

Research Article

Recognition of Psychological Characteristics of Students' Behavior Based on Improved Machine Learning

Mingchao Li 

School of Marxism Studies, Zhuhai College of Science and Technology, Zhuhai 519000, China

Correspondence should be addressed to Mingchao Li; limingchao@jluzh.edu.cn

Received 26 August 2021; Revised 6 October 2021; Accepted 23 October 2021; Published 9 November 2021

Academic Editor: Mu Zhou

Copyright © 2021 Mingchao Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Contemporary classroom teaching requires the combination of students' classroom behavior and their psychological activities and appropriately changes the teaching mode according to students' psychological characteristics. This paper analyzes the traditional characteristic recognition algorithm, and after improving its deficiencies, an improved characteristic extraction algorithm is proposed, based on the actual situation of classroom learning. This new algorithm can effectively improve the students' psychological feature prediction; with the support of this algorithm, a comprehensive analysis model with classroom behavior recognition and psychological feature recognition is constructed; also, the functional structure of the system is built up. Through experimental research, the model proposed in this paper is analyzed, and the experimental data has approved that the systemic model could play an important role in classroom teaching.

1. Introduction

Effective mental health education has become the top priority of every school's education work. In many research results, there is a common point of view that everyone recognizes, which is to take the mental activity course as an important or primary way to implement mental health education. Regardless of the form, effect, and degree of popularity among students, the actual effect of the mental activity course is obvious. The implementation of the course must be achieved through standardized teaching. In order to make the mental activity course really play its role in group tutoring, it is very necessary to do an in-depth study on the teaching design of the mental activity course. In reviewing previous studies, we found that there are many studies on the goals, functions, principles, historical development, and approaches of the mental activity course. However, there are obviously fewer studies on the age-specific mental activity course of high school students, and there are very few researches on how to design a mental activity course teaching model suitable for student development according to the features of high school students. Therefore, combining the author's own daily education practice and analyzing the

design of high school mental activity course from a microerspective are the research purpose of this paper [1].

How the level of teaching design of mental activity course will directly affect the tutoring effect of the class. Moreover, effective and successful teaching design is an important prerequisite and guarantee for successful teaching. Therefore, to carry out related research on the teaching design of mental activity course has very important theoretical and practical significance for improving the teaching effect and tutoring effect of mental activity course. How to develop a mental activity course from a more detailed perspective, such as the analysis of activity concepts, the positioning of activity goals, and the design of activity links, are all related content that a good instructional design should involve. As far as psychological counseling teachers are concerned, only with a well-designed teaching design can they better manage the entire class from a big idea. Therefore, for a successful mental activity course, the teacher first calms down and makes meticulous instructional design is an important first step. Effective classroom management is an important guarantee for a good lesson. Only by controlling and managing the classroom can the effectiveness of a lesson be guaranteed. This is true for classes in other

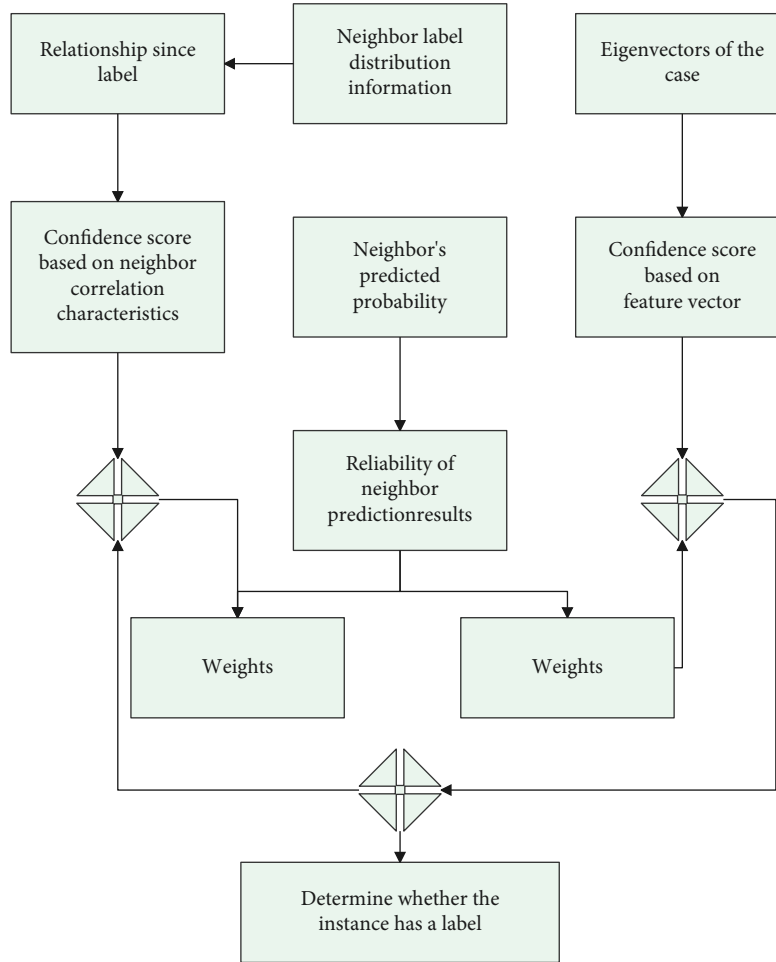


FIGURE 1: Design of MLNB algorithm.

subjects, especially for mental activity courses. Teachers in mental activity classrooms must be aware of classroom management. In view of the excessiveness of each link, the feedback of students answering questions, the construction of classroom group motivation, etc., all these require the teachers of the mental activity classroom to do well [2]. In classroom activities, different stages of teaching activities constitute the entire classroom. A mental activity course can be divided into a group warm-up phase, a group conversion phase, a group work phase, and a group end phase. Moreover, each stage has its tasks and requirements. At the same time, only when the activities of several stages are thoroughly developed and excavated can a section of mental activity course that is effective and touches students' hearts can be formed [3].

According to the actual needs of current teaching, based on machine learning algorithms, this paper combines the needs of teaching mental activity analysis to link mental activities and students' classroom behaviors to construct a research model of the correlation between students' classroom behaviors and mental activities based on improved machine learning and explores the correlation between students' classroom behaviors and mental activities.

2. Related Work

This paper uses improved machine learning as the basic algorithm to analyze the correlation between students' classroom behaviors and mental activities. There are many related researches on machine learning in factor correlation analysis, and these literatures are analyzed below.

The literature [4] aggregated user characteristic information from Flickr, Twitter, and Coca Cola to prove that its crossdomain user modeling strategy has a great impact on the recommendation quality in a cold start environment. This is one of the earliest studies to integrate user data from multiple fields to improve recommendation performance. Due to the heterogeneity of user preferences in different fields, literature [5] unified the social labels and semantic concepts of users in different systems. The literature [6] constructed a user relationship network and then used a random walk to identify the neighbor users of the target user in the source field, which is used for collaborative filtering recommendation in the target field. The literature [7] mapped labels to emotional categories and then used traditional content-based recommendation methods to make crossdomain recommendations. The literature [8] aggregated the

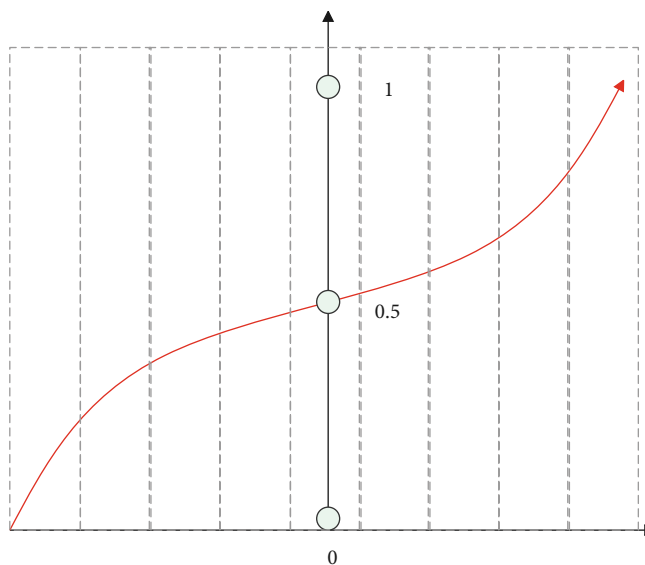


FIGURE 2: The process of sigmoid-fitting real value being mapped to probability.

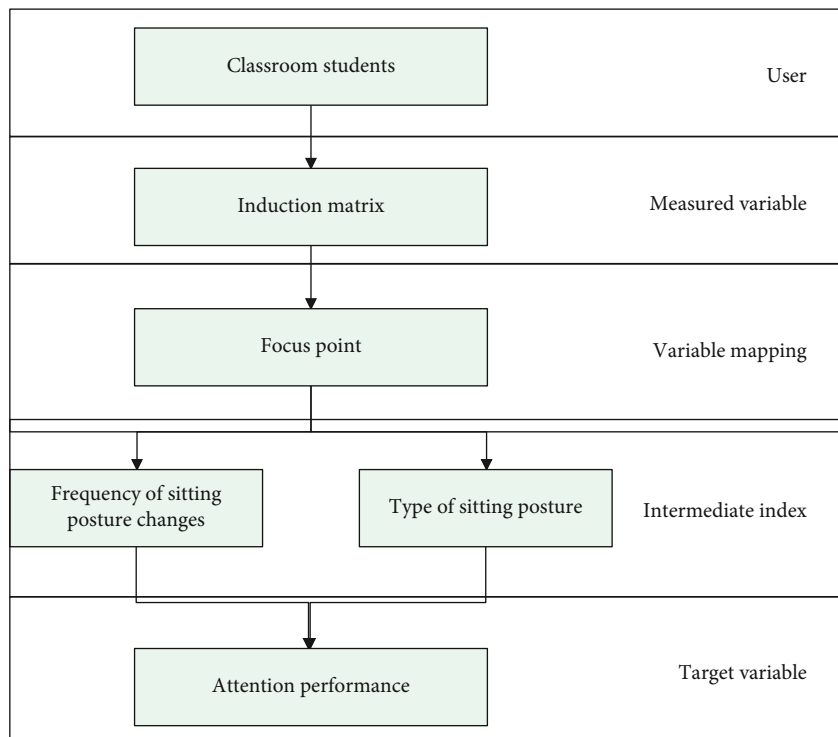


FIGURE 3: Evaluation model of mental activity behavior.

K-nearest neighbors of the target user in each field by weighting, which is used for collaborative filtering recommendation in the target field.

The literature [9] proposed the “Code Book Transfer (CBT)” method to achieve crossdomain collaborative filtering. Moreover, by compressing the cluster-level user item matrix scoring mode into a compact form, it obtained knowledge from the evaluation matrix of the auxiliary domain. In addition, it reconstructed the sparse matrix of the target field by extending the codebook and applied it to

the real-world data set and proved that the recommendation effect of the “codebook migration method” has obvious advantages. The literature [10] proposed a three-factor factorization model based on joint nonnegative matrix. The model uses an effective alternate minimization algorithm to enhance the crossdomain recommendation effect, and it can not only learn the scoring mode shared between domains but also flexibly control the sharing level. The experimental results obtained on real-world data sets show that the proposed model outperforms many existing

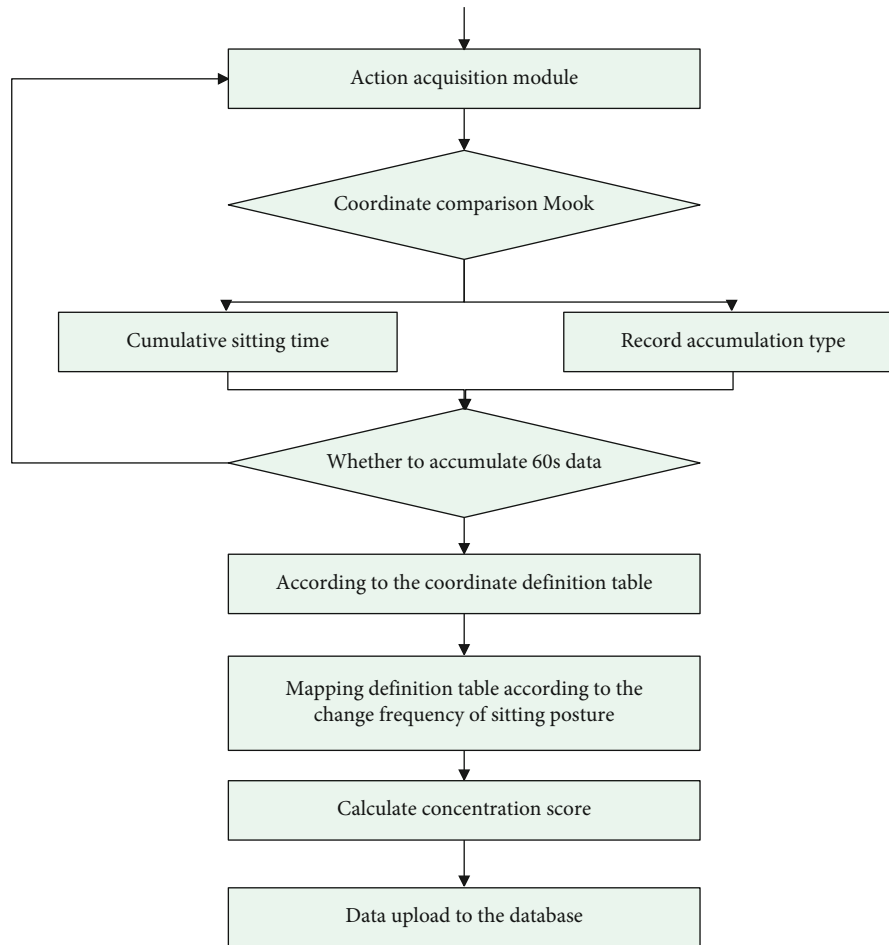


FIGURE 4: Flow chart of calculation algorithm for individual mental activity level of students.

methods in crossrecommendation tasks. The literature [11] proposed a crossdomain model that uses Factorization Machine (FM) to obtain additional knowledge from auxiliary domains. The model has the ability to encode domain-specific knowledge with real-valued feature vectors and can make better use of the interactive mode in the auxiliary domain, and its result is more advantageous than the existing crossdomain recommendation methods. Aiming at the crossdomain recommendation problem between users and item overlapping systems, literature [12] proposed a matrix factorization method to learn the potential feature vectors of users and items. This method assumes that there is potential consistency between systems; for example, the potential vectors in different systems all obey the same Gaussian distribution. The experimental results obtained on the rating data sets of movies, books, and electronic products show that the results obtained by this method are better than those obtained by the previous algorithms.

The literature [13] proposed a universal Crossdomain Triadic Factorization (CDTF) model based on the “user-product-domain” triad, which can better capture the relationship between user factors and project factors in a specific field. Moreover, it has designed two algorithms to take advantage of users’ explicit and implicit feedback and at the same time introduces genetic algorithms to optimize

the influence relationship between domains. The experimental results obtained on two real data sets verify the performance of the proposed CDTF model. The literature [14] proposed a universal crossdomain recommendation framework that combines social network information with crossdomain data. This framework uses the combination of tensor factorization and social relationship regularization to form a clustering hierarchical tensor of multiple domains. The results of experiments conducted on real-world data sets show that the proposed framework is better than previous methods. The literature [15] proposed a general framework for content-based crossdomain recommendation, which is an effective feature enhancement method. It uses the user’s metadata features to achieve domain adaptation and conducts experiments on the LinkedIn data set to give users work suggestions. In 2016, literature [16] proposed an adaptive model based on transfer learning to predict user behavior in different fields. Moreover, experiments were conducted on tweets in four different fields of IMDB, YouTube, Goodreads, and Pandora to find actively participated tweets. The results show that transfer learning can improve performance.

In real-world data sets, the basic challenge of multilabel classification is the effective use of label correlation. Labels often appear in pairs, and from the perspective of learning

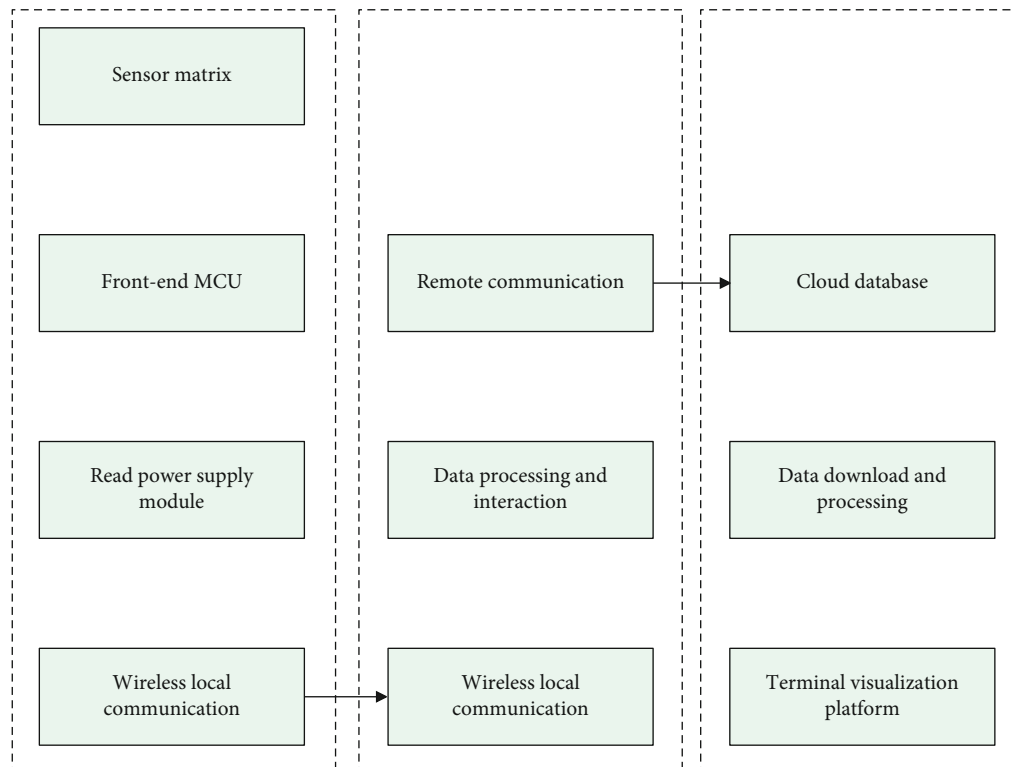


FIGURE 5: Overall framework of the system.

and prediction, their correlation provides useful information other than basic information for the sample. Therefore, considering the correlation of labels is beneficial to the accuracy of the algorithm and improves the hit rate of prediction. Literature [17] proposed that label correlation can improve the accuracy of multilabel classifiers. The more relevant the labels considered, the higher the complexity of the model. If only part of the label correlation is considered, the true dependency relationship cannot be captured. If all the correlations are considered, the complex relationship of the labels is more difficult to deal with.

3. Comprehensive Label Correlation Features and Sample Features of Student Mental Activities and Classroom Behavior

The two conversion strategies for multiple labels are as follows: (1) LP method: it changes the multilabel problem into the multiclass problem, regards the different label sets as the new category, converts the data into the multiclass set, and uses the multiclass classification to solve the problem; (2) BR method: it transforms multiple labels into binary classification problem solving for each single label. For each label, the instance associated with the label is placed in one class and the rest in another class. Since the binary classification algorithm can be processed in parallel when the high-order label correlation is added, which is beneficial to improve the computational efficiency, the strategy of adding correlation to the binary classification prediction results is adopted.

BR method can select different algorithms, such as decision tree, random forest, SVM, and neural network.

If the intersection of similar samples is relatively small, then the similarity is not reliable. In order to improve the reliability of the prediction results, it is necessary to reduce the impact of the unreliable results. Therefore, this paper adds weight to the neighbor prediction results.

The MLNB-BR multilabel classification algorithm is proposed in this paper, and the high-order label correlation is added to the feature-based binary classification method. The overall design of the algorithm is shown in Figure 1. The BR method is used to classify the feature data set. First, the method predicts the probability p_f of the existence of label l_k based on the feature vector and reasonably adds the label correlation information to the result of feature vector prediction to improve the classification accuracy. Due to the different reliability of neighbor prediction results, the classification results based on neighbor correlation features are dynamically adjusted, and the neighbor label dependency features are added to the features to modify the classification results [18].

When the classification based on neighbor features is unreliable, the uncertainty of the neighbor prediction result is calculated, and the uncertainty probability of the neighbor prediction is used to correct the classification result of the neighbor instance. For different labels and different samples, the weights based on the original features and the weights based on the neighbor features are dynamically adjusted. If the reliability of classification based on neighbor features is high, choosing a larger value for the prediction result weight

ω_1 of neighbor features can help preserve good neighbor label relationships p_r [1]. At the same time, the result p_f is adjusted based on the original feature. If the reliability of classification based on neighbor features is low, the prediction result weight ω_2 of the feature vector can increase the influence of the original feature and help correct the error of the neighbor feature. Therefore, the correlation of neighbor labels can be integrated with the classification results based on features to improve classification performance.

$$\begin{aligned} p &= \omega_1 \cdot p_r + \omega_2 \cdot p_f, \\ \omega_1 + \omega_2 &= 1. \end{aligned} \quad (1)$$

The greater the probability that the neighbor predicts the occurrence or nonoccurrence of the label l_k , the less the uncertainty of the result of the neighbor classification, the greater the probability that the label is correctly predicted, and the more reliable the result based on the neighbor classification. On the contrary, the smaller the probability of occurrence or nonoccurrence of the neighbor prediction label l_k , the greater the uncertainty of the result of the neighbor classification, the lower the probability that the label is correctly predicted, and the lower the reliability of the classification result based on the neighbor. For each sample, the neighbor information can be counted according to the principle of maximum posterior probability, and the probability of occurrence of label l_k can be obtained. For each sample x and each label l_k , the reliability of neighbor information is calculated according to the probability of the occurrence and nonoccurrence of the neighbor's predicted label l_k . The uncertainty of event occurrence can be measured by information entropy. The calculation formula of information entropy is

$$H(x) = - \sum_{x \in X} p(x) \log p(x). \quad (2)$$

The tests on a large number of data sets show the validity of the maximum posterior probability. The formulas for calculating the probability of occurrence and nonoccurrence of tags are, respectively [19]

$$\begin{aligned} p_1(x) &= p(H_1^l) p(E_{\vec{c}_i(l)}^l | H_1^l), \\ p_0(x) &= p(H_0^l) p(E_{\vec{c}_i(l)}^l | H_0^l). \end{aligned} \quad (3)$$

MLKNN algorithm that combines the traditional KNN method and Bayesian method to deal with multilabel classification problem is proposed. The MLKNN method mainly extracts label information from neighbors to predict sample labels and counts the accuracy of neighbor prediction to evaluate the probability of label occurrence and nonoccurrence. The main idea of this method is to get the label information of unknown instance K -nearest neighbor samples by using the KNN method and to get the label information of unknown instance K -nearest neighbor samples by counting the label distribution information according to the neighbor label set. After that, it counts the label distribution information based

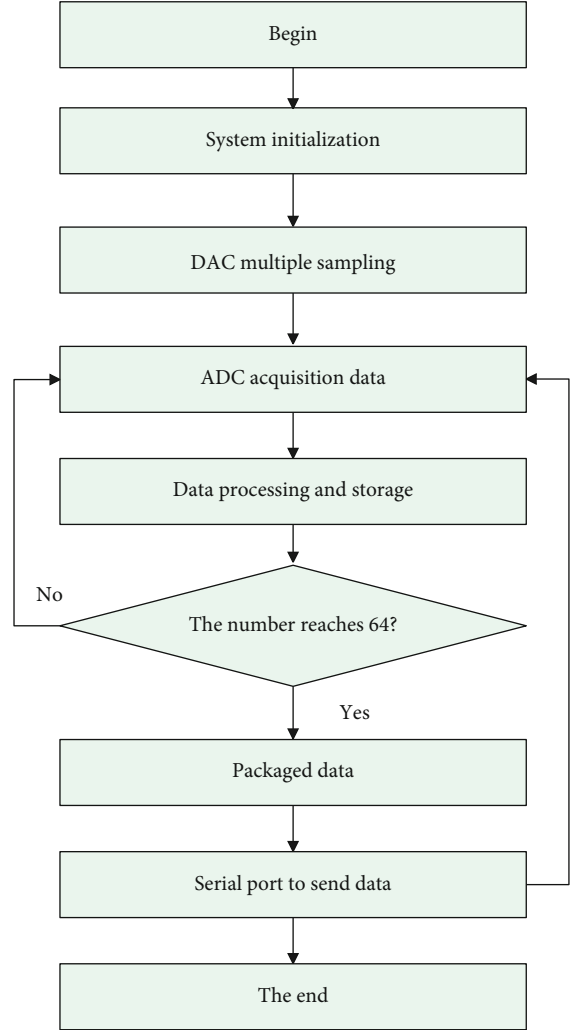


FIGURE 6: The working flow chart of the master single-chip microcomputer.

on the neighboring label sets and uses the probability of occurrence and nonoccurrence of each label l_k in the neighboring samples as the basis for classification. This paper considers that the MLKNN algorithm can only reflect the distribution features of neighbor samples, and it can be regarded as an attribute of neighbor samples. In this paper, the maximum a posteriori probability is used to count the probability of occurrence and nonoccurrence of neighbor prediction label, and the uncertainty of neighbor information is obtained through information entropy.

The modeling process of MLKNN is as follows: we give an instance X and its corresponding label set $Y_x \in Y$, Y is the set of all labels, and the total number of labels is defined as n . If there is a label l_k in Y_x , then $Y_x(l) = 1$, otherwise, $Y_x(l) = 0$. $N(x)$ is the set of K -nearest neighbors of instance X . For instance X , there are $\vec{c}_x(l)$ nearest neighbors of X with label l_k . The calculation formula is

$$\vec{c}_x(l) = \sum_{a \in N(x)} \vec{y}_a(l), \quad l \in Y. \quad (4)$$

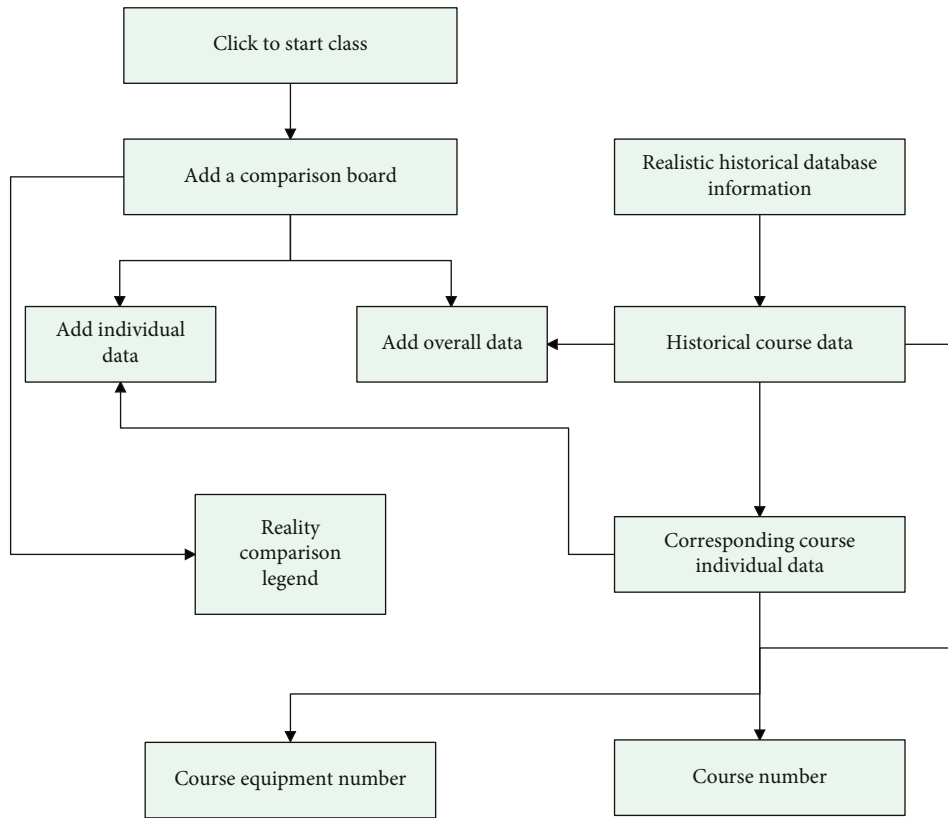


FIGURE 7: Flow chart of historical data module.

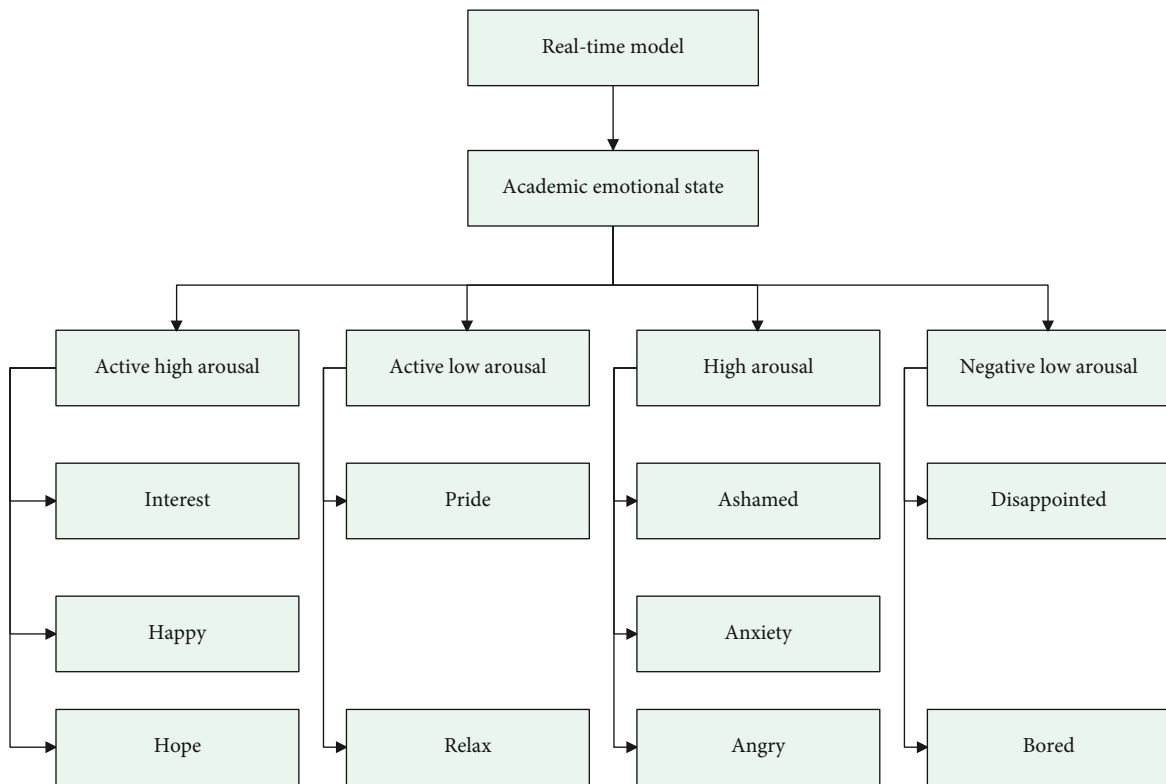


FIGURE 8: Real-time learner model.

TABLE 1: Statistical table of students' classroom behavior feature recognition and mental activity feature recognition.

No.	Behavioral feature recognition	Mental feature recognition	No.	Behavioral feature recognition	Mental feature recognition	No.	Behavioral feature recognition	Mental feature recognition
1	76.8	80.5	29	86.3	75.9	57	87.5	73.6
2	75.7	74.9	30	81.5	70.9	58	82.8	81.5
3	73.4	81.8	31	76.7	74.2	59	77.4	78.1
4	79.6	70.8	32	71.6	81.6	60	78.2	80.8
5	77.9	78.4	33	85.6	75.4	61	73.4	74.0
6	74.5	80.3	34	81.8	73.5	62	76.8	72.5
7	85.8	78.5	35	84.6	73.9	63	87.1	71.9
8	85.3	78.9	36	83.2	79.0	64	77.9	77.4
9	72.1	70.1	37	80.6	79.1	65	74.2	79.6
10	80.1	70.3	38	77.3	71.9	66	84.3	72.7
11	71.8	78.9	39	86.2	70.5	67	76.1	73.7
12	83.7	74.8	40	85.8	80.7	68	75.3	75.6
13	87.7	80.4	41	77.3	71.6	69	87.1	72.4
14	71.0	74.5	42	72.9	70.1	70	78.3	75.8
15	79.9	76.5	43	88.6	73.0	71	88.6	79.6
16	77.1	78.5	44	84.3	78.1	72	80.7	78.8
17	84.6	80.7	45	75.8	74.8	73	87.1	78.0
18	83.3	78.3	46	84.8	78.2	74	74.2	75.7
19	81.1	79.1	47	87.9	69.3	75	77.4	76.6
20	72.8	76.3	48	74.1	78.3	76	82.6	70.4
21	84.1	70.4	49	81.4	70.8	77	73.0	78.3
22	74.5	71.4	50	74.1	72.1	78	80.2	75.5
23	73.8	80.2	51	86.3	70.6	79	86.9	71.7
24	71.0	81.9	52	72.3	72.7	80	80.4	79.4
25	83.3	75.6	53	84.1	75.6	81	71.4	76.3
26	82.3	71.7	54	78.2	77.4	82	73.1	81.6
27	81.2	75.9	55	76.4	70.5	83	85.8	80.3
28	78.8	75.0	56	73.3	70.3	84	77.5	75.7

For the test instance t , the K -nearest neighbor set $N(t)$ is first obtained. For the label l_k , if there is a label l_k in the instance t , it is recorded as H_1^l , and if there is no label l_k in the instance t , it is recorded as H_0^l . We define E_j^l as samples that contain label l in K neighbors. Based on the vector \vec{c}_x , the formula for obtaining whether the instance t has a label l_k through the maximum posterior probability criterion and the Bayes criterion is

$$\vec{y}_t(l) = \arg \max_{b \in \{0,1\}} p(H_b^l | E_{\vec{c}_t(l)}^l), \quad l \in Y. \quad (5)$$

After using the Bayesian formula to transform, we get

$$\begin{aligned} \vec{y}_t(l) &= \arg \max \frac{p(H_b^l) p(E_{\vec{c}_t(l)}^l | H_b^l)}{p(E_{\vec{c}_t(l)}^l)} \\ &= \arg \max_{b \in \{0,1\}} p(H_b^l) p(E_{\vec{c}_t(l)}^l | H_b^l). \end{aligned} \quad (6)$$

Among them, $p(H_b^l)$ represents the prior probability of whether the instance t has a label l_k . When $b = 1$, $p(H_1^l)$ is equal to the number of samples with label l_k divided by the total number of samples.

$$p(H_1^l) = \frac{\sum_{i=1}^m \vec{y}_{x_i}(l)}{m}. \quad (7)$$

When $b = 0$, $p(H_0^l)$ is equal to the number of samples without label l_k divided by the total number of samples.

$$p(H_0^l) = 1 - p(H_1^l). \quad (8)$$

The posterior probability is calculated as follows:

$$p(E_{\vec{c}_t(l)}^l | H_1^l) = \frac{c^{[j]}}{\sum_{p=0}^k c^{[p]}}. \quad (9)$$

First, for each label l , we count the number $c^{[j]}$ ($j = 1, 2, 3, \dots$) of j neighbors with label l_k on the entire data set and

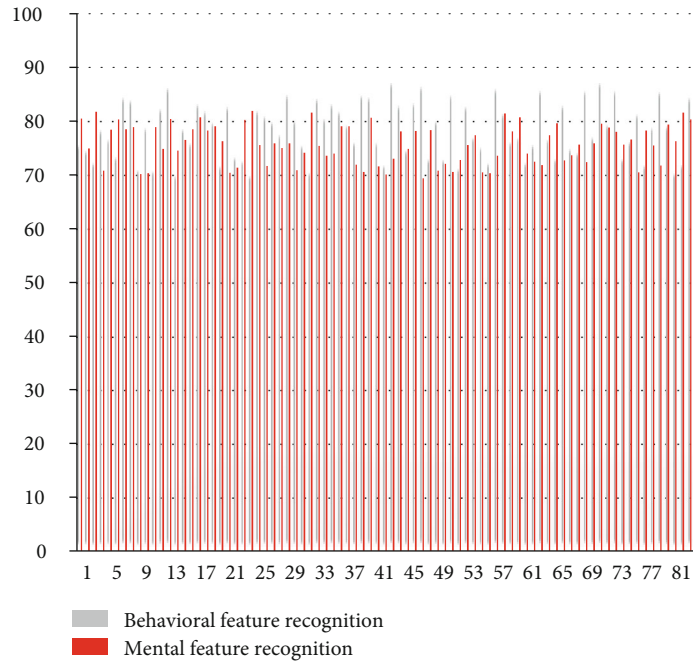


FIGURE 9: Statistical diagram of students' classroom behavior feature recognition and mental activity feature recognition.

the sample also with label l_k . If there are j neighbors in k neighbors with label l_k , then $c[j] = c[j] + 1$. Then, we count the proportion of j neighbors with label l_k and their own have label l_k in the overall sample.

Since it is necessary to synthesize neighbor correlation features and sample features and convert the output result of the classifier into a probability, different confidence scores can be synthesized according to the weight. This paper uses Platt's sigmoid-fitting method to map the real-valued output to the $[0, 1]$ probability space. The method of sigmoid-fitting real-valued output into probability is as follows Figure 2:

The function that maps real values to the range of $[0, 1]$ is

$$p = \frac{1}{1 + e^{(Af+B)}}. \quad (10)$$

Parameters A and B are the parameters to be fitted, used to adjust the size of the output value, and f is the SVM output value. On the training set, the maximum likelihood estimation is used to fit the parameters A and B in the formula, and the probability value on the training set is defined as

$$t_i = \frac{y_i + 1}{2}. \quad (11)$$

After smoothing, N_+ represents the number of positive samples, and N_- represents the number of negative samples:

$$t_i = \begin{cases} \frac{N_+ + 1}{N_- + 2} & \text{if } y_i = +1, \\ \frac{1}{N_- + 2} & \text{if } y_i = -1. \end{cases} \quad (12)$$

By minimizing the negative log likelihood function, A and B are solved to obtain the mapping function:

$$\min - \sum_i t_i \log(p_i) + (1 - t_i) \log(1 - p_i). \quad (13)$$

The prediction results of SVM based on feature vector and SVM based on label correlation are

$$\begin{aligned} p_r &= \omega_l^T \cdot x + b_l, \\ p_f &= w * x + b^*. \end{aligned} \quad (14)$$

According to the probability output function corresponding to the SVM, the standard SVM output value p_f is converted into p_f , and p_r is converted into p_r . The weight of neighbor feature classification is calculated as

$$\begin{aligned} \omega_1 &= -(p_1(x) \log p_1(x) \mp (x) \log p_0(x)), \\ \omega_2 &= 1 - \omega_1. \end{aligned} \quad (15)$$

Among them,

$$\begin{aligned} p_1(x) &= p(H_1^l) p(E_{\bar{c}_i(l)}^l | H_1^l), \\ p_0(x) &= p(H_0^l) p(E_{\bar{c}_i(l)}^l | H_0^l). \end{aligned} \quad (16)$$

In the original feature classification result, the neighbor label correlation is added to obtain the final comprehensive score:

$$p_x = \omega_1 \cdot p_f + \omega_2 \cdot p_r. \quad (17)$$

According to the comprehensive score, it is judged whether there is a label l_k .

$$y_l = \begin{cases} +1 & \text{if } p_x \geq \frac{1}{2}, \\ -1 & \text{if } p_x < \frac{1}{2}. \end{cases} \quad (18)$$

4. Analysis Model of the Correlation between Students' Classroom Behaviors and Mental Activities Based on Improved Machine Learning

Attention, as an important measure of student classroom effectiveness, is what the classroom behavior analysis designed for this paper expects to capture, quantify, and apply to group studies. Using the research method of this paper to monitor and quantify the students' attention level in class not only solves the cost of manpower and material cost of traditional attention measurement methods but also helps to record and quantify the attention data in the classroom in real time. Embodied cognition theory points out the interaction between body posture and cognitive state; that is, physiological indicators of body behavior can represent different cognitive levels. At present, many studies at home and abroad are exploring the use of different behavioral indicators as intermediate variables to characterize the level of attention. Based on the previous research background and related theories, this study uses two indicators of sitting posture type and sitting posture change frequency to represent attention performance. The evaluation model is shown in Figure 3.

The mapping calculation algorithm of the attention level of this research is detailed in flow Figure 4.

Based on the previous design of classroom behavior analysis, the specific implementation process of the system is described in detail. The implementation part of the technology includes the general description of the development environment and tools, the overall architecture of the system, the specific implementation scheme of each module of the system, and the corresponding user use process for the system application in the next chapter. The overall framework of the system is shown in Figure 5.

The work flow chart of the master single-chip micro-computer is shown in Figure 6.

TABLE 2: Statistical table of correlation between students' classroom behaviors and mental activities.

No.	Correlation	No.	Correlation	No.	Correlation
1	72.1	29	71.4	57	66.4
2	69.1	30	70.6	58	66.7
3	67.0	31	71.7	59	70.1
4	72.6	32	68.0	60	72.3
5	74.4	33	69.6	61	73.9
6	72.1	34	65.5	62	63.5
7	71.0	35	71.1	63	72.4
8	67.1	36	65.7	64	69.4
9	74.3	37	67.9	65	68.0
10	63.4	38	64.6	66	67.0
11	69.7	39	71.5	67	74.0
12	66.1	40	70.7	68	66.7
13	67.3	41	70.7	69	74.8
14	69.2	42	71.0	70	63.4
15	72.6	43	68.1	71	73.7
16	71.2	44	70.5	72	65.2
17	67.4	45	68.5	73	69.6
18	67.8	46	71.8	74	72.2
19	70.6	47	72.7	75	70.0
20	72.2	48	72.1	76	67.8
21	71.6	49	72.8	77	66.3
22	72.0	50	67.9	78	74.5
23	72.7	51	63.4	79	70.0
24	68.8	52	74.6	80	63.1
25	68.1	53	68.9	81	71.2
26	71.7	54	65.7	82	65.8
27	74.2	55	69.9	83	71.8
28	74.5	56	71.1	84	70.2

According to the work flow chart 6 of stm32l433c6, when the individual action collection module starts to work, the system is initialized first, and then, the analog signal is amplified by OPAMP operational amplifier in the MCU. At the same time, by cooperating with the DAC peripheral multichannel sampling, the dynamic adjustment of the level is realized. Then, it uses ADC to collect data and transfers the data from RX data register to memory through USART DMA transceiver driver. In the data acquisition process of ADC, the data of 64 pressure points is taken as a cycle. When the data amount reaches, it is packaged according to the MQTT gateway protocol and sent to the local field processing module through the serial port, and the next ADC data acquisition is carried out, and the operation is reciprocating.

Users can also view the historical data of all recorded courses through the terminal software platform, including the overall data of the classroom and the corresponding individual data, and get the visual legend by adding comparison, so as to conduct in-depth comparative analysis of the classroom situation. The process is shown in Figure 7.

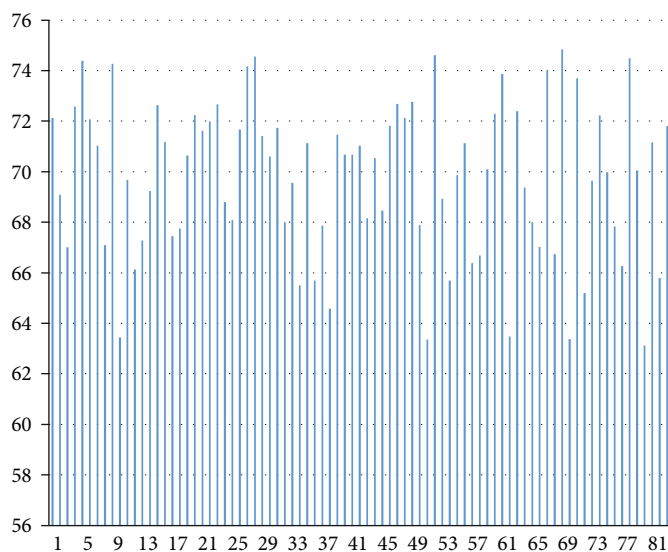


FIGURE 10: Statistical diagram of correlation between students' classroom behaviors and mental activities.

The real-time behavior data of learners are processed by using academic emotion theory, and the real-time learner model is constructed. Academic emotional state is mainly divided into positive and negative states. The positive state includes positive high arousal and positive low arousal, and the negative state includes negative high arousal and negative low arousal, as shown in Figure 8.

5. Correlation Analysis between Students' Classroom Behaviors and Mental Activities Based on Improved Machine Learning

After constructing the correlation analysis system between students' classroom behaviors and mental activities based on improved machine learning, this paper verifies the performance of the system and analyzes whether there is correlation between students' mental activities and their behaviors. In this paper, students' behavior features are identified and recorded, and then, students' mental activities are recorded through the system constructed in this paper. Firstly, the data collection performance of the system is evaluated, and the results are shown in Table 1 and Figure 9.

Through the above analysis, we can see that the correlation analysis system of students' classroom behaviors and mental activities based on improved machine learning constructed in this paper has high accuracy in data collection of students' classroom behavior features and students' mental activity features. On this basis, this paper studies the correlation between students' classroom behaviors and mental activities, and the results are shown in Table 2 and Figure 10.

From the above analysis results, we can see that there is a significant correlation between students' classroom behaviors and their 2-dimensional activities. Therefore, teachers can judge students' mental activities according to students' classroom behaviors, which is convenient for timely adjustment of teaching strategies and maximum improvement of classroom teaching effect.

6. Conclusion

In this paper, we improve the traditional feature recognition algorithm and show how to extract the label correlation features of neighboring instances and view the neighboring instance label distribution information as neighboring features and then use the neighboring features to calculate the label correlation. Moreover, this paper proposes a method that integrates label relevance features and sample features. The method obtains the reliability of the neighbor information by the magnitude of the probability of occurrence or nonoccurrence of the neighbor prediction labels and corrects the results of the neighbor features.

This study designs and implements analysis system of the correlation between students' classroom behaviors and mental activity based on improved machine learning, which consists of multiple smart hardware devices, a LAN off device, and a terminal visualization and analysis software. Moreover, the system can collect real-time data on classroom behavioral changes of multiple individuals in classroom teaching, transform them into group attention levels through calculation, and visualize them through terminal software. In addition, the system can be used as a real-time supplemental monitoring tool for instructors on classroom conditions, as well as a long-term digital record of classroom instruction and behavior analysis for educators and schools.

Data Availability

The labeled data set used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares no competing interests.

Acknowledgments

This study is sponsored by Research on the Issue of National Identity Education in Guangdong Colleges and Universities Promoting the Return of Young Students in Hong Kong Under the Perspective of Branding Vision, Research Project of Party Building in Study Conference of Party Building in Guangdong Colleges and Universities in 2019 (project No. 2020MB015), and Analysis on the Brand Development Strategy of Independent Colleges Guided by Party Construction, Key Research Topics of Party Building and Ideological and Political Education of Zhuhai College of Jilin University in 2019 (project No. 2019JZDJ001).

References

- [1] G. Wang, J. Sun, J. Ma, K. Xu, and J. Gu, "Sentiment classification: the contribution of ensemble learning," *Decision Support Systems*, vol. 57, pp. 77–93, 2014.
- [2] S. Vishwakarma and S. Kumar, "A lexical approach for tweets sentiment classification," *Journal of Applied Physics*, vol. 76, no. 3, pp. 1980–1982, 2015.
- [3] L. Dong, F. Wei, S. Liu, M. Zhou, and K. Xu, "A statistical parsing framework for sentiment classification," *Computational Linguistics*, vol. 41, no. 2, pp. 293–336, 2015.
- [4] F. Fang, K. Dutta, and A. Datta, "Domain adaptation for sentiment classification in light of multiple sources," *INFORMS Journal on Computing*, vol. 26, no. 3, pp. 586–598, 2014.
- [5] S. Zhou, Q. Chen, and X. Wang, "Fuzzy deep belief networks for semi-supervised sentiment classification," *Neurocomputing*, vol. 131, pp. 312–322, 2014.
- [6] S. M. Liu and J. H. Chen, "A multi-label classification based approach for sentiment classification," *Expert Systems with Applications*, vol. 42, no. 3, pp. 1083–1093, 2015.
- [7] L. Shang, Z. Zhou, and X. Liu, "Particle swarm optimization-based feature selection in sentiment classification," *Soft Computing*, vol. 20, no. 10, pp. 3821–3834, 2016.
- [8] A. Moreo Fernández, A. Esuli, and F. Sebastiani, "Distributional correspondence indexing for cross-lingual and cross-domain sentiment classification," *Journal of Artificial Intelligence Research*, vol. 55, pp. 131–163, 2016.
- [9] S. Harer and S. Kadam, "Sentiment classification and feature based summarization of movie reviews in mobile environment," *International Journal of Computer Applications*, vol. 100, no. 1, pp. 30–35, 2014.
- [10] Z. Zhang, Z. Wang, C. Gan, and P. Zhang, "A double auction scheme of resource allocation with social ties and sentiment classification for device-to-device communications," *Computer networks*, vol. 155, pp. 62–71, 2019.
- [11] Y. Li, J. Wang, S. Wang, J. Liang, and J. Li, "Local dense mixed region cutting + global rebalancing: a method for imbalanced text sentiment classification," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 7, pp. 1805–1820, 2019.
- [12] V. N. Phu, N. D. Dat, V. T. Ngoc Tran, V. T. Ngoc Chau, and T. A. Nguyen, "Fuzzy C-means for English sentiment classification in a distributed system," *Applied Intelligence*, vol. 46, no. 3, pp. 717–738, 2017.
- [13] M. Huang, Q. Qian, and X. Zhu, "Encoding syntactic knowledge in neural networks for sentiment classification," *ACM Transactions on Information Systems*, vol. 35, no. 3, pp. 1–27, 2017.
- [14] Y. Liu, J. W. Bi, and Z. P. Fan, "Multi-class sentiment classification: the experimental comparisons of feature selection and machine learning algorithms," *Expert Systems with Applications*, vol. 80, pp. 323–339, 2017.
- [15] M. Iqbal, A. Karim, and F. Kamiran, "Balancing prediction errors for robust sentiment classification," *ACM Transactions on Knowledge Discovery from Data*, vol. 13, no. 3, pp. 1–21, 2019.
- [16] M. E. Basiri, A. Kabiri, M. Abdar, W. K. Mashwani, N. Y. Yen, and J. C. Hung, "The effect of aggregation methods on sentiment classification in Persian reviews," *Enterprise Information Systems*, vol. 14, no. 9–10, pp. 1394–1421, 2020.
- [17] D. Wang, B. Jing, C. Lu et al., "Coarse alignment of topic and sentiment: a unified model for cross-lingual sentiment classification," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 2, pp. 736–747, 2021.
- [18] G. Rao, W. Huang, Z. Feng, and Q. Cong, "LSTM with sentence representations for document-level sentiment classification," *Neurocomputing*, vol. 308, pp. 49–57, 2018.
- [19] G. Lee, J. Jeong, S. Seo, C. Y. Kim, and P. Kang, "Sentiment classification with word localization based on weakly supervised learning with a convolutional neural network," *Knowledge-Based Systems*, vol. 152, pp. 70–82, 2018.