Mental Disorder Classification Prediction Based on AdaBoost Algorithm Optimized by Long Short-Term Memory Network

Yucong Li[1,a,*], Yulin Jin[2]

[1]College of Engineering, Northeastern University, Manitoba Province, City of Winnipeg, R3C 1W7, Canada.

[2]Department of Economics, University of California, Berkeley, State of California, Berkeley, 94704, USA

a. li.yuco@northeastern.edu

*corresponding author

Abstract: This paper discusses the parameter optimization of AdaBoost algorithm through long short-term memory network (LSTM) to improve the performance of mental disorder classification model. This innovative combination not only enhances the model's ability to recognize complex patterns, but also improves its robustness to noise and outliers. We introduced the model in the experiment and observed that the accuracy curve showed that the model's prediction accuracy improved significantly from the initial 20% to nearly 100%, and remained stable at this level. At the same time, the value of the loss function also decreases from the initial 1.4 to the final 0.1, indicating that the model tends to converge and learn effectively. On the training set, through confusion matrix analysis, we found that the prediction results of manic bipolar disorder, depressive bipolar disorder, major depression and normal individuals were 100% accurate, indicating that the model performed well in the training stage. However, on the test set, while 27 patients were correctly classified, there were still eight patients who were incorrectly classified, resulting in a test set prediction accuracy of 77.14%. This shows that although the model performs very well on the training set, it still needs to be further optimized in practical applications to improve its adaptability to new data. In summary, this study demonstrates the positive effects of combining LSTM and AdaBoost algorithm, which provides an efficient and reliable method for classifying patients with mental disorders. Future work will focus on improving the model to enhance its ability to generalize in real-world applications to better serve clinical diagnosis and treatment.

Keywords: Long Short-term Memory Network, AdaBoost, Patients with mental disorders.

1.   Introduction

Mental disorders are a group of diseases that affect an individual's thinking, feelings and behavior and often have a significant impact on the patient's daily life. According to the World Health Organization (WHO), approximately one in four people worldwide will experience a mental health problem at some point in their lives [1]. There are many types of mental disorders, including depression, anxiety, schizophrenia, bipolar disorder, and others. Due to the complexity and diversity of these diseases, traditional diagnostic methods often rely on clinical experience and subjective judgment, which can lead to misdiagnosis or missed diagnosis. Therefore, developing more objective

and effective classification methods has become an important research direction in the field of mental health.

In recent years, with the development of biomedical technology, researchers have begun to explore the combination of biomarkers, genomic data and neuroimaging information in order to establish a more scientific classification system. However, these methods usually face challenges such as large data volume and complex features. In this context, machine learning, as a powerful data analysis tool, has gradually attracted attention [2].

Machine learning is a method that enables computer systems to learn and improve through experience. Because of its excellent performance in processing large-scale, high-dimensional data, it is widely used in the medical field, especially in the classification of patients with mental disorders [3].

First, machine learning is able to efficiently process data from different sources, such as clinical questionnaires, brain imaging results, physiological indicators, and so on. These data can be integrated through algorithms to extract underlying patterns and features [4]. For example, by analyzing brain imaging data from a large number of people with depression, patterns of activity in certain regions can be found to correlate with depressive symptoms. Such findings not only help improve diagnostic accuracy, but also provide insight into disease mechanisms.

Second, machine learning algorithms such as support vector machines (SVM) [5], random forests [6], and deep learning can be used to build efficient predictive models. These models are able to learn from historical data and quickly classify new patients. For example, in one study, by training a deep learning-based model, patients with depression and anxiety were effectively distinguished with significantly higher accuracy than traditional assessment methods. This suggests that machine learning can not only improve classification, but also provide decision support for clinicians.

In addition, machine learning is adaptive and can constantly update and optimize the model as new data is added. This is particularly important in the field of mental disorders, where patient populations are heterogeneous and different individuals may exhibit different combinations of symptoms. By dynamically adjusting model parameters, machine learning can help identify new subtypes, thereby facilitating the development of personalized treatment options. In this paper, parameters of AdaBoost algorithm were optimized based on long short-term memory network to improve the model's effectiveness in classifying patients with mental disorders.

2. Sources of data sets

This paper uses open data set, which comes from the open source data set contributed by Kaggle. It can be seen that predecessors have done a lot of work on this data set, including the classification of basic machine learning algorithms, such as decision tree, random forest and neural network, but the classification effect is poor. This paper attempts to apply AdaBoost algorithm to this data set based on long short-term memory network for the first time. In order to display the data intuitively, we selected part of the data for display, as shown in Table 1.

Table 1: Partial data set.

Exhausted	Sleep dissorder	Sexual Activity	Concentration	Optimisim	Expert Diagnose
Sometimes	Sometimes	3 From 10	3 From 10	4 From 10	Bipolar Type-2
Usually	Sometimes	4 From 10	2 From 10	5 From 10	Depression
Sometimes	Sometimes	6 From 10	5 From 10	7 From 10	Bipolar Type-1
Usually	Most-Often	3 From 10	2 From 10	2 From 10	Bipolar Type-2
Sometimes	Sometimes	5 From 10	5 From 10	6 From 10	Normal

A total of four categories of psychiatric patients were included in the dataset, which collected 120 psychological patients with 17 basic symptoms for the diagnosis of manic bipolar disorder, depressive bipolar disorder, major depressive disorder, and normal individuals. The dataset contains 17 basic symptoms that psychiatrists use to diagnose the described disease. Behavioral symptoms are considered to be the patient's level of sadness, exhaustion, euphoria, sleep disturbances, mood swings, suicidal thoughts, anorexia, anxiety, attempts to explain, nervous breakdown, ignoring and moving on, admitting mistakes, overthinking, aggressive reactions, optimism, sexual activity, and concentration.

3.   Method

3.1.   Long short-term memory network

Long short-term memory network (LSTM) is a special type of recurrent neural network (RNN), which aims to solve the problem of gradient disappearance and explosion faced by traditional RNN when processing long sequence data [7]. LSTM captures long-term dependencies in time series data by introducing gating mechanisms that enable networks to selectively remember and forget information efficiently. Its core structure includes input gate, forget gate and output gate. These gate control units ensure that important information can be retained and unnecessary information can be forgotten by regulating the flow of information. The structure diagram of LSTM is shown in Figure 1.
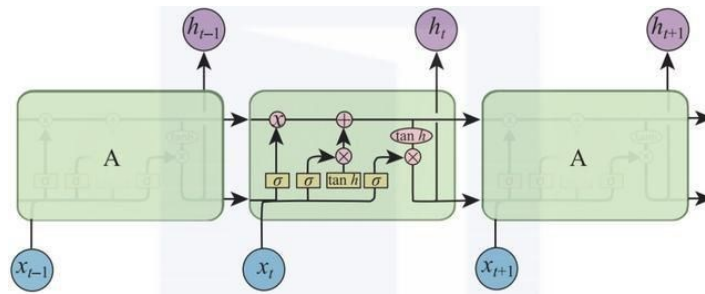

Figure 1: Partial data set.

Specifically, the input gate is responsible for deciding what new information needs to be added to the cell state; Oblivion gates control what old information should be discarded; The output gate determines which parts of the current cell state will be passed on to the next layer or as the final output. This design allows LSTMS to remain sensitive to critical information over a long time span, while avoiding the learning difficulties that traditional RNNS are prone to in processing long sequences. In addition, LSTM uses the concept of cell states to transmit information in a relatively constant way, making it excellent at handling complex time series tasks.

LSTM is widely used in natural language processing, speech recognition, time series prediction and other fields. For example, in machine translation, LSTM can effectively capture the context of sentences to generate more accurate translation results [8]. In speech recognition, it helps the model understand the time pattern in the speech signal and improve the recognition accuracy. With the development of deep learning, LSTM has become an important part of many advanced models and has driven advances in artificial intelligence technology.

3.2.   AdaBoost

AdaBoost is an ensemble learning method that improves the overall performance of a model by combining multiple weak classifiers. The core idea is to combine multiple simple classifiers (usually decision stumps) into one strong classifier. During the training process, AdaBoost dynamically

adjusts the weights of the samples according to the classification results of each sample. Specifically, for each iteration, the algorithm focuses on the samples that were misclassified by the previous round of classifiers, increasing their weight so that subsequent weak classifiers pay more attention to these hard-to-classify samples. This adaptive adjustment mechanism enables AdaBoost to effectively reduce bias and variance and improve model performance on training and test sets [9]. The structure diagram of AdaBoost model is shown in Figure 2
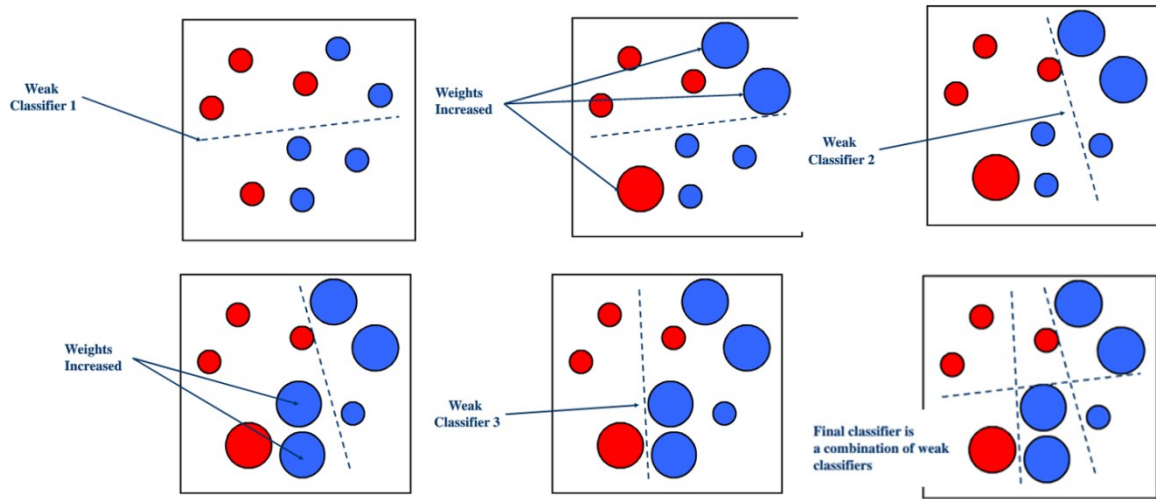

Figure 2: The structure diagram of AdaBoost model.

In practice, AdaBoost first initializes the weights of all training samples as equivalent, and then gradually builds a weak classifier through an iterative process. In each round, the algorithm calculates the weighted error rate of the current weak classifier on the training set and updates the sample weights based on that error rate. Weak classifiers with a lower error rate will have a higher weight in the final combination, while those with a higher error rate will have a lower weight. Finally, all the weak classifiers are combined by weighted voting or weighted average to form the final strong classifier. This method not only improves the accuracy of the model, but also enhances its robustness to noise and outliers.

3.3. AdaBoost based on long short term memory network optimization

When LSTM is combined with AdaBoost, AdaBoost is used to perform a weighted combination of multiple weak classifiers that can be built based on LSTM. Specifically, the LSTM model is first used to train the time series data to extract potentially important features. These features are then passed as input to multiple simple models (weak classifiers), such as decision trees or linear regression models [10]. Through the AdaBoost algorithm, these weak classifiers will be combined into a strong classifier. Each weak classifier is given different weights based on its performance on the training set, resulting in a more accurate and robust overall model.

This combination can not only improve the recognition ability of the model for complex patterns, but also enhance the resistance to noise and outliers. Because AdaBoost itself has good generalization ability, combined with strong feature extraction ability, the final model can maintain good performance in the face of new samples. Combining long short-term memory network with AdaBoost can give full play to their respective advantages, effectively extract important information from time series data, and improve the overall performance by using integrated learning method.

4.   Experiments and Results

In terms of experimental parameter setting, the optimizer used in this paper is gradient descent algorithm, the maximum training times are set to 500, the initial learning rate is set to 0.01, the learning rate decline factor is set to 0.1, and the learning rate is 0.01*0.1 after 400 training times. In terms of experimental hardware Settings, 4090 graphics card with 32G memory was used in the experiment, and matlab was used for all code model experiments.

First of all, the model was imported for experiment. The experiment was divided into training process and test process. The changes of accuracy and loss values of the training process model were recorded, and the change curves of accuracy and loss were drawn, as shown in Figure 3.
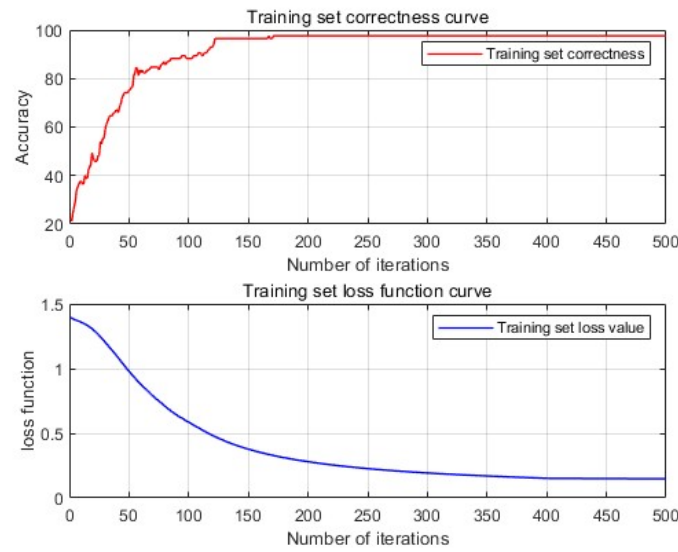

Figure 3: The changes of accuracy and loss values of the training process model.

According to the accuray curve, the accuracy of the model's prediction has increased from 20% at the beginning to about 100%, and continues to stay at this value. It can be seen from the loss curve that the loss value of the model changes from the initial 1.4 to the final 0.1 and remains at this value, tending to converge.

Test the prediction effect of the model using the test set, and output the confusion matrix of the training set and test set prediction, the results are shown in Figure 4 and Figure 5.
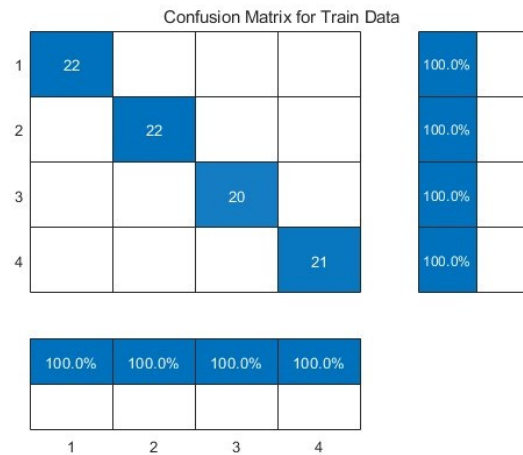

Figure 4: The confusion matrix.

According to the confusion matrix of the training set, in the training set, the prediction results of manic bipolar disorder, depressive bipolar disorder, major depression and normal individuals were all correct, and the prediction accuracy was 100%.
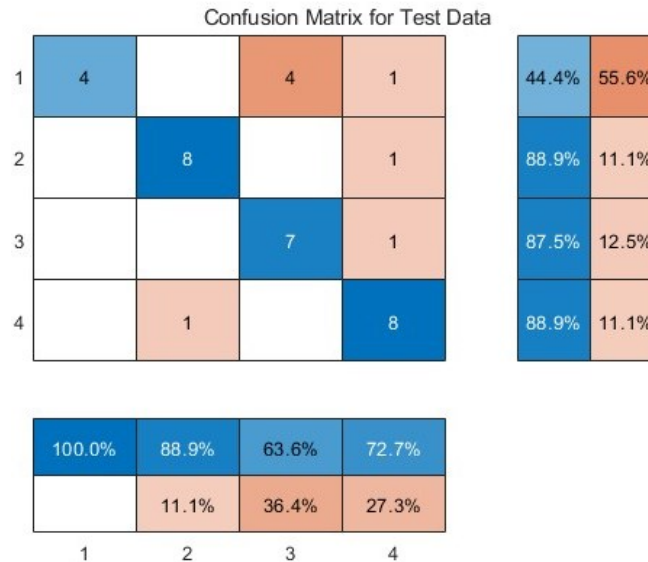

Figure 5: The confusion matrix.

According to the confusion matrix of the test set, in the test set, 27 patients were correctly classified and 8 patients were wrongly classified, with a prediction accuracy of 77.14%. The performance of the model in the test set was slightly.

5. Conclusion

In this paper, the parameters of AdaBoost algorithm were optimized based on long short-term memory network (LSTM) to improve the performance of the model in the classification of patients with mental disorders. This combination not only enhances the model's ability to recognize complex patterns, but also improves its resistance to noise and outliers. By introducing LSTM for feature extraction, combined with AdaBoost's weighted ensemble learning strategy, we can process data with temporal features more effectively, thus achieving higher classification accuracy.

During the experiment, we imported the optimized model, and observed through the accuracy curve that the accuracy of the model prediction gradually increased from the initial 20% to nearly 100%, and remained at this level steadily. At the same time, it can be seen from the loss curve that the loss value of the model decreases from the initial 1.4 to 0.1 and tends to converge, showing a good training effect. These results show that the optimized model performs well on the training set, and the prediction results of all categories (including manic bipolar disorder, depressive bipolar disorder, major depressive disorder, and normal individuals) are 100% accurate, fully demonstrating the effectiveness of the method in the task of classifying mental disorders.

However, performance on the test set was slightly worse than on the training set. In the test set, a total of 27 patients were correctly classified, while 8 patients were incorrectly classified, resulting in a final prediction accuracy of 77.14%. Nevertheless, the results still show a high overall accuracy, which indicates that our model has a strong generalization ability. By comparing the confusion matrix of the training set and the test set, it can be seen that although there is a certain degree of overfitting, the overall performance is still satisfactory.

In conclusion, by combining LSTM with AdaBoost, this study effectively improved the model performance in the classification task of patients with mental disorders. Future work can further explore data preprocessing, feature selection, and other methods of combining deep learning techniques with AdaBoost, in order to further improve the adaptability and accuracy of the model in different application scenarios. Such research not only provides new technical means for the field of mental health, but also provides important support for the early identification and intervention of related diseases.

References

[1] Yu, Yong, et al. "A review of recurrent neural networks: LSTM cells and network architectures." Neural computation 31.7 (2019): 1235-1270.

[2] Staudemeyer, Ralf C., and Eric Rothstein Morris. "Understanding LSTM--a tutorial into long short-term memory recurrent neural networks." arxiv preprint arxiv:1909.09586 (2019).

[3] Zhao, Zheng, et al. "LSTM network: a deep learning approach for short-term traffic forecast." IET intelligent transport systems 11.2 (2017): 68-75.

[4] Gers, Felix A., Nicol N. Schraudolph, and Jürgen Schmidhuber. "Learning precise timing with LSTM recurrent networks." Journal of machine learning research 3.Aug (2002): 115-143.

[5] Sundermeyer, Martin, Hermann Ney, and Ralf Schlüter. "From feedforward to recurrent LSTM neural networks for language modeling." IEEE/ACM Transactions on Audio, Speech, and Language Processing 23.3 (2015): 517-529.

[6] Ramakrishna, Mahesh Thyluru, et al. "Homogeneous adaboost ensemble machine learning algorithms with reduced entropy on balanced data." Entropy 25.2 (2023): 245.

[7] Lan, Yifu, Youqi Zhang, and Weiwei Lin. "Diagnosis algorithms for indirect bridge health monitoring via an optimized AdaBoost-linear SVM." Engineering Structures 275 (2023): 115239.

[8] Wen, Lifeng, et al. "Predicting the deformation behaviour of concrete face rockfill dams by combining support vector machine and AdaBoost ensemble algorithm." Computers and Geotechnics 161 (2023): 105611.

[9] Wang, Rui, et al. "AdaBoost-driven multi-parameter real-time warning of rock burst risk in coal mines." Engineering Applications of Artificial Intelligence 125 (2023): 106591.

[10] Wen, Lifeng, et al. "Predicting the deformation behaviour of concrete face rockfill dams by combining support vector machine and AdaBoost ensemble algorithm." Computers and Geotechnics 161 (2023): 105611.