Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025
Paper ID #48608

A CNN-Driven Hybrid Classification Model to Predict Students' Academic Performance

Dr. Sujan Poudyal, Cleveland State University

Dr. Sujan Poudyal is an Assistant College Lecturer in the Department of Computer Science at Cleveland State University. He earned his Ph.D. in Electrical and Computer Engineering from Mississippi State University in 2022, where his research focused on applying hybrid machine learning techniques in Educational Data Mining. Dr. Poudyal has over eight years of academic experience, including roles as a Lecturer in Nepal and a Graduate Teaching and Research Assistant at Mississippi State University. His expertise includes machine learning, academic integrity detection, and innovative teaching methodologies. Dr. Poudyal actively publishes and presents on machine learning and educational analytics.

A CNN-Driven Hybrid Classification Model to Predict Students' Academic Performance

Abstract - Large amount of data is obtained from online courses, e-learning platforms, and institutional technologies. Educational Data Mining (EDM) leverages these data to make informed decisions that improve the educational experience, student outcomes, and institutional efficiency. As academic achievement is one of the key aspects for assessing quality education, predicting students' academic performance has become a focus for research in this field. Several traditional machine learning models like Naïve Bayes, K-Nearest Neighbors (KNN), Logistic Regression and a few more are most widely used in EDM field to predict students' academic performance. Though CNN are widely used in several other domains, they have not been extensively studied for educational data mining. In this study, we proposed a novel hybrid model that combines the CNNs' feature extraction strength with the traditional classification model. We first converted our 1D numerical student data into 2D image representations. This allows 2D-CNN applications to extract important features from the data. The extracted features are then fed into traditional classification models, including Naïve Bayes, K-Nearest Neighbors, and Logistic Regression. The performance of the hybrid model is evaluated in a pass/fail classification scenario. Experimental results show that our proposed CNN-based hybrid classification model outperforms the standalone traditional model in terms of classification accuracy. This study introduces an innovative approach in the educational domain, demonstrating that CNNs can provide a more robust and reliable method for predicting student performance, especially when predicting binary results like pass or fail.

1. Introduction

Educational institutions view their students as valuable assets and are committed to fostering their academic success. Academic performance is a key metric in determining students' future opportunities, including access to higher education, career advancement, and improved quality of life. A longitudinal study highlighted that students' high school GPAs strongly correlate with their college success and earning potential later in life [1]. Consequently, accurately predicting students' academic performance is critical for implementing early interventions and ensuring better outcomes.

Data mining, a well-established method for extracting meaningful patterns from large datasets, has gained traction in various domains such as healthcare [2], fraud detection [3], bioinformatics [4], and education [5-8]. Within the educational context, the rapid adoption of technology—through online courses, e-learning platforms, and institutional systems—has resulted in a massive amount of data. This raw data often contains hidden insights that can be uncovered using Educational Data Mining (EDM). EDM specializes in analyzing educational data to reveal actionable insights, such as identifying at-risk students, improving course content, and predicting academic performance [9].

Traditionally, machine learning algorithms like Naïve Bayes, K-Nearest Neighbors (KNN), Logistic Regression, and Decision Trees have been employed to predict students' academic performance [10-12]. While these models are effective, their reliance on handcrafted features can limit their adaptability to complex datasets. Deep learning approaches, particularly

Convolutional Neural Networks (CNNs), have demonstrated superior capabilities in automatically extracting features from raw data, making them highly effective for tasks such as image classification and object detection [13,14]. However, CNN applications in the EDM domain remain underutilized, as educational datasets are primarily one-dimensional numerical data rather than two-dimensional image-like data.

In this study, we address this limitation by transforming one-dimensional numerical student data into two-dimensional image representations, enabling the application of CNN architectures. We propose a novel hybrid classification model that combines the strengths of CNN-based feature extraction with traditional classifiers, such as Naïve Bayes, KNN, and Logistic Regression. This hybrid approach leverages CNNs' ability to extract essential features while utilizing traditional models' effectiveness for classification tasks. Our research focuses on predicting students' academic performance in a binary pass/fail classification scenario.

The experimental results demonstrate that the CNN-based hybrid classification model significantly outperforms standalone traditional models, achieving higher classification accuracy. By introducing a novel application of CNNs in EDM, this study contributes to advancing the predictive capabilities in the educational sector and addresses the following research question:

Does combining CNN-based feature extraction with traditional classification models lead to better prediction accuracy for students' academic performance compared to standalone traditional models?

This research underscores the potential of combining deep learning and traditional machine learning techniques to enhance the prediction of academic performance, paving the way for more effective and data-driven interventions in education.

2. Data and Experimentation

We utilized the Open University Learning Analytics Dataset (OULAD) [15] for our research. This dataset contains data on 32,593 students enrolled in 22 courses at the Open University during 2013 and 2014. Figure 1 illustrates the database schema of the OULAD dataset.

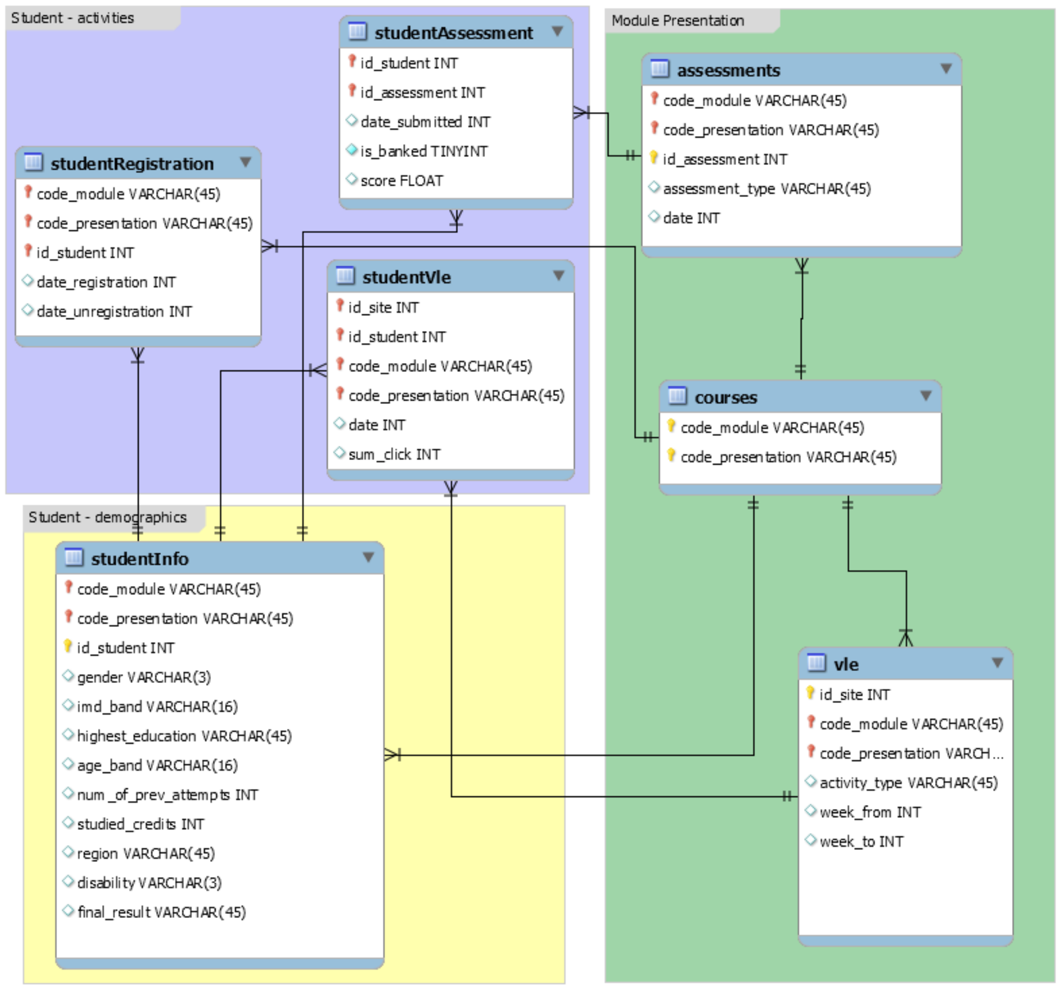

Figure 1. Database schema of OULAD Dataset [15].

Figure 1 highlights that the OULAD dataset comprises seven distinct files. The course file includes details about the modules (or courses) offered in 2013 and 2014. The assessments file provides information about the various assessments associated with each module, while the VLE file outlines the resources available in the virtual learning environment (VLE). The studentInfo file contains demographic information about the students. The studentRegistration file records data on when students registered for and withdrew from their courses. Assessment results are stored in the studentAssessment file, and the studentVle file tracks students' engagement with the materials available in the VLE. Each of these files required preprocessing to compile a dataset suitable for our analysis.

2.1. Data Preprocessing

As outlined on section 2, the dataset includes seven files, each requiring individual processing. We utilized Python to process these files and consolidate them into a single CSV file. This final dataset incorporated student demographic information, daily interactions with the university's VLE, assessment outcomes, and the students' final results. The final student results were initially categorized as *Distinction* (N=3,024), *Pass* (N=12,361), *Fail* (N=7,052), and *Withdraw* (N=10,156). For binary classification purposes, we merged *Distinction* and *Pass* into a single category labeled as *Pass* (N=15,385), and combined *Fail* and *Withdraw* into a single category labeled as *Fail* (N=17,208). Additionally, categorical variables in the dataset were converted to numerical form using one-hot encoding, ensuring all variables were numerical for further analysis. Following the encoding process, we finalized a dataset consisting of 32,593 students. Since our study involves the use of a 2D-CNN, the data, which was initially in a 1D format, it was reshaped into a 2D structure compatible with the 2D-CNN architecture [16].

2.2. Approach

Unlike image data, which is naturally spatially related, the organized tabular data that makes up the OULAD dataset does not. Still, prior studies have demonstrated that 2D CNNs can handle tabular data efficiently when converted to a 2D format [16]. In our method, we converted the feature vectors of each student into an 8x5 matrix so that CNN filters could extract complex feature interactions. Patterns that might not be well defined in a 1D format can be detected by the CNN because of this transformation.

Despite the fact that a 1D CNN is often used for sequential data, our dataset lacks a naturally occurring sequential relationship between features. As an alternative, 2D CNNs use spatial filters to allow hierarchical pattern extraction. Our results empirically validate the usage of a 2D CNN for feature extraction by showing that the hybrid 2D CNN-traditional model performs better than standalone traditional classifiers. Below is a detailed explanation of our CNN model and the feature extraction process.

2.2.1. CNN Model Description:

The input to the CNN is a 2D array with a shape of W×H×1, where W and H represent the width and height of the input image. It is passed to the first convolutional layer that uses 32 filters of size 3×3 with a stride of 1. The convolution operation extract features from the input image and produces a feature map. The size of the output feature map is calculated by using equations (1-3) [17].

$$W_c = \frac{W - F}{S} + 1 \tag{1}$$

$$H_c = \frac{H - F}{S} + 1 \tag{2}$$

$$D_c = K \tag{3}$$

where,

S is a stride step (if s=1, we move the filter one pixel at a time on the image), K is number of filters used, F is filter size, W and H are width and height of input image.

The output of the first convolutional layer is passed to a max pooling layer, which reduces the spatial dimensions while retaining essential features. The pooling window is of size 2×2 with a stride of 1. The dimensions of the output feature map are computed using equations (4-6).

$$W_p = \frac{W - F}{S} + 1 \tag{4}$$

$$H_p = \frac{H - F}{S} + 1 \tag{5}$$

$$W_p = K \tag{6}$$

where,

S is stride step, F is filter size, and W, H are parameters of the input image.

The next layer is another convolutional layer, which consists of 32 filters of size 1×1. This layer further extracts refined features from the pooled feature map, preserving the spatial dimensions while reducing the depth. After the convolutional layers, the feature map is passed through a flatten layer. This layer converts the 2D feature maps into a 1D feature vector, preparing the data for the fully connected layers. The first fully connected (dense) layer processes the feature vector and outputs a representation of size 8. The next dense layer, referred to as the "feature layer," produces a feature vector of size 6. Finally, the last dense layer, with a softmax activation function, outputs the class probabilities for the given input. We used activation function in our model for the non-linearity transformation for the input signal. We used Rectified Linear Unit (ReLU) for all the convolutional layer and the first dense layer in our model. Mathematically, ReLU is defined as y=max(0,x). It converts all the negative values to zero in the feature map. For the last dense layer, we used sigmoid activation function to add the non-linearity to the input signal.

2.2.2. Feature Extraction and Classification

Learning the spatial relationships between transformed features is a necessary step in the 2D CNN feature extraction process. The input 2D representation in our case is structured as an 8x5 image, with an element of the original 1D feature vector in each row. CNN filters can learn significant interactions within and across rows due to this transformation, which ensures that features stay grouped logically even though it doesn't maintain the strict spatial correlations found in traditional image data.

By using convolutional and pooling layers, the CNN finds important feature dependencies that help predict student academic performance and recovers localized patterns from the 2D representation. Max-pooling reduces noise while maintaining the most important information by further refining these derived features. The final classification is then determined by passing these extracted feature maps to traditional classification algorithm such as Naïve Bayes, KNN, and Logistic Regression.

Prior study has demonstrated that CNNs are effective for structured tabular data, despite their primary use for image data [16]. CNNs improve predictive performance by capturing abstract correlations between features through the use of hierarchical feature extraction. Our hybrid model effectively combines CNN feature extraction with the simplicity and interpretability of traditional classifiers. The performance of the hybrid model was evaluated based on classification accuracy using Naïve Bayes, KNN, and Logistic Regression. Our results demonstrate that this approach enhances prediction accuracy, confirming the effectiveness of CNN-based feature extraction for academic performance prediction.

2.3. Experimentation and Evaluation

We split our dataset to seventy percentage as training test and the remaining thirty percentage as the test set. We used the accuracy as our performance metrics. We did the parameter tuning to obtain the best results.

We used different optimizers to see the effect on classification accuracy by the traditional model. We used learning rate of 0.001 and the epochs as 100 for all the experiments to see the effect of different optimizers on test accuracy. We used Stochastic Gradient Descent (SGD), Ada delta, Root Mean Square Propagation (RMSprop)), and adaptive moment (Adam) as the optimizers. The accuracy of different classifiers is shown in Table 1.

Table 1. Classification Accuracy using different optimizers

Hybrid CNN using / Optimizers	KNN	NB	LR
SGD	75.6	74.6	78.8
RMSprop	86.6	87.1	88.6
Ada delta	75.3	70.1	69.3
Adam	86.7	88.7	89.1

Table 1 shows the test accuracy for using different optimizers. It is seen that when we used the SGD optimizer, the model performed the worst with test accuracy of 75.6, 74.6, and 78.8 using traditional classification algorithm, KNN, NB, and LR respectively. The Adam performed the best with test accuracy of 86.7, 88.7, and 89.1 using traditional classification algorithm, KNN, NB, and LR respectively.

2.4. Evaluation with Baseline Model

The baseline models were also used to predict the academic performance of the students. These models followed the traditional approach without utilizing CNN for feature extraction. The performance of our hybrid CNN model was compared to these baseline traditional models based on classification accuracy. The results are shown in Figure 2.

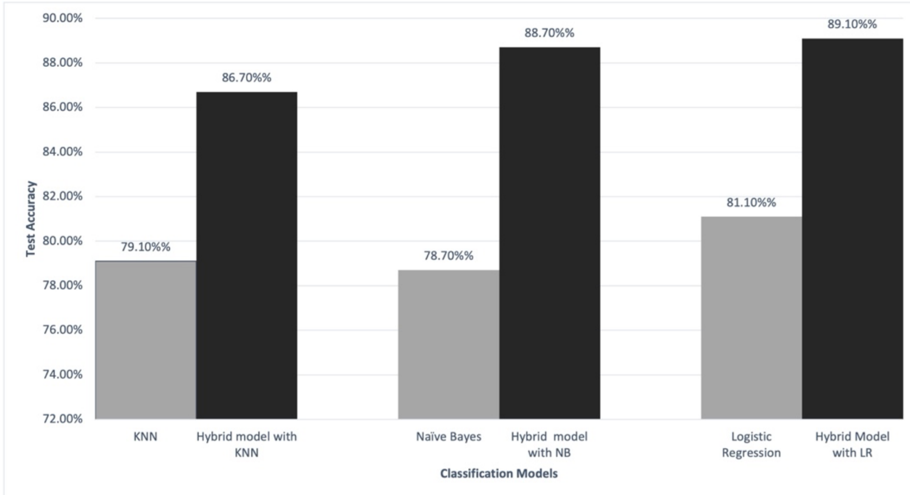

Figure 2. Comparison of Hybrid CNN model with baseline models

Figure 2 shows that the hybrid model outperformed the traditional baseline models in terms of classification accuracy. The percentage improvement in the classification accuracy is highlighted in the Table 2.

Table 2. Model comparison with the accuracy improvement

Model	Standalone Accuracy	Hybrid CNN Accuracy	Accuracy Improvement (%)
KNN	79.10	86.70	09.60
Naïve Bayes	78.70	88.70	12.70
Logistic Regression	81.10	89.10	09.90

Table 2 shows that an accuracy improvement of 12.70% was achieved when using our hybrid CNN model with Naïve Bayes. In contrast, the classification accuracy improved the least, by 9.6%, when using the KNN model.

3. Implication of Results

Educational data mining has traditionally relied on various methods, ranging from traditional machine learning techniques to deep learning approaches, to predict students' academic performance. However, limited research has explored the integration of CNN-based feature extraction with traditional classifiers, particularly in the educational domain. In this study, we introduced a novel hybrid 2D-CNN model that combines CNN's feature extraction capability with traditional classification algorithms. Our results demonstrate that the hybrid model significantly outperformed baseline traditional models in terms of accuracy, as presented in the section 2.3. We conducted parameter tuning to evaluate the model's performance under different configurations. We tested various optimizers to analyze their impact on classification accuracy.

Among the optimizers studied—SGD, RMSprop, AdaDelta, and Adam—our hybrid model achieved the best results with Adam, as highlighted in Table 1. The findings suggest that Adam optimizes the learning process effectively for our hybrid model.

Since our input data, transformed from 1D numerical to 2D image representation, is relatively small in size (8x5), we limited the architecture to a single pooling layer per CNN model. Despite this, our approach achieved an accuracy as high as 89%, outperforming standalone traditional models. This demonstrates that CNNs, typically used for image data, can be effectively applied to 1D numerical educational datasets by transforming them into 2D representations. These findings provide compelling evidence that hybrid CNN models can enhance the accuracy of predicting students' academic performance. This innovative approach addresses the gap in leveraging deep learning for educational data mining and opens new avenues for utilizing CNNs in domains beyond their conventional applications.

4. Conclusion

The rapid growth of online learning platforms, institutional technologies, and e-learning resources has generated vast amounts of data. This data provides educators with an opportunity to analyze and better understand students' learning behaviors. While raw data from the educational domain can reveal basic outcomes, such as pass or fail rates, EDM techniques extract deeper insights and help predict these outcomes with greater accuracy.

In this study, we developed a hybrid 2D CNN-based model to predict whether students would pass or fail. Our approach involved converting numerical 1D student data into 2D to leverage the feature extraction capabilities of CNNs. The hybrid model achieved a high accuracy of 89%, surpassing traditional baseline models, as demonstrated in Table 2. The results validate our research question by confirming the effectiveness of CNNs in this context. To optimize performance, we conducted extensive parameter tuning, including experiments with various batch sizes, optimizers, and epochs. The results, illustrated in Table 1, highlight that the best performance was achieved with Adam optimizers for 100 epochs. The comparison in Figure 2 further confirms that our hybrid model significantly outperformed standalone traditional models in terms of classification accuracy.

Even though results show how effective our approach is, there are significant ethical concerns raised by using AI to predict students' academic performance, especially with regard to bias and fairness. The OULAD dataset might have inherent biases related to demographics, socioeconomic status, or institutional regulations because it is based on real student records. Machine learning models run the risk of sustaining current educational disparities if these biases are not addressed properly. A major concern is that the training data may overrepresent or underrepresent particular student groups, which could result in predictions that are more accurate for some groups than others. As a result, interventions based on model recommendations may differ. A smaller number of students from disadvantaged backgrounds in the dataset, for instance, could make it more difficult for the model to predict their academic success, which could result in unfair treatment. While bias mitigation strategies aren't specifically addressed in our current study, future work could investigate strategies like explainability techniques and fairness-aware machine learning algorithms to ensure that AI-driven educational decisions are transparent and equal. To promote equitable and welcoming learning environments, it is essential to ensure ethical AI practices in educational data mining.

Our research showcases the potential of applying CNNs to numerical educational datasets, providing a robust method for predicting academic performance. Given the small size of our transformed 2D data, future work could explore larger image-based datasets from the educational domain to fully understand the potential of deep learning in EDM. Furthermore, a direct comparison of 1D CNN and 2D CNN techniques would shed more light on how well-suited each is for educational data mining. Additionally, using explainable AI techniques like SHAP (SHapley Additive Explanations) could improve model transparency, and fairness-aware algorithms should be explored to reduce any potential biases.

The findings suggest that educational institutions can utilize CNN-based models to predict student outcomes and implement targeted interventions to improve academic success. The promising results of this study motivate further exploration of deep learning techniques in the pedagogical field, aiming to enhance the predictive accuracy and broaden the scope of EDM applications.

References

[1] J. Burrus, A. Betancourt, S. Holtzman, *et al.*, "The predictive validity of high school GPA for postsecondary outcomes: A meta-analysis," *Journal of Educational Psychology*, vol. 115, no. 4, pp. 672-690, 2023.

[2] A. Esteva, A. Robicquet, B. Ramsundar, *et al.*, "Deep learning in healthcare: Current applications and emerging trends," *Nature Medicine*, vol. 29, no. 3, pp. 283-298, 2023.

[3] P. Kaur, C. Chauhan, and K. Singh, "Fraud detection in financial systems: Machine learning and beyond," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 53, no. 2, pp. 1-11, 2023.

[4] K. Lee, H. Yang, and T. Liu, "Exploring bioinformatics datasets with deep learning: A comprehensive review," *Bioinformatics Reviews*, vol. 39, no. 5, article btaa931, 2023.

[5] S. Poudyal, M. Nagahi, M. Nagahisarchoghaei, and G. Ghanbari, "Machine learning techniques for determining students' academic performance: A sustainable development case for engineering education," in *Proc. 2020 Int. Conf. Decision Aid Sciences and Application (DASA)*, Sakheer, Bahrain, 2020, pp. 920-924, doi: 10.1109/DASA51403.2020.9317178.

[6] S. Poudyal, M. J. Mohammadi-Aragh, and J. E. Ball, "Data mining approach for determining student attention pattern," in *Proc. 2020 IEEE Frontiers in Education Conf. (FIE)*, Uppsala, Sweden, 2020, pp. 1-8, doi: 10.1109/FIE44824.2020.9274061.

[7] M. Nagahi, R. Jaradat, M. Nagahisarchoghaei, G. Ghanbari, S. Poudyal, and S. Goerger, "Effect of individual differences in predicting engineering students' performance: A case of education for sustainable development," in *Proc. 2020 Int. Conf. Decision Aid Sciences and Application (DASA)*, Sakheer, Bahrain, 2020, pp. 925-931, doi: 10.1109/DASA51403.2020.9317157.

[8] S. Poudyal, M. J. Mohammadi-Aragh, and J. E. Ball, "Hybrid feature extraction model to categorize student attention pattern and its relationship with learning," *Electronics*, vol. 11, no. 9, article 1476, 2022, doi: 10.3390/electronics11091476.

[9] M. Pechenizkiy, *et al.*, "Educational data mining: Recent advances and future directions," *Int. J. Data Mining in Education*, vol. 9, no. 1, pp. 15-28, 2023.

[10] S. Hussain, O. Daramola, and R. Rauf, "Machine learning models for predicting academic performance: A review," *Computers & Education*, vol. 180, article 104430, 2022.

[11] S. Helal, M. Habib, and Z. Hussain, "Identifying vulnerable students using demographic and academic data: A machine learning perspective," *Applied Soft Computing*, vol. 131, article 109744, 2022.

[12] E. Fernandes, M. Rocha, and C. Silva, "Predicting student success through gradient boosting techniques: Insights from a multi-institutional dataset," *Decision Support Systems*, vol. 163, article 113775, 2023.

[13] W. Rawat and Z. Wang, "Deep convolutional neural networks for image recognition: An updated review," *Neural Computing and Applications*, vol. 35, no. 7, pp. 543-562, 2023.

[14] M. Isra, *et al.*, "Adapting CNN architectures for tabular data: Opportunities and challenges," *IEEE Access*, vol. 11, pp. 45988-46001, 2023.

[15] J. Kuzilek, M. Hlosta, and Z. Zdrahal, "Open University Learning Analytics dataset," *Sci. Data*, vol. 4, article 170171, 2017.

[16] S. Poudyal, M. J. Mohammadi-Aragh, and J. E. Ball, "Prediction of student academic performance using a hybrid 2D CNN model," *Electronics*, vol. 11, no. 7, article 1005, 2022, doi: 10.3390/electronics11071005.

[17] "Understand the architecture of CNN," [Online]. Available: https://towardsdatascience.com/understand-the-architecture-of-cnn-90a25e244c7. [Accessed: Jan. 4, 2024].