ISSN: 0974-8571 Vol.13 No. 2 December, 2021

International Journal of Computational Intelligence in Control

Optimization based Neural Network Approach for Student Performance Classification

K. LAKSHMI
Research Scholar, PG and Research Department of Computer Science, Jamal Mohamed College (Autonomous) (Affiliated to Bharathidasan University), Tiruchirappalli, Tamilnadu, India.

Dr. M. RAJAKUMAR
Assistant Professor, PG and Research Department of Computer Science, Jamal Mohamed College (Autonomous) (Affiliated to Bharathidasan University), Tiruchirappalli, Tamilnadu, India.

Date of Submission: 15th August 2021 Revised: 27th October 2021 Accepted: 18th December 2021

***Abstract* - Educational Data Mining contributes cutting-edge approaches, strategies, and applications to the advancement of the learning environment. By investigating and utilising educational data using machine learning and data mining approaches, the current advancement gives significant tools for evaluating the student learning environment. Academic institutions today operate in a complicated and competitive environment. Universities have numerous issues, including analysing performance, offering high-quality education, developing ways for evaluating students' achievement, and planning future activities. In order to solve challenges that students have during their studies, these colleges must create student intervention strategies. The Optimized Siamese Neural Network (CA-SNN) Classification model is proposed in this study. The Cultural Algorithm is used to optimise the SNN classifier. The use of CA in SNN classifier is proposed, which considerably reduces the challenge of training SNN using a local search-based learning method. SNN is a deep learning architecture that can handle complex data and produce the best results for a particular dataset. Various metrics and other classifiers are used to evaluate the proposed CA-SNN classification model's performance.**

Index Terms - Feature Selection, Classification, Optimization, Neural Network, Educational Data Mining, Student's Performance Evaluation.

INTRODUCTION

Educational Data Mining has had a considerable influence on current educational advances (EDM). A wide range of studies have identified and implemented new possibilities and opportunities for technologically enhanced learning systems that are tailored to the needs of students. The EDM's cutting-edge approaches and application strategies are crucial in improving the learning environment. For example, by analysing both the educational setting and machine learning approaches, the EDM is crucial in determining the student learning environment. According to [1,] the EDM discipline is concerned with the exploration, study, and implementation of Data Mining (DM) methodologies. For its success, the DM discipline uses multi-disciplinary methodologies. Educational data is evaluated using machine learning and statistical methodologies to find relevant patterns that improve students' knowledge and academic institutions in general.

Modern educational institutions operate in a complicated and competitive environment. As a result, most universities nowadays confront issues such as analysing performance, delivering high-quality education, devising systems for evaluating students' performance, and predicting future needs. Universities use student intervention plans to help students overcome difficulties during their education. Student performance prediction at the entry level and over time assists universities in effectively developing and evolving intervention programmes, with both management and educators benefiting from the students' performance prediction plans. Students enrol in online courses in e-learning, which is a fast increasing and advanced kind of education. E-learning platforms including Intelligent Tutoring Systems (ITS), Learning Management Systems (LMS), and Massive Open Online Courses (MOOC) make extensive use of EDM in the development and construction of automatic grading systems, recommender systems, and adaptive systems. These platforms employ clever tools that collect vital user data such as the frequency with which a student logs into the e-learning system, the precision with which the student answers questions, and the amount of hours spent reading texts and watching video lectures [2].

The collected data is processed and analysed over time using various machine learning algorithms in order to improve usability and construct interactive tools on the learning platform. "Research employing Machine Learning (ML) is part of Artificial Intelligence (AI), attempting to impart knowledge to computers through data, observations, and intimate interaction with the world," says Dr. Yoshua Bengio [3] of the University of Monreal. The accumulated knowledge enables the computer to correctly generalise to new settings." Machine learning is a branch of artificial intelligence in which computers learn from data, evaluate patterns, and anticipate outcomes. The regeneration of the machine from just a pattern recognition algorithm to Deep Learning (DL) methods is due to rising volumes of data, cheaper storage, and robust computer systems. ML models can analyse larger and more complicated data automatically and fast, producing accurate findings and avoiding unanticipated dangers.

RELATED WORKS

Giannakas, F., et al [4] suggested a Deep Neural Network (DNN) framework for binary classification with two hidden layers for early software engineering team performance prediction. Different activation functions (Sigmoid, ReLu, and Tanh) and optimizers were used to test the framework (Adagrad and Adadelta).

Sokkhey, Phauk, and Takeo Okazaki [5] To improve classification performance by solving the misclassification problem, researchers combined a hybrid approach of principal component analysis (PCA) with four machine learning (ML) algorithms: random forest (RF), C5.0 of decision tree (DT), nave Bayes (NB) of Bayes network, and support vector machine (SVM). Three datasets were utilised to test the proposed models' resilience.

Waheed, Hajra, et al [6] used a deep artificial neural network trained on a set of handmade attributes taken from clickstream data from virtual learning environments to forecast at-risk pupils and provide early intervention strategies. The suggested model has a classification accuracy of 84 percent to 93 percent, according to the data.

Pallathadka, Harikumar, et al [7] Students should be advised ahead of time to focus their efforts on a specific area in order to boost their academic performance. This type of research can help an institution reduce its failure rates. This study predicts students' achievement in a course based on their previous performance in related courses. Data mining is a set of techniques for discovering hidden patterns in huge volumes of data. These patterns could be useful for analysis and forecasting. Data mining applications in the realm of education are referred to as education data mining.

Bhutto, Engr Sana, et al [8] introduced a model for predicting students' academic achievement that uses supervised machine learning algorithms such as support vector machine and logistic regression. The results of several trials utilising various technologies are compared, and it is discovered that the sequential minimal optimization technique exceeds logistic regression in terms of accuracy. And the knowledge gained from this study can assist educational institutions in predicting students' future conduct and categorising their performance as excellent or bad.

Ha, Dinh Thi, et al [9] Machine learning techniques were used to estimate students' ultimate Grade Point Average based on personal variables (such as gender and living location), university application scores, gap year, and first- and second-year academic performance. The data was compiled by merging information from a three-year survey of graduate students with information from the university's student management information system.

Rastrollo-Guerrero, Juan L., Juan A. Gómez-Pulido, and Arturo Durán-Domínguez [10] Almost 70 publications were evaluated in this study to highlight numerous modern methodologies extensively used for predicting students' performance, as well as the goals they must achieve in this sector. Machine Learning, Collaborative Filtering, Recommender Systems, and Artificial Neural Networks are just a few of the Artificial Intelligence techniques and methodologies available.

Lin, Yongzheng, et al [11] brought machine learning-based methodologies into the examination of students' academic performance, and used rank models' learning-based approaches to create a framework for analysing students' learning abilities. The capability of numerous media, including image and video, has been inserted into the proposed architecture by using imaging sensors. Meanwhile, the authors have used the pipeline for different kids' academic performance to demonstrate academic commonalities among students as well as the varying influences of basic courses on students' future development between groups.

CULTURAL ALGORITHM

Cultural Algorithms (CA) are a sophisticated method evolved from nature's cultural growth mechanism [12]. A CA is an evolutionary computational classification based on knowledge. Its fundamental understanding is the integration of knowledge mechanisms into advanced computational systems. Its simulations are divided into two stages of development: population space and belief space. A categorical communication protocol composed of an acceptance function and an influence function, denoted here as Accept () and Influence (), respectively, connects the two spaces. The cultural algorithm contains the following characteristics:

- Dual Evolutionary: Inborn parent data exists in the population and belief space .
- The belief in the escort of space knowledge protects population space evolution.

- Assist the hierarchy of population space and belief space.
- Assisting two-space adaptive evolution;
- Different speeds can be used to carry out different types of space evolution;
- Encourage the use of a variety of algorithms to solve the problem;
- Within a model, "cultural" change can be articulated in a variety of ways.

SIAMESE NEURAL NETWORK CLASSIFICATION

The siamese neural network algorithm was first introduced by Bromley et al. [10] to detect forged signatures in 1994. Before that, Baldi and Chauvin [11] introduced a similar artificial neural network able to recognize fingerprints, though by a different name. In the study by Bromley et al. [10], by comparing two handwritten signatures, this Siamese Neural Network was able to state if the two signatures were both original or if one was a forgery.

An artificial neural network is a machine learning model made of several layers of neurons, which are processing units inspired by biological neurons [15]. Typical supervised neural networks are called feedforward, and are based on the perceptron model [16]. In a feedforward neural network, each neuron of the first layer reads in an input real value, multiplies it by a weight, and sends the result to all the neurons in the following layer. In common neural network representations, scientists consider the left-most layer as the first layer and the right-most layer as the output layer (Fig. 1). The neurons of each layer beyond the input layer do the same job and send the results to the neurons of the next layer, until the final layer (the process goes from left to right). In a supervised neural network, the final layer then sends its results to a one-neuron output layer that produces the neural network real valued result.

During training, the neural network compares the values produced by the neural network with its corresponding ground truth, and computes the statistical error (usually the mean square error or the cross-entropy error). Afterwards, the neural network sends the error back to the previous layers and updates its neuron weights accordingly, through a technique called error back-propagation [17]. In representation terms, back-propagation happens from right to left (Fig. 1). The training stops when the neural network reaches the maximum number of iterations initially set. Once the model is trained, it can then be applied to the test set: the neural network will process each test data instance (only once, from the first layer on the left to the last layer on the right) and will generate a predicted value. Once the neural network has generated a predicted value for each test set instance, the programmer can use it to compute a confusion matrix.

Feedforward neural networks with error back-propagation are employed in siamese neural networks, as well. The siamese neural network architecture, in fact, contains two identical feedforward neural networks joined at their output (Fig. 1), which work parallelly in tandem. Each neural network contains a traditional perceptron model. During training, each neural network reads a profile made of real values, and processes its values at each layer. The neural network activates some of the neurons based upon these values, updates its weights through error back-propagation [14], and at the end it generates an output profile that is compared with the output of the other neural network.

The algorithm compares the output of the upper neural network and the output of the lower neural network through a distance metric (Fig. 1): a cosine distance in the original model [13].

Through this similarity measure, the neural network states that the two profiles are different (cosine similarity value in the [_1, 0] interval) or similar (cosine similarity value in the [0, +1] range). The algorithm then labels the data instance as positive if in the former case, or as negative if in the latter case. The final output value can finally be compared with its corresponding ground truth value; all the classified outputs can be employed to generate a confusion matrix.

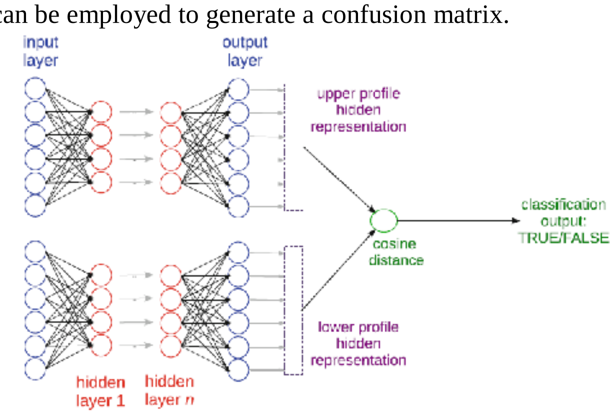

FIGURE 1: SIAMESE NEURAL NETWORK MODEL REPRESENTATION

PROPOSED OPTIMIZATION BASED SIAMESE NEURAL NETWORK (CA-SNN)

In this work, the Siamese Neural Network (SNN) classifier is optimized with Genetic algorithm. CA is used in SNN classifier is proposed which significantly overcomes the problem of using local search-based learning algorithm to train SNN. SNN is a deep learning architecture which is used to handle complex data and to generate the optimal result for the given dataset. Siamese Neural Network (SNN) architecture consists of two neural networks that share a identical weights and are joined at one or more layers. SNN make use of twin network to realize a non-linear embedding (it's usually a Deep Neural Network) from its input domain. Weight sharing ensures that two similar samples will not map to different parts of a input space as each part uses the same functionality. The network is symmetric, meaning that it does not matter how an input pair is fed to the network. CA is an intelligent probabilistic search algorithm which can be applied to a variety of combinational optimization problem. CA stimulates the processes by taking an initial population of individuals and applying CA operators in each generation. Each individual is encoded as a chromosome which is a solution to the problem. A chromosome is a collection of genes, means an individual is made up of genes. The fitness of each individual is calculated by objective function. Highly fit individuals are given chances for reproduction, in crossover procedure. Mutation is optional for changing some of genes in individual to avoid duplicity.

T

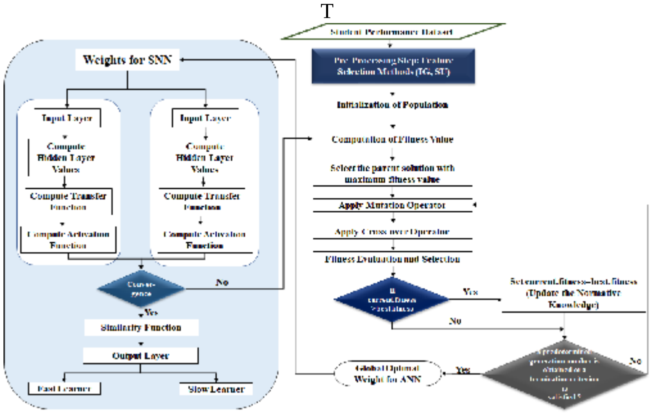

FIGURE 2: FLOWCHART OF THE PROPOSED CULTURAL ALGORITHM BASED SIAMESE NEURAL NETWORK (CA-SNN) CLASSIFICATION METHOD

Step by Step Procedure for Proposed Weight Optimization of SNN with CA

Input: Student Performance Dataset, Cultural Algorithm Operators, Number of hidden layers and its neurons.

Output: Class (Fast Learner or Slow Learner)

Step 1: Generate initial population of chromosomes.

Step 2: Evaluate the fitness function of the population. RMSE is considered as the fitness function.

Step 2.1: Select the parent solution with maximum fitness value. Decoding of the chromosome and building of SNN classifier based on the input data.

Step 2.2: Choosing those classifiers whose weight is bigger than λ to construct the classifier ensemble.

Step 2.3: Calculating the classification accuracy (i.e., fitness of the i -th chromosome) of the testing data using the generated classifier ensemble.

Step 2.4: Find the chromosome with highest fitness among the population.

Step 3: Are the optimization criteria met? If YES, go to step 9. If NO, go to step 4.

Step 4: Generate new population using the selection operator.

Step 5: Perform the crossover operator according to the crossover probability.

Step 6: Perform the mutation operator according to the mutation probability.

Step 7: Evaluate fitness of each new chromosome.

Step 7.1: Repeat step 2.1 to step 2.3 for each new chromosome.

Step 7.2: Find the chromosome with highest fitness and worst one.

Step 7.3: Store the best optimal solution using Normative Knowledge.

Step 7.4: Call influence function to update the belief space.

Step 7.5: Update the population space by using update function.

Step 8: Find the best chromosome during the evolution history and guarantee its survival to the next generation, i.e., comparing the chromosomes fitness for the generations, if the fitness of new chromosome is greater than old chromosome, then replace latest chromosome value with old one; otherwise replace new one with old value. Go to Step 3.

RESULT AND DISCUSSION

In this research work, the student performance dataset is considered from UCI repository [18]. From the dataset link, the student performance in the Mathematic subject is considered for this research work. Table 1 depicts the dataset features used in this research work.

Table 1: Student Performance Dataset and its feature values

Attribute Name	Values
School	student's school (binary: 'GP' - Gabriel Pereira or 'MS' - Mousinho da Silveira)
Sex	student's sex (binary: 'F' - female or 'M' - male)
Age	student's age (numeric: from 15 to 22)
Address	student's home address type (binary: 'U' - urban or 'R' - rural)
Famsize	family size (binary: 'LE3' - less or equal to 3 or 'GT3' - greater than 3)
Pstatus	parent's cohabitation status (binary: 'T' - living together or 'A' - apart)
Medu	mother's education (numeric: 0 - none, 1 - primary education (4th grade), 2 - 5th to 9th grade, 3 - secondary education or 4 - higher education)
Fedu	father's education (numeric: 0 - none, 1 - primary education (4th grade), 2 - 5th to 9th grade, 3 - secondary education or 4 - higher education)
Mjob	mother's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), 'at_home' or 'other')
Fjob	father's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), 'at_home' or 'other')
Reason	reason to choose this school (nominal: close to 'home', school 'reputation', 'course' preference or 'other')
Guardian	student's guardian (nominal: 'mother', 'father' or 'other')
Traveltime	home to school travel time (numeric: 1 - <15 min., 2 - 15 to 30 min., 3 - 30 min. to 1 hour, or 4 - >1 hour)

Studytime	weekly study time (numeric: 1 - <2 hours, 2 - 2 to 5 hours, 3 - 5 to 10 hours, or 4 - >10 hours)
Failures	number of past class failures (numeric: n if 1<=n<3, else 4)
Schoolsup	extra educational support (binary: yes or no)
Famsup	family educational support (binary: yes or no)
Paid	extra paid classes within the course subject (Math or Portuguese) (binary: yes or no)
Activities	extra-curricular activities (binary: yes or no)
Nursery	attended nursery school (binary: yes or no)
Higher	wants to take higher education (binary: yes or no)
Internet	Internet access at home (binary: yes or no)
Romantic	with a romantic relationship (binary: yes or no)
Famrel	quality of family relationships (numeric: from 1 - very bad to 5 - excellent)
Freetime	free time after school (numeric: from 1 - very low to 5 - very high)
Gout	going out with friends (numeric: from 1 - very low to 5 - very high)
Dalc	workday alcohol consumption (numeric: from 1 - very low to 5 - very high)
Walc	weekend alcohol consumption (numeric: from 1 - very low to 5 - very high)
Health	current health status (numeric: from 1 - very bad to 5 - very good)
Absences	number of school absences (numeric: from 0 to 93)
G1	first period grade (numeric: from 0 to 20)
G3	final grade (numeric: from 0 to 20, output target) (A – 16-20, B – 11-15, C – 6-10, D – 0-5)

Table 2 depicts the performance metrics used in this research work to evaluate the performance of the proposed CA-SNN, SNN, ANN and RF using original dataset and feature selection methods like Information Gain (IG) and Symmetrical Uncertainty (SU) processed datasets [19][20][21].

Table 2: Performance Metrics used in this research work

Metrics	Equation
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Sensitivity	$\frac{TP}{TP+FN}$
Specificity	$\frac{TN}{TN+FP}$
Precision	$\frac{TP}{TP+FP}$
False Positive Rate	1-Specificity
Miss Rate	1-Sensitivity
False Discovery Rate	1-Precision

Table 3 gives the Classification Accuracy (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets. Figure 3 depicts the graphical representation of the classification accuracy of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets. From the table 4 and figure 3, it is clear that the proposed CA-SNN classification gives more accuracy when it is compared other classification methods using original dataset and feature selection processed datasets.

TABLE 3: Classification Accuracy (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets

Feature Selection Methods	Classification Accuracy (in %) by Classification Techniques			
	Proposed CA-SNN	**SNN**	**ANN**	**RF**
Original Dataset	69.48	67.35	65.42	60.28
IG	78.57	76.65	75.62	63.24
SU	83.84	81.58	72.21	44.32

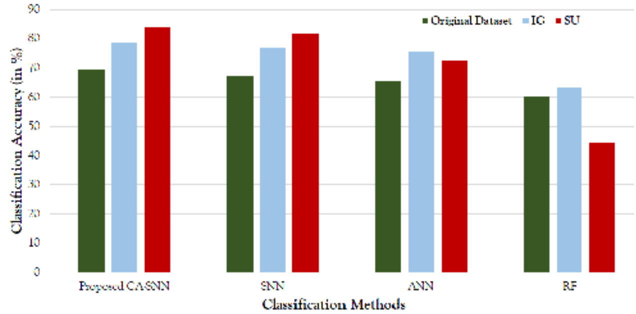

FIGURE 3: Graphical representation of the Classification Accuracy (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets

Table 4 gives the Sensitivity (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets. Figure 3 depicts the graphical representation of the classification accuracy of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets. From the table 4 and figure 4, it is clear that the proposed CA-SNN classification gives more sensitivity when it is compared other classification methods using original dataset and feature selection processed datasets.

TABLE 4: Sensitivity (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets

Feature Selection Methods	Sensitivity (in %) by Classification Techniques			
	Proposed CA-SNN	**SNN**	**ANN**	**RF**
Original Dataset	68.75	66.98	64.31	59.37
IG	75.57	72.35	62.95	61.72
SU	76.88	74.67	60.59	46.64

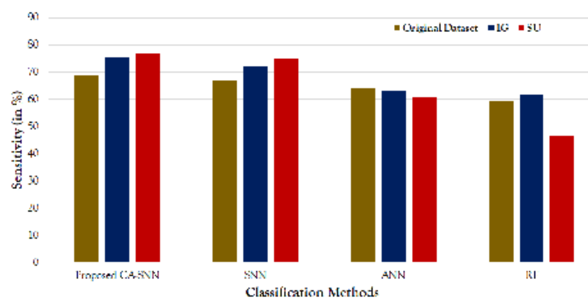

FIGURE 4: Graphical representation of the Sensitivity (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets

Table 5 gives the Specificity (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets. Figure 5 depicts the graphical representation of the Specificity of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets. From the table 5 and figure 5, it is clear that the proposed CA-SNN classification gives more Specificity when it is compared other classification methods using original dataset and feature selection processed datasets.

TABLE 5: Specificity (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets

Feature Selection Methods	Specificity (in %) by Classification Techniques			
	Proposed CA-SNN	SNN	ANN	RF
Original Dataset	68.94	65.65	63.42	58.46
IG	73.67	71.45	68.23	52.27
SU	76.25	72.41	65.35	36.91

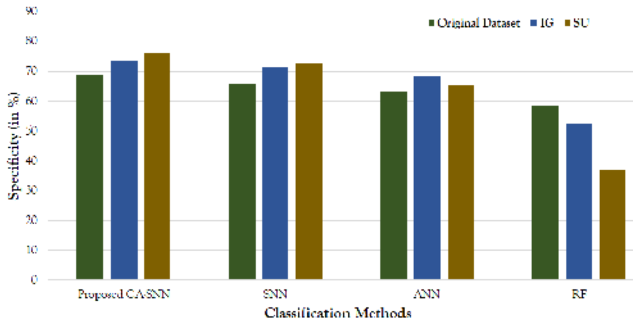

FIGURE 5: Graphical representation of the Specificity (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets

Table 6 gives the Precision (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets. Figure 6 depicts the graphical representation of the Precision of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets. From the table 6 and figure 6, it is clear that the proposed CA-SNN classification gives more precision when it is compared other classification methods using original dataset and feature selection processed datasets.

TABLE 6: Precision (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets

Feature Selection Methods	Precision (in %) by Classification Techniques			
	Proposed CA-SNN	SNN	ANN	RF
Original Dataset	65.28	62.35	61.33	58.88
IG	73.81	70.22	63.49	51.46
SU	74.58	70.79	58.13	48.54

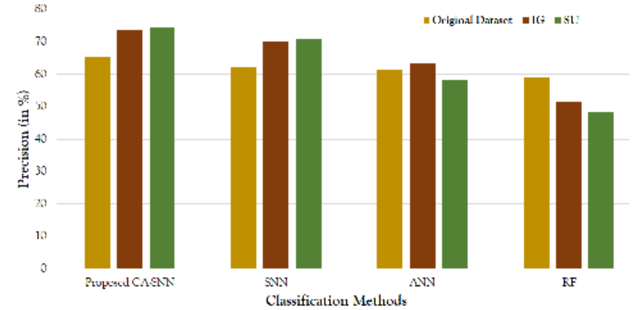

FIGURE 6: GRAPHICAL REPRESENTATION OF THE PRECISION (IN %) OF THE PROPOSED CA-SNN, SNN, ANN AND RF USING ORIGINAL DATASET, IG AND SU PROCESSED DATASETS

Table 7 gives the False Positive Rate (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets. Figure 7 depicts the graphical representation of the False Positive Rate (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets. From the table 7 and figure 7, it is clear that the proposed CA-SNN classification reduces the FPR when it is compared other classification methods using original dataset and feature selection processed datasets.

TABLE 7: False Positive Rate (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets

Feature Selection Methods	False Positive Rate (in %) by Classification Techniques			
	Proposed CA-SNN	SNN	ANN	RF
Original Dataset	31.06	34.35	36.58	41.54
IG	26.33	28.55	31.77	47.73
SU	23.75	27.59	34.65	63.09

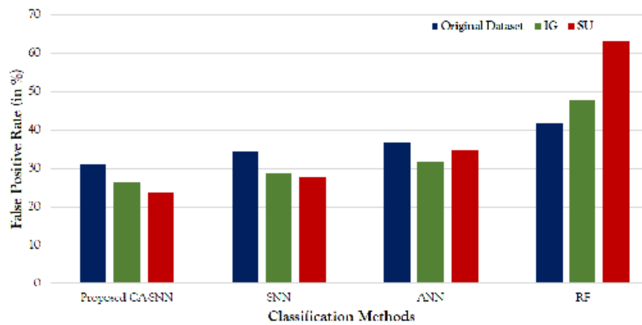

FIGURE 7: GRAPHICAL REPRESENTATION OF THE FALSE POSITIVE RATE (IN %) OF THE PROPOSED CA-SNN, SNN, ANN AND RF USING ORIGINAL DATASET, IG AND SU PROCESSED DATASETS

Table 8 gives the Miss Rate (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets. Figure 8 depicts the graphical representation of the Miss Rate (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets. From the table 8 and figure 8, it is clear that the proposed CA-SNN classification reduces the miss rate when it is compared other classification methods using original dataset and feature selection processed datasets.

TABLE 8: Miss Rate (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets

Feature Selection Methods	Miss Rate (in %) by Classification Techniques			
	Proposed CA-SNN	SNN	ANN	RF
Original Dataset	31.25	33.02	35.69	40.63
IG	24.43	27.65	37.05	38.28
SU	23.12	25.33	39.41	53.36

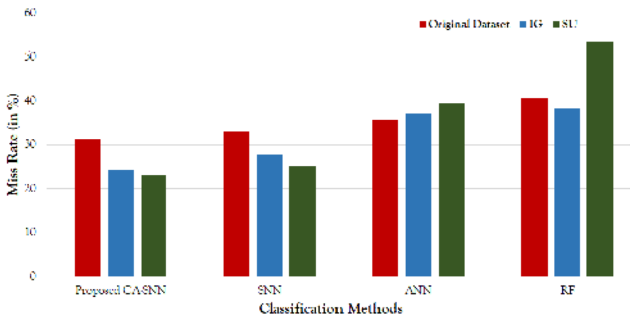

FIGURE 8: GRAPHICAL REPRESENTATION OF THE MISS RATE (IN %) OF THE PROPOSED CA-SNN, SNN, ANN AND RF USING ORIGINAL DATASET, IG AND SU PROCESSED DATASETS

Table 9 gives the False Discovery Rate (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets. Figure 9 depicts the graphical representation of the False Discovery Rate (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets. From the table 9 and figure 9, it is clear that the proposed CA-SNN classification reduces the FDR when it is compared other classification methods using original dataset and feature selection processed datasets.

TABLE 9: False Discovery Rate (in %) of the Proposed CA-SNN, SNN, ANN and RF using original dataset, IG and SU processed datasets

Feature Selection Methods	False Discovery Rate (in %) by Classification Techniques			
	Proposed CA-SNN	SNN	ANN	RF
Original Dataset	34.72	37.65	38.67	41.12
IG	26.19	70.22	36.51	48.54
SU	25.42	29.21	41.87	51.46

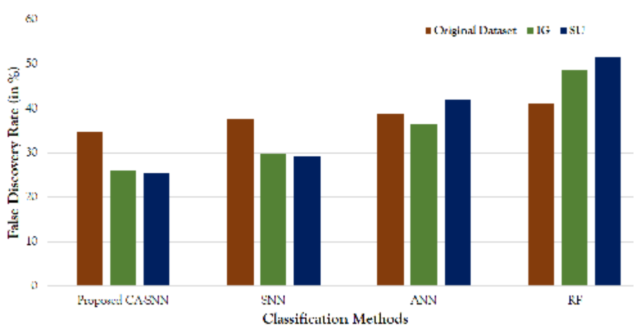

FIGURE 9: GRAPHICAL REPRESENTATION OF THE FALSE DISCOVERY RATE (IN %) OF THE PROPOSED CA-SNN, SNN, ANN AND RF USING ORIGINAL DATASET, IG AND SU PROCESSED DATASETS

CONCLUSION

Due to the huge amount of data in educational databases, predicting the performance of students has become more difficult. The shortage of an established framework for evaluating and tracking the success of students also isn't currently being considered. In this paper, an optimization based deep learning classification method is proposed to improve the accuracy in the prediction of slow and fast learners among the students. Cultural Algorithm is used to optimize the weights of the Siamese Neural Network (SNN). From the results obtained, it is shown that the proposed CA-SNN classification method performs better in terms of accuracy, sensitivity, specificity and precision. It is also reduced the error rates like FPR, miss rate and FDR with the feature selection processed datasets than the other classifiers like SNN, ANN and RF.

REFERENCES

[1] Hashim, Ali Salah, Wid Akeel Awadh, and Alaa Khalaf Hamoud. "Student performance prediction model based on supervised machine learning algorithms." *IOP Conference Series: Materials Science and Engineering*. Vol. 928. No. 3. IOP Publishing, 2020.

[2] Rastrollo-Guerrero, Juan L., Juan A. Gómez-Pulido, and Arturo Durán-Domínguez. "Analyzing and predicting students' performance by means of machine learning: A review." *Applied sciences* 10.3 (2020): 1042

[3] Alshabandar, Raghad, et al. "Students performance prediction in online courses using machine learning algorithms." *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2020.

[4] Giannakas, F., et al. "A deep learning classification framework for early prediction of team-based academic performance." *Applied Soft Computing* 106 (2021): 107355.

[5] Sokkhey, Phauk, and Takeo Okazaki. "Hybrid machine learning algorithms for predicting academic performance." *Int. J. Adv. Comput. Sci. Appl* 11.1 (2020): 32-41.

[6] Waheed, Hajra, et al. "Predicting academic performance of students from VLE big data using deep learning models." *Computers in Human behavior* 104 (2020): 106189.

[7] Pallathadka, Harikumar, et al. "Classification and prediction of student performance data using various machine learning algorithms." *Materials Today: Proceedings* (2021).

[8] Bhutto, Engr Sana, et al. "Predicting students' academic performance through supervised machine learning." 2020 International Conference on Information Science and Communication Technology (ICISCT). IEEE, 2020.

[9] Ha, Dinh Thi, et al. "An empirical study for student academic performance prediction using machine learning techniques." *International Journal of Computer Science and Information Security (IJCSIS)* 18.3 (2020).

[10] Rastrollo-Guerrero, Juan L., Juan A. Gómez-Pulido, and Arturo Durán-Domínguez. "Analyzing and predicting students' performance by means of machine learning: A review." *Applied sciences* 10.3 (2020): 1042.

[11] Lin, Yongzheng, et al. "Machine Learning-Based Classification of Academic Performance via Imaging Sensors." *IEEE Sensors Journal* 21.22 (2020): 24952-24958.

[12] Kechid, Amine, and Habiba Drias. "Cultural coalitions detection approach using GPU based on hybrid Bat and Cultural Algorithms." *Applied Soft Computing* 93 (2020): 106368.

[13] Bromley, Jane, et al. "Signature verification using a" siamese" time delay neural network." *Advances in neural information processing systems* 6 (1993).

[14] Wang, Juan, et al. "A multi-resolution approach for spinal metastasis detection using deep Siamese neural networks." *Computers in biology and medicine* 84 (2017): 137- 146.

[15] Huixian, Jiang. "The analysis of plants image recognition based on deep learning and artificial neural network." *IEEE Access* 8 (2020): 68828-68841.

[16] Muminov, Bakhodir, et al. "Localization and Classification of Myocardial Infarction Based on Artificial Neural Network." *2020 Information Communication Technologies Conference (ICTC)*. IEEE, 2020.

[17] Bhattacherjee, Aindrila, et al. "Classification approach for breast cancer detection using back propagation neural network: a study." *Deep learning and neural networks: concepts, methodologies, tools, and applications*. IGI Global, 2020. 1410-1421.

[18] https://archive.ics.uci.edu/ml/datasets/student+performance

[19] Poornappriya, T. S., and M. Durairaj. "High relevancy low redundancy vague set- based feature selection method for telecom dataset." Journal of Intelligent & Fuzzy Systems Preprint: 1-18.

[20] Durairaj, M., and T. S. Poornappriya. "Choosing a spectacular Feature Selection technique for telecommunication industry using fuzzy TOPSIS MCDM." International Journal of Engineering & Technology 7.4 (2018): 5856-5861.

[21] Durairaj, M., and T. S. Poornappriya. "Why Feature Selection in Data Mining Is Prominent? A Survey." International Conference on Artificial Intelligence, Smart Grid and Smart City Applications. Springer, Cham, 2019.