitter	BILSTM	Precision Recall F1-Score Accuracy	0.85 0.85 0.85 0.85

3. Dataset

In this section, we present the different used datasets to perform the evaluation.

3.1. Big Five Dataset

The Big Five personality test is an all-encompassing personality inventory that draws upon decades of psychological research. Through extensive study, psychologists and academic researchers have consistently observed that individuals' personality variances tend to cluster into five overarching dimensions, commonly known as the Big Five. In this work, we use Essays [24]. It is a well-know personality dataset. It is composed by 2,468 anonymous self-reported journals for different individuals labeled by the participant's personality traits [9]. It presents the different personality traits using five independent metrics: which are cEXT, cNEU, cAGR, cCON and cOPN. Where:

- **cEXT**: It corresponds to "Extroversion". It is used to evaluate a person's curiosity, imagination, creativity and open-mindedness.
- **cNEU**: It corresponds to "Neuroticism". It is used to evaluate the emotionality, anxiety, emotional instability and the propensity for negative feelings.
- **cAGR**: It corresponds to "Agreeableness". It is used to evaluate the tendency of a person to be altruistic, compassionate, cooperative and friendly.

- **cCON**: It corresponds to ”Conscientiousness”. It is used to evaluate a person’s reliability, responsibility, self-control and organizational ability.
- **cOPN**: It corresponds to ”Extraversion”. It is used to evaluate the sociability, enthusiasm, social warmth and stimulation seeking.

There is also another column:

- **TEXT**: It contains the texts or messages written by the authors. These texts serve as text samples for our analysis of personality traits. Each text may vary in length and content, as it represents the individual expression of each author.

Figure 1 presents an excerpt of the dataset content.

	#AUTHID	TEXT	cEXT	cNEU	cAGR	cCON	cOPN
0	1997_504851.txt	Well, right now I just woke up from a mid-day ...	n	y	y	n	y
1	1997_605191.txt	Well, here we go with the stream of consciousn...	n	n	y	n	n
2	1997_687252.txt	An open keyboard and buttons to push. The thin...	n	y	n	y	y
3	1997_568848.txt	I can't believe it! It's really happening! M...	y	n	y	y	n
4	1997_688160.txt	Well, here I go with the good old stream of co...	y	n	y	n	y
...	...	...	...	...	...	...	...
2462	2004_493.txt	I'm home. wanted to go to bed but remembe...	n	y	n	y	n
2463	2004_494.txt	Stream of consiousnessskdj. How do you s...	y	y	n	n	y
2464	2004_497.txt	It is Wednesday, December 8th and a lot has be...	n	n	y	n	n
2465	2004_498.txt	Man this week has been hellish. Anyways, now i...	n	y	n	n	y
2466	2004_499.txt	I have just gotten off the phone with brady. I...	n	y	y	n	y

2467 rows × 7 columns

Figure 1. The structure of the used dataset.

3.2. NRC Word-Emotion

NRC (https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm, accessed on 01/12/2023) Emotion Lexicon is a list of English words. Each word is meticulously associated with an emotion, facilitating the extraction of emotional nuances from text data. They are eight basic emotions : anger, fear, anticipation, trust, surprise, sadness, joy, and disgust and two dimensions which are negative and positive. The process of annotating these associations was carried out manually through crowdsourcing. This dataset empowers our project with the capability to unearth subtle emotional cues that intertwine with the broader personality traits. By aligning the emotional dimensions with the Big Five traits, we enrich our analysis and contribute to a more comprehensive understanding of personality traits inferred from text.

4. Data Augmentation

In our study, we employed a range of data augmentation techniques to enhance the diversity and size of our dataset. The following techniques were specifically utilized:

- **Paraphrasing**: This technique involves rephrasing a sentence or text while retaining its original meaning. We leveraged automated paraphrasing algorithms and lexical resources to identify equivalent or similar expressions. Hence, we generated new instances of essays while preserving the overall contextual meaning of the sentences.

Figure 2 presents an example of paraphrasing.

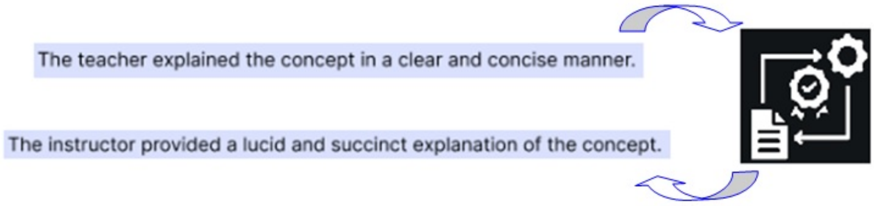

Figure 2. Example of paraphrasing.

- **Synonym Replacement**: In this approach, certain words are substituted with their synonyms. This introduces textual variation while maintaining the global sentence meaning. By replacing specific words with their synonyms, we enriched the diversity of our training data.

Figure 3 presents an example of synonym replacement.

Figure 3. Example of synonym replacement.

- **Back Translation**: This technique entails translating the text into another language and then translating it back to the original language. This process introduces variations in sentence formulation and structure. By employing this technique, we created new essay instances with linguistic variations while preserving a similar contextual sense.

Figure 4 presents an example of back translation where we used the french as an intermediate language.

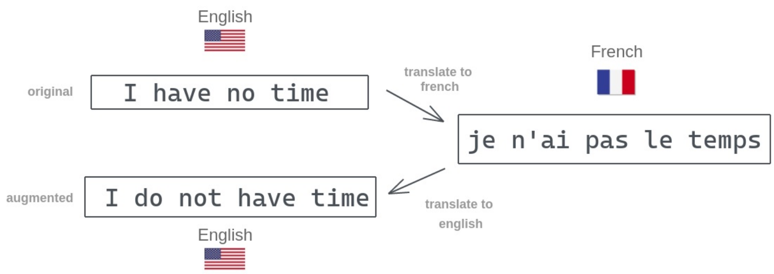

Figure 4. Example of back translation.

Through the application of these data augmentation techniques, we expanded the size of the dataset from 2,467 to 4,934. The newly generated essay instances featured linguistic and contextual variations while retaining the same personality trait classifications as those in the original dataset.

5. Preprocessing

Data preprocessing stands as a pivotal phase in data analysis, aiming to clean and transform raw data into a format suitable for analysis and modeling. It involves several steps:

- **Lowercasing**: All text was converted to lowercase to ensure consistency during analysis and to treat uppercase and lowercase letters uniformly.
- **Removal of Non-Alphabetic Characters**: Non-alphabetic characters, such as punctuation and special symbols, were removed from the text. This step aimed to focus solely on the alphabetic content of the text.
- **Tokenization**: It is employed to divide the text into individual words or tokens. This facilitated further analysis at the word level.

- **Stopword Removal**: Stopwords which corresponds to commonly used words that contribute little meaningful information, were eliminated from the text to reduce noise and concentrate on more significant words for analysis.
- **Lemmatization**: Lemmatization was applied to revert words to their base or dictionary form. This standardization process helped reduce variations among words.

6. Evaluation

To compare the performance of the proposed approaches, we use Accuracy (ACC). It is a metric used to evaluate the classification of models by measuring the ratio of correctly predicted instances over the total number of evaluated instances [25]. Its formula is as following:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

where:

- TP (True Positive): It implies the actual value is positive and the model predicts a positive value.
- TN (True Negative): It implies the actual value is negative and the model predicts a negative value.
- FP (False Positive): It implies the actual value is negative but the model predicts a positive value.
- FN (False Negative): It implies the actual value is positive but the model predicts a negative value. Confusion Matrix is used to calculate the following metrics.

7. Implementation and Discussion

Three Different approaches were implemented and evaluated to extract the most performed one according to the used dataset. In the first one, we implemented CNN with BERT then with XLNET, and LSTM with BERT then with XLNET. In the second one, we exploited the second Dataset: NRC Emotion Lexicon (as presented in Section 3.2) to extract more features from the first dataset and we integrated the results within the two models CNN (using BERT then W2V) and LSTM (using BERT then W2V). The last approach is about using the staking to create Ensemble model. In our work we implemented two Ensemble models.

7.1. First Approach

7.1.1. CNN and BERT

Table 2 presents the different accuracy values for each personality trait corresponding to the implementation of CNN with BERT.

Table 2. Accuracy values for BERT-CNN.

	Train	**Validation**	**Test**
cEXT	62%	51%	49%
cNEU	58%	50%	50%
cAGR	57%	53%	52%
cCON	66%	54%	55%
cOPN	63%	56%	57%

Figure 5 presents the histogram generated from the previous table (Table 2) where we display the different values for an easier interpretation.

The model appears to have good performance on the training set, with high accuracy for some traits, such as cCON and cOPN. However, it presents less satisfactory results on the validation set and the test set, indicating possible overfitting to the training data.

The difference between the performances on the training set and the validation/test set may indicate that the model has memorized the training data rather than generalizing effectively to the new data. The accuracy of the model on the test set is close to that of the validation set, suggesting that the model does not perform much better on unknown examples.

Overall, although the model showed encouraging performance on the training set, there are signs of overfitting which could be corrected by better regularization of the model. This encourages us to explore other combinations of models and embeddings to see if we can improve generalization and accuracy on test sets.

In conclusion, the BERT-CNN model showed promising performance on the training set, but it needs improvement to better generalize to new data.

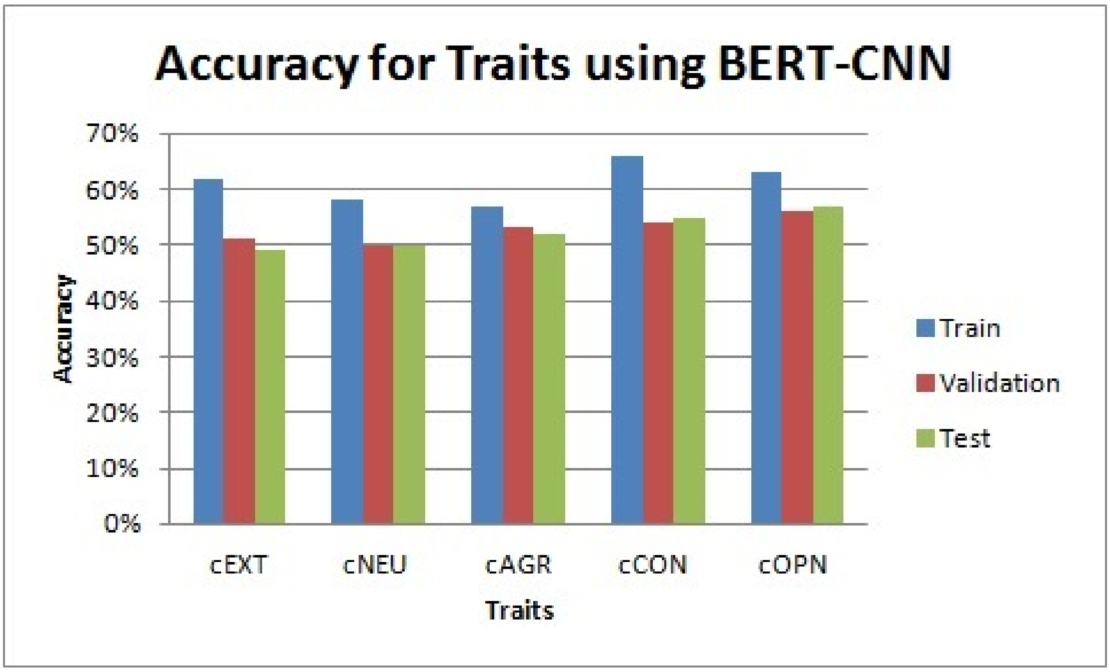

Figure 5. Generated histogram from Table 2.

7.1.2. CNN and XLNET

Table 3 presents the different accuracy values for each personality trait corresponding to the implementation of CNN with XLNET.

Table 3. Accuracy values for XLNET-CNN.

	Train	**Validation**	**Test**
cEXT	68%	52%	53%
cNEU	67%	52%	54%
cAGR	66%	54%	51%
cCON	57%	50%	51%
cOPN	66%	57%	55%

Figure 6 presents the histogram generated from the previous table (Table 3) where we display the different values for an easier interpretation.

The XLNet-CNN model shows higher performance than the previous model (BERT and CNN) on the training set, with higher accuracy for all features. This suggests that using CNN architecture in combination with XLNet embedding resulted in more discriminative representations for personality traits.

However, despite high performance on the training set, the model shows mixed performance on the validation set and test set. The difference between the training and validation/test scores further indicates possible overfitting to the training data.

The model's accuracy on the test set is slightly higher than that of the validation set, suggesting that the model generalizes slightly better to unknown examples. However, these performances could be improved to obtain better generalization.

In conclusion, the XLNet-CNN model shows promising performance on the training set, but also shows signs of overfitting.

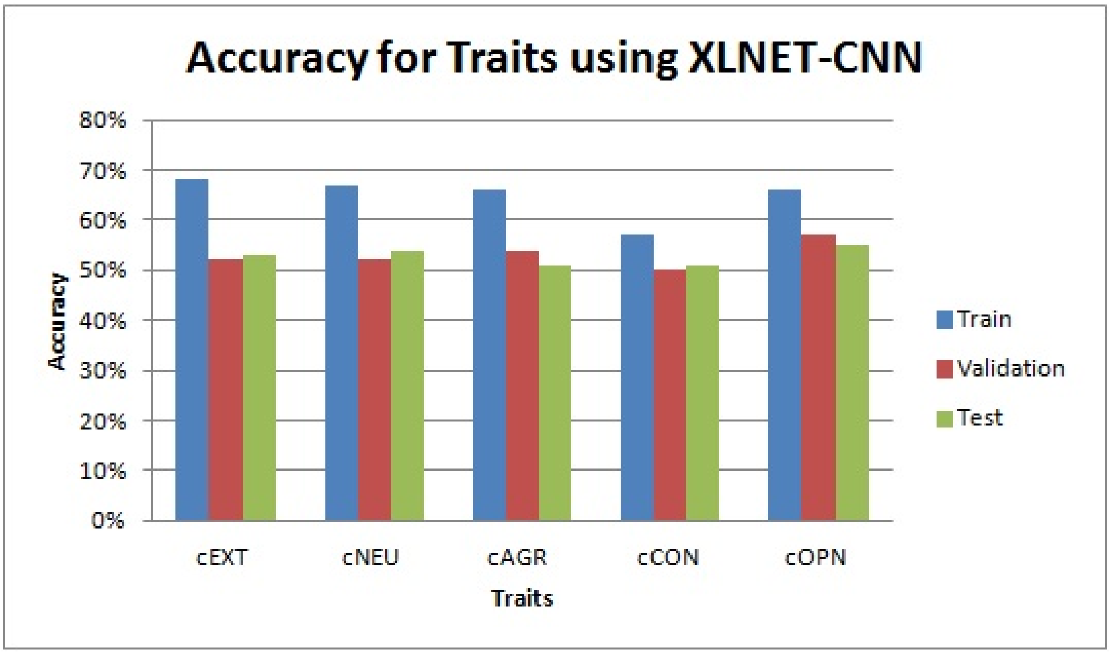

Figure 6. Generated histogram from Table 3.

7.1.3. LSTM and BERT

Table 4 presents the different accuracy values for each personality trait corresponding to the implementation of LSTM with BERT.

Table 4. Accuracy values for BERT-LSTM.

	Train	**Validation**	**Test**
cEXT	52.11%	50.51%	53.19%
cNEU	50.17%	50.51%	48.94%
cAGR	53.47%	52.78%	52.18%
cCON	50.59%	49.37%	52.79%
cOPN	52.42%	50.76%	49.24%

Figure 7 presents the histogram generated from the previous table (Table 4) where we display the different values for an easier interpretation.

The BERT-LSTM model showed moderate performance in personality trait classification. However, it presents similar results on the training, validation and test sets, suggesting good generalization of the model to unknown data. This may indicate that the model was successful in learning meaningful information from the training data and applying it effectively to new examples, furthermore the LSTM-BERT model no longer appears to suffer from overfitting.

Compared to previous approaches (BERT-CNN and CNN-XLNet), the BERT-LSTM model seems to generalize better to test data, with more stable performance across different sets. However, the overall accuracy remains moderate and there is still possible to improve the results.

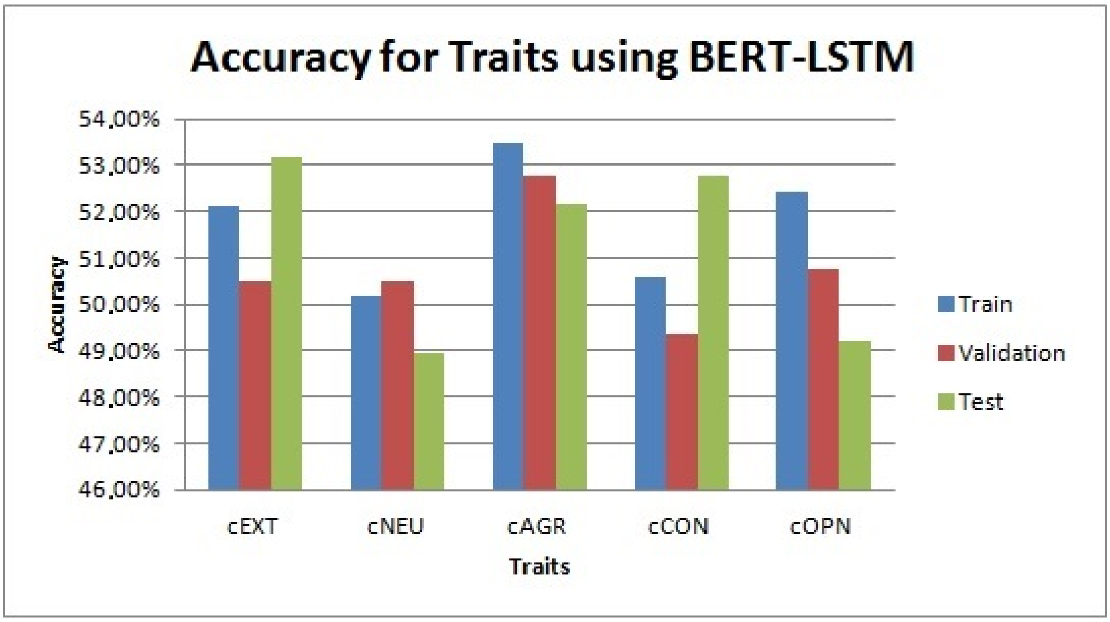

Figure 7. Generated histogram from Table 4.

7.1.4. LSTM and XLNET

Table 5 presents the different accuracy values for each personality trait corresponding to the implementation of LSTM with XLNET.

Table 5. Accuracy values for XLNET-LSTM.

	Train	**Validation**	**Test**
cEXT	52.01%	51.14%	50.76%
cNEU	50.21%	50.51%	48.94%
cAGR	53.47%	52.78%	52.18%
cCON	49.29%	52.41%	46.40%
cOPN	50.97%	52.91%	52.08%

Figure 8 presents the histogram generated from the previous table (Table 5) where we display the different values for an easier interpretation.

The XLNet-LSTM model shows similar performance to the previous models that we tested (BERT-CNN, XLNET-CNN, etc.).However, its performance is not very interesting.The accuracy obtained for each trait on the training, validation, and testing set are close to each other, suggesting that the model does not suffer from overfitting. However, the overall results remain relatively modest.

In conclusion, although the XLNET-LSTM model shows a lack of overfitting and stable performance, there are still opportunities for improvement.

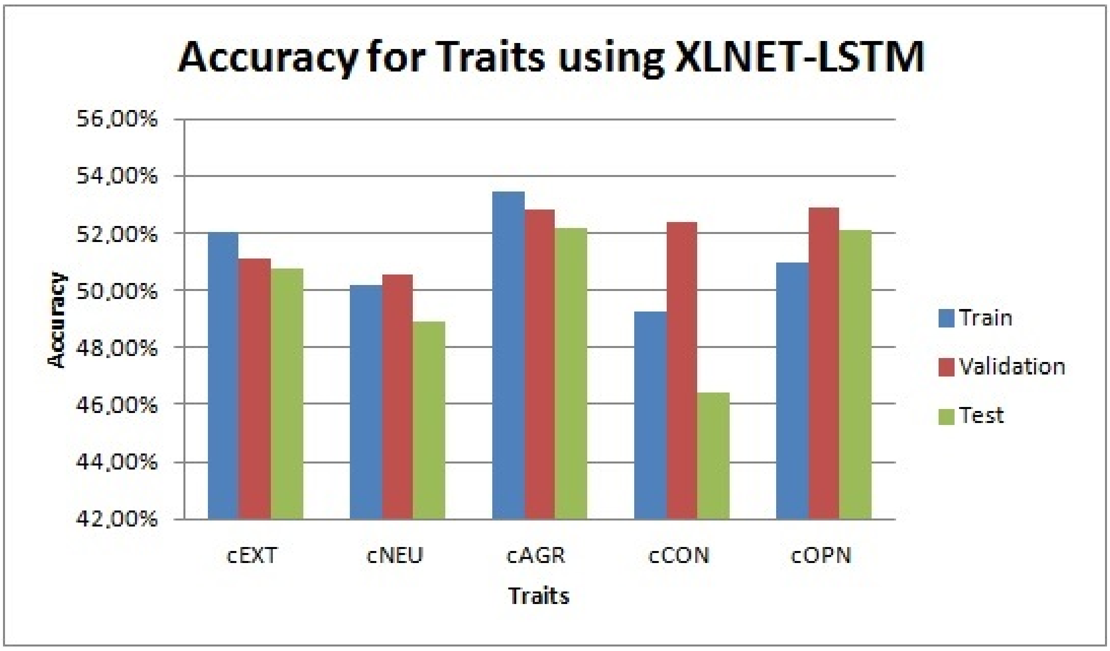

Figure 8. Generated histogram from Table 5.

7.2. Second Approach

Compared to the previous approach, in this one, we use the second dataset to extract more features that will be passed to the implemented models in order to improve the accuracy values.

7.2.1. CNN-BERT with Extracted Features

Table 6 presents the different accuracy values for each personality trait corresponding to the implementation of CNN, BERT with the extracted features.

Table 6. Accuracy values for BERT-CNN with extracted features.

	Train	**Validation**	**Test**
cEXT	52.07%	55.44%	45.69%
cNEU	57.02%	56.33%	53.39%
cAGR	57.99%	54.81%	46.61%
cCON	58.63%	58.10%	47.01%
cOPN	64.08%	61.27%	60.18%

Figure 9 presents the histogram generated from the previous table (Table 6) where we display the different values for an easier interpretation.

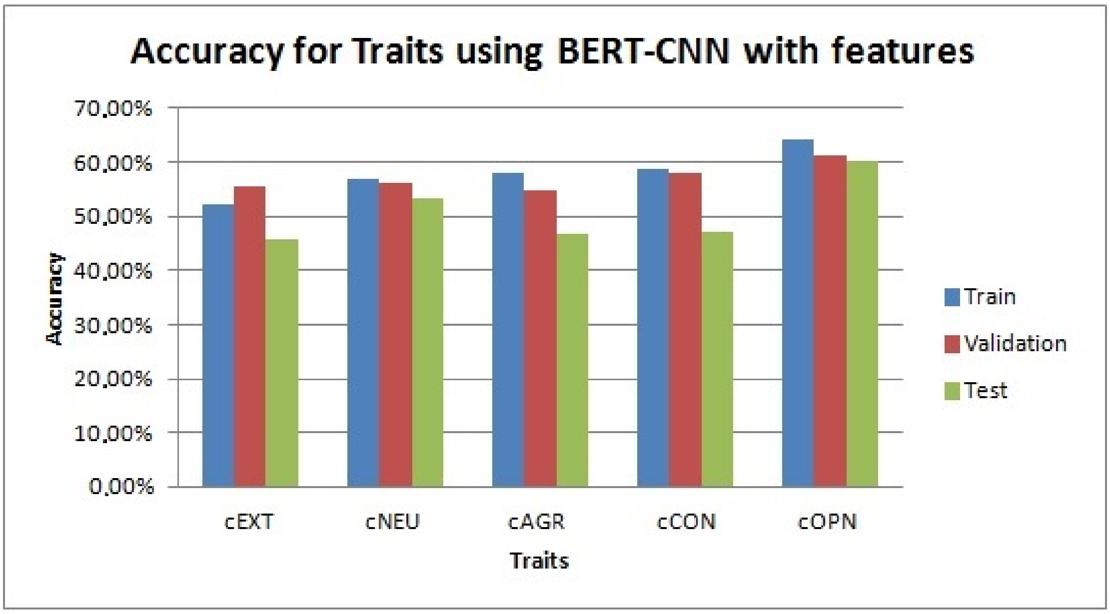

Figure 9. Generated histogram from Table 6.

The BERT with the extracted features and CNN model shows encouraging performance with accuracy generally above 0.50 for all personality traits on the training, validation and test set. This shows that the model is able to capture meaningful relationships between personality traits and texts.

However, we see that the performance on the validation and test set is relatively close to that of the training set, indicating a potential risk of overfitting. Although the model showed good ability to generalize to new data, it can be improved to avoid overfitting. The addition of the extracted features and the CNN layer improved the performance compared to the previous combination (BERT-CNN), especially on the validation set. This confirms our hypothesis that taking features into account allows us to better understand the linguistic aspects of the text and improve the overall representation.

7.2.2. CNN-W2V with Extracted Features

Table 7 presents the different accuracy values for each personality trait corresponding to the implementation of CNN, W2V with the extracted features.

Table 7. Accuracy values for W2V-CNN with extracted features.

	Train	**Validation**	**Test**
cEXT	100%	48.86%	46.91%
cNEU	87.39%	52.15%	54.81%
cAGR	96.26%	50.76%	47.82%
cCON	96.55%	53.04%	47.42%
cOPN	81.47%	54.68%	54.51%

Figure 10 presents the histogram generated from the previous table (Table 7) where we display the different values for an easier interpretation.

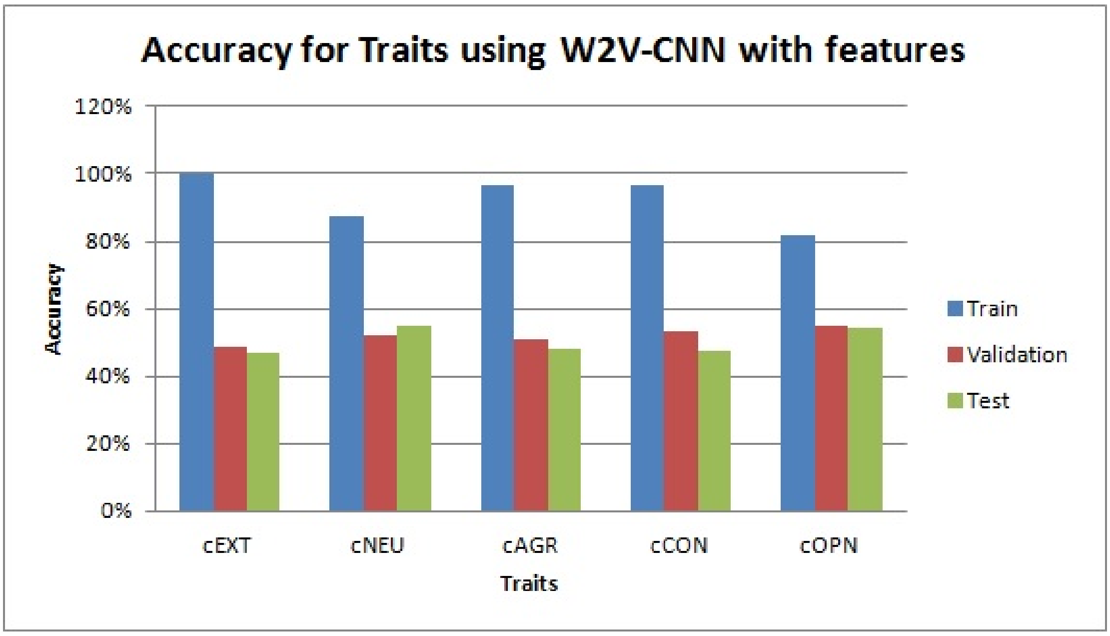

Figure 10. Generated histogram from Table 7.

The model using Word2Vec with the extracted features and convolution network (CNN) shows remarkable performance on the training set, with an accuracy of 1.0 for the cEXT feature and high accuracy values for the other features. These results indicate that the model successfully learned the training data well and captured the relationships between textual features and personality traits.

However, the performance on the validation set and test set is relatively lower. This suggests that the model has difficulty generalizing to new data, which can be attributed to possible overfitting to the training data.

7.2.3. LSTM-BERT with Extracted Features

Table 8 presents the different accuracy values for each personality trait corresponding to the implementation of LSTM, BERT with the extracted features.

Table 8. Accuracy values for BERT-LSTM with extracted features.

	Train	**Validation**	**Test**
cEXT	58.66%	51.90%	49.04%
cNEU	56.51%	55.82%	51.87%
cAGR	56.32%	54.68%	44.68%
cCON	57.27%	56.46%	45.29%
cOPN	62.12%	57.47%	54.41%

Figure 11 presents the histogram generated from the previous table (Table 8) where we display the different values for an easier interpretation.

The LSTM model with BERT word embeddings and extracted features shows interesting performance on the training set, with high accuracies for some features such as cEXT, cNEU and cOPN. This indicates that the model is successful captured some trends and correlations within the training data. However, the performance on the validation set and test set is relatively lower. This implies that the model has some difficulty to be applied to new data, which can be attributed to possible overfitting to the training data.

The small gap between performance on the validation set and the test set indicates that the model behaves similarly on unknown examples, but overall the results on the test set remain moderate.

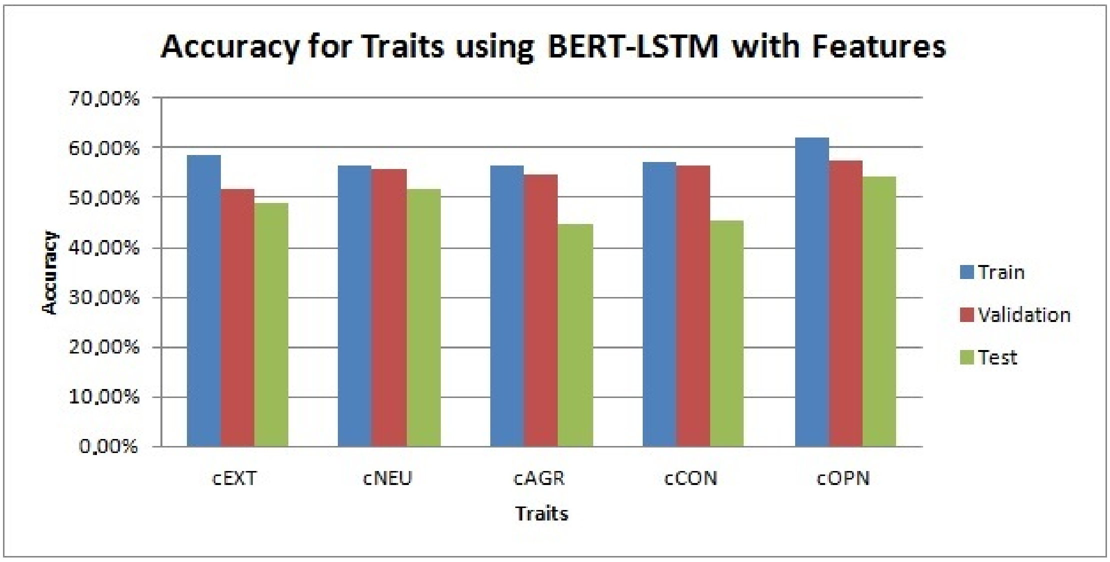

Figure 11. Generated histogram from Table 8.

7.2.4. LSTM-W2V with Extracted Features

Table 9 presents the different accuracy values for each personality trait corresponding to the implementation of LSTM, W2V with the extracted features.

Table 9. Accuracy values for W2V-LSTM with extracted features.

	Train	**Validation**	**Test**
cEXT	63.19%	51.52%	50.56%
cNEU	56.38%	54.05%	51.57%
cAGR	62.24%	55.06%	49.65%
cCON	63.32%	54.18%	50.46%
cOPN	69.88%	55.82%	53.90%

Figure 12 presents the histogram generated from the previous table (Table 9) where we display the different values for an easier interpretation.

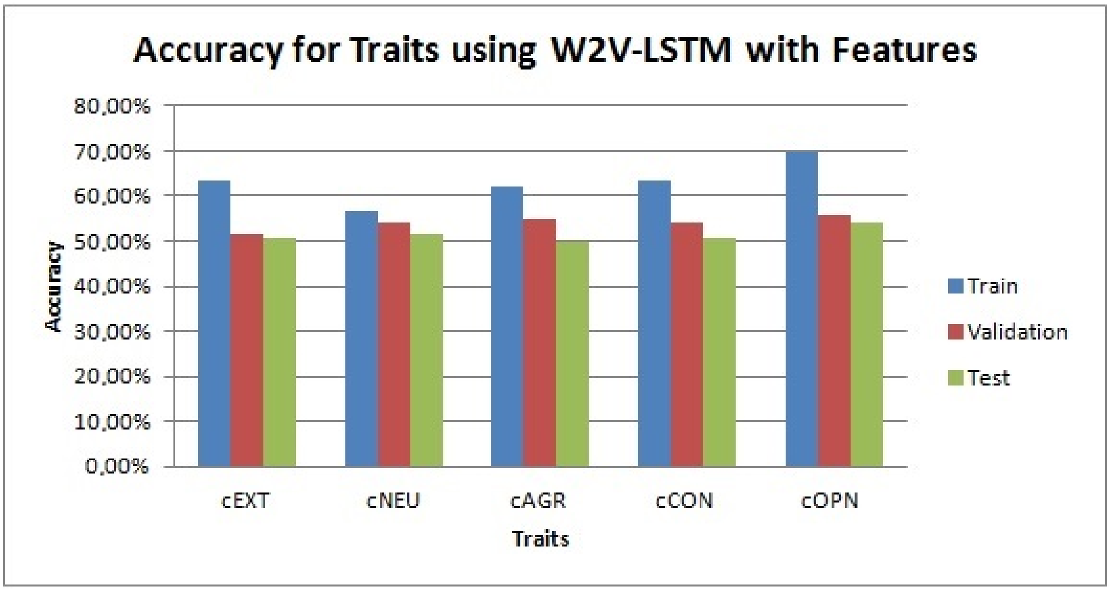

Figure 12. Generated histogram from Table 9.

The LSTM with Word2Vec and the extracted features model shows encouraging performance on the training set, with high accuracy values for some features such as cOPN, cCON and cEXT. This indicates that the model can capture certain trends and correlations within the training data, and it is able to predict these traits with high accuracy on this data. However, the performance on the validation set and test set is relatively lower. This suggests that the model may have difficulty generalizing to new data and may suffer from overfitting to training data.

7.3. Third Approach

In our ensemble learning approach, we used the stacking technique with logistic regression as a meta-learner model.

Stacking (Figure 13) is a more advanced ensemble method which consists of combining several models in a hierarchical manner. Instead of simply aggregating the predictions from individual models, stacking uses a meta-learner model to take the predictions from the base models as input and produce a final prediction.

The meta-learner model is typically trained using a supervised approach on a validation dataset. Stacking is effective for combining the strengths of different models and improving the generalization of the final model.

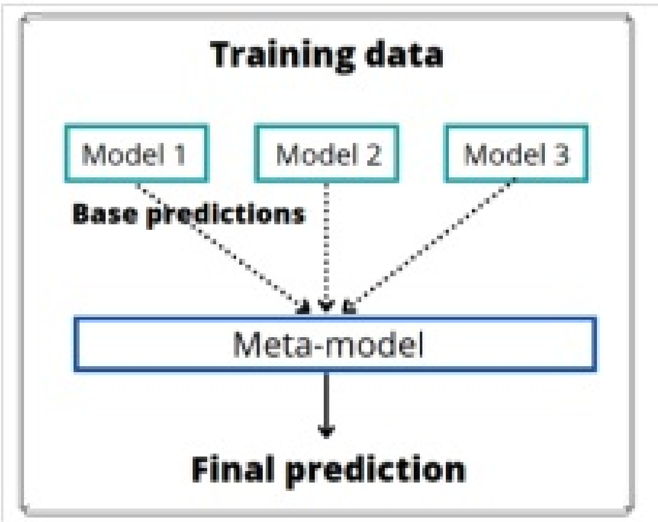

Figure 13. General architecture of the stacking.

7.3.1. First Ensemble

The first Ensemble combines CNN, BERT and the extracted features with LSTM, BERT, and the extracted features. It uses as meta-model the logistic regression as presented in Figure 14.

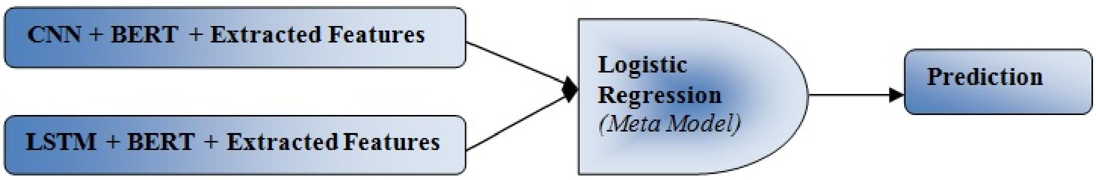

Figure 14. The structure of the first Ensemble.

Table 10 presents the different accuracy values for each personality trait corresponding to the implementation of the first Ensemble.

Table 10. Accuracy values for the Ensemble 1.

	Leaky ReLU	Tanh	Sigmoid
cEXT	53.09%	54.71%	53.84%
cNEU	58.29%	57.95%	58.09%
cAGR	50.20%	55.01%	54.04%
cCON	55.46%	56.83%	55.66%
cOPN	55.26%	61.09%	61.94%

Figure 15 presents the histogram generated from the Table 10. It corresponds the accuracy values for different activation functions used in the first Ensemble.

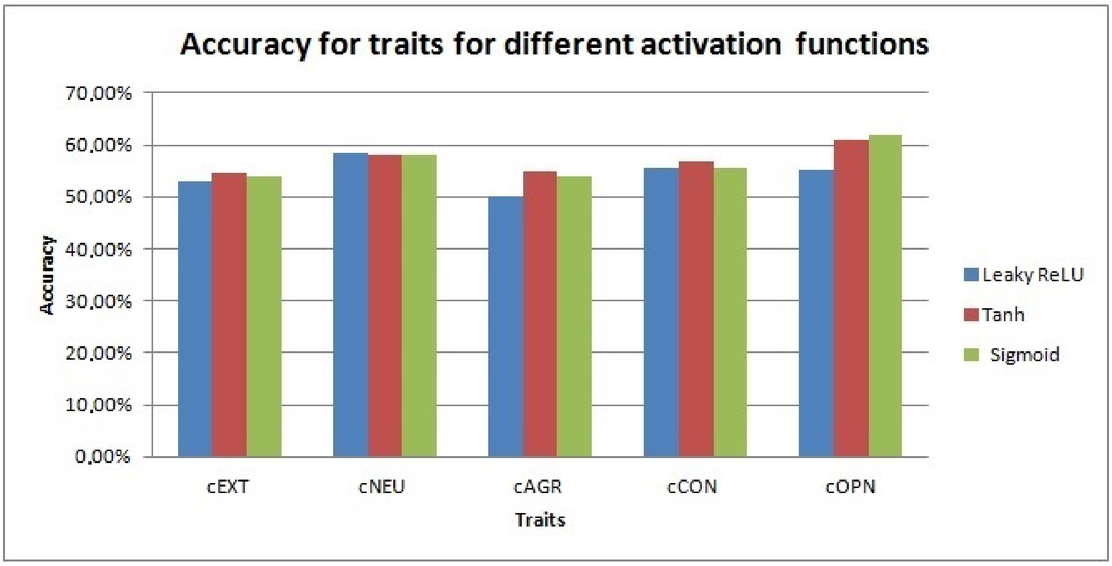

Figure 15. Generated histogram from Table 10.

We can notice that the performance varies depending on the activation function used in the models. Each activation function has its advantages and disadvantages, and it can have a significant impact on the model's ability to capture trends and patterns in the data. In our case, the Tanh activation function has the best results for the personality traits cEXT, cOPN and cCON, with accuracies above 55%. This suggests that Tanh allows the model to better represent nonlinear relationships in the data, which is important for capturing the complex characteristics of personality traits. On the other hand, the Sigmoid activation function gave the best performance for the cNEU and cAGR personality traits, with accuracies slightly above 58%. The Sigmoid function is often used for binary classification, and it seems to work well for certain personality traits that can be more easily differentiated into two distinct categories. In contrast, the Leaky ReLU activation function performed worse than the other two functions for all personality traits, with accuracies ranging from 50% to 56%. This may indicate that Leaky ReLU does not allow the model to effectively capture some nuances and variations in the data. In conclusion, the choice of activation function can have a significant impact on model performance. In our case, the Sigmoid activation function gave the best results for different personality traits.

7.3.2. Second Ensemble

The second Ensemble combines CNN, W2V and the extracted features with LSTM, W2V, and the extracted features. It uses as meta-model the logistic regression as presented in Figure 16.

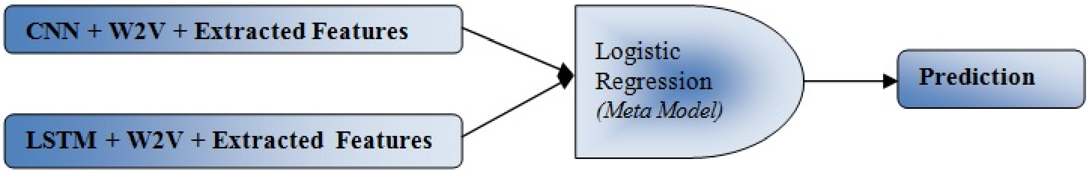

Figure 16. The structure of the second Ensemble.

Table 11 presents the different accuracy values for each personality trait corresponding to the implementation of the second Ensemble.

Table 11. Accuracy values for the Ensemble 2.

	Leaky ReLU	**Tanh**	**Sigmoid**
cEXT	51.77%	53.29%	51.26%
cNEU	55.92%	57.95%	53.79%
cAGR	55.72%	52.88%	52.58%
cCON	53.90%	57.34%	54.00%
cOPN	55.52%	55.72%	54.60%

Figure 17 presents the histogram generated from the Table 11. It corresponds the accuracy values for different activation functions used in the second Ensemble.

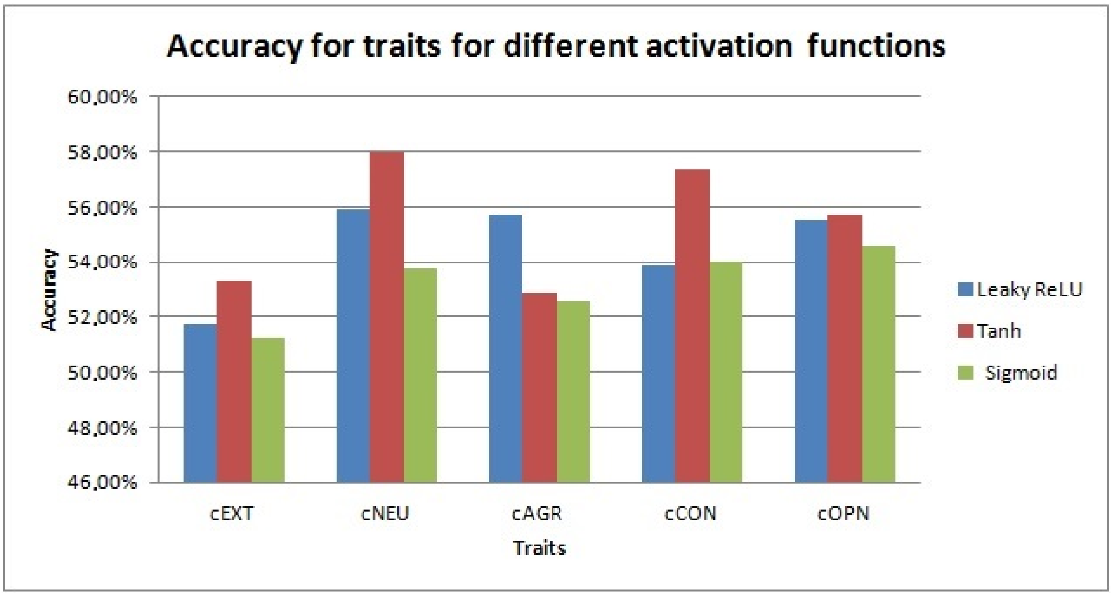

Figure 17. Generated histogram from Table 11.

7.4. Discussion

In this section, we will compare the different scenarios based on the obtained results.

So, we conclude that the best model is W2V-CNN that achieved the highest values for the different personality traits (cEXT: 100%, cNEU: 87.39%, cAGR: 96.26%, cCON : 96.55%, and cOPN: 81.47%). It has even better results compared to the state of the art works. Indeed, W2V presents several advantages compared to BERT. We can mention its simplicity and rapidity. It can provide meaningful embeddings even with smaller training datasets, whereas BERT requires a substantial amount of training data to perform well. Also, Word2Vec embeddings are relatively interpretable, as each word is represented as a fixed-size vector. This can make it easier to understand and debug the representations learned by the model. Also, CNN presents some advantages compared to LSTM such as the parallelization since CNN is highly parallelizable. The convolution layers operate independently on different regions of the input which lead to faster training and inference times compared to LSTM.

We can conclude, also, that the choice of activation function can have a significant impact on model performance where the value of accuracy change according the used function.

8. Deployment

To deploy our solution, we used Django (https://www.djangoproject.com/,accessed on 05/12/2023) which is a high-level Python web framework.

The simple and user-friendly interface of this page (Figure 18) makes it easy to choose between real-time text entry or file loading for personality analysis based on the big five traits. This efficient design aims to make the classification process quick and accessible, while providing a variety of options to fit your specific needs.

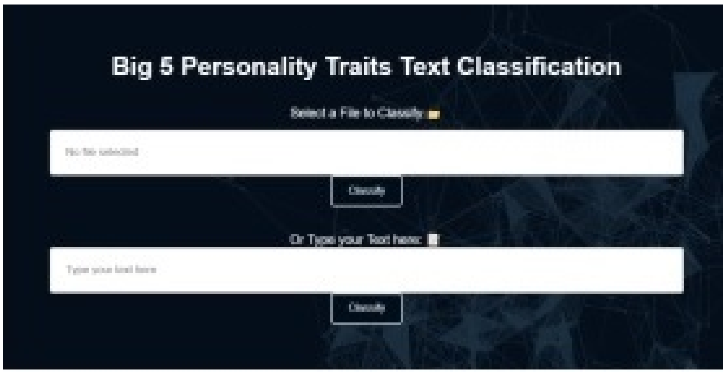

Figure 18. The first user-friendly interface.

After submitting the text for classification, the second page of results is displayed, where a detailed analysis of the five personality traits is presented. Here are the key elements of this page:

- **Detailed Explanation of Traits (Figure 19)**: For each personality trait, an in-depth explanation is given. It explores the associated characteristics and behaviors. These explanations provide a comprehensive overview of each trait for better understanding.

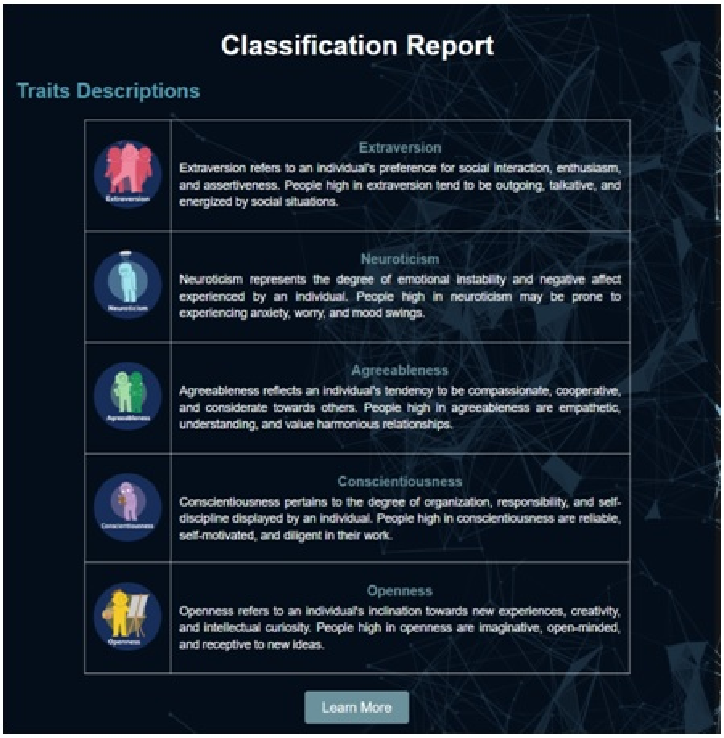

Figure 19. Detailed explanation of traits Interface.

- **Classification Table (Figure 20)**: A table displays the percentages of each personality trait for the submitted text. This clear visualization allows to more understand the distribution of lines in the input text (Figure).

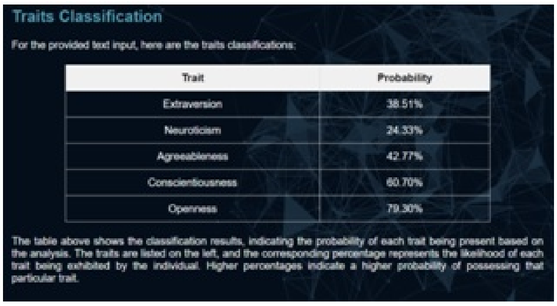

Figure 20. Classification table Interface.

- **Trait Radar Chart (Figure 21)**: A radar chart is also presented, graphically illustrating the levels of each personality trait. This visual representation helps quickly grasp the relative strengths and weaknesses of each trait.

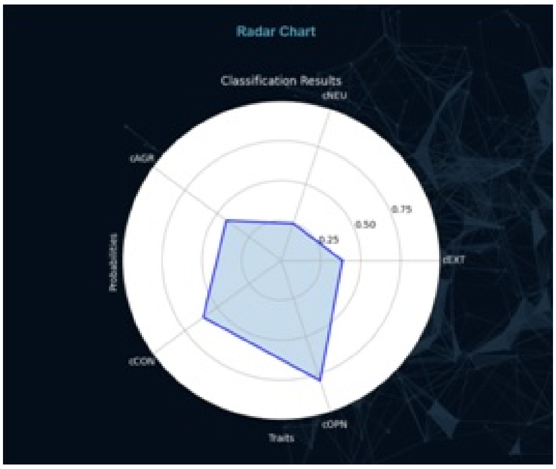

Figure 21. Trait radar chart Interface.

- **Classification Pie Chart (Figure 22)**: A pie chart shows the overall distribution of the five personality traits for the entire text. This offers a global perspective on the personality analyzed.

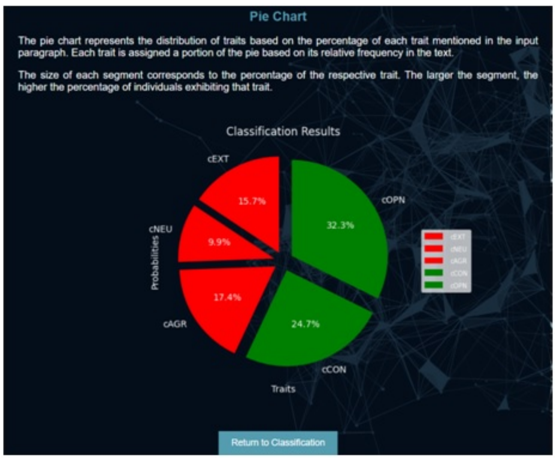

Figure 22. Classification pie chart Interface.

9. Conclusions

Personality trait detection have gained significant importance in both academic research and practical applications across various fields It can be useful in a verity of domains such as the marketing where it gives the possibility to personalize the products according to the client' personality.

To achieve this goal, we propose combining NLP, with DL algorithms which are two branches of artificial intelligence offering several methods, models and algorithms to facilitate the user work. The first one satisfies the requirement for natural language interaction between the user and the machine. The second one is based on multi-layered neural networks to simulate the human brain function recognize complex patterns in structured and unstructured data.

We used for this purpose a well-known dataset that organized the personality traits into five groups Openness, Conscientiousness, Extroversion, Agreeableness and Neuroticism that are called OCEAN in abbreviate.

Concerning the DL algorithms, we evaluated three different approaches. In the first one we implemented : BERT-CNN, BERT-LSTM, W2V-CNN and W2V-LSTM. In the second one, we took the models of the first approach and we added more features that were extracted from NRC Word-Emotion. The last approach is based on Ensemble models.

We concluded that Word2Vec-CNN has the highest accuracy value, also, the choice of activation function can have a significant impact on model performance. Compared to many literature works, our model ensures good results.

As perspectives, we will extend this work to deal with Arabe standard moderne datasets, and some Arabic dialects such as Tunisian dialect.

Author Contributions

The authors contributed equally. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding

Conflicts of Interest

Author declares no conflict of interest

Data Availability Statement

The data that support the findings of this study are openly available in:

- Essays: https://www.kaggle.com/datasets/manjarinandimajumdar/essayscsv
- NRC Emotion Lexicon: https://www.kaggle.com/datasets/wjburns/nrc-emotion-lexicon

References

1. Lauriola, I., Lavelli, A., and Aiolli, F. (2022). An introduction to deep learning in natural language processing: Models, techniques, and tools. *Neurocomputing, 470*, 443-456.
2. Khurana, D., Koli, A., Khatter, K., and Singh, S. (2023). Natural language processing: State of the art, current trends and challenges. *Multimedia tools and applications, 82*(3), 3713-3744.
3. Corr, P. J., and Matthews, G. (Eds.). (2020). *The Cambridge handbook of personality psychology*. Cambridge University Press.
4. Feizi-Derakhshi, A. R., Feizi-Derakhshi, M. R., Ramezani, M., Nikzad-Khasmakhi, N., Asgari-Chenaghlu, M., Akan, T., and Jahanbakhsh-Naghadeh, Z. (2022). Text-based automatic personality prediction: A bibliographic review. *Journal of Computational Social Science*, 5(2), 1555-1593.
5. Cattell, R. B., Eber, H. W., and Tatsuoka, M. M. (1970). *Handbook for the sixteen personality factor questionnaire (16 PF) in clinical, educational, industrial, and research psychology, for use with all forms of the test*. Ill Institute for Personality and Ability Testing.
6. Eysenck, H., Eysenck, S. B., and Giuliett. (1975). *Manual of the Eysenck personality questionnaire*. Hodder and Stoughton.
7. Briggs, K. C., and Myers, I. B. (1976). *Myers–Briggs type indicator*. Consulting Psychologists Press.
8. John, O. P., and Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives—PsycNET. In *Handbook of personality: Theory and research* (pp. 102–138). Guilford Press.
9. Mehta, Y., Majumder, N., Gelbukh, A., and Cambria, E. (2020). Recent trends in deep learning based personality detection. *Artificial Intelligence Review*, 53, 2313 - -2339.
10. Yang HC, Huang ZR (2019) Mining personality traits from social messages for game recommender systems. Knowledge-based Systems 165:157–168

11. Yin H, Wang Y, Li Q, Xu W, Yu Y, Zhang T (2018) A network-enhanced prediction method for automobile purchase classification using deep learning
12. Majumder N, Poria S, Peng H, Chhaya N, Cambria E, Gelbukh A (2019) Sentiment and sarcasm classification with multitask learning. IEEE Intel Syst 34(3):38–43
13. Mohades D. F., Sadr, H., and Tarkhan, M. (2022). Contextualized multidimensional personality recognition using combination of deep neural network and ensemble learning. *Neural Processing Letters*, 54(5), 3811-3828.
14. Kunte, A., and Panicker, S. (2019). Personality prediction of social network users using ensemble and XGBoost classifiers. In *2nd international conference on computing analytics and networking (ICCAN), Bhubhaneshwar*. 5, 133–140
15. Remaida, A., Abdellaoui, B., Moumen, A., and El Idrissi, Y. E. B. (2020). Personality traits analysis using Artificial Neural Networks: A literature survey. In *2020 1st international conference on innovative research in applied science, engineering and technology (IRASET)* (pp. 1-6). IEEE.
16. Sadr, H., Pedram, M. M., and Teshnehlab, M. (2019). A robust sentiment analysis method based on sequential combination of convolutional and recursive neural networks. *Neural processing letters*, 50, 2745-2761.
17. Sadr H, Pedram MM, Teshnehlab M (2020) Multi-view deep network: a deep model based on learning features from heterogeneous neural networks for sentiment analysis. IEEE Access 8:86984–86997
18. Rahman, M. A., Al Faisal, A., Khanam, T., Amjad, M., and Siddik, M. S. (2019). Personality detection from text using convolutional neural network. In 2019 1st international conference on advances in science, engineering and robotics technology (ICASERT) (pp. 1–6). IEEE.
19. Sun, X., Liu, B., Cao, J., Luo, J., and Shen, X. (2018). Who am I? Personality detection based on deep learning for texts. In *2018 IEEE international conference on communications (ICC)*(pp. 1–6). IEEE.
20. Kazameini, A., Fatehi, S., Mehta, Y., Eetemadi, S., and Cambria, E. (2020). Personality trait detection using bagged svm over bert word embedding ensembles. *arXiv preprint arXiv:2010.01309*.
21. Deilami, F. M., Sadr, H., and Nazari, M. (2022). Using machine learning based models for personality recognition. *arXiv preprint arXiv:2201.06248.*
22. Yang, F., Quan, X., Yang, Y., and Yu, J. (2021). Multi-document transformer for personality detection. In *Proceedings of the AAAI conference on artificial intelligence* 35(16), pp. 14221-14229).
23. Asghar, J., Akbar, S., Asghar, M. Z., Ahmad, B., Al-Rakhami, M. S., and Gumaei, A. (2021). Detection and classification of psychopathic personality trait from social media text using deep learning model. *Computational and Mathematical Methods in Medicine*, 1-10.
24. Pennebaker, J. W., and King, L. A. (1999). Linguistic styles: language use as an individual difference. *Journal of personality and social psychology*, 77(6), 1296.
25. Agrawal,S. K. (2024) Metrics to Evaluate your Classification Model to take the right decisions. https://www.analyticsvidhya.com/blog/2021/07/metrics-to-evaluate-your-classification-model-to-take-the-right-decisions/#:~:text=for%20classification%20problems.-,Accuracy,the%20total%20number%20of%20predictions.cited05/03/2024.
26. Mkhini,M., Arfaoui, N., Sidibe. A. S., Baron, B., Serawork Wallelign, S. (2023) Automatic Personality Trait Recognition based on Deep Learning Algorithms. 23rd International Conference on Intelligent Systems Design and Applications (ISDA 2023)