A Convolutional Neural Network-Based Predictive Model for Assessing the Learning Effectiveness of Online Courses Among College Students

Xuehui Zhang, Lin Yang*
School of Electronic Information and Automation, Guilin University of Aerospace Technology, Guilin, China

***Abstract*—With the development of artificial intelligence (AI) technology, higher education institutions usually consider both online courses and offline classrooms in the course design process. To verify the effectiveness of online courses, this study designed a deep learning model to analyze the learning behavior of online course users (college students) and predict their final grades. Firstly, our method summarizes several learning features that are used in machine learning models for predicting student grades, including the performance of users (college students) in online courses and their basic information. Based on nutcracker optimization algorithm (NOA), we designed a multi-layer convolutional neural network (CNN) and developed an improved NOA (I-NOA) to optimize the internal parameters of the CNN. Prediction mainly includes two steps: firstly, analyzing users' emotions based on their comments in online course forums. Secondly, predict the final grade based on the user's emotions and other quantifiable learning features. To validate the effectiveness of INOA-Based CNN (I-NOA-CNN) algorithm, we evaluated it using a dataset consisting of five different online courses and a total of 120 students. The simulation results indicate that compared with existing methods, the I-NOA-CNN algorithm has higher prediction accuracy, and the proposed model can effectively predict the learning effect of users.**

Keywords*—Convolutional neural network; nutcracker optimization algorithm; assessment of learning effectiveness; college students; online courses

I. Introduction

With the development of AI, online course platforms are widely adopted by higher education institutions. These free online courses not only promote the dissemination of knowledge, but also enhance the flexibility of users (college students) in learning. In addition, online courses can reduce the demand for hardware resources such as classrooms and laboratories for higher education institutions, thereby saving costs [1]. However, without face-to-face interaction, users may feel isolated and find it difficult to maintain learning motivation and engagement, especially for students with weaker self-learning abilities [2]. Therefore, it is necessary to evaluate the learning effectiveness of users in stages during their learning process, in order to identify problems and improve students' learning plans in a timely manner, thereby providing personalized learning support for students. Overall, evaluating the effectiveness of online course learning not only helps to improve the quality of education, but also enhances the transparency and trust of education, making it an indispensable part of the modern education system.

Predictive Learning Analysis (PLA) technology is of great significance for improving teaching methods, optimizing learning resources, and enhancing learning efficiency. PLA utilizes user learning performance and other related data to predict learning effectiveness through established models, such as logistic regression and deep neural networks. PLA is widely used in predicting grades, developing personalized learning paths, and predicting dropout rates. It plays a crucial role in analyzing users' learning processes, predicting grades, and identifying factors that may affect learning outcomes in advance. Based on PLA technology, Chen et al. designed a machine learning model that takes students' learning behavior as an input layer and evaluates the learning effectiveness of online courses by predicting their final grades [3]. Reference [4] also designed a machine learning model aimed at evaluating the learning effectiveness of users using cloud platforms for online learning, with the aim of detecting users' learning risks in the early stages.

Martínez-Caro analyzed the factors that affect user learning outcomes during online learning, including teaching content and methods, students' emotions, and personal factors such as their adaptability and self-management abilities [5]. Although the above factors may affect users' learning outcomes, students' adaptability and self-management abilities cannot be quantified. In addition, the large number of college students is not conducive to the statistical analysis of the above information. Therefore, Chen et al. established a machine learning model that takes quantifiable content such as the user's family size, family education support, and learning time as inputs, and the user's final grade as output. This provides a new approach for evaluating user learning effectiveness for online courses [6]. Hassan et al. designed a clustering prediction model with the aim of predicting students' final grades based on their behavior [7]. Although students' emotions directly affect the effectiveness of learning, the above studies did not consider using students' emotions as input variables to predict the final grades of online course platform users.

In response to the problem that existing data prediction models do not consider the influence of user emotions, this study designs a CNN based student performance prediction model, which is divided into two parts. In the first part, the

*Corresponding Author.
This work was sponsored by 2024 Guangxi University Young and Middle-aged Teachers' Research Basic Ability Enhancement Project (Project Number: 2024KY0790).

model evaluates user emotions through CNN algorithm based on student comments in online course forums. After obtaining the user's emotional results, the user's emotions, along with other behavioral characteristics such as learning duration and comment frequency, are used as input variables to predict the user's final grade. Fig. 1 shows the framework designed by this research institute. At present, predicting students' emotions through their facial expressions has become a mainstream emotion analysis method [8]. However, emotion analysis frameworks based on facial expressions not only require a large amount of computing resources, but also pose a challenge in terms of data privacy protection. Online course platforms often provide a forum for student communication. Therefore, other studies have adopted methods based on comment datasets for predicting student emotions [9]. In this study, we employed a sentiment analysis model based on a comment dataset.

Considering that some existing online course users' grade prediction models are only applicable to a single subject and do not take into account the model's generalization ability in other subjects. Therefore, in order to improve the generalization ability of the model, this study selected courses from five different disciplines and a dataset of 60 students to validate the designed model, aiming to support personalized course design while enhancing students' learning experience. The main contributions of this article are summarized as follows:

- This study designed a multi-layer CNN framework for predicting the final grades of online course users. At the same time, a prediction scheme was designed that first predicts emotions and then predicts grades.
- In order to improve the prediction accuracy of CNN algorithm, an optimization algorithm based on particle swarm optimization algorithm and NOA algorithm was designed to optimize the output layer weights of ConvNet.
- A dataset consisting of 5 different courses, 60 students, and a time span of 15 weeks was used to test the performance of the I-NOA-CNN algorithm. The results showed that compared with other existing methods, the I-NOA-CNN algorithm had the highest prediction accuracy.

The rest of this article is arranged as follows. In Section II, literature related to student performance prediction is reviewed. The method proposed in this article is presented in Section III. Section IV presents the results. Finally, Section V summarizes the entire text.

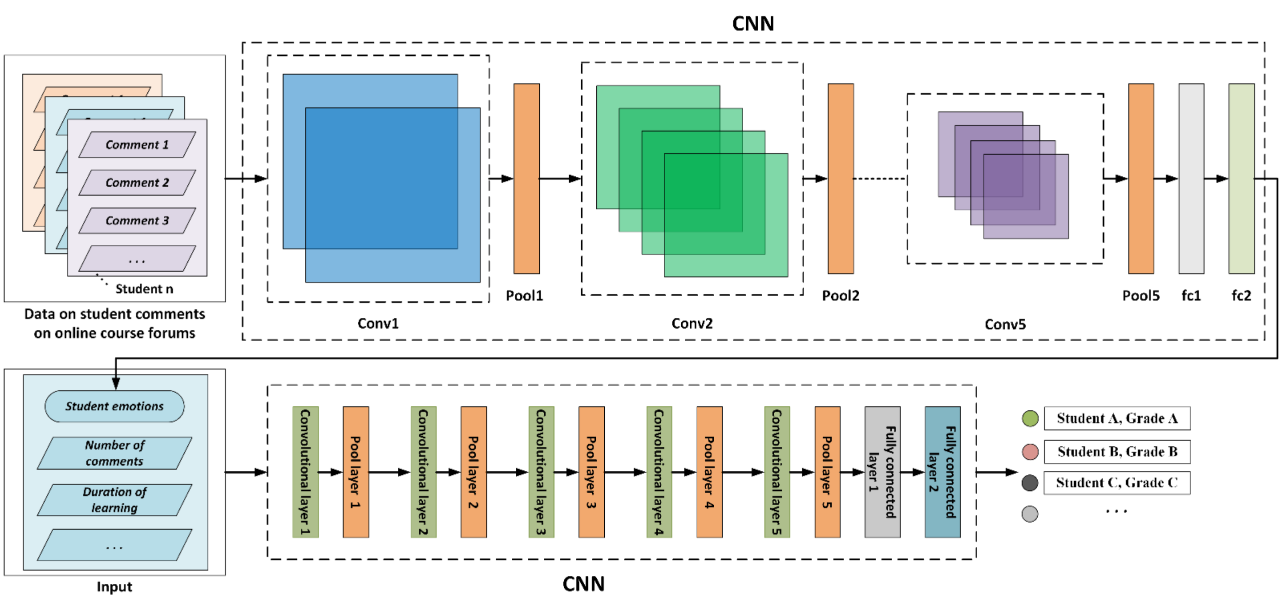

Fig. 1. A CNN-based architecture for predicting final grade of online course users.

II. Related Work

The rapid development of digital education and distance learning has made predicting user grades for online courses an important research area. The main purpose of this field is to use data analysis and machine learning techniques to predict students' performance in the course in advance, thereby helping educators and students themselves better adjust learning strategies and teaching methods [10]-[11]. As this study focuses on the analysis of online course user grades, a CNN architecture based on the I-NOA algorithm, and a scheme for predicting sentiment before predicting academic performance are designed. Therefore, this article focuses on reviewing relevant literature on neural network-based grade prediction, CNN algorithm, and text-based sentiment prediction.

A. Neural Network-based Prediction of Academic Performance

With the development of artificial intelligence technology, especially the widespread application of artificial neural networks, researchers have begun to explore how to use these advanced computational models to predict students' grades [12]. Neural networks, especially deep learning models, have been widely used for predicting academic performance due to their high efficiency and accuracy in processing large-scale datasets. Deep learning models are able to extract deep level features from students' online learning behavior, thereby predicting their future performance [13]. The author in [14] developed an early warning system tailored to the issue of whether students in specific subject courses can successfully pass the course assessment, which is essentially similar to the final grade prediction model. This study proposes a weighted voting combination strategy to improve the accuracy of predictions. By comprehensively utilizing time series data and personalized characteristics of students, this method significantly enhances the ability to predict the risk of students failing course assessments, providing education workers with a powerful tool to implement timely and effective teaching interventions. Gupta et al. also explored the method of using educational data generated by online learning platforms using machine learning technology to warn students of their academic performance. This study used Hidden Markov Models to analyze these numbers and ultimately established an efficient modeling framework that provides data-driven decision-making guidance for higher education institutions to achieve more sustainable educational development [15].

Xu et al. designed a machine learning model based on students' online behavior to address the problem of predicting their academic performance in online courses. This model can accurately predict students' final grades and design learning environments and activities that are suitable for students [16]. The author in [17] aims to predict students' mastery of knowledge by establishing a refined model of their learning process using conversation and relationship graphs. Additionally, graph convolutional networks are used to analyze students' acquisition of knowledge status, ultimately achieving the goal of predicting students' grades. The author in [18] designed a learning achievement prediction model based on reinforcement learning, which takes students' homework texts as input to predict their knowledge mastery. Another study combined artificial neural networks and fuzzy systems for a student performance evaluation system, and the results showed that the model had higher prediction accuracy than traditional statistical methods [19]. Although there has been some progress in predicting student academic performance based on artificial neural networks, there are still some challenges, such as data diversity and quality, model generalization ability, and interpretability issues. Therefore, this study focuses on improving the robustness and adaptability of predictive models.

B. *Convolutional Neural Network*

Although numerous studies have designed student performance prediction models, the accuracy of these models remains a challenge so far. Nie et al. designed a prediction model based on fuzzy reasoning theory [20], which combines meta heuristic algorithms with fuzzy logic to improve the prediction accuracy of the model. Lu et al. designed an improved CNN algorithm to analyze students' behavioral characteristics and predict their final grades [21]. The author in [22] also developed an improved CNN model aimed at evaluating students' learning outcomes and satisfaction. The author in [23] designed an improved artificial neural network aimed at analyzing the behavior of research-oriented students in higher education institutions, in order to identify factors that affect student learning outcomes.

The CNN algorithm has also been widely applied in other fields, which provides a broad idea for the algorithm design in this article. Song et al. designed an improved CNN algorithm to improve the prediction accuracy of wind power generation [24]. Naulia et al. designed a CNN framework based on optimization algorithms, also aimed at improving the prediction accuracy of CNN, and demonstrated through simulation examples that the improved CNN framework has potential advantages compared to gradient descent methods [25]. The author in [26] focuses on the problem of mechanical fault diagnosis and designs a lightweight CNN algorithm aimed at improving the prediction efficiency and accuracy of CNN.

At present, in some of the latest research, relevant researchers have begun to use optimization algorithms to optimize the parameters of CNNs. In order to solve the problem that the computation time of CNN increases exponentially with the size of the problem during the prediction process, [27] adopted the method of first using particle swarm optimization algorithm to reduce the solution space and then using CNN for prediction. This method significantly shortens the computation time of CNN in computing high-dimensional problems. Li et al. used differential evolution (DE) algorithm to optimize the parameters of CNN, aiming to improve the accuracy of CNN in music sentiment analysis problems [28]. The author in [29] establishes a deep learning model based on metaheuristic optimization algorithm, aiming to use the optimization algorithm to formulate the optimal learning strategy for the deep learning model. Meanwhile some state-of-the-art metaheuristics likewise provide ideas for algorithm improvement [30]-[31].

C. *Text-based Sentiment Prediction*

Lin et al. designed a multimodal learning model for sentiment analysis using text and images. Compared to models that only use images for sentiment analysis, the multimodal learning model performed better [32]. Reference [33] also designed a method for sentiment analysis using text and images, with a significant contribution being the improvement of the problem of modality loss. Alshaikh et al. designed a text-based sentiment analysis system to address the issue of sentiment analysis in education. This system is used to identify valuable information while providing personalized learning for students [34]. The author in [35] conducted research on sentiment analysis of comments on social platforms. This study aims to identify the emotions conveyed by comments by extracting comment information from social platforms and optimizing deep neural networks through gradient descent algorithm.

III. Proposed Framework

In this study, we first improved Abdel-Basset et al.'s NOA algorithm and designed an I-NOA algorithm [31]. Then, we use the I-NOA algorithm to optimize the weights of the output layer of the CNN algorithm. The architecture designed for this study is mainly divided into two layers. In the first layer, the I-NOA-CNN algorithm is used to analyze students' emotions based on the text they comment on in the online course platform's comment section. In the second layer, we predict students' final grades based on their emotions, number of comments, course content, and class time.

A. Improved Nutcracker Optimization Algorithm

The traditional NOA algorithm imitates the process of Nutcracker collecting food, which is divided into four stages: exploration stage, storage stage, cache and search stage, and recovery stage. We have focused on improving the exploration phase and storage stage phase, aiming to enhance the convergence speed and accuracy of the NOA algorithm.

1) Exploration stage: In the exploration phase of the I-NAO algorithm it is necessary to initialize the parameters, which include the maximum number of iterations k_{max}, the number $J=\{1,2,\cdots,j_{max}\}$ of individuals, the dimensions $G=\{1,2,\cdots,g_{max}\}$ of the individuals and the positions $\psi_{g,j}$, $\forall g \in G, \forall j \in J$ of the individuals.

Due to the fact that traditional NOA algorithms rely on randomly selected individuals and the average value of the g-th dimension of all individuals during the individual position update process in the exploration phase, it is not conducive to quickly finding the region where the optimal solution is located in the later optimization process. Therefore, in this study, the update strategy of particle swarm optimization (PSO) algorithm and the exploration strategy of NOA algorithm are integrated to improve convergence speed and accuracy. Randomly generate a random number R_{exp} in the [0,1] interval, and generate a random number ϕ that decreases with the number of iterations. If $R_{exp} \le \phi$, update the position according to the particle swarm optimization exploration strategy.

$$\vec{E}_j^{k+1} = \begin{cases} \vec{E}_{j,g}^k, & R_1 \le R_2 \\ \begin{cases} E_{m,g}^k + \beta\left(\vec{E}_{ra,g}^k - \vec{E}_{rb,g}^k\right) \\ +\alpha\left(R_1 \times \left(M_g - N_g\right)\right), \ k \le \dfrac{k_{max}}{2} \\ c \times R_2 \times \left(\vec{E}_{best,g}^k - \vec{E}_{j,g}^k\right), \ k > \dfrac{k_{max}}{2} \end{cases} & , R_1 > R_2 \end{cases} \tag{1}$$

where, $\vec{E}_{j,g}^k$ is the g-th dimension of the j-th individual in the k-th iteration process. $\vec{E}_{ra,g}^k$ and $\vec{E}_{rb,g}^k$ are the g-th dimension of randomly selected individuals. $\vec{E}_{m,g}^k$ is the average of the g-th dimension of all individuals. $\vec{E}_{best,g}^k$ is the g-th dimension of the optimal individual among all. R_1 and R are random numbers on the interval [0,1]. c is the learning factor of PSO. β is a random flight step size. α is a random number that follows a normal distribution. M_g and N_g are two vectors.

2) Storage stage: This stage imitates the process of Nutcracker storing food. In this process, we introduced the hunting phase of the grey wolf optimization algorithm and designed a storage strategy based on grey wolf hunting. Before the start of this stage, we generated random numbers R_3, R_4, R_5 in the [0,1] interval. Based on the Levy flight strategy, we generated random number L_{evy}. If $R_{exp} > \phi$, perform storage operations based on grey wolf hunting according to (2).

$$\vec{E}_g^{k+1} = \begin{cases} \vec{E}_g^k + \beta \times \left(\vec{E}_{best,g}^k - \vec{E}_g^k\right) \times \left|L_{evy}\right| + R_1 \times \left(\vec{E}_{ra,g}^k - \vec{E}_{rb,g}^k\right), & R_3 \le R_4 \\ \vec{E}_{best,g}^k + \beta \times \left(\vec{E}_{ra,g}^k - \vec{E}_{rb,g}^k\right), & R_3 \le R_5 \\ \dfrac{\vec{E}_{ra,g}^k + \vec{E}_{rb,g}^k + \vec{E}_{rc,g}^k}{3}, & otherwise \end{cases} \tag{2}$$

3) Cache and search stage: The cache search phase is designed to further expand the algorithm's exploration of the solution space. This stage relies on the reference points (candidate solutions) in the Nutcracker solution space. The definition of reference point (RP) is as follows:

$$RP = \begin{bmatrix} \overrightarrow{RP}_{j,1}^k & \overrightarrow{RP}_{j,2}^k \\ \vdots & \vdots \\ \overrightarrow{RP}_{j\,max,1}^k & \overrightarrow{RP}_{j\,max,2}^k \end{bmatrix} \tag{3}$$

The calculation method of $\overrightarrow{RP}_{j,g}^k$ is as follows.

$$\overrightarrow{RP}_{j,g}^k = \vec{E}_{j,g}^k + \chi \times cos(\theta) \times \left(\vec{E}_{ra,g}^k - \vec{E}_{rb,g}^k\right) \tag{4}$$

where, χ linearly decreases with the number of iterations from 1 to 0.

4) Recovery stage: The recovery phase involves evaluating and selecting existing solutions based on corresponding strategies. Fig. 2 shows the strategy used in the recovery phase of the I-NOA algorithm. When the maximum number of iterations is reached, the recovery phase no longer selects the optimal solution and directly outputs the final solution.

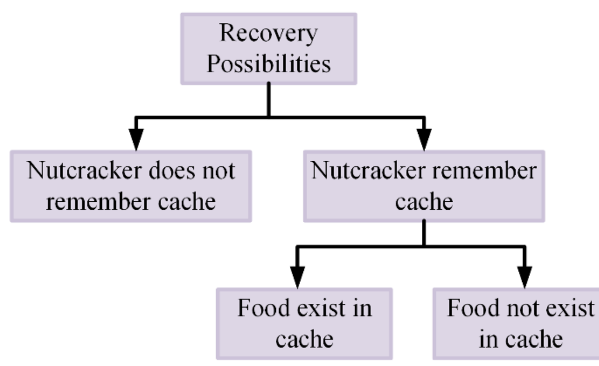

Fig. 2. Diagram of updates during the recovery phase.

B. Improved Convolutional Neural Network

At this step, we mainly calculate the loss function of the training phase of the training model. In this study, the loss function is defined as the difference between the predicted value of the CNN algorithm and the true value. The CNN algorithm in this study includes five convolutional layers, five pooling layers, and two fully connected layers. The convolutional layer is mainly responsible for extracting features from students' relevant data. The fully connected layer is based on the activation function to classify students' emotions and grades. Fig. 3 shows the framework of a fully connected layer. The loss function used in this article is shown below.

$$F_{lc}\left(\hat{f},f\right)=-\sum_{i} f_i \times log\left(\hat{f}_i\right)+\left(1-f\right)_i \times log\left(1-\hat{f}_i\right),\ \forall i \in I \tag{5}$$

where, f is the true value of the label. $\hat{f}$ is the probability that the CNN algorithm predicts the current label correctly.

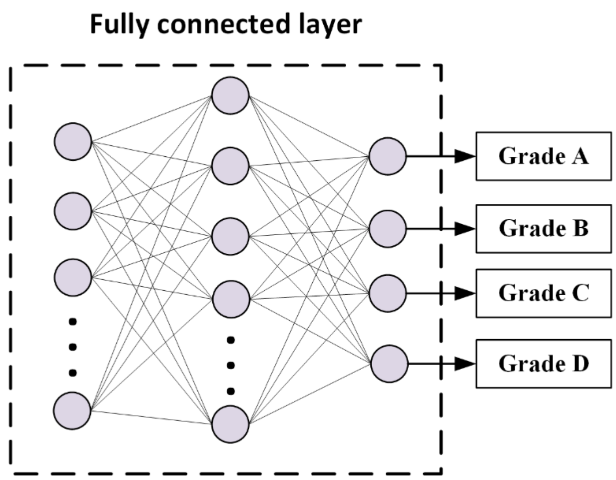

Fig. 3. Schematic diagram of the structure of the fully connected layer.

IV. Results and Discussion

To validate the performance of the I-NOA-CNN algorithm, we conducted simulation experiments based on a dataset of 60 students in five different online courses. The time span of the dataset is 15 weeks, and in this study, we use every three weeks as a time node. The data includes the content of online courses, the amount of homework for online courses, students' class time, the text of student comments on online course platforms, and the frequency of comments. The student's comment text is used for student sentiment analysis.

During the experiment, the number of individuals in the I-NOA algorithm was 20, and the maximum number of iterations was 200. The dropout rate of CNN algorithm is 0.5, and the learning rate is 0.001. Out of the data of 60 students, 15 were used as the training set, while the data labels of the remaining 45 students were used as the testing set. In addition, in order to verify the prediction accuracy and robustness of the I-NOA-CNN algorithm, some recently developed algorithms such as the classic CNN algorithm [21], psoCNN [27], and NOA-CNN were selected for comparison with the I-NOA-CNN algorithm. During the calculation process, each algorithm is independently run 30 times.

A. Student's Emotional Prediction

Fig. 4 and 5, respectively show the prediction results of CNN algorithm and I-NOA-CNN algorithm on students' emotions in weeks 15 of Course 1.

From Fig. 4 and 5, we can see that the I-NOA-CNN algorithm has higher accuracy than the CNN algorithm in predicting students' emotions. Table I shows the accuracy of each algorithm's predictions over 30 runs. The I-NOA-CNN algorithm has the highest prediction accuracy, while the CNN algorithm has the lowest. Although the psoCNN and NOA-CNN algorithms have better prediction accuracy than the CNN algorithm, they are far lower than the I-NOA-CNN algorithm.

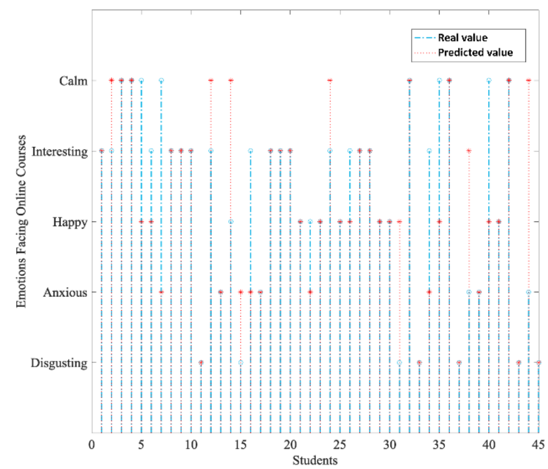

Fig. 4. The results of CNN algorithm for predicting student emotions.

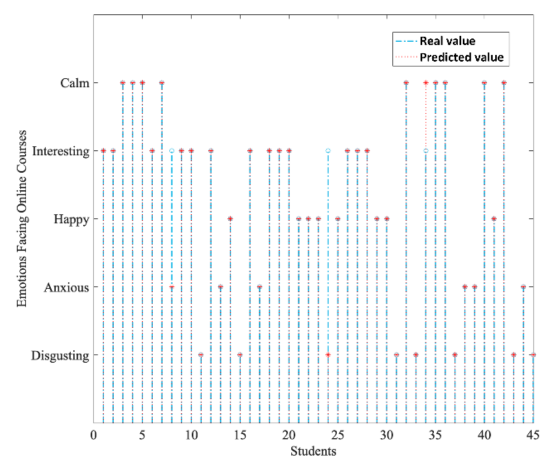

Fig. 5. The results of I-NOA-CNN algorithm for predicting student emotions.

TABLE I. The Accuracy of Each Algorithm's Prediction of Student Emotions Based on the Data from the 15th Week

Algorithm	Course 1	Course 2	Course 3	Course 4	Course 5
CNN	65.57%	71.47%	75.12%	81.83%	74.28%
psoCNN	80.62%	81.16%	85.42%	82.99%	79.19%
NOA-CNN	85.76%	84.57%	92.69%	83.08%	82.17%
I-NOA-CNN	93.57%	96.41%	97. 59%	96.18%	95.07%

B. Student's Final Grade Prediction

Fig. 6 and 7, respectively show the prediction results of CNN algorithm and I-NOA-CNN algorithm on students' final grades at weeks 15 in course 1. In the final grade prediction process, we divide the final grade into five levels: A, B, C, D, and E.

From Fig. 6 and 7, it can be seen that at week 15, the I-NOA-CNN algorithm had higher accuracy in predicting students' final grades in course 1 than the CNN algorithm. Fig. 8 and 9 respectively show the prediction accuracy of four algorithms in different courses and time nodes during multiple runs. From Fig. 8, it can be seen that the I-NOA-CNN algorithm has good prediction accuracy (>90%) in different courses and has strong generalization ability. From Fig. 9, it can be seen that the I-NOA-CNN algorithm also has the best prediction performance at different time nodes. In addition, due to the expansion of the time range, the amount of data increases, and as the time range expands, the prediction accuracy of each algorithm shows an upward trend.

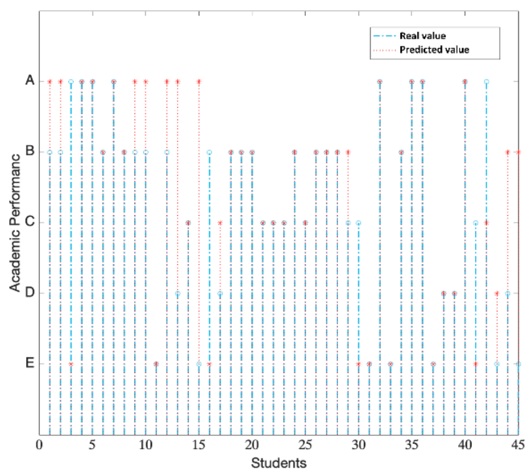

Fig. 6. The prediction results of CNN algorithm on students' final grades.

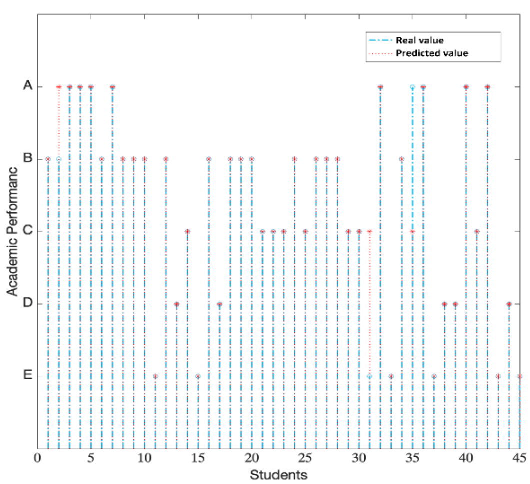

Fig. 7. The prediction results of I-NOA-CNN algorithm on students' final grades.

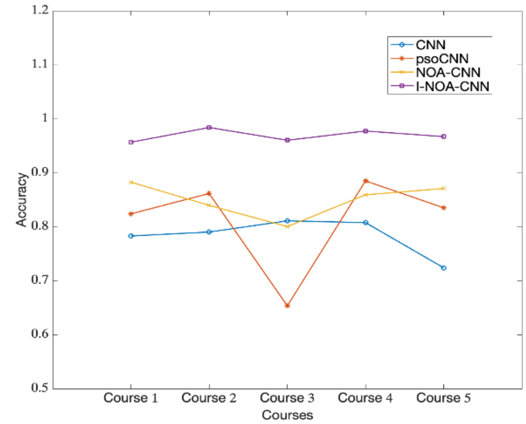

Fig. 8. Different algorithms predict the accuracy of student grades in different course.

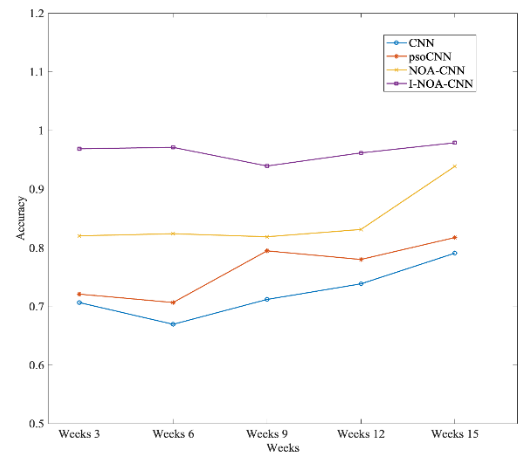

Fig. 9. The accuracy of different algorithms in predicting student grades at different time points.

V. Conclusion

This study focuses on the problem of predicting student grades for online courses and mainly designs an improved CNN algorithm. The grade prediction model mainly consists of two steps. The first step is to predict students' emotions based on an improved CNN algorithm. The second step is to predict students' final grades based on their emotions, course content, and review data. Compared with some of the latest algorithms, this algorithm has higher prediction accuracy and better robustness. The data for this study covers a time span of 15 weeks and does not take into account some short-term online courses. In future research, we will focus on studying the performance prediction of online courses with shorter time spans (less than one week).

References

[1] C. Müller, T. Mildenberger, and D. Steingruber, "Learning effectiveness of a flexible learning study programme in a blended learning design: why are some courses more effective than others?," International Journal of Educational Technology in Higher Education, vol. 20, no. 1, pp. 10, 2023.

[2] T. Soffer and R. Nachmias, "Effectiveness of learning in online academic courses compared with face-to-face courses in higher education," Journal of Computer Assisted Learning, vol. 34, no. 5, pp. 534-543, 2018.

[3] W. Chen, C. G. Brinton, D. Cao, A. Mason-Singh, C. Lu and M. Chiang, "Early Detection Prediction of Learning Outcomes in Online Short-Courses via Learning Behaviors," IEEE Transactions on Learning Technologies, vol. 12, no. 1, pp. 44-58, Jan.-March 2019.

[4] Arul Leena Rose. P. J and Ananthi Claral Mary.T, "An Early Intervention Technique for At-Risk Prediction of Higher Education Students in Cloud-based Virtual Learning Environment using Classification Algorithms during COVID-19" International Journal of Advanced Computer Science and Applications (IJACSA), 13(1), 2022.

[5] E. Martínez-Caro, "Factors affecting effectiveness in e-learning: An analysis in production management courses," Computer Applications in Engineering Education, vol. 19, no. 3, pp. 572–581, 2011.

[6] Yanli Chen and Ke Jin, "Educational Performance Prediction with Random Forest and Innovative Optimizers: A Data Mining Approach" International Journal of Advanced Computer Science and Applications (IJACSA), 15(3), 2024.

[7] Y. M. I. Hassan, A. Elkorany and K. Wassif, "Utilizing Social Clustering-Based Regression Model for Predicting Student's GPA," in IEEE Access, vol. 10, pp. 48948-48963, 2022.

[8] G. Tonguc and B. O. Ozkara, "Automatic recognition of student emotions from facial expressions during a lecture," Computers and Education, vol. 148, pp. 103797, 2020.

[9] K. F. Hew, X. Hu, C. Qiao, and Y. Tang, "What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach," Computers and Education, vol. 145, pp. 103724, 2020.

[10] D. Baneres, M. E. Rodríguez-González and A. E. Guerrero-Roldán, "A Real-Time Predictive Model for Identifying Course Dropout in Online Higher Education," IEEE Transactions on Learning Technologies, vol. 16, no. 4, pp. 484-499, Aug. 2023, doi: 10.1109/TLT.2023.3267275.

[11] A. S. Aljaloud et al., "A Deep Learning Model to Predict Student Learning Outcomes in LMS Using CNN and LSTM," IEEE Access, vol. 10, pp. 85255-85265, 2022.

[12] H. Waheed, S.-U. Hassan, R. Nawaz, N. R. Aljohani, G. Chen, and D. Gasevic, "Early prediction of learners at risk in self-paced education: A neural network approach," Expert Systems With Applications, vol. 213, pp. 118868, 2023.

[13] Mayanda Mega Santoni, T. Basaruddin and Kasiyah Junus, "Convolutional Neural Network Model based Students' Engagement Detection in Imbalanced DAiSEE Dataset" International Journal of Advanced Computer Science and Applications (IJACSA), 14(3), 2023.

[14] R. Alcaraz, A. Martínez-Rodrigo, R. Zangróniz and J. J. Rieta, "Early Prediction of Students at Risk of Failing a Face-to-Face Course in Power Electronic Systems," IEEE Transactions on Learning Technologies, vol. 14, no. 5, pp. 590-603, Oct. 2021.

[15] A. Gupta, D. Garg and P. Kumar, "Mining Sequential Learning Trajectories With Hidden Markov Models For Early Prediction of At-Risk Students in E-Learning Environments," IEEE Transactions on Learning Technologies, vol. 15, no. 6, pp. 783-797, Dec. 2022.

[16] Z. Xu, H. Yuan and Q. Liu, "Student Performance Prediction Based on Blended Learning," in IEEE Transactions on Education, vol. 64, no. 1, pp. 66-73, Feb. 202.

[17] Z. Wu, L. Huang, Q. Huang, C. Huang, and Y. Tang, "SGKT: Session graph-based knowledge tracing for student performance prediction," Expert Systems With Applications, vol. 206, pp. 117681, 2022.

[18] Q. Liu et al., "EKT: Exercise-Aware Knowledge Tracing for Student Performance Prediction," in IEEE Transactions on Knowledge and Data Engineering, vol. 33, no. 1, pp. 100-115, 1 Jan. 2021.

[19] O. Taylan and B. Karagözoğlu, "An adaptive neuro-fuzzy model for prediction of student's academic performance," Computers & Industrial Engineering, vol. 57, no. 3, pp. 732–741, 2009.

[20] J. Nie and H. Ahmadi Dehrashid, "Evaluation of student failure in higher education by an innovative strategy of fuzzy system combined optimization algorithms and AI," Heliyon, vol. 10, no. 7, pp. e29182, 2024.

[21] X. Lu, Y. Zhu, Y. Xu, and J. Yu, "Learning from multiple dynamic graphs of student and course interactions for student grade predictions," Neurocomputing, vol. 431, pp. 23–33, 2021.

[22] M. Lohakan and C. Seetao, "Large-scale experiment in STEM education for high school students using artificial intelligence kit based on computer vision and Python," Heliyon, vol. 10, no. 10, pp. e31366, 2024.

[23] A. Rivas, A. González-Briones, G. Hernández, J. Prieto, and P. Chamoso, "Artificial neural network analysis of the academic performance of students in virtual learning environments," Neurocomputing, vol. 423, pp. 713–720, 2021.

[24] Y. Song, D. Tang, J. Yu, Z. Yu and X. Li, "Short-Term Forecasting Based on Graph Convolution Networks and Multiresolution Convolution Neural Networks for Wind Power," IEEE Transactions on Industrial Informatics, vol. 19, no. 2, pp. 1691-1702, Feb. 2023.

[25] P. S. Naulia, J. Watada and I. A. Aziz, "A Mathematically Inspired Meta-Heuristic Approach to Parameter (Weight) Optimization of Deep Convolution Neural Network," IEEE Access, vol. 12, pp. 83299-83322, 2024.

[26] Z. Zhao and Y. Jiao, "A Fault Diagnosis Method for Rotating Machinery Based on CNN With Mixed Information," IEEE Transactions on Industrial Informatics, vol. 19, no. 8, pp. 9091-9101, Aug. 2023.

[27] J. Tmamna, E. Ben Ayed, R. Fourati, A. Hussain, and M. Ben Ayed, "A CNN pruning approach using constrained binary particle swarm optimization with a reduced search space for image classification," Applied Soft Computing, vol. 164, pp. 111978, 2024.

[28] M. Omar, F. Yakub, S. S. Abdullah, M. S. A. Rahim, A. H. Zuhairi, and N. Govindan, "One-step vs horizon-step training strategies for multi-step traffic flow forecasting with direct particle swarm optimization grid search support vector regression and long short-term memory," Expert Systems With Applications, vol. 252, pp. 124154-, 2024.

[29] J. Li, S. Soradi-Zeid, A. Yousefpour, and D. Pan, "Improved differential evolution algorithm based convolutional neural network for emotional analysis of music data," Applied Soft Computing, vol. 153, pp. 111262, 2024.

[30] Ikpo C V, Akowuah E K, Kponyo J J and Boateng K O, "Ọdigbo Metaheuristic Optimization Algorithm for Computation of Real-Parameters and Engineering Design Optimization" International Journal of Advanced Computer Science and Applications (IJACSA), 14(1), 2023.

[31] M. Abdel-Basset, R. Mohamed, M. Jameel, and M. Abouhawwash, "Nutcracker optimizer: A novel nature-inspired metaheuristic algorithm for global optimization and engineering design problems," Knowledge-based systems, vol. 262, pp. 110248, 2023.

[32] R. Lin and H. Hu, "Multi-Task Momentum Distillation for Multimodal Sentiment Analysis," IEEE Transactions on Affective Computing, vol. 15, no. 2, pp. 549-565, 2024.

[33] H. Liu, K. Li, J. Fan, C. Yan, T. Qin and Q. Zheng, "Social Image–Text Sentiment Classification With Cross-Modal Consistency and Knowledge Distillation," IEEE Transactions on Affective Computing, vol. 14, no. 4, pp. 3332-3344, 2023.

[34] K. A. Alshaikh, O. A. Almatrafi and Y. B. Abushark, "BERT-Based Model for Aspect-Based Sentiment Analysis for Analyzing Arabic Open-Ended Survey Responses: A Case Study," IEEE Access, vol. 12, pp. 2288-2302, 2024.

[35] P. Durga and D. Godavarthi, "Deep-Sentiment: An Effective Deep Sentiment Analysis Using a Decision-Based Recurrent Neural Network (D-RNN)," IEEE Access, vol. 11, pp. 108433-108447, 2023.