The enhancement of Personality Assessment and Detection using Machine Learning Techniques

Shunyu Chen[1], Yichen Liu[2], Tianqin Meng[3], Sibo Wang[4, *]

[1]College of Mechanical and Electrical Engineering, Shanghai Jian Qiao University, Shanghai 201306, China
[2]Regis Jesuit High School, CO 80016, USA
[3]Merivale High School, Canada
[4]College of Artificial Intelligence, Tianjin University, Tianjin 300072, China
*Corresponding author: 3019244093@tju.edu.cn

Abstract. The rapid advancements in science and technology have had a profound impact on how people perceive themselves and communicate with others. As a result, personality tests have become increasingly popular for individuals seeking self-awareness and a deeper understanding of others. This study focuses specifically on the Myers-Briggs Type Indicator (MBTI) personality test and utilizes machine learning techniques to enhance its effectiveness. The research begins by exploring various methods for conducting personality tests, including Logistic Regression (LR), Support Vector Machine (SVM), and Gradient Boosting (GB). These methods are compared, and based on their performance, Gradient Boosting is identified as the most promising approach. Further optimization is carried out, resulting in a final model capable of accurately predicting personality traits. The precision and accuracy of the model meet the desired requirements, showcasing its potential for practical applications. Moving forward, the study highlights the importance of future improvements and refinements to enhance the overall performance of the model. By continually advancing and refining machine learning techniques in the context of personality assessment, individuals can gain valuable insights into themselves and others, leading to personal growth and improved communication. In summary, this research demonstrates the significant role of machine learning in improving personality tests like MBTI, empowering individuals to develop self-awareness and foster meaningful connections with others.

Keywords: MBTI, Logistic Regression, SVM, Gradient Boosting.

1 Introduction

In the modern era, understanding human personalities has become increasingly important. MBTI is a well-known personality profiling system that categorizes individuals into 16 different personality types based on a series of guiding questions [1]. Recently, several college application systems have incorporated this personality profiling test to gain a better understanding of applicants and help students find their

P. Kar et al. (eds.), *Proceedings of the 2023 International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2023)*, Advances in Computer Science Research 108,
https://doi.org/10.2991/978-94-6463-300-9_12

interests. This study aims to address the following key questions: Firstly, to evaluate the accuracy of machine learning algorithms in predicting MBTI personality types using text-based data. Secondly, to explore the contribution of machine learning models in enhancing the understanding of MBTI profiling predictions. Lastly, to identify linguistic cues and patterns that can improve the accuracy of personality prediction and optimize the performance of machine learning algorithms in this field.

This study endeavors to make significant contributions in various areas. From a practical aspect, showcasing the effectiveness of machine learning algorithms in predicting MBTI personalities has direct implications for industries such as recruitment, team formation, and personalized marketing strategies. Additionally, this study advances our understanding of human behavior and personality traits by identifying patterns and language indicators associated with specific MBTI personality types. By leveraging machine learning algorithms, this study enhances the field of personality evaluation and analysis, offering a methodological approach to predicting personalities.

2 Related work

In recent years, the field of predicting human personality types has made great progress, with various methods proposed for personality categorization. Tandera built a system to predict a person's personality based on Facebook user information [2]. Using deep learning architectures, specifically multi-layer perceptron (MLP) models, and compared the approach with traditional machine learning algorithms such as Naive Bayes, SVM, LR, GB, and Linear Discriminant Analysis (LDA) [3, 4]. The results demonstrated the superior accuracy of the deep learning approach, achieving an average accuracy of 74.17%.

Cui and Qi introduced the SoftMax classifier as a baseline model and discussed the Naive Bayes method for text classification and SVM for separating classes by seeking hyperplanes [5]. Their research provided insights into the methodology and approaches used in predicting personality types using machine learning techniques.

Amirhosseini and Kazemian focused on machine-learning approaches for predicting personality types based on the MBTI [6]. They explored classic machine learning methods, neural networks, SVM, grey prediction models, and recurrent neural networks (RNN) for building classifiers to predict personality types [7, 8]. Their study highlighted the successful application of these techniques in predicting personality using social media platforms like Twitter.

Garg and Garg compared several machine learning algorithms for content-based personality resolution of tweets [9]. The research employed classifiers such as SVM, Decision Tree (DT), K-Nearest Neighbors (KNN), LR, Random Forest (RF), and Extreme Gradient Boosting(XGB) [10, 11]. These classifiers have been extensively used in personality type prediction research and have demonstrated clear advantages.

Murari and Bharathi employed the Binary-Partitioning Transformer (BPT) technique, which divides input sequences into multiple scale spans via binary partitioning [12]. This approach improves the effectiveness of the self-attention system and captures the links between words in a sentence for model building. Pablo Sánchez-Fernández introduced Artificial Neural Networks (ANN) as a more advanced technique for redirecting input information [13]. Their study investigated multiple techniques for evaluating MBTI personality prediction.

These studies contribute to the existing body of research on personality prediction by exploring various machine learning algorithms, deep learning architectures, and natural language processing techniques. They provide insights into the accuracy, performance, and methodology of different models and offer valuable guidance for further research in predicting human personality types.

3 Methodology

This study presents a methodology that combines the strengths of machine learning algorithms and domain expertise to achieve the objectives outlined in the preceding section. The following section provides a comprehensive description of the methodology, highlighting its key steps and critical strategies.

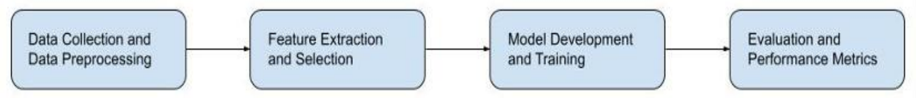


Fig. 1. Personality Classification process flow chart (Photo/Picture credit: Original)

Figure 1 shows the personality classification process, the first phase entails gathering a thorough dataset of people's personality attributes, including their online behaviors. To guarantee data quality and consistency, this dataset was carefully selected and prepossessed. The data is lemmatized, and punctuation, numbers, and stopwords in the data are deleted. The raw data is then transformed into meaningful and instructive features using feature extraction techniques. In this case, the features are the words per comment, the variety of word counts, website posts, and inquiries, which are extracted from the raw data of posts based on individuals of different personality traits. Important personality factors like communication style, level of involvement, and information-seeking behavior are captured by these attributes.

Different machine learning algorithms suitable for personality assessment and detection are investigated for model building and training, including LR, SVM, and GB. The preprocessed dataset is used to train the chosen models. Accuracy, precision, recall, and F1-score are just a few of the performance metrics that are used to assess the trained models.

3.1 Data Preprocessing

The textual analysis problem being tackled by machine learning involves categorizing personalities using the characteristics of post count and word count. However, using just these features has limitations in terms of being able to capture a wider variety of personality traits and can result in problems with the class imbalance that affect the model's performance.

To get over these restrictions, a data pretreatment procedure was used to include two more traits that were thought to be useful for personality categorization. First, a limited amount of data from an analysis of the words used in each comment for each personality type is testified from the dataset. Due to an imbalance in the datasets, the neural network may have been biased towards examples with a huge amount of data. By combining inputs from other datasets, it is hoped to equalize the data for each label and lessen the impact on recall in the classification confusion matrix.

In addition, two new headings—website articles and inquiries—were added by the algorithm. The objective is to record a wider range of personality traits and determine their averages, which will be utilized for categorization, by increasing the variability in features.

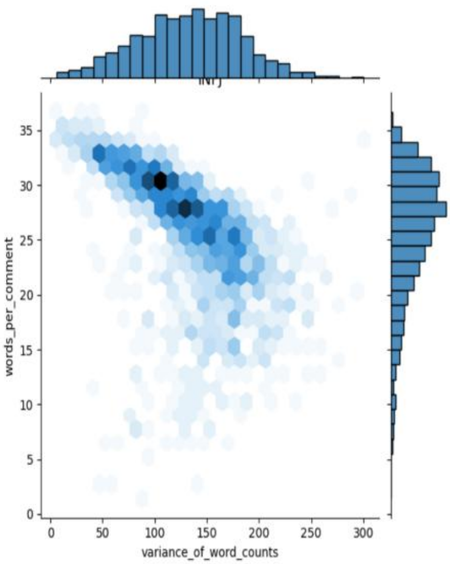


Fig. 2. Relationship between each label (Photo/Picture credit: Original)

Figure 2 shows the link between the Average Words Per Comment and the Variance of Word Counts for each label. The relationship between these variables and their associations with other personality labels are identified through this analysis as potential trends or patterns.

Figure 3 shows the frequency of words in MBTI, the textual data involved in identifying the most descriptive words for each personality type. Commonly used words such as "think" and "thank" were excluded from the feature set as they lacked discriminative power for this specific classification task. These data pre-processing techniques aim to enhance the overall quality of the dataset, mitigate class imbalance issues, and improve the accuracy and precision of machine learning models in predicting MBTI personality types.

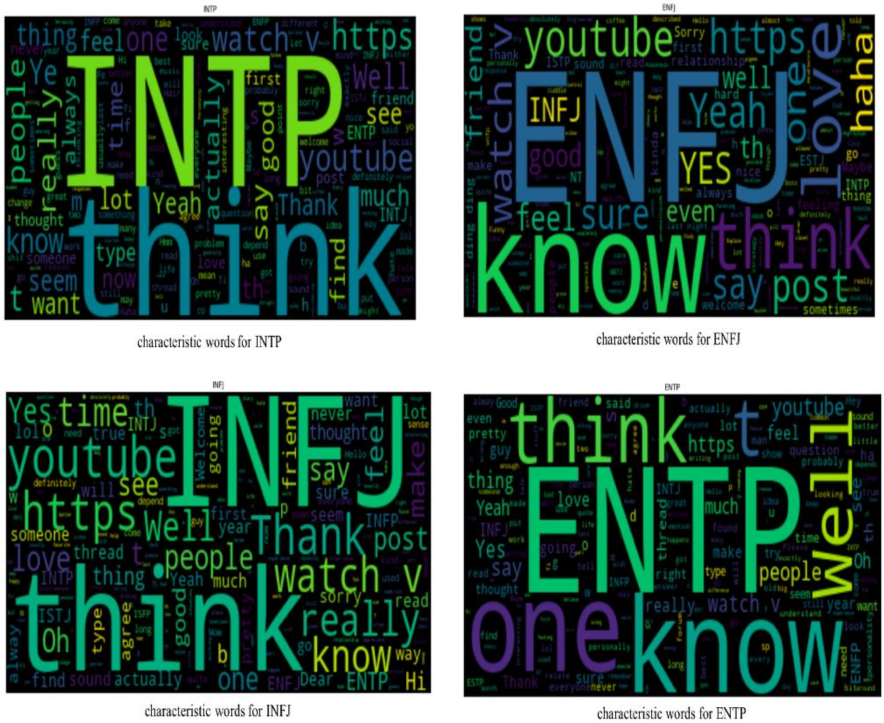


Fig. 3. The frequency of words in MBTI (Photo/Picture credit: Original)

In the final stage of the process, certain irrelevant words, such as links and symbols, are removed from the text. This step helps improve the precision of model building by eliminating noise and focusing on the most relevant words. The relationship between the number of words in the text and their occurrence frequency is depicted in Figure 4.

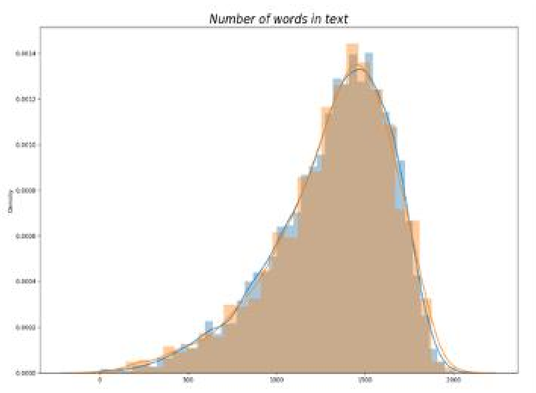


Fig. 4. Number of words in the text (Photo/Picture credit: Original)

3.2 Model selection

LR, SVM, and GB were adapted for the personality classification text. Logistic Regression has proven to be highly effective in the field of classification, which aims to distinguish one type of data from another type. Under binary classification, taking a statistical point of view of the perceptron, logistic regression predicts the probability of an input belonging to a certain class.

It is based on the logistic function which maps any real-valued number to a value between 0 and 1.

$$\sigma(x) = \frac{1}{1+e^{-x}},\ \sigma: R \to (0, 1) \tag{1}$$

It uses a very steep activation function 'o':

$$O\,(W^{T}x + w_{0}) = \begin{cases} 1, W^{T}x + w_{0} > 0 \\ -1, otherwise \end{cases} \tag{2}$$

The likelihood of a given data point (x_n, t_n) is:

$$P\,(t_n = 1 \mid x_n, W, w_0) = \sigma(t_n(W^{T}x + w_0)) \tag{3}$$

The loss function of the logistic regression is the negative log-likelihood function of the dataset, which is defined as:

$$L\,(W, w_0) = -\,\Sigma\, log\sigma\,(t_n(W^{T}x + w_0)) \tag{4}$$

To minimize the negative log-likelihood, the gradient descent method is used to optimize a random weight initialization. The key step is to compute the gradient for the loss function. Here is the result:

$$\frac{dL}{dW} = -\,\Sigma\,[1 - \sigma(t_n W^{T}x_n)](t_n x_n) \tag{5}$$

During the process of gradient descent optimization, each iteration follows a specific procedure. Firstly, the current loss is computed based on the provided dataset. Then, the gradient is calculated, representing the direction of the steepest ascent. Finally, a small step is taken in the opposite direction of the gradient to update the model parameters. This process is repeated iteratively until convergence is achieved.

In the provided code structure, the main procedure of gradient descent has been implemented. It encompasses the steps of computing the loss, calculating the gradient, and updating the model parameters. This ensures that the optimization process progresses towards minimizing the loss and improving the model's performance.

3.3 Model training

Before running the final model, it is important to clarify the requirements and objectives of the experiment. In addition to implementing the desired functionality, specific targets need to be achieved. In this case, the optimization demand is to achieve a precision of at least 80 percent. Maintaining a precision level below this threshold may lead test participants to doubt the accuracy of their character classification, thereby diminishing the significance of the experiment. Therefore, it is evident that the current model based on Gradient Boosting is not meeting the desired criteria. To reach our target, further optimization is necessary.

Figure 5 shows the model optimization process. To optimize the model, parameter adjustment is a direct and convenient approach regardless of the model type. The main objective of this optimization is to address bias and variance, thereby mitigating the issues of underfitting and overfitting.

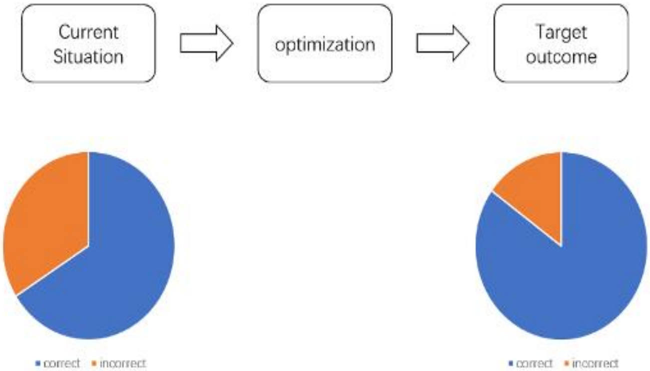


Fig. 5. Model optimization process (Photo/Picture credit: Original)

To mitigate underfitting, the model complexity was increased, and relevant parameters were adjusted. Conversely, to combat overfitting, the training data was expanded. Alternative datasets were also experimented with during the optimization process, but the resulting improvements were minimal and deemed insignificant. In the whole process, the original datasets were retained. During the parameter adjustment phase, the parameters were manually tuned by exploring different combinations using the sklearn library and GridSearchCV. These widely used libraries provide convenient and comprehensive functionality. After completing these steps, the model was retrained and evaluated to assess the impact of the optimizations. Figure 6 shows the model rebuilding process.

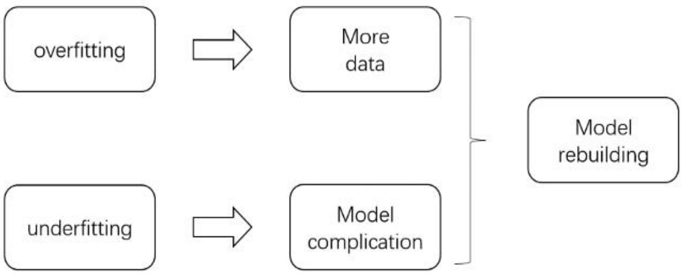


Fig. 6. Model rebuilding process (Photo/Picture credit: Original)

3.4 Evaluation and performance metrics

Evaluating the accuracy of a classifier is crucial to assess the predictive performance of a machine learning classification model on unseen data. This is particularly important when dealing with skewed data distributions, as the model may tend to predict everything as belonging to the biased label in the case of personality classification, as observed in the original database.

To effectively summarize the performance of a classification model in a binary classification scenario, a confusion matrix, as presented in Table 1, is commonly used. The confusion matrix provides valuable insights into the model's performance by showcasing the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

Table 1. Confusion Matrix in Binary Classification

	Predicted label		
True label		+	-
	+	TP	FN
	-	FP	TN

Form the confusion matrix, the following indexes are calculated:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

$$Recall = \frac{TP}{TP + TN} \tag{7}$$

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

$$F1\ score = \frac{2 * Recall * Precision}{Recall + Precision} \tag{9}$$

Accuracy measures the overall correctness of the model's predictions. It calculates the ratio of correctly predicted samples to the total number of samples. Accuracy is commonly used to evaluate classification models when the classes are balanced. However, it may not be suitable for imbalanced datasets where the number of samples in different classes varies significantly.

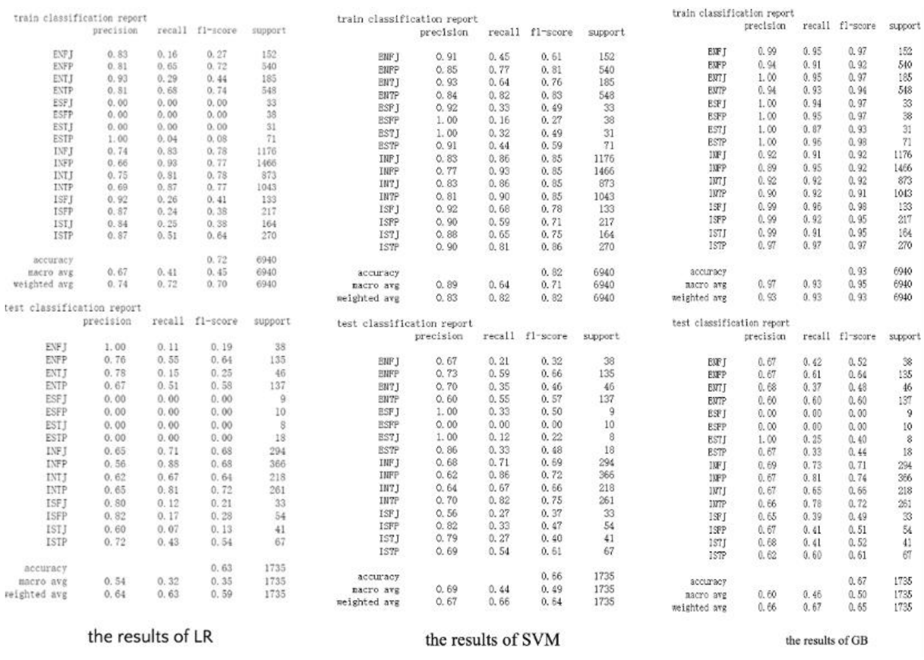
Recall focuses on the positive samples. The denominator represents the actual number of positive samples. For example, when detecting highly contagious diseases, if there are many patients, it is important that all of them are identified. In this case, the emphasis should be on recall because the goal is to identify as many positive samples as possible. The true positive rate should ideally be 1.

Precision focuses on the positive predictions (or charges) made by the model. For example, if you want the judicial system to judge whether a person is guilty as accurately as possible, you do not want them to make too many false charges as it would impact someone's life. Precision represents the number of correct predictions out of all predictions made as positive.

The F1 score considers the balance between recall and precision. Since recall and precision contradict each other, it is necessary to adjust the threshold to balance the two. If precision is very low, the value of 1/ precision approaches infinity, resulting in a low F1 score. The F1 score can tell us whether the model has achieved a good balance between recall and precision.

4 Results

Before retraining the model, it is designed to evaluate the model’s performance. The desired precision is 80 percent, and if the final results exceed this threshold, the experiment can be considered successful. The training and testing results of the models are shown in Figure 7.



train classification report

	precision	recall	f1-score	support
ENFJ	0.83	0.16	0.27	152
ENFP	0.81	0.65	0.72	540
ENTJ	0.93	0.29	0.44	185
ENTP	0.81	0.68	0.74	548
ESFJ	0.00	0.00	0.00	33
ESFP	0.00	0.00	0.00	38
ESTJ	0.00	0.00	0.00	31
ESTP	1.00	0.04	0.08	71
INFJ	0.74	0.83	0.78	1176
INFP	0.66	0.93	0.77	1466
INTJ	0.75	0.81	0.78	873
INTP	0.69	0.87	0.77	1043
ISFJ	0.92	0.26	0.41	133
ISFP	0.87	0.24	0.38	217
ISTJ	0.84	0.25	0.38	164
ISTP	0.87	0.51	0.64	270
accuracy			0.72	6940
macro avg	0.67	0.41	0.45	6940
weighted avg	0.74	0.72	0.70	6940

test classification report

	precision	recall	f1-score	support
ENFJ	1.00	0.11	0.19	38
ENFP	0.76	0.55	0.64	135
ENTJ	0.78	0.15	0.25	46
ENTP	0.67	0.51	0.58	137
ESFJ	0.00	0.00	0.00	9
ESFP	0.00	0.00	0.00	10
ESTJ	0.00	0.00	0.00	8
ESTP	0.00	0.00	0.00	18
INFJ	0.65	0.71	0.68	294
INFP	0.56	0.88	0.68	366
INTJ	0.62	0.67	0.64	218
INTP	0.65	0.81	0.72	261
ISFJ	0.80	0.12	0.21	33
ISFP	0.82	0.17	0.28	54
ISTJ	0.60	0.07	0.13	41
ISTP	0.72	0.43	0.54	67
accuracy			0.63	1735
macro avg	0.54	0.32	0.35	1735
weighted avg	0.64	0.63	0.59	1735

the results of LR

train classification report

	precision	recall	f1-score	support
ENFJ	0.91	0.45	0.61	152
ENFP	0.85	0.77	0.81	540
ENTJ	0.93	0.64	0.76	185
ENTP	0.84	0.82	0.83	548
ESFJ	0.92	0.33	0.49	33
ESFP	1.00	0.16	0.27	38
ESTJ	1.00	0.32	0.49	31
ESTP	0.91	0.44	0.59	71
INFJ	0.83	0.86	0.85	1176
INFP	0.77	0.93	0.85	1466
INTJ	0.83	0.86	0.85	873
INTP	0.81	0.90	0.85	1043
ISFJ	0.92	0.68	0.78	133
ISFP	0.90	0.59	0.71	217
ISTJ	0.88	0.65	0.75	164
ISTP	0.90	0.81	0.86	270
accuracy			0.82	6940
macro avg	0.89	0.64	0.71	6940
weighted avg	0.83	0.82	0.82	6940

test classification report

	precision	recall	f1-score	support
ENFJ	0.67	0.21	0.32	38
ENFP	0.73	0.59	0.66	135
ENTJ	0.70	0.35	0.46	46
ENTP	0.60	0.55	0.57	137
ESFJ	1.00	0.33	0.50	9
ESFP	0.00	0.00	0.00	10
ESTJ	1.00	0.12	0.22	8
ESTP	0.86	0.33	0.48	18
INFJ	0.68	0.71	0.69	294
INFP	0.62	0.86	0.72	366
INTJ	0.64	0.67	0.66	218
INTP	0.70	0.82	0.75	261
ISFJ	0.56	0.27	0.37	33
ISFP	0.82	0.33	0.47	54
ISTJ	0.79	0.27	0.40	41
ISTP	0.69	0.54	0.61	67
accuracy			0.66	1735
macro avg	0.69	0.44	0.49	1735
weighted avg	0.67	0.66	0.64	1735

the results of SVM

train classification report

	precision	recall	f1-score	support
ENFJ	0.99	0.95	0.97	152
ENFP	0.94	0.91	0.92	540
ENTJ	1.00	0.95	0.97	185
ENTP	0.94	0.93	0.94	548
ESFJ	1.00	0.94	0.97	33
ESFP	1.00	0.95	0.97	38
ESTJ	1.00	0.87	0.93	31
ESTP	1.00	0.96	0.98	71
INFJ	0.92	0.91	0.92	1176
INFP	0.89	0.95	0.92	1466
INTJ	0.92	0.92	0.92	873
INTP	0.90	0.92	0.91	1043
ISFJ	0.99	0.96	0.98	133
ISFP	0.99	0.92	0.95	217
ISTJ	0.99	0.91	0.95	164
ISTP	0.97	0.97	0.97	270
accuracy			0.93	6940
macro avg	0.97	0.93	0.95	6940
weighted avg	0.93	0.93	0.93	6940

test classification report

	precision	recall	f1-score	support
ENFJ	0.67	0.42	0.52	38
ENFP	0.67	0.61	0.64	135
ENTJ	0.68	0.37	0.48	46
ENTP	0.60	0.60	0.60	137
ESFJ	0.00	0.00	0.00	9
ESFP	0.00	0.00	0.00	10
ESTJ	1.00	0.25	0.40	8
ESTP	0.67	0.33	0.44	18
INFJ	0.69	0.73	0.71	294
INFP	0.67	0.81	0.74	366
INTJ	0.67	0.65	0.66	218
INTP	0.66	0.78	0.72	261
ISFJ	0.65	0.39	0.49	33
ISFP	0.67	0.41	0.51	54
ISTJ	0.68	0.41	0.52	41
ISTP	0.62	0.60	0.61	67
accuracy			0.67	1735
macro avg	0.60	0.46	0.50	1735
weighted avg	0.66	0.67	0.65	1735

the results of GB

Fig. 7. Train and test results of the models (Photo/Picture credit: Original)

Based on the evaluation results, the precision, recall, and F1-score results are listed in Table 2. The Gradient Boosting model outperformed the other algorithms in this experiment. Therefore, the Gradient Boosting model was selected as the final model for this study.

Table 2. Model comparison table

	precision	recall	F1-score
LR train	0.74	0.72	0.70
LR test	0.64	0.63	0.59
SVM train	0.83	0.82	0.82
SVM test	0.67	0.66	0.64
GB train	0.93	0.93	0.93
GB test	0.66	0.67	0.65

After rebuilding, Figure 8 shows a significant improvement in the model precision. The dataset was divided into two subsets, with one achieving 88 percent precision and the other reaching 80 percent. Both subsets meet the required precision threshold, indicating that the model is suitable for character testing.

```
              precision    recall  f1-score   support

           0       0.88      0.96      0.91      3478
           1       0.80      0.56      0.66      1076

    accuracy                           0.86      4554
   macro avg       0.84      0.76      0.79      4554
weighted avg       0.86      0.86      0.85      4554
```

Fig. 8. Rebuilding model results (Photo/Picture credit: Original)

The results reveal a substantial enhancement in precision and f1-score. The datasets utilized were sufficiently large to disregard data-related errors, ensuring the reliability of the outcomes. However, when the model processed data from the second subset, the recall was lower, suggesting that there is still room for improvement in the model's learning ability.

The before and after comparison of the model is shown in Figure 9, it demonstrates significant improvements in precision, recall, and f1-score after the optimization of the algorithm. These enhancements validate the effectiveness of the optimization process.

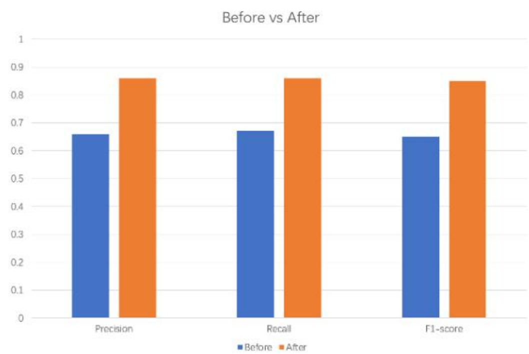


Fig. 9. Before and after comparison of the model (Photo/Picture credit: Original)

5 Conclusion

This study has investigated the application of machine learning techniques to enhance the effectiveness of the MBTI personality test. In the experiment, the paper compared three algorithms: Logistic Regression, SVM, and Gradient Boosting, to predict MBTI personality types using text-based data. The results unveiled that the Gradient Boosting algorithm showcased superior performance, solidifying its position as the most effective choice for this task. Through rigorous model optimization and parameter adjustments, significant improvements were achieved in precision and f1-score, fulfilling the predefined objectives of the experiment. However, challenges remain in handling specific datasets, warranting further refinement and adaptability enhancements for the model. Future efforts will be directed towards improving the model's versatility across diverse datasets and addressing its learning capacity for complex data patterns. Additionally, a user-friendly web interface will be developed to facilitate convenient personality assessment for individuals, promoting personal growth and meaningful connections. In summary, this study highlights the crucial role of machine learning in advancing personality tests like MBTI, by continuously advancing machine learning techniques for personality assessment, individuals can gain profound self-awareness and cultivate meaningful relationships, ultimately fostering personal growth and improved communication. The significance of this research lies in its potential to empower individuals in their journey towards better understanding themselves and others.

Acknowledgment

All the authors contributed equally and their names were listed in alphabetical order.

References

1. I. B. Myers. The Myers-Briggs Type Indicator: Manual [J]. 1962.
2. T. Tandera, D. Suhartono, R. Wongso, et al. Personality prediction system from Facebook users[J]. Procedia computer science 116, 604-611 (2017).
3. G. I. Web, E. Keogh, R. Miikkulainen. Naïve Bayes[J]. Encyclopedia of machine learning, 15, 713-714 (2010).
4. S. Suthaharan. Support vector machine[J]. Machine learning models and algorithms for big data classification: thinking with examples for effective learning, 207-235 (2016).
5. B. Cui, C. Qi. Survey analysis of machine learning methods for natural language processing for MBTI Personality Type Prediction[J]. Final Report Stanford University (2017).
6. M. H. Amirhossein, H. Kazemian. Machine learning approach to personality type prediction based on the myers–briggs type indicator[J]. Multimodal Technologies and Interaction, 4(1), 9 (2020).
7. H. Salehinejad, S. Sankar, J. Barfett, et al. Recent advances in recurrent neural networks[J]. arXiv preprint arXiv:1801.01078 (2017).

8. Y. C. Hu. Constructing grey prediction models using grey relational analysis and neural networks for magnesium material demand forecasting[J]. Applied Soft Computing, 93, 106398 (2020).
9. S. Garg, A. Garg. Comparison of machine learning algorithms for content-based personality resolution of tweets[J]. Social Sciences & Humanities Open, 4(1), 100178 (2021).
10. Zhongheng Zhang. Introduction to machine learning: k-nearest neighbors. Annals of translational medicine 4.11 (2016).
11. Y. C. Chang, K. H. Chang, G. J. Wu. Application of eXtreme gradient boosting trees in the construction of credit risk assessment models for financial institutions[J]. Applied Soft Computing, 73, 914-920 (2018).
12. M. D. Kamalesh, B. Bharathi. Personality prediction model for social media using machine learning Technique[J]. Computers and Electrical Engineering, 100, 107852 (2022).
13. P. Sánchez-Fernández, L. G. B. Ruiz, and M. C. P. Jiménez. Application of classical and advanced Machine Learning models to predict personality on social media. Expert Systems with Applications, 119498 (2023).