Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 5 (2023)

Prediction Analysis of Students' Performance on the Hybrid EDA-SVR Model

[1]M.A. Arul Rozario , [2]R.GunaSundari

[1]Department of Computer Science Karpagam Academy of Higher Education Coimbatore 641021, TamilNadu, India
[2] Department of Computer Applicatios. Karpagam Academy of Higher Education Coimbatore 641021,TamilNadu, India

Abstract: The performance of the students is a significant aspect in determining how well a college's instructors are doing. This report's objective is to suggest a new, clever method for predicting students' performance using support vector regression (SVR) that is improved by a dual algorithm (EDA). To the best of our knowledge, there aren't many papers published that have been established to forecast students' performance based on student behaviour, therefore the novelty of this investigation is to provide a fresh hybrid intelligent strategy in this area. The EDA-SVR scenario explicitly exceeded the other approaches, based on the data, by achieving smaller mean square error (MSE). In other words, EDA-SVR, with an MSE of 0.0091, performs better than DT, SVR, ANN, and PSO-SVR, which all have MSEs of 0.0315, 0.0242, 0.0232, and 0.0116, respectively. Other parameter estimation techniques, such as the direct method was evaluated, grid selection strategy, GA, FA, and PSO, are utilised in a comparison research to examine the effectiveness of EDA. The findings demonstrate that the EDA algorithm may successfully avoid local optima and blindness search, as well as accelerate resolution to the optimal method.

Keywords: students Performance, EDA-SVR Model, Smaller Mean Square Error (MSE), Blindness Search

1. Introduction

The use of computers in teaching has increased significantly in recent years. It has always been crucial to forecast students' academic achievement in the classroom. The primary benchmark for assessing students' degree of knowledge is still their achievement, which is also a significant indicator of how well schools and teachers are able to impart information. The expansion of teachers and the number of students is out of step with the increase in enrollment, which has an impact on the effectiveness of instruction and the academic achievement of the students. Therefore, it is crucial to correctly predict pupils' performance when managing education. The forecasting of students' performance enables teachers to modify their instruction and help pupils work better.

Two methods can be made out of the popular evaluation prediction methods. Creating mathematical tools like multiple regression analysis and sparse factor analysis is the initial step. By using a linear regression model predicted the performance of the pupils (Sravani and Bala 2020). The second is founded on techniques for data-driven predictive analysis, including logistic regression (LR) (Janssens et al 2005), Naive Bayes (Rish, Irina 2001), decision trees (Quinlan and Ross 1986) artificial neural networks (Agatonovic-Kustrin and Rosemary Beresford 2005) support vector regression [SVR], and soon. These techniques merely extract the model from the pertinent data, not the involvement of experts. The evaluation of popular machine learning techniques is shown in Table 1. The time and spatial complexity of popular machine learning techniques is shown in Table 2. Table 2 demonstrates that SVR works well in both time and spatial difficulty in random studies. To forecast the pupils' academic success(Borkar and Rajeswari 2014) used educational data mining and artificial neural networks. When predicting student achievement, (Ghorbani and Ghousi 2020) examined the effectiveness of different machine learning techniques such random forest, k-nearest neighbour, support vector machine, and decision tree. The productivity of the pupils is determined by a range of studying habits, and it differs widely from person to person. Consequently, the conventional statistical model could occasionally be unsuccessful. The data-driven method makes an effort to directly anticipate student performance using data on their behaviour.Just enough achievement data must be gathered in order to create a data-driven model for predicting student achievement. Second, SVR is more suited than other standard data-driven models for assessing the student behaviour data from Tables 1 and 2. SVR is chosen to forecast the students' achievement in this paper as a result.

Table 1: Comparison of common machine learning methods.

Model	*Advantages*	*Disadvantages*
Logistic regression	1. Simple calculation and fast speed 2.Avoid overfitting through regularization	The performance is poor when faced with the multivariate or nonlinear decision boundary
Naive Bayes	Perform well on small-scale data	Very sensitive to the expression of input data
Decision tree	1. Able to apply to samples with missing attribute values 2. Strong robustness to outliers	Easy to overfit
Artificial neural networks	Perform well on nonlinear data	1. Long training time 2. The computational complexity is proportional to the network Complexity
Support vector regression	1. Strong generalization ability 2. Can apply to high-dimensional nonlinear data 3. Low computational complexity	Sensitive to the selection of parameters and kernel function

Table 2: The time and space complexity of common machine learning methods.

Model	*Time complexity*	*Space complexity*
Logistic regression	$O(n * m)$	$O(m)$
Naive Bayes	$O(n * m * c)$	$O(m * c)$
Decision tree	$O(n * log(n) * d)$	$O(p)$
Artificial neural networks	$O(t * \sum n_1 n_2 + n_2 n_3 + \cdots)$	$O(t * \sum (n_1 n_2 + n_2) + (n_2 n_3 + n_3 + \cdots)$
Support vector regression	$O(n^2)$	$O(n_{sv})$

2. Related Work

More researchers are now choosing the SVR model's parameters using an intelligent optimization technique. A novelngraph-based ELM (G-ELM) for the identification of unbalanced epileptic EEG signals was proposed (Zhou et al. 2021). The suggested approach makes use of graph theory to build a relationship graph between samples in accordance with dataset. Then, a model that combines the relationship graph and ELM is created; it retains the ELM's quick learning and strong generalisation characteristics and boosts classification results. The efficiency and scalability of the suggested technique were proved through tests on a genuine unbalanced epileptic EEG dataset.

In order to diagnose atrophic gastritis (AG), (Zhang et al. 2021) coupled deep learning-based image recognition techniques with serological specific markers. This work may serve as a model for future clinical applications of artificial intelligence identification technology.

By anticipating the starting positions of shapelets of high quality, we provide a pruning approach to decrease the number of shape lets in Yan et al. In order to keep the diversity of shapelets, a novel approach for shapelet selection is also put out to exclude similar shapelets. Last but not least, the experimental outcomes on

16 benchmark datasets demonstrate that the proposed strategy outperforms state-of-the-art for early categorization on time series.

TBOPE, a brand-new ensemble approach built on SAX and presented (Bai et al. 2021), is based on multi-feature database modeling and clustering algorithms. To be more precise, we extract the mean feature and trend characteristic of time series first. Second, we build several single classifiers using histograms of two different feature types based on the Bag-of-Feature method.

A learning algorithm based on functional- gradient boosting techniques for logistic regression was created and an empirical evaluation on standard data sets shown that it was superior to previous learning techniques for LR (Ramanan et al 2020).

Support vector regression (SVR), an artificial intelligence technique is used to forecast the vertical load capacity of driven piles in cohesionless soils (GA) (Luo et al 2021). To the best of our knowledge, no research has previously been conducted using the GA-SVR model to forecast the vertical load capacity of driven piles over a range of timescales. The novelty of our study is the creation of a fresh hybrid intelligent methodology in this area.

In their study, (Huang et al 2020) propose a multi-objective optimization methodology based on artificial intelligence to make it easier to identify the best mix design for SFRC. From earlier literature, a sizable dataset was compiled, including 299 instances for the uniaxial compressive strength (UCS) test and 269 instances for the flexural strength (FS) test. In order to forecast UCS and FS for SFRC, a support vector regression (SVR) model was used. A sensitivity analysis was done to determine the impact of the inputs on the output variables for the algorithms as well as the hyper parameters of SVR models, which were tuned using a firefly algorithm (FA).

This area is investigated by Liu et al (2021) who also offer a solution. First, the theory and training of the PSO-SVR model are examined in this study. Based on this, this research also investigates the correlation between output frequency difference data and the corresponding yarn tension applied to the SAW yarn tension sensor. The PSO-SVR model is then trained and used to forecast the output tension of the sensor using the frequency difference data as input and corresponding tension as output.

It has been demonstrated that intelligent algorithms are efficient in resolving parameter optimization issues. It provides significant robustness to the many types of problems and is independent of the problem's unique area. An efficient global optimization algorithm is the DA algorithm. Individuals continue to develop and approach the ideal answer to the problem after the initial combat between them. As a result, in this paper, the

DA algorithm is chosen to jointly optimise the SVR model parameters and the features.

3. Methodology

This section will introduce the necessary background knowledge and the proposed model. First, the SVR model and DA algorithm are elaborated. Next, the proposed EDA-SVR model is described in detail.

3.1 Support Vector Regression.

SVR is the application of a support vector machine (SVM) in regression learning.Suppose $(x1,y1),\dots,(xn,yn), xi \in Rm, yi \in R,$ are the sample data. Such a linear function, namely SVR function, is as follows:

$$f(x) = \omega^T \varphi(x) + b \quad (1)$$

where $\omega = (\omega_1, \omega_2, \dots, \omega_m)^T$is a vector normal to the maximum-margin hyperplane and $b$is the deviation. $\varphi(.)$is a nonlinear mapping.

The problem can be treated as the following optimization problem:

$$\min \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}(\leq \xi_i^+ + \xi_i^-)$$

$$s.t\begin{cases}\omega^T\varphi(x) + b - y_i \leq \varepsilon + \xi_i^+ \\ y_i - \omega^T\varphi(x) - b \leq \varepsilon + \xi_i^+ \\ \xi_i^+, \xi_i^- \geq 0, \quad i = 1,2,\dots,l.\end{cases} \quad (2)$$

where C is the regularization factor and ξ_i^+ and ξ_i^- are slack variables representing lower and upper constraints on the outputs of the model. ε is a positive constant. Errors are calculated only if the deviation between $f(x)$ and y_i is greater than ε.,

The Kalush–Kun–Tuck (KKT) optimum conditions are both necessary and sufficient for the aforementioned issue, which is a quadratic problem with linear constraints. The solution is a linear combination of a subset of sample points termed support vectors (s.v.) as follows:

$$\omega = \sum_{S.V}\beta_i\varphi(x_i) \Longrightarrow f_{\omega,b}(x) = \sum_{S.V}\beta_i\langle\varphi(x_i),\varphi(x)\rangle + b \quad (3)$$

Let $\kappa(xi,xj) = \langle\varphi(xi),\varphi(xj)\rangle$, which is called the kernel function. It can map points from low- to high-dimensional space. Then, equation (3) can be rewritten as follows:

$$\omega = \sum_{S.V}\beta_i\varphi(x_i) \Longrightarrow f_{\omega,b}(x) = \sum_{S.V}\beta_i k\langle x_i,x\rangle + b \quad (4)$$

Kernel selection is one of the most important methods for improving SVR's capabilities. The radial basis function is used in this paper, as seen in the equation:

$$K(x_i,x_j) = e^{-\sigma\|x_i-x_j\|^2} \quad (5)$$

The prediction accuracy of the SVR model depends on the good settings of the hyperparameters C and ε and the kernel parameter σ. Therefore, the selection of the parameters is an important issue. Next, we will introduce the EDA algorithm to optimize SVR parameters.

3.2 DA Algorithm

Duelist algorithm (DA) is a new algorithm based on a genetic algorithm proposed by Biyanto [19] from the perspective of human combat and learning ability.The process of the DA algorithm is shown in Figure 1.

Encoding. In this paper, the encoding of the DA algorithm is composed of parameters (C, ε, σ) and feature subsets, as shown in Figure 2

$b_C^1 \sim b_C^{n_C}$, $b_\varepsilon^1 \sim b_\varepsilon^{n_\varepsilon}$, $b_\sigma^1 \sim b_\sigma^{n_\sigma}$ and $b_f^1 \sim b_f^{n_f}$ are the binary strings of parameters C,ε,σ, and features, respectively. n_C, n_ε, n_σ, and n_f are the numbers of binary digits of C, ε, σ, and features, respectively.

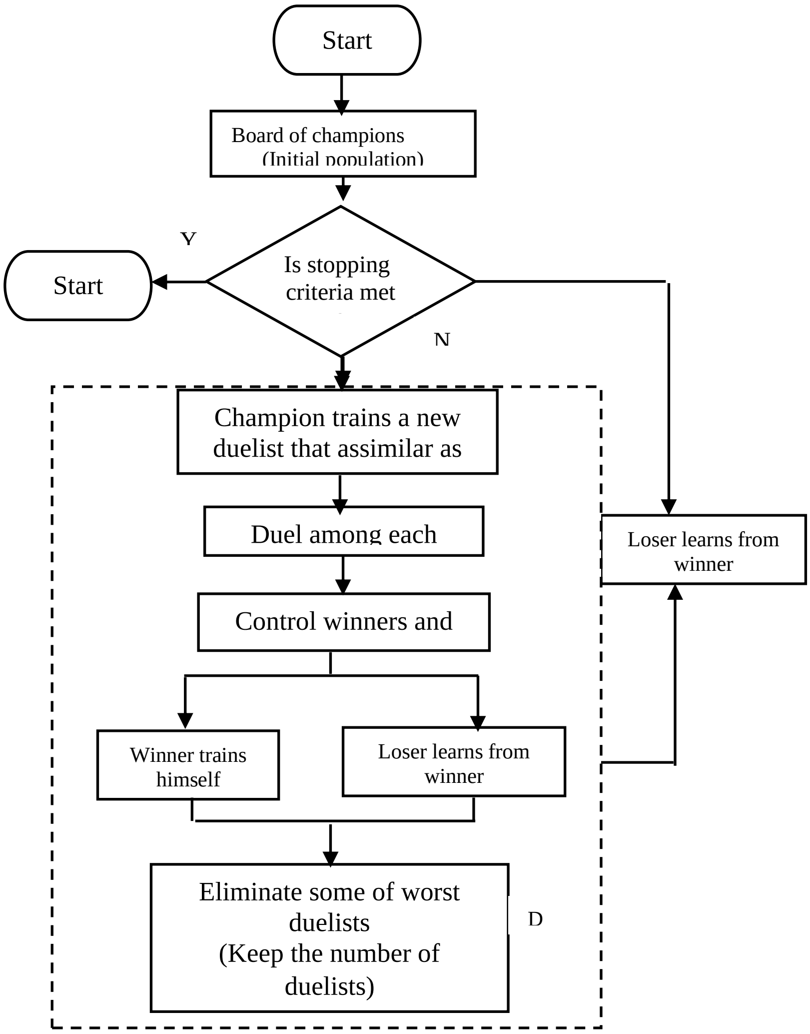

Fig 1: Flowchart of DA algorithm.

3.3 Fighting Capability Function

In this paper, we take the mean squared error (MSE) as the fighting capability. Let yi be the observed values and yi be the predicted values, then

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (6)$$

The number of samples in this case is n. A lower MSE score reflects a stronger fighting capacity.

Scheduling duel times between duelists. By pitting duelists against each other one-on-one, the DA algorithm optimises the solution. Algorithm 1 displays the duel process' pseudocode.

Duelist's Development Both the loser and the winner must amp up their fighting skills in this step. Algorithm 2 shows the pseudocode of the duelist's improvement.

ALGORITHM 1: Purpose of the winner and the loser.

```
(1) Duelist X and Y, Luk_coeff
(2) FC = Fighting capability; LC = Luk coeff
(3) X(Luk) = X(FC) ∗ (LC + (ran(0, 1) ∗ LC));
(4) Y(Luk) = Y(FC) ∗ (LC + (ran(0, 1) ∗ LC));
(5) if ((X(FC) + X(Luck)) > = (Y(FC) + Y(Luk)))
(6) X(Winner) = 1;
(7) Y(Winner) = 0;
(8) else
(9) X(Winner) = 0;
(10) Y(Winner) = 1;
(11) end if
```

ALGORITHM 2: Duelist's Development.

```
(1) Duelist X and Y, Duelist_len, Prob_inn; Prob_learn
(2) if X(Winner) = 1
(3) for i = 1:(Duelist_len)
(4) r = ran(0,1)
(5) if r < Prob inn
(6) if X[i] = 1
(7) X[i] = 0
(8) else X[i] = 1
(9) end if
(10) end if
(11) end for
(12) else
(13) fori = 1: (Duelist length)
(14) r = ran(0, 1)
(15) if r < Prob_learn
(16) Y[i] = X[i]
(17) end if
(18) end for
(19) end if
```

3.4 EDA-SVR Model

After in-depth analysis, it is found that the DA algorithm has four shortcomings.

(1) A random number generator produces the initial solution's value. The original population's uniform distribution and individual quality cannot be guaranteed by the random procedure. Some of the answers are very far from being the best option.

(2) After thoroughly examining the DA algorithm's entire workflow, we can say that the luck coefficient significantly affects the algorithm's performance. The new person is more random the higher the luck coefficient. As a result, the solution's fitness fluctuation grows and the rate of convergence to the ideal solution slows down. Weaker new individuals will have more randomness, which slows down the process of finding the best solution. Hence, the luck parameter setting is critical to the algorithm's success.

(3) Following the combat, each duelist is divided into winners and losers. Each loser in a battle receives instruction from the winner in order to improve their fighting skills, while the winners develop independently.

As a result, it is clear that the loser's improvement is dependent on two people exchanging information, which will cause the algorithm's convergence to occur slowly.

(4) The DA method is susceptible to local optimization and has limited search accuracy, just like other swarm intelligent optimization techniques.

The DA algorithm's shortcomings have led to the following improvements in this paper:

(1) The chaotic sequence is used to launch the population. By utilising chaos, it not only broadens the diversity of the population but also preserves the fundamental randomness of the optimization process. Numerous mathematical ideas can be used to create chaotic sequences. In this paper, a logistic equation is used to generate chaotic sequences as seen below:

$$x(t+1) = \mu x(t)\big(1 - x(t)\big), t = 1,1,2, \ldots, \quad (7)$$

where μ is the control parameter. When $0 < x(0) < 1$ and $\mu = 4$, the logistic equation is in a complete chaotic state. In this case, $x(t)$ is chaotic and in the interval $(0,1)$. Given the initial value $x(0) \in (0,1)$, the time series $x(1), x(2), \ldots$, can be generated.

(2) The statistical principle states that there are more opportunities to look for additional ideal solutions close to the best one. In other words, we can start by slightly increasing the luck coefficient. The solution will thus be more illogical, making it simpler to identify the optimum approach. When it is near the optimum answer, a little luck coefficient enables the algorithm to look for further optimal solutions in its vicinity. The adaptive luck coefficient c is defined as a result of the analysis that has just been don.:

$$c = \frac{i_{max}}{\lambda(i_{max} + i + 1)} \quad (8)$$

Here, $i_{max}$is the total number of iterations, i is the current iteration number, and λ is the adjustment coefficient of step length, which is determined according to the feasible regions of different optimizationproblems.

(3) Each loser in a battle receives training by observing one of the winners in order to become better. One uses the roulette approach to decide who will win and who will lose..

(4) The chaotic sequence search method is used to generate the neighbourhood solutions. The algorithm may deviate from the local optimum due to the ergodicity and randomness of the chaotic variables. This improves the global searching capability of the algorithm. First, based on the current ideal position, equation (7) generates the chaotic sequence. Then the dual position is replaced with the ideal position of the chaotic sequence. The aforementioned processes can produce neighbourhood solutions of the local optimal solution during the iteration, allowing the current solution to diverge from the local optimal solution.

The four strategies are intended to enhance the algorithm in distinct, non-overlapping steps. From the outset, strategy (a) is a step forward. In approach (b), the lucky coefficient changes adaptively, hastening the convergence of the algorithm to the ideal state. The goal of approach (c) in the duelist improvement step is to increase the variety of solutions. If method (d) is applied, the newly generated solution can depart from the local optimum. The algorithm's forecast accuracy and rate of convergence to the ideal solution are both ensured by the four strategies outlined above.

The improved DA method will then be used to optimise the SVR parameters. Figure 2 shows the flowchart for the hybrid EDA-SVR model developed in this study.

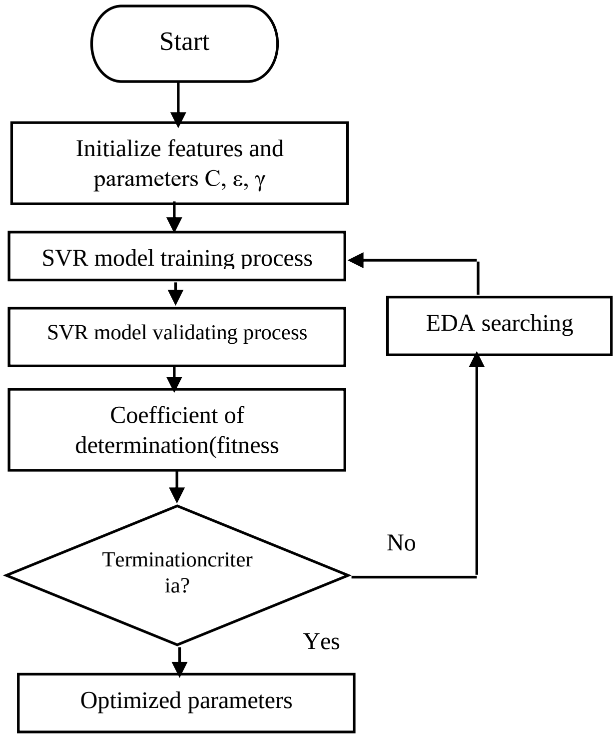

Fig 2: Flowchart of the EDA-SVR model.

EDA: Enhanced Duelist algorithm SVR : Support Vector Regression

4. Experimental Study

The arithmetic performance metrics of 240 kids from five classes in grade two at a college serves as the research object in this paper. The remaining 60 samples will be used as testing data, and 180 samples will be utilised as training data. The 18 features identified in each sample are listed in Table 3.

Table 3: The features of each student sample

Category	*No.*	*Feature*	*Feature*
	1	Sex	{1,2}
	2	Native place	{1,2,3,4,5}
Basic information	3	Semester	{1,2}
	4	Education level of parents	{1,2,3,4,5}
	5	Work as a student cadre	{0,1}
Interest	6	Interest in the course	{1,2,3,4}
	7	The degree of keeping up with the class	{1,2,3,4}
	8	Learning initiative	{1,2,3,4}
Behavior in class	9	Number of absenteeism	{0,25}
	10	Frequency of distraction	{5,47}
	11	Average number of hands raised	{0,12}
	12	Number of questions answered	{0,9}
	13	Number of assignments submitted	{0,20}
	14	Number of interactions between teachers and students	{0,17}
	15	Number of group discussions attended	{0.17}
Behavior outside class	16	Time of study this course outside the class	{5,65}
	17	Study extracurricular material time	{0.12}
	18	Online viewing time	{0,60}

The PC used for all experiments has an Intel Core i5-1035 8 GB processor, Windows 10 as its operating system, and PyCharm 2021.1 as its programming environment for Python 3.6.6. The parameter settings are presented in Table 4. Table 4 show the outcomes of predicting the students' performance using the EDA-SVR model.

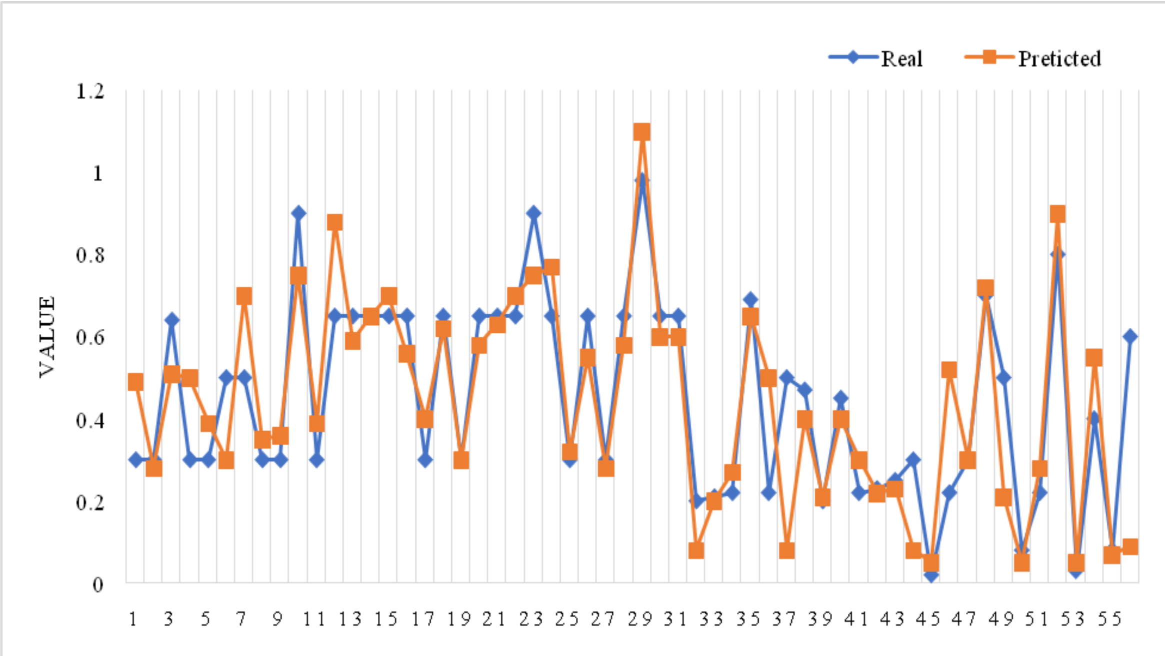

Fig 3: The prediction results of students' performance based on the EDA-SVR model

IDA-SVR:Enhanced Duelist algorithm - Support Vector Regression

Table 4: The results of EDA-SVR in students' performance prediction.

Model	*MSE*	*Selected feature*
EDA-SVR	0.0091	7

An study of the SVR-based prediction model for student performance shows that there are 153 support vectors used in the model. It can be demonstrated that in the sample set of 240 students, a performance prediction can be made using only the data from 153 students. Due to the length of the topic, two examples are presented below for analysis.

Take the example of a student with a performance in the 90th percentile represented by a 36-index support vector. The top five crucial learning behaviour variables in the model are the number of assignments turned in, the typical number of hands raised, the amount of time spent studying for this course outside of class, the frequency of distractions, and the amount of absenteeism. The proper feature weight vector for this sample is [1.0335, 0.8327, 0.8133, 0.7415, 0.5448]. It is evident that the number of assignments turned in significantly affects the final grade. We'll examine a student who has experienced academic difficulty next.

We'll now examine the support vector that contains the number 228, which stands for a student who had a score of 72 out of 100. The top five crucial learning behaviour factors in the model are work as a student cadre, absenteeism, distraction frequency, number of assignments turned in, and the average number of hands raised. The proper feature weight vector for this sample is [1.2814, 0.9738, 0.7122, 0.5415, 0.4388]. This student is less significant than the student with index number 36 in terms of the quantity of assignments turned in and the typical number of hands raised.

Therefore, the two qualities mentioned above have a significant effect on pupils' academic progress. Additionally, absence is one of the main reasons why students perform poorly in class. The teaching recommendations that follow are based on the model analysis that came before. Teachers should look into specific techniques to improve student performance in light of the findings of the experimental research and the current context. In order to increase students' motivation and learning engagement during the teaching process and hence boost efficiency, teachers should employ appropriate teaching tactics.

Four rival models—basic SVR, PSO-SVR, decision tree (DT), and artificial neural networks—are contrasted with the EDA-SVR model to show its superiority (ANN). The created ANN model has three hidden layers with the ReLU activation function. The decision tree algorithm's maximum decision tree depth is set to 5. Table 6 provide an overview of the results of the suggested approach and the comparison methods for the prediction of student performance problem. The proposed model outperforms the listed comparable approaches in terms of prediction accuracy.

Note: Pop is the population size of the intelligent algorithm, and iter denotes the quantity of times the intelligent algorithm has been iterated. According to Table 5, the proposed EDA-SVR model has the highest time complexity and running time, which suggests that it trades speed for accuracy. We plan to investigate ways to shorten runtime and temporal complexity in future research. The duration could be further reduced, first, by looking into more computationally effective SVR algorithms and a quicker parameter adjusting mechanism. In order to improve learning outcomes and reduce model processing costs, parallelization techniques and methodologies are also worth researching and employing.

Table 5: Comparison of performance between EDA-SVR and other models.

Model	*MSE*	*Time Complexity*	*Running Time(s)*
SVR	0.0242	$O(n^2)$	0.32
PSO-SVR	0.0116	$O(pop * iter * n^2)$	89.67
DT	0.0315	$O(n * \log(n) * d)$	1.3
ANN	0.0232	$O(t * \sum n_1 n_2 + n_2 n_3 + \cdots)$	3.48
EDA-SVR	0.0091	$O(pop * iter * n^2)$	102.21

SVR : Support Vector Regression
PSO-SVR : Particle swarm optimization - Support Vector Regression
DT : Decision Tree
ANN : Artificial Neural Network
EDA-SVR : Enhanced Duelist algorithm - Support Vector Regression

5. Conclusion

This research suggests a hybrid EDA-SVR model to forecast student performance. The main contributions of this study can be summed up as follows: (1) A novel intelligent strategy is put forth to forecast students' academic achievement based on their behaviour using support vector regression. Several of the experiments make use of students' arithmetic test results. The experimental results show that the suggested model does a great job of overcoming the difficulty of predicting student performance. (2) The enhanced duel technique is applied to SVR's kernel parameters optimization and feature selection. The EDA algorithm may successfully avoid local optima and blindness search while greatly accelerating convergence to the ideal solution as compared to other parameter optimization techniques. The method put out in this study aims to address the problem of predicting students' performance. However, it can be applied to address issues in different fields. Due to the fact that the suggested hybrid approach is essentially a labelled small sample data prediction technique. It is applicable to any industry that meets the aforementioned requirements, including forecasting economic indicators, environmental indicators, aberrant ECG signal identification, circuit failure diagnostics, and more.

The proposed model performs better than many others, yet it still has certain drawbacks. The updated DA algorithm is insecure to start with. The effects of different initial values on the outcomes will vary, for instance, if the initial values of the parameters to be optimised are picked at random. Furthermore, even though the enhanced DA algorithm supports global search, it cannot ensure that it will eventually converge to the best solution worldwide. Second, compared to other algorithms, SVR can produce noticeably superior results with a limited sample training set. However, when the sample dimension is big, the SVR's time complexity grows, considerably decreasing the predictor's performance.Third, the improved DA algorithm optimises the SVR parameters by training individuals on the training set and assessing their performance on the testing set. There are more optimization iterations when the accuracy is higher. In other words, the suggested method significantly trades off speed for accuracy. To alleviate the aforementioned restrictions, our findings can be broadened in the following future research directions.

Recommendations

As computer technology develops, neural networks can accommodate more layers, and deep learning techniques now perform better than machine learning in many fields. The objective function, constraint conditions, and kernel function of the SVR model must be enhanced based on the problem itself to increase SVR performance.

References

[1] Sravani, Boddeti, Myneni Madhu Bala 2020. "Prediction of student performance using linear regression." In 2020 *International Conference for Emerging Technology (INCET*), 1-5. IEEE, 2020.

[2] Janssens A, Cecile JW, Yazhong Deng, Gerard JJM Borsboom, Marinus JC Eijkemans, J. Dik F. Habbema, Ewout W. Steyerberg 2005. "A new logistic regression approach for the evaluation of diagnostic test results." *Medical decision making* 25 (2): 168-177.

[3] Rish, Irina 2001. "An empirical study of the naive Bayes classifier." In IJCAI 2001 workshop on empirical methods in artificial intelligence, 3(22): 41-46.

[4] Quinlan J. Ross 1986. "Induction of decision trees." *Machine learning,* 1(1): 81-106.

[5] Agatonovic-Kustrin S, Rosemary Beresford 2000. "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research." *Journal of pharmaceutical and biomedical analysis,* 22 (5): 717-727.

[6] Suchita, Borkar, K. Rajeswari 2014. "Attributes Selection for Predicting Students' Academic Performance using Education Data Mining and Artificial Neural Network." *International Journal of Computer Applications*, 86(10):2-8.

[7] Ghorbani, Ramin, Rouzbeh Ghousi 2020. "Comparing different resampling methods in predicting students' performance using machine learning techniques." *IEEE Access,* 8: 67899-67911.

[8] Zhou, Jie, Xiongtao Zhang, Zhibin Jiang 2021. "Recognition of imbalanced epileptic EEG signals by a graph-based extreme learning machine." Wireless Communications and Mobile Computing

[9] Zhang, Jianhai, Jianhong Yu, Suna Fu, Xinhua Tian 2021. "Adoption value of deep learning and serological indicators in the screening of atrophic gastritis based on artificial intelligence." *The Journal of Supercomputing,* 77(8): 8674-8693.

[10] Yan, Wenhe, Guiling Li, Zongda Wu, Senzhang Wang, Philip S. Yu 2020. "Extracting diverse-shapelets for early classification on time series." *World Wide Web,*23(6): 3055-3081.

[11] Bai, Bing, Guiling Li, Senzhang Wang, Zongda Wu, and Wenhe Yan 2021. "Time series classification based on multi-feature dictionary representation and ensemble learning.*Expert Systems with Applications* 169: 114162.

[12] Ramanan, Nandini, Gautam Kunapuli, Tushar Khot, Bahare Fatemi, Seyed Mehran Kazemi, David Poole, Kristian Kersting, Sriraam Natarajan 2021. "Structure learning for relational logistic regression: An ensemble approach." *Data Mining and Knowledge Discovery*, 35(5): 2089-2111.

[13] Luo, Zhenyan, Mahdi Hasanipanah, Hassan Bakhshandeh Amnieh, Kathirvel Brindhadevi, M. M. Tahir 2021. "GA-SVR: a novel hybrid data-driven model to simulate vertical load capacity of driven piles.*Engineering with Computers,* 37(2): 823-831.

[14] Huang, Yimiao, Junfei Zhang, Foo Tze Ann, and Guowei Ma 2020. "Intelligent mixture design of steel fibre reinforced concrete using a support vector regression and firefly algorithm based multi-objective optimization model. *Construction and Building Materials,* 260: 120457.

[15] Liu, Shoubing, Peng Xue, Jinyan Lu, Wenke Lu 2021."Fitting analysis and research of measured data of SAW yarn tension sensor based on PSO–SVR model." *Ultrasonics,* 116: 106511.

[16] Sun, Yuting, Shifei Ding, Zichen Zhang, Weikuan Jia 2021. "An improved grid search algorithm to optimize SVR for prediction." *Soft Computing,* 25(7): 5633-5644.

[17] Zhang, Huan, Liangxiao Jiang, Liangjun Yu 2021."Attribute and instance weighted naive Bayes.*Pattern Recognition*, 111: 107674.

[18] Schidler, André, Stefan Szeider 2021. "SAT-based decision tree learning for large data sets." In Proceedings of AAAI, 21.

[19] T. Biyanto, H. Fibrianto, E. Listijorini, and T. Budiati 2015. “Duelist algorithm: an algorithm inspired by how duelist improve their capabilities in a duel,” in Proceedings of the Seventh International Conference on Swarm Intelegence, Bali, Indonesia.