Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 31 No. 6s (2024)

A Multi-View Deep Learning Approach for Enhanced Student Academic Performance Prediction

Mrs.V.Bakyalakshmi[1], Dr.S.Kanchana[2]

[1]Ph.D Scholar, Department of Computer Science, PSG College of Arts & Science, vbakyalakshmi@yahoo.co.in.

[2]Associate Professor, Department of Software Systems, PSG College of Arts & Science, kanchana@psgcas.ac.in.

Article History:

Received: *06-06-2024*

Revised: *08-07-2024*

Accepted: *28-07-2024*

Abstract:

Educational institutions are utilizing Deep Learning (DL) techniques to develop predictive systems that identify students at risk of underperforming based on historical academic data patterns, thereby enhancing their educational outcomes through targeted interventions. From this outlook, an Ensemble Generative Adversarial Network with a Student Accomplishment prediction using the Distinctive DL (EGAN-SADDL) model was designed to generate large-scale student data and predict their academic achievements. However, integrating heterogeneous kinds of student data into the SADDL model is a complex task that, if not executed properly, may result in the model failing to capture crucial data relationships, leading to lower performance. Hence, this paper proposes an EGAN with Improved SADDL (EGAN-ISADDL) model based on multi-view learning for predicting student academic performance. The main aim of this model is to learn features from multiple sources, including academic records, demographic information, and social media activity, using the multi-view learning approach. First, the academic and demographic attributes of students are collected, along with the physiological features extracted from the information posted on social media by students. Second, the Long Short-Term Memory with Deep Convolutional Neural Network (LSTM-DCNN) and Recursive Neural Network (ReNN) models receive these features in parallel, extracting intermediate features in multiple views. Third, a multi-view classifier jointly learns each set of features to predict students' academic performance, enabling early identification of at-risk students with high accuracy. Finally, experiments conducted on a dataset of 50,000 student records demonstrate that EGAN-ISADDL attains 96.28% accuracy compared to the existing single-view learning models.

Keywords: Academic performance prediction, EGAN-SADDL, Heterogeneous data, Multi-view learning, Recursive neural network.

1. Introduction

Evaluating learning achievement is a pivotal indicator of academic accomplishment, as poor grades can result in stress, despair, and detrimental conduct. Students may encounter challenges such as mental health disorders, familial or societal difficulties, or an absence of direction, all of which might result in their absence from classes [1]. As a result, their academic progress is in jeopardy, and educational institutions must promptly discover and intervene. Educational Data Mining (EDM) is a significant area of research that focuses on forecasting students' academic performance and mitigating undesirable effects such as poor grades or student attrition [2]. Early identification of learners with lower grades enables investigation to enhance their academic performance and assure their success. This is particularly well-suited for high school and college students [3]. High school kids, specifically, are an optimal demographic to focus on due to the significant influence their academic achievements have on their future educational opportunities. Through the examination of a dataset containing socioeconomic and educational data of high school students, it is feasible to discern pupils who exhibit subpar academic performance [4]. Developing precise criteria can aid in identifying students who are likely to fail in the final semester. Assessing all students' academic performance is difficult for teachers and educational institutions due to the large number of students and limited resources.

Many experts in the field have suggested different Machine Learning (ML) algorithms as a potential solution to this issue [5]. Various domains have extensively utilized these algorithms, such as identifying at-risk students, predicting final exam scores, forecasting placement rates, and detecting early failures [6]. As a result, the research community has focused a lot of effort on identifying and effectively addressing students who are at risk [7]. However, the accuracy of early student risk prediction heavily depends on the dataset's characteristics, which are meticulously annotated with expert knowledge. These features are complex and varied, which presents a significant difficulty in extracting attributes of high quality while managing several data formats and a substantial amount of data. To effectively predict learners' achievements, it is essential to investigate their mental health traits in addition to commonly studied aspects such as personal, academic, and economic issues.

As a result, the SADDL algorithm has been developed [8], which utilizes a range of student variables, encompassing physiological, demographic, and academic factors. Information is gathered from the API services of Twitter, Facebook, and Instagram and then subjected to pre-processing utilizing Natural Language Processing (NLP) techniques. The Linguistic Inquiry and Word Count (LIWC) and Latent Dirichlet Allocation (LDA) methods are employed to extract features about mental health and changes in attitudes. The LSTM, DCNN, and Multi-Layer Perceptron (MLP) algorithms combine attributes with common properties to provide the DDL. The LSTM model captures temporal patterns in the time-series data and converts them into a tensor of features. This tensor is then fed into the multidimensional DCNN to analyze the correlations between these properties. The temporal, correlation, and student demographic information are combined to form a cohesive attribute vector. Furthermore, the MLP classifier accurately forecasts students' academic achievement. However, the accuracy of predicting student achievement is still limited because there is not enough data available on students. Collecting further data proved to be a laborious and monotonous undertaking.

To combat this challenge, the EGAN-SADDL model [9] has been developed to forecast student performance. The EGAN comprises two types of GANs: Divergence GAN (DivGAN) minimizes the gap between latent elements and observed data, while Success-aware GAN (Suc-GAN) decreases the imbalance between raw and synthetic data using Gumbel-Softmax Relaxation (GSR). Tuna Swarm Optimization (TSO) stabilizes learning of the EGAN, which balances data quality and diversity. The generated student data is input into the SADDL model for predicting student achievement. Conversely, the diverse nature of input data, stemming from various sources like learning management systems, surveys, and administrative records, can significantly impact the SADDL classifier model's performance. This heterogeneous data includes numerical data (e.g., grades, attendance records), categorical data (e.g., demographics, course types), and text data (e.g., feedback comments, survey responses). Integrating these different data types or sources into a unified model is a complex task that, if not executed properly, may result in the model failing to capture crucial data relationships, leading to lower performance.

Therefore, this manuscript proposes the EGAN-ISADDL model based on the multi-view learning approach for student performance prediction. SADDL can automatically extract features from student data, making it possible to use these extracted features as appropriate inputs. The intermediate representations obtained from these networks can serve as valuable features. SADDL extracts temporal and correlation features from student data, while multi-view learning focuses on learning features from multiple perspectives. Multi-view learning leverages the diverse properties of input data, learning features from each view and combining them to improve prediction accuracy. The proposed ISADDL combines deep features extracted from student data by heterogeneous deep neural networks and classifies them using a multi-view classifier. This approach utilizes intermediate features extracted from LSTM and ReNN for classification tasks. The main contribution of this manuscript can be summarized as follows:

1. Initially, the academic and demographic attributes of students are collected, along with information from Twitter, Facebook, and Instagram, using API services to extract physiological features through LDA approach.
2. The LSTM-DCNN and ReNN models are used to generate intermediate features from the input attributes simultaneously. These features are temporal and correlational, treated as distinct views.
3. A multi-view classifier instead of MLP processes each set of features as a separate view and trains them jointly to predict students' academic performance, enabling early identification of at-risk students with high accuracy.
4. Extensive experiments demonstrate that the EGAN-ISADDL model outperforms traditional models in terms of outcomes.

The remaining manuscript is structured in the following manner: Section 2 discusses earlier studies. Section 3 explores the EGAN-ISADDL model, while Section 4 evaluates its efficiency. Section 5 summarizes the findings along with potential enhancements.

2. Literature Survey

Forecasting students' academic achievement is a critical issue in higher education. To tackle this problem, many academics have utilized educational data mining technology to make informed decisions and predict student success. This section provides a brief summary of the current literature on the use of DL models to predict student performance and identify high-risk students at an early stage.

Arashpour et al. [10] combined Support Vector Machine (SVM) and Artificial Neural Network (ANN) using Teaching-Learning-Based Optimization (TLBO) to predict student performance. They used TLBO to select features for both algorithms and determine the optimal combination of input variables. The ANN architecture was also determined using TLBO. Different hybrid models were established using a dataset with anonymized information on discrete and continuous variables for learning analytics. However, the inadequate amount of available data limited the accuracy.

Al-Azazi et al. [11] developed a day-wise multi-class framework using ANN-LSTM to predict student performance based on demographic and clickstream data. They used the LSTM network instead of the ANN's input layer for time-series analysis to capture latent dependencies among features. However, the small number of records limited the model's performance metrics, such as recall and F1-score.

Albreiki et al. [12] developed a hybrid model using graph convolution networks and ML to predict academic performance, focusing on at-risk students. They utilized Euclidean and cosine distances to establish similarities between student data and construct a graph. Topological features were extracted from the graph to capture structural correlations, which were then learned by ML algorithms for prediction. However, it needs additional factors related to student behavior on online platforms to improve prediction accuracy.

Liu et al. [13] developed the MCAG framework to forecast student achievement. This framework utilized the Maximum Information Coefficient (MIC), DCNN, attention strategy, and Gated Recurrent Unit (GRU). The MIC (Multiple Indicator link) was employed to ascertain the link between influential factors and student accomplishment. Subsequently, the DCNN-attention model was employed to extract high-dimensional features and allocate attention weights to distinct aspects. The GRU was employed to capture temporal features. Ultimately, the characteristics extracted from both the DCNN and the GRU were merged to forecast student performance. However, the limited availability of student records and other relevant factors compromised accuracy.

Sun et al. [14] developed a novel paradigm for forecasting college scholar achievement based on multi-attribute merging and attention mechanisms. They focused on examining previous academic rankings among college scholars across multiple scopes to uncover correlation qualities between individuals and courses. They employed an attention method to identify correlations between multidimensional aspects. However, due to a lack of available student data, the framework's accuracy was low, prompting a proposal to address student behavioral variables to increase prediction accuracy.

Albahli [15] used random forest and decision tree algorithms with Synthetic Minority Over-sampling Technique (SMOTE) to handle imbalanced data, as well as a Bayesian optimizer-based hyperparameter tuning to improve student academic prediction performance. However, the small number of student records limited the accuracy. Vives et al. [16] employed an LSTM network to forecast students' academic performance in the programming fundamentals course, pinpointing those at risk of failing. They employed GAN and SMOTE to balance the data. However, they found that incorporating additional attributes related to competencies and learning styles could enhance prediction accuracy.

Noviandy et al. [17] developed a stacked classifier using Light Gradient Boosting Machine (LightGBM), random forest, and logistic regression to predict dropout risks and academic success. Only considering static academic and socio-economic features at enrollment limited the study's accuracy, potentially missing dynamic changes during the student's educational path. Baniata et al. [18] developed the DL model using the GRU network to predict students' academic performance. However, a small dataset limited the model efficiency, necessitating further exploration to incorporate diverse data for improved prediction accuracy.

Fazil et al. [19] developed a new Attention-aware Stacked convolutional Stacked BiLSTM (ASIST) network for student representation learning to predict their performance. ASIST jointly learns about student representation based on sequential behavior vectors such as student demographic data, previous academic performance, continuous assessment results, and student interaction with the virtual learning environment. It processes sequential behavior vectors using stacked BiLSTM networks and uses the DCNN model to capture diurnal weekly interaction behavior. Additionally, it employs an attention mechanism to assign weights to features based on their importance. The encoded feature vectors are then concatenated with assessment information, and a softmax layer predicts the performance categories of students. However, the lack of demographic and contextual factors limited the model's accuracy.

Huang and Zeng [20] developed a framework using Dual-Graph Neural Network (DGNN) to predict academic performance. They utilized an interaction-based graph neural network to learn local academic performance representations from online interaction activities and an attribute-based graph neural network to learn global academic performance representations from student attributes. They combined these representations to generate predicted academic performances. However, the model relies on the availability and quality of online learning logs, which may vary across courses and platforms. While it can handle dynamic networks, it may not fully capture the complexity of student interactions and attributes.

2.1 Research Gap

The literature review highlights the limitations of single-view learning methods in predicting student academic performance. Current models rely on a single perspective, which can lead to reduced accuracy when dealing with diverse data types like academic records, demographic information, and physiological characteristics on social media sites. Single-view techniques struggle to incorporate temporal and correlational features when analyzing data from multiple perspectives simultaneously. This research gap emphasizes the need for multi-view learning strategies that integrate various data types, extract complementary features from different viewpoints, capture intricate relationships across multiple data sources, and enhance prediction accuracy.

3. Proposed Methodology

This section presents the EGAN-ISADDL model to predict learner achievements, as illustrated in Figure 1. Initially, student datasets are compiled by gathering various types of data, including academic records, demographic information, and online posts. The LDA is utilized to extract physiological attributes, such as mood changes and mental health, from students’ online posts. Subsequently, all attributes are inputted into the LSTM-DCNN and ReNN models simultaneously to capture temporal and correlational features from diverse perspectives. These features are then learned by the multi-view classifier to forecast the student’s academic performance categories: pass, fail, distinction, and high distinction.

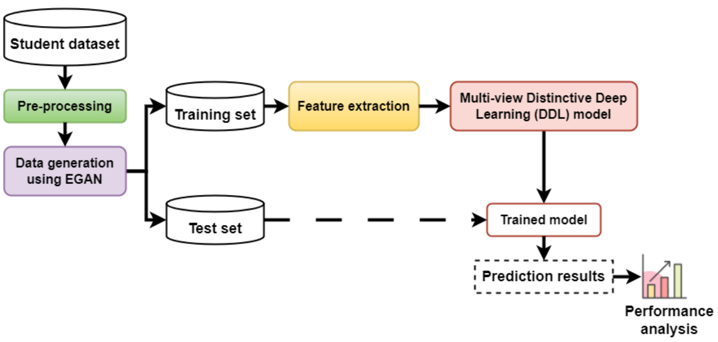

Figure 1. Schematic Representation of the Proposed Work

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 31 No. 6s (2024)

3.1 Dataset Details

A dataset consisting of 50,000 data instances has been curated by gathering student documents from both government and self-financed engineering universities in Coimbatore, Tamil Nadu. Data was collected between June 2022 and December 2022. This dataset comprises 39 attributes, encompassing students' name, age, sex, programs, types of college (medical or engineering and government or self-financed), locality, family type (nuclear or joint), family traits such as profession and qualification of family members, time spent watching TV, college traits, economic traits, social traits, electrical appliances, individual traits, and learning traits. Locality elements refer to the geographical positioning of students' residences, educational institutions, and universities, which include rural, urban, or semi-urban settings. College traits include study resources such as lecture notes and books, teaching methodologies, class sizes, smartphone usage policies, and other related factors. Social traits include parental guidance for homework, the number of friends, and academic achievements. Furthermore, there exists an additional dataset comprising diverse information disclosed by 15,000 students regarding their educational aptitudes on Twitter, Facebook, and Instagram. Next, the dataset undergoes pre-processing using several NLP approaches [8] to exclude undesirable data.

The EGAN model is used to generate synthetic student data by combining DivGAN and Suc-GAN [9]. It generates 30,000 data instances for the student dataset and 50,000 for the social media dataset. This produces an augmented student dataset with 80,000 instances and a social media dataset with 65,000 instances. The LDA feature extraction method is then applied to extract physiological features related to the student's mental health and mood changes from the online information dataset [8].

3.2 Multi-View Distinctive Deep Learning Model for Student Accomplishment Prediction

The proposed multi-view deep network consists of three main components. Initially, feature extractors such as LSTM-DCNN and ReNN simultaneously process the input data. Next, the extracted intermediate representations from both deep learning models are inputted into a multi-view classifier as two distinct feature sets for predicting students' academic performance. Figure 2 illustrates the overall structure of the proposed model.

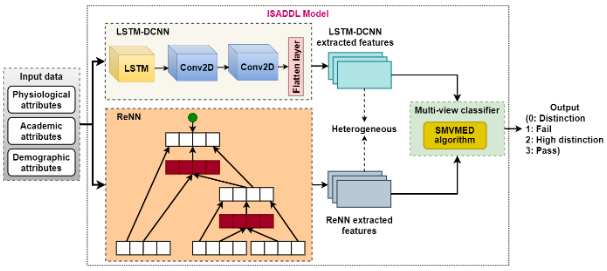

Figure 2. Overall Architecture of Proposed Model

3.2.1 Extracting Features from LSTM-DCNN

To produce temporal correlation features from the input data, the LSTM-DCNN is applied [8]. The proposed model utilizes extracted intermediate variables as individual views instead of passing output features to a perceptron. This approach is based on the idea that these variables capture the most characteristics from the LSTM-DCNN. By treating these variables as separate views, the efficiency of perceptron can be leveraged for basic linear calculations, making them suitable features for classification. In subsequent steps, supervised multi-view classifiers train these intermediate variables as the initial view.

3.2.2 Extracting Features from ReNN

The proposed ReNN model enhances the standard ReNN by incorporating a nested neural layer to compute a new feature for the parent vector. This new feature, combined with the children's vectors, forms the parent's vector representation. Each node uses a softmax classifier to predict its label. Consider V_{c1} and V_{c2} are children's vectors and V_m is the new feature that is calculated by $V_m = f\left(w^1 \begin{bmatrix} V_{c1} \\ V_{c2} \end{bmatrix} = b\right)$. So, V_p is the parent vector and calculated by $V_p = f\left(w^2 \begin{bmatrix} V_{c1} \\ V_{c2} \\ V_m \end{bmatrix} = b\right)$. In this stage, the extracted intermediate features represent the most significant characteristics. In the next steps, supervised multi-view classifiers will train these features as a second exclusive view.

3.2.3 Multi-View Classification

The intermediate features extracted from previous layers are treated as two distinct views. This enables the training of the multi-view classifier on student documents and their academic performance labels from various viewpoints. By leveraging the heterogeneous properties of the dataset through multi-view learning, a function can be learned on each view and trained jointly to improve overall performance. The Soft Margin Consistency-based Multi-View Maximum Entropy Discrimination (SMVMED) algorithm is selected for multi-view classification in this context. The SMVMED aims to achieve margin consistency in a flexible manner by maximizing the relative entropy between the fundamental parameters of two view margins. This approach allows for a trade-off parameter to balance between a large margin and margin consistency, achieved by minimizing the KL-divergence between the fundamental parameters of the margin views.

The SMVMED algorithm deals with a multi-view dataset $\{X_t^1, X_t^2, y_t | 1 \le t \le n\}$, where X_t^1 and X_t^2 are the t^{th} samples from the first and second views, respectively, and y_t is their corresponding labels. It aims to learn two discriminant functions, $L_1(X_t^1|\theta_1)$ and $L_2(X_t^2|\theta_2)$ for the first and second views, respectively. The parameters θ_1 and θ_2 are the classifier parameters. The algorithm assumes dependent distributions $p(\theta_1)$ and $p(\gamma)$ for the first view and $q(\theta_1)$ and $q(\gamma)$ for the second view. The joint distributions are $p(\theta_1, \gamma) = p(\theta_1)p(\gamma)$ and $q(\theta_2, \gamma) = q(\theta_2)q(\gamma)$, where $\gamma = \{\gamma_t | 1 \le t \le n\}$ is the margin parameter. Additionally, $p(\theta_1)$ and $q(\theta_2)$ are the posteriors of θ_1 and θ_2, respectively, and $p(\gamma)$ and $q(\gamma)$ are the posteriors of the margins from the first and second views, respectively. Thus, the optimization problem is formulated as follows:

$$\min_{p(\theta_1,\gamma)q(\theta_2,\gamma)} KL(p(\theta_1)\|p_0(\theta_1)) + KL(q(\theta_1)\|q_0(\theta_1)) + (1-\alpha)KL(p(\gamma)\|p_0(\gamma)) + (1-\alpha)KL(q(\gamma)\|q_0(\gamma)) + \alpha KL(p(\gamma)\|q(\gamma)) + \alpha KL(q(\gamma)\|p(\gamma)) \quad (1)$$

$$\text{Such that} \int p(\theta_1,\gamma)[y_t L_1(X_t^1|\theta_1) - y_t]d\theta_1 d\gamma \ge 0$$

$$\int q(\theta_2,\gamma)[y_t L_2(X_t^2|\theta_2) - y_t]d\theta_2 d\gamma \ge 0, \text{ where } 1 \le t \le n$$

In Eq. (1), α is a parameter that balances the large margin and soft margin consistency. When solving for the partial derivatives of the Lagrangian in Eq. (1) becomes challenging $(p(\theta_1,\gamma) = q(\theta_2,\gamma) = 0)$, an iterative approach is used. The solution is found by updating $p^m(\theta_1, \gamma)$ and $q^m(\theta_2, \gamma)$ by solving two problems:

$$p^m(\theta_1, \gamma) = \min_{p^m(\theta_1, \gamma)} KL(p^m(\theta_1) \| p_0(\theta_1)) + (1-\alpha) KL(p^m(\gamma) \| p_0(\gamma)) + \alpha KL(p^m(\gamma) \| q(\gamma)) \quad (2)$$

$$\text{Such that } \int p^m(\theta_1, \gamma)[y_t L_1(X_t^1|\theta_1) - y_t] d\theta_1 d\gamma \geq 0\,, 1 \leq t \leq n$$

Also,

$$q^m(\theta_2, \gamma) = \min_{q^m(\theta_2, \gamma)} KL(q^m(\theta_2) \| q_0(\theta_2)) + (1-\alpha) KL(q^m(\gamma) \| q_0(\gamma)) + \alpha KL(q^m(\gamma) \| q(\gamma)) \quad (3)$$

$$\text{Such that } \int q^m(\theta_2, \gamma)[y_t L_2(X_t^2|\theta_2) - y_t] d\theta_2 d\gamma \geq 0\,, 1 \leq t \leq n$$

The solution to this optimization problem typically takes on the following common forms:

$$p^m(\theta_1, \gamma) = \frac{1}{Z_1^m(\lambda_1^m)} p_0(\theta_1)[p_0(\gamma)]^{1-\alpha}[q^{m-1}(\gamma)]^{\alpha} exp\left(\sum_{t=1}^{n} \lambda_{1,t}^m \, [y_t L_1(X_t^1|\theta_1) - y_t]\right) \quad (4)$$

$$q^m(\theta_2, \gamma) = \frac{1}{Z_2^m(\lambda_2^m)} q_0(\theta_2)[q_0(\gamma)]^{1-\alpha}[p^{m-1}(\gamma)]^{\alpha} exp\left(\sum_{t=1}^{n} \lambda_{2,t}^m \, [y_t L_2(X_t^2|\theta_2) - y_t]\right) \quad (5)$$

In Eqns. (4) and (5), $Z_1^m(\lambda_1^m)$ and $Z_2^m(\lambda_2^m)$ are the normalization constants and λ_1^m and λ_2^m are set by finding the maximum of the following objective functions:

$$J_1^m(\lambda_1^m) = -\log Z_1^m(\lambda_1^m) \quad (6)$$

$$J_2^m(\lambda_2^m) = -\log Z_2^m(\lambda_2^m) \quad (7)$$

The convergence of the iteration process is determined by calculating the relative error between values of Eqns. (6) and (7) from two successful iterations. The iteration ends when the relative errors calculated by Eqns. (8) and (9) are both less than the tolerance (ϵ).

$$\frac{J_1^m(\lambda_1^m) - J_1^{m-1}(\lambda_1^{m-1})}{J_1^{m-1}(\lambda_1^{m-1})} \quad (8)$$

$$\frac{J_2^m(\lambda_2^m) - J_2^{m-1}(\lambda_2^{m-1})}{J_2^{m-1}(\lambda_2^{m-1})} \quad (9)$$

After obtaining $p(\theta_1)$ and $q(\theta_2)$, the label of a new input sample (X^1, X^2) is predicted by the following Eqns. (10) and (11):

$$\hat{y}_1 = sign(\int p(\theta_1) L_1(X^1|\theta_1) d\theta_1) \quad (10)$$

$$\hat{y}_2 = sign(\int p(\theta_2) L_2(X^2|\theta_2) d\theta_2) \quad (11)$$

By integrating the two views, the absolute prediction rule is as follows:

$$\hat{y} = sign\left(\frac{1}{2}\int p(\theta_1) L_1(X^1|\theta_1) d\theta_1 + \frac{1}{2}\int p(\theta_2) L_2(X^2|\theta_2) d\theta_2\right) \quad (12)$$

Accordingly, the multi-view classifier, using the SMVMED algorithm, can predict a student's academic performance by learning heterogeneous characteristics from the input data.

4. Experimental Results

This section presents the EGAN-ISADDL model's effectiveness and compares it with previous models, including SADDL [8], EGAN-SADDL [9], GRU [18], ASIST [19], and DGNN [20], in MATLAB 2019b. To achieve this, all of the existing models were also tested using the dataset presented in Section 3.1. Table 1 presents the parameter settings for the proposed and existing models.

Table 1. Parameter Settings

Parameters	Range
Learning rate	0.0001
Optimizer	Adam
Epochs	120
Batch size	64
Dropout	0.5
Momentum	0.9

The performance evaluation metrics encompass the following measures:

- **Accuracy:** It is the proportion of an accurate expectation of scholars' results to the overall amount of data tested.

$$Accuracy = \frac{True\ Positive\ (TP) + TrueNegative\ (TN)}{TP + TN + False\ Positive\ (FP) + False\ Negative\ (FN)} \tag{13}$$

TP is the number of +ve records (pass) correctly expected as pass, TN is the number of -ve records (fail) correctly expected as fail, FP is the number of -ve records incorrectly expected as pass, and FN is the number of +ve records incorrectly expected as fail.

- **Precision:** It is determined by

$$Precision = \frac{TP}{TP + FP} \tag{14}$$

- **Recall:** It is determined by

$$Recall = \frac{TP}{TP + FN} \tag{15}$$

- **F-measure:** It is determined as:

$$F - measure = 2 \times \frac{Precision \cdot Recall}{Precision + Recall} \tag{16}$$

- **RMSE:** It is the discrepancy of actual grade (y_g) and predicted grade (p_g).

$$RMSE = \sqrt{\sum (y_g - p_g)^2 / N} \tag{17}$$

In Eq. (17), N is the total number of actual performance values.

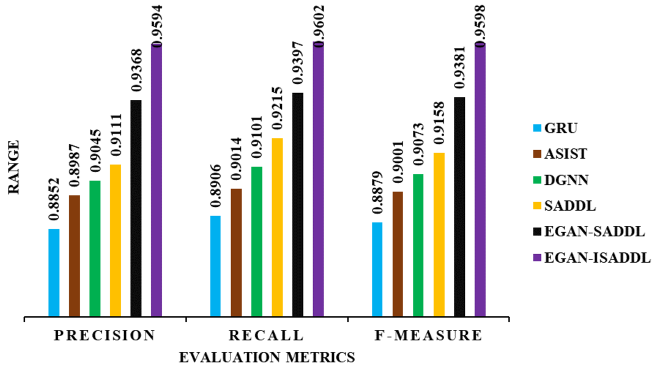

Figure 3. Precision, Recall, and F-measure of Different Models for Student's Academic Performance Prediction

Figure 3 shows the performance metrics of different models for predicting student academic performance. The EGAN-ISADDL model outperforms others because it uses multi-view learning. Compared to GRU, ASIST, DGNN, SADDL, and EGAN-SADDL, the proposed EGAN-ISADDL achieves higher precision, recall, and F-measure scores. Specifically, EGAN-ISADDL improves precision by 8.38%, 6.75%, 6.07%, 5.3%, and 2.41% over the other models, while also increasing recall by 7.81%, 6.52%, 5.5%, 4.2%, and 2.18%, and F-measure by 8.1%, 6.63%, 5.79%, 4.8%, and 2.31%, respectively.

Figure 4 shows the accuracy of various models for predicting academic achievements. The EGAN-ISADDL model outperforms other models, including GRU, ASIST, DGNN, SADDL, and EGAN-SADDL, by 8.89%, 8%, 6.16%, 4.98%, and 1.64%, respectively. This indicates that the EGAN-ISADDL model excels at accurately predicting student achievements through multi-view learning compared to single-view learning used by other models.

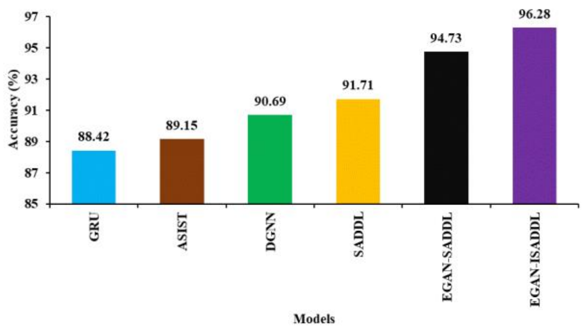

Figure 4. Accuracy of Different Models for Student's Academic Performance Prediction

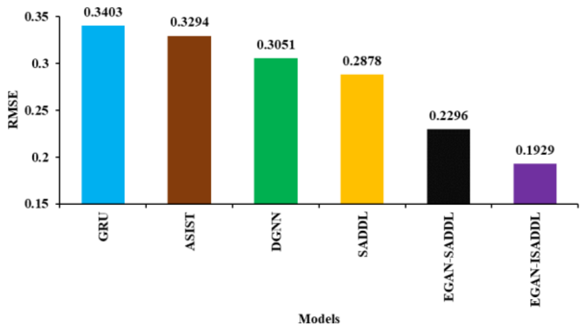

Figure 5. RMSE of Different Models for Student's Academic Performance Prediction

Figure 5 shows a comparison of RMSE for different student academic performance prediction models. The EGAN-ISADDL model outperforms the GRU, ASIST, DGNN, SADDL, and EGAN-SADDL models, reducing the RMSE by 43.31%, 41.44%, 36.77%, 32.97%, and 15.98%, respectively. This indicates that the EGAN-ISADDL model effectively improves the accuracy of predicting student academic achievement by using multi-view learning.

5. Conclusion

This paper presents the EGAN-ISADDL model, a novel multi-view learning approach for predicting student academic performance. By integrating features extracted from the LSTM-DCNN and ReNN models and using the SMVMED algorithm for multi-view classification, the proposed model effectively handles heterogeneous student data and captures complex relationships. Extensive experiments demonstrated the superiority of EGAN-ISADDL over existing models, achieving 96.28% accuracy and 0.1929 RMSE. These results highlight the model's potential as a valuable tool for educational institutions in the early identification of at-risk students and the implementation of targeted interventions. The EGAN-ISADDL model represents a significant advancement in student performance prediction, offering improved accuracy and robustness in handling diverse data sources. It has the potential to enhance educational support systems and contribute to better academic outcomes for students.

References

[1] Mori, Y., Tiiri, E., Khanal, P., Khakurel, J., Mishina, K., & Sourander, A. (2021). Feeling Unsafe at School and Associated Mental Health Difficulties among Children and Adolescents: A Systematic Review. *Children*, *8*(3), 232.

[2] Zhang, Y., Yun, Y., An, R., Cui, J., Dai, H., & Shang, X. (2021). Educational Data Mining Techniques for Student Performance Prediction: Method Review and Comparison Analysis. *Frontiers in Psychology*, *12*, 698490.

[3] Khan, I., Ahmad, A. R., Jabeur, N., & Mahdi, M. N. (2021). An Artificial Intelligence Approach to Monitor Student Performance and Devise Preventive Measures. *Smart Learning Environments*, *8*, 1-18.

[4] Bressane, A., Zwirn, D., Essiptchouk, A., Saraiva, A. C. V., de Campos Carvalho, F. L., Formiga, J. K. S., ... & Negri, R. G. (2024). Understanding the Role of Study Strategies and Learning Disabilities on Student Academic Performance to Enhance Educational Approaches: A Proposal Using Artificial Intelligence. *Computers and Education: Artificial Intelligence*, *6*, 100196.

[5] Pallathadka, H., Wenda, A., Ramirez-Asís, E., Asís-López, M., Flores-Albornoz, J., & Phasinam, K. (2023). Classification and Prediction of Student Performance Data Using Various Machine Learning Algorithms. *Materials Today: Proceedings*, *80*, 3782-3785.

[6] López-García, A., Blasco-Blasco, O., Liern-García, M., & Parada-Rico, S. E. (2023). Early Detection of Students' Failure Using Machine Learning Techniques. *Operations Research Perspectives*, *11*, 100292.

[7] Yang, S. J., Ogata, H., Matsui, T., & Chen, N. S. (2021). Human-Centered Artificial Intelligence in Education: Seeing the Invisible Through the Visible. *Computers and Education: Artificial Intelligence*, *2*, 100008.

[8] Venkatachalam, B., & Sivanraju, K. (2023). Predicting Student Performance Using Mental Health and Linguistic Attributes with Deep Learning. *Revue d'Intelligence Artificielle*, *37*(4), 889-899.

[9] Baghyalakshmi, V., & Sivanraju, K. (). Data Augmentation for Enlarging Student Feature Space for Deep Learning-based Student Accomplishment Prediction.

[10] Arashpour, M., Golafshani, E. M., Parthiban, R., Lamborn, J., Kashani, A., Li, H., & Farzanehfar, P. (2023). Predicting Individual Learning Performance Using Machine-Learning Hybridized with the Teaching-Learning-Based Optimization. *Computer Applications in Engineering Education*, *31*(1), 83-99.

[11] Al-Azazi, F. A., & Ghurab, M. (2023). ANN-LSTM: A Deep Learning Model for Early Student Performance Prediction in MOOC. *Heliyon*, *9*(4), e15382.

[12] Albreiki, B., Habuza, T., & Zaki, N. (2023). Extracting Topological Features to Identify At-Risk Students Using Machine Learning and Graph Convolutional Network Models. *International Journal of Educational Technology in Higher Education*, *20*(1), 23.

[13] Liu, Y., Hui, Y., Hou, D., & Liu, X. (2023). A Novel Student Achievement Prediction Method Based on Deep Learning and Attention Mechanism. *IEEE Access*, *11*, 87245-87255.

[14] Sun, D., Luo, R., Guo, Q., Xie, J., Liu, H., Lyu, S., ... & Song, S. (2023). A University Student Performance Prediction Model and Experiment Based on Multi-Feature Fusion and Attention Mechanism. *IEEE Access*, *11*, 112307-112319.

[15] Albahli, S. (2024). Efficient Hyperparameter Tuning for Predicting Student Performance with Bayesian Optimization. *Multimedia Tools and Applications*, *83*(17), 52711-52735.

[16] Vives, L., Cabezas, I., Vives, J. C., Reyes, N. G., Aquino, J., Cóndor, J. B., & Altamirano, S. F. S. (2024). Prediction of Students' Academic Performance in the Programming Fundamentals Course Using Long Short-Term Memory Neural Networks. *IEEE Access*, *12*, 5882-5898.

[17] Noviandy, T. R., Zahriah, Z., Yandri, E., Jalil, Z., Yusuf, M., Yusof, N. I. S. M., ... & Idroes, R. (2024). Machine Learning for Early Detection of Dropout Risks and Academic Excellence: A Stacked Classifier Approach. *Journal of Educational Management and Learning*, *2*(1), 28-34.

[18] Baniata, L. H., Kang, S., Alsharaiah, M. A., & Baniata, M. H. (2024). Advanced Deep Learning Model for Predicting the Academic Performances of Students in Educational Institutions. *Applied Sciences*, *14*(5), 1963.

[19] Fazil, M., Rísquez, A., & Halpin, C. (2024). A Novel Deep Learning Model for Student Performance Prediction Using Engagement Data. *Journal of Learning Analytics*, 1-19.

[20] Huang, Q., & Zeng, Y. (2024). Improving Academic Performance Predictions with Dual Graph Neural Networks. *Complex & Intelligent Systems*, 1-19.