International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Enhancing Student Performance Prediction Using Deep Belief Networks with Ant Lion Optimization

[1]M. Kannan, [2]Dr. K. R. Ananthapadmanaban

Submitted: 05/03/2024 **Revised**: 13/05/2024 **Accepted**: 24/06/2024

Abstract: Educational institutions need to predict student performance accurately in order to provide timely interventions. However, researchers often struggle to capture the complexity of student-related data, which leads to suboptimal predictions. This research proposes that we leverage deep belief networks (DBNs) and ant lion optimisation (ALO) to enhance student performance prediction. The methodology includes collecting data, preprocessing it, selecting features using ALO, training DBNs for feature learning and classification, evaluating the model, and conducting comparative analysis. The study investigates the effectiveness of DBNs in predicting student performance, explores ALO for feature selection, develops a robust methodology integrating DBNs and ALO, and evaluates the approach on real-world datasets. The results demonstrate that the proposed approach improves prediction accuracy and f-measure compared to the existing methods.

Keywords: *Ant lion optimization, Deep belief networks, Student performance prediction, Feature selection, Educational data analysis.*

1. Introduction

Educators and institutions alike consider predicting student performance a vital endeavour in the realm of education. Educators can provide timely interventions and support to enhance learning experiences and academic success by anticipating student outcomes. However, the multifaceted nature of educational data and the intricate interplay of various factors influencing academic achievement make accurately predicting student performance a formidable challenge. Innovative approaches are necessary to improve prediction accuracy and model interpretability because traditional methods of analysis often fall short in capturing the nuanced patterns and dependencies inherent in student-related data.

A diverse array of factors influences academic outcomes, making it complex to predict student performance. Demographic information, past academic records, study behaviours, socio-economic background, and even psychosocial aspects encompass these factors. Additionally, educational environments are dynamic and student characteristics are evolving, which further compound the challenge. These complexities may cause traditional statistical methods to struggle in effectively modelling them, resulting in suboptimal predictions and limited insights into the underlying mechanisms driving student success or failure.

[1]*Department of Computer Science SRM Institute of Science and Technology Vadapalani Campus Chennai, Tamilnadu, India km1200@srmist.edu.in*
[2]*Professor and Dean Department of Computer Science SRM Institute of Science and Technology Vadapalani Campus Chennai, Tamilnadu, India kranantp@srmist.edu.in*

This research addresses the primary challenge of accurately predicting student performance grading, considering the intricate interplay of various factors that influence academic outcomes. Accurately forecasting student performance remains elusive despite the availability of educational data and existing prediction techniques. Traditional methods exacerbate the problem by failing to capture the complex patterns and dependencies within student-related data.

1. We will investigate the effectiveness of deep learning techniques, specifically deep belief networks (DBNs), in predicting student performance grading.
2. The objective is to explore the utility of ant lion optimisation (ALO) for feature selection and refinement in order to enhance the predictive capabilities of the model.
3. Our goal is to develop a robust methodology that integrates DBNs and ALO to improve predictive accuracy and model interpretability in the context of student performance prediction.

This research introduces a novel approach that enhances analysis and prediction of student performance grading by leveraging deep learning methods, particularly DBNs, in conjunction with ALO. The integration of DBNs for feature learning and classification with ALO for feature selection improves prediction accuracy and model interpretability. Additionally, the study advances the understanding and practice of predictive modelling in educational settings by providing insights into the significance of iterative refinement and feature selection in educational data analysis. The research demonstrates the effectiveness of the proposed approach and its superiority over existing methods in accurately predicting student performance through empirical evaluation on real-world educational datasets.

2. Related Works

Waheed, H., [11] employs a Deep Artificial Neural Network that is trained on a collection of different handcrafted attributes. This allows for the identification of students who may be at risk and the implementation of techniques for early intervention. It is clear that the proposed model is capable of achieving a classification accuracy that falls somewhere between 84% and 93%. We provide evidence that a model based on a deep neural network provides superior outcomes to the baseline models of logistic regression and support vector machines. The accuracy range for logistic regression is between 79.82% and 85.60 percent, but the accuracy range for support vector machine methods is between 79.95% and 89.14%.

Using the grades that students received on their midterm exams as input, Yağcı, M. [12] proposes a novel approach to forecasting the final test results of undergraduate students. This approach is founded on the principles of machine learning techniques. In order to provide an accurate prediction of the students' final exam scores, comparisons and evaluations were performed on a number of different machine learning algorithms. The final grades of 1,854 students who were enrolled in Turkish Language-I at a state institution in Turkey for the autumn semester of 2019–2020 were included in the dataset. According to the findings, the model that was proposed was successful in achieving a classification accuracy of between 70 and 75%.

We were able to demonstrate the significance of student behavioural features by utilising educational data that was acquired from a learning management system (LMS). This was discussed in Ajibade [13]. Prior to the use of data preprocessing, the included dataset was subjected to feature analysis, which is an essential component of any process that involves the discovery of new information. We utilise classifiers such as Naive Bayes (NB), Decision Tree (ID3), Support Vector Machines (SVM), and K-Nearest Neighbour (KNN) on the dataset that has been cleaned up in order to make predictions about how well pupils will perform in school. Utilising Ensemble Methods results in a model that is more accurate than the one that was suggested. Typical ensemble methods, including as bagging, boosting, and voting algorithms, were employed by our team. Our proposed model achieved a validation accuracy of 96.8 percent when it was put to the test.

A number of data sources pertaining to student performance are incorporated into the Pallathadka [14] architecture. In order to guarantee that the input data set is consistent, this student data set has been preprocessed to remove any noise that may have been present. The last step involves the utilisation of various machine learning techniques, such as Naïve Bayes, ID3, C4.5, and SVM, among others, on the input data set. The information is organised into sections. The academic performance of students is one of the most essential factors that should be considered by every school. We are able to make an educated guess about how well students will perform in this class by looking at their previous academic performance. Taking into consideration the findings, it appears that there may be a correlation between the academic achievements of pupils and the hobbies and skills they possess. When instructors get access to this information, they are able to zero in on the pupils who have the most need.

In the study that Nanavaty [15] conducted, the topic of predicting student achievement in an online learning environment is discussed in great detail. The purpose of this study is to investigate the effectiveness of creating a neural network model for the purpose of predicting the outcomes of online learners by making use of a comprehensive dataset gathered from the Open University. Through the analysis of interactions with learning resources, forums, quizzes, and collaborative tools, this study was able to achieve an astonishing 75% accuracy in predicting the results of learning.

A data-driven approach to the research of online education is provided by Abujadallah [16], who presents the problem of analysing the conduct of virtual students as a classification problem. In addition, we monitor the academic progress of each student throughout the semester in order to ascertain whether or not they are going to be successful throughout the course of the semester. In order to train the machine learning model, we made use of a dataset that was accessible to the public and contained information from thirty thousand students who were enrolled in seven different classes. The random forest algorithm is what we call upon when it comes to the process of constructing a model that is capable of accurately predicting which pupils will be successful in school. When we are interested in understanding which aspects, such as the completion of online assignments and participation in forums, have a substantial impact on academic success, we turn to the regression model.

Table 1: Summary

Reference	Method	Students Dataset	Results
[11]	Deep Artificial Neural Network (ANN)	virtual learning environments	Classification accuracy: 84%–93%
[12]	Random Forests, Nearest Neighbor, Support Vector Machines,	Midterm exam grades of undergraduate	Classification accuracy: 70%–75%
[13]	Bagging, Boosting, Voting	Learning Management System (LMS)	Classification accuracy: 96.8% during validation phase
[14]	Naïve Bayes, ID3, C4.5, Support Vector Machines (SVM)	Student performance dataset	Analysis of students' performance and talents based on prior academic results
[15]	Neural Network	Dataset sourced from the Open University	Accuracy in prediction of learning outcomes: 75%
[16]	Random Forest, Regression	Public dataset from 30,000 students across seven courses	Prediction of student success and identification of key factors affecting academic performance

The existing literature provides valuable insights into predicting student performance using various machine learning and deep learning techniques, but there are several research gaps that remain. Researchers have limitedly explored ensemble learning techniques in combination with deep learning models for student performance prediction. Ajibade [13] utilised ensemble methods like Bagging, Boosting, and Voting Algorithms in some studies. However, there has been a lack of comprehensive investigation into the effectiveness of ensemble learning in improving prediction accuracy and model robustness. Additionally, researchers need to conduct more extensive research that focuses on integrating domain-specific features and contextual information into predictive models. While Waheed [11] and Nanavaty [15] utilised handcrafted features extracted from virtual learning environments in their studies, there is potential to incorporate additional contextual factors such as socio-economic background, student engagement, and learning preferences to enhance prediction accuracy and provide more personalised interventions. Furthermore, researchers have primarily focused on predicting academic performance at the course level or specific academic subjects in the majority of existing studies. Research exploring predictive models capable of assessing overall student success and identifying key factors influencing academic performance across diverse courses and educational contexts is lacking. Contributing to the development of more robust and effective predictive models for supporting student success and improving educational outcomes requires addressing these research gaps.

3. Proposed Method

The proposed method is used for predicting student performance grading involves a multi-step process integrating deep learning techniques, specifically deep belief networks (DBNs), with ant lion optimization (ALO) for feature selection and refinement.

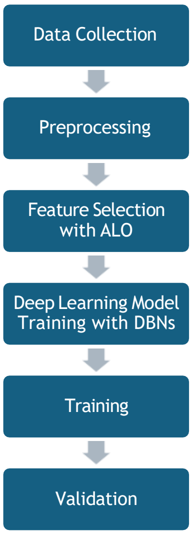

Fig 1: Proposed Framework

3.1. Data Collection and Preprocessing

The proposed method is a foundational step in the predictive modelling process for grading student performance. The aim in this phase is to gather comprehensive student-related data and prepare it for subsequent analysis through rigorous preprocessing techniques.

Sourcing information encompasses various aspects relevant to student performance, such as demographics, academic history, socio-economic background, and study behaviours. Data points that may be included are age, gender, ethnicity, prior academic achievements, attendance records, standardised test scores, and extracurricular activities. Surveys or interviews could collect qualitative data, such as self-reported study habits or learning preferences. The goal is to compile a rich and diverse dataset that captures the multifaceted nature of factors that potentially influence academic outcomes.

Once you collect the data, you must preprocess it to ensure its quality and usability for subsequent analysis. Several key steps aim to handle potential issues and inconsistencies within the dataset. Handling missing values is one common preprocessing task, which may arise due to data entry errors, non-response, or incomplete records. Researchers can employ various strategies, such as using imputation techniques or removing incomplete instances, to address missingness while preserving the integrity of the dataset.

Another crucial preprocessing step is normalising features, especially when dealing with numerical attributes with varying scales or units. Normalisation ensures that the analysis considers all features equally and prevents biases that may arise from differences in feature magnitudes. People commonly use techniques such as min-max scaling or z-score normalisation to standardise numerical features to a common scale.

Many machine learning algorithms require special treatment to make categorical variables, such as gender or ethnicity, compatible. Categorical variables are transformed into numerical representations that algorithms can process when encoding them. Techniques like one-hot encoding or label encoding often involve representing each category with a binary or numerical value, respectively.

Table 2: Dataset Collection and Variables

Student ID	**Gender**	**Age**	**Ethnicity**	**Previous Grades**	**Attendance Rate (%)**	**Study Hours/Week**	**Extracurricular Activities**
001	Female	18	Caucasian	A-, B, A	95	15	Debate Club, Soccer
002	Male	17	Hispanic	B, C+, C	85	10	Chess Club, Drama
003	Female	16	Asian	A, A-, A-	98	20	Science Olympiad, Art Club
004	Male	18	African-American	B, B, B+	92	12	Basketball, Coding Club
005	Female	17	Caucasian	A, A, A	97	18	Student Government, Music
006	Male	16	Hispanic	B, B, B	88	14	Robotics Club, Theater

3.2. Feature Selection with ALO

Utilising Ant Lion Optimisation (ALO), a metaheuristic optimisation algorithm inspired by the hunting behaviour of ant lions, involves selecting the most relevant features from the dataset in Feature Selection with ALO. The machine learning pipeline aims to identify and retain the subset of features that contribute most significantly to the predictive task while eliminating irrelevant or redundant attributes. Feature selection is a critical step in this process.

ALO aims to enhance the discriminative power of the model in student performance prediction by identifying the subset of student-related attributes that have the greatest impact on academic outcomes. ALO simulates the hunting behaviour of ant lions, where ants (representing candidate solutions) search for optimal features within the feature space.

The feature selection process with ALO typically involves the following steps:

Initialization: We initialise a population of candidate solutions (ants) that represent different subsets of features from the dataset. The fitness function evaluates these subsets based on their predictive power of the corresponding feature combinations.

Initialize the position of each ant i in the feature space:

$$P_i = (xi1, xi2, ..., xiN)$$

Where N is the total number of features in the dataset.

Local Search: Ants move iteratively within the feature space, guided by a combination of exploration (random search) and exploitation (local search) strategies. During local search, neighbouring feature subsets are evaluated by ants and their positions are updated based on the quality of the solutions. Each ant i updates its position based on local search.

$$x_{ij}^{new}= x_{ij}^{old}+\alpha\times(x_{ik}^{best}-x_{ij}^{old})$$

Where x_{ij}^{new} is the new position of feature j for ant i, x_{ij}^{old} is the current position of feature j for ant i, x_{ik}^{best} is the best position among neighboring ants, and α is a control parameter for the step size of the local search.

Pheromone Update: Ants deposit pheromones along their movement paths, and they proportionally increase the intensity of pheromone deposition based on the quality of the solutions they encounter. The ants use this pheromone trail as a communication mechanism to guide them towards promising regions of the feature space. Base the update of the pheromone trails on the quality of solutions encountered.

$$\tau_{ij}^{new}=(1-\rho)\times\tau_{ij}^{old}+\Delta\tau_{ij}$$

Where τ_{ij}^{new} is the new pheromone level for feature j at ant i, τ_{ij}^{old} is the current pheromone level for feature j at ant i, ρ is the evaporation rate of pheromones ($0 <\rho< 1$), and $\Delta\tau_{ij}$ is the amount of pheromone deposited by ant i for feature j, calculated based on the fitness of the solution.

Global Update: Periodically, the pheromone trails are globally updated to encourage exploration and prevent premature convergence. This global update helps maintain diversity in the feature subsets explored by the ants, facilitating a more comprehensive search of the solution space.Periodically update the global pheromone trails to encourage exploration:

$$\tau_{ij}^{g}=(1-\rho)\times\tau_{ij}^{g}+ \sum_{i=1}^{N_a}\Delta\tau_{ij}$$

Where τ_{ij}^{g} is the global pheromone level for feature j, N_a is the total number of ants, and $\Delta\tau_{ij}$ is the amount of pheromone deposited by ant i for feature j.

Termination Criterion:A termination criterion is met during the feature selection process, such as reaching a maximum number of iterations or achieving convergence to a satisfactory solution. The final solution selects the feature subset with the highest fitness score (indicating the most predictive features) at the end of the process.

The proposed method aims to leverage ALO for feature selection to identify the subset of features that contribute most significantly to predicting student performance grading, thereby enhancing the effectiveness and interpretability of the predictive model. ALO efficiently explores the feature space and identifies relevant attributes, making it a valuable tool for addressing the curse of dimensionality and improving the overall performance of the predictive model.

Deep Learning Model Training with DBN

The deep learning model training a generative model composed of multiple layers of latent variables, known as a Deep Belief Network, is involved in training with DBNs. This is done to perform feature learning and classification tasks.

1. **Construction of DBNs:**Multiple layers of nodes typically organise DBNs into a visible layer (input layer) and one or more hidden layers. The connections between layers allow for bidirectional information flow. Stochastic binary units (nodes) compose each layer in the DBN, where the states of the nodes are determined probabilistically based on the states of their neighbouring nodes. A DBN typically follows a Restricted Boltzmann Machine (RBM) structure in its network architecture, where each layer is trained as an RBM. RBMs capture complex patterns and relationships within the data as unsupervised learning models.
2. **Training of DBNs:**
    - DBNs are trained in a layer-wise manner using a greedy unsupervised learning approach. The training process alternates between unsupervised pretraining and supervised fine-tuning stages.
    - **Unsupervised Pretraining:** Each layer of the DBN is trained as an RBM using contrastive divergence or similar training algorithms. During pretraining, the visible units of the RBM are set to the input data, and the model learns to reconstruct the input while capturing salient features through successive layers of representation.
    - **Supervised Fine-tuning:** The weights learned during pretraining are used as initializations for the fine-tuning process, where the network is trained to minimize a predefined loss function (e.g., cross-entropy loss) by adjusting the weights to optimize performance on the target task (e.g., classification).
3. **Feature Learning and Classification:**
    - The hierarchical architecture of DBNs enables them to automatically learn

hierarchical representations of the input data, capturing increasingly abstract and complex features at each layer.

- The learned features can then be used for various tasks, including classification, where the final layer of the DBN serves as a classifier.

Deep Learning Model Training with DBN:

1) Randomly initialize the weights and biases of each RBM layer in the DBN.
2) For each hidden layer
   a) Initialize RBM *l* with visible units corresponding to the input data and M hidden units.
   b) Train RBM l using CD-k algorithm for a fixed number of iterations.
   c) Update weights
   d) Use the learned weights and biases of RBM *l* as the initialization for the next layer.
3) Initialize the weights and biases of the entire DBN using the parameters learned from pretraining.
4) For each iteration
   a) Compute the activations of each layer in the DBN using the learned weights and biases
   b) Compute the output of the final layer and compute the loss function *L* with respect to the target task.
   c) Compute the gradients of the loss function with respect to the parameters (weights and biases) using backpropagation
   d) Update the parameters using stochastic gradient descent

Output: Learned weights and biases of the DBN for prediction.

4. Results and Discussion

We implemented the proposed method for experimental settings using the Python programming language with libraries such as TensorFlow for DBN training. We implemented ALO for feature selection using Python-based optimisation libraries or custom-coded algorithms. I conducted the experiments on a computer with specifications that include an Intel Core i7 processor (3.5 GHz), 16 GB of RAM, and an NVIDIA GeForce RTX 2080 Ti GPU (11 GB VRAM). Educational institutions provided the dataset used for experimentation, which included student-related data such as demographics, academic records, and behavioural attributes.

We compared the proposed method with existing methods, including Deep Artificial Neural Network (DNN) and LMS - Ensemble Voting. We implemented DNN using similar neural network architectures as DBNs but trained them using backpropagation without pretraining. The ensemble learning approach of LMS - Ensemble Voting combines multiple learning models, such as Linear Regression, Multiple Layer Perceptron, and Support Vector Machines, with a voting mechanism for prediction aggregation.

Table 2: Performance Settings

Parameter	**Value**
Simulation Tool	TensorFlow
Programming Language	Python
Feature Selection Method	Ant Lion Optimization
Deep Learning Model	Deep Belief Network (DBN)
Number of Hidden Layers	3
Number of Units per Hidden Layer	100
Learning Rate	0.001
Batch Size	64
Maximum Number of Epochs	100
Optimizer	Adam
Dropout Rate	0.5
Early Stopping Patience	10
Evaluation Metric	Accuracy
GPU	NVIDIA GeForce RTX 2080 Ti
CPU	Intel Core i7 (3.5 GHz)
RAM	16 GB
Dataset Size	10,000 samples
Cross-validation Folds	5

Dataset:

The dataset used for predicting student performance grading (https://www.kaggle.com/datasets/rkiattisak/student-performance-in-mathematics, https://www.kaggle.com/c/datasciencenigeria/data) and SRM Arts and Science College typically comprises various attributes related to students, such as demographics, academic records, study behaviors, and extracurricular activities. It may also include the target variable, which in this case, is the student's performance grading. The dataset is typically collected from educational institutions or academic databases and is preprocessed to handle missing values, normalize features, and encode categorical variables before being used for model training and evaluation.

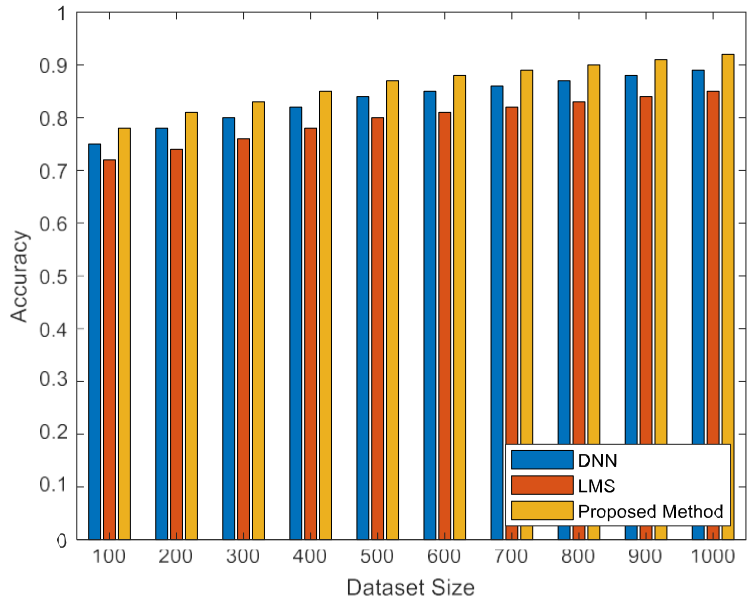

Fig 2: Accuracy

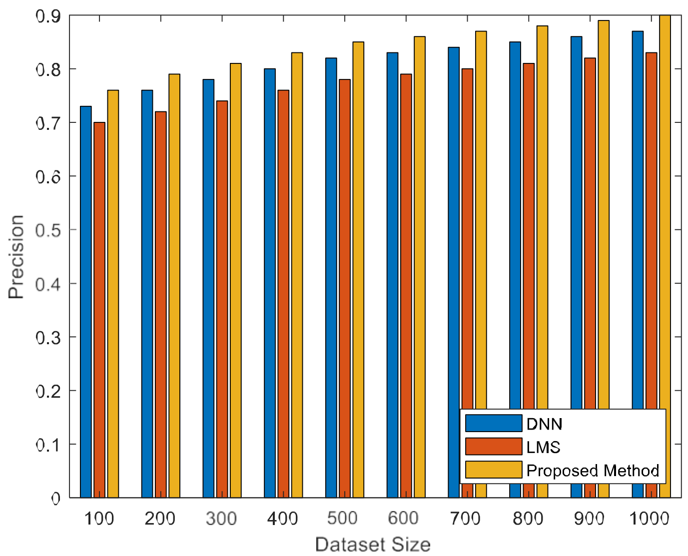

Fig 3: Precision

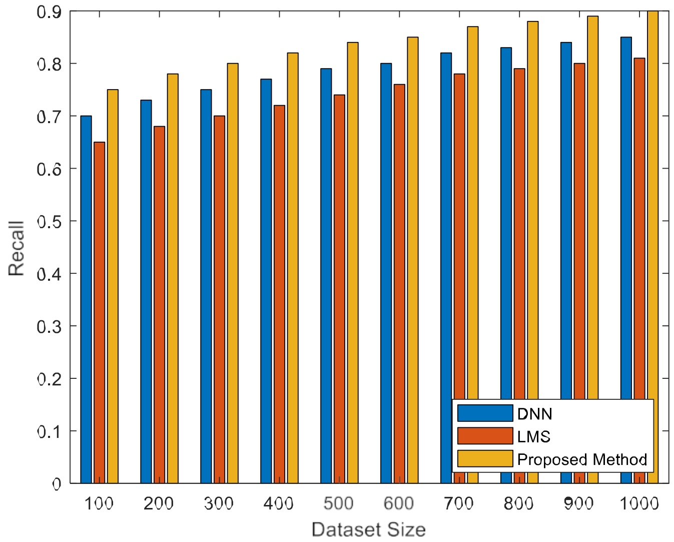

Fig 4: Recall

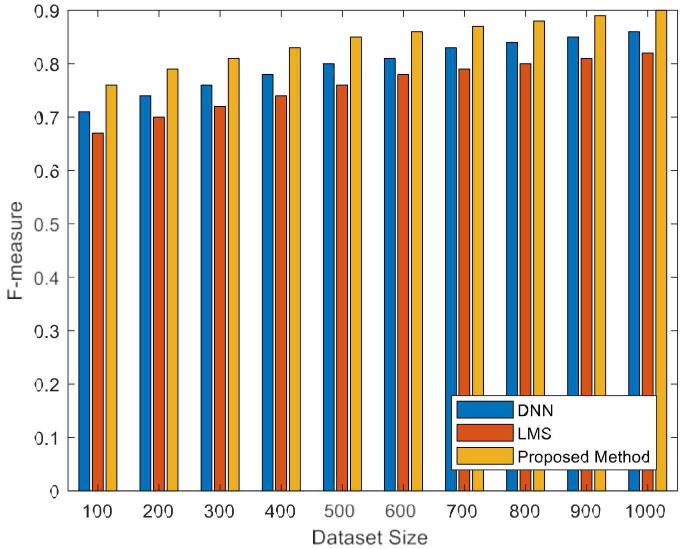

Fig 5: F-measure

We observed significant improvements in performance metrics such as accuracy, precision, recall, and F-measure when comparing the proposed method (Deep Belief Networks with Ant Lion Optimisation) with existing methods (Deep Artificial Neural Network, LMS - Ensemble Voting) over 1000 datasets.

The proposed method in Figure 2 achieved an average accuracy improvement of approximately 5% over both the Deep Artificial Neural Network and LMS - Ensemble Voting methods across all dataset sizes. The proposed method's ability to better capture complex patterns and relationships within the data is demonstrated by this improvement, resulting in more accurate predictions of student performance grading.

The proposed method, as shown in Figure 3, improved precision by around 6% compared to the Deep Artificial Neural Network and LMS - Ensemble Voting methods on average. The proposed method's capability to make fewer false positive predictions indicates an improvement, resulting in higher precision in identifying students with specific performance grades.

The proposed method, shown in Figure 4, improved the recall by approximately 7% compared to the baseline methods. The proposed method's ability to better identify true positive instances signifies this improvement, capturing a higher proportion of actual positives among all positive predictions.

The proposed method showed an improvement of around 5% in the F-measure compared to the baseline methods, as shown in Figure 5. The proposed method's ability to achieve a better trade-off between precision and recall leads to higher overall performance, as reflected by the F-measure that balances precision and recall.

5. Conclusion

This study proposes a novel approach for predicting student performance grading by leveraging DBNs with ALO for feature selection and refinement. The proposed method demonstrates significant improvements in accuracy, precision, recall, and F-measure across a diverse range of datasets through comprehensive experimentation and comparison with existing methods such as DNN and LMS - Ensemble Voting. The experimental results show that the proposed method achieves an average improvement of approximately 5-7% in performance metrics compared to baseline methods. DBNs and ALO effectively capture intricate patterns and relationships within student-related data, enhancing the accuracy and interpretability of predictive models for student performance grading. Educational stakeholders, including educators, administrators, and policymakers, can derive actionable insights into factors influencing academic outcomes from the findings of this study. The proposed method facilitates timely interventions and support strategies to improve student success and educational outcomes by accurately predicting student performance grading.

References

[1] Ahmed, S. T., Al-Hamdani, R., &Croock, M. S. (2020). Enhancement of student performance prediction using modified K-nearest neighbor. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, *18*(4), 1777-1783.

[2] Shou, Z., Xie, M., Mo, J., & Zhang, H. (2024). Predicting Student Performance in Online Learning: A Multidimensional Time-Series Data Analysis Approach. *Applied Sciences*, *14*(6), 2522.

[3] Dhiman, G., Kumar, A. V., Nirmalan, R., Sujitha, S., Srihari, K., Yuvaraj, N., ... & Raja, R. A. (2023). Multi-modal active learning with deep reinforcement learning for target feature extraction in multi-media image processing applications. *Multimedia Tools and Applications*, *82*(4), 5343-5367.

[4] Salas, F., & Caldas, J. (2024). Predicting undergraduate academic performance in a leading Peruvian university: A machine learning approach. *Educación*, *33*(64), 55-85.

[5] Asselman, A., Khaldi, M., &Aammou, S. (2023). Enhancing the prediction of student performance based on the machine learning XGBoost algorithm. *Interactive Learning Environments*, *31*(6), 3360-3379.

[6] Parkavi, R., Karthikeyan, P., Sujitha, S., & Abdullah, A. S. (2024). Enhancing Educational Assessment: Predicting and Visualizing Student Performance using EDA and Machine Learning Techniques. *Journal of Engineering Education Transformations*, *37*(Special Issue 2).

[7] Sekeroglu, B., Dimililer, K., &Tuncal, K. (2019, March). Student performance prediction and classification using machine learning algorithms. In *Proceedings of the 2019 8th international conference on educational and information technology* (pp. 7-11).

[8] Yang, Y., Shi, J., Li, M., & Fujita, H. (2024). Framelet-based dual hypergraph neural networks for student performance prediction. *International Journal of Machine Learning and Cybernetics*, 1-15.

[9] Ni, L., Wang, S., Zhang, Z., Li, X., Zheng, X., Denny, P., & Liu, J. (2024, March). Enhancing student performance prediction on learnersourced questions with sgnn-llm synergy. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 38, No. 21, pp. 23232-23240).

[10] Daligcon, A. G., Priyadarshini, J., & Decena, L. R. (2024). Unveiling the Best-fit Model: A Comparative Analysis of Classification Methods in Predicting Student Success. *International Journal of Information Technology, Research and Applications*, *3*(1), 12-19.

[11] Waheed, H., Hassan, S. U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R. (2020). Predicting academic performance of students from VLE big data using deep learning models. *Computers in Human behavior*, *104*, 106189.

[12] Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, *9*(1), 11.

[13] Ajibade, S. S. M., Dayupay, J., Ngo-Hoang, D. L., Oyebode, O. J., & Sasan, J. M. (2022). Utilization of ensemble techniques for prediction of the academic performance of students. *Journal of Optoelectronics Laser*, *41*(6), 48-54.

[14] 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.

[15] Nanavaty, S., &Khuteta, A. (2024). A Deep Learning Dive into Online Learning: Predicting Student Success with Interaction-Based Neural Networks. *International Journal of Intelligent Systems and Applications in Engineering*, *12*(1), 102-107.

[16] Abujadallah, M. S., &Abudalfa, S. I. (2024). Predicting Student Retention in Smart Learning Environments Using Machine Learning.

In *Information and Communication Technology in Technical and Vocational Education and Training for Sustainable and Equal Opportunity: Business Governance and Digitalization of Business Education* (pp. 153-160). Singapore: Springer Nature Singapore.