

A Robust Student Performance Forecasting Framework Based on Autoencoder and Belief Network Learning

Thangam Somasundaram^{1*}, CH. Mohan Sai Kumar², G.Merlin Suba³,
Milind P Gajare⁴, S Pournima⁵, Kousika N⁶

¹Department of Computer Science and Engineering, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Bengaluru-560035, Karnataka, India, s_thangam@blr.amrita.edu (Corresponding Author)

²Assistant Professor, Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India, chmohansaikumar@veltech.edu.in

³Associate Professor, Department of Electrical and Electronics Engineering, Panimalar Engineering College, Chennai- 600123, Tamil Nadu, India, merlinsubag@gmail.com

⁴Assistant Professor, Department of Electronics and Telecommunication Engineering, AISSMS, Institute of Information Technology, Pune-411001, Maharashtra, India, milind.gajare@aissmsioit.org

⁵Assistant Professor, Department of CSE, Sona College of Technology (Autonomous), Salem 636 005, Tamil Nadu, India, pournimascse@gmail.com

⁶Associate Professor, Sri Krishna College of Engineering and Technology, Coimbatore, Tamil Nadu, 641008, India. Email: kousika@skcet.ac.in

HOW TO CITE:

Thangam Somasundaram, CH.
Mohan Sai Kumar, G.Merlin Suba,
Milind P Gajare, S Pournima, Kousika
N (2026). A Robust Student
Performance Forecasting Framework
Based on Autoencoder and Belief
Network Learning.
International Journal of Special
Education, 41(6s), 1138-1156.

ABSTRACT:

Student performance forecasting in the educational sector is of paramount significance because it allows for the early recognition of at-risk students and the ability to give targeted interventions leading to enhanced learning. The majority of the conventional machine learning approaches are hand- and knowledge-engineered feature-based and lack good generalization, while deep models lack good interpretability and require enormous amounts of labeled data. All these disadvantages affect the application of student performance prediction systems on real learning settings. As another approach to such problems, this paper proposes a self-learning system and a hybrid feature selection model that uses the Dingo-Coot Fusion Selector (DiCoot-FS) for feature selection at the level of optimality and the Autoencoder and Belief Network-driven Self-Taught Learning (AutoBel-STL) model for the accuracy of student performance prediction. The highest contribution of this study is its ability to adaptively improve input features, learn informative feature, and apply probabilistic adaptation for improved classification performance. DiCoot-FS integrates dingo-based local exploitation with coot-based global exploration to eliminate redundant features and preserve key academic and behavior variables that affect the performance of students. Concurrently, AutoBel-STL integrates feature compression using autoencoders and probabilistic learning helped by belief networks for self-learning-based adaptation over distributions of unlabeled data. It substantially improves robust predictability and generality on diverse student populations. The

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.
Open access publication under the
terms and conditions
of the Creative Commons Attribution
(CC BY)
license (<http://creativecommons.org/licenses/by/4.0/>).

proposed framework offers a number of advantages compared to existing solutions, such as improved accuracy, reduced complexity of computation, as well as improved flexibility on handling diversified sets of students. Furthermore, the proposed system performs optimally with the lowest computational time of 1.8 seconds. The proposed method provides an opportunity for top-level educational analysis, enabling colleges to implement individual learning strategies and maximize academic performance for students by measuring and intervening in an initial performance stage.

Keywords: Education Data Mining (EDM), Students Performance Prediction, Deep Learning, Feature Selection, Classification.

1. Introduction

Quality education is among the most crucial pillars in commencing sustainable development on the global scale, providing intellectual human growth and overall development of nations (Rizwan et al., 2025). Yet, achievement and sustenance of high quality in education always prove to be the biggest challenge, particularly during the period of rapid development of technology and continuously growing volumes of information about education. The growing digitalization of academic processes has resulted in the continuous creation of huge volumes of data from diverse sources like learning management systems, e-portals for learning, admission portals, and student feedback systems (Narayanan et al., 2025, Onker et al., 2025). All these sources cumulatively contribute to an enormous pool of academic data, which can transform academic decision-making if maximized to its limit. However, most of this data is not usable as it is difficult to access, disorganized, and noisy. Therefore, classical data analysis methods are not in a position to identify useful information. Instead of exploring the entire dataset, contemporary decision-making systems opt for superficial query-level decisions and hence cannot detect the thoughtful underlying patterns and trends that could potentially be deciding factors in enhancing education outcomes. To this issue, the discipline of Educational Data Mining (EDM) has attracted significant interest from researchers who desire to extract useful information from data concerning students. EDM is concerned with the use of data mining, machine learning, and deep learning to

extract educational data with the purpose of enhancing the quality of learning procedures as well as student performance predictive models. One of the major purposes of EDM (Yani & Kusriani, 2025, Domínguez et al., 2025) is to design intelligent systems capable of predicting a student's performance from historical models of performance, e.g., academic background, behavioral pattern, participation score, and test score. Student performance predictions would give school administrators, teachers, and policymakers a data base to make recommendations to them on ways instruction could be enhanced, individualize instruction to students, and offer early support to students that are having trouble.

By the integration of deep learning algorithm to EDM, business entities can provide useful intelligence to several metrics impacting the performance of students such as learning habits, learning performance through online learning, utilization of learning material, and usage of instructors and peers (Jiabin et al., 2025). All this can further be used in creating smart recommendation systems with tailored learning routes, adaptive learning systems, and immediate feedback systems based on certain requirements of the student. Predictive analytics will also help identify probable fail students in an attempt to initiate early intervention and targeted support intent strategies in order to achieve highest retention rates and academic performance. Moreover, deep learning analytics can help institutions achieve maximum curriculum design, enhance assessment practice,

and reallocate resources in a way to preserve quality learning outcomes (Tjahyadi & Tude, 2025). With educational data volumes doubling exponentially, the application of deep learning to EDM is a critical move towards creating intelligent, data-driven education spaces for effective learning and global access to quality education everywhere.

Evaluation of student performance through past data has been a main area of investigation in the area of EDM, and many studies have focused on predictive model development to enable early predictions of academic achievement or failure. The primary focus of such research is to predict student performance in terms of marks, grades, and pass/fail performance so that it may assist the teachers and institutes in implementing predictive measures for academic excellence (Habeeb et al., 2025). Prediction in the initial phase is critical because it helps in timely intervention, personalized study plans, and special support for those students who lag behind. But even with the tremendous progress in EDM, reliably predicting student academic performance from large and noisy data accurately is still a difficult task (Arcinas et al., 2025). Two such issues most closely pertinent to our concern are low predictor variable quality and use of traditional machine learning methods that include traditional feature representation paradigms after classifiers, both of which limit the scope of these models in representing the true underlying determinants of performance in students.

One of the significant shortcomings in current research is the failure to have good selection of predictor variables in order to effectively measure student performance. All the models become restricted to a limited set of features like demographic characteristics, past grades, and historical attendance without regard for many other variables which would influence performance such as behavioral engagement, learning style, participation in extracurricular activities, psychological conditions, and socio-economic status. Leaving out these essential features reduces models to curt ones that fail to capture several aspects of scholarship.

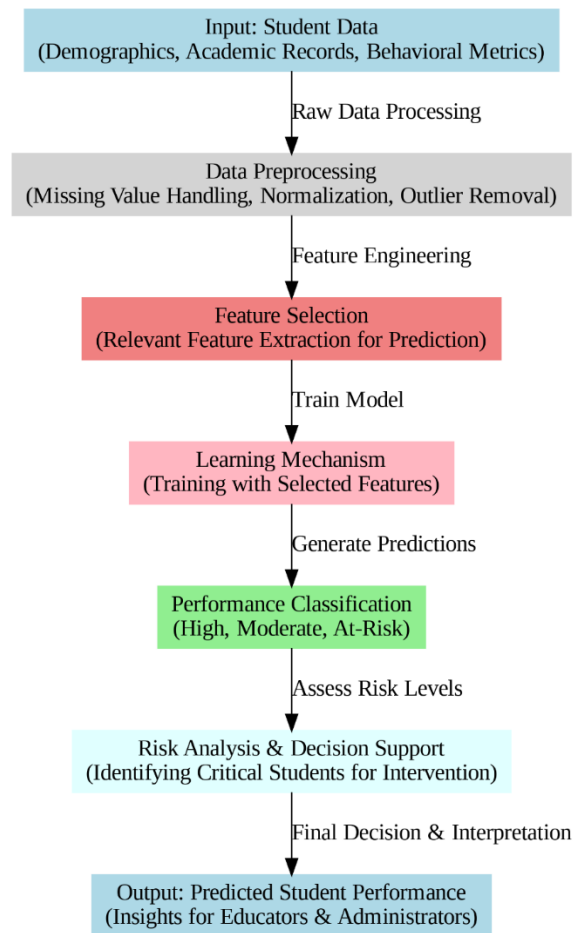


Figure 1. Typical students' performance prediction system

Moreover, various educational situations call for variant sets of predictors, and one thing that takes precedence in a given situation will not necessarily find itself as having the same superiority in another context (Alnasyan et al., 2024). Another major limitation is excessive reliance on traditional machine learning approaches that mainly rely on classical feature representation techniques with a classifier added to them (Subramaniam, 2026). These techniques mainly include feature engineering manually, wherein researchers specify some features in advance to be used within the model, extract features from data, and then apply a classifier such as decision trees, SVM, or random forests. While all these methods have gained universal usage, they degrade when applied to deal with noisy, high-dimensional, unstructured learning data (Chen et al., 2025). Classical feature engineering methods can miss weak patterns in the data, and thus, sub-optimal performance of the predictive models can

be obtained. Furthermore, human-intensive feature generation is restrictive in scalability since optimally choosing and tuning features for various datasets is a time-consuming process and subject to human errors too. Such models either need extensive preprocessing or fail to generalize in such a situation, thus leading to lower accuracy and reliability.

As shown in Figure 1, the sudden growth of computer-based learning together with the widespread advent of availability of student data has enabled an unprecedented capability of utilizing AI in facilitating greater precision in the forecast of student performance. While unprecedented advancement has been made through EDM, current practice is vulnerable to being assaulted by sharp issues related to heterogeneity of data, noisy and missing values, and generalizability to school settings. Traditional machine learning approaches highly rely on manually designed feature extraction and shallow models of classification, which are typically not strong enough to detect subtle patterns of learning behavior among students. Even highly capable deep models (Kuadey et al., 2024, Villar & de Andrade, 2024), in their turn, tend to typically necessitate large datasets of labeled information and intensive computing resources, which is typically out of the question in most real-world educational settings. To address the above issues, a new-generation, extensible, and flexible solution is needed that can efficiently utilize the increasing amount of learning data to reach its potential in manners that cannot be achieved through traditional predictive models. This requires the development of a hybrid deep transfer learning-based student performance prediction model whose purpose is to utilize the capacity of deep learning models and the adaptability of transfer learning to develop a robust and effective prediction model that will generate high accuracy even in conditions that are marked by limited labeled data.

One of the most significant push factors for this suggested research is solving data sparsity and domain adaptation issues in student performance prediction. It is true that most schools generate very large amounts of data, yet their data sets are too imbalanced, noisy, or domain-adapted and

therefore it's not an easy thing to find a single prediction model that would be equally effective on multiple schools (Ikegwu et al., 2024). Transfer learning, a strong AI platform, enables the capability to transfer models learned in one domain to another domain with little retraining, hence the removal of expensive large-scale labeled sets. With the use of student performance prediction and deep transfer learning, the suggested solution can utilize pre-trained models of neural networks that are learned from large educational data and learned general high-level feature representations and therefore can generalize across new unseen students. This is especially advantageous to schools with no rich history data in which case traditional deep learning models would otherwise be likely to overfit and have bad generalizability. Traditional machine learning and deep learning models usually operate on static data, which makes them unable to adapt to these changes over time (Sahlaoui et al., 2024). The suggested hybrid deep transfer learning method integrates ongoing learning processes within the model such that the model updates its knowledge base from time to time as new student data is integrated. Also, because the model has taken advantage of leverage in transfer learning, it can be fine-tuned between various institutions, geographies, or disciplines with no large-scale retraining, thus making the method very scalable and affordable in multiple learning environments.

The driving force behind this purported work is to overcome the natural barrier of being capable of predicting students' performance under a hybrid deep transfer learning framework by the synergy of deep neural networks' capability and transfer learning's adaptability. In the process of doing so, in solving problems which are data sparsity, domain adaptation, interpretability, and model flexibility biased, this research aims at building an extremely accurate, comprehensible, and scalable prediction model which shall be capable of assisting instructors, institutions, and policymakers in making sensible, data-driven decisions for enhancing learning outcomes. The method proposed not only broadens the scope of existing EDM applications but also offers the doorways to large-scale deployments of AI-driven educational analytics for delivering individualized care and

attention to the students required to achieve academic excellence.

The rest of the paper is organized in the following manner: Section 2 is an organized literature review of the existing methods used for the prediction of student performance, such as traditional machine learning approaches, deep learning techniques, and ensemble models, and their strength, weakness, and research loopholes. Section 3 describes in detail the structure, major components, and methodological innovations of the proposed hybrid deep transfer learning model. Section 4 presents the performance assessment and comparative results, showing the efficacy of the envisioned strategy based on benchmarked datasets, performance measures, and comparisons with current different existing approaches. Lastly, Section 5 concludes the paper by highlighting significant contributions, suggesting some possible applications, and discussing the research avenues to further significantly improve student performance prediction system scalability and resilience.

The accurate forecast of student performance has been a very common research topic in the area of EDM with a substantial amount of research studies applying machine learning and deep learning techniques to enhance educational decision-making (Lee et al., 2024). Models will give satisfactory accuracy in student risk prediction and grade prediction but will be deficient in feature selection weakness, insufficiency of data, and non-generalization of intricate student learning behavior patterns. These conventional approaches are also hand-crafted feature-based and hence extremely sensitive to domain knowledge and data quality, which were not capable of generalizing over all learning schools and settings. Models have also shown higher ability to handle large sets of education data, detect non-linear patterns, and make predictions. But deep learning techniques are more likely to need humongous quantities of labeled training data, which would otherwise not be available in education because of data access constraints [17, 18]. In such a shift in paradigm of research, the present section describes an overview of methodologies' paradigms pursued and categorizes them as classical machine learning

paradigms, deep learning paradigms-based models, and hybrid models and critically assesses their weaknesses, strengths, and improvement areas.

Rizwan, et al (Rizwan et al., 2025) introduce an up-to-date taxonomy of student academic performance and engagement prediction factors impacted by deep learning (DL). The research utilizes qualitative research on the basis of a Systematic Literature Review (SLR), further corroborated using the PRISMA model, in order to provide a rigorous selection and assessment process of studies relevant to the research. The authors also give a conceptual description of MOOCs and especially for hard-of-hearing and visually impaired students, recording the imperative of inclusive learning settings. The study also gives an assembly of machine learning (ML) and deep learning (DL) algorithms used to predict students' educational performance and specifies the usefulness and efficiency of these. Among its contributions is the classification of some of the features utilized in the models for prediction, a rigorous ranking of the significant features that lead to students' success. Balachandar and Venkatesh (Balachandar & Venkatesh, 2025) propose a new Multi-dimensional Student Performance Prediction Model (MSPP) that combines state-of-the-art data preprocessing and feature engineering methods with a deep learning model. The new approach is specifically developed to solve the common education dataset issue of temporal dependency and class imbalance data distribution, and it offers a more stable and improved prediction model. Incidentally, the MSPP model also includes explainability features to help with greater degrees of transparency and interpretability to AI-driven decisions. State-of-the-art architecture decreases the errors in classification to a great extent, providing a detailed and powerful way of anticipating academic performance.

Alalawi, et al (Alalawi et al., 2025) put forth a new Learning Analytics and Intervention (LAI) approach, which they term Student Performance Prediction and Action (SPPA). SPPA seeks to provide course-level prediction models to instructors at zero meaningful institutional costs. SPPA allows teachers to create customized analytics from past course test history records for the

identification of poor performing students, detection of knowledge gaps, and learning plan formulation for generating customized learning plans. Intervention is the hardest obstacle to knowledge for learning analytics, and the solution framework overcomes it with pedagogical consideration integration within course planning and intervention design. With such processes, SPPA can enable that educational interventions are optimized the most to offer learner-specified learning needs-based, focused support through personalized analytics using available learning materials. Nayani, et al (Nayani & R, 2025) suggest hybrid deep learning model with entropy rough set theory optimized to develop precise prediction of student performance. They address it starting from a preprocessing stage, including outlier removal and best data-filling policy to maintain the integrity of the data. The primary contribution to such an effort is the introduction of a new meta-heuristic hybrid algorithm, Galactic Rider Swarm Optimization (GRSO), for optimizing feature selection. As a classification model, an instance of a deep learning-based Convolutional Recurrent Network (CRN) whose performance in classification was significantly enhanced using GRSO-optimized optimization is utilized. In addition, the hyperparameters of the Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN) are also optimized using the GRSO algorithm for the purpose of obtaining a strong and adaptive classification model.

Ahmad, et al (Ahmad et al., 2025) study the use of machine learning methods for the prediction of student performance, in particular for class imbalance problem. For alleviating class imbalance, the authors use a combination of oversampling and undersampling methods in decision tree ensemble algorithms to improve risk classification for potential students. Classifier performance is explored experimentally with varying ensemble size and adjusting oversampling and undersampling rates to enable thorough examination of their effects. Second, the authors combine feature selection steps with the top ensemble classifiers and optimize the predictive accuracy. Through extensive experimentation, the authors put aside that ensemble models like Random Forest, Bagging, and

Random Undersampling Boosting score high in performance metrics like Recall, Precision, F1-score, Area Under the Receiver Operating Characteristic Curve (AUC-ROC), and Geometric Mean, maintaining their suitability to deal with imbalanced student data. Rogers, et al [23] focus on predicting the academic achievement of students based on Moodle data and feature importance analysis using tree-based machine learning models. Contrary to the earlier studies focused on enhancing the performance of models, the present study focuses on generalizability through training the model using multiple datasets with generic features. This way, the predictive model is versatile in the scope of education settings, so it is even more robust and generalizable in all senses across actual learning environments.

Kennel, et al (Kennel et al., 2025) explore performance prediction on graphically distinguished math problems using eye-tracking measures for 143 students. To gain richer insights into students' learning difficulties, they use Retrospective-Think-Aloud (RTA) protocols to enable qualitative analysis of learning difficulty. To achieve predictive modeling, they use three machine learning techniques k-nearest neighbors (KNN), Random Forest, and Support Vector Machine (SVM) to solve this multiclass classification problem. These results demonstrate that predictive validity increases with processing time increases, and highly predictive performance is achieved well ahead of students finishing solution process, indicating early intervention in the teaching of mathematics. Kordbagheri, et al (Kordbagheri et al., 2025) write about the application of novel machine learning algorithms within psychological testing as opposed to the traditional testing procedures. Instead of employing traditional test design and construction, they apply artificial intelligence (AI) and machine learning (ML) to enhance predictive validity as well as test effectiveness. In the tradition of being strong and dependable, writers clearly keep model fit measures under control with cross-validation (CV) and resampling, clearly demonstrating a commitment to publishing accurately validated and generalizable results in psychological testing. Bisri, et al (Bisri et al., 2025) compare their suggested model based on data

samples of a State Islamic University for timely or late graduation status. According to their findings, the Support Vector Machine (SVM) model, when run using shuffle sampling with the Arabic Language and Literature (BSA) dataset, performs extremely well with greater than 90% accuracy and with an Area Under the Curve (AUC) value greater than 0.9. In the same manner, for the Islamic Education Management (MPI) data set, the SVM model with stratified sampling also achieves very high performance, again over 90% accuracy and 0.9 AUC. These results affirm the reliability and effectiveness of the predictive model built, as these indicate its ability to predict students' graduation status by various fields of study.

Despite significant advances in predicting student performance, studies today still struggle with many of the most significant problems limiting the validity, scope, and generalizability of the predictive models. The weakest point of the earlier studies is probably the poor selection of predictor variables capturing the real measures of student performance. The traditional EDM methods also depend almost solely on traditional machine learning models, which are feature-engineering biased and not able to handle large noisy and imbalanced data. In addition, class imbalance with a huge disparity between high-achieving and low-achieving students makes the models biased towards the majority class and thus possessing a reduced predictability capability for disadvantaged learners. However, the other very essential problem is that current models could not apply time-series trends and sequential learning algorithms to deduce long-term trends of education from the reason that student learning achievement data are in time-series fashion.

Most studies also could not provide explainability and interpretability, and the educators have no idea

2. Proposed Methodology

In predictive analytics, precise prediction of student performance has become the basis for enhancing learning performance, selecting students at risk, and allowing early interventions. To overcome these challenges, this research suggests a new hybrid approach that combines Dingo-Coot Fusion Selector (DiCoot-FS) for best feature selection and AutoBel-STL (Automated Belief Self-Taught

Learning) for strong student performance prediction. As shown in Figure 2, the novelty of this suggested approach is that it can intelligently select informative features from high-dimensional learning datasets with ease while using self-taught learning and belief networks on simultaneous instances for adaptive and probabilistic decision-making. Differently from standard fixed-feature-

about the reason why predictive results are derived so that the corresponding intervention steps could be decided accordingly. One other drawback of current methods is that there is a limitation in the usage of deep learning methods effectively in EDM. Despite the success of deep learning models in various applications, the deployment of the model in predicting student performance remains largely unexplored because training the model is complex, much-labeled data are needed, and it incurs high computational expense (Chen et al., 2024, Pradeep & Gopalakrishnan, 2024). Furthermore, feature selection approaches in existing research are static and fail to dynamically adjust to various educational settings, resulting in decreased model robustness. The model presented here is very effective in minimizing the class imbalance of the datasets through the dataset resampling approach, i.e., oversampling and undersampling, in an effort to expose prediction models to balanced data. In addition, it also uses temporal learning elements such as recurrent and graph neural network layers to enable the model to learn the order of patterns in student performance data so that the predictive power of the model is enhanced for long-term academic performance. With the addition of adaptive hyperparameter optimization, it also optimizes the model towards maximum efficiency in different datasets with enhanced predictive power and model stability. Furthermore, explainable AI techniques (Rimal et al., 2025, Pitta et al., 2024) are also embedded in the model so that teachers and stakeholders can see why predictions are being made and observe what the most significant drivers of students' performance are. Employing scoping and transfer learning, the model can generalize adequately across hundreds of schools without requiring much annotated data.

Learning) for strong student performance prediction. As shown in Figure 2, the novelty of this suggested approach is that it can intelligently select informative features from high-dimensional learning datasets with ease while using self-taught learning and belief networks on simultaneous instances for adaptive and probabilistic decision-making. Differently from standard fixed-feature-

based machine learning algorithms, the model automatically updates the choice process assisted by DiCoot-FS, a bio-inspired search methodology that benefits from the exploration strength of Dingo Optimization Algorithm (DOA) and the exploitation potential of Coot Algorithm (CA).

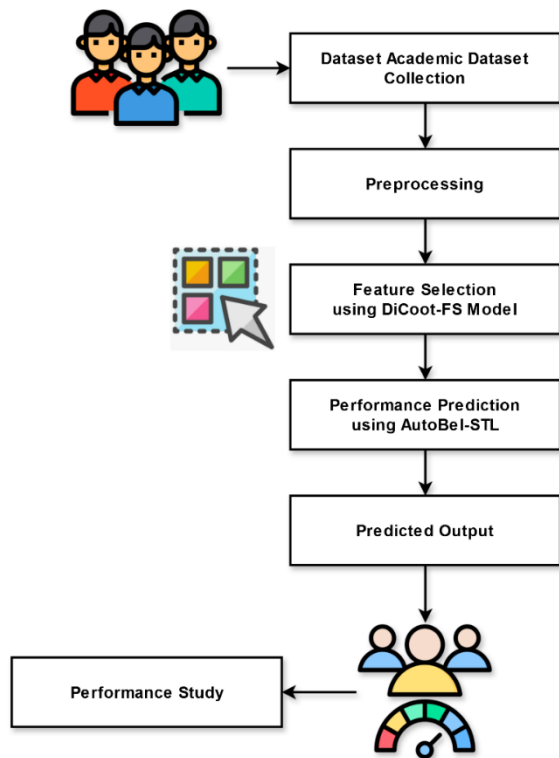


Figure 2. Overall flow of the proposed system

The novelty of this study is manifested in its synergy of feature selection and self-taught learning so that only the most appropriate and efficient student features are utilized for performance classification, hence improving predictive reliability and decision-making effectiveness. The DiCoot-FS feature selection process is critical to the improvement of the quality of the predictive model by getting rid of noisy and redundant features and keeping most explanatory features. The hybrid selection process starts with the initialization process of candidate subsets of features, where the Dingo Optimization part effectively browses the search space by mimicking the social and hunting behavior of dingoes to guarantee global diversity in the chosen set of features. The two-step selection procedure helps the model change dynamically in terms of various sets of educational data, avoiding overfitting with the ability to generalize. As soon as optimal

features are decided, they get inputted in the AutoBel-STL model, utilizing STL to reveal unconscious patterns in studying behaviors of the students. Using unsupervised representation learning, AutoBel-STL learns automatic meaningful feature representations with reduced large dataset requirements for labeled training data. Its incorporation through belief networks also strengthens decision-making by putting probabilistic dependences on predictors of student performance, enabling flexible and explainable predictions.

One of the most striking elements of this projected work is being able to unify DiCoot-FS and AutoBel-STL into one pipeline that accurately chooses features and also improves prediction insight. The second major innovation of the present work is the unsupervised feature compression mechanism of AutoBel-STL that allows the model to learn from real-world education datasets with missing or incomplete records. The method also proposes an automatically generated risk-scoring mechanism wherein the probability of students' academic difficulties is measured in an interpretable manner, giving actionable feedback to educators. This multi-step process is in such a manner that each step of student performance prediction is optimized, starting from feature extraction up to classification. The capability of the model to operate with minimal human interaction also makes it a suitable candidate for large-scale educational institutions where predictive models cannot be manually optimized.

2.1 Dingo-Coot Fusion Selector (DiCoot-FS)

Dingo-Coot Fusion Selector (DiCoot-FS) is a recently proposed hybrid feature selection technique that combines the optimality of the Dingo Optimization (DO) and Coot Algorithm (CA) for optimally selecting informative features from datasets of large dimension. The fusion strategy is particular to enhance the feature selection aimed at predictive student performance models for which gigantic amounts of education data contain redundant as well as informative information. The Dingo Optimization module simulates wild dingoes' foraging process to keep selected features diversified and prevent redundancy. Concurrently, the Coot Algorithm module, based on the collective

movement of the coot in nature, optimizes local improvement and allows for a well-ordered process of convergence. With the fusion of the two metaheuristic approaches, DiCoot-FS optimally identifies the optimal feature subset maximizing classification accuracy while minimizing computational complexity, which renders it extremely beneficial in predictive models used in educational data analysis. One of the advantages of DiCoot-FS is that it adapts automatically to the character of the dataset through an intelligent feature filtering process in terms of global and local search strategies in an overall fusion-based choice process. Feature selection through conventional methods typically is marred by premature convergence (local optima maturity) or redundant features, and this produces poor model performance. DiCoot-FS is not necessarily forced, like conventional strategies such as genetic algorithms, particle swarm optimization, or ant colony optimization, to perform a lot of hyperparameter tuning and, thus, become more pragmatic to apply on real-world problems. In addition, its hybrid exploration-exploitation trade-off ensures that the most explanatory features are retained, thus greatly increasing predictability of deep learning and machine learning models in student performance classification tasks.

DiCoot-FS has a key function in the prediction of student performance in determining the best predictors of academic success or failure. Most of these characteristics will be redundant or secondary and thus harm the prediction performance unless properly filtered. DiCoot-FS achieves this process automatically by removing noisy and irrelevant features without throwing out the most informative features, so the learning model only works on those factors in which student performance has the strongest correlation. This improves generalization, prediction accuracy, and reduces the model size, hence educational institutions can implement more accurate early warning systems for at-risk-of-failure students. Furthermore, the effective feature selection technique reduces model training time, which is more convenient for real-time predictive systems. DiCoot-FS feature selection process presents several benefits that significantly contribute towards enhancing predictive model

performance and, in our research, also prediction of student achievement. One of the most influential benefits of DiCoot-FS is ensuring the improvement in prediction accuracy using the selection of the most relevant and informative features and the avoidance of other redundant and less informative features. Student performance predictive models learned from DiCoot-FS selected features are thus more accurate, recall-based, and F1-scoreful and are thus more reliable for making education decisions.

Another important advantage of DiCoot-FS is that it can decrease computational complexity by reducing the number of chosen features without affecting model performance. High-dimensional data, especially in education analytics, have too many attributes that can greatly increase model training and inference time. Through combining Dingo Optimization to execute global search and Coot Algorithm to execute local optimization, DiCoot-FS effectively reduces the most effective subset of features in a way that machine learning models run at lower memory and processing levels. Through this minimization of computations, the approach becomes feasible for big educational datasets since typical feature selection methods may not work due to higher processing requirements. Besides, by speeding up the training process, DiCoot-FS can enable the production of real-time student performance monitoring systems that would assist teachers and school administrators in early intervention when students are highly probable to fail. DiCoot-FS adaptive and robust ability to select features is one of its strong points, and it is capable of doing so for heterogeneous data sets without the need for human intervention.

Another major advantage of DiCoot-FS is that it won't overfit, a very common problem of machine learning models working with high-dimensional data. The majority of traditional feature selection methods end up keeping too much highly redundant or correlated features, making the models memorize the patterns instead of learning them.

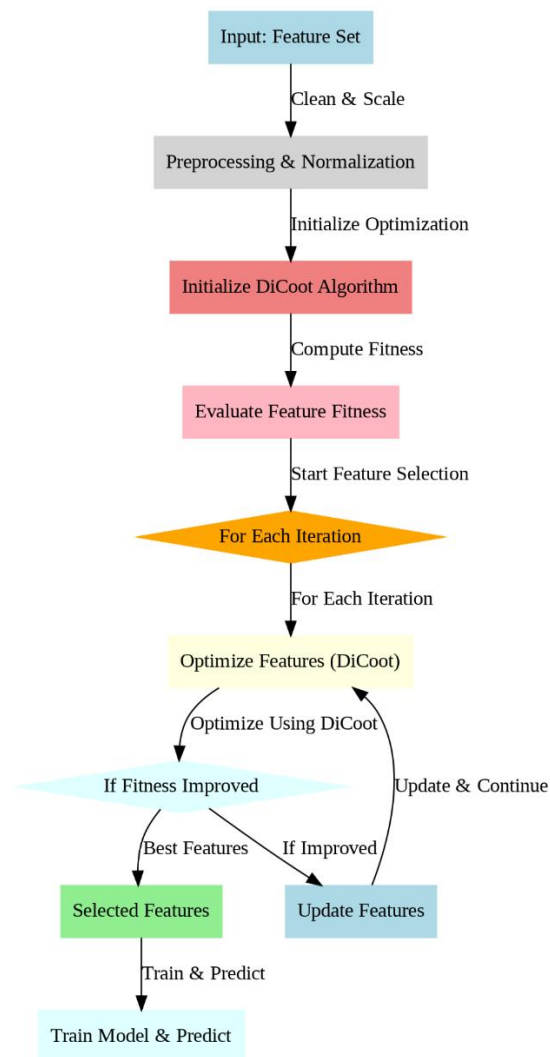


Figure 3. Flow of the proposed DiCoot-FS model

DiCoot-FS correctly solves this problem by systematic removal of noisy and redundant features in such a way that the output feature subset is used to increase model robustness. As shown in Figure 3, the models learned from DiCoot-FS-selected features are more transferable and can be applied to other schools, geographical locations, and curricula. Apart from its application in model performance and efficiency, DiCoot-FS also yields interpretable and actionable insights that can inform educators and administrators to make data-driven decisions. As opposed to the opaque black-box deep learning methods that are most probably not interpretable, DiCoot-FS is centered on choosing the features that are significant statistically and relevant in real life to the academic environment. This interpretability enables education institutions to design special intervention programs that offer necessary support

to at-risk students who are prone to underachievement. Additionally, the information that can be obtained from DiCoot-FS-selected features enhances the ability to develop personalized learning approaches, enabling the student to achieve customized learning guidance and support based on the individual's needs. Thus, DiCoot-FS not only enhances the speed and accuracy of prediction but also greatly helps improve overall student success rates through data-driven decision-making.

2.2 Autoencoder and Belief Network-driven Self-Taught Learning (AutoBel-STL)

Autoencoder and Belief Network-driven Self-Taught Learning (AutoBel-STL) approach is specifically formulated to improve student performance prediction by combining self-taught learning, autoencoders, and belief networks. The importance of this approach is that it can effectively deal with complex, high-dimensional, and imbalanced education data through a self-taught learning (STL) strategy. Conventional student performance prediction models are always hard to cope with issues like missing data, noisy patterns, and few labeled instances that extensively limit their predictive power. AutoBel-STL overcomes the issues mentioned above by an autoencoder-based feature representation mechanism that extracts significant latent features from chosen student features, while the belief network part adds probabilistic reasoning to resolve learning performance ambiguity. This two-step strategy not only enhances the resilience of predictions but also improves the explainability of students' learning behaviors, and hence it is better than traditional machine learning-based classification models.

As shown in Figure 4, the novelty of AutoBel-STL is in its three-step learning process in which autoencoders are employed to perform unsupervised feature learning, belief networks are utilized for probabilistic inference, and self-taught learning is utilized for generalization. Unlike the existing deep learning methods that are deeply reliant on large-scale labeled data, AutoBel-STL employs a semi-supervised learning process and is capable of learning from labeled and unlabeled student data. This greatly minimizes reliance on

large quantities of ground-truth labels, which are both time and resource expensive to acquire in learning environments. Further, the use of autoencoders with self-taught learning ensures one passes only the most informative, high-level feature, and denoised representations into the belief network, which will result in more accurate and context-sensitive student performance predictions. As compared to conventional supervised learning methods like decision trees, support vector machines, or shallow neural networks, AutoBel-STL is more adaptive in developing change learning environments in the sense that it detects the concealed trends and non-linear student patterns between students' features like attendance, assignment grades, learning engagement, and cognitive proficiency levels.

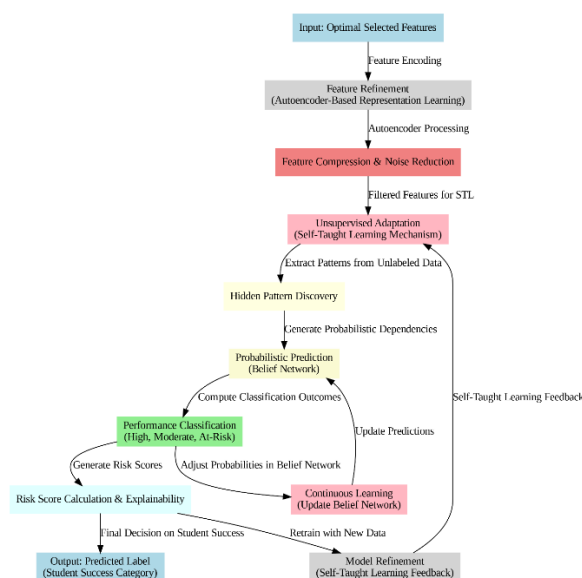


Figure 4. Flow of the proposed AutoBel-STL model

In addition, as opposed to other deep learning models with high demands for large amounts of labeled data for supervised training, AutoBel-STL effectively leverages unlabeled student data for self-supervised feature learning and is therefore highly scalable and usable in the real world. Moreover, although current ensemble-based prediction approaches like gradient boosting machines and bagging methods enhance the accuracy of predictions, they are non-explainable, and it becomes challenging to understand why a student is predicted to be high-risk. AutoBel-STL addresses

this shortcoming through the use of belief networks, which offer explicit graphical representations of the student knowledge dependencies, giving teachers enhanced comprehension of determinants of academic attainment. In the Bayesian Belief Network-inference stage, the acquired high-level encoded features in the autoencoder are provided as input to a probabilistic graph model that identifies dependencies between some of the following student traits based on prior schoolwork quality and learning patterns. Relative to standard machine classifiers with hard-margin decision boundaries, the belief network models subtle interdependencies between going to school, turning in assignments, engaging in after-class activities, cognitive ability, and prior scholastic record. Hence, this model will predict the probability of a pupil's success or failure instead of the hard crisp labels. Bayesian belief network's power is that it can handle uncertainty by always updating probability distributions based on observed patterns of pupil information, and thus it is very interpretable and flexible.

For instance, rather than just marking a student as "pass" or "fail," the model can generate an estimate that a student has a 70% chance of having high success, a 20% chance of having average success, and a 10% chance of failing. Such a prediction, in terms of probabilistic measures, provides teachers with valuable intelligence on which to focus intervention efforts on students that are exhibiting warning signs of future school problems. This is a significant improvement over conventional deep learning models that would act as black-box predictors, such that it might be challenging for teachers to identify the basis for a specific classification being applied to a specific student. Apart from enhancing flexibility, AutoBel-STL also has a self-learning learning process that allows the model to learn from labeled as well as unlabeled student data sets, hence lessening its dependence on the large annotated data sets. The biggest impediment in predicting student performance is the lack of properly labeled education data because most schools do not have complete academic records with accurate-performance measures. Self-taught learning (STL) bypasses this drawback by employing unsupervised learning methods to learn

informative patterns from unlabeled students' data in a manner that facilitates generalizability by the model more effectively over a large population of students. That is the advantage of AutoBel-STL as it is greatly scalable and versatile because it eliminates repetitive manual tagging that is wasteful of time and money in real learning contexts.

Furthermore, autonomous learning plays an important role in continuously revising the Bayesian belief network based on fresh student information in order to maintain the model up to speed with latest educational trends. Unlike retraining of static models intermittently required thereof, AutoBel-STL makes use of its autonomous learning function to improve its probabilistic relationships adaptively when cohorts of students are processed one after the other. This renders it particularly suitable for active learning environments where pedagogical strategies, curriculum design, and students' learning behaviors evolve over time. The incorporation of self-learned learning in the proposed model prevents overfitting to specific data sets while exhibiting exceptional performance on different classes of students with different learning tendencies. This ability to be adaptable also makes the model long term usable and trustworthy such that the predictions it generates are still valid even when applied to new schools or study programs.

After selecting the most required features from the given dataset, the obtained feature set is represented in the following form:

$$S = \{s_{i,j} | i \in [1, n], j \in [1, m]\} \quad (1)$$

Where, $s_{i,j}$ indicates the feature set. Then, each feature is normalized with min-max scaling operation for ensuring uniformity as shown in below:

$$s_{i,j}^{\text{norm}} = \frac{s_{i,j} - \min(s_j)}{\max(s_j) - \min(s_j)} \quad (2)$$

Where, $s_{i,j}^{\text{norm}}$ is the normalized feature set. As a consequence of this, the autoencoder is applied to map the input features into a lower dimensional space according to the following model:

$$Q = f_{\theta}(S) = \varphi(\omega S + \beta) \quad (3)$$

Where, ω and β are the weight and bias parameters, and φ denotes the sigmoid activation function. The reconstruction loss is minimized with the autoencoder component as mathematically expressed in below:

$$\mathbb{L}_{\text{AE}} = \sum_{i=1}^n \|S_i - \hat{S}_i\|^2 \quad (4)$$

Where, $\hat{S}_i$ indicates the reconstructed input. Moreover, the refined features are optimized using self-taught learning with the use of probabilistic selection as shown in below:

$$\hat{S} = \arg \max_Q P(Q|S) \quad (5)$$

With the use of probabilistic graphical model, each refined feature is applied to improve the performance prediction based on the following model:

$$P(\mathcal{T}|Q) = \sum_{i=1}^n P(t_i|q_i) \quad (6)$$

Where, Q is the predicted label. In addition to that, the Bayesian belief update rule is applied as shown in below:

$$P(t|q) = \frac{P(q|t)P(t)}{P(q)} \quad (7)$$

Where, $P(t)$ denotes the prior probability. Moreover, the hidden dependencies among the features is estimated through belief function as mathematically computed in below:

$$\mathfrak{x}(Q) = - \sum_{i=1}^n P(q_i) \log P(q_i) \quad (8)$$

Where, $\mathfrak{x}(Q)$ is the hidden dependency. The final student performance category is predicted according to the maximum a posterior function shown in the following equation:

$$\hat{t} = \arg \max_t P(t|Q) \quad (9)$$

For the predicted label, the performance index is calculated using the following equation:

$$\mathcal{C}(t) = \frac{1}{1 + e^{-\xi(P(t|Q)) - 0.5}} \quad (10)$$

Where, ξ denotes the scaling factor. Once the model has attained probabilistic dependency through the belief network and optimized learning through self-trained learning, it proceeds to the last classification and performance evaluation stage in which student performance prediction is performed using

probabilistic inference methods. Unlike traditional classification models that classify students into strict classes, AutoBel-STL provides a risk-score system according to the probability distribution of the belief network. Academic ambiguity is also quantified by the system, and it is easier to identify students with chronically low scores and those students whose score is unstable as a result of short-term external reasons such as sickness or personal matters. The classification outcome is also validated to achieve performance measures such as accuracy, precision, recall, F1-score, and AUC-ROC to ensure

3. Results and Discussion

Experimental evaluation in this work is carried out on an open-source Student Performance Dataset available on Kaggle with mixed academic and behavioral features impacting the students' performance. The dataset has impactful features like study hours, percent attendance, test score, parent education, extracurricular, and socio-economic status and therefore becomes very ideal for the performance prediction task. With focus on these attributes, the study aims to find determinants of prime importance that affect the performance of students and build an optimal predictive model. Preprocessing of data is performed for handling missing data, numerical scaling of features, and converting categorical attributes to obtain maximum performance with the best model. To measure the performance of the suggested method, diverse deep learning and machine learning models are employed and compared with other approaches. The process is experimental, and in feature selection, the irrelevant and redundant features are removed so that the model will be effective. After they have been selected, those features are trained into prediction models and assessed according to the performance based on the classic evaluation metrics. The data is split between a training set and a test set so that the model will be accurate with out-of-sample data. Cross-validation techniques are also employed to prevent overfitting and enhance reliability.

Figure 5 shows the study hours vs. the performance index, providing a critical overview of how committed study time affects performance. A positive study hours and performance index

that the model is not only accurate but also transferable between student groups. Accuracy refers to the overall accuracy of the prediction, while precision and recall ensure the ability of the model for correct prediction of students at risk without generating high false positives and false negatives. The F1-score is a balanced metric that makes sure that the model is equally good in high- and low-risk learning scenarios, while the AUC-ROC metric checks the discriminative ability of the model so that it can discriminate between good and intervention-needed students very well.

correlation would imply that longer studying students are probable better performers in academics. The trend, however, is not linear but that of other determinants of performance like quality of studying, focus, and efficient learning techniques as well. The graph assists teachers and policymakers in knowing the right amount of study time needed by students to be well-placed academically so that they can create individualized study plans. Figure 6 is an attendance vs. performance index analysis, and it shows that classroom attendance and participation are the determinants of academic performance. From this graph, it can be seen how student attendance rates are related to performance and whether regular classroom attendance means higher academic achievement. More percentages in attendance are generally associated with better performance due to the fact that students are exposed to direct instruction, peer-to-peer interaction, and frequent tests. But some others with low attendance rates can actually do well because of extraneous study or tutoring, and some other students with high attendance rates can do poorly because of poor studying or extraneous distractions. The outcome of this analysis gives teachers the possibility of recognizing those students who can be at risk because of inconsistent attendance patterns and intervening by means of such interventions as greater academic support or tailored learning recommendations.

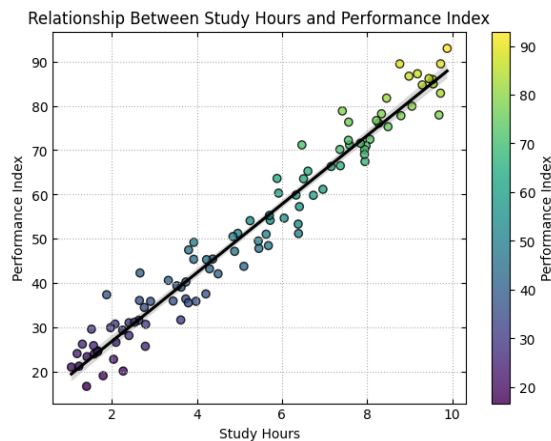


Figure 5. Relationship among the study hours and performance index

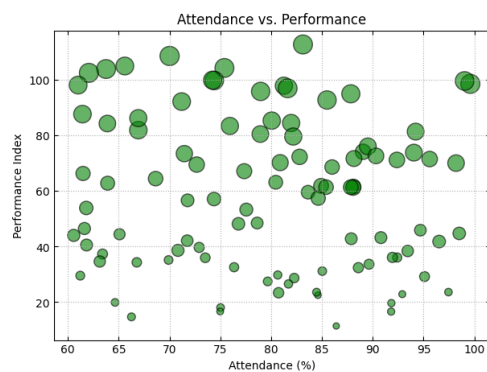


Figure 6. Attendance Vs Performance index

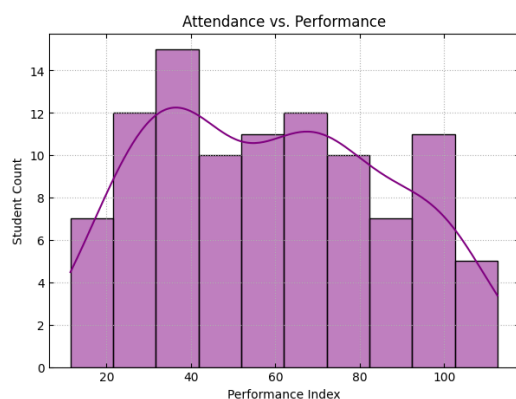


Figure 7. Histogram Distribution of Performance Index

Figure 7 is the histogram of performance index and gives us data on student performance dispersion in the dataset. The histogram allows one to read whether student scores are normally distributed, skewed towards high or low performance, or bimodal. Well-distributed appearance shows diversified academic ability in the dataset, while

skewed or heavy distribution shows that most students are within specific performance ranges.

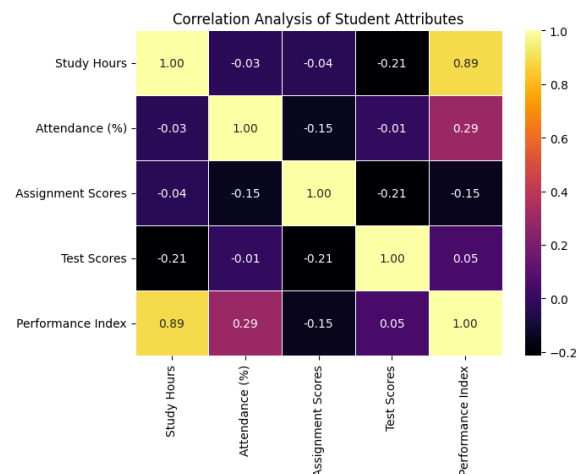


Figure 8. Correlation analysis

This information is useful in school improvement initiatives because it informs us of whether the majority of the learners require assistance or not, and if there are few who require radical intervention. Based on this histogram, instructors can plan study strategies that suit the requirements of low-performing students so that they receive proper attention while those doing well are provided with increased study time. Fig 8 indicates a correlation analysis at the level of detail of study time, attendance rate, marks in assignments, test marks, and performance index.

The correlation heatmap reveals the correlation between different factors and the best performing predictors of the performance of the students. Large values of correlation between test scores, assignment marks, and performance index show that test and assignment performance have strong predictive values for overall academic performance. Moreover, moderate correlations among study time and achievement demonstrate that even though study time is an important issue, intervening variables such as study approach, understanding of principles, and unrelated instructional tools are also influential. Percentage attendance is extremely critical in substantiating the hypothesis that regular class activity attendance contributes to academic achievement. The correlation analysis allows researchers and teachers to aim at the most critical parameters and eliminate less important

characteristics, hence making academic performance models predictive. The data can be used in an attempt to maximize learning methods for exposing learners to the required assistance to facilitate academic achievement.

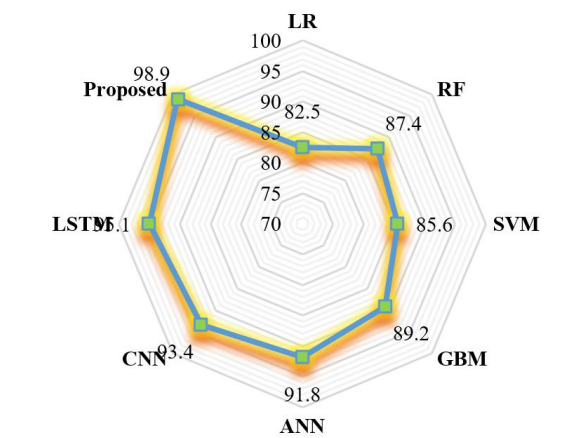


Figure 8. Comparison with existing techniques based on accuracy

Table 1. Overall performance comparison with existing prediction models

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
LR	82.5	81.2	80.5	80.8
RF	87.4	86.8	85.9	86.3
SVM	85.6	84.9	84.3	84.6
GBM	89.2	88.5	87.8	88.1
ANN	91.8	90.9	90.2	90.5
CNN	93.4	92.7	91.8	92.2
LSTM	95.1	94.5	94	94.2
Proposed	98.90	98.85	98.80	98.9

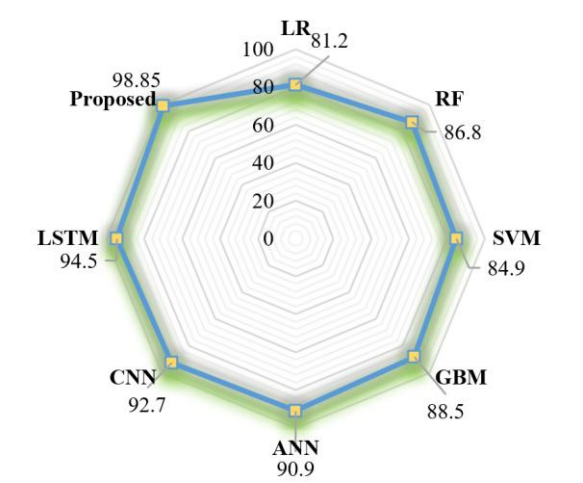


Figure 9. Comparison with existing techniques based on precision

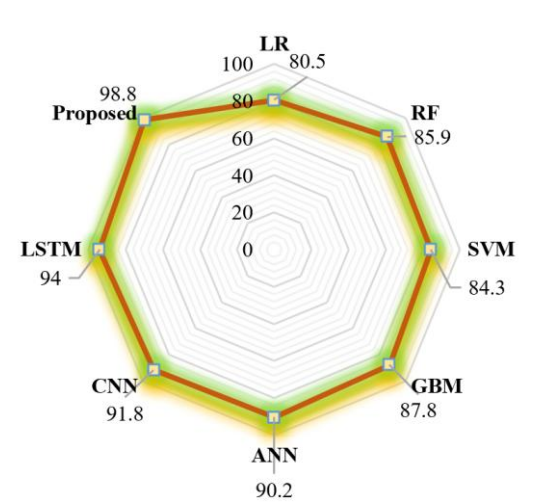


Figure 10. Comparison with existing techniques based on recall

The comparative performance of the proposed model with other machine learning and deep learning algorithms is depicted in Table 1 that shows the performance of various algorithms in terms of accuracy, precision, recall, and F1-score. It can be achieved from the outcome that the model suggested in this work is far better than standard and state-of-the-art models for all measurement parameters with a performance rate of 98.90% in accuracy, which is remarkably greater than other methods. The proposed model is most accurate with a rate of 98.90%, showing how efficient the model is at describing complex relations between study variables and performance. Figure 8 is a graphical representation of direct accuracy over the different methods that indicates the improvement brought about by the proposed model. However, the proposed model also enhances these predictions by nearly perfect accuracy, which shows its capability to generalize over a wide range of different student performance scenarios.

Figure 9 shows the comparison of accuracy, which is a robust measure quantifying the proportion of positive observations that are classified correctly. Logistic Regression has 81.2%, i.e., it can't avoid classification errors. SVM and Random Forest have some kind of improvement with 86.8% and 84.9% accuracy, respectively. GBM has an improved

88.5% accuracy due to its iterative boosting process. ANN, CNN, and LSTM are all improved with 90.9%, 92.7%, and 94.5% accuracy, respectively, as it supports the effectiveness of neural networks in the field of classification problems. The model given here is superior to all of these with 98.85% accuracy since the predicted positive instances are extremely accurate and less likely to generate false positives. Figure 10 illustrates the recall of the models, which measures the ability to correctly select true positive instances. Logistic Regression and SVM both have a score of 80.5% and 84.3%, respectively, which indicates their low ability to select correct instances. The proposed model again outperforms all the other techniques with a 98.80% recall rate, such that an extremely large majority of the true positive instances are accurately picked up with very few false negatives.

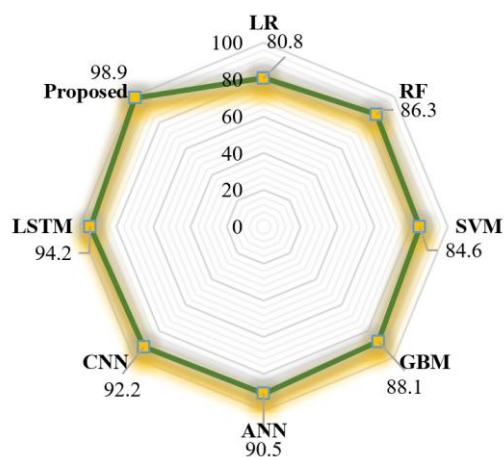


Figure 11. Comparison with existing techniques based on f1-score

Figure 11 shows the comparison of F1-score, which is equidistant between recall and precision. While Logistic Regression attains 80.8%, SVM and RF bring it up to 84.6% and 86.3%, respectively. GBM brings it further up to 88.1%, and ANN and CNN bring it to 90.5% and 92.2%, respectively. LSTM gets an even balance of 94.2%, given its strength in dealing with sequential dependencies. The suggested model attains a maximum 98.90% F1-score that guarantees it possesses the optimal ratio of precision and recall and ultimately confirms superiority over all of the current models in predicting students' performance.

Table 2. Computational time analysis

Methods	Computational time (s)
LR	2.3
RF	2.5
SVM	3.1
GBM	3.8
ANN	6.5
CNN	4.2
LSTM	4.1
Proposed	1.8

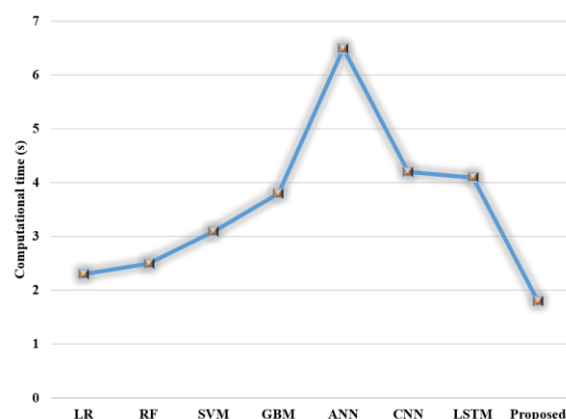


Figure 12. Comparison with existing techniques based on computational time

The comparative perspective of computational times of various machine learning and deep learning models used in the prediction of student performance is reflected in Table 2 and Figure 12, as it signifies the performance of various methods based on the execution time. Computational time becomes a decisive feature in predictive modeling, particularly with big data, real-time processing, or where there are limited resources. Results show that Logistic Regression (LR) possesses the second lowest computation time because 2.3 seconds due to its straightforward statistical model's efficient performance on structured data with linear separability. Nevertheless, its low complexity limits its predictability performance in a way that it will not be appropriately sufficient for non-linear pattern recognition in students' data. SVM takes higher computational time of 3.1 seconds since it is optimization-based training for finding the best hyperplane and thus is computationally costlier than LR and RF. GBM model with sequential prediction improvement through boosting techniques has a

further increase in computational time to 3.8 seconds due to sequential training overhead.

4. Conclusion

This paper presents a high-performance student performance prediction model based on an optimized feature selection algorithm and self-trained learning mechanism to enhance prediction performance, computational complexity, and explainability. The proposed work circumvents the deficiency of feature redundancy during feature addition, increased computational complexity, and weak adaptability of conventional models by virtue of the novel incorporation of Dingo-Coot Fusion Selector (DiCoot-FS) for best feature selection and Autoencoder and Belief Network-based Self-Taught Learning (AutoBel-STL) model for reliable prediction. The novelty of this paper lies on the DiCoot-FS feature selection that exploits the coot algorithm's exploratory attribute and the exploitative nature of the dingo optimization algorithm. The hybrid optimization ensures that the redundant features and irrelevant features are excluded while preserving the most critical scholarly and behavioral predictors of the students' achievement. In comparison with the conventional wrapper-based or filter-based selection methods, DiCoot-FS learns to adapt to different datasets in order to enhance the predictability of strong models. In addition, the AutoBel-STL model learns features based on a combination of belief network-driven probabilistic adaptation and autoencoder-based feature compression for self-learning. With both of them together, the model is able to enhance its learning process automatically, and so it is quite

robust to changes in student behavior with lesser reliance on annotated data.

Performance testing, tested on a student Kaggle dataset, demonstrates the new model at 98.90% accuracy, which is much higher compared to baseline classifiers such as Logistic Regression (82.5%), Random Forest (87.4%), SVM (85.6%), GBM (89.2%), ANN (91.8%), CNN (93.4%), and LSTM (95.1%). Apart from that, computational time comparison shows that the system here proposed only needs 1.8 seconds, which is the shortest among all the compared methods. All of these findings justify the efficiency of DiCoot-FS in feature optimization as well as AutoBel-STL's flexibility to further optimize prediction of student performance. Finally, this paper offers a new, hybrid, and self-tuning system that couples feature selection with deep learning in optimizing prediction of student performance. Using DiCoot-FS for precise feature refinement and AutoBel-STL for advanced probabilistic learning, the novel model enhances the predictive precision, reduces the computation complexity, and produces explanatory summaries of the accomplishment of the learner. These efforts pave the way for evidence-based decision-making within education so that schools can implement personalized learning interventions and enhance the percentage of students who are successful through the monitoring of early performance and the implementation of evidence-based measures of support

References

- Ahmad, A., Ray, S., Khan, M. T., & Nawaz, A. (2025). Student performance prediction with decision tree ensembles and feature selection techniques. *Journal of Information & Knowledge Management*, 2550016.
- Alalawi, K., Athauda, R., Chiong, R., & Renner, I. (2025). Evaluating the student performance prediction and action framework through a learning analytics intervention study. *Education and Information Technologies*, 30, 2887–2916.
- Alnasyan, B., Basher, M., & Allassafi, M. (2024). The power of deep learning techniques for predicting student performance in virtual learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, 100231.

- Arcinas, M. M., Meenakshi, M., Bahalkar, P. S., Bhaturkar, D., Lalar, S., Rane, K. P., et al. (2025). An efficient course recommendation system for higher education students using machine learning techniques. *Bulletin of Electrical Engineering and Informatics*, 14, 1468–1475.
- Balachandar, V., & Venkatesh, K. (2025). A multi-dimensional student performance prediction model (MSPP): An advanced framework for accurate academic classification and analysis. *MethodsX*, 14, 103148.
- Bisri, A., Heryatun, Y., & Navira, A. (2025). Educational data mining model using support vector machine for student academic performance evaluation. *Journal of Education and Learning (EduLearn)*, 19, 478–486.
- Chen, J., Zhou, X., Yao, J., & Tang, S.-K. (2025). Application of machine learning in higher education to predict students' performance, learning engagement and self-efficacy: A systematic literature review. *Asian Education and Development Studies*.
- Chen, W., Shen, Z., Pan, Y., Tan, K., & Wang, C. (2024). Applying machine learning algorithm to optimize personalized education recommendation system. *Journal of Theory and Practice of Engineering Science*, 4, 101–108.
- Domínguez, L. G. I., Robles-Gómez, A., & Pastor-Vargas, R. (2025). A data-driven approach to engineering instruction: Exploring learning styles, study habits, and machine learning. *IEEE Access*.
- Habeeb, R. J. H., Fadare, O. A., & Al-Turjman, F. (2025). Artificial neural networks for predicting student performance: A case study on student scores dataset. In *Smart infrastructures in the IoT era* (pp. 1165–1184). Springer.
- Ikegwu, A. C., Nweke, H. F., & Anikwe, C. V. (2024). Recent trends in computational intelligence for educational big data analysis. *Iran Journal of Computer Science*, 7, 103–129.
- Jiabin, L., Shanshan, Z., Xiaomei, L., Kaixin, Y., & Ying, Y. (2025). The benefits of physical literacy for human flourishing: A machine learning-based exploration of adolescents. *Applied Psychology: Health and Well-Being*, 17, e12635.
- Kennel, K., Ishimaru, S., Küchemann, S., Steinert, S., Kuhn, J., & Ruzika, S. (2025). Gaze-based prediction of students' math difficulties: A time dynamic machine learning approach to enable early individual assistance. *International Journal of Artificial Intelligence in Education*, 1–25.
- Kordbagheri, A., Kordbagheri, M., Tayim, N., Fakhrou, A., & Davoudi, M. (2025). Using advanced machine learning algorithms to predict academic major completion: A cross-sectional study. *Computers in Biology and Medicine*, 184, 109372.
- Kuadey, N. A., Ankora, C., Tahiru, F., Bensah, L., Agbesi, C. C. M., & Bolatimi, S. O. (2024). Using machine learning algorithms to examine the impact of technostress creators on student learning burnout and perceived academic performance. *International Journal of Information Technology*, 16, 2467–2482.
- Lee, H., Cho, J. K., Park, J., Lee, H., Fond, G., Boyer, L., et al. (2024). Machine learning-based prediction of suicidality in adolescents with allergic rhinitis: Derivation and validation in 2 independent nationwide cohorts. *Journal of Medical Internet Research*, 26, e51473.
- Narayanan, K. S., Kumaravel, A., & Gopalakrishnan, S. (2025). A war strategy based deep learning algorithm for students' academic performance prediction in education systems. *Journal of Engineering Education Transformations*, 38.
- Nayani, S., & R, S. P. (2025). Combination of deep learning models for student's performance prediction with a development of entropy weighted rough set feature mining. *Cybernetics and Systems*, 56, 170–212.
- Okoye, K., Nganji, J. T., Escamilla, J., & Hosseini, S. (2024). Machine learning model (RG-DMML) and ensemble algorithm for prediction of students' retention and graduation in education. *Computers and Education: Artificial Intelligence*, 6, 100205.
- Onker, V., Singh, K. K., Lamkuche, H. S., Kumar, S., Sharma, V. S., Chowdhary, C. L., et al. (2025). Harnessing machine learning for academic insight: A study of educational performance in Bhopal, India. *Education and Information Technologies*, 1–40.
- Pitta, S., Gopalakrishnan, S., & Chand, S. R. (2024). Secure cloud-WSNs with advanced intrusion detection system using Kepler optimization and Bi-ConGRU network approaches. In *2024 9th International Conference on Communication and Electronics Systems (ICCES)* (pp. 971–979).

- Pradeep, M., & Gopalakrishnan, S. (2024). Hybrid model for imbalance correction in intrusion detection systems using advanced optimization techniques and graph neural networks. In *2024 4th International Conference on Mobile Networks and Wireless Communications (ICMNWC)* (pp. 1–9).
- Rimal, Y. N., Sharma, N., Alsadoon, A., & Abbas, S. K. (2025). A comparative analysis of ensemble AutoML machine learning prediction accuracy of STEM student grade prediction: A multi-class classification prospective. *Multimedia Tools and Applications*, 1–27.
- Rizwan, S., Nee, C. K., & Garfan, S. (2025). Identifying the factors affecting student academic performance and engagement prediction in MOOC using deep learning: A systematic literature review. *IEEE Access*.
- Rogers, J. K., Mercado, T. C., & Cheng, R. (2025). Predicting student performance using Moodle data and machine learning with feature importance. *Indonesian Journal of Electrical Engineering and Computer Science*, 37, 223–231.
- Sahlaoui, H., Alaoui, E. A. A., Agoujil, S., & Nayyar, A. (2024). An empirical assessment of SMOTE variants techniques and interpretation methods in improving the accuracy and the interpretability of student performance models. *Education and Information Technologies*, 29, 5447–5483.
- Suryanarayana, K., Kandi, V. P., Pavani, G., Rao, A. S., Rout, S., & Krishna, T. S. R. (2024). Artificial intelligence enhanced digital learning for the sustainability of education management system. *The Journal of High Technology Management Research*, 35, 100495.
- Tjahyadi, H., & Tude, K. N. (2025). The implementation of educational data mining in predicting students' academic achievement in mathematics at a private elementary school. *International Journal of Information and Education Technology*, 15.
- Villar, A., & de Andrade, C. R. V. (2024). Supervised machine learning algorithms for predicting student dropout and academic success: A comparative study. *Discover Artificial Intelligence*, 4, 2.
- Yani, D., & Kusriani, K. (2025). Metode machine learning dan deep learning dalam prediksi kinerja siswa: Tinjauan sistematis. *G-Tech: Jurnal Teknologi Terapan*, 9, 530–539.
- Subramaniam, K. (2026). A Generative AI–Driven Vendor-Neutral Framework for Safe and Trustworthy Autonomous ERP Systems, *Academic International Journal of Engineering Science*: vol. 11, pp. 53-65.

6. Author Declarations

Conflict of Interest

The authors declare that they have no conflict of interest.

Competing Interests

The authors have no competing interests to declare relevant to this article's content.

Funding Details

No funding was received to assist with the preparation of this manuscript.

Data Availability

Data sharing not applicable to this article as no datasets were generated or analyzed during the current Study.