Automated Identification Of Student Engagement In Online Learning: A Bagging Ensemble Deep Learning Method

Madhavi R K
Senior Scale Lecturer, Dept of Computer Science and Engineering
Govt. Polytechnic, Chintamani

Abstract- *The COVID-19 pandemic has transformed education, shifting learning from traditional classroom settings to online platforms. While this transition offers benefits, such as freeing learning from time and location limitations, and allowing education to take place anytime, anywhere, it also presents a challenge: identifying student engagement in online environments where interaction is limited. Student engagement, defined as the active participation of learners in the educational process, is a crucial factor in shaping the learning experience. This study addresses this issue by introducing a model that employs bagging (bootstrap aggregating) ensemble learning, applied to 1-dimensional convolutional neural networks (1D CNN), 1-dimensional residual networks (1D ResNet), and hybrid ensemble deep learning models. Using the DAiSEE dataset, our results demonstrate that the bagging ensemble of the 1D CNN model reaches 93.25% accuracy, outperforming the standalone model by 3.25%. The deep learning ensemble bagging method achieves 93.75% accuracy, exceeding the performance of the 1D ResNet model by 3.5%. Furthermore, the hybrid ensemble bagging approach delivers the highest accuracy of 94.25%, showing a 1% improvement over the 1D CNN model and a 0.5% increase over the 1D ResNet model.*

Keywords- Bootstrap aggregating, convolutional neural networks, deep learning, student engagement detection, ensemble models, virtual learning, residual neural networks.

I. INTRODUCTION

The COVID-19 pandemic has generated complex, multifaceted challenges worldwide. It has posed an unprecedented obstacle for the education sector, bringing about unexpected consequences and significant social changes. Educational institutions have been forced to address urgent issues in ensuring the continuity of learning. However, amid these difficulties, the present era of technological advancement offers a promising solution. The swift progress in communication technologies provides an alternative means for sustaining education during the ongoing COVID-19 crisis.[1][2][3]

As a product of digital technology, online learning is shaping both current and future educational frameworks. It breaks the boundaries of physical classrooms and eliminates time restrictions.

In addition to the flexibility that online learning provides, its ability to digitally document materials, including recorded lessons, allows for easy access and review at any time.[4][5]

Interaction, a critical factor for effective learning, involves dynamic exchanges between teachers and students and can occur in both offline and online environments. However, achieving such interaction in online settings is more challenging. One major obstacle is the limited information available to teachers about students' engagement and participation levels. As D'Errico et al. note, engagement is an internal state built from various cues, which may not always be visually evident. Engagement involves behavioural, emotional, and cognitive dimensions, making it vital to develop methods for detecting student engagement to improve the learning process. [6][7]

Deep learning, which represents advancement in neural network methodologies for tackling various issues, incorporates convolutional neural networks (CNNs) specifically designed for classifying image or video data. CNNs have effectively resolved numerous classification challenges by automatically extracting high-level features without requiring explicitly defined attributes. Ensemble learning, a technique that merges multiple algorithms to achieve more accurate predictions, has demonstrated its effectiveness, as noted by Yu et al. The study introduced an ensemble learning approach aimed at predicting a person's emotional expressions, highlighting that this method can enhance the performance of the solution. The purpose of ensemble learning is to create a model that delivers greater accuracy than that achievable by any single algorithm or model. Thus, ensemble deep learning can significantly boost the efficacy of deep learning techniques.[8][9]

II. RELATED WORKS

A. Detection of Student's Engagement

Chen et al. predicted student engagement in collaborative learning through the use of computer vision. Their research presents a multi-modal deep neural network (MDNN) that combines facial expressions and gaze direction as two essential elements for forecasting student engagement in collaborative settings. At the same time, Buono et al. proposed a deep learning technique utilizing LSTM to predict student engagement in online learning, incorporating features such as eye gaze, head position, and facial action units. In another study, Ikram et al. employed a primary dataset consisting of 32 in-person classroom videos to assess student engagement levels using a deep learning approach based on VGG16.[11][12][13]

Nezami et al. developed an engagement model utilizing deep learning, which was trained in two separate stages. Initially, a deep learning model focused on training facial expressions, providing a thorough representation of facial features. The weights from this facial expression training model were then used to initialize another model designed to recognize student engagement during the learning process. The findings of this study include a comparative analysis of the engagement model against deep learning models such as CNN and VGGNet, as well as traditional learning models like HOG + SVM. Remarkably, the engagement model achieved higher accuracy compared to the other three algorithms.Additionally, Pabba and Kumar introduced an intelligent system for monitoring student engagement in classroom settings through facial expression recognition. This system attained training and testing accuracies of 78.70% and 76.90%, respectively.[14][15]

B. Ensemble Learning

To enhance the performance of deep learning classification, one effective approach is to implement ensemble learning. Ensemble models can lead to improved evaluation results in engagement detection. Several studies have demonstrated that incorporating bagging techniques in deep learning, particularly with CNNs, can boost classification performance. For instance, Zhang et al. reported that their proposed deep learning bagging method achieved a diagnostic accuracy of 98.89% for COVID-19. In contrast, the ResNet algorithm, when used without ensemble learning, attained an accuracy of 97.22%. This indicates that deep learning bagging can significantly enhance classification accuracy.Similarly, research by Deng et al. found that employing ensemble bagging strategies yielded the highest accuracy compared to using a single classification model. The proposed ensemble bagging model is not only robust but also enhances the overall accuracy of the model while simultaneously reducing classification errors associated with a standalone model.[16]

III. METHODOLOGY

The proposed bagging ensemble learning model is illustrated in Figure 1. For a comprehensive understanding of how to address the imbalance in the DAiSEE dataset, please refer to our previous work. Each video from the DAiSEE dataset is extracted into 300 frames. The OpenFace library is utilized for feature extraction from each frame, resulting in a numerical vector comprising 709 facial feature values. These features include facial landmark detection, head pose estimation, eye gaze estimation, and facial expressions represented as facial action units (AUs).From the 709 facial features, a careful feature selection process is conducted to identify the key attributes that accurately represent the dataset, thereby enhancing the prediction model's precision. This selection process employs Singular Value Decomposition (SVD). The optimal number of components is determined based on the trade-off in variance produced by SVD. Additionally, a data augmentation phase is implemented to increase both the quantity and diversity of the dataset.

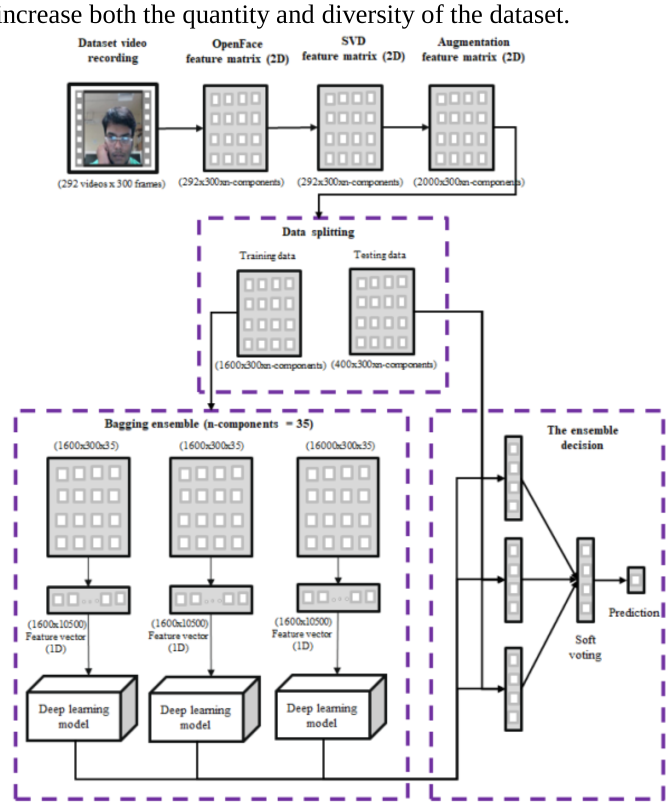

Figure 1: Learning model of Ensemble learn

This study introduces a bagging ensemble learning approach that requires additional resources; therefore, the data are transformed from multidimensional arrays into 1D vectors following the bagging process to improve computational efficiency. Any spatial information that may be lost during this transformation will be recovered during the facial feature extraction phase using the OpenFace library.

A. Dataset

The DAiSEE dataset is a multi-label video classification collection consisting of 9,068 ten-second videos recorded from 112 students, aimed at identifying students' affective states, including boredom, confusion, engagement, and frustration. Boredom is defined as a state of feeling tired or restless due to a lack of interest. Confusion refers to a noticeable lack of understanding, while engagement is characterized by a state of interest resulting from participation in an activity. Frustration is described as a feeling of displeasure or annoyance.Table 1 provides an overview of the data labeling for each video in the DAiSEE dataset. Each video indicates the level of the student's affective state using numerical values: 0 for very low, 1 for low, 2 for high, and 3 for very high. For instance, in Table 1, the video with clipID 1100011002.avi has a boredom level of 0 (very low), an engagement level of 2 (high), a confusion level of 0 (very low), and a frustration level of 0 (very low). This research specifically focuses on the engagement aspect of students' affective states.[10]

ClipID	Boredom	Engagement	Confusion	Frustation
1100011002.avi	0	2	0	0
1100011003.avi	0	2	0	0
1100011004.avi	0	3	0	0
1100011009.avi	0	2	1	0
1100011011.avi	0	3	0	0
1100011012.avi	0	2	2	0

Table 1: Data labelling of daisee dataset

B. Approach of Bagging Ensemble

Bagging, which stands for bootstrap aggregating, is a form of ensemble learning that utilizes multiple models of the same algorithm, all trained on the same dataset. The prediction outcomes from each training model are combined using soft voting techniques. An overview of the bagging process is depicted in Figure 2.

In this study, three bagging training subsets are employed for ensemble learning. Sampling with replacement is utilized to create these three subsets, allowing for instances to be sampled multiple times for the same classifier. The final decision of the three-bagging ensemble learning is determined through the soft voting technique. This technique incorporates two methods: averaging soft voting and maximum soft voting, as illustrated in Figure 3.

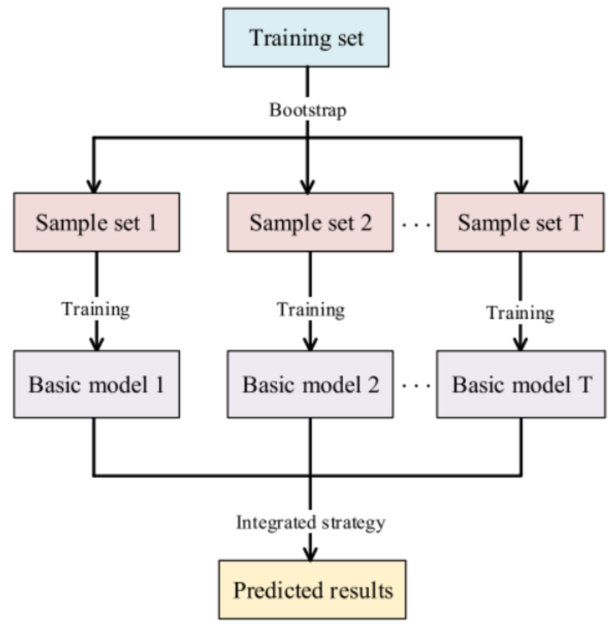

Figure 2: Bagging Ensemble

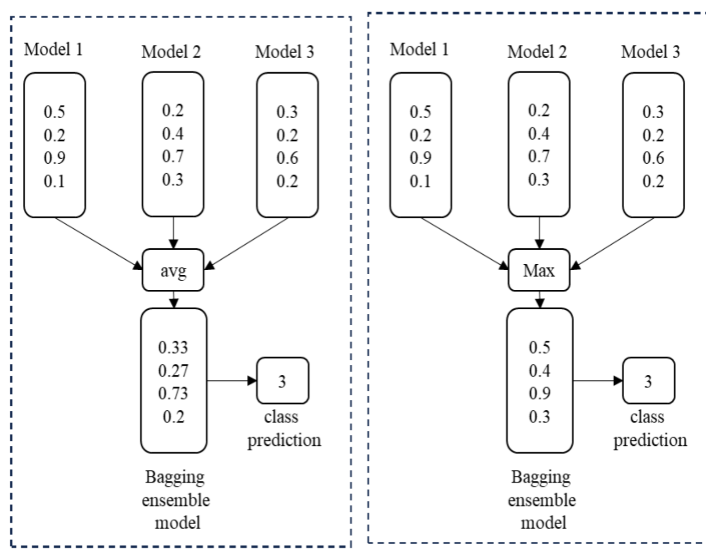

Figure 3: Soft Voting in Ensemble Decision: Left for Averaging Soft Voting, Right for Maximum Soft Voting

Averaging soft voting determines the average probability for each class across all bagging models. As shown in Figure 5, model 1 assigns probabilities of 0.5 to class 1, 0.2 to class 2, 0.9 to class 3, and 0.1 to class 4. If relying solely on model 1, the prediction would be class 3, as it has the highest probability. This approach is similarly applied to models 2 and 3. In ensemble bagging, the probabilities from each model are averaged for every class, resulting in a new probability distribution that is used for prediction. In this case, the ensemble averaging soft voting predicts the data as class 3. In contrast, maximum soft voting makes the ensemble decision by selecting the highest probability from each model. In this example, maximum soft voting also predicts the data as class 3.

IV. CONCLUSION

In summary, the application of ensemble learning, particularly through bagging techniques, significantly enhances the accuracy and reliability of predictions in various contexts, including student engagement detection. By utilizing multiple models and integrating their outputs through methods such as averaging soft voting and maximum soft voting, the ensemble approach effectively captures diverse insights from the data. The ability to transform multidimensional data into 1D vectors without losing critical spatial information further streamlines the computational process, ensuring efficient analysis. The DAiSEE dataset serves as a valuable resource for understanding affective states, with the implemented bagging ensemble learning model demonstrating promising results in predicting student engagement. Ultimately, the advancements in deep learning methodologies and the innovative strategies employed in this study pave the way for more effective educational interventions and enhanced learning experiences for students.

REFERENCES

[1] K. V. Elumalai, J. P. Sankar, J. A. John, N. Menon, M. S. M. Alqahtani, and M. A. Abumelha, ‘‘Factors affecting the quality of e-learning during the COVID-19 pandemic from the perspective of higher education students,’’ J. Inf. Technol. Educ., Res., vol. 19, pp. 731–753, Sep. 2020, doi: 10.28945/4628.

[2] V. J. García-Morales, A. Garrido-Moreno, and R. Martín-Rojas, ‘‘The transformation of higher education after the COVID disruption: Emerging challenges in an online learning scenario,’’ Frontiers Psychol., vol. 12, pp. 1–6, Feb. 2021, doi: 10.3389/fpsyg.2021.616059.

[3] O. B. Adedoyin and E. Soykan, ‘‘COVID-19 pandemic and online learning: The challenges and opportunities,’’ Interact. Learn. Environments, vol. 31, no. 2, pp. 863–875, Feb. 2023, doi: 10.1080/10494820.2020. 1813180.

[4] A. Haleem, M. Javaid, M. A. Qadri, and R. Suman, ‘‘Understanding the role of digital technologies in education: A review,’’ Sustain. Operations Comput., vol. 3, pp. 275–285, Jan. 2022, doi: 10.1016/j.susoc.2022.05.004.

[5] C. Greenhow, C. R. Graham, and M. J. Koehler, ‘‘Foundations of online learning: Challenges and opportunities,’’ Educ. Psychologist, vol. 57, no. 3, pp. 131–147, Jul. 2022, doi: 10.1080/00461520.2022. 2090364.

[6] K. Aldrup, B. Carstensen, and U. Klusmann, ‘‘Is empathy the key to effective teaching? A systematic review of its association with teacher– student interactions and student outcomes,’’ Educ. Psychol. Rev., vol. 34, no. 3, pp. 1177–1216, Sep. 2022, doi: 10.1007/s10648-021-09649-y.

[7] F. D’Errico, M. Paciello, and L. Cerniglia, ‘‘When emotions enhance students’ engagement in e-learning processes,’’ J. E-Learn.Knowl. Soc., vol. 12, no. 4, pp. 9–23, 2016, doi: 10.20368/1971-8829/ 1144.

[8] I. Dubovi, ‘‘Cognitive and emotional engagement while learning with VR: The perspective of multimodal methodology,’’ Comput. Educ., vol. 183, Jul. 2022, Art.no. 104495, doi: 10.1016/j.compedu.2022.

[9] J. Yu, Z. Cai, P. He, G. Xie, and Q. Ling, ‘‘Multi-model ensemble learning method for human expression recognition,’’ 2022, arXiv:2203.14466.

[10] A. Gupta, A. D’Cunha, K. Awasthi, and V. Balasubramanian, ‘‘DAiSEE: Towards user engagement recognition in the wild,’’ 2016, arXiv:1609.01885

[11] Y. Chen, J. Zhou, Q. Gao, J. Gao, and W. Zhang, ‘‘MDNN: Predicting student engagement via gaze direction and facial expression in collaborative learning,’’ Comput. Model. Eng. Sci., vol. 136, no. 1, pp. 381–401, 2023, doi: 10.32604/cmes.2023.023234.

[12] P. Buono, B. De Carolis, F. D’Errico, N. Macchiarulo, and G. Palestra, ‘‘Assessing student engagement from facial behavior in on-line learning,’’ Multimedia Tools Appl., vol. 82, no. 9, pp. 12859–12877, Apr. 2023, doi: 10.1007/s11042-022-14048-8.

[13] S. Ikram, H. Ahmad, N. Mahmood, C. M. N. Faisal, Q. Abbas, I. Qureshi, and A. Hussain, ‘‘Recognition of student engagement state in a classroom environment using deep and efficient transfer learning algorithm,’’ Appl. Sci., vol. 13, no. 15, p. 8637, Jul. 2023, doi: 10.3390/ app13158637.

[14] O. M. Nezami, M. Dras, L. Hamey, D. Richards, S. Wan, and C. Paris, ‘‘Automatic recognition of student engagement using deep learning and facial expression,’’ 2018, arXiv:1808.02324.

[15] C. Pabba and P. Kumar, ‘‘An intelligent system for monitoring students’ engagement in large classroom teaching through facial expression recognition,’’ Exppert Syst., vol. 39, no. 1, pp. 1–28, Jan. 2022, doi: 10.1111/exsy. 12839.

[16] W. Deng, Q. Xu, J. Liu, Y. Lu, M. Fan, and X. Liu, ‘‘Image classification method of longhorn beetles of Yunnan based on bagging and CNN,’’ in Proc. 5th Int. Conf. Pattern Recognit. Artif.Intell. (PRAI), Aug. 2022, pp. 198–203, doi: 10.1109/PRAI55851.2022.9904155.