THE SCIENTIFIC TEMPER (2026) Vol. 17 (2): 5682-5700
E-ISSN: 2231-6396, ISSN: 0976-8653
Doi: 10.58414/SCIENTIFICTEMPER.2026.17.2.11
https://scientifictemper.com/
RESEARCH ARTICLE

Logistic Elitist Liquid Neural Network For Student Dropout Prediction

S V Arulvani[1*], Dr. C. Jayanthi[2]

Abstract

Student dropout prediction involves detecting students who are at risk of exiting their studies before completion. This process utilizes the data on academic performance, engagement, demographic factors, and other relevant indicators to predict possible dropouts, allowing organizations to execute timely interventions to improve maintenance rates.Different machine learning and deep learning methods are developed for the early identification of students at risk of dropping out has gained a lot of interest recently. But, the accurate dropout prediction with minimal error is a major concern. A novel Logistic regressive Elitist Optimized Liquid Neural Network (LOGEO-LNNet) is developed for categorizing the student dropout with high accuracy and minimal error rate. The proposed LOGEO-LNNet consists of four major steps namely data acquisition, preprocessing, feature selection and classification. To begin with, the LOGEO-LNNet model performs data acquisition by collecting the studentdata samples from a dataset. Followed by, data preprocessing is employed which includes missing data imputation and outlier deletion. Then the combinatorial Target projection matching based feature selection process is employed for identifying the morerequired features. With these selected features, the system proceeds to the classification process by employing Liquid Neural Networks to distinguish the more than two categories by employing Polytomous logistic regression. Fine-tuning the layers of liquid neural network is a vital step using the adaptive elitist shuffled shepherd metaheuristic algorithm thereby minimizing errors and increasing the accuracy of the student dropout prediction. Experimental evaluation of LOGEO-LNNet model is carried out with different factors such as accuracy, precision, sensitivity, F1 score, specificity and student dropout predictiontime, confusion matrix with respect to a number of student data.The experimental result reveal that the proposed LOGEO-LNNet model consistently achieves superior student dropout predictionaccuracy performance, exhibiting lower error rates and reduced prediction time compared to existing deep learning methods.

Keywords: Student dropout prediction, Combinatorial target projection matching based feature selection, Liquid neural networks, Fine-tuning, Adaptive elitist shuffled shepherd metaheuristic algorithm

[1]Ph.D Research Scholar(Full Time), PG & Research Department of Computer Science, Government Arts College (Autonomous), Karur-5, Affiliated to Bharathidasan University,Tiruchirappalli, Tamil Nadu, India

[2]Associate Professor, PG & Research Department of Computer Science, Government Arts College (Autonomous), Karur-5, Affiliated to Bharathidasan University, Tiruchirappalli, Tamil Nadu, India

***Corresponding Author:** S V Arulvani, Ph.D Research Scholar(Full Time), PG & Research Department of Computer Science, Government Arts College (Autonomous), Karur-5, Affiliated to Bharathidasan University,Tiruchirappalli, Tamil Nadu, India, E-Mail: arulphdcs@gmail.com

How to cite this article: Arulvani, S.V., Jayanthi, C. (2026). Logistic Elitist Liquid Neural Network For Student Dropout Prediction. The Scientific Temper, **17**(2):5682-5700.

Doi: 10.58414/SCIENTIFICTEMPER.2026.17.2.11

Source of support: Nil

Conflict of interest: None.

Introduction

Student dropout is a most significant challenge facing higher education organizations worldwide for both individual students and institutional efficiency. Dropout not only reflects a student's academic performance but also affects a range of skills, attributes, and competencies,courses across various disciplines. By integrating the advanced data analytics and deep learning techniques, student's dropout prediction process is employed by considering the various academic, behavioral, and demographic factors.Explainable Relational Graph Convolutional Network (ERGCN) model was developed in [1] for student dropout prediction with better high accuracy and transparency based on multidimensional relationship between students, courses, and behaviors. However, the storage and computational efficiency of the model did not improve as the number of students continued to grow. Multilayer Perceptron (MLP) model was introduced in [2] to predict student suspension and dropout with higher accuracy performance. However, the model did not explore

© The Scientific Temper. 2026

Received: 20/01/2026 **Accepted:** 15/02/2026 **Published:** 25/02/2026

feature augmentation or cost-sensitive learning to further improve class separability and prediction precision.

A CatBoost algorithm was designed in [3] to detect the dropout and failure prediction based on multi-objective hyperparameter optimization. However, timely interventions in student dropout prediction remained major issue. Machine learning algorithms were designed in [4] for early prediction model ofpotential dropout students by means of demographic data with a learning analytics approach. However, it failed to explore the incorporation of qualitative data and behavioral patterns to improve the model's interpretability and address emerging risk factors. Educational Data Mining(EDM) algorithms were developed in [5] for predicting the student dropout based on factors influencing academic performance. However, efficient deep learning model was not employed to enhance the performance of student dropout prediction. Preprocessed Kernel Inducing Points data distillation technique (PP-KIPDD) was introduced in [6] for student dropout prediction based on selected significant features. However, different samples were not considered to identify more key features for student dropout predictions.

Decision tree and logistic regression model were developed in [7] to predict the student dropout and accurately predict at risk student. However, the designed model displays high computational complexity, particularly when handling large-scale datasets. XGBoost classifier model was introduced in [8] to decrease dropout rates and improve student retention. However, deep learning and machine learning models were not employed for predicting the student dropout. Random Forest algorithm was designed in [9] to classify students at risk of dropping out by analyzing various behavioral data and facilitating timely interventions to support student maintenance and success in academic setting.However, it did not focus on enhancing the system's predictive capabilities by incorporating additional data to improve the effectiveness of dropout prevention measures. Dilated Convolutional Attention Network was developed in [10] for Dropout Prediction based on Lie Group Features. However, it failed to enhance the model's relational interpretation capabilities and scalability while simultaneously further enhancing the accuracy and reducing computational complexity.

Neural Network based model was developed in [11] for student dropout prediction. But attention mechanisms or gradient-based interpretability techniques were not employed for accurate student dropout prediction. Bayesian modelling of the Weibull distribution model was introduced in [12] to predict student dropout rates for analyzing university dropout rates. However, it failed to explore the factors contributing to dropout in better intensity.Machine learning models were developed in [13] for predicting dropout within the higher education system with a reduced set of features. However, it failed to enhance the effectiveness of the Machine learning models. K-Means and C4.5 algorithms were developed in [14] identify students at high risk of dropping outbased on academic performance. However, the model did not efficiently reducethe issue of computational complexity. Artificial neural networks (ANNs) were developed in [15] for dropout prediction within rural higher education based on socioeconomic, academic performance. However, the complexity of dropout intentions among rural students was not minimized in virtual higher education.

Major Contributions

The key contributions of the LOGEO-LNNet model are outlined as follows:

- To design a novel model called LOGEO-LNNet for accurate student dropout prediction by implementing four different processes namely data pre-processing, feature selection, classification and fine tuning.
- To minimize the accurate student dropout prediction, LOGEO-LNNet model is developed which includes data pre-processing and relevant feature selection based on combinatorial Target projection matching. This process reduces the time consumption of student dropout prediction.
- To improve the accuracy of student dropout prediction, a Liquid Neural Network deep learning model is employed to analyze data samples using a logistic regression function. Moreover, the adaptive elitist shuffled shepherd metaheuristic algorithm is applied for hyperparameter fine-tuning in Liquid Neural Network to further minimize the error rate and increases the sensitivity and specificity.
- The comparative analysis of the proposed LOGEO-LNNet model has been evaluated based on conventional deep learning methods with different performance metrics.

Related works

Blending ensemble learning (BEL) was introduced in [16] to improve the performance of the student dropout prediction from online courses. However, more efficient and effective approach was not developed to reduce the data complexity.Stock-and-flow model were developed in [17] for dropout prediction from the higher education students. However, deep learning model was not employed for error handling capabilities of the classifier model. Machine Learning (ML) techniques were developed in [18] to predict early dropout of students based on demographic, academic, and behavioral factors to identify students at risk. However, dimensionality reduction techniques were not implemented to address the high dimensionality and improve model efficiency. Explainable AI (XAI) techniques were developed in [19] for improving dropout factors and most influential attributes. However, the technique incurred high computational costs in student dropout prediction. Binary classification models were developed in [20] for student dropout predictionto obtain the higher F1 score and accuracy. However developing robust, precise and efficient student dropout prediction has become a critical issue.

Cluster-based machine learning models were developed in [21] forpredicting dropout risk based onfinal course outcomes to detect at-riskstudents. However, it did not create a more adaptable, robust, framework for accurate predictive challenges. Different machine learning techniques were introduced in [22] for accurate dropout prediction based on internal academic criteria such as completed courses, internship report status, and thesisproposal. However, advanced algorithms suchas deep learning model were developed to reduce the misclassification in accurate dropout prediction.Random Forest classifier model was developed in [23]for predicting the university dropouts. However, it did not integrating deep learning techniques to further improve accuracy and model performance. Multinomial logistic regression analysis was introduced in [24] integrated with finite mixture models for predicting dropout from higher education. However, the time complexity of the dropout prediction was not reduced. Stacked ensemble machine learning model was developed in [25] for predicting student dropout risk from the online learning courses. However, hyperparameter adjustment was not employed to achieve high accuracy.

Cat Boost Classifier (CATC) model with fine-tuning model was developed in [26]forpredicting the student graduation and dropout patterns. However, it failed to enhance the quality within education at a larger scale. Ensemble-SHAP Explainable Student Prediction framework was introduced in [27] to predictstudent dropout prediction. However, Deep learning approaches were not employed for automatic feature extraction and complex pattern recognition ineducational data. Clustering-based approach was introduced in [28] for classifying students based on their academic performance based on learning management systems. However, more robust predictive model was not employed. An integration of metaheuristic optimization techniques with ensemble technique was developed in [29] for student dropout identification to enhance the accuracy and robustness during identifying high risk students. However, scalability and efficiency of the approach were not improved. Random forest (RF) and feature tokenizer transformer (FTT) model were developed in [30] for academic dropout prediction. However, the model did not effectively minimize the misclassification rate.

Proposal Methodology

A student dropout refers to the event where students terminate their education before finishing their course, which occurs at any educational stage from high school to higher education. This performance has important allegations for both individual students as well as educational organizations. This behavior entails a complete approach that includes various aspects, including academic performance, socio-economic condition, institutional characteristics, psychological factors, and so on.High dropout rates are major issues in the educational domain, disturbing a wide variety of courses across various regulations. These rates varied severely depending on the region,institution, and specific course type. Therefore student dropout is a notable social issue and general implications for individuals and society. However, identifying students at high risk of dropping out is a foremost challenge for sustainable education.To deal with this need, a novel LOGEO-LNNet model is introduced specifically designed to improve the accuracy of student dropout predictions. The overall processing diagram includes four main stages namely data acquisition, preprocessing, significant feature collection, and classification. These different phases work together to efficiently distinguish the students into multiple classes.

Figure 1 illustrates the overall structural design of the proposed LOGEO-LNNet model for student dropout predictions. The above construction comprises of four important processes namely Data acquisition, pre-processing, feature selection, and classificationto enhance the accuracy of dropout predictions. In the beginning, student data samples aregathered from the dataset for accurate predictions. The most important step of proposed model entails data pre-processing or cleaning to change unprocessed dataset into more appropriate structure by the means of missing data imputation and outlierdata deletion.Followed by, the important feature selectionis carried out by utilizing the function for accurate student dropout prediction with minimal time consumption. As a final point, the proposed model executes the dropout prediction through classification attaining higher accuracy with smaller errors. These dissimilar primary processes of the proposed LOGEO-LNNet model are explained in the succeeding sectors.

Data acquisition

The foremost process of proposed LOGEO-LNNet model is the data acquisition which entails collecting the raw data samples from the dataset. In this section, student-related data is collected from Predict Students' Dropout and Academic Success dataset [https://www.kaggle.com/datasets/mattop/predict-students-dropout-and-academic-success].The collected dataset used in this analysis was collected from a higher education organization. It encloses records of students registered in different undergraduate programs, including agronomy, design, education, nursing, journalism, management, social services, and technology. The dataset comprises information available at the time of enrollment, such as academic background, demographic

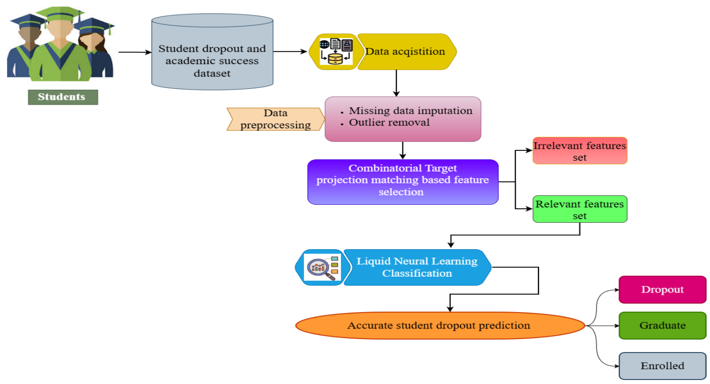


Figure 1: Architecture diagram of proposed LOGEO-LNNet model

characteristics, and socio-economic status. Moreover, it also incorporates academic performance outcomes from the first and second semesters. This dataset forms the basis for developing classification models to predict student dropout risk and academic achievement. Overall, the dataset comprises 4,424 student records with 37 attributes. Table 1 reviews the academic attributes used to predict whether a student dropout or not.

Data preprocessing

The second vital step of proposed LOGEO-LNNet model is the data preprocessing helps to generate a cleaned and reliable dataset for accurate students' dropout prediction. Since raw agriculture data collected from outer repositories comprises a noise, missing data, and irregular samples. This irregularity of the dataset directs to high prediction error. Therefore, the proposed model executes data preprocessing step before applying any machine learning or deep learning model. The preprocessing stage of proposed model primarily comprises of two key processes namely missing data imputation and outlier data deletion.

Let us consider the Students' Dropout and Academic Success dataset 'DS' and data samples 'AS' as well as features $\{f_1, f_2, \ldots, f_m\}$ are arranged in the form of matrix. Therefore, the input matrix is formulated as given below,

Table 1: Attribute information

S. No	*Features*	*S. No*	*Features*
1	Marital status	20	Age at enrollment
2	Application mode	21	International
3	Application order	22	Curricular units 1st sem (credited)
4	Course	23	Curricular units 1st sem (enrolled)
5	Daytime/evening attendance	24	Curricular units 1st sem (evaluations)
6	Previous qualification	25	Curricular units 1st sem (approved)
7	Previous qualification (grade)	26	Curricular units 1st sem (grade)
8	Nacionality	27	Curricular units 1st sem (without evaluations)
9	Mother's qualification	28	Curricular units 2nd sem (credited)
10	Father's qualification	29	Curricular units 2nd sem (enrolled)
11	Mother's occupation	30	Curricular units 2nd sem (evaluations)
12	Father's occupation	31	Curricular units 2nd sem (approved)
13	Admission grade	32	Curricular units 2nd sem (grade)
14	Displaced	33	Curricular units 2nd sem (without evaluations)
15	Educational special needs	34	Unemployment rate
16	Debtor	35	Inflation rate
17	Tuition fees up to date	36	GDP
18	Gender	37	Target—(dropout, graduate, enrolled)
19	Scholarship holder		

$$IM = \begin{bmatrix} A_1 & A_2 & \dots & A_m \\ Ds_{11} & Ds_{12} & \dots & Ds_{1n} \\ Ds_{21} & Ds_{22} & \dots & Ds_{2n} \\ \vdots & \vdots & \dots & \vdots \\ Ds_{m1} & Ds_{m2} & \dots & Ds_{mn} \end{bmatrix} \quad n = rows, m = column \tag{1}$$

Where, IM symbolizes an input data matrix where each column indicates a number of features $A_1, A_2, \dots A_m$, each row represents a number of records or data samples or records' $Ds = \{Ds_1, Ds_2, \dots, Ds_n\}$ ' respectively. First, missing values (often represented as none, or blanks are located in given input dataset) are determined. After detecting the missing values, the polynomial interpolation function is employed for determining the missing values with the dataset based on other known existing data samples. The interpolation function is expressed as follows,

$$P(Ds) = \vartheta_0 Ds^0 + \vartheta_1 Ds^1 + \vartheta_2 Ds^2 + \dots + \vartheta_n Ds^n \tag{2}$$

Where, $P(Ds)$ specifies a missing data samples, $\vartheta_0, \vartheta_1, \vartheta_2 \dots \vartheta_n$ indicates a regression coefficient, Ds indicates a number of available data samples. In this way, all the missing data are handled within the dataset.

After handling the missing data, the outlier's data is identified and removal from the dataset for accurate student dropout predictionby using Peirce statistical decisive factor. It is a method used to find out the two or more outliers within the dataset. This factor is calculated based on the absolute variation between the data samples and their mean valuewhich is greater than the product of the highest acceptable deviation.

$$Q = \sum_{i=1}^{n} |Ds_i - \mu| \tag{3}$$

$$\mu = \frac{1}{n} \sum_{i=1}^{n} Ds_i \tag{4}$$

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Ds_i - \mu)^2} \tag{5}$$

$$Z = \begin{cases} Q > \sigma_{max} & ; Outlier \\ Otherwise & ; no\,outlier \end{cases} \tag{6}$$

Where, Z denotes the Peirce statistical decisive factor outcome, μ denotes a mean (or average) of the dataset, σ_{max} defines the maximum allowable deviation from the mean, σ denotes a standard deviation of the dataset. If the determined ' Q ' value is larger than ' σ_{max} ', then the data is identified as outlier. Otherwise the data is considered as a normal. In this way, all the missing data are determined and removed. The processing algorithm is given below,

// Algorithm 1: Data pre-processing
Input: dataset ' D ', features $A_1, A_2, \dots A_m$, data samples $Ds = \{Ds_1, Ds_2, \dots, Ds_n\}$
Output: Pre-processed dataset
Begin 1. For each dataset ' D ' with features ' A ' do 2. Formulate input vector matrix ' IM ' using (1) 3. If any missing value in ' IM ' then 4. Apply polynomial interpolation using (2) 5. Load the missing value to the particular cell 6. End if 7. For each data samples do 8. Compute the deviation between the mean and data using (3) 9. if ($Q > \sigma_{max}$) then 10. Data samples is identified as outlier 11. else 12. Data samples is identified as normal 13. End if 14. Remove outlier data samples from the dataset 15. Return (preprocessed dataset) 16. End for 17. End for End

Algorithm 1 given above explains the step by step process of dataset preprocessing for accurate student dropout prediction. To begin with, a number of data samples are collected from the dataset and create the input matrix. Afterward, proposed model perform the missing data handling applying polynomialinterpolation function. After imputing the missing data, the outlier data is identified and removed. First, the variance between the mean and the data samples are calculated. If the estimated deviation is greater than maximum allowable difference, then the input data sample is considered as anoutlier. Otherwise, it is considered as an outlier data. Therefore, the preprocessed results are obtained for minimizing the student dropout prediction time.

Combinatorial Target projection matching based feature selection

After the data pre-processing, the proposed LOGEO-LNNet model performs the significant feature selection to reduce the dimensionality of the original dataset. This is done by applying a Combinatorial Target projection matching which is the process of choosing the subset of the relevant features from the dataset aiming to reduce the time consumption of student dropout prediction. The Combinatorial Target projection matching is a machine learning method for visualizing high-dimensional features where the relevant features are identified by Sokal–Michener's correlation matching.

The Sokal–Michener's correlation matching is a statistical analysis which employed to analyze the features and target variables based on the correlation measure. The correlation matching process is expressed as follows

$$H = CORR\left(T, A_j\right) \tag{7}$$

$$CORR\,(T, A_j) = 1 - \sum_{j=1}^{m} \frac{|T\,\Delta\,A_j|}{m} \tag{8}$$

Where, H indicates an analysis outcomes, $CORR\left(T, A_j\right)$ designates a correlation between the target samples 'T' and features 'A_j',' m' indicates a number of features, $T\,\Delta\,A_j$ denotes a variation between the features and target samples. Based on the Sokal–Michener's simple matching method, the correlation offers the similarity value returns from '0' to '1.

Subsequently, the target projection function is employed to project the features from high dimensional space into low dimensional.

$$K = arg\min\left|\delta - CORR\left(T, A_j\right).\varphi\right| \tag{9}$$

Where, K denotes an output of target projection function, δ indicates a threshold, $CORR\left(T, A_j\right)$ represents the correlation function, φ represents a projection matrix that transforms the more features into the target space, $\arg min$ denotes a argument of minimum function. The correlated features closer to the threshold are selected and removed the remaining features, thus in turn minimize the student dropout prediction.

Algorithm 2 given above describes the various steps involved in feature selection using Combinatorial Target projection matching aimed to improve student dropout prediction while minimizing time consumption. First, the preprocessed dataset is taken as input for this analysis. Then,correlations between the target and features samples are calculated. The target projection function analyzes the correlation values with the threshold function. Based on analyzes, the relevant and irrelevant features are distinguished with higher accuracy. Finally, the more significant relevant features are chosen for accurate student dropout prediction in educational institutions applications.

Algorithm 2: Combinatorial Target projection matching

Input: preprocessed Datasets 'D', features $A_1, A_2, \ldots A_m$, data samples $Ds = \{Ds_1, Ds_2, \ldots, Ds_n\}$

Begin
1. Collect the preprocessed dataset 'D' as input
2. For each feature 'A_j'
3. Measure the correlation function using (8)
4. Apply target projection function using (9)
5. if ($arg\min\left|\delta - CORR\left(T, A_j\right).\varphi\right|$) then
6. Project relevant feature into low dimensional space
7. else
8. Project irrelevant features
9. End if
10. Return the relevant features vector
11. end for

End

Logistic regressive Elitist stochastic optimized Liquid Neural Networks based classification

The last process of the proposed LOGEO-LNNet model involves the classification with the selected relevant features for predicting the student dropout. In this context, the Liquid Neural Network (LNN) is a class of neural networks that display dynamic manners and have the ability to process and learn thecontinuous streams of data efficiently than the traditional neural networks. Itis types of recurrent neural networks (RNNs) and utilizes a dynamic network of neurons to process sequential data. The structural design of this LNN model is illustrated in Figure 2.

Figure 2 demonstrates the schematic structure of a liquid neural network for accurate student dropout prediction. The structure consists of three key layers such as an input layer, a liquid or reservoir layers, and an output layer. Each layer consists of individual computational unitcalled as spiking neurons or nodes, which are devised to track the dynamic and adaptive performance of biological neural organizations. Each neuron in the layer receives and processes the given input data samples and transmits the results to the successive layers.

As shown in figure 2, input layer provides as the starting point for raw input data samples with the selected relevant features.This input data samples is then transferred into the liquid (or reservoir) layer by the means of spiking neurons. In that layer, computation process is executed by employing the logistic regression analysis.Each neuron in liquid layer measured the weighted sum, enabling the network to identify the neuron activation for processing the given input data samples.

$$W = \sum_{i=1}^{n} Ds_i * \alpha_{IR} + B_{IR} \tag{10}$$

Where, W denotes a weighted sum output, α_{IR} represent synaptic weights between the input and reservoir layer, Ds indicates a number of data samples, 'B_{ir}' indicates a bias.Within the reservoir layer, spiking neurons construct the liquid behaviour of the network in dynamic neural connections. This layer processes the given input samples by simultaneously exploring multiple interpretations of the input.

The polytomous logistic regression model is employed in reservoir layer to predict a multiclass outcome (i.e. more

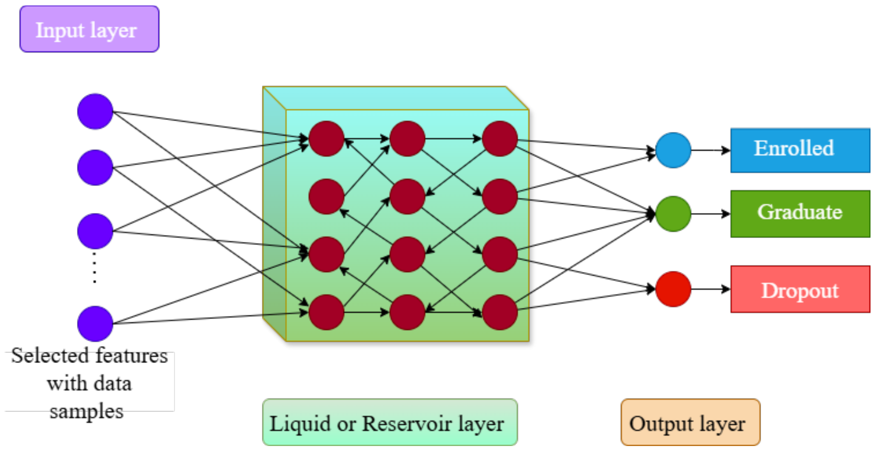


Figure 2: Schematic structure of Liquid Neural Network

than one output)such as dropout, enrolled and graduate. Polytomous logistic regression model is a supervised machine learning algorithm employed to model the relationship between a set of independent variables (selected features with data samples) and a dependent variable (target outcome). In this context, the model includes numerous input features cooperatively contribute to predicting the student dropout.Therefore, the relationship between the dependent and independent variables is expressed as follows,

$$LR = \frac{1}{1 + \exp - \left(c_0 + c_1.A_1 + c_2.A_2 + \ldots + c_k.A_k \right)} \quad (11)$$

Where, LR indicates a polytomous logistic regression outcomes, $c_0, c_1, c_2 \ldots . c_k$ represents a coefficients for the respective features, $A_1, A_2 \ldots . A_k$ represents a features. The outcomes provide outcomes from 0 to 1.For a given input, it computes a score for each of the three classes. The class with the highest regressionoutcomes is chosen as the final prediction. The state of each spiking neuron in the reservoir layer at time't' is formulated as follows,

$$R(t) = f\left(\sum_{i=1}^{n} Ds_i * \alpha_{IR} + \alpha_R R_{t-1} + B_R \right) \quad (12)$$

Where $R(t)$ symbolizes a state of each neuron in the reservoir layer at time 't', f symbolizes a the liquid activation function which measures the output of each neuron in the reservoir layer by employing a weighted sum of the input samples 'Ds_i' and the previous state R_{t-1}, α_{IR} indicates a weights between input and reservoir layer, α_R designates a recurrent weights within the liquid layer, R_{t-1} symbolizes a previous liquid layer output, B_R symbolizes a bias term in reservoir layer. This dynamic system permits the network to efficiently process the sequential data samples and preserve the classification output. Based on the results, the error occurred during the multiclass classificationis expressed as follows,

$$CR = \left[\frac{Number\ of\ misclassified\ data\ samples}{Total\ number\ of\ data\ samples} \right] \quad (13)$$

Where CR represents the classification error.To reduce the classification error, fine tuning process is employed by adjusting hyperparameter(i.e. weights) between the layers to increase the accuracy of student dropout prediction through Nesterov Momentum method.

$$\alpha_{t+1} = \alpha_t - \eta\ F_t \quad (14)$$

$$F_t = \tau F_{t-1} + (1 - \tau) \frac{\partial CR}{\partial \alpha_t} \quad (15)$$

Where, α_{t+1} represents an updated weight, α_t denotes a current weight, η specifies a learning rate$(\eta < 1)$, '$\frac{\partial CR}{\partial \alpha_t}$' specifies a partial derivative of classification error 'CR' with respect to weight 'α_t', τ represents default value 0.9. From the updating process, various counts of weights are obtained. Among the multiple weights, optimal one is chosen by applying anadaptive elitist shuffled shepherd metaheuristic algorithm for reducing the error and increasing the accuracy of student dropout prediction.

The proposed optimization algorithm is inspired by shepherd behavior, particularly the process of guiding, controlling and directing the scattered into optimal structures. The main aim of this algorithm is sheep moved through the search space influenced by fining the best sheep in their herd.In this optimization algorithm, sheep's are related to the multiple weights updated from the Nesterov Momentum method. To begin with the algorithm, population of weights (i.e. sheep) is randomly initialized within the search space.

$$\alpha_k = \{\alpha_1, \alpha_2, \ldots, \alpha_k\} \quad (16)$$

After initializing the population of weights, the fitness function is computed. It is a mathematical expression that estimates the quality of a solution in terms of the main objective.

$$\omega = \arg\min CR \tag{17}$$

Where, ω indicates a fitness function, CR indicates classification error rate. After computing the fitness,

As a result, the weights with better fitness are selected as current best than the others in the total populations. After finding the current best, total population is partitioned into number of groups called as herd. For each herd, the best sheep is assigned to direct others for searching grazing area. Therefore, the sheep's position in the particular herd is updated by the following expression.

$$P_i^{t+1} = P_i^t + q.0.5\left|P_{best}{}^t - P_i^t\right| + d.r\left(0,1\right) \tag{18}$$

Where, P_i^{t+1} specifies a updated position of the sheep's within the specific herd, P_i^t indicates a current position of the sheep's, q,d indicates a parameters that control exploration and exploitation, $\left|P_{best}{}^t - P_i^t\right|$ denotes a divergence between the best solution and current solution, $r\left(0,1\right)$ indicates a random number between 0 and 1, $P_{best}{}^t$ represents a current best solution in particular herd which has better fitness.

After that, the shuffling process is executed by selecting the sheep's from the multiple herds randomly as follows,

$$S = rand\left(\alpha_k\right) \in H \tag{19}$$

Where, α denotes a set of sheep'srandomly selected from herd H. After that, multiple herds are combined and identify the best weight among the othersby applying Elitist approach.

$$E = \begin{cases} \omega(\alpha_k) > \omega(\alpha_r) & select\,best\,weight \\ Otherwise\;\; ; & Remove\,the\,others \end{cases} \tag{20}$$

Where, E indicates an Elitist selection results, $\omega(\alpha_k)$ denotes a fitness of one weight $\omega(\alpha_r)$ denotes a fitness of another weight.Following the merging step, the global best position is determined based on their fitness scores. If the fitness value of a newly updated position higher than the previous one, it replaces the new one. This process is continued until the predefined iterations get reached. Finally, the most optimal solution weight is chosen to refining the reservoir layer for minimizing the error rate. Finally, the newly accurate multiclass classification results are obtained at output layer.

$$Y = f_{softmax}\;(\alpha_{RO}\,R_t) \tag{21}$$

Where, Y represents a final classification output generated at output layer, R_t represents a reservoir layer output, α_{Ro} indicates the weights between reservoirand output layer, a softmax activation function '$f_{softmax}$' in the output layer for multi class output. In this way, strident dropout prediction is achieved at the output layer.

Algorithm 3 outlines a complete procedure to improve the accuracy of student dropout prediction by utilizing a Logistic regressive Elitist stochastic optimized Liquid Neural Networks. The approach initiates by applying the selected relevant features with their data samples into the input layer of the LNN. These features are then sent to the liquid or reservoir layer, where logistic regression model is applied for deeply analysis the relationship between the features and obtain multiclass output. Once multi class classification is completed, the error rate is measured. To improve the accuracy by minimizingthe classification error, the adaptive elitist shuffled shepherd metaheuristic algorithm is employed to discover optimal weight values using a fitness-based assessment. This optimization algorithm then estimatesdifferent behavioral factors and updates the weights consequently. This process is continued until the predefined maximum number of iterations getsachieved. Finally, an optimal weight is determined to minimize classification errors and improve the overall accuracy of student dropout prediction.

Experimental Setup

In this subsection, the proposed LOGEO-LNNet model and two existing methods, ERGCN[1] and MLP [2], are implemented using Python high level programming language.In order to conduct the experiment, the Students' Dropout and Academic Success dataset is taken from the kaggle repositoryhttps://www.kaggle.com/datasets/mattop/predict-students-dropout-and-academic-success. The primary objective of this dataset is to recognize patterns related to student dropout and academic success in a higher education institution. It includes data on students registered in various undergraduate programs, such as agronomy, design, education, nursing, journalism, management, social services, and technology. Overall, the dataset consists of 4,424 student instances described by 37 attributes. Table 1 presents the academic-related features that are used to categorize students into three outcomes namely graduates, dropouts, and enrolled students.

Implementation results

The LOGEO-LNNet model is broadly examined to compute its efficiency in students' dropout prediction. The assessment comprises numerous stages, including data acquisition, preprocessing, feature selection, and classification, utilizing the Students' Dropout and Academic Success dataset taken from Kaggle repository.

First, the 4424 data sample are collected from the dataset are shown in figure 3 for evaluating the performance of LOGEO-LNNet model.

Figure 4 describes the class distribution dataset. The class distribution of 4,424 instances labeled data samples,

Algorithm 3: Logistic regressive Elitist stochastic optimized Liquid Neural Networks

Input: Number of selected features '$A_1, A_2, \ldots A_k$', data samples $Ds = \{Ds_1, Ds_2, \ldots, Ds_n\}$

Output: Enhance accuracy of student dropout prediction

Begin
1: Collect selected features '$A_1, A_2, \ldots A_k$' at input layer
2: For each A_k do ----- liquid layer
3:Apply regression analysis using (11)
4. End for
5. if $(LR = 1)$ then
6. Data samplesis classified into particular class
7. End if
8. Obtain the recurrent reservoir layer output using (12)
9. For each classification result
10. Compute the classification error using (13)
11. Update the weights using (14) and (15)
12. End for
13. Initialize the population of the weights $\alpha_1, \alpha_2, \ldots, \alpha_k$
14. While (t< t_{max}) do
15. foreach weight 'α_k'
16. Compute the fitness 'ω' using (17)
17. End for
18. Sort the weights based on fitness
19. Select the current best
20. Partitions weights into multiple herds
21. For each weight α_k in herd 'H'
22. Update the position using (18)
23. End for
24. Perform shuffling process using (19)
25. For each weight α_k in merged herd 'H
26. Apply Elitist function toverify the fitness using (20)
27.End for
28. if $(\omega(\alpha_k) > \omega(\alpha_r))$ then\
29. Replace α_k is a global best
30. End if
31. t= t+1
32. Go to step 14
33. End while
34. Return (optimal weight)
35. Apply softmax activation function using (21) --- output layer
36. Return (multiclass classification output)
End

which are distributed among 3 different classes. This class distribution guarantees that the model receives anillustration of data samples.

The chart illustrates the distribution of students across three academic outcomes namely graduate, dropout, and enrolled.On the x-axis, the 3 class labels are represented. The y-axis indicates the count of samples for each class. Graduates form the largest group, indicating that a majority of students successfully complete their studies. Dropouts represent a considerable portion of the population, highlighting that a significant number of students leave before finishing their programs. The enrolled group is the minimum, signifyingless students are currently in progress.

After data acquisition, preprocessing stageisperformed, including handling missing values and removing outliers. Initially, the total dataset size is 501KB, which was reduced to 463KBafter preprocessing. The preprocessing results are shown in Figure 5.

The LOGEO-LNNet model focuses on identifying more significant features that enhance the accuracy of student dropout prediction. Heatmap similarities between the features are demonstrated in Figure 6.

Figure 7 shows selected 15 features by using Target projection matching. To achieve this, it combinatorial Target projection matching techniques to select 15 key features which has similarity score close to '1' from the 37 features. Finally, the classification is carried out using Logistic regressive Elitist stochastic optimized Liquid Neural Networks. Figure 8 portrays the classification results.

In Figure 8, Logistic regressive Elitist stochastic optimized Liquid Neural Networks is trained using 15 selected relevant features from data samples for achieving classification results to classify the data samples into graduate, dropout and enrolled, thereby minimizing the classification error.

Performance Comparison Analysis

In this section, performance assessment between the proposed LOGEO-LNNet model and two existing approaches, ERGCN[1] and MLP [2]are analyzed using numerous metrics, including student dropout prediction accuracy, precision, sensitivity, F1 score, Specificity, Confusion matrix and student dropout prediction time.

Performance Assessment of Student Dropout Prediction Accuracy

It represents as the proportion of correctly predicted dropout data samplesfrom the total number of input student data samples. The formula for calculating the prediction accuracy is mathematically expressed as follows:

$$SDPA = \frac{TP + TN}{TP + TN + FP + FN} * 100 \quad (22)$$

Where, '$SDPA$ denotes student dropout prediction accuracy, 'TP' represents atrue positive, TN denotes a true negative,'FP' indicates a false positive, FN indicates a false negative. It is measured in terms of percentage (%). Table 2 shows the validation results of student dropout prediction accuracy for proposed LOGEO-LNNet and existing ERGCN [1] and MLP [2] respectively.

Student Dropout and Academic Success Data Acquisition

```
Marital status,Application mode,Application order,Course,Daytime/evening attendance,Previous qualification,Previous qualification (grade),Naciona
1,17,5,171,1,1,122.0,1,19,12,5,9,127.3,1,0,0,1,1,0,20,0,0,0,0,0,0.0,0,0,0,0,0,0.0,0,10.8,1.4,1.74,Dropout
1,15,1,9254,1,1,160.0,1,1,3,3,3,142.5,1,0,0,0,1,0,19,0,0,6,6,6,14.0,0,0,6,6,6,13.66666667,0,13.9,-0.3,0.79,Graduate
1,1,5,9070,1,1,122.0,1,37,37,9,9,124.8,1,0,0,0,1,0,19,0,0,6,0,0,0.0,0,0,6,0,0,0,0.0,0,10.8,1.4,1.74,Dropout
1,17,2,9773,1,1,122.0,1,38,37,5,3,119.6,1,0,0,1,0,0,20,0,0,6,8,6,13.42857143,0,0,6,10,5,12.4,0,9.4,-0.8,-3.12,Graduate
2,39,1,8014,0,1,100.0,1,37,38,9,9,141.5,0,0,0,1,0,0,45,0,0,6,9,5,12.33333333,0,0,6,6,6,13.0,0,13.9,-0.3,0.79,Graduate
2,39,1,9991,0,19,133.1,1,37,37,9,7,114.8,0,0,1,1,1,0,50,0,0,5,10,5,11.85714286,0,0,5,17,5,11.5,5,16.2,0.3,-0.92,Graduate
1,1,1,9500,1,1,142.0,1,19,38,7,10,128.4,1,0,0,1,0,1,18,0,0,7,9,7,13.3,0,0,8,8,8,14.345,0,15.5,2.8,-4.06,Graduate
1,18,4,9254,1,1,119.0,1,37,37,9,9,113.1,1,0,0,0,1,0,22,0,0,5,5,0,0.0,0,0,5,5,0,0.0,0,15.5,2.8,-4.06,Dropout
1,1,3,9238,1,1,137.0,62,1,1,9,9,129.3,0,0,0,1,0,1,21,1,0,6,8,6,13.875,0,0,6,7,6,14.14285714,0,16.2,0.3,-0.92,Graduate
1,1,1,9238,1,1,138.0,1,1,19,4,7,123.0,1,0,1,0,0,0,18,0,0,6,9,5,11.4,0,0,6,14,2,13.5,0,8.9,1.4,3.51,Dropout
1,1,1,9670,1,1,139.0,1,38,19,5,7,130.6,1,0,0,1,0,0,18,0,0,6,6,6,12.33333333,0,0,6,7,5,14.2,0,13.9,-0.3,0.79,Graduate
1,1,1,9500,1,1,136.0,1,19,38,9,9,119.3,1,0,0,1,0,1,18,0,0,8,8,7,13.21428571,0,0,8,8,7,13.21428571,0,12.7,3.7,-1.7,Graduate
1,1,2,9853,1,1,133.0,1,19,37,4,9,130.2,1,0,0,1,0,0,19,0,0,6,6,0,0.0,0,0,6,0,0,0.0,0,12.7,3.7,-1.7,Dropout
1,53,1,9254,1,42,110.0,1,1,1,4,7,111.8,1,0,0,1,0,1,21,0,0,6,7,6,10.57142857,0,0,6,8,5,11.0,0,8.9,1.4,3.51,Graduate
1,1,1,9085,1,1,149.0,1,38,37,5,5,137.1,1,0,0,1,0,1,18,0,0,5,7,4,13.25,0,0,5,5,5,12.0,0,10.8,1.4,1.74,Graduate
1,1,1,9773,1,1,127.0,1,19,37,9,3,120.7,1,0,0,1,0,0,20,0,0,6,6,5,13.2,0,0,6,7,0,0.0,0,15.5,2.8,-4.06,Dropout
1,18,1,9238,1,1,137.0,1,19,38,5,8,137.4,1,0,0,1,0,0,18,0,0,6,10,1,12.0,0,0,6,14,2,11.0,0,10.8,1.4,1.74,Enrolled
1,17,2,9500,1,1,135.0,1,19,1,5,4,127.3,1,0,0,1,0,0,18,0,0,7,8,7,13.30625,0,0,8,8,8,14.545,0,15.5,2.8,-4.06,Graduate
1,1,1,9130,1,1,137.0,1,3,19,3,5,136.3,1,0,0,1,0,0,20,0,0,5,8,4,12.5,1,0,5,8,4,12.25,2,10.8,1.4,1.74,Graduate
1,1,1,9853,1,1,140.0,1,19,19,7,7,124.6,1,0,0,1,0,0,18,0,0,7,7,6,11.66666667,0,0,7,8,6,13.5,0,16.2,0.3,-0.92,Enrolled
1,1,3,171,1,1,122.0,1,1,1,9,8,120.3,0,0,0,1,0,1,21,0,0,0,0,0,0.0,0,0,0,0,0,0.0,0,11.1,0.6,2.02,Graduate
```

Figure 3: Sample data collection from the dataset

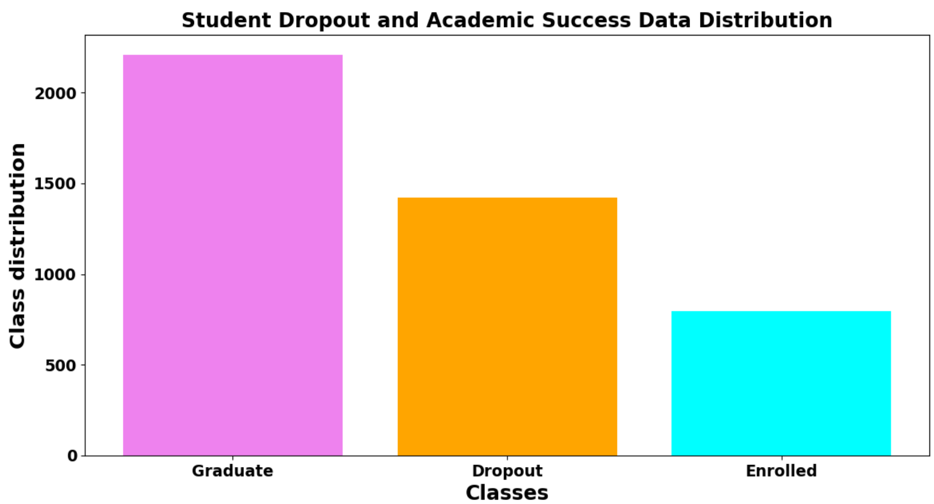


Figure 4: Class distribution of dataset

Outlier Data Removal

```
Marital status,Application mode,Application order,Course,Daytime/evening attendance,Previous qualification,Previous qualification (grade),Nacionality,
1,17,5,171,1,1,122.0,1,19,12,5,9,127.3,1,0,0,1,1,0,20,0,0,0,0,0,0.0,0,0,0,0,0,0.0,0,10.8,1.4,1.74,Dropout
1,15,1,9254,1,1,160.0,1,1,3,3,3,142.5,1,0,0,0,1,0,19,0,0,6,6,6,14.0,0,0,6,6,6,13.66666667,0,13.9,-0.3,0.79,Graduate
1,1,5,9070,1,1,122.0,1,37,37,9,9,124.8,1,0,0,0,1,0,19,0,0,6,0,0,0.0,0,0,6,0,0,0,0.0,0,10.8,1.4,1.74,Dropout
1,17,2,9773,1,1,122.0,1,38,37,5,3,119.6,1,0,0,1,0,0,20,0,0,6,8,6,13.42857143,0,0,6,10,5,12.4,0,9.4,-0.8,-3.12,Graduate
2,39,1,9991,0,19,133.1,1,37,37,9,7,114.8,0,0,1,1,1,0,50,0,0,5,10,5,11.85714286,0,0,5,17,5,11.5,5,16.2,0.3,-0.92,Graduate
1,1,1,9500,1,1,142.0,1,19,38,7,10,128.4,1,0,0,1,0,1,18,0,0,7,9,7,13.3,0,0,8,8,8,14.345,0,15.5,2.8,-4.06,Graduate
1,18,4,9254,1,1,119.0,1,37,37,9,9,113.1,1,0,0,0,1,0,22,0,0,5,5,0,0.0,0,0,5,5,0,0.0,0,15.5,2.8,-4.06,Dropout
1,1,3,9238,1,1,137.0,62,1,1,9,9,129.3,0,0,0,1,0,1,21,1,0,6,8,6,13.875,0,0,6,7,6,14.14285714,0,16.2,0.3,-0.92,Graduate
1,1,1,9238,1,1,138.0,1,1,19,4,7,123.0,1,0,1,0,0,0,18,0,0,6,9,5,11.4,0,0,6,14,2,13.5,0,8.9,1.4,3.51,Dropout
1,1,1,9670,1,1,139.0,1,38,19,5,7,130.6,1,0,0,1,0,0,18,0,0,6,6,6,12.33333333,0,0,6,7,5,14.2,0,13.9,-0.3,0.79,Graduate
1,1,1,9500,1,1,136.0,1,19,38,9,9,119.3,1,0,0,1,0,1,18,0,0,8,8,7,13.21428571,0,0,8,8,7,13.21428571,0,12.7,3.7,-1.7,Graduate
1,1,2,9853,1,1,133.0,1,19,37,4,9,130.2,1,0,0,1,0,0,19,0,0,6,6,0,0.0,0,0,6,0,0,0.0,0,12.7,3.7,-1.7,Dropout
1,53,1,9254,1,42,110.0,1,1,1,4,7,111.8,1,0,0,1,0,1,21,0,0,6,7,6,10.57142857,0,0,6,8,5,11.0,0,8.9,1.4,3.51,Graduate
1,1,1,9085,1,1,149.0,1,38,37,5,5,137.1,1,0,0,1,0,1,18,0,0,5,7,4,13.25,0,0,5,5,5,12.0,0,10.8,1.4,1.74,Graduate
1,1,1,9773,1,1,127.0,1,19,37,9,3,120.7,1,0,0,1,0,0,20,0,0,6,6,5,13.2,0,0,6,7,0,0.0,0,15.5,2.8,-4.06,Dropout
1,18,1,9238,1,1,137.0,1,19,38,5,8,137.4,1,0,0,1,0,0,18,0,0,6,10,1,12.0,0,0,6,14,2,11.0,0,10.8,1.4,1.74,Enrolled
1,1,1,9130,1,1,137.0,1,3,19,3,5,136.3,1,0,0,1,0,0,20,0,0,5,8,4,12.5,1,0,5,8,4,12.25,2,10.8,1.4,1.74,Graduate
1,1,1,9853,1,1,140.0,1,19,19,7,7,124.6,1,0,0,1,0,0,18,0,0,7,7,6,11.66666667,0,0,7,8,6,13.5,0,16.2,0.3,-0.92,Enrolled
1,1,3,171,1,1,122.0,1,1,1,9,8,120.3,0,0,0,1,0,1,21,0,0,0,0,0,0.0,0,0,0,0,0,0.0,0,11.1,0.6,2.02,Graduate
```

Figure 5: Data preprocessing

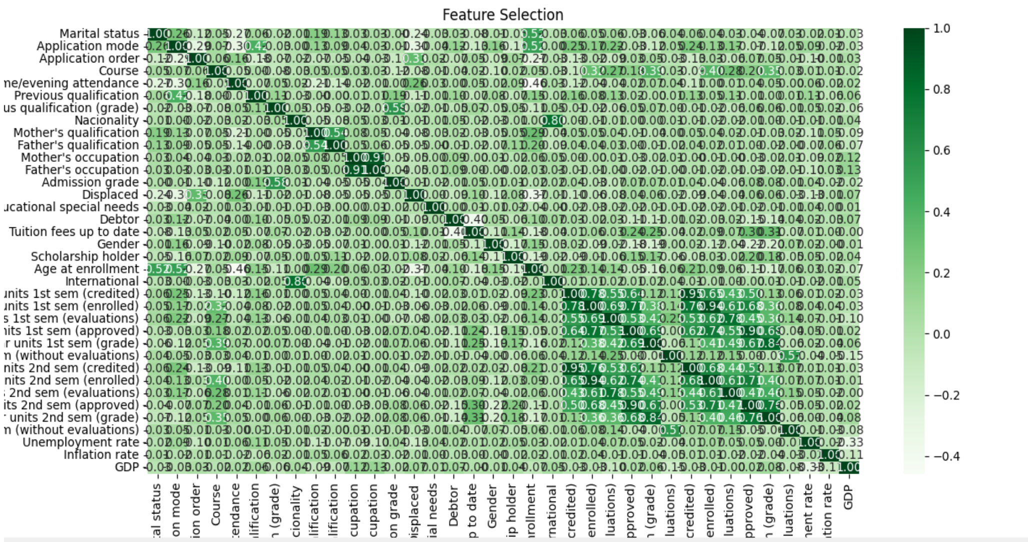


Figure 6: Heatmap similarities between the features

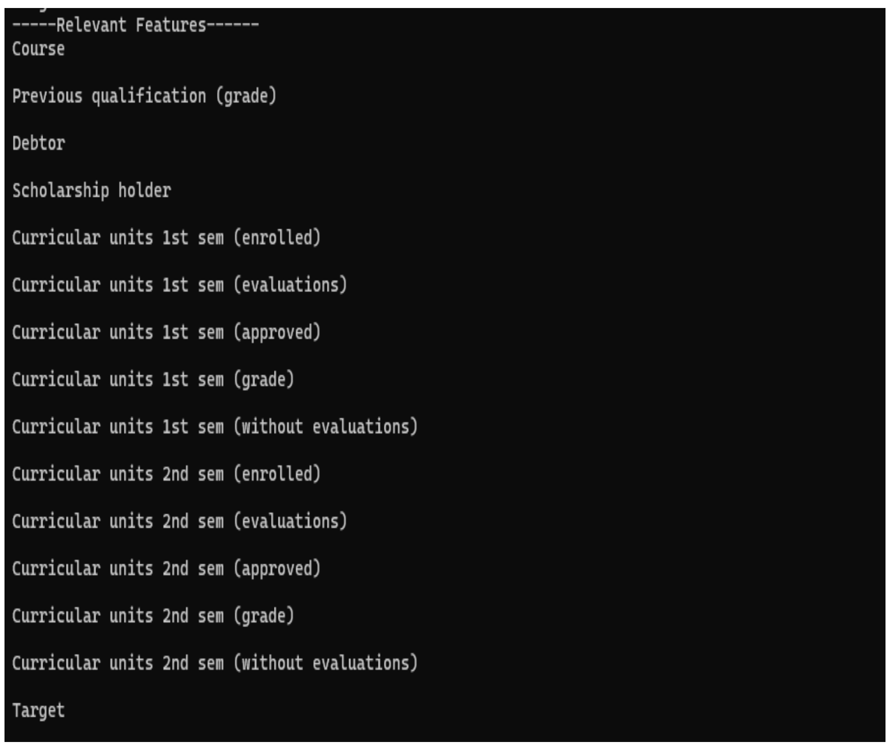
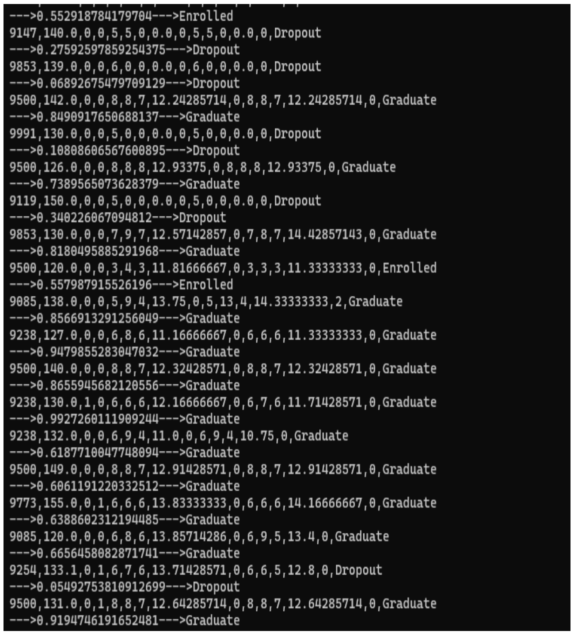


Figure 7: Selected 15 features

Figure 9 demonstrates a graphical representation of student dropout prediction accuracy using three approaches namely the proposed LOGEO-LNNet model and two existing approaches, ERGCN [1] and MLP [2]. The horizontal axis specifies the number of data samples, ranging from 440 to 4400, while the vertical axis demonstrates the corresponding prediction accuracy. Among the three methods, the LOGEO-LNNet model consistently shows superior performance in student dropout prediction accuracy. For example, with an initial set of 440data samples, LOGEO-LNNet model achieved an accuracy of 97.33%, whereas the [1] and [2] approaches observed to be93.77% and 92.22% accuracy, respectively. As the data samples count increases, varying resultswere observed across all three methods, enabling a complete performance assessment. From the results, LOGEO-LNNet model demonstrates an accuracy improvement of approximately 4% than [1] and 6% than [2]. This improvement is achieveddue to the integration of



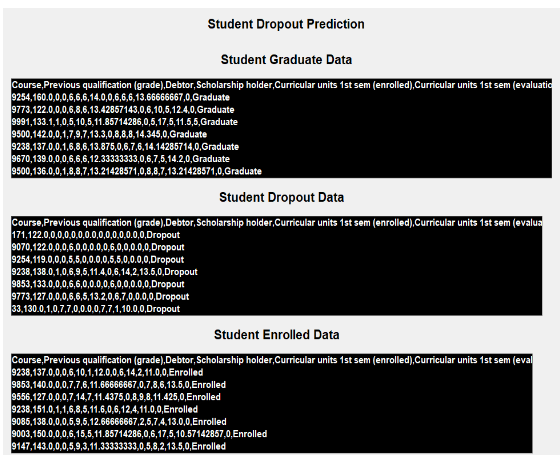


Figure 8: Classified results

Liquid Neural Networks based classification which analyzes the features, using Logistic regression resulting it enhances the accuracy. Furthermore, the fine-tuning process of the proposed deep learning architecture utilizes adaptive elitist shuffled shepherd metaheuristic algorithm to decrease classification error by increasing the true positive and true negative rates, while reducing the false positives and false negatives in the more accurate student dropout prediction.

Performance Assessment of Precision

It refers to the proportion of number of data samples correctly identified positive cases (true positives) by the total number of data samples predicted as positive, which consist of both true positives as well as false positives. The precision metric is calculated as follows

$$PS = \left(\frac{TP}{TP + FP} \right) \quad (23)$$

Where, PS represents a precision, TP indicates the true positive, FP represents the false positive. Table 3 demonstrates the result of precision for proposed LOGEO-LNNet and existing ERGCN [1] and MLP [2] respectively.

Figure 10 demonstrates the graphical diagram of precision with respect to the number of data samples collected from the dataset ranges from 440 to 4400. The overall performance of precision in student dropout prediction is measured using three methods namely LOGEO-LNNet model and two existing approaches, ERGCN [1] and MLP [2].The horizontal axis represents the number data samples, while the vertical axis of the graph designates the overall precision performance. Let us consider440 datasamples to calculate the precision. In order to apply the LOGEO-LNNet model achieved a precision of 0.979, while the existing methods [1] and [2] observed to be0.956

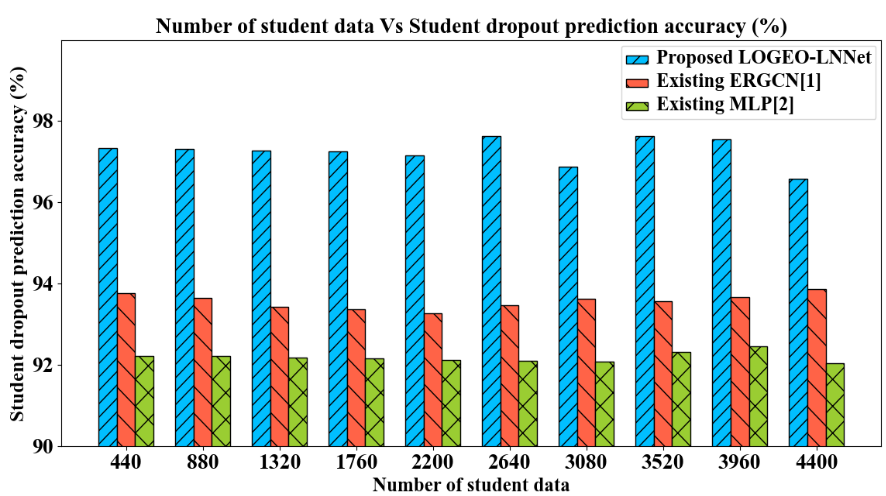


Figure 9: Graphical analysis of student dropout prediction accuracy

Table 2: Comparison of student dropout prediction accuracy

Number of student data	*Student dropout prediction accuracy (%)*		
	Proposed LOGEO-LNNet	*Existing ERGCN [1]*	*Existing MLP [2]*
440	97.33	93.77	92.22
880	97.31	93.65	92.21
1320	97.28	93.42	92.18
1760	97.26	93.36	92.16
2200	97.16	93.28	92.13
2640	97.63	93.46	92.11
3080	96.88	93.62	92.08
3520	97.63	93.57	92.32
3960	97.55	93.66	92.45
4400	96.59	93.86	92.04

Table 3: Comparison of precision

Number of student data	*Precision*		
	Proposed LOGEO-LNNet	*Existing ERGCN [1]*	*Existing MLP [2]*
440	0.979	0.956	0.942
880	0.975	0.954	0.94
1320	0.972	0.952	0.933
1760	0.976	0.948	0.938
2200	0.978	0.953	0.935
2640	0.969	0.956	0.941
3080	0.974	0.957	0.938
3520	0.977	0.955	0.937
3960	0.975	0.953	0.942
4400	0.973	0.960	0.952

and 0.942 respectively. Likewise, dissimilar precision results were achieved with respect to various counts of input data samples. As a final point, the observed results of LOGEO-LNNet model are compared to the existing methods [1] and [2]. The average of ten comparison results specifies that the LOGEO-LNNet model demonstrates an improvement in precision of approximately 2% compared to [1] and compared to4% over method [2]. This improved performance of LOGEO-LNNet model is achieved by integrating the liquid neural network in combination with the logistic regression for data analysis based on feature evaluation. Moreover, fine-tuning of liquid neural network is done through the integration of Nesterov Momentum method and adaptive elitist shuffled shepherd metaheuristic algorithm to minimize the training and validation errors. This algorithmefficientlyincreases theperformance of precision by enhancing the true positive rate while reducing false positives in student dropout prediction.

Performance Assessment of Sensitivity

It also known assensitivity, measures the ability of a model to correctly identify all actual positive cases. It refers to the ratio of true positive predictions to the total number of actual positive instances, which comprises both true positives and false negatives.

$$ST = \left(\frac{TP}{TP + FN}\right) \tag{24}$$

Where, ST indicates a sensitivity, TP represents the true positive, FN denotes the false negative. The results of sensitivity for proposed LOGEO-LNNet and existing ERGCN [1] and MLP [2] are described in Table 4.

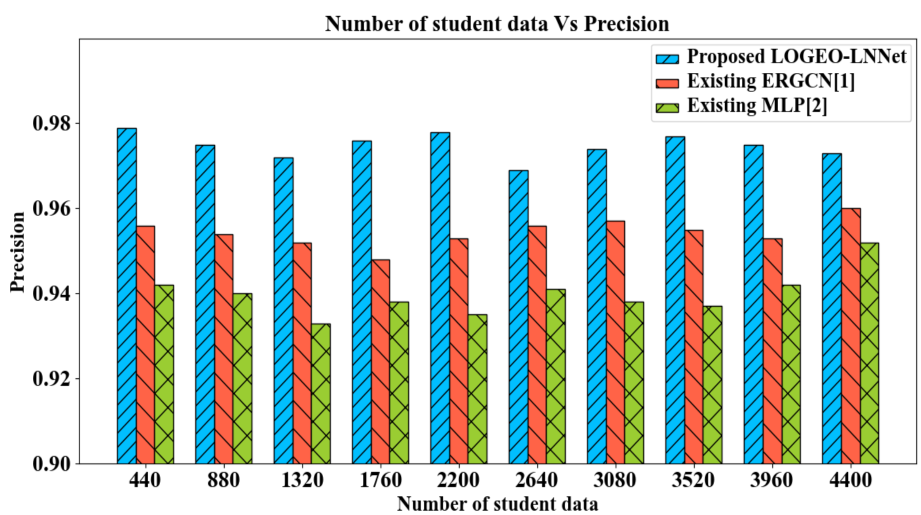


Figure 10: Graphical analysis of precision

Table 4: Comparison of sensitivity

Number of student data	sensitivity		
	Proposed LOGEO-LNNet	Existing ERGCN [1]	Existing MLP [2]
440	0.985	0.961	0.955
880	0.983	0.96	0.952
1320	0.984	0.958	0.945
1760	0.986	0.956	0.943
2200	0.985	0.953	0.946
2640	0.986	0.955	0.945
3080	0.984	0.958	0.947
3520	0.983	0.961	0.949
3960	0.985	0.962	0.953
4400	0.986	0.968	0.954

Figure 11 illustrates an overall comparison analysis of sensitivity performance with the number of data samples collected from the dataset. The Proposed LOGEO-LNNet model exhibits higher recall performance in sensitivitycompared to the existing approaches [1] and [2]. For example, with 440 data samples, LOGEO-LNNet model achieved a recall of 0.985, while methods [1] and [2] observed to be 0.961 and 0.955, respectively. Likewise, various performance results were analyzedalong with different counts of input samples. The experimental analysis demonstrates that the average sensitivity of LOGEO-LNNet model improved by approximately 3% compared to [1] and 4% compared to [2].The greater performance of LOGEO-LNNet model is achieved because of the fine-tuning process applied within the liquid neural network. By utilizing a liquid neural network, the model reduces the sensitivity between predicted and actual outcomes in multiclass output through optimal weight adjustment using the adaptive elitist shuffled shepherd metaheuristic algorithm. This process getsiterated continue the error gets minimized, resulting in minimal false negatives and enhance in true positives for precise student dropout predictionin educational sector.

Performance Assessment of F1-score

This metric integrates both precision as well as recall into one value, providing a balanced evaluation of model efficiency.

$$F1score = 2 * \left(\frac{PS * ST}{PS + ST} \right) \tag{25}$$

Where, PS indicates a precision, ST indicates a sensitivity. The results of F1-score for proposed LOGEO-LNNet and existing ERGCN [1] and MLP [2] are portrayed in Table 5.

The graphical representation of F1 score of three different methods namely LOGEO-LNNet model and two

Table 5: Comparison of F1-score

Number of student data	F1-score		
	Proposed LOGEO-LNNet	Existing ERGCN [1]	Existing MLP [2]
440	0.981	0.958	0.948
880	0.978	0.956	0.945
1320	0.977	0.954	0.938
1760	0.980	0.951	0.940
2200	0.981	0.953	0.940
2640	0.977	0.955	0.942
3080	0.978	0.957	0.942
3520	0.979	0.957	0.942
3960	0.979	0.957	0.947
4400	0.979	0.963	0.952

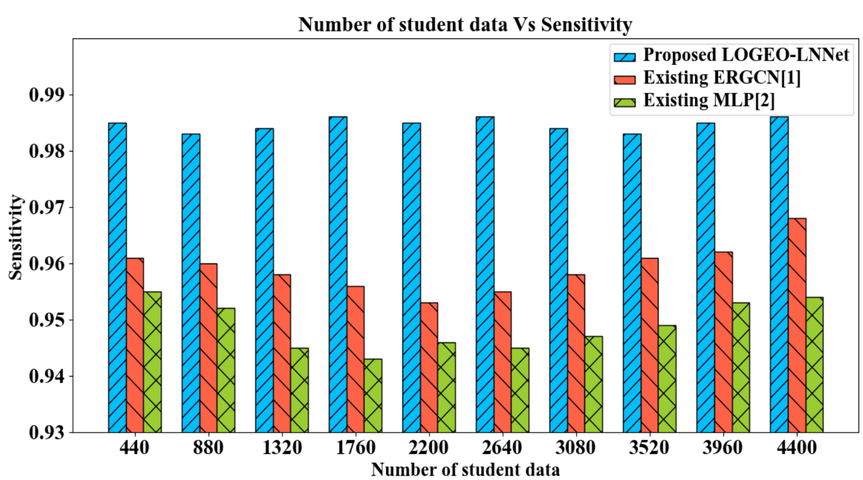


Figure 11: Graphical analysis of sensitivity

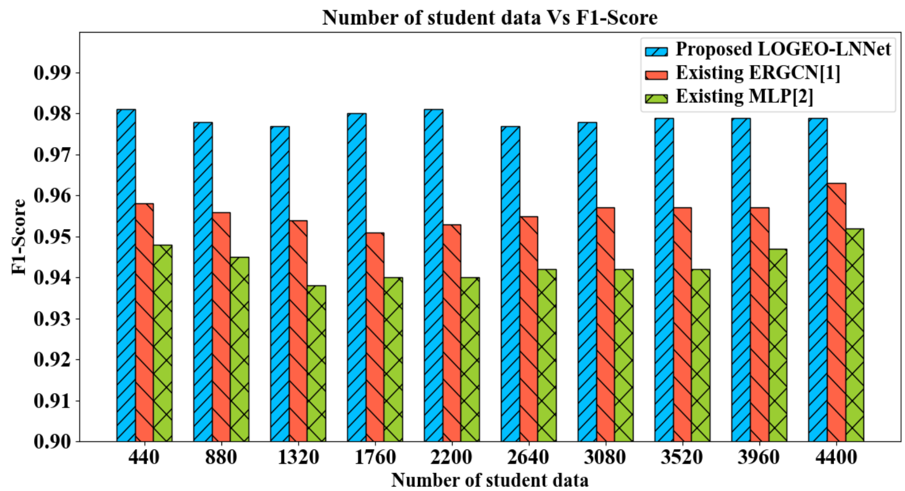


Figure 12: Graphical analysis of F1-score

existing approaches, ERGCN [1] and MLP [2] are shown in figure 12. The results obviously allocate that the higher F1 score is achieved using LOGEO-LNNet model compared to other methods. In the first iteration with 440data samples, the F1 score of proposed LOGEO-LNNet model was observed to be 0.981, whereas the F1 score of existing methods [1] and [2] were 0.958 and 0.948 respectively. These results validate that the proposed LOGEO-LNNet model significantly enhances the F1 score compared to existing techniques. This improvement is achieved by the LOGEO-LNNet model increases the precision as well as recall in the student dropout prediction.Therefore, the performance of LOGEO-LNNet model was compared with the numerous results observed from the existing methods. The overall performance results designate that LOGEO-LNNet model increases the F1 score by 2% and 4% compared to [1] and [2] respectively.

Table 6: Comparison of specificity

Number of student data	*Specificity*		
	Proposed LOGEO-LNNet	*Existing ERGCN [1]*	*Existing MLP [2]*
440	0.936	0.863	0.818
880	0.933	0.862	0.826
1320	0.932	0.858	0.811
1760	0.928	0.855	0.823
2200	0.924	0.846	0.818
2640	0.916	0.839	0.812
3080	0.918	0.842	0.806
3520	0.906	0.837	0.807
3960	0.885	0.833	0.796
4400	0.838	0.761	0.714

Performance Assessment of Specificity

It exposes the model capability to correctly distinguish between different classes of student dropout. The mathematical formula is used for calculating the specificity as follows,

$$SP = \frac{TN}{TN + FP} \tag{26}$$

Where, 'SP' denotes a specificity, TN indicates a true negative and'FP' designates a false positive. Specificity result shown in Table 6 by using proposed LOGEO-LNNet and existing ERGCN [1] and MLP [2].

Figure 13 illustrate the specificity performance against number of data samples, ranging from 440 to 4400, derived from the Dataset. In this figure 13, the horizontal axis indicates the number of data samples, while the vertical axis indicates the corresponding specificity values in student dropout prediction. The experimental results designate that the proposed LOGEO-LNNet model consistently achieved higher specificity than the two existing deep learning models. For example, 440data samples were considered in first iteration. The LOGEO-LNNet model attained a specificity of 0.936, outperforming the existing methods, which recorded values of 0.863 and 0.818, respectively. These results were numerically compared across dissimilar number of data samples to evaluate specificity in student dropout prediction. The comparative analysis demonstrates that the LOGEO-LNNet model improved specificity by 9% and 14% whencompared to [1] and [2].

Performance Assessment of Student Dropout Prediction Time

It refers to the amount of time taken by the algorithm to predict the dropout studentsfrom the given input data samples. The overall prediction time is expressed as follows,

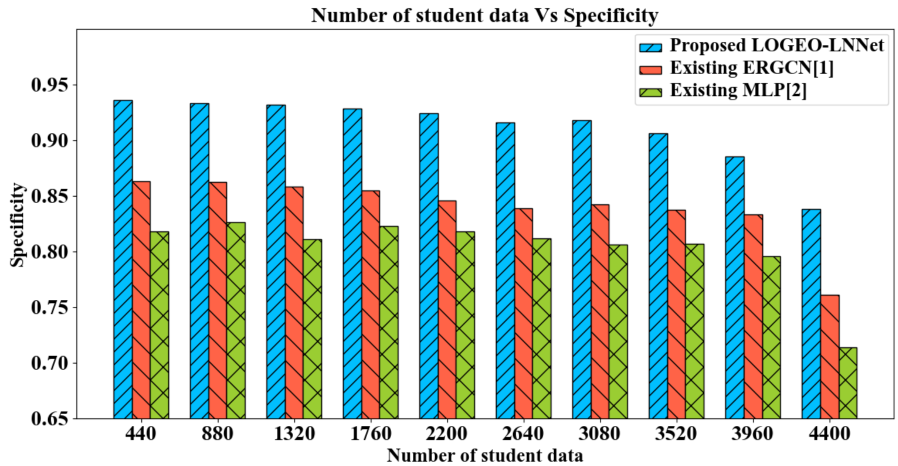


Figure 13: Graphical analysis of specificity

Table 7: Comparison of student dropout prediction time

Number of student data	*Student dropout prediction time (ms)*		
	Proposed logeo-lnnet	*Existing ergcn [1]*	*Existing mlp [2]*
440	22	26.4	30.8
880	25.3	29.6	32.6
1320	28.2	31.5	34.5
1760	30.4	33.4	35.4
2200	32.7	34.8	37.8
2640	34.2	37.2	40.2
3080	36.8	38.4	42.6
3520	38.4	40.9	44.4
3960	40.1	43.6	47.2
4400	42.6	47.5	50.8

$$SDPT = \sum_{i=1}^{n} Ds_i * TME(SDP)] \quad (27)$$

Where '$SDPT$ represents student dropout prediction time, n indicates the number of student data samples 'Ds', $TME(SDP)$ indicates a time for predicting the dropout.The time consumption is measured in terms of milliseconds (ms). Table 7 portrays the student dropout prediction time result using three methods.

Figure 14 illustrates the performance analysis of student dropout prediction time using the proposed LOGEO-LNNet model and two existing approaches, ERGCN [1] and MLP [2]. The detailed graphical results are observed against the number of input data samples, ranging from 440 to 4400. As the number of data sample increases, the crop recommendation time for all three methods increases accordingly. The x-axis indicates the number of samplescollected from the dataset, while the y-axis indicates the corresponding student dropout prediction time.In an experiment with 400 data samples, LOGEO-LNNet model achieved a student dropout prediction time of $22ms$, while methods [1] and [2] observed a prediction time of $26.4ms$ and $30.8ms$ respectively.Among three methods, the LOGEO-LNNet model constantly decreases thestudent dropout prediction timeby 9% and 16% compared to the two existing models. This efficiency is achieved due to its data pre-processing and feature selection process. In the data preprocessing, Polynomial interpolation function is employed for handling the missing data. Followed by, Peirce statistical decisive factor is used for outlier data removal. Moreover, the combinatorial Target projection matching based feature selectionprocess is carried out in LOGEO-LNNet model continually to select the more relevant features for accurate student dropout prediction. Overall, the analysis confirms that the LOGEO-LNNet model reduced computation time in student dropout prediction.

Performance Analysis of Confusion Matrix

The confusion matrix is performance assessment tool in student dropout predictionmodel, as it facilitates a complete assessment between predicted outcomes and actual class labels. In this analysis, confusion matrices are employed to estimate and compare the performance of three various approaches namely the proposed LOGEO-LNNet model and two existing approaches, ERGCN [1] and MLP [2]. Based on a dataset comprising 4400student data samples, these matrices visually exposes the performance of each model detects a student dropout. Each confusion matrix table provides the classification results with four main components namely True Positives (TP), False Positives (FP), False Negatives (FN), and True Negatives (TN) to increase the effectiveness of each method.

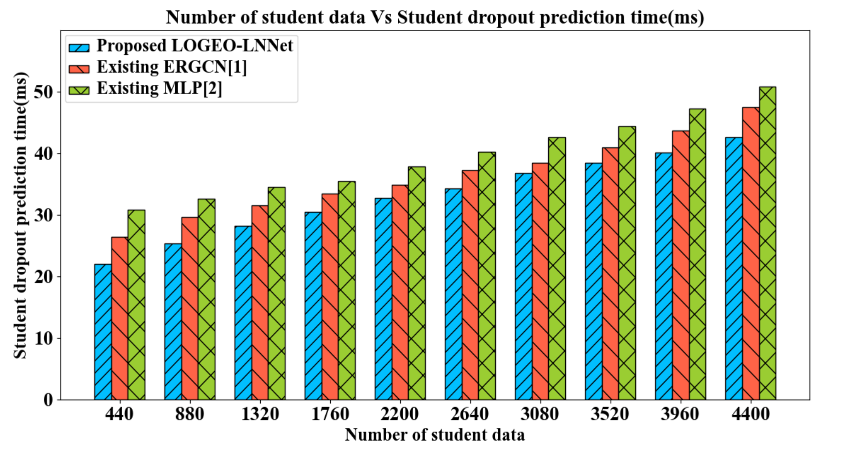


Figure 14: Graphical analysis of student dropout prediction time

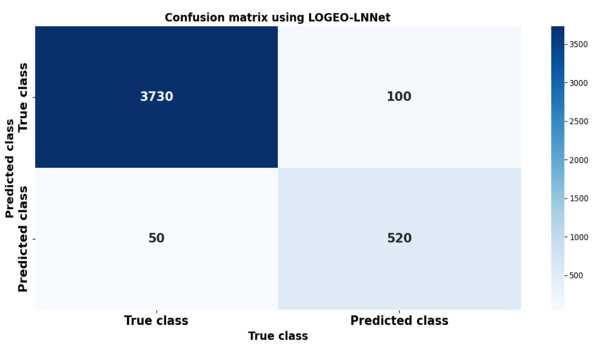


Figure 15: Confusion matrix using LOGEO-LNNet model

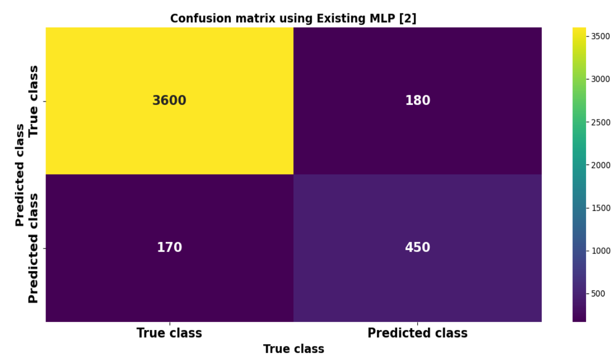


Figure 17: Confusion matrix using ERGCN [1]

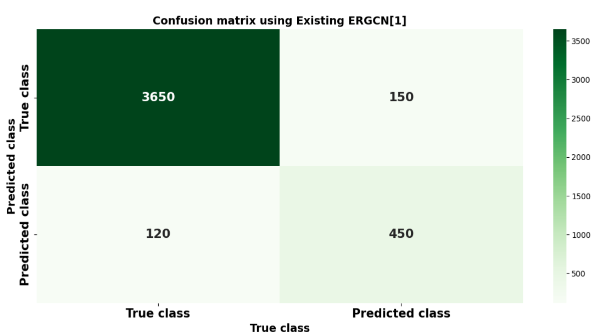


Figure 16: Confusion matrix using ERGCN [1]

Figures 15, 16, 17 present performance of confusion matrices for the LOGEO-LNNet model and two existing approaches, ERGCN [1] and MLP [2].The above figure illustrates the results of the classification algorithms. Instudent dropout prediction, these matrices help distinguish between multiple classes namely dropout, enrolled and graduate. The output samples in these matrices provide insight into LOGEO-LNNet model performance, where high TP and TN combined with low FP and FN indicates better classification accuracy. In this analysis, the confusion matrices analysis provide as a essential tool for demonstrating the model's efficiency, highlighting its ability to reduce error and enhance the student dropout predictioncompared to existing approaches.

Conclusion

Student dropout prediction plays a vital role in higher education for determining academic success in specific institutions. This paper introduces a novel LOGEO-LNNet model specifically developed to accurately predicting the student dropout. The proposed model initiates with a data preprocessing and significant feature selection aimed to minimize the overall student dropout prediction time. It also incorporates a liquid neural network modelto facilitate efficient samples analysis based on regression analysis to enhance prediction accuracy. To further enhance the model's performance, hyperparameter tuning is employed to minimize the error rates associated with student dropout prediction.A comprehensive assessment was conducted using several performance metrics, including accuracy, precision, sensitivity, F1-score, specificity,student dropout prediction time analysis across different student data samples. The quantitatively analyzed results exhibit that the LOGEO-LNNet model outperforms traditional deep learning methods, achieving higher accuracy in student dropout prediction and minimizing the time consumption as well as error rate.

Acknowledgement

None.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Conflict of Interest

The authors declared that they have no conflicts of interest in this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Availability of Data and Material

Predict Students' Dropout and Academic Success dataset is used and publicly available within the article. It is taken from https://www.kaggle.com/datasets/mattop/predict-students-dropout-and-academic-success

Code Availability

Not applicable.

Author Contributions

The corresponding author claims the significant contribution of the paper, including formulation, analysis, and editing. The co-author provides guidance to verify the analysis result and manuscript editing.

References

Yang Guo andYijia He, "MOOC Dropout Prediction Using Explainable Relational Graph Convolution", IEEE Access,Volume 13, 2025, Pages 204759 – 204772. **DOI:** 10.1109/ACCESS.2025.3637528

Yu-Huei Cheng, Mu-Hsin Shih,Che-Nan Kuo, "Early Prediction of Student Suspension and Dropout for Counseling Resource Optimization Using Multilayer Perceptron", IEEE Access, Volume 13, 2025, Pages 119810 – 119819. DOI: 10.1109/ACCESS.2025.3587484

MarksonRebeloMarcolino, Thiago Reis Porto, Tiago Thompsen Primo, Rafael Targino, Vinicius Ramos, Emanuel Marques Queiroga, Roberto Munoz &CristianCechinel, "Student dropout prediction through machine learning optimization: insights from moodle log data", Scientific Reports, volume 15, 2025, Pages 1-16.https://doi.org/10.1038/s41598-025-93918-1

SekarGestiAmaliaUtami, HaryonoSetiadi, ArifRohmadi, "Comparative Analysis of Machine Learning Algorithmswith RFE-CV for Student Dropout Prediction", Jurnal Teknik Informatika (JUTIF), Volume 6, Issue 3, 2025, Pages 1319-1338. https://doi.org/10.52436/1.jutif.2025.6.3.4695

ÁdámKocsis a and GyöngyvérMolná, "Factors influencing academic performance and dropout rates in higher education", Oxford Review of Education, Volume 51, Issue 3, 2025, Pages 414-432. https://doi.org/10.1080/03054985.2024.2316616

Helai Liu ,Mao Mao,XiaLi,JiaGao, "Model interpretability on private-safe oriented student dropout prediction", PLoS ONE ,Volume 20, Issue 3, 2025, Pages 1-23.https://doi.org/10.1371/journal.pone.0317726

NurhanaRoslan, JastiniMohdJamil , IzwanNizalMohdShaharanee, Sultan Juma Sultan Alawi, "Prediction of Student Dropout in Malaysian's Private Higher Education Institute using Data Mining Application",Journal of Advanced Research in Applied Sciences and Engineering Technology, Volume45, Issue 2, 2025, Pages 168-176. https://doi.org/10.37934/araset.45.2.168176

Blanca Carballo-Mendívil, Alejandro Arellano-González, Nidia Josefina Ríos-Vázquez, and María del PilarLizardi-Duarte, "Predicting Student Dropout from Day One: XGBoost-Based Early Warning System Using Pre-Enrollment Data", Applied Science, Volume 15, Issue 16, 2025, Pages1-29.https://doi.org/10.3390/app15169202

Zhiqiang Zhao, Ping Ren, "Random Forest-Based Early Warning System for Student Dropout Using Behavioral Data", Bulletin Of Education and Psychology, Volume 5, Issue 5,2025, Pages 1-21.https://hal.science/hal-05102483

Yinxu Liu, ChengjunXu, DeshengYang andYunchengShen, "MOOC Dropout Prediction via a Dilated Convolutional Attention Network with Lie Group Features", Informatics, Volume 12, Issue 4,2025, Pages 1-17.https://doi.org/10.3390/informatics12040127

ParthaGhosh, ArnabCharit, Hindol Banerjee, DebanwesaBandhu, AgnivGhosh, Ankita Pal, TakaakiGoto&SoumyaSen, "DropWrap: A Neural Network Based Automated Model for Managing Student Dropout", Volume 13, 2025, Pages 1-14. https://doi.org/10.1007/s44227-025-00058-z

Govinda Prasad Dhungana, Arun Kumar Chaudhary, Ramesh Prasad Tharu, Vijay Kumar, "Bayesian Modelling of Weibull Distribution for Predicting Student Dropout Rates: A Case of Higher Education Institutions in Nepalese University", ZhongguoKuangyeDaxueXuebao, Volume 30, Issue 1,2025,Pages 48-59. https://zkdx.ch/journal/zkdx/article/view/200

André Menolli, Gustavo MarcelinoDionísio, Alan Silva da Paz Floriano, Thiago Adriano Coleti, "Exploring Feature Reduction for Dropout Predicting in Higher Education in Brazil", RevistaBrasileira de InformáticanaEducação, Volume 33, 2025,Pages 106–129. https://doi.org/10.5753/rbie.2025.4619

Syahrul Anwar, "Application of K-Means and C4.5 Algorithms for dropout risk prediction in vocational high schools", Indonesian Journal of Multidisciplinary Science, Volume 4,Issue 5, 2025, Pages 428–435. https://doi.org/10.55324/ijoms.v4i5.1102

Alfredo Guzman Rincon, "Dropout in rural higher education in virtual modality: a neuralnetwork approach", Cogent Education, Volume 12, Issue 1, 2025, Pages 1-17.https://doi.org/10.1080/2331186X.2025.2493445

Muhammad Ricky Perdana Putra, EmaUtami, "A Blending Ensemble Approach to Predicting Student Dropout in Massive Open Online Courses (MOOCs)", JUITA: JurnalInformatikae, Volume 13, Issue 1, March 2025, Pages 11-18.https://doi.org/10.30595/juita.v13i1.24061

Sandra Patricia Barragán Moreno & Alfredo GuzmánRincón, "Digital divide as an explanatory variable for dropout in higher education", International Journal of Educational Technology in Higher Education, Springer, Volume 22, 2025, Pages 1-26.https://doi.org/10.1186/s41239-025-00550-0

Guilherme Antonio Borges, Carolina Filipa Dos Santos Pedro, Julio Cesar Santos Dos Anjos, André Rodrigues, Fernando Boavida, Jorge Sá Silva, "A Platform for Early Class Dropout Prediction of University Students", IEEE Access ,Volume 13, 2025, Pages 109116 - 109133. DOI: 10.1109/ACCESS.2025.3581751

Francisco da Conceição Silva, Andre Macedo Santana, Rodrigo Miranda Feitosa, "An Investigation Into Dropout Indicators in Secondary Technical Education Using Explainable Artificial Intelligence", IEEE RevistaIberoamericana de Tecnologias del Aprendizaje, Volume20, 2025, Pages 105 – 114. DOI: 10.1109/RITA.2025.3566095

HyeonGyu Kim,Yoohyun Park, "Calibrating F1 Scores for Fair Performance Comparison of Binary Classification Models With Application to Student Dropout Prediction", IEEE Access, Volume13, 2025, Pages 136554 – 136567. DOI: 10.1109/ACCESS.2025.3594735

Yuan Lu,SoonjaYeom, RiseulRyu, Soo-Hyung Kim, "Course-Level Clustering to Enhance Dropout Prediction Accuracy", IEEE Access, Volume 13, 2025, Pages 204996 – 205013. DOI: 10.1109/ACCESS.2025.3639083

Aini, Q., Rahajeng, E., Tiohandra, M., Pratama, H. A., & Hammad, J. (2025). Predictive modeling of student dropout using academic data and machine learning techniques. Applied Information System and Management (AISM), 8(2), 203-212.

Johannes Berens, Andrea Diem, Luis Rumert, Kerstin Schneider & Stefan C. Wolter, "Crossing individual university boundaries: a comprehensive approach to predicting dropouts in the higher education system", Higher Education, Springer, 2025, Pages 1-22.https://doi.org/10.1007/s10734-025-01509-w

Marco Forti, Massimo Bilancia, Barbara Cafarelli, EmilianodelGobbo & Andrea Nigri, "Predicting dropout from higher education with multinomial finite mixture models: empirical evidence from Italy", Quality & Quantity, Springer, 2025, Pages 1-15. https://doi.org/10.1007/s11135-025-02135-5

OlusolaOlajideAjayi, "Predicting Student Dropout Risk in Online Learning using Stacked Ensemble Machine Learning and Explainable AI Techniques", International Journal of Computer Applications, Volume 187 , Issue 40, 2025, Pages 26-29. DOI: 10.5120/ijca2025925707

BehrouzAlefy ,VahidBabazadeh, "Analyzing Student Graduation and Dropout Patterns Using Artificial Intelligence and Survival Strategies", Journal of Artificial Intelligence and System Modelling, Volume 03, Issue 02, 2025, Pages 32-46. DOI: 10.22034/jaism.2025.509969.1113

Ziyang Liu, Xiang Zhou and Yijun Liu, "Student Dropout Prediction Using Ensemble Learningwith SHAP-Based Explainable AI Analysis", Journal of Social Systems and Policy Analysis, Volume 2, Issue 3, 2025, Pages 111-132.DOI: 10.62762/JSSPA.2025.321501

JankaPecuchova, Martin Drlik, "Enhancing the Early Student Dropout Prediction Model Through Clustering Analysis of Students' Digital Traces", IEEE Access , Volume 12, 2024, Pages 159336 – 159367. DOI: 10.1109/ACCESS.2024.3486762

RadicGoran, Luka Jovanovic,NebojsaBacanin,Miloš S. Stanković, Vladimir Simic, Milos Antonijevic, "Identifying and Understanding Student Dropouts Using Metaheuristic Optimized Classifiers and Explainable Artificial Intelligence Techniques", IEEE Access ,Volume 12, 2024, Pages 122377 – 122400. DOI: 10.1109/ACCESS.2024.3446653

Andrea Zanellati, Stefano PioZingaro, Maurizio Gabbrielli, "Balancing Performance and Explainability in Academic Dropout Prediction", IEEE Transactions on Learning Technologies, Volume 17, 2024, Pages 2086 – 2099. DOI: 10.1109/TLT.2024.3425959