A Comparative Study of Personality Prediction Using Machine Learning and Deep Learning Approaches

Atharva Chouthai
Computer Science
Virginia Tech
Falls Church
atharvachouthai@vt.edu

Anvita Karne
Computer Science
Virginia Tech
Falls Church
anvitakarne@vt.edu

1. ABSTRACT

Understanding personality through text has become increasingly important in areas such as personalized marketing, recruitment, and mental health assessment. This project is motivated by the need for systems that can naturally and accurately analyze personality traits from written language, especially in contexts like social media and emotionally nuanced content. Traditional machine learning models such as Logistic Regression, Naive Bayes, and Random Forest have demonstrated some success in this area but rely heavily on manually engineered features, limiting their ability to capture the complex, contextual, and emotional nuances of text. In this study, we explore and compare traditional models with advanced deep learning architectures, including Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), and DistilBERT. We enhance these models with techniques like pre-trained embeddings (e.g., GloVe) and emotion lexicons (e.g., EmoLex) to capture semantic and emotional richness. Additionally, we integrate attention mechanisms into Bi-LSTM to enable the model to focus on critical features in the text. Using the MBTI dataset, which includes essays and social media posts, we aim to identify patterns between writing styles and personality traits. Our experiments reveal that deep learning models significantly outperform traditional approaches. In particular, the Bi-LSTM + Attention model achieved the highest accuracy of 96% and lowest validation loss of 13% out of all the other three deep learning models that we implemented, demonstrating its ability to capture context and emotional depth using the attention mechanism. DistilBERT, despite computational constraints, also showcased strong potential. This work highlights the effectiveness of modern deep learning techniques in personality prediction and provides insights into their applicability in real-world scenarios.

CCS CONCEPTS

Computing methodologies → Natural language processing → Text analysis
Computing methodologies → Machine learning → Deep learning → Neural networks
Information systems → Information retrieval → Text representation and semantics
Applied computing → Psychology → Personality prediction
Computing methodologies → Artificial intelligence → Knowledge representation and reasoning

KEYWORDS

Personality prediction, Text Classification, Machine Learning, Deep Learning, LSTM, BERT, Attention mechanisms, Linguistic Patterns

ACM Reference format:

Atharva Chouthai, Anvita Karne. 2024. A Comparative Study of Personality Prediction Using Machine Learning and Deep Learning Approaches.

2. INTRODUCTION

Personality prediction from text is an intriguing and rapidly evolving field, combining the intricacies of human psychology with the power of machine learning and natural language processing (NLP). The ability to infer personality traits—such as openness, conscientiousness, extraversion, agreeableness, and neuroticism—from written words has vast applications across industries. From enhancing user experiences with personalized recommendations to streamlining recruitment processes by identifying ideal candidates, this task offers immense potential for transforming how we interact with and understand individuals.

Historically, traditional machine learning approaches attempted to solve this problem using handcrafted linguistic features and statistical techniques. While they provided valuable initial insights, these models lacked the capacity to capture the nuanced, contextual, and emotional intricacies of human language. The advent of deep learning has revolutionized personality prediction, enabling models to uncover deeper patterns in text. Techniques such as Long Short-Term Memory (LSTM) networks, Bidirectional LSTMs, and transformer-based architectures like DistilBERT have brought unprecedented accuracy and contextual understanding to this domain.

In this project, we embarked on a journey to combine the best of traditional and modern approaches to tackle the challenges of personality prediction. By integrating advanced feature extraction techniques such as pre-trained word embeddings (e.g., GloVe), and modern deep learning architectures, we aimed to create a system capable of accurately capturing both the semantic and emotional dimensions of text. Our exploration extended to techniques like Bi-LSTM with attention mechanisms, which allowed the model to focus on the most relevant parts of the input, and DistilBERT, which brought the power of transformer-based learning into the task despite computational constraints.

Beyond accuracy, our focus was on addressing critical challenges such as interpretability and class imbalance. To ensure fairness across all personality types, we employed SMOTE to mitigate imbalances and used regularization techniques like dropout and early stopping to prevent overfitting. These efforts not only improved the generalization of our models but also provided insights into how emotional and contextual cues influence personality traits.

This project has been a valuable learning experience, showcasing the trade-offs between computational efficiency, predictive accuracy, and model interpretability. By combining classical machine learning techniques with state-of-the-art deep learning architectures, we aim to contribute to the growing body of knowledge in personality prediction. The insights gained here have far-reaching implications across industries such as mental health, personalized marketing, recruitment, and human-computer interaction. This report captures our progress, the lessons learned, and the potential future directions for further exploration in this fascinating intersection of deep learning, NLP, and personality analysis.

3. LIST OF CONTRIBUTIONS

- **Comprehensive Comparison of Deep Learning Models**: We evaluated advanced architectures such as LSTM, Bi-LSTM, Bi-LSTM with Attention, and DistilBERT, providing insights into their effectiveness for personality prediction tasks and highlighting the strengths of contextual and sequential learning.
- **Integration of Attention Mechanisms**: The inclusion of attention mechanisms in Bi-LSTM improved the model's focus on linguistically significant features, resulting in enhanced accuracy and better performance for underrepresented classes.
- **Addressing Class Imbalances**: By employing SMOTE for balancing class distributions, we ensured fair evaluation across all personality types, significantly improving predictions for minority classes.
- **Transformer-Based Model Evaluation**: We fine-tuned DistilBERT on a subset of the dataset, demonstrating its strong potential even under computational constraints and showcasing the scalability of transformer-based architectures.
- **Effective Overfitting Prevention and Future Directions**: Regularization techniques like dropout and early stopping ensured better generalization for complex models. Additionally, we identified future opportunities to explore large language models (LLMs) for improved contextual understanding.

4. RELATED WORK

The task of predicting personality traits from textual data has gained significant momentum in recent years, driven by advancements in machine learning and the increasing availability of digital text data. Foundational work by Digman [4] introduced the Five-Factor Model, which provided the basis for personality classification. Building on this, Mairesse et al. [5] explored linguistic cues for automatic personality recognition, but their reliance on handcrafted features and traditional machine learning methods limited their ability to capture complex nuances in language.

The advent of deep learning revolutionized personality prediction, enabling models to extract richer and more contextual patterns from text. Poria et al. [3] demonstrated the effectiveness of combining Convolutional Neural Networks (CNNs) with Long Short-Term Memory (LSTM) networks to model intricate dependencies in text. Expanding on this, other works, such as those by Gjurković and Šnajder [1], utilized social media platforms like Reddit to uncover valuable insights into personality traits, while Tandera et al. [2] leveraged Facebook data to build systems capable of predicting user personalities with psycholinguistic and behavioral features. Transformer-based models have further elevated the field by introducing unparalleled contextual understanding. Devlin et al. [12] introduced BERT, a pre-trained transformer model, which significantly advanced tasks requiring deep semantic understanding. Similarly, Vaswani et al. [13] proposed the groundbreaking "Attention is All You Need" architecture, which forms the basis of many state-of-the-art transformer models. Goldberg and Levy [11] emphasized the role of embeddings such as Word2Vec, which capture word semantics effectively, providing a strong foundation for downstream tasks.

Innovative methods integrating emotional and contextual features have also shown promise. Mohammad and Turney [7], [8] developed the Word-Emotion Association Lexicon (EmoLex), which associates words with specific emotions, enhancing text analysis for personality prediction. Jing et al. [9] explored MBTI traits using social media data, demonstrating the value of user-generated content for such tasks. Meanwhile, Patra et al. [10] presented multimodal frameworks, combining text with other data modalities to enrich personality detection. Addressing model optimization, Zeiler [6] introduced ADADELTA, an adaptive learning rate method, which has been pivotal in improving training efficiency for deep learning models. Poria et al. [14] further explored hybrid models combining psycholinguistic and deep learning features, demonstrating improved performance in sentiment and personality tasks.

Our approach builds on these advancements by employing LSTM, Bi-LSTM, and DistilBERT models, integrating attention mechanisms to capture nuanced relationships in text and leveraging techniques like SMOTE to address class imbalance. This comprehensive strategy not only enhances prediction accuracy but also advances the understanding of personality traits in diverse textual contexts.

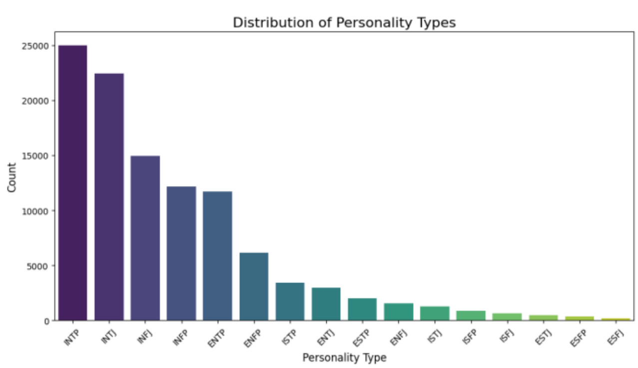
5. DATASET DESCRIPTION

In this project, we utilize the MBTI Personality Types 500 Dataset, which comprises approximately 100,000 preprocessed records of social media posts labeled with the corresponding Myers-Briggs Type Indicator (MBTI) personality types.

The dataset categorizes individuals into 16 distinct personality types, such as INTJ, INFP, and ENFJ, based on four dichotomies: Introversion (I) vs. Extraversion (E), Intuition (N) vs. Sensing (S), Thinking (T) vs. Feeling (F), and Judging (J) vs. Perceiving (P).

A notable challenge with this dataset is the imbalance in class distribution, where certain personality types are underrepresented. This imbalance can adversely affect the performance of machine learning models, leading to biased predictions toward the majority classes. To address this issue, we employ the Synthetic Minority Over-sampling Technique (SMOTE) to augment the minority classes, thereby promoting a more balanced and fair evaluation across all personality types.

By leveraging this enriched dataset, our objective is to capture the linguistic and emotional features that correlate with specific personality traits, ultimately enhancing the predictive performance and generalizability of our models. Below is the figure showcasing the distribution of personalities.



6. MODEL DESCRIPTION

To address the task of personality prediction, we employed a combination of classical machine learning and advanced deep learning techniques. Starting with Naive Bayes as a baseline, we progressively explored neural network architectures, including LSTM, Bidirectional LSTM, and models enhanced with attention mechanisms. Each model was designed to evaluate the effectiveness of different approaches in capturing the intricate patterns present in textual data.

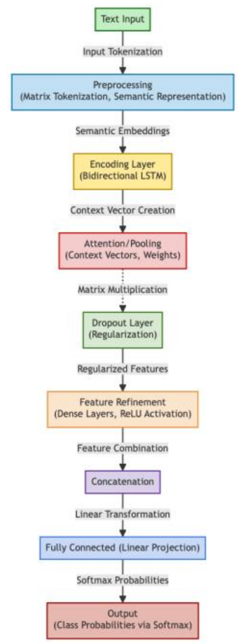


Fig 1: Conceptualized diagram of personality prediction system

6.1 Naive Bayes (Baseline): A Multinomial Naive Bayes variant was employed, as it is well-suited for text data where features represent the frequency of words or terms. Preprocessed text data was vectorized, and the model computed the probability of each word belonging to a particular personality class. The final prediction was made by selecting the class with the highest posterior probability.

6.2 Long Short-Term Memory (LSTM): The Long Short-Term Memory (LSTM) network was implemented as the first deep learning approach to capture sequential dependencies in the textual data. This architecture was selected for its ability to handle long-term dependencies, making it well-suited for personality prediction tasks involving linguistic patterns.

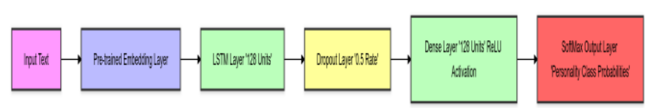


Fig 2: Lstm Architecture

In the LSTM architecture, the input text is tokenized and transformed into dense vector representations using a Pre-trained Word Embedding Layer (dimension: 300). We chose pre-trained embeddings, such as GloVe, to leverage semantic relationships between words, providing a strong foundation for downstream learning. The embeddings are then processed by an LSTM Layer with 128 units, which we selected for its ability to capture long-term dependencies critical for personality prediction. A Dropout Layer with a rate of 0.5 is applied to reduce overfitting, especially given the imbalanced dataset. The refined output is passed to a Dense Layer with 128 units and ReLU activation, which learns higher-level patterns essential for classification. Finally, the SoftMax Layer computes probabilities for the 16 personality types. The model is trained using the Adam Optimizer, chosen for its adaptive learning rate and efficient convergence. We employed Sparse Categorical Cross entropy as the loss function, which is ideal for integer-encoded class labels, reducing computational overhead while maintaining precise gradient calculations. This architecture effectively captures sequential dependencies while ensuring computational efficiency.

6.3 Bi-directional LSTM: To improve upon the sequential modeling capabilities of LSTMs, we implemented a Bidirectional LSTM (Bi-LSTM) network. This approach allowed us to capture context from both past and future sequences, providing a more comprehensive understanding of the text data.

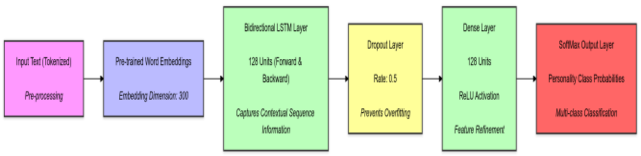


Fig 3: Bi-Lstm Architecture

The Bi-LSTM architecture enhances the LSTM model by incorporating a Bidirectional LSTM Layer, allowing the model to process input sequences in both forward and backward directions. Tokenized text is first passed through a Pre-trained Word Embedding Layer (dimension: 300) to encode semantic relationships. The Bidirectional LSTM Layer, with 128 units, captures contextual dependencies from preceding and succeeding words, offering a richer representation of the input sequence. A Dropout Layer with a rate of 0.5 is applied to prevent overfitting, followed by a Dense Layer with 128 units and ReLU activation to refine features. The final SoftMax Layer generates probabilities for each personality class. The Adam Optimizer was used to stabilize and accelerate training, while Categorical Cross entropy served as the loss function to handle one-hot encoded labels effectively. This architecture improves the understanding of complex linguistic structures, making it suitable for text with nuanced relationships.

6.4 Bi-Lstm with Attention Mechanism: To enhance the Bi-LSTM architecture, we integrated an attention mechanism to focus on the most relevant parts of the input sequence.

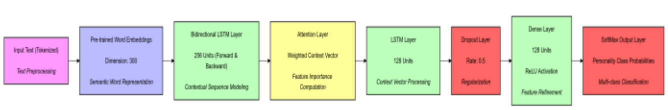


Fig 4: Bi-Lstm with attention mechanism architecture

To enhance the Bi-LSTM model's performance and interpretability, we incorporated an Attention Layer into the architecture. The input text is first tokenized and passed through a Pre-trained Embedding Layer (dimension: 300), which encodes semantic information. The embeddings are processed by a Bidirectional LSTM Layer with 128 units to capture both forward and backward contextual dependencies. The Attention Layer computes a weighted context vector, allowing the model to emphasize the most significant parts of the input text, which is particularly beneficial for identifying personality traits expressed through nuanced or emotionally charged language. A Dropout Layer (rate: 0.5) regularizes the output, which is then refined by a Dense Layer (128 units, ReLU activation). The final SoftMax Layer computes class probabilities. We used the Adam Optimizer for efficient training and Categorical Cross entropy Loss to evaluate predictions against the one-hot encoded labels. The attention mechanism enhanced the model's ability to prioritize critical linguistic features, improving performance and interpretability. These model configurations made this bi-lstm with attention our best performing model for this task.

6.5 DistilBert (Transformer Architecture): To leverage state-of-the-art transformer-based architectures, we implemented DistilBERT, a lightweight version of BERT, for personality prediction. This model was chosen for its ability to handle complex language representations with reduced computational overhead.

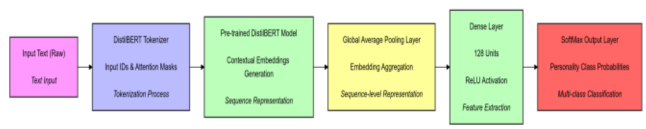


Fig 5: DistilBert (Transformer architecture)

The DistilBERT architecture leverages transformer-based contextual embeddings for personality prediction, offering a balance between computational efficiency and contextual understanding. Tokenized input text is processed through the DistilBERT Encoder, which generates contextualized embeddings for each token. To summarize the sequence, we used a Global Average Pooling Layer, which aggregates the embeddings into a fixed-length representation while retaining essential information. The pooled embeddings are refined using a Dense Layer (128 units, ReLU activation) and then passed to a SoftMax Layer for personality classification. The AdamW Optimizer was employed for fine-tuning the model, ensuring stability and convergence, while Categorical Cross entropy Loss was used to handle the one-hot encoded labels. Despite being trained on a subset of the data due to computational constraints, DistilBERT demonstrated competitive performance, highlighting the potential of transformer-based architectures for text-based personality prediction. We think that the transformer-based model will work more efficiently when trained on a huge amount of dataset.

7. EVALUATION

In the evaluation of these deep learning models, we employed a variety of metrics to provide a comprehensive understanding of their performance. Accuracy gives a straightforward measure of how often the model's predictions match the true labels, making it a useful initial indicator. Precision, Recall, and F1-score delve deeper into the quality of the predictions: Precision focuses on how many of the predictions labeled as a certain class are actually correct, Recall measures how many of the true instances of a class the model successfully detects, and the F1-score, which is the harmonic mean of Precision and Recall, balances the trade-offs between the two. Together, these metrics offer a nuanced view of model effectiveness, particularly in tasks where misclassifications can carry different levels of severity.

When training these classification models, categorical_crossentropy loss was used. This loss function is standard for multi-class classification problems because it directly measures the divergence between the predicted probability distribution over classes and the actual distribution (represented as a one-hot vector for the correct class). Minimizing categorical_crossentropy encourages the model to assign higher probabilities to the correct class and reduce uncertainty in its predictions. In doing so, it supports learning feature representations that lead to improved accuracy, precision, recall, and F1-scores over training iterations.

$$L(y, y\hat{}) = -\sum_{k=1}^{k} (y_k * \log(\hat{y}_k))$$

K: This is the total number of personality categories into which you are classifying individuals. For example, if you have 4 personality categories (A, B, C, D), then $K = 4$

y: This is the true label, represented as a one-hot encoded vector of length K. For the correct personality category, say category$j$, $yj = 1$ and for all other categories $yk = 0$ where $k != j$. Essentially, y is telling us which personality type the individual actually belongs to.

$y\hat{}$ **:** This is the predicted probability distribution over the K personality categories that your model outputs. Each $y\hat{k}$ is the model's estimated probability that the individual belongs to category k . All predicted probabilities sum to 1.

$\log(y\hat{k})$**:** Taking the natural logarithm of the predicted probability emphasizes the model's confidence. If the correct category has a high predicted probability,$y\hat{k}$ is close to 1, and $\log(y\hat{k})$is close to 0 (since $\log(1) = 0$), resulting in a smaller loss. If the correct category is predicted with a very low probability (close to 0), the log term becomes a large negative number, increasing the overall loss.

7.1 EXPERIMENTAL RESULTS

7.1.1 Naive Bayes: The Naive Bayes classifier achieved a validation accuracy of 76%, with macro-averaged F1, precision, and recall scores of 0.77. While it performed well on dominant classes such as ESFJ and ESTP, it struggled with underrepresented classes like INFP and INFJ. The results indicate that Naive Bayes, while efficient, is limited by its independence assumption, which prevents it from capturing deeper relationships in the data. Table 1 highlights the evaluation metrics.

Metrics	Score
Validation Accuracy	0.76
Precision	0.77
Recall	0.76
F1-Score	0.77

7.1.2 Long Short-Term Memory: The LSTM model was the first deep learning approach implemented. It captured sequential dependencies in the text data but struggled with imbalanced classes. The training and validation performance metrics are summarized in Table 2 and the figure displays the training and validation loss for LSTM.

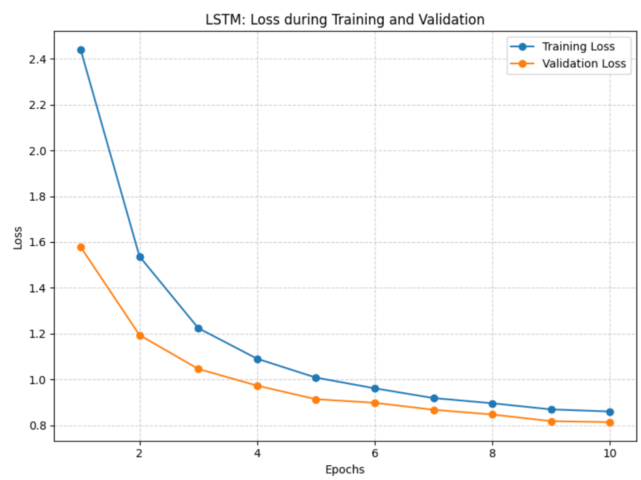


Fig 6: LSTM model's validation loss progression

Metric	Score
Validation Accuracy	0.70
Validation Loss	0.84
Precision	0.67
Recall	0.70
F1-Score	0.68

Table 2: LSTM Evaluation metrics

The LSTM performed well for dominant classes like ESFJ (F1: 1.00) and ISFJ (F1: 0.98), but its performance for minority classes such as INFJ (F1: 0.22) and INTP (F1: 0.18) revealed limitations in handling class imbalances.

7.1.3 Bi-directional LSTM: The Bi-LSTM model significantly outperformed LSTM by capturing bidirectional context. Its metrics are presented in Table 3, and the figure displays the training and validation loss for Bi-LSTM.

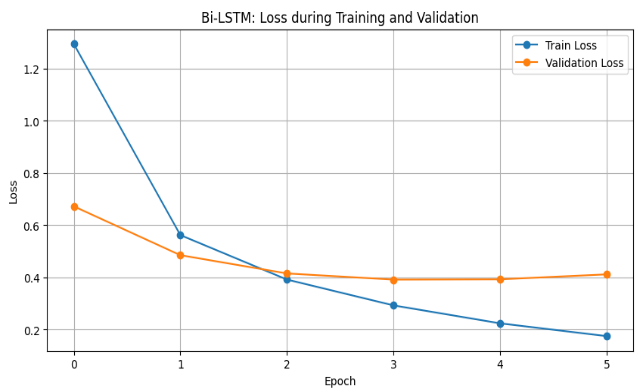


Fig 7: Bi-LSTM model's validation loss progression

Metrics	Score
Validation Accuracy	0.88
Validation Loss	0.41
Precision	0.88
Recall	0.88
F1-Score	0.88

Table 3: Bi-Lstm Evaluation Metrics

Bi-LSTM excelled across most classes, achieving F1-scores of 0.99 for ESFJ and ESTP. Its bidirectional context enabled better performance for minority classes compared to the standard LSTM.

7.1.4 Bi-directional LSTM With Attention Mechanism: Adding attention to the Bi-LSTM architecture refined its ability to focus on important parts of the input sequence, improving both interpretability and performance. Table 4 highlights its metrics, and the figure displays the training and validation loss for Bi-LSTM.

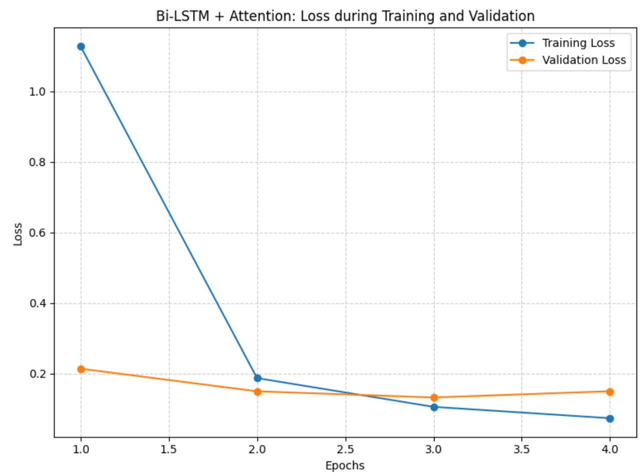


Fig 8: Bi-LSTM + attention model's validation loss progression

Metrics	Score
Validation Accuracy	0.96
Validation Loss	0.13
Precision	0.96
Recall	0.96
F1-Score	0.96

Table: 4 Bi-Lstm with attention mechanism evaluation metrics

The attention layer improved interpretability and performance for underrepresented classes like INFP (F1-score: 0.88) and INFJ (F1-score: 0.87). This demonstrates the importance of focusing on key parts of the sequence for better class representation and prediction which the attention mechanism provides.

7.1.5 DistilBert (Transformers model): After achieving unmatched accuracy and minimal validation loss with our Bi-LSTM + Attention model, we decided to explore a transformer-based architecture to further enhance performance. We implemented DistilBERT, a lightweight and computationally efficient variant of the larger BERT model, which balances power and efficiency. Due to computational constraints, we fine-tuned DistilBERT on a subset of 50,000 data points out of the available 400,000, using the GPU resources provided by Google Colab. Despite this limitation, DistilBERT demonstrated competitive performance, showcasing the robustness and effectiveness of transformer-based architectures in capturing complex language patterns. Table 5 highlights its metrics.

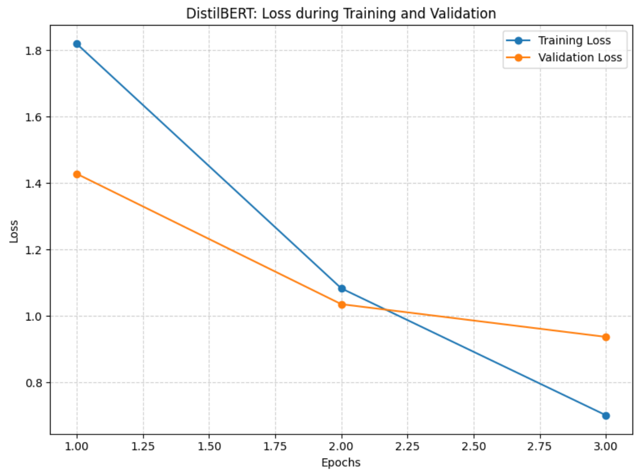


Fig 7: DistilBERT model's validation loss progression

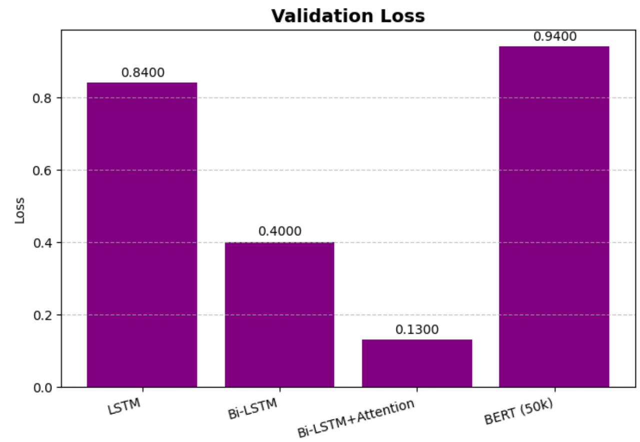


Fig 9: Validation Loss of all models

Metrics	Score
Validation Accuracy	0.72
Validation Loss	0.94
Precision	0.73
Recall	0.72
F1- Score	0.72

Table 5: DistilBERT Evaluation Metrics

To provide a clear comparison of model performance, we visualized the training and validation accuracies, along with validation loss, for all the models. These plots illustrate the progression of accuracy and reduction in loss across different architectures. This comparative analysis helped us evaluate the effectiveness of each model and identify the trade-offs between accuracy and computational efficiency.

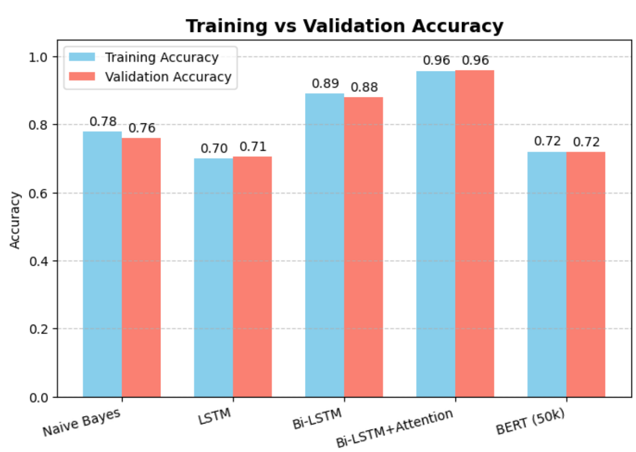


Fig 8: Training vs Validation Accuracy of all models

The improvement in results across models can be attributed to the key advancements in architecture and feature processing at each stage. With Naive Bayes, we relied on simple probabilistic assumptions that could not account for word order or relationships, limiting performance. Moving to LSTM allowed us to capture sequential dependencies in the data, leveraging the context in word sequences. However, standard LSTM models struggled with long-term dependencies and imbalanced classes. The introduction of Bi-LSTM addressed this limitation by processing sequences bidirectionally, thereby capturing both past and future context. This bidirectional approach provided richer sequence representations, significantly boosting generalization and reducing overfitting.

Adding an attention mechanism further enhanced performance by enabling the model to focus on the most critical parts of the input sequence, improving predictions for underrepresented classes. Attention not only improved interpretability but also reduced the model's reliance on irrelevant features, resulting in the lowest validation loss among all tested models.

Finally, DistilBERT marked a significant shift towards transformer-based architectures, which excel in modeling complex relationships within the data. Despite being fine-tuned on a limited subset due to computational constraints, DistilBERT showcased its robustness and potential for scaling with larger datasets. These advancements highlight the importance of progressively incorporating architectural innovations to optimize performance while addressing the specific challenges posed by the dataset.

8. BROADER IMPACTS / DISCUSSION

The potential applications of our personality prediction model span multiple domains, reflecting its versatility and societal relevance. In Human Resources, our work can help recruiters gain deeper insights into candidates' personalities through text-based analysis, enabling more tailored interview processes and team compositions. By identifying traits like conscientiousness or openness, hiring managers can make more informed decisions, fostering better workplace dynamics.

In the world of Targeted Marketing, personality insights can revolutionize personalized marketing strategies. For example, understanding a user's preferences and behavior patterns allows businesses to craft more compelling and relevant campaigns. This approach not only enhances user engagement but also improves conversion rates by aligning messages with individual personalities.

Our project also holds promise in Education and Counseling, where understanding students' or clients' personality traits can guide tailored teaching methods or therapy approaches. For instance, counselors could adapt their sessions based on traits like introversion or neuroticism, fostering better communication and outcomes. Similarly, educators might adjust their instructional strategies to cater to diverse learning styles, making education more inclusive and effective.

Recommender Systems are another area where personality-based predictions can make a significant impact. Platforms for entertainment, learning, or shopping can refine their algorithms by incorporating personality profiles, delivering recommendations that resonate with individual users. Whether it's suggesting movies, books, or courses, such systems could create more engaging and satisfying user experiences.

Through this project, we've not only explored the technical aspects of personality prediction but also recognized its far-reaching potential to influence industries that rely on understanding human behavior. These broader impacts motivate us to continue refining our models to ensure accuracy and usability in real-world applications.

9. CONCLUSION

In this project, we leveraged both classical and deep learning approaches, including LSTM, Bi-LSTM, Bi-LSTM with Attention, and DistilBERT, to predict personality traits from textual data. By incorporating pre-trained word embeddings like GloVe and advanced tokenization techniques, we ensured that the models could capture semantic relationships and contextual dependencies in text effectively. The integration of attention mechanisms further enhanced the ability to prioritize key features, while careful selection of optimizers, such as Adam, allowed for stable and efficient training. While our results demonstrated the effectiveness of these architectures, the emergence of Large Language Models (LLMs) like GPT and advanced BERT variants holds immense promise for the future. These models, with their ability to deeply understand linguistic nuances and scale to large datasets, could address limitations such as data imbalance and enhance the interpretability of predictions. Future iterations of this work could explore fine-tuning LLMs and combining them with innovative embeddings and tokenization strategies to unlock new applications in domains like human resources, education, and personalized marketing. This project not only underscored the current capabilities of NLP techniques but also highlighted the transformative potential of state-of-the-art language models in personality prediction.

5. WORK DISTRIBUTION

Phase	Teammate A (Atharva Chouthai)	Teammate B (Anvita Karne)
Phase 1: Research	Researched classical ML models and proposed topics.	Finalized the topic and researched advanced deep learning models like LSTM and BERT.
Phase 2: Data Collection & Preprocessing	Handled tokenization of the dataset for model input.	Integrated GloVe (300d) embeddings for pre-trained word representations.
Phase 3: Classical ML Model Implementation	Reviewed and analyzed classical ML results to establish baselines.	Implemented and evaluated Naive Bayes.
Phase 4: Deep Learning Model Implementation	Implemented Bi-LSTM with Attention and DistilBERT architectures, focusing on attention mechanisms and transformer efficiency.	Implemented LSTM and Bi-LSTM architectures, optimizing for long-term sequence dependencies.
Phase 5: Model Evaluation	Evaluated Bi-LSTM with Attention and DistilBERT models, visualizing performance trends.	Evaluated Naive Bayes, LSTM, and Bi-LSTM models, focusing on accuracy and loss metrics.
Phase 6: Related Work	Drafted portions of the related work focusing on classical ML and deep learning integration.	Drafted portions on transformer-based models and their applications.
Phase 7: Documentation & Reporting	Drafted the methodology, introduction, and related work sections to align with the project goals.	Drafted results, discussion, conclusion, and broader impacts sections, highlighting key insights.

Phase 8: Presentation	Prepared slides on Bi-LSTM with Attention and DistilBERT models, emphasizing their innovations.	Prepared slides on GloVe embeddings, LSTM, and Naive Bayes, ensuring comprehensive coverage of foundational approaches.

6. REFERENCES

[1]. Gjurković, Matej & Šnajder, Jan. (2018). Reddit: A Gold Mine for Personality Prediction. 87-97. 10.18653/v1/W18-1112.

[2]. Tommy Tandera, Hendro, Derwin Suhartono, Rini Wongso, Yen Lina Prasetio, Personality Prediction System from Facebook Users, Procedia Computer Science, Volume 116, 2017, Pages 604-611, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2017.10.016.

[3]. Mehta, Yash & Majumder, Navonil & Gelbukh, Alexander & Cambria, Erik. (2020). Recent trends in deep learning-based personality detection. Artificial Intelligence Review. 53. 10.1007/s10462-019-09770-z.

[4]. Digman, J. M. (1990). "Personality Structure: Emergence of the Five-Factor Model." Annual Review of Psychology, 41(1), 417–440.

[5]. Mairesse, Francois & Walker, Marilyn & Mehl, Matthias & Moore, Roger. (2007). Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text. J. Artif. Intell. Res. (JAIR). 30. 457-500. 10.1613/jair.2349.

[6]. Zeiler, M. D. (2012). "ADADELTA: An Adaptive Learning Rate Method." Computing Research Repository (CoRR).

[7]. Mohammad, S. M., & Turney, P. D. (2013). "Crowdsourcing a Word–Emotion Association Lexicon." Computational Intelligence, 29(3), 436–465.

[8]. Mohammad, S. M., & Turney, P. D. (2010). "Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon." Proceedings of the NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, 26–34.

[9]. Jing, L., Li, C., & Liang, Y. (2017). "Investigating MBTI Personality Traits Through Online Social Media Usage." Procedia Computer Science, 122, 648–654. DOI: 10.1016/j.procs.2017.11.430.

[10]. Patra, B. G., Das, D., & Bandyopadhyay, S. (2016). "Multimodal Mood Classification Framework for Hindi Songs." Computación y Sistemas, 20(3), 515–526.

[11]. Goldberg, Y., & Levy, O. (2014). "Word2Vec Explained: Deriving Mikolov et al.'s Negative-Sampling Word-Embedding Method." arXiv preprint arXiv:1402.3722.

[12]. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT).

[13]. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N.& Polosukhin, I. (2017). "Attention Is All You Need." Advances in Neural Information Processing Systems, 30, 5998–6008.

[14]. Poria, S., Cambria, E., & Gelbukh, A. (2016). "Deep Convolutional Neural Network Textual Features and Multiple Kernel Learning for Utterance-level Multimodal Sentiment Analysis." Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT).