The Seybold Report
ISSN 1533-9211
DOI 10.5281/zenodo.6697697

WEIGHED QUANTUM PARTICLE SWARM OPTIMIZATION TECHNIQUE TO MEASURE THE STUDENT PERFORMANCE

THENMOZHI L

Assistant Professor, MGR College, Hosur, Tamilnadu, India. Email: thenmozhilakshmanan@gmail.com

Dr. N.CHANDRAKALA

HOD, SSM College of Arts and Science, salem,Tamilnadu,India. Email: nchandrakala15@gmail.com

Abstract

Forecasting each student's unique achievement could provide useful information about which students are most likely to fail or drop out, as well as which qualities hunt the student's educational career. Data mining (DM) gives the tools required to deal with this educational data in the hunt for knowledge and patterns. Inability to forecast students' marks, this research employs an educational dataset retrieved from the UCI Machine Learning (ML) repository connected to students' grades in India as well as offers methods based on absolute learning machine networks, ensemble learning, and particle swarm optimization. As a result, a storage facility strategy is required, in which we might efficiently organize information that will then be processed in such a manner that we'll have perfect knowledge and which various other operations like mining and artificial intelligence can be performed. Furthermore, two generated data sets were employed to check the reliability of the results produced using the suggested regression models. Hypothesis tests were used to verify the accuracy of the results after collecting the error value for each suggested method. The results show that the method that integrates ensemble techniques, Weighed Quantum Particle Swarm Optimization (WQPSO) performs better.

Keywords: Data mining; Machine Learning; PSO; Regression Models; Education field

1. Introduction

Several things could hunt a student's educational career and, as a result, his achievement [1]. The capacity to anticipate each student's performance independently could provide helpful information regarding which students are at risk of failing or dropping out. Forecasting student achievement at an institution could serve as a guide in this way, enabling visualization of elements affecting the teaching and learning process and how they affect student achievement [2]. As a result, it is easy to quickly evaluate and analyze which aspects must be tracked in favour of the performance rate to achieve the target level. It is projected that the volume of information in the globe doubles every 20 months, making the development of new ideas and methods suitable for contributing to the purpose of removing meaningful information from these enormous amounts of digital data increasingly essential [3-4]. DM is the study of extracting knowledge from huge datasets, and it employs knowledge and techniques from a variety of fields, including statistics, deep learning, pattern classification, as well as artificial intelligence.

Computer science, engineering, mathematics, physics, neuroscience, and cognitive science are just a few of the domains where DM has been used [5]. Educational data mining [6] has sparked a growing interest in employing data mining to study scientific concerns through

The Seybold REPORT
ISSN 1533-9211
DOI 10.5281/zenodo.6697697

educational research in recent years. Educational data mining (EDM) would be defined as a scientific field devoted to the development but also application of ML processes to particular types of data derived from educational settings in an attempt to face significant educational problems, as well as the use of these techniques to better understand the factors that influence the teaching process and student learning [7]. Neural Networks (NN) are one of the techniques employed in the DM stage because of their capacity to estimate complex operations as well as create appropriate for a wide variety of challenging natural and artificial events.

The lack of quicker training methods is a disadvantage of NNs [8]. Traditional learning methods are generally much slower than required and aim to modify the network's variables to attain the optimum modelling of the curve that best describes the problem in question. As an example, gradient descent approaches are mainly used for training neural networks, but they are slow and prone to converge to local minimums [9]. Proves, on the other hand, that a neural network with only one hidden layer may learn N different observations using practically any nonlinear activation function and only a small number of N neurons in the hidden layer [10]. The PSO functions by inverting outcome matrices from the hidden layer of the neural network and conducts training thousands of times quicker than conventional methods but also obtains higher generalization was used to build this type of network.

2. Related Works

Several issues arise when learning methods produce only one hypothesis, such as the possibility of getting caught in local minimums or the necessity to properly map the search space to achieve the optimum solution based on only a small collection of training [11]. There could be numerous solutions that fulfil the specified criteria in this fashion, but only one would be used as a return for the learning procedure. An ensemble method solves stability concerns and improves the outcome by learning numerous instances of the same fundamental learning algorithm, such as bootstrap aggregation [12], which comprises of training numerous instances of the same learning algorithm using multiple samples from the training dataset. Likewise, when working with PSO keep in mind the random factor induced by the creation of the hidden layer's weights as well as thresholds. One way to get around this negative point is to adjust the network parameters so these configurations could be defined in an optimized way, for example, the number of neurons in the hidden layer, thus limiting the effects of the arbitrary component current in the PSO network on the outcome [13-14]. As a result, an optimization problem could be modelled that tries to alter the variables of the PSO network according to a price measure, such as the Mean Absolute Error (MAE).

The goal of this study should be to integrate PSO NN with the ensemble learning method to produce regression models that advantage from the positive attributes of both methods, as well as to use the PSO to reduce the effects of the PSO technique randomness [15], define the number of estimators used in the ensemble efficiently, as well as modify the optimal combination of the PSOM network to achieve accurate prediction at the end. This model was tested on simulated datasets to ensure that the results were consistent, and then it was

ISSN 1533-9211

DOI 10.5281/zenodo.6697697

implemented to educational data from high school, particularly from the India subject, to estimate student achievement through the final grade, assisting in decision-making and the advancement of actions as part of the educational atmosphere's cooperation [16]. The results show that the algorithms based on ensemble employed in combination with the PSO are better in terms of forecast precision and reliability.

3. Materials and Methods

Ensemble learning and WQPSO was discussed to educational issues such as student achievement, school dropout, failure rate, and acceptance. Its goal is to see how accurate ensemble approaches are at predicting student success in a four-year undergraduate engineering course. Boosting, Bagging, and Random Forest was among the ensemble methods utilized, and they were contrasted to base techniques that did not use ensemble methods. Finally, the authors found that the ensemble approaches outperformed the base methods in terms of outcomes [17]. Develops and implements ensemble regression approaches to educational data to estimate the dropout rate. The results show that adopting ensemble methods improves the outcomes, which might assist educational system managers in making decisions.

Due to WQPSO's quick training, it aims to forecast dropout rates for online classes using a combination of PSOM Networks and decision trees. When contrasted to typical machine learning methods, the final results of the suggested method reported significant improvements. In the realm of EDM, swarming the particles is used to categorize problems on cognitive levels [18]. When the particle cluster classification method was contrasted to seven other machine learning techniques, advances in results could be shown, reaffirming the increase that could be produced through this type of approach. Proposes a classification approach that could be used to forecast student official outcomes using the particle swarm optimization algorithm predict on discrete spaces. The model's efficiency was significantly improved when compared to previous classification techniques.

The writer utilizes WQPSO to minimize the dimensionality of the data set before using classification algorithms to forecast the achievement of higher education students by forecasting their marks. Following the studies, it was discovered that the suggested framework could aid in the improvement of academic value and decision-making in the educational system. The difference between this work and the others described in this area is the unification of WQPSO [19]. Although PSO drawbacks, such as random weight production, could be mitigated by using ensemble techniques, which involve training various versions of the same method to decrease stability, WQPSO rapid training speed allows for more efficient ensemble training. Furthermore, using the WQPSO to specify the number of items in the WQPSOs hidden layer as well as the number of estimators in the ensemble aims to get an optimization approach and, as a result, an enhancement in the final result.

The ML technique that has been created to train hidden layer feed forward neural networks thousands of times quicker than classic approaches such as the back propagation algorithm while also improving generalization capabilities [20]. The approach was predicated on the

ISSN 1533-9211

DOI 10.5281/zenodo.6697697

concept that may estimate any function using random values for both the input layer's weights and each activation threshold in the network's hidden layer. WQPSO 's main feature has been that the hidden neuron layer does not need to be modified repeatedly, as well as the training error ||Kβ−y|| and the weight standard ||β|| are both minimized.

$$\sum_{x=1}^{m} \beta_x \left(w_x i_y + b_x\right), 1 \leq y \leq N \qquad (1)$$

The output weights are βx the activation function is f, the input weights are wx, and the activation threshold is bi. Provided that the model accurately captures the information, the connection could be stated as K = j, with

$$K = \begin{bmatrix} f(w_1.i_1+b_1)\ldots f(w_m.i_1+b_x) \\ f(w_1.i_n+b_1)\ldots f(w_m.i_n+b_x) \end{bmatrix} \qquad (2)$$

This set of methods integrates the results of preset machine learning techniques from a set of techniques to reduce forecast mistakes or reduce error rates. These ML algorithms utilized in groups were referred to as weak learners in this situation since the ensemble technique is used to enhance the systems' forecasts. Among the ensemble approaches, bootstrap gathering, boosting, and stacking come to mind. To achieve the final result, these ensemble approaches resemble the training dataset that is sent on to the machine learning techniques that are utilized as weak learners.

Eberhart and Kennedy invented the PSO method, which comprises a community of particles that are initialized in a particular search space and have their velocity data modified at each moment to their best values of p-best and g-best. A random factor should be used to weight the acceleration, with different random integers being created for acceleration towards p-best and g-best. Equations 1 and 2 are used to modify the particle's velocity and direction, respectively

$$U_{xd} = w * u_{xd} + c_1 * \text{rand} * (p_{xd} - i_{xd}) + c_2 * \text{rand} * \left(p_{gd} - i_{xd}\right) \qquad (3)$$

$$i_{xd} = i_{xd} + u_{xd} \qquad (4)$$

Where w represents the moment of inertia, c1 and c2 are two positive constants, and Rand () or Rand () are two random functions in the [0, 1] range.

4. Methodology

This section explains the process in depth. The approach is divided into four parts, as shown in Figure 1: dataset definition, modelling, calculation assessment, and analysis of the findings.

DOI 10.5281/zenodo.6697697

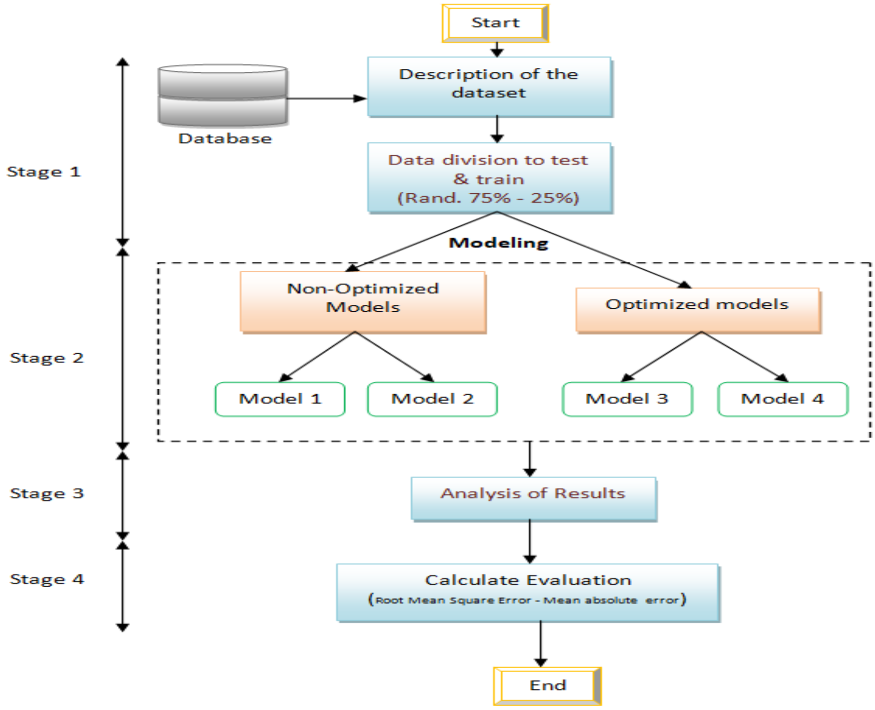

Figure 1: Flow of proposed diagram

4.1 Data sets

In this study, the models were first tested on two generated data sets to ensure that the findings acquired from each model were consistent. The following linear regression function was used to generate these data sets:

$$j = \beta_0 + \beta_1 i + e \qquad (5)$$

0 and 1 were produced using a random uniform dispersion with mean = 0 and standard deviation = 1, i was produced using a random uniform dispersion with mean = -1, and e was produced using a normal distribution with mean = 0 and standard deviation = 1. The parameter to be forecasted is y, and the input to the regression models is x, with the values to be approximated being 0 and 1. One of the data sets was supplied using the rule specified by Equation 5, and this data set was referred to as a data set without noise in this paper. In the other data set, the noise was also introduced to the y value to make forecasting more difficult.

The educational data set used in this study was taken from UCI of ML and the issue is the estimation of high school grades in India. Student marks, as well as demographic and school-related social variables, are included in the statistics. And it was gathered through school reports and questionnaires; in this effort, a dataset of student development in India was utilized. The base had 32 columns at first, but after the binary, category, and textual elements were removed, the base was reduced to 14 columns. Table 1 lists the columns that would be used in the final database, as well as their meanings.

The Seybold Report
ISSN 1533-9211
DOI 10.5281/zenodo.6697697

Table 1: Data set and its description

Column Name	Description
G3	Final grade
Absences	No. of school absences
Health	Current health status
Walc	Weekendalcohol consumption
Dalc	Workday alcohol consumption
Goout	Going out with friends
Freetime	Free time after school
Age	Student's age
Medu	Mother's education
Fedu	Father's education
Traveltime	Home to school travel time
Studytime	Weekly study time
Failures	No. of past class failures
Famrel	Quality of family relationships

In this case, the parameter to be forecasted was column G3, which contains information on the students' final grade, while the rest of the columns are used as input for the algorithm. First, the dataset was split into training and validation sections, after which it would be utilized as input for four different learning methods, which are detailed in the next chapter.

4.2 Model development

A WQPSO network was utilized in combination with the WQPSO in this design, optimizing parameter values in an attempt to optimize the final results. Each particle in this model connects a WQPSO network, which is then used to compute the absolute mean error to be minimized, which determines the particle population's direction and acceleration. Table 2 lists the variables that were selected to be changed in this model, as well as a short description of each one:

Table 2: Variables of Model

Parameter	Description
Alpha	Mixture coefficient for distances and scalar product input activations.
Number of hidden neurons	Number of units to generate the hidden layer.

This is the proposed method for integrating the ensemble with the WQPSO; it employs a Bagging ensemble with the WQPSO as the basis predictor. At each iteration of the algorithm, each WQPSO component was responsible for acquiring the absolute mean bagging error and attempting to reduce this quantity, which is received at the end of each WQPSO. Table 3 lists the variables that were changed in this model, as well as a short description of each:

The Seybold Report
ISSN 1533-9211
DOI 10.5281/zenodo.6697697

Table 3: Variables

Variable	Description
Alpha	Mixture coefficient for distances & scalar product input activations
No. of hidden neurons	No. of units to generate the hidden layer.
No. of estimators	No. of base estimators utilized to build the ensemble

5. Implementation

As a result, the WQPSO is first initialized with 30 particles, each of which has a position array containing values that represent the values of the parameters to be changed. For models 1 and 2, Table 4 illustrates the ranges selected for random initialization of each of the attribute values.

Table 4: Period initialization with weights

Variable	Initialization interval
Alpha	0.02 to 2
No. of hidden neurons	20 to 600
No. of estimators	25 to 300

The absolute mean error was the cost function to be optimized in both models. The following is the pseudo code that describes the WQPSO.

Allocation Pseudo code for Proposed System

The QoS-based allocation (QBA) method is comprised of the following steps:

Step 1: An incoming user task comes from any of the queues.

Step 2: If taskTi∈TQ1 (i.e., it is of high priority), then

Step3: Ti to VMG1 should be assigned (i.e. platinum class)

Step4: If TiTQ2 is true (i.e. medium priority)

Then

Step 5: Ti should be assigned to VMG2 (i.e. gold class)

Else

Step 6:VMG3 should be assigned task Ti (i.e. silver class)

The Seybold Report
ISSN 1533-9211
DOI 10.5281/zenodo.6697697

WQPSO WITH QUALITY OF SERVICE ALGORITHM

BEGIN TO CALCULATE A VM'S LOAD AND VOLUME.

For each machine, calculate p_best and g_best.

Do

Update the VM's load and volume.

Calculate the value of each machine's future resource needs. For each machine, update p_best.

For each machine, update g_best.

Move the job from the overloaded machine to the low-loaded machine.

Do

For each of the particle

Calculate the fitness real value of resource use, storage waste, and cost and efficiency responsiveness.

If the fitness function is better than the existing average fitness value (p_best), set the minimum values to that of the new p_best.

While the necessity for termination is not broken.

End

The requirement for termination has not been met.

End

Predict future VM resource requirements.

End

Each particle carries information about its position in the form of a vector made up of N values, the number of which is determined by the number of variables to be changed. Model 3 has a vector made up of two positions, whereas model 4 has a vector made up of three positions. Equation (6) was used to iteratively change this vector for each particle, with 30 components and 15 PSO iterations. These parameters are then utilized to train a WQPSO and a bagging type assembly using WQPSO as a basic technique. The absolute mean error is then calculated, which is then utilized as a cost function to determine the best overall price to direct component speed and position adjustments. The method, which begins with the division of the dataset into training and validation and ends with the extraction of each one's forecasts, is then repeated 300 times to acquire the average of each cycle's final mistake, ensuring robustness and confidence in the final results.
It's a metric that determines the difference between expected and actual values. The equation defines the mean absolute error.

ISSN 1533-9211
DOI 10.5281/zenodo.6697697

$$\text{Mean Absolute Error} = \frac{1}{n}\sum_{x=1}^{n}[j_x - \hat{j}_x] \qquad (6)$$

The root of mean squared error (RMSE)
It is a metric created by taking the square root of the mean square difference between the observed and forecasted numbers. This metric could be created using the formula of mean square error.

$$\text{Root Mean Square Error} = \sqrt{\frac{1}{n}\sum_{x=1}^{n}[(j_x)^2 - (\hat{j}_x)^2]} \qquad (7)$$

5.1 Results and Discussions
Table 5 presents correspondingly the prediction errors achieved for each of the developed regression models, for the generated datasets without noise and with noise.

Table5: With and without noise

Model	MAE				RMSE			
	With Noise	Without Noise	Errors	P-Value	With Noise	Without Noise	Errors	P-Value
Model 1 – Basic PSOM with PSO	0.1887	0.079	2.245	8.5xE-10	0.3512	0.092	2.15	8.5xE-11
Model 2 – Bagging PSOM with PSO	0.1898	0.080	2.245	8.5xE-13	0.3612	0.093	2.19	8.5xE-17

It is clear that, for both sets of simulated data, the algorithms that apply WQPSO for parameter optimization outperformed the non-optimized methods, with fewer forecast errors in both situations. When employing the optimized methods, the result generated on the generated data sets provides a guide on the reliability of the mistakes. The algorithms, as previously indicated, were deployed to the educational database to predict the G3 variable, which relates to the students' final marks. Figure 2 depicts the regression curves with the calculated values on the y axis as well as the real numbers on the x-axis for easier visualization.

DOI 10.5281/zenodo.6697697

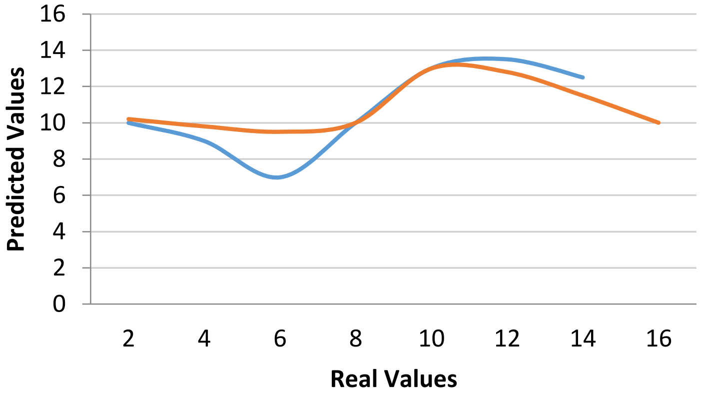

Figure 2: Regression curve

The curves depict the behaviour of each model prediction when compared to actual values. Although this perfect partner happens when a curve defines its trajectory by generating identical angles for both axes, it's easy to see how Model 2 maintains a more consistent trajectory with less abrupt curves, demonstrating better consistency than the other designs. When comparing the optimized methods, it's clear that the model that integrates the bagging ensemble and the WQPSO outperforms Model 1, which simply employs the PSO optimized with WQPSO. The decreased errors produced by Model 2 for both MAE and RMSE give additional assurance in the bagging ensemble's employment and reinforce its power in decreasing errors.

When comparing the outcomes of all models, it is clear that the proposed Model 4 is capable of achieving good outcomes than all of the others, since it has the lowest MAE and RMSE errors of all of the designs. As a result, it is essential to underline the efficiency of bagging and WQPSO in achieving superior educational performance from the moment they are utilized together. Statistical tests were done using the errors collected to gain a higher degree of dependability, and then hypothesis tests were constructed to support the findings collected. The normality of the data was first investigated using the Shapiro-Wilk test; however, it was discovered that the information did not follow a normal distribution, so the Wilcoxon hypothesis test was adopted with a significance level of 5%. Model 4's results were compared to those of all other models to ensure that they were accurate.

6. Conclusion

The use of EDM to reach data from educational contexts has been demonstrated to be an effective instrument, because it can model the issue to acquire data that could aid in decision making, as well as to identify possible points that require additional attention to mitigate an aim problem in question, using robust methods and refined methods. The issue in this work is to predict students' grades in India's high school discipline and thus to assess performance by focusing on assessment procedures to forecast student achievement and suggesting a model

The Seybold REPORT

ISSN 1533-9211

DOI 10.5281/zenodo.6697697

that combines WQPSO with NN, the ensemble bagging technique, and the WQPSO algorithm to improve the results acquired. Addressing the issue of student performance prediction has numerous advantages, as it allows for better decision-making and resource distribution in the educational environment in a secure, smart, as well as trustworthy manner. The models developed here would be applied to huge educational datasets in the future to analyze student achievement and aid in policy assessment and decision making. It is also intended to make model improvements, such as replacing the basic PSO with a more robust WQPSO model and evaluating the results with different ensemble methods, to get stronger stability and good outcomes.

References

[1] Miranda LJ. PySwarms: a research toolkit for Particle Swarm Optimization in Python. Journal of Open Source Software. 2018 Jan 11;3(21):433.

[2] Xue Y, Bai C, Qiu D, Kong F, Li Z. Predicting rockburst with database using particle swarm optimization and extreme learning machine. Tunneling and Underground Space Technology. 2020 Apr 1;98:103287.

[3] Latchoumi, T. P., Ezhilarasi, T. P., & Balamurugan, K. (2019). Bio-inspired weighed quantum particle swarm optimization and smooth support vector machine ensembles for identification of abnormalities in medical data. SN Applied Sciences, 1(10), 1-10.

[4] Wang B, Moayedi H, Nguyen H, Foong LK, Rashid AS. Feasibility of a novel predictive technique based on artificial neural network optimized with particle swarm optimization estimating pullout bearing capacity of helical piles. Engineering with Computers. 2020 Oct;36(4):1315-24.

[5] Sakri SB, Rashid NB, Zain ZM. Particle swarm optimization feature selection for breast cancer recurrence prediction. IEEE Access. 2018 Jun 4;6:29637-47.

[6] Yagiz S, Ghasemi E, Adoko AC. Prediction of rock brittleness using genetic algorithm and particle swarm optimization techniques. Geotechnical and Geological Engineering. 2018 Dec;36(6):3767-77.

[7] Nagra AA, Han F, Ling QH, Mehta S. An improved hybrid method combining gravitational search algorithm with dynamic multi swarm particle swarm optimization. IEEE Access. 2019 Mar 5;7:50388-99.

[8] Devaraj AF, Elhoseny M, Dhanasekaran S, Lydia EL, Shankar K. Hybridization of firefly and improved multi-objective particle swarm optimization algorithm for energy-efficient load balancing in cloud computing environments. Journal of Parallel and Distributed Computing. 2020 Aug 1;142:36-45.

[9] Hasheminejad SM, Sarvmili M. S3PSO: Students' performance prediction based on particle swarm optimization. Journal of AI and Data Mining. 2019 Jan 1;7(1):77-96.

[10] Latchoumi TP, Balamurugan K, Dinesh K, Ezhilarasi TP. Particle swarm optimization approach for waterjet cavitation peening. Measurement. 2019 Jul 1;141:184-9.

[11] Ranjeeth S, Latchoumi TP, Paul PV. A survey on predictive models of learning analytics. Procedia Computer Science. 2020 Jan 1;167:37-46.

[12] Latchoumi, T. P., Swathi, R., Vidyasri, P., & Balamurugan, K. (2022, March). Develop New Algorithm To Improve Safety On WMSN In Health Disease Monitoring. In 2022 International Mobile and Embedded Technology Conference (MECON) (pp. 357-362). IEEE.

[13] Latchoumi, T. P., Reddy, M. S., & Balamurugan, K. (2020). Applied machine learning predictive analytics to SQL injection attack detection and prevention. European Journal of Molecular & Clinical Medicine, 7(02), 2020.

ISSN 1533-9211
DOI 10.5281/zenodo.6697697

[14] Du SY, Liu ZG. Hybridizing Particle Swarm Optimization with JADE for continuous optimization. Multimedia Tools and Applications. 2020 Feb;79(7):4619-36.

[15] Latchoumi TP, Ezhilarasi TP, Balamurugan K. Bio-inspired weighed quantum particle swarm optimization and smooth support vector machine ensembles for identification of abnormalities in medical data. SN Applied Sciences. 2019 Oct;1(10):1-0.

[16] Karnan, B., Kuppusamy, A., Latchoumi, T. P., Banerjee, A., Sinha, A., Biswas, A., & Subramanian, A. K. (2022). Multi-response Optimization of Turning Parameters for Cryogenically Treated and Tempered WC–Co Inserts. Journal of The Institution of Engineers (India): Series D, 1-12.

[17] Banu, J. F., Muneeshwari, P., Raja, K., Suresh, S., Latchoumi, T. P., & Deepan, S. (2022, January). Ontology Based Image Retrieval by Utilizing Model Annotations and Content. In 2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 300-305). IEEE.

[18] Garikapati P, Balamurugan K, Latchoumi TP, Malkapuram R. A Cluster-Profile Comparative Study on Machining AlSi 7/63% of SiC Hybrid Composite Using Agglomerative Hierarchical Clustering and K-Means. Silicon. 2021 Apr;13:961-72.

[19] Agrawal AP, Kaur A. A comprehensive comparison of an ant colony and hybrid particle swarm optimization algorithms through test case selection. InData engineering and intelligent computing 2018 (pp. 397-405). Springer, Singapore.

[20] Shukla AK, Pippal SK, Chauhan SS. An empirical evaluation of teaching–learning-based optimization, genetic algorithm, and particle swarm optimization. International Journal of Computers and Applications. 2019 Nov 6:1-5.