International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Artificial Intelligence based Student Proctoring in Online Examination andGrade Prediction

Dr. Preethi D.[1], Dr. Reshma V. K.[2], Dr. Vigneash L.[3], Dr. P. Divya[4], Dr. Senthil Ganesh R.[5], Dr. Sivakumar S. A.[6]

Submitted: 23/09/2023 **Revised**: 25/10/2023 **Accepted**: 12/11/2023

Abstract: Numerous fields are utilizing and getting nurtured by the usage of Machine Learning (ML) algorithms owing to its simplicity in implementation and suitable accuracy. As the various online teaching and learning tools are booming post COVID - 19 to promote remote education, proctoring and assessing students' performance is a major challenge. This situation fits into data mining where students can be categorized based on their level of learning. Hence this promotes early attention for the slow learner students. This paper aims to provide solution to both proctoring and pre-assessing their grades through ML algorithms. The abnormal scores in final exams over internal assessments are compared and identified as outliers and it's been resolved using neural networks along with anomaly detection algorithms for proctoring. A comparison is made on various existing algorithms such as k-NN, Naive Bayes, SVM and Lion Optimization to explore the data in university records and exploit them for the purpose of judging their grades and ranks in advance. It is observed that the simulation results indicate that the Lion Optimization with anomaly detection performs better when compared with the other ML algorithms with an accuracy of 95.2%.

Keywords: *k-Nearest Neighbor, Lion Optimization, ML algorithms, Naïve Bayes, Proctoring and Pre- assessing, Support Vector Machine*

1. Introduction

According to the article published by UNICEF, around 825 million children across the world are lacking in the skillsets expected by the recruiting companies post COVID-19. Schools and Higher Education Institutes (HEIs) adopted many critical pathways and means to promote quality education through digital platforms. The online teaching and learning platforms adopted to reach education and keep the students equipped with skills remotely is a major transformation in the era. Also, it has paved way for numerous young children to continue their education online. As the mode of teaching has transformed from conventional to digital learning, innovative teaching as well as assessment techniques have been developed to help equip millions of young people with essential skills. Preserving the academic integrity of students is a major challenge in such online tool utilization.

There are four types of assessment techniques which is widely followed by all the schools and HEIs:

a. Diagnostic Assessment: Academicians will use this for analysing whether the student's' have the ability to cope up with the course plan and its prerequisites. It is informal and will not be considered for the final grade.

b. Formative Assessment: It is an integral part of assessment and used for obtaining feedback from the students' regarding their understandability of lessons. This assists the teachers to reform their mode of subject delivery. This also does not have any influence on the final grade.

c. Continuous and Summative Assessments: Continuous assessment is usually conducted periodically over the course and summative is conducted at the end of the course. These assessments are used for final grading.

d. Authentic Assessment: This assessment is used as a metric to decide how well a student can apply the acquired skills and knowledge outside the classroom.

All the above assessments will be deputed with supervisors to directly monitor the students in classrooms but during and after post COVID-19 numerous online courses have sprinted. The dearth of candid supervision by examiners during online examinations poses a significant disorderly

*[1]Associate Professor, Department of ECE, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai
prettz_d@yahoo.in[1]*
[2]Associate Professor, Computer Science and Engineering, Sri Krishna College of Engineering and Technology, Coimbatore
reshmavk@skcet.ac.in[2]
[3]Associate Professor, Electronics and Communication Engineering, Arjun College of Technology, Coimbatore.
lvigneash2504@gmail.com[3]
[4]Assistant Professor (SS), Dr NGP Institute of Technology, Coimbatore
divya@drngpit.ac.in[4]
[5]Associate Professor, Electronics and Communication Engineering, Sri Krishna College of Engineering and Technology, Coimbatore
drsenthilganesh@gmail.com[5]
[6]Associate Professor, Electronics and Communication Engineering, Dr NGP Institute of Technology, Coimbatore
drsasivakumar@gmail.com[6]
*Corresponding Author: prettz_d@yahoo.in[1]

conduct of students during such cases. Though there are various techniques to combat this mishap using Respondus, Web Camera's, Turnitin and SafeAssign, etc., still few students tend to use third party Ghostwriter such as WeChat according to the report released by University of New South Wales. One of the major objective(s) in this paper is to analyse their genuiness in exam once the summative assessment is over using ML algorithms based on the behaviour of their grades over the entire course.

Later, the student repository pertaining to their assessment records are collected from the Institution to predict the learner characteristics and patterns. This in turn helps the faculty as well as the student to improve their performance and score better. Hence, the second objective is to identify the various features pertaining to the ML algorithms such as k-NN, Naive Bayes, SVM and Random Forest to enhance their accuracy.

The paper organizations is as follows: Section 2 provides the insight to various literature work pertaining to the work and Section 3 demonstrates the proposed model followed by experimentation results in Section 4. The statistical experiment and its corresponding performance measures achieved using various ML algorithms are compared to conclude the better solution.

2. Literature Survey

2.1. Proctoring Students in Online Exams

[1] has experimented the online cheating of students in publications and also has categorized their levels. Carnevale has presented an article about the various ways and means of misconduct of students during exams. Martin et.al., has addressed the issues of educators due to lack of direct control on students. New digital monitoring and validation techniques to support the above-mentioned challenges are discussed. Watson et.al., has analysed the subjective measures rather than objective to avoid the student interaction during online exams. As misconduct in exams is more often in Online Assessment Platforms such as Massive Online Open Courses (MOOC) rather than traditional assessment, different proctoring methods such as keystroke recognition and Web camera are discussed by Xiong et.al. In addition, shuffling of questions and answers are also proposed in their work., has addressed the threat of conducting exams online dueto lack of proctoring tools and techniques.

[2] proposed model for authentication of log in user/student using knowledge-based technique with passwords. But the major drawback of this method being third party hackers who maliciously enter the host to boost the score of the student. The same scenario is also supported by [3]., as even the low secure passwords might lead to disruption of the system. Also presents biometric based authentication using keystroke with respect to fixed text using classification techniques. Rodrigues et.al., experimented with 800 samples using Markov models as classifier. [4]., has demonstrated using Neural Network (NN) to calculate the latencies to distinguish between valid and invalid users. [5]., has proposed hybrid algorithm combining Genetic Algorithm and SVM to select the features from enrolled users. [6]., included height, weight and gender as a feature to reduce False Rejection of valid users against the fake users.

2.2. Predicting Student Grade in Final Exams

[6] analysed the ML algorithms using a dataset of 117 students. With NB and ANN being implemented on the same data, ANN produced the best results. Video learning analysis was applied for dataset of 772 students by Raza et al., for early prediction of students' performance. When compared with the features for determining student performance of ANN with Random Forest (RF), RF is better than any other algorithm. An average accuracy of 83.48% was achieved by [7]., on carrying out experimentation with dataset having 1020 students. [11]., obtained 75.55% accuracy in predicting the final score of students in online exams using ML algorithms over a dataset having 460 students. Highest accuracy of 98.5% is achieved by [8]., using EMT to predict students' performance showed its superiority compared to the most commonly used ML algorithms for the samepurpose.

[9] demonstrated the application of NB classifier for determining the student dropout in schools or HEIs. This technique also showcases the root cause analyses in academic program with grading and GPA calculation of students. Exploration of performance on four different classifiers such as DT, SVM, ANN and DA with limelight to various performance metrics. They also have utilized EDM modules for data preparation and evaluation of grades. [8]., has proposed the online examination grading prediction using unsupervised learning technique such as clustering. Pradeep et.al., applied the samples in the database to pre determine the dropout of students. Classifier features such as attendance, marks, gender and physical traits were considered in their experimentation.

3. Ai Based Algorithms and their Mathematical Modelling

Based on either of two approaches supervised or unsupervised learning in system design for extracting grades from the students record, results in automatic detection process. In supervised case, the grades are labelled manually segregated from collected record and classified for every new entry recorded while unsupervised techniques are more flexible and do not require the reference grading label. Therefore, the unsupervised classification methods are used most frequently in grading applications.

Any classifier has two intermediate steps such as training phase and testing phase is carried out for most informative and discriminative features. Hence with the trained dataset, extraction and recognition of student performance can be characterized.

3.1 k-Nearest Neighbor

The labelling of test samples is allocated nearest in thenon-parametric machine learning process known as k- Nearest Neighbour (k-NN). The vector in equation (1) is denoted by Q if "S" is the sample in the training data with "q" attributes of independent variables. The representation is denoted as:

$$T = (q_1, q_2, .., q_n) \quad \ldots$$
(Eq.1)

A scalar function 'z' can be used to tie the variable 'p' to the attribute 'q', "p=z(q)". Assigning the set witha sample Q=t for which the relevant class must be identified, with the condition that 'z' is known. In contrast, the 'k' samples with q=t must be assigned to any class 'c' from the values of 'p' must be classed if 'z'is unknown.

Euclidian distance is used to calculate the minimum distance between the training data and the reference data set in order to assess the distance between the samples. In some cases, if any two data have the same distance, the nearest label will be selected at random. Equation (2) is used to determine the distance between two points 'l' and 'm', and the results are listed in ascending order.

$$\boldsymbol{d(t,p) = \sqrt{\sum_{i=1}^{n}(l_i - m_i)}^2} \quad \ldots$$
(Eq. 2)

Accuracy and reaction time can be compromised. Additionally, due to smoothing, the likelihood of over-fitting diminishes as k grows. When k = n, as in this and similar cases, a random decision is made.

Pseudocode 1 kNN algorithm

1: Import necessary modules from dataset
2: Loading data
3: Create feature and target arrays
4: Split the dataset for training and few as test dataset

```
M_train, M_test, N_train, N_test = train_test_split (M, N, test_size = 0.17, random_state=47)

knn = KNeighborsClassifier(n_neighbors=6)

knn.fit (M_train, N_train)
```

5: Calculate the accuracy of the model

```
print (knn.score (M_test, N_test))
```

3.2. Naive-Bayes Classifier

The Naive-Bayes Classifier (NB) has its roots in Bayesian theory of probability. Although Thomas Bayes developed the Bayes ratio formula during the course of the 18th century, Bayesian networks and classifiers didn't begin to effectively use it until the 1980s and 1990s. Since the deterministic relationship between each instance and its class is not taken into account by this classifier, it calculates conditional class probability, it is extensively utilized for a variety of applications thanks to its simplicity, learning speed, classification speed, and storage space. Because the input will not be encoded in the actual situation, conditional class probability improves the classifier's performance.

As soon as each class is determined for the provided features, the probabilistic classifier with the highest chance of correctly predicting the unknown class membership is assigned. From the training set, a linear decision boundary is encountered raising the likelihood within the class and instance area. If the specified class is left unclassified, the alternative classes will be assigned depending on the highest rank determined by the evaluation function.

The probability model of a classifier is a conditional model and is written as p(C|F_1,...,F_n) on a specific dependent class C, which has a small number of classes and is dependent on a variety of feature variables, including F_1,...,F_n. Despite the enormous number of features, this model on 'n' probability is realistically challenging. Consequently, this model is reorganized as

$$p\left(C|F_{-1},\ldots,F_{n}\right) = \frac{p(C)p(F_{-1},\ldots,F_{-n}|C)}{p(F_{-1},\ldots,F_{-n})} \quad \ldots$$
(Eq.3)

The above equation can be written as

$$posterior = \frac{priori \times likelihood}{evidence} \quad \ldots \text{ (Eq.4)}$$

Using the definition of conditional probability for repeated applications, it can be rewritten as

$$p(C, F_{-1}, \ldots, F_{-n}) = p(C)p(F_{-1}, \ldots, F_{-n}|C).. \quad \ldots$$
(Eq.5)

$$p(C, F_{-1},, F_{-n}) = p(C)p(F_{-1}|C)p(F_{-2},.F_{-n}|C, F_{-1}) \quad \ldots$$
(Eq.6)

$$p(C, F_{-1}, \ldots, F_{-n} = p(C)p(F_{-1}|C)p(F_{-2}|C, F_{-1})p(F_{-3},., F_{-n}|C, F_{-1}, F_{-2})$$
…(Eq.7)

$$p(C, F_{-1}, .., F_{-n}) = p(C)p(F_{-1}|C)p(F_{-2}|C, F_{-1}).. (F\_4., F_{-n}|C, F_{-1}, F_{-2}, F_{-3}) \quad \ldots$$
(Eq.8)

$$\boldsymbol{p(C, F_{-1}, \ldots\ldots, F_{-n}) = p(C)p(F_{-1}|C)p(F_{-2}|C, F_{-1})..p(F_{-n}|C, F_{-1}, F_{-2}, F_{-3}, .F_{n-1})}$$ …

(Eq.9) The "naive" conditional independence assumptions now enter the picture. It implies that by guaranteeing that each feature F__m is conditionally independent of each feature F__n, given m and n and not equal.

$$p(F_{-m}|C, F_{-n}) = p(F_{-m}|C), for\ \ m \neq n \qquad \text{… (Eq.10)}$$

Therefore, the joint model can be mathematically expressed as

$$\boldsymbol{p(C|F_{-1}, \ldots, F_{-n}) = \frac{1}{q} p(C) \prod_{i=1}^{n} p(F_{-i}|C)} \qquad \text{… (Eq.11)}$$

where '$\boldsymbol{q}$' is known as scaling factor which is dependent on $\boldsymbol{F_{-1}, \ldots, F_{-n}}$, (i.e.) a constant value, if the feature variable values are known. This model is paired with a decision rule for the NB classifier. The most likely hypothesis is chosen using the Maximum A Posteriori (MAP) decision rule. The function 'classify' serves as the appropriate classifier and is defined as follows:

$$\boldsymbol{classify(f_{-1}, .., f_{-n})} =$$

$$\underset{c}{argmax}\ \ p(C = c) \prod_{i=1}^{n} p(F_{-i} = f_{-i}|C = c)$$
...(Eq.12)

This classifier's shortcoming is its sensitivity lag, which occurs anytime characteristics remain unconnected inside a class.

Pseudocode 2 NB algorithm

1: Data preparation with set of features in the instance allotted with a class label

2: Determine prior probability of each class label training set upon dividing the frequency by number of instances.

3**:** Group the data rows under each class and find

the likelihood of the features.

4: Calculate posterior probability using bayes and

obtain new class instance with highest probability

#70% of data is training data and 30% is test data used for testing

5: prepare model and test model

predictions = getPredictions(info, test_data)

accuracy = accuracy (test_data, predictions)

print("Accuracy of the model")

3.3. Support Vector Machine Classifier

Support Vector Machines (SVM), a statistical supervised technique created by Boser and first shown in 1992, can be used for both classification and regression. This can be used for binary and multiclass segmentation, the two main types of segmentation, depending on the goal. Hyper plane decision boundaries can be used to separate classes based on the training dataset. Since the noise functions in the extracted biomedical data using the probability that are non-linear, logistic regression P(n=1/m) for n equal to 1 for a given input m of function h(m), which is based on anticipation and prediction of the exact class based on the threshold function: h(m) greater than or less than 0.5 as in equation (Eq. 13).

The assigned values are therefore either +1 or -1. In circumstances of perfect separability, the hyper plane decision boundary's distance is maximized to the nearest training data point; in all other cases, it is minimized to the nearest misclassified data point, but it is maximized to the nearest training data point that is classified. This technique can also be used to categorize dynamic noise sources. The SVM classifier is used to overcome the bottlenecks that were previously stated since it defines the margin as being closer to the decision surface for lower confidence and farther away for higher confidence[10].

$$\boldsymbol{P}(n = 1/m) = h(m) = \sigma(\beta^T m)$$ …
(Eq.13)

Let's say the training dataset ranges from '1 to L' of the type {m_i, n_i} for the values of 'i', the binary class of choice n_i falls into either of two groups '-1' or '+1', where 'L' represents the number of training points. The hyper plane, '$w = mx + b = 0$', shows the ideal distance between groups when the unit vector across 'b' is perpendicular to the origin and the value of 'w' is normal to the margin. The choice is made according to equation (Eq. 14) for the prediction of the noise source at any given moment in time.

$$m_i.w + b \geq +1 \quad for\ \ n_i = +1 \quad [HI]$$

$$m_i.w + b \leq -1 \quad for\ \ n_i = -1 \quad [H2]$$ …
(Eq.14)

Pseudocode 3 SVM algorithm

1: Plot the hyper plane $P_{-ij} = y_{-i}y_{-j}.x_{-i}x_{-j}$

2: Fitness function to be determined with 'i' ranges between 1 to L : $\sum_{i=1}^{L} \alpha_{-i} - \frac{1}{2}\alpha^T H_{-}\alpha$

3: Weight is determined for vector values having $\alpha_i > 0$: $\sum_{i=1}^{L} \alpha_i n_i m_i$

4: Taking average of all the collected support vectors to determine the orientation,

$$b = \frac{1}{T_s}\sum_{s\in S}(n_s - \sum_{m\in S}\alpha_{-m}n_{-m}m_{-m}.m_{-s})$$

where 'α'represents Lagrange multiplier and 'm'being the input.

5: During termination of each iteration, a new value of 'm' is obtained and once again a new decision of 'n' is formulated until constraint is reached.

Since the data's are more dynamic in nature, elaborate decision boundary is required rather than simple hyper plane.

3.4. Lion Optimization Algorithm

Nomads and groups make up the lion community (pride). Each member of the group is regarded as one of the numerous subsets of a set represented by [L_1,L_2,L_(3....) L_N], where 'N' is the total number of lions. The algorithm's pseudo code is presented in the following manner, step by step:

Step 1: Initialize the lions' positions at random in order to encircle the prey. This algorithm was inspired by the fact that the resistance of the prey to a lion group attack is represented by the notation L ((l_1),(l_2),(l_3),...(l_N)), where l_i[a_i,b_i] is an optimal solution across a space limit of 'N' for real co-ordinates 'a' and 'b', respectively.

Step 2: Assume a prey at random with an initial position of

$$PREY = \sum Hunters(x_{-1}, x_{-2}, \dots x_{-NVar}) / No.of\ Hunters$$
… (Eq.16)

For every new change in position of prey is given by,

$$PREY' = PREY + rand(0,1) * \%\ of\ Improve * (PREY - Hunter)$$
.. (Eq.17)

Step 3: The change in position or alignment of hunters for encircling the prey is,

$$Hunter' = \begin{cases} rand(Hunter, PREY), Hunter < PREY \\ rand(Hunter, PREY), Hunter > PREY \end{cases}$$
… (Eq.18)

where rand(a, b) represents the lower and upper bound. The success rate at which an optimum solution is attained is defined by,

$$S(i,t,p) = \begin{cases} 1, Best_{i,p}^{t}, p < Best_{i,p}^{t-1} \\ 0, Best_{i,p}^{t} = Best_{i,p}^{t-1} \end{cases}$$
…(Eq.19)

Pseudocode 4 LOA algorithm

1: Generate random samples of lion as initial population.

2: Initiate prides and nomad lions with %N and %S respectively.

3: In each pride some percentage of female lion go for hunting and the rest will mate with resident lion for strong cubs. If male lion is weak, it will send out of pride as nomad

4: For nomads both male and female lion will be roaming in search space. If the nomads' mate strong cubs will come or else prides might be attacked by nomads

5: some female with immigration rate would roam as nomads.

6: if no termination condition achieved, go to step 3.

4. Experimentation And Results

The simulation results are carried out using 630 samples taken from an UCI repository having atleast 33 feature attributes including age, gender, pass percentage, fail percentage and attendance, etc., In addition it is observed that the histogram is evenly distributed as the data is non-linear in nature. The confusion matrix is generated for the proposed and the compared algorithms. A detailed comparison is made on the performance metrics such as accuracy, precision, recall factor, F-measure and Kappa.

In addition, the student questionnaire as shown in Table 1, is circulated to collect responses and based on the answers a real time dataset is prepared which is then later used for classification and prediction of the student performance before examination. The questionnaire is circulated amongst all students for data population.

Table 1 Real time dataset collection using Student Questionnaire.

Factors	Description	Domain Values
1. Learning Techniques		

Visual	visual patterns are used for cognition	Value varies between 1-10
Physical	Activity based learning for students	Value varies between 1-10
Aural	Includes both the techniques Listening and studying	Value varies between 1-10
2. Personal Information		
Gender	Student's sex	{M,F}
Age	Below 18, 19-20, 21-22, 23 and Above	{I, II, III, IV}
Degree Type	Engineering/ Arts	{E, A}
Annual Income	Financial Status of student's family	{<20k, (20k-50k),(50k-80k), >80k}
3. Student Behavior		
Dependability Mode	Internet/faculty	Value varies between 1-10
Veracity	Attitude/Reflex	Value varies between 1-10
4. Intellectual Factors		
Inter-personal	Interaction with others	Value varies between 1-10
Intra-Personal	Enthusiasm	Value varies between 1-10
Logical	Problem Solving Nature	Value varies between 1-10
5. Comprehensive Ability		
Thinking ability	Innovative thinking	Value varies between 1-10
Attention	Mental Ability to focus	Value varies between 1-10
Memory	Protocols **used**	Value varies between 1-10

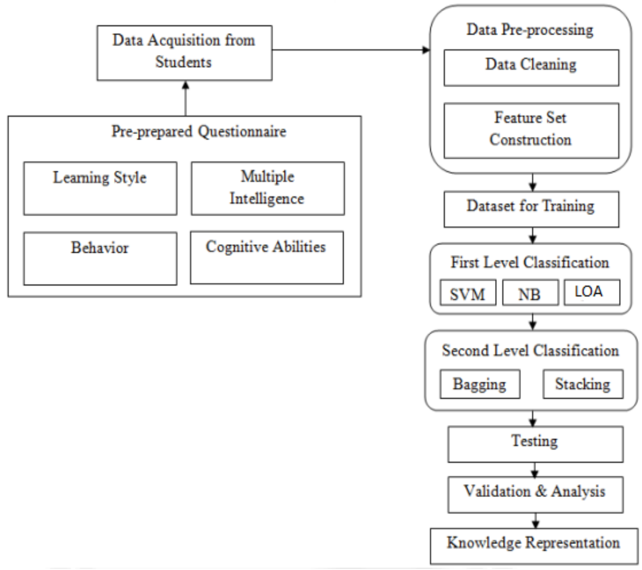

Fig.1 Flowchart depicting overall process of the online grade prediction

The features are given to the classifier for training in order to get the right outcomes. With the help of a trained network, student samples are categorized into classes like P, Q, R, and S, where P represents samples of students with good behavior and academic performance, Q represents samples of students with average academic performance and good behavior, R represents samples of students with average results and average behavior, and S represents samples of students with poor academic performance. Further student development initiatives are implemented in order to improve outcomes based on the classification results.

The bootstrap total has been used to calculate packing capacity as bagging is chosen in second level of classification. The outfit model's base classifier model is taken into consideration and assigned equal loads to each person. Additionally, the final classification brings about packing model is selected using the voting notion, which is illustrated visually in Fig. 1.

As soon as there are 'N' learner models and 'M' attributes, classification starts. By categorizing samples and subsets, the attribute set is used to broaden the form. The exercise data is used to determine the attribute that produces the best divide. This occurs frequently in all procedures that are simultaneously trained. are provided based on the combination of predictions made by each classifier. The datasets have been preserved as well as converted from Comma Separated Value organization to ARFF in order to make them usable with the WEKA instrument. When the results from various boundaries were analyzed, it was found that when the model was refined and used in learning the board frameworks at regular intervals based on machine learning techniques, the results had a higher precision rate.

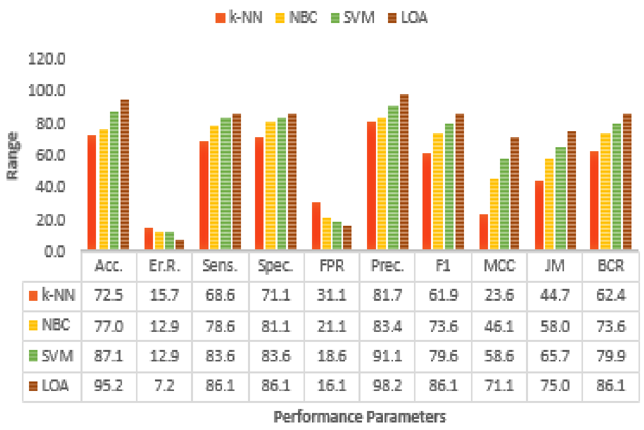

Fig.2 Performance Parameters of Classifier for sample size 150

The amount of participation in live meetings, the amount of participation in course archiving, and the amount of time spent on the material all contain more commitment than the other factors when the commitment of the input within the artificial neural system towards the output variables is examined. The evaluation parameters considered for simulation results are classification accuracy, precision rate, recall value, error rate and processing time. The LOA for sample size of 50 performs well with an accuracy of 93.4% and sensitivity of 88.5%, which are 30.0% and 32.5% higher than the basic k-NN classifier as shown in Fig. 2. This is based on analysis of accuracy and simulation time for the classifiers for classifying different student categories.

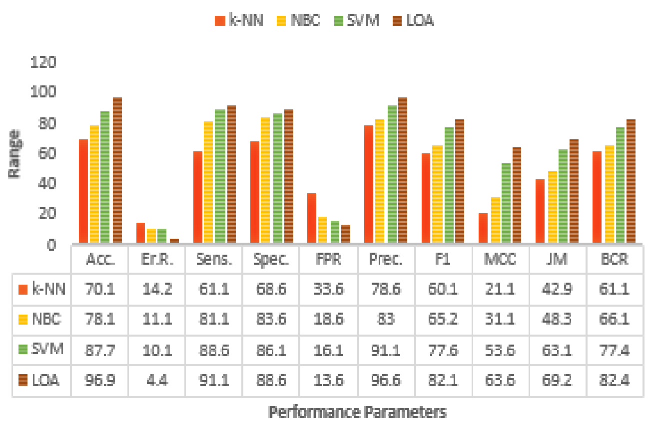

Fig.3 Performance Parameters of Classifier for sample size 350

It is observed that the LOA for sample size of 150 performs well with an accuracy of 95.2% and sensitivity of 93.5% which is 32.5% and 31.7% higher than the basic k-NN classifier as shown in Fig.3. It is observed that the LOA for sample size of 350 performs well with an accuracy of 96.9% and sensitivity of 91.1% which is 34.5% and 33.5% higher than the basic k-NN classifier as shown in Fig.3.

The use of LOA to successfully classify students into the appropriate categories within the allotted time period is seen in Table 2. Additionally, based on the average of the times for the four classifiers with the best performance, the range of time 0.1; between 0.1-1, 1-10, and 10-20 and the point in time during which the algorithm resumes are computed. Although all classifiers are able to identify student performance, each classifier has a different average completion time. Due to the capacity to work in parallel and handle multiple tasks, LOA execution time efficiency has significantly increased. The abortion times for k-NN and NB are seen to be 10.90 ms and 25.12 ms, respectively.

Table 2 Simulation Time of classifiers

Classifier	**Time for classification (ms)**				
	<0.1	**0.1-1**	**1-10**	**10-20**	**Abort**
k-NN	3.48	3.35	4.90	6.01	10.90
NB	2.56	2.07	3.99	3.87	25.12
SVM	1.89	1.56	2.78	2.66	26.67
LOA	1.09	1.20	2.01	2.45	36.98

5. Conclusions

This paper discusses the Accurate Student Classification Model for evaluating student academic achievement and proctoring, as well as numerous types of student behaviors and characteristics that influence the student's outcome. Several data analysis processes are performed with the goal of identifying hidden knowledge using predictive modelling and pattern recognition. In order to lower student failure rates and raise pass rates, many higher education institutions rely on the retention strategy. Therefore, the suggested effort entails creating a novel categorization model by successfully using machine learning techniques and data mining models in order to improve student exam performance and accelerate their rate of learning.

Additionally, the information is sent to the tutors so they can improve their performance effectively. The model's outcomes significantly outperform the other studies that were being analyzed. The evaluation's findings indicate that the proposed classification model, which employs many layers of classification, has a better accuracy rate than the other models that were also evaluated, averaging approximately 95.2%. The suggested model also achieves minimal computational and time complexity with minimal inaccuracy.

References

[1] Law, M.H., Figueiredo, M.A. and Jain, A.K. 'Simultaneous feature selection and clustering using mixture models', IEEE transactions on pattern analysis and machine intelligence., Vol.26(9), pp.1154-1166,2014.

[2] Lee, C.H., Lee, G.G. and Leu, Y,' Application of automatically constructed concept map of learning to conceptual diagnosis of e-learning', Expert Systems with applications., Vol.36(2),pp.1675-1684,2009.

[3] Li, Y., Gou, J. and Fan, Z.,'Educational data mining for students' performance based on fuzzy C-means clustering', The Journal of Engineering., Vol.11, pp.8245-8250,2019.

[4] Lima, M.N., Soares, W.L., Silva, I.R. and Fagundes, R.A.D.A. 'A Combined Model based on Clustering and Regression to Predicting School Dropout in Higher Education Institution' International Journal of Computer Applications., Vol.176, pp.1-8,2019.

[5] Ramos, J.L.C., e Silva, R.E.D., Silva, J.C.S., Rodrigues, R.L. and Gomes, A.S.(2016). A comparative study between clustering methods in educational data mining. IEEE Latin America Transactions., 14(8):3755-3761.

[6] Zaffar, M., Hashmani, M.A. and Savita, K.S. 'Performance analysis of feature selection algorithm for educational data mining', IEEE Conference on Big Data and Analytics., pp. 7-12,2017.

[7] Li, J., Zhao, J., Xue, G., 'Design of the index system of the college teachers' performance evaluation based on AHP approach', International Conference on Machine Learning and Cybernetics, IEEE, p. 995-1000,2018.

[8] Butz, BP, Duarte M & Miller SM 2, 'An Intelligent Tutoring System for Circuit Analysis', IEEE Transactions on Education, Vol. 49, No. 2, pp. 216-223,2006.

[9] De Moura FF, Franco LM, De Melo SL & Fernandes MA, 'Development of learning styles and multiple intelligences through particle swarm optimization', IEEE In Systems, Man, and Cybernetics (SMC), pp. 835-840,2006.

[10] Hua, Z., Xue-qing, L., Jie-cai, Z., & Jiang-man, X., 'Research and implementation of Course Teaching learning Process Management System', IEEE International Symposium on IT in Medicine & Education, Vol. 1, pp. 865 – 871,2009.

[11] Sharma, R., Dhabliya, D. A review of automatic irrigation system through IoT (2019) International Journal of Control and Automation, 12 (6 Special Issue), pp. 24-29.

[12] Timande, S., Dhabliya, D. Designing multi-cloud server for scalable and secure sharing over web (2019) International Journal of Psychosocial Rehabilitation, 23 (5), pp. 835-841.