HindiPersonalityNet: Personality Detection in Hindi Conversational Data using Deep Learning with Static Embedding

Akshi Kumar[1], Dipika Jain[2*], Rohit Beniwal[2]
[1]Dept. of Computing & Mathematics, Manchester Metropolitan University, United Kingdom
[2]Dept. of Computer Science & Engineering, Delhi Technological University, New Delhi, India
dipikajain_2k20phdco505@dtu.ac.in

Abstract: Personality detection along with other behavioural and cognitive assessment can essentially explain why people act the way they do and can be useful to various online applications such as recommender systems, job screening, matchmaking, and counselling. Additionally, psychometric NLP relying on textual cues and distinctive markers in writing style within conversational utterances reveal signs of individual personalities. This work demonstrates a text-based deep neural model, HindiPersonalityNet of classifying conversations into three personality categories {ambivert, extrovert, introvert} for detecting personality in Hindi conversational data. The model utilizes GRU with BioWordVec embeddings for text classification and is trained/tested on a novel dataset, शख्सियत (pronounced as Shakhsiyat) curated using dialogues from an Indian crime-thriller drama series, Aarya. The model achieves an F1-score of 0.701 and shows the potential for leveraging conversational data from various sources to understand and predict a person's personality traits. It exhibits the ability to capture semantic as well as long-distance dependencies in conversations and establishes the effectiveness of our dataset as a benchmark for personality detection in Hindi dialogue data. Further, a comprehensive comparison of various static and dynamic word embedding is done on our standardized dataset to ascertain the most suitable embedding method for personality detection.

Keywords: Personality, low-resource, deep learning, word embeddings, NLP

Taxonomy used in Paper:

Abbreviation	Definition
NLP	Natural Language Processing
MBTI	Myers Briggs Type Indicator
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
CNN	Convolutional Neural Network
BiLSTM	Bidirectional Long Short-Term Memory
GRU	Gated Recurrent Unit
GPT	Generative Pre-trained Transformer
RoBERTa	Robustly Optimized BERT Pretraining Approach
ELECTRA	Efficiently Learning an Encoder that Classifies Token Replacements Accurately
T5	Text-to-Text Transfer Transformer
XLM	Cross-lingual Language Model

1. Introduction

NLP (Natural Language Processing) can be used to analyse sentiments and emotions in real-time language data and give insights to an individual's communication style in real-time, such as the use of assertive or passive language, the frequency of interruptions, or the level of engagement in the conversation. Psychometric NLP is a relatively new area of research that has gained popularity in recent years due to the increasing availability of large-scale textual data and the advancement of natural language processing techniques. Application of psychometric NLP to conversational data such as chat logs or transcripts of spoken conversations, has the potential to measure psychological constructs such as personality traits, emotions, and mental health conditions.

Personality traits constitute the most tangible psychometrics for psychological explanation of human behaviour. The study of personality traits using NLP provides a non-invasive, valuable tool for understanding individual differences and developing interventions to promote psychological well-

being in various settings. There are multiple personality dimension theories in literature that offer distinct frameworks to comprehend personality and behaviour. Examples of such theories include the Big Five [1,2], MBTI [3-7], HEXACO [8,9] Personality Inventory, Cattell's Sixteen Personality Factor [10], Eysenck's Personality Theory [11], Enneagram [12], and the Five-Factor Personality Model [13], among others. While these theories may differ in their specific dimensions and categories, they all recognize the importance of introversion, extroversion, and ambiversion as key aspects of personality. The concept of introversion/extroversion was introduced in 1910 by Carl Gustav Jung [14], existing as part of a continuum with each personality type at separate ends of the scale. In 1947, psychologist Eysneck [11] added another category as *ambivert personality* for individuals showing traits of introverted personality in some situations and in others, they behave as an extrovert personality type.

Psychometric NLP can analyse language data to identify linguistic cues associated with the extroversion, introversion, and ambiversion dimensions of personality. Here are some examples of linguistic cues associated with each of these dimensions:

- *Extroversion:* People who score high on the extroversion dimension tend to be outgoing, social, and assertive. Linguistic cues that may indicate high extroversion include the use of more positive emotion words, more words related to social activities, and more assertive language.
- *Introversion:* People who score high on the introversion dimension tend to be quiet, reserved, and reflective. Linguistic cues that may indicate high introversion include the use of more negative emotion words, more words related to introspection and reflection, and fewer words related to social activities.
- *Ambiversion:* Ambiverts have characteristics of both extroversion and introversion and may display a balance of social and introspective behaviours. Linguistic cues that may indicate ambiversion include the use of a mix of positive and negative emotion words, a moderate use of words related to social activities, and a moderate use of introspective language.

. Conversational data can provide valuable insights into a person's personality traits, as the way someone speaks or communicates can reflect their personality. Conversational data can be in the form of chat transcripts, emails, social media messages, or any other written or spoken communication between individuals. By analysing conversational data, it is possible to identify patterns in the language used, such as the frequency of certain words or phrases, the tone of the conversation, and the overall sentiment. These patterns can be used as features to train machine learning models that can predict personality traits. Especially as within the field of psycholinguistic, which studies the interrelation between linguistic factors and mental aspects, researchers have confirmed that individuals' personalities can be revealed by their written language and styles. For example, a person who frequently uses words related to aggression or dominance in their conversation may be classified as having a more dominant personality trait. Similarly, a person who uses more positive words and phrases may be classified as having a more positive and outgoing personality. While several pertinent efforts made to develop datasets and benchmarks for low-resource NLP task including sentiment analysis [15] and sarcasm detection [16], few attempts have recently been made to address the personality detection challenges in languages other than English. These include indigenous and low-resource ones such as Arabic [17], Persian [18,19], Bahasa Indonesia [20], and Indian languages like Hindi [3, 21], and Bengali [22]. Most of the Indian languages are low resource, which means that they have relatively lesser data available for training NLP systems, especially conversational systems. Moreover, the datasets created are based on MBTI and Big-5 frameworks

As individual tendency towards the outer world (extroversion) or the inner world (introversion) can have a significant influence on career choice, relationships and overall lifestyle, this research puts forward a personality detection model, HindiPersonalityNet, for assorting dialogues into three personality categories {ambivert, extrovert, introvert}. A conversational dataset called 'शख्सियत' (pronounced as Shakhsiyat) is created using dialogues from an Indian crime-thriller drama series in Hindi, Aarya. We demonstrate the use of deep neural network with static embedding that can learn and generalize patterns in the data, resulting in accurate and efficient personality detection. We run intensive simulations on the dataset and evaluate several performance metrics. The major contributions of this work can be summarised below:

- Creation of a novel text-based Hindi dataset, शख्सियत dataset (pronounced as Shakhsiyat) with three class labels {extrovert, introvert, ambivert} which can be used as a benchmark dataset for personality detection in Hindi language.
- Implementation and evaluation of HindiPersonalityNet model that uses BioWordVec embeddings with GRU for detecting personality traits in the conversational Hindi dataset.
- Comprehensive comparison of word embeddings for personality detection in Hindi dataset using Psychometric natural language processing.

The rest of the paper is organized as follows. Section 2 discusses some recent work done in the field of personality detection, specifically in Indian languages. Section 3 provides an overview of the fundamental concepts related to static and dynamic word embeddings, as well as deep learning models that are commonly used for NLP tasks. Section 4 delves into the specifics of the Hindi personality detection dataset that was curated for our study and presents the details of HindiPersonalityNet model for the task of personality detection in Hindi conversational data. The experimental results obtained are discussed in Section 5. Section 6 presents the concluding remarks and the scope for future improvement for this work.

2. Related Work

Significant literature account for studies using machine learning (including the state-of-the-art transformer-based models) to detect and predict various psychological traits such as personality, behaviour as well as issues like depression and anxiety using linguistic cues from text. But a large body of work has been reported only in English language, especially from social media postings. Specific to personality detection public datasets and benchmark models have been well-reported and evaluated using MBTI and Big-5 personality traits [23,24]. Recent works also report few studies on personality detection other than English. In 2018, Adi et al. [20] presented optimization techniques for machine learning (ML)-based Bahasa Indonesian automatic personality detection. Fatehi et al. [18] put forward a deep neural network using fastText embeddings with LSTM and attention layers to automate a personality detection system in Persian language. The dataset was collected from Twitter and annotated into MBTI personality traits. The authors also used an interpretability technique to explain the results. In 2022, Anari et al. [19] proposed a lightweight deep convolutional neural network for categorizing personality into 9 types from handwritten Persian text. Salem et al. [17] proposed a ML technique for Egyptian dialect automatic personality detection.

Lately, a small number of studies have been reported on automatic personality detection in Indian languages. In one of the earliest works in personality detection in Hindi text, Singh et al. [25] introduced a psycho-lexical approach where the data from different Hindi novels and sources were collected and further studied based on three major trigunas such as sattvic, rajasic, and tamasic. In 2017, Singh and Raad [26] extended this work with the participation of 1250 different Hindi speaking people of young generation and recorded the personality traits by observing the tri-gunas and the big-5 personality traits. In 2020. Khan et al. [21] proposed a multimodal Hindi conversational dataset with audio, video, and Hinglish utterances (Hindi+English). One of our recently published works [3] reports the use of an ensemble of SVM kernels with soft voting for MBTI personality detection in a novel Hindi text-dataset विशेष चरित्र_MBTI (pronounced as vishesh charitr). Apart from Hindi language, in 2020, Rudra et. al. [22] has reported the use of various ML and deep learning models to detect Big-5 personality traits for informal Bangla transcripts. Our approach demonstrates the effectiveness of deep neural networks with static embeddings for personality detection in Hindi conversational data. GRU is used to model the temporal relationships between the words in the input text, as it allows to capture the nuanced and subtle patterns associated with specific personality traits in textual data. Simultaneously, BioWordVec embeddings are trained on a large corpus of biomedical literature, including scientific texts, articles, and abstracts, which enables them to learn the specific language used in the biomedical domain. This makes them effective at capturing the nuances of language use in specific domains, which can be beneficial for personality detection in conversational data.

3. Preliminaries

Conversational AI is becoming more prevalent, with chatbots and virtual assistants being used for customer service, sales, and other tasks. NLP is a key component of conversational AI, as it enables these systems to understand natural language inputs and generate appropriate responses. In the past, NLP models relied heavily on handcrafted features and rule-based systems, which were limited in their ability to understand natural language. However, with the advent of word embeddings and deep learning, NLP models can now learn from large amounts of data and automatically extract meaningful features that capture the nuances of human language.

3.1. Word Embedding

Word embedding refers to the process of representing words as dense, low-dimensional vectors in a way that captures their semantic and syntactic properties. There are two main categories of word embeddings used in natural language processing (NLP): static and dynamic embeddings. Static embeddings, such as GloVe [27, 28], Crawl [29], Wikipedia [30], BioWordVec [31-32], GoogleNews [33], and PubMed [34], are pre-trained on a large corpus and remain fixed throughout the NLP task. They provide a fast and efficient approach to encoding words with semantic and syntactic information.

In contrast, dynamic embeddings, such as ElMo [35], fastText [36-37], and BERT [4], are generated during the training of a neural network for a specific NLP task and evolve over time. They can capture more context-dependent information, including the latest language trends, and are optimized for the specific NLP task at hand. However, they are computationally more expensive and require more training data than static embeddings. In this study, we perform a comprehensive comparison of word embeddings (static and dynamic) for personality detection in Hindi dataset using Psychometric natural language processing.

3.2. Deep Learning Models

Deep learning models have become increasingly popular in the field of natural language processing (NLP) due to their ability to handle large amounts of data and learn complex relationships between words and phrases. Some commonly used deep learning models for NLP include [38, 39]:

- CNN (Convolutional Neural Network): A neural network model that is commonly used for image recognition but can also be applied to NLP tasks such as text classification and sentiment analysis. The model uses convolutional filters to identify local patterns in the input text.
- GRU (Gated Recurrent Unit): A type of recurrent neural network (RNN) that is designed to overcome the vanishing gradient problem of traditional RNNs. GRUs use gating mechanisms to selectively update and reset the hidden state of the network.
- LSTM (Long Short-Term Memory): Another type of RNN that is designed to handle the problem of vanishing gradients and address the issue of forgetting long-term dependencies. LSTMs use a memory cell that can selectively retain or forget information over time.
- BiLSTM (Bidirectional Long Short-Term Memory): A variant of LSTM that processes input sequences in both forward and backward directions, allowing the model to capture both past and future information.
- Attention: A mechanism that can be added to various neural network architectures, including RNNs and transformers, to allow the model to selectively focus on specific parts of the input text. Attention has been shown to improve the performance of many NLP tasks, including machine translation and sentiment analysis.
- Transformer-based models: The transformer architecture has revolutionized the field of NLP, and researchers are constantly working to develop new and more powerful models. Some of the most recent examples include: GPT (Generative Pre-trained Transformer), RoBERTa (Robustly Optimized BERT Pretraining Approach), T5 (Text-to-Text Transfer Transformer), ELECTRA (Efficiently Learning an Encoder that Classifies Token Replacements Accurately), XLM (Cross-lingual Language Model), RoBERTa (Robustly Optimized BERT Pretraining Approach), XLNet

In this research, we utilized GRU [35, 40] and performed a comparative analysis involving LSTM, BiLSTM and CNN. Fig. 1 showcases the diverse combinations of embeddings and deep learning models employed in this study.

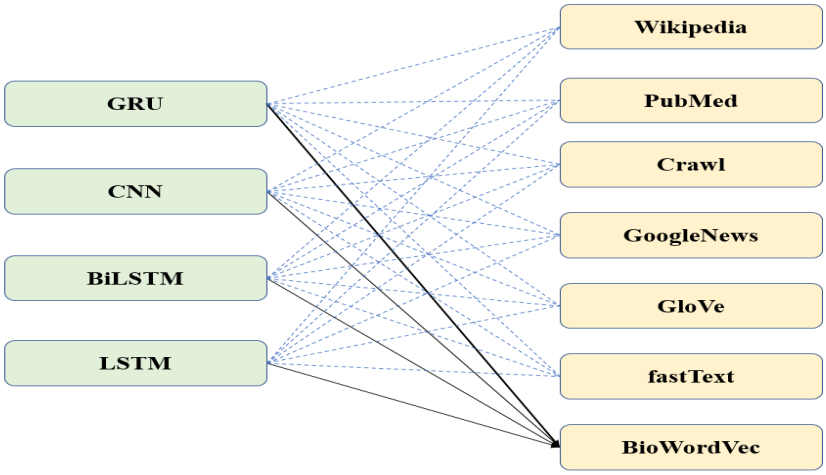

Fig.1. Deep learning and embedding employed

4. Methodology

This section presents the methodology employed to curate the 'Shakhsiyat' dataset and develop the 'HindiPersonalityNet' model, contributing to the advancement of natural language processing (NLP) in the Hindi language and personality detection tasks.

4.1 शख्सियत (pronounced as Shakhsiyat) Dataset

To create the "Shakhsiyat" dataset, which means "personality" in Hindi, we transcribed the dialogues of the Indian crime-thriller drama series called Aarya in Hindi. We manually classified the dialogues into three personality categories: introvert, extrovert, and ambivert, using various differentiators and context to guide our classifications. In order to ensure the accuracy of the annotations, we conducted a validation process in collaboration with a team of two highly knowledgeable and qualified post-graduate students specializing in Psychology from a reputable public university located in Delhi, India. Table 1 provides statistical details of the dataset.

Table 1. Shakhsiyat (शख्सियत) Dataset statistics

Description	Statistics
Total number of instances	6734
Total number of "Extrovert" instances	1710
Total number of "Introvert" instances	495
Total number of "Ambivert" instances	4529

The dataset comprises 6734 rows and 3 columns, with the first row serving as the header. The first column contains the speaker's name, such as Aarya, Veer, Tej, Adi, etc. The second column contains the actual utterance spoken by the speaker, while the last column provides the label for the utterance (extrovert, ambivert, or introvert). Table 2 illustrates some sample dialogues along with their corresponding categories. Our proposed Shakhsiyat (शख्सियत) dataset can serve as a benchmark for personality detection in the Hindi language, as well as other text classification or NLP-based tasks related to Hindi language.

Table 2. Snapshot of Shakhsiyat (शख्सियत) dataset

Speaker (Character)	Utterances (Hindi)	Label
AARYA	**हाय स्वीटी।**	Extrovert
	Hi sweetie.	
ADI	**-माँ, वीर मुझे परेशान कर रहा है!**	Introvert

	Mother, Veer is troubling me!	
AARYA	**मैं जानता हूँ। घबराओ मत।**	Ambivert
	I know. Do not panic.	
AARYA	**घबराना बंद करो।**	Introvert
	stop panicking	
VEER	**आदि, बैठो।**	Ambivert
	Adi, sit down.	
AARYA	**आप अभी उठे हो**	Extrovert
	you just woke up	
VEER	**पापा, उससे कहो कि मुझे दे दो**	Introvert
	papa, tell him to give it to me	
VEER	**वहाँ बैठो!**	Ambivert
	sit there!	
ARU	**माँ, मुझे पैसे चाहिए।**	Extrovert
	Mother, I need money.	

The data extracted from the transcripts of the Hindi series contains various special characters, URLs, etc. This part of the transcript adds non-pertinent noise to the dialogues, and they need to be removed. To tackle this, we perform various text-cleaning techniques used extensively in Natural Language Processing. Since our dataset is in the Hindi language, we also use iNLTK library to better process the obtained Hindi dialogues. The iNLTK library is publicly available with basic in-built functions for Natural Language Processing in Indian Languages.

4.2 The HindiPersonalityNet Model

The model uses BioWordVec embedding to train GRU (Gated Recurrent Unit). BioWordVec embeddings are trained on a large corpus of biomedical literature, including scientific texts, articles, and abstracts, which enables them to learn the specific language used in the biomedical domain. This makes them effective at capturing the nuances of language use in specific domains, which can be beneficial for personality detection in conversational data. GRU is a type of recurrent neural network (RNN) architecture that is capable of processing sequential data. It is designed to overcome some of the limitations of traditional RNNs, such as the vanishing gradient problem, which can make it difficult to train deep models. GRUs use a gating mechanism to selectively update the hidden state at each time step, allowing them to capture long-term dependencies in the input sequence. To build the NLP model, we first initialize the GRU with random weights and then train it on the BioWordVec embeddings. The model takes in a sequence of word embeddings as input, which are processed by the GRU layer. The output of the GRU layer is then fed into a fully connected layer for the classification tasks. Fig.2 depicts the system architecture of the proposed model.

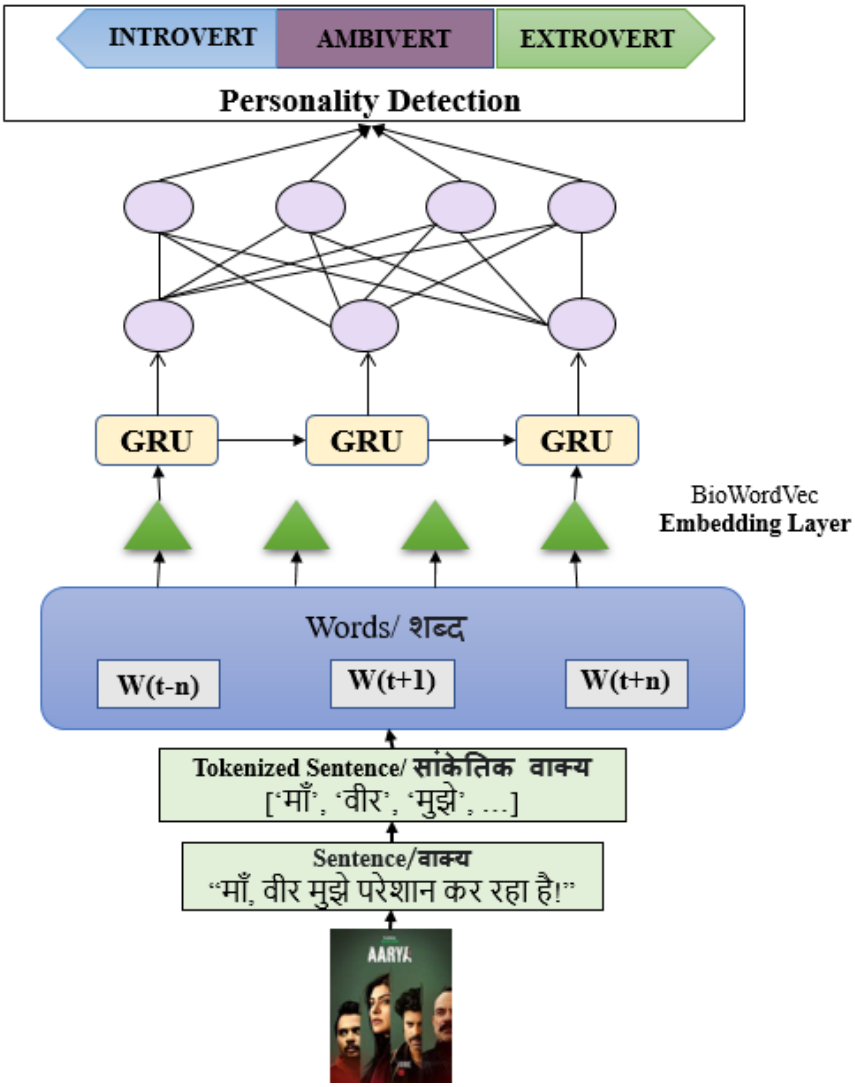

Fig.2. The HindiPersonalityNet Model

5. Results and Discussion

The evaluation metrics utilized to estimate the performance of the proposed HindiPersonalityNet model include accuracy, precision, recall, and F1-score [3]. These metrics provide a comprehensive measure of the model's efficacy. While accuracy is considered, F1-score is also incorporated to address the sensitivity of accuracy towards imbalanced datasets and to assess the classifier's performance in terms of specificity and sensitivity. To establish a benchmark, the results of the proposed model are compared with those of the state-of-the-art model. Table 3 showcases the performance results of the proposed model, providing a comprehensive overview of its effectiveness.

Table 3. Comparative Performance Analysis of Hindi Personality Detection Models

Model	Dataset	Accuracy	F1-score	Precision	Recall
HindiPersonalityNet	शख्सियत dataset	0.739	0.701	0.704	0.739
KBSVE-P [3]	विशेष चरित्र_MBTI dataset	0.668	0.679	0.674	0.701

This model achieved the highest accuracy of 0.739 and recall of 0.739, indicating that it performed well in correctly identifying positive instances. It also had a relatively high F1-score of 0.701 and precision of 0.704. These results suggest that the HindiPersonalityNet model with GRU-BioWordVec

embeddings is effective in capturing the patterns and features necessary for accurate predictions in the given task.

Table 4 presents a comparative performance analysis of two different models for Hindi and Bangla Personality Detection. The table provides information on the models, the datasets they were trained on, and the corresponding evaluation metrics including accuracy, F1-score, precision, and recall.

Table 4. Comparative Analysis of Hindi and Bangla Personality Detection Models

Model	Dataset	Accuracy	F1-score	Precision	Recall
HindiPersonalityNet	शख्सियत dataset	0.739	0.701	0.704	0.739
C-LSTM [22]	Bangla Personality Trait dataset	0.370	0.370	0.370	0.360

By comparing the performance metrics of the two models, it can be observed that the HindiPersonalityNet model achieved higher accuracy, F1-score, precision, and recall values compared to the C-LSTM model trained on the Bangla Personality Trait dataset. This analysis provides insights into the relative performance of the two models for personality detection in Hindi and Bangla languages.

Fig.3 presents the performance comparison of all 3 Indian language personality detection models.

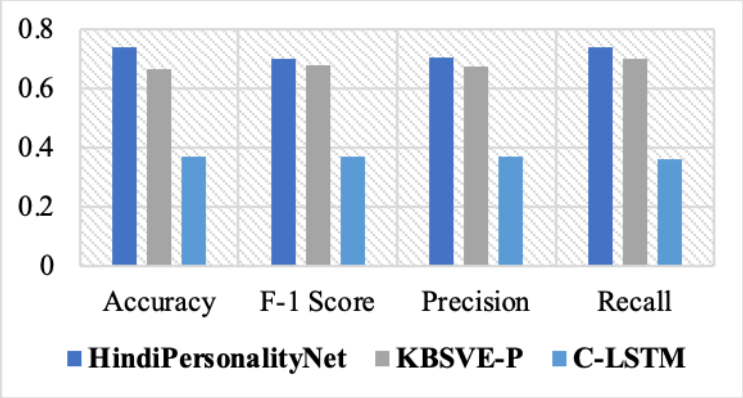

Fig.3. Performance comparison of Indian Language personality detection

The results presented in the table 5 compare the performance of different models using a combination of embedding and deep learning techniques, as illustrated in Fig 1.

Table 5. Comparative Analysis of various deep learning models with BioWordVec

Model	Accuracy	F1-score	Precision	Recall
HindiPersonalityNet (GRU-BioWordVec)	0.739	0.701	0.704	0.739
CNN BioWordVec	0.677	0.680	0.684	0.677
BiLSTM BioWordVec	0.679	0.675	0.674	0.679
LSTM BioWordVec	0.665	0.662	0.659	0.665

The CNN-based model achieved an accuracy of 0.677, which is lower than that of HindiPersonalityNet. However, it demonstrated a balanced F1-score of 0.680, precision of 0.684, and recall of 0.677. These results indicate that the CNN model performed reasonably well, but it fell short of the performance achieved by the HindiPersonalityNet model. Both BiLSTM BioWordVec and LSTM BioWordVec models exhibited similar performance. The BiLSTM BioWordVec achieved an accuracy of 0.679, an F1-score of 0.675, precision of 0.674, and recall of 0.679. On the other hand, the LSTM BioWordVec had an accuracy of 0.665, an F1-score of 0.662, precision of 0.659, and recall

of 0.665. These models performed slightly lower than the HindiPersonalityNet and CNN models but still provided reasonably good results.

Table 6 presents a detailed comparative analysis of the HindiPersonalityNet model when combined with various deep learning models (CNN, LSTM, BiLSTM, GRU) and word embeddings (Crawl, fastText, GloVe, GoogleNews, PubMed, Wikipedia). The training and testing accuracies are reported for each combination, providing insights into how well the model performs on both the training and testing datasets.

Table 6. Accuracy comparison of embedding and deep learning combinations

Embedding + Deep Learning Model	Training Accuracy	Testing Accuracy
Crawl CNN	0.689	0.719
Crawl LSTM	0.735	0.724
Crawl GRU	0.742	0.721
Crawl BiLSTM	0.725	0.727
fastText CNN	0.757	0.724
fastText LSTM	0.821	0.722
fastText GRU	0.737	0.731
fastText BiLSTM	0.751	0.731
Glove CNN	0.707	0.727
Glove LSTM	0.749	0.725
Glove GRU	0.737	0.738
Glove BiLSTM	0.745	0.732
GoogleNews CNN	0.757	0.692
GoogleNews LSTM	0.853	0.698
GoogleNews GRU	0.766	0.725
GoogleNews BiLSTM	0.737	0.717
PubMed CNN	0.755	0.719
PubMed LSTM	0.784	0.727
PubMed GRU	0.752	0.731
PubMed BiLSTM	0.748	0.734
Wikipedia CNN	0.762	0.724
Wikipedia LSTM	0.752	0.712
Wikipedia GRU	0.764	0.721
Wikipedia BiLSTM	0.780	0.716

The table offers insights into the performance of different combinations of deep learning models and word embeddings for the HindiPersonalityNet model. It helps identify the most effective combinations in terms of accurately classifying instances and provides valuable information for selecting appropriate models and embeddings for similar tasks. By analysing the testing accuracies across different models, it is observed that fastText LSTM achieves the highest testing accuracy of 0.722, indicating its effectiveness in accurately classifying instances. The table also highlights the influence of different word embeddings on the model's performance. For instance, when combined with the GloVe embedding, the GRU model achieves a relatively high testing accuracy of 0.738. This suggests that the GloVe embedding provides valuable semantic information for the model to make accurate predictions.

The graph in fig.4 depicts the accuracy comparison of the seven embedding (Crawl, fastText, GloVe, GoogleNews, PubMed, Wikipedia and BioWordVec) when used with GRU. It can be observed that the HindiPersonalityNet model that uses BioWordVec with GRU outperforms the other models.

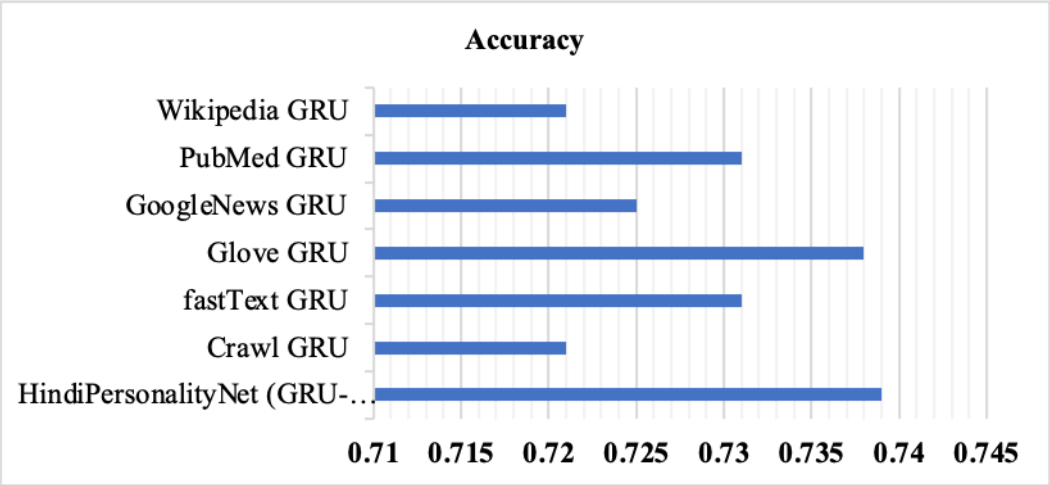

Fig.4. Accuracy comparison for embeddings with GRU

6. Conclusion and future work

In recent years, the field of natural language processing (NLP) has witnessed significant advancements in various domains and languages. In the context of Hindi language analysis, the development of comprehensive datasets and effective models becomes essential for enabling sophisticated NLP applications. This paper presents two significant contributions: the 'Shakhsiyat' dataset and the 'HindiPersonalityNet' model. The 'Shakhsiyat' dataset is curated by transcribing dialogues from the popular Indian crime-thriller drama series 'Aarya' in Hindi, meticulously classified into introvert, extrovert, and ambivert personality categories. With 6734 instances, this dataset provides a benchmark for personality detection in the Hindi language. Additionally, the 'HindiPersonalityNet' model utilizes BioWordVec embeddings and a GRU architecture, trained on a large corpus of biomedical literature, to capture domain-specific language nuances. This model demonstrates its efficacy in personality detection tasks and showcases its potential for conversational data analysis. Together, the 'Shakhsiyat' dataset and the 'HindiPersonalityNet' model contribute to the advancement of NLP in the Hindi language and offer valuable resources for researchers and practitioners in this field.

While this paper presents significant contributions to NLP in Hindi, there are several avenues for future work and improvement. The 'Shakhsiyat' dataset provides a solid foundation, but its further expansion with additional diverse dialogues and annotations from different sources can enhance its representativeness and applicability. Also, the 'HindiPersonalityNet' model shows promising results, there is scope for improvement through architectural modifications, exploring different deep learning models, or incorporating more advanced techniques such as attention mechanisms or transformer-based architectures. Further, integrating other modalities such as audio and visual cues along with text can lead to more comprehensive personality detection models, enabling a deeper understanding of human behaviour in conversations. The current classification into introvert, extrovert, and ambivert categories provides a broad categorization and therefore as another direction of future work could focus on developing models that can detect and differentiate more nuanced personality traits and dimensions. Lastly, the practical deployment of the HindiPersonalityNet model in real-world applications, such as chatbots, virtual assistants, or social media analysis, warrants further investigation to assess its effectiveness and usability.

References

1. Yoneda, T., Lozinski, T., Turiano, N., Booth, T., Graham, E. K., Mroczek, D., & Terrera, G. M. (2023). The Big Five personality traits and allostatic load in middle to older adulthood: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, 105145.
2. Ong V, Rahmanto ADS, Williem, Suhartono D, Nugroho AE, Andangsari EW, et al. Personality Prediction Based on Twitter Information in Bahasa Indonesia. *In: Proceedings of the 2017 Federated Conference on Computer Science and Information Systems*. Prague, Czech Republic; 2017. p. 367–372

3. Kumar, A., Beniwal, R., & Jain, D. (2023). Personality Detection using Kernel-based Ensemble Model for leveraging Social Psychology in Online Networks. *ACM Transactions on Asian and Low-Resource Language Information Processing*.
4. Jain, D., Kumar, A., & Beniwal, R. (2022, July). Personality BERT: A Transformer-Based Model for Personality Detection from Textual Data. In *Proceedings of International Conference on Computing and Communication Networks: ICCCN 2021* (pp. 515-522). Singapore: Springer Nature Singapore.
5. Ren, Z., Shen, Q., Diao, X., & Xu, H. (2021). A sentiment-aware deep learning approach for personality detection from text. *Information Processing & Management*, *58*(3), 102532.
6. Cerkez, N., Vrdoljak, B., & Skansi, S. (2021). A Method for MBTI Classification Based on Impact of Class Components. *IEEE Access*, *9*, 146550-146567.
7. Shafi, H., Sikander, A., Jamal, I. M., Ahmad, J., & Aboamer, M. A. (2021). A Machine Learning Approach for Personality Type Identification using MBTI Framework. *Journal of Independent Studies and Research Computing*, *19*(2), 6-10.
8. Ashton, M. C., & Lee, K. (2007). Empirical, Theoretical, and Practical Advantages of the HEXACO Model of Personality Structure. Personality and Social Psychology Review, 11(2), 150–166. https://doi.org/10.1177/1088868306294907
9. Aghababaei, N., & Arji, A. (2014). Well-being and the HEXACO model of personality. *Personality and Individual Differences*, *56*, 139-142.
10. Ross, C., Orr, E. S., Sisic, M., Arseneault, J. M., Simmering, M. G., & Orr, R. R. (2009). Personality and motivations associated with Facebook use. *Computers in human behavior*, *25*(2), 578-586.
11. Davidson IJ. The ambivert: A failed attempt at a normal personality. *Journal of the History of the Behavioral Sciences.* 2017 Sep;53(4):313-31.
12. Riso, D. R., & Hudson, R. (2000). Understanding the enneagram: The practical guide to personality types. *Houghton Mifflin Harcourt*.
13. Montag, I., & Levin, J. (1994). The five-factor personality model in applied settings. *European Journal of Personality*, 8(1), 1-11.
14. Guilford JP, Braly KW. Extroversion and introversion. *Psychological Bulletin*. 1930 Feb;27(2):96.
15. Kumar, A., & Albuquerque, V. H. C. (2021). Sentiment Analysis Using XLM-R Transformer and Zero-shot Transfer Learning on Resource-poor Indian Language. *Transactions on Asian and Low-Resource Language Information Processing*, *20*(5), 1-13.
16. Jain, D., Kumar, A., & Garg, G. (2020). Sarcasm detection in mash-up language using soft-attention based bi-directional LSTM and feature-rich CNN. *Applied Soft Computing*, *91*, 106198.
17. Salem MS, Ismail SS, Aref M. Personality traits for Egyptian twitter users dataset. In Proceedings of the *2019 8th international conference on software and information engineering 2019* Apr 9 (pp. 206-211).
18. Fatehi S, Anvarian Z, Madani Y, Mehditabar M, Eetemadi S. MBTI Personality Prediction Approach on Persian Twitter. [28_Paper.pdf (winlp.org)]
19. Anari MS, Rezaee K, Ahmadi A. TraitLWNet: a novel predictor of personality trait by analyzing Persian handwriting based on lightweight deep convolutional neural network. *Multimedia Tools and Applications*. 2022 Mar;81(8):10673-93.
20. Adi GY, Tandio MH, Ong V, Suhartono D. Optimization for automatic personality recognition on Twitter in Bahasa Indonesia. *Procedia Computer Science*. 2018 Jan 1; 135:473-80.
21. Khan SN, Leekha M, Shukla J, Shah RR. Vyaktitv: A multimodal peer-to-peer hindi conversations-based dataset for personality assessment. In2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM) 2020 Sep 24 (pp. 103-111). IEEE.
22. Rudra U, Chy AN, Seddiqui MH. Personality traits detection in Bangla: A benchmark dataset with comparative performance analysis of state-of-the-art methods. *In2020 23rd International Conference on Computer and Information Technology (ICCIT) 2*020 Dec 19 (pp. 1-6). IEEE.
23. Mehta Y, Majumder N, Gelbukh A, Cambria E. Recent trends in deep learning-based personality detection. *Artificial Intelligence Review*. 2020 Apr; 53:2313-39.
24. Vásquez RL, Ochoa-Luna J. Transformer-based Approaches for Personality Detection using the MBTI Model. *In2021 XLVII Latin American Computing Conference (CLEI) 2021 Oct 25* **(**pp. 1-7). IEEE.
25. Singh JK, Misra G, De Raad B. Personality structure in the trait lexicon of Hindi, a major language spoken in India. *European Journal of Personality*. 2013 Nov;27(6):605-20.
26. Singh JK, De Raad B. The personality trait structure in Hindi replicated. *International Journal of Personality Psychology.* 2017 Jun 29;3:26-35.
27. Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. *In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532-1543).

28. Kumar, A., Srinivasan, K., Cheng, W. H., & Zomaya, A. Y. (2020). Hybrid context enriched deep learning model for fine-grained sentiment analysis in textual and visual semiotic modality social data. *Information Processing & Management*, 57(1), 102141.
29. Sarma, P. K., Liang, Y., & Sethares, W. A. (2018). Domain adapted word embeddings for improved sentiment classification. *arXiv preprint arXiv:1805.04576*.
30. Sheehan, E., Meng, C., Tan, M., Uzkent, B., Jean, N., Burke, M., ... & Ermon, S. (2019, July). Predicting economic development using geolocated wikipedia articles. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 2698-2706).
31. Zhang, Y., Chen, Q., Yang, Z., Lin, H., & Lu, Z. (2019). BioWordVec, improving biomedical word embeddings with subword information and MeSH. *Scientific data*, 6(1), 52.
32. Wang, S., Tseng, B., & Hernandez-Boussard, T. (2021). Development and evaluation of novel ophthalmology domain-specific neural word embeddings to predict visual prognosis. *International journal of medical informatics*, *150*, 104464.
33. Wang, Y., Liu, S., Afzal, N., Rastegar-Mojarad, M., Wang, L., Shen, F., ... & Liu, H. (2018). A comparison of word embeddings for the biomedical natural language processing. *Journal of biomedical informatics*, *87*, 12-20.
34. Sharma, S., & Daniel Jr, R. (2019). Bioflair: Pretrained pooled contextualized embeddings for biomedical sequence labeling tasks. *arXiv preprint* arXiv:1908.05760.
35. Kumar, A., & Sachdeva, N. (2022). A Bi-GRU with attention and CapsNet hybrid model for cyberbullying detection on social media. *World Wide Web*, 25(4), 1537-1550.
36. Jain, D. K., Kumar, A., & Sangwan, S. R. (2022). TANA: The amalgam neural architecture for sarcasm detection in indian indigenous language combining LSTM and SVM with word-emoji embeddings. *Pattern Recognition Letters*, 160, 11-18.
37. Hu, S., Kumar, A., Al-Turjman, F., Gupta, S., & Seth, S. (2020). Reviewer credibility and sentiment analysis based user profile modelling for online product recommendation. *IEEE Access*, 8, 26172-26189.
38. Ni, J., Young, T., Pandelea, V., Xue, F., & Cambria, E. (2023). Recent advances in deep learning based dialogue systems: A systematic survey. *Artificial intelligence review*, 56(4), 3055-3155.
39. Dodge, J., Gurevych, I., Schwartz, R., Strubell, E., & van Aken, B. (2023). Efficient and Equitable Natural Language Processing in the Age of Deep Learning (Dagstuhl Seminar 22232). *In Dagstuhl Reports* (Vol. 12, No. 6). Schloss Dagstuhl-Leibniz-Zentrum für Informatik.
40. Dey, R., & Salem, F. M. (2017, August). Gate-variants of gated recurrent unit (GRU) neural networks. *In 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS*) (pp. 1597-1600). IEEE.