Journal of Science and Technology
ISSN: 2456-5660 Volume 7, Issue 05 (JULY 2022)
www.jst.org.in *DOI:https://doi.org/10.46243/jst.2022.v7.i05.pp20- 31*

Prediction System for Student's Academic Performance to increase University Admission System and Cumulative Grade Point Average Credits

P.Mohana Priya[1]
[1](*Department of Information Technology, SASTRA Deemed to be University,Thanjavur,India)*
Corresponding Author: mohanapriyatce@gmail.com

To Cite this Article
P.Mohana Priya[1] **Prediction System for Student's Academic Performance to increase University Admission System and Cumulative Grade Point Average Credits "***", Journal of Science and Technology, Vol. 07, Issue 05,-July 2022, pp020-31***
Article Info
Received: 13-05-2022 *Revised: 2-06-2022* *Accepted: 1015-07-2022* *Published: 18-07-2022*

Abstract:

Education sector is a big boon for society and it is utmost important to strengthen the university admission system by constructing basic eligibility criteria in order to maintain consistent results and to analyze students' performance in the forthcoming semesters. This research work incorporates two prediction systems in which prediction system 1 consists of supervised machine learning classification algorithms such as Support Vector Machine, Random Forest, Naïve Bayes, Artificial Neural Networks Multi-Layer Perceptron and prediction system 2 is feeded with unsupervised clustering algorithms such as KNN, K-Means, DBSCAN and Agglomerative hierarchical clustering algorithms that have been trained with students' academic and personal details. It is found that 98% of detection accuracy is yielded as the result of supervised classification algorithms. Data is an important asset for every organization and hence this article is proposed to secure data from common breaches in software defined network. In this article, hybrid cipher model is proposed to safeguard the communication of data transmitted among the layers in software defined networks. The logic of hybrid cipher model is incorporated in software defined controller which encrypts open flow request and response messages. Software Defined Network is adapted for implementing hybrid cipher model as the network provides customizable platform and act as a unmanned security featured software controller. The proposed Hybrid Diagonal Transposition algorithm is incorporated with software defined wireless sensing node for encrypting user's data. Hence the unmanned security featured wireless sensing node is situation-aware, it detects malicious traffic flows and encrypts user's data. Hybrid Diagonal Transposition algorithm prevents data breaches in Software Defined Networks. Results are interpreted for various network and sensor metrics such as routing hops, participating node temperature, battery voltage, humidity, lights, received packets per node, number of network hops, power consumption, radio duty cycle, temperature of sensors, beacon interval, network hops, routing metric and the same work will be extended in future for comparative results.

Key Word*:Prediction System;University Admission System; Machine Learning;Students Academic Performance; Supervised and Unsupervised Learning Algorithms*

Journal of Science and Technology
ISSN: 2456-5660 Volume 7, Issue 05 (JULY 2022)
www.jst.org.in *DOI:https://doi.org/10.46243/jst.2022.v7.i05.pp20- 31*

I. Introduction

Education sector is a big boon for the society with respect to various metrics such as inculcating discipline, knowledge, technological advancements, life-skills and about handling situations among various students' community. It is utmost important to strengthen the university admission system[1] before pre-admission of a student by considering the basic eligibility criteria's such as high-school board marks, engineering entrance examination scores that is related with All India Engineering Entrance Examination (AIEEE), Joint Entrance Examination (JEE) and also with students schooling demography information. The above said are the basic criteria's that needs to be assured before admitting a student in order to analyze his/her performance and to consistently maintain good Grade Point Average (GPA) for all students until degree completion. Most of the higher education institutions are facing challenges while the huge education dataset been analyzed to predict the student performance for the forthcoming semesters.

In this dynamic real-life scenario like trending COVID pandemic has entirely changed the routine of a human and his lifestyle while pursuing a career-based goals. The COVID pandemic situation has become quite a major challenge for all education institutions in various aspects like student admissions, conducting physical classes, lab experimentations and examination. These stated scenarios lead to blended learning environment[2] to have a keep-in-touch platform with all students regardless of year of joining in order to carry out all activities without any interruptions.

This research work focuses mainly on predicting student academic performance[3] to increase university admission rate[4] and Cumulative Grade Point Average (CGPA)[5] of all students in both normal and COVID pandemic situations using machine learning algorithms[6]. Prediction begins with training the system to self-learn and to analyze the forth-feeding inputs for evaluating students' performance.

Student performance dataset is collected from UCI Machine Learning[7] repository which consists of students personal and academic information. The student performance dataset is tested with various supervised machine learning classification algorithms such as Random Forest[8], Naïve Bayes[9], Artificial Neural Network[10] and Support Vector Machine[9] and the same is tested against unsupervised machine learning clustering algorithms[11] namely KNN, DBSCAN and Agglomerative (or) hierarchical clustering methods. Further, additional attributes are incorporated namely social network-usage[12] and extra-curricular activities. Predicting GPA of the student for the forthcoming semesters is based on the pre-requisite courses that has been studied by the student. Toppers of the department can also be identified based on the resources referred by the student, participating in various discussions and also with student attendance.

Most of the existing research work concentrates on prediction of student performance using student demography[13], social network interactions[14] through supervised machine learning algorithms. This research article compares the results with both supervised and unsupervised machine learning algorithms such as classification and clustering. Finally, it decides which category of learning algorithm better predicts the student's performance.

The objective is to train the university admission system with various student academic attributes and personal information. The research work addresses the feasibility of implementing various machine learning prediction algorithms either supervised (or) unsupervised and finally will compare the efficiency of those algorithms with respect to evaluation metrics such as accuracy, precision, recall and F1-measure.

The article is organized in such a way that chapter 2 discusses about related state of the art methods for predicting student academic performance, chapter 3 deals with the proposed prediction system, chapter 4 discusses about an environmental setup, chapter 5 discusses about results and its discussions and chapter 6 concludes with future research directions.

II. Literature Review

Pranav Dabhade et al[15] have proposed a prediction approach which incorporates various machine learning algorithms that uses educational data for self-learning about a student performance. Hence, prediction algorithms are feeded with a questionnaire-survey and chosen academic-institution. The educational data of students have been pre-processed before it is feeded to the learning algorithms and hence irrelevant data is removed whereby data dimensionality is reduced considered as the advantage of this research work. The proposed prediction approach fails to concentrate on ANN[16] and regression techniques whereby the target is classified only as labels.

Harikumar Pallathadka et al [17] have maken use of machine-learning based classification algorithms to predict students' and faculties performance. In this research work, student performance is evaluated based on the past academic reports, student talents and interests whereas faculties performance is evaluated based on the submitted student feedback about a course and its respective staff-handled. Machine learning algorithms such as Naïve Bayes[18], ID3[19], C4.5[17] and Support Vector Machine (SVM)[17] are used to classify students who belongs to toppers category, average-grade category and below-average category. Among the interpreted results, SVM classifier outperforms rest of the algorithms used and the drawback of this work is considered as the efficiency of the algorithm is evaluated only with classification accuracy and error rate.

Shaik et al[20] have maken use of logistic regression and neural network algorithms to predict student's performance. This research work concentrates on students' co-curricular activities and its impact on academic scores. It is found from the findings that students involved in various co-curricular have utmost skills and can perform better in academics. Co-curricular, extra-curricular activities make the student to get involved in various activities such as group discussions, meetings and team work. It is found that the student success is directly linked to both the curricular and extra-curricular activities.

Vasiliki Matzavela et al[21] proposed decision-tree based learning and datamining for predicting student performance in a non-traditional environment. Most of the existing methods either concentrates on numerical data (or) categorical data while predicting student's performance. In this research work, each node of decision tree is weighted based on the knowledge of the student that belongs to various metrics. The proposed system uses both classification and prediction model using DT-Quest algorithm for enhancement of student academic performance.

Ahajjam Tarik et al[22] proposed an intelligent system which is based on Artificial Intelligence technique to predict student's performance during COVID-19. Various AI methods are used in which random-forest algorithm is found to be the best prediction algorithm which is based on the data values inferred from all three algorithms.

Xing Xu et al[23] has proposed a prediction model which consists of decision tree, neural network and support vector machine to predict academic performance of the student considering internet usage behaviors. Internet usage behavior is determined by various set of features such as online duration, Internet traffic and internet connection frequency. The prediction model consists of decision tree, neural network and support vector machine to analyze internet behavior features. Internet usage data is capable of differentiating and predicting students' performance.

Hassan Zein eddin et al[24]proposed AutoML approach to predict student performance. The AutoML approach concentrates on both behavioral data, academic data and also it includes statistical data to predict the grade of the student before pre-admission to the university. AutoML includes various ML algorithms namely ANN, K-Nearest Neighbor (KNN), K-Means, Naïve Bayes, SVM, Logistic regression, decision tree and its hyper-parameters are also tuned to provide best (or) accurate classification results. These algorithms are used to analyze the psychographic variables to predict student career success.

Maheswari et al[25] has analyzed the student performance factors using KNN algorithm. Most of the existing research work concentrates on the academic attributes where as in this research work various student factors such as his/her own interest on subject, parent's financial status, parent's involvement and interest towards their wards education, teachers involvement and interest, support from friends, neighbors, lifestyle of the family, student awareness on technologies and other activities.

Hajra waheed et al[26] deploys deep learning algorithm on the e-learning environments to analyze and predict student performance. It has set of unique hand-crafted features those are extracted from the e-learning environments with click-stream data. The proposed method also determines the efficiency of the deep learning model in its early prediction in its decision-making process.

El-Sayed Atlam et al[27] have maken use of descriptive statistics and machine learning to find the impact of COVID-19 pandemic on educational systems. In this research work, questionnaire-based survey is collected in which the first part of survey deals with demographic and educational characteristics and second part of the survey deals with collecting primary information of the student.

III. Proposed Prediction System

The proposed prediction system is feeded with student personal and academic details which serveas an input to the prediction systems 1 and 2 as shown in Fig. 1 and Fig. 2.

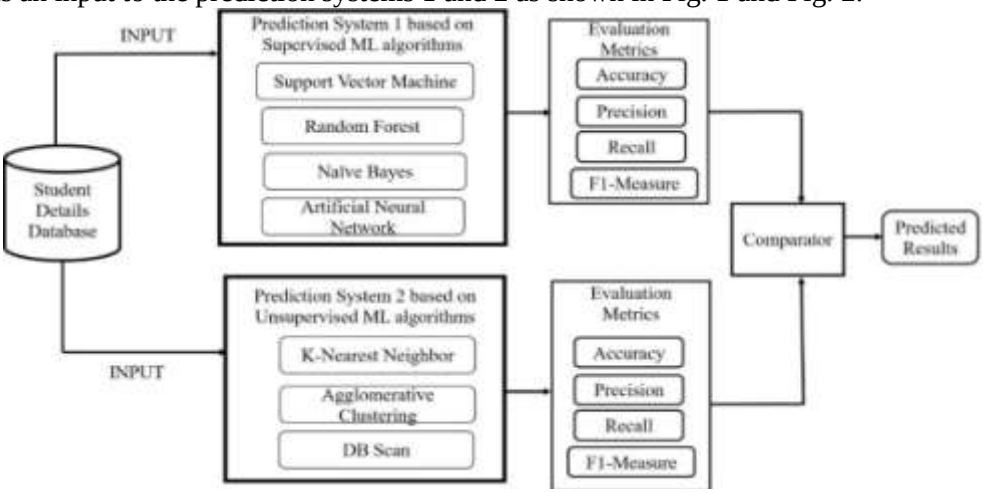

Fig. 1: Block Diagram of the Proposed Prediction System

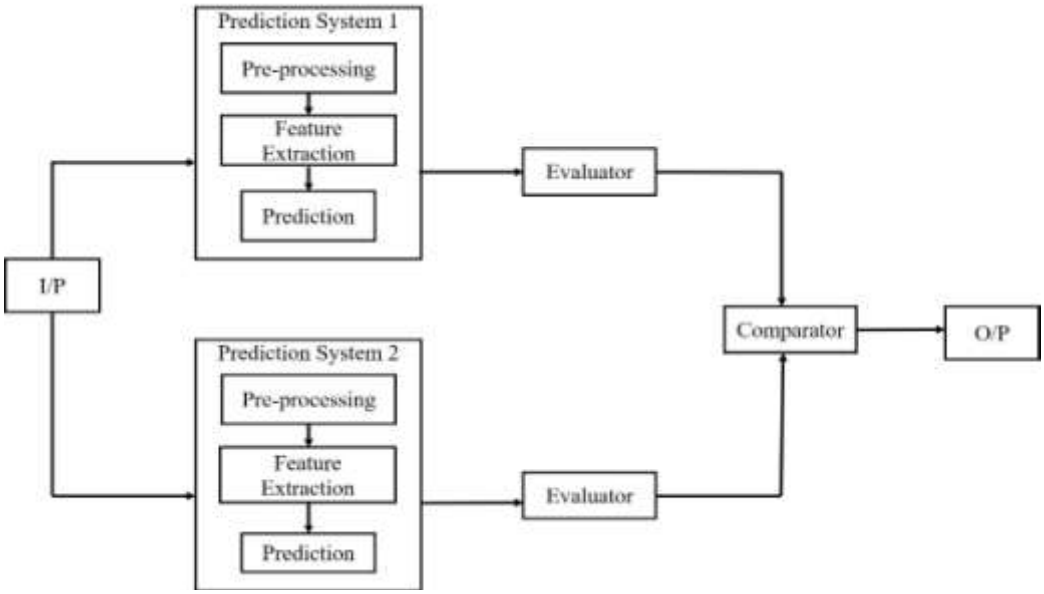

Fig. 2: Internal Working Module of the Proposed Block Diagram

Prediction system 1 is deployed with various supervised machine learning classification algorithms and prediction system 2 is deployed with various clustering algorithms. In both systems, input features are extracted followed by pre-processing of irrelevant and noisy data. Classifying the target is done with algorithms such as SVM, RF, NB and MLP of ANN. The feeded values are grouped together based on the clustering algorithms such as DBSCAN, agglomerative and KNN.

Classification Algorithms

A. SVM creates a hyper plane that separates outliers in which various set of students can be categorized based on the obtained total scores in high school. The hyperplane divides region of surface and the highest distance of various points from the plane reduces the error rate of the results. SVM is used in many of the real time applications and considered to be the reputed algorithm for the classification task. In this research work linear SVM is used for classifying the target label. Linear SVM hyperplanes predict performance of students using the below given formula.

$$\boldsymbol{y = ax + b} \tag{3.1}$$

where,

a = male top scorers

b = female top scorers

$$h(x_i) = \begin{cases} +1, & ax + b > 0 \\ -1, & ax + b < 0 \end{cases} \tag{3.2}$$

B. Random Forest is a machine learning technique which consists of multiple decision trees by considering a student ID as the root node. The Random Forest algorithm constructs multiple decision trees based on the chosen cross-validation, percentage splits, a confusion matrix is constructed for the target label and performance metrics such as True Positives (TP) and False Positives (FP) are calculated. Here, Random Forest Classifier is used to predict the student performance for forthcoming semesters by removing impurities and to gain ratio using Gini Index and entropy mathematical formulations.

$$Giniindex = 1 - \sum_{i=1}^{c} P_i^2 \quad (3.3)$$

$$Entropy = \sum_{i=1}^{c} P_i * log_2(P_i) \quad (3.4)$$

C. Naïve Bayes algorithm works based on the conditional probability distribution (or) probabilistic classifier which relies on posterior probability distribution. Posterior probability distribution is applied to the target labels such as total, male and female to segregate the information of how many male and female top scorers are there in the particular section. The posterior probability distribution for the given target labels can be calculated using the following equation.

$$P\left(\frac{c}{d}\right) = \frac{P\left(\frac{c}{d}\right)P(c)}{P(d)} \quad (3.5)$$

where,

$$P\left(\frac{c}{d}\right) = P\left(\frac{c_1}{d}\right) X\ P\ \left(\frac{d_2}{c}\right) X \ldots . X\ P\left(\frac{d_n}{c}\right) X\ P(c)$$

c = male top scorers

d = female top scorers

D. Multi-Layer Perceptron of ANN

Multi-Layer Perceptron (MLP) consists of input (visible layer), hidden layer and output layers. MLP classifier is activated with softmax function to classify the details of male and female top scorers based on students' academic and personal details. MLP performs binary classification using step function to categorize the provided values and it is shown as follows:

$$h^1 = step(Z^1) = step(w^1 x + b^1) \quad (3.6)$$

$$Y = step(Z^2) = step(w^2 h^1 + b^2) \quad (3.7)$$

where,

w = Weight function between neurons.

b = Bias Values.

Z = Step function.

x = Input Layer.

h = Hidden Layer.

E. Clustering Algorithms

Clustering algorithms group similar class of target labels which is closest to the centroid point "K". Euclidean distance between the centroid and groups are determined using the following formula. KNN clustering algorithm also construct clusters for the unknown data which is unsupervised. KNN construct clusters both for the feeded data and also for the new knowledge base of data. In this way, top scorers details are clustered together.

$$d(c,d) = \sqrt{(c_1 - d_1)^2 + (c_2 - d_2)^2} \quad (3.8)$$

if c = (c1,c2) and d = (d1,d2)

Assigning each point to the nearest cluster on KNN clustering algorithms is given as follows:

$$\arg min_{c_i \in C}\, dist(c_i, x)^2 \quad (3.9)$$

where,

ci = Clustered instances

C = Clusters and

x = Data points and

dist () is the euclidean distance.

Finding new centroid from the clustered group of points is determined using,

$$C_i = \frac{1}{|S_i|} \sum_{x_i \in s_i} x_i \quad (3.10)$$

The working principle of Density based Spatial Clustering of applications with Noise (DBSCAN) algorithm is based on both created clusters altogether with anomalous data found. This algorithm groups the similar groups of clustered values from the radius point. Minimum number of data values are clustered and DBSCAN categorizes the data points into three categories such as Core points, Border points and Outlier. Based on the terminologies provided above, clusters are created for male and female toppers data point values.

Agglomerative clustering algorithm clusters hierarchically using bottom-up approach and it works with its own cluster and concentrates on various minimum distance points to determine the difference between given data values.

IV. Experimental Setup

The proposed modules of machine learning classification and clustering algorithms are implemented using both python 3.10 version[28] and Waikato Environment for Knowledge Analysis (WEKA) tool[29]. The student-performance dataset is collected from UCI machine learning repository for the purpose of training and testing the classification algorithms. Machine learning algorithms are trained and tested with the ratio of 60:40. In python, Machine learning open-source libraries are collected from scikit-learn repository, from sklearn.svm, SVC is imported for implementing svm machine learning algorithms, from sklearn.ensemble, RandomForestClassifier is imported for implementing random forest classification algorithm, from sklearn.naive_bayes, Gaussian NB is imported for implementing naïve

bayes classification algorithm, from sklearn.neural_network.MLPRegressor, Multilayer perceptron regressor is imported for implementing regressor values. For clustering the student performance as male toppers and female toppers, from sklearn. neighbors, KNeighbors Regressor is imported for implementing K-Nearest Neighbor algorithm, from sklearn.cluster DBSCAN algorithm is imported and from sklearn.cluster agglomerative clustering is imported. 10 fold Cross-validation is implemented using train_test_split ratio. For plotting the graphical illustration of performance of the algorithms, from matplotlib.pyplot, plt is imported. All classification and clustering algorithms are implemented for 10 iterations in the testing phase, and those values are considered to be final for analyzing accuracy and correctness of algorithms.

V. Results and Discussions

This section discusses about the obtained results of all classification and clustering algorithms. Table 5.1 is tabulated with computed True Positive (TP) and False Positive (FP) values. Table 5.2 and Fig.5.1 discusses about the performance metrics such as precision, recall and F1-measure for SVM, RF, NB, ANN (MLP). All performance metrics are calculated using below given formulae.

$$Precision = \frac{TP}{(TP+FP)} \quad (5.1)$$

$$Recall = \frac{TP}{TP+FN} \quad (5.2)$$

$$F1 - Measure = 2\ X\ Recall\ X \frac{Precision}{Recall\ +Precision} \quad (5.3)$$

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (5.4)$$

Table 5.1: True Positive Rates and False Positive Rates of Supervised Classification Algorithms

Sl.No	Classification Algorithms	TP	FP
1	SVM	0.590	0.296
2	RF	0.771	0.137
3	Naïve Bayes	0.713	0.148
4	Artificial Neural Network	0.977	0.012

Table 5.2: Performance of Prediction Techniques based on Classification

Classification Techniques	Precision	Recall	F1-Measure
SVM	0.652	0.590	0.565
RF	0.771	0.771	0.770
NB	0.715	0.713	0.706
ANN (MLP)	0.977	0.977	0.977

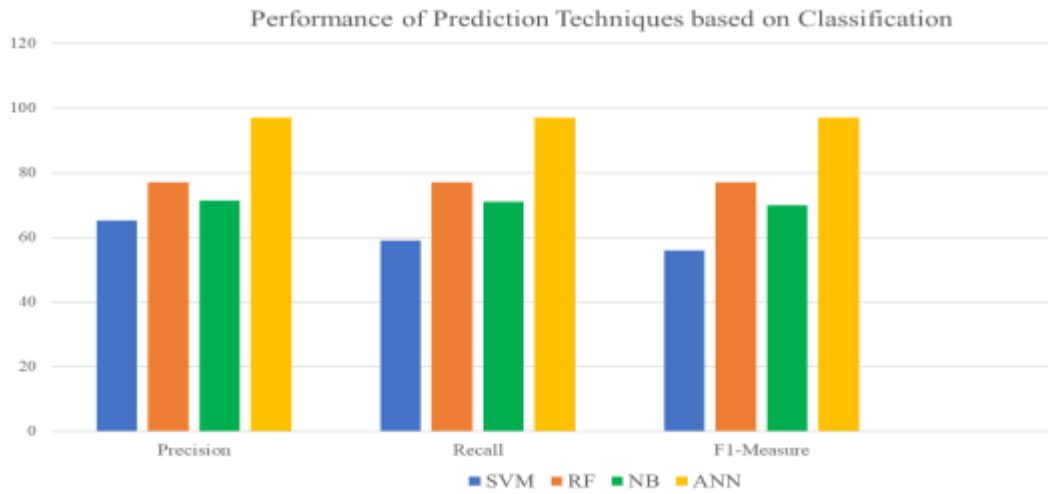

Fig.5.1. Performance of Prediction Techniques

Table 5.3 shows the confusion matrices of classification techniques. Table 5.3.1 is tabulated with the confusion matrix of SVM, Table 5.3.2 is tabulated with the confusion matrix of RF, Table 5.3.3 is tabulated with the confusion matrix of NB and Table 5.3.4 is tabulated with the confusion matrix of ANN (MLP), which consists of expected and predicted values of each category of student's performance data.

Table 5.3. Confusion Matrices of Classification Techniques

Table5.3.1 Confusion Matrix of SVM

A	B	C
57	68	2
12	183	16
2	97	43

Table5.3.2 Confusion Matrix of RF

A	B	C
110	15	2
23	161	27
0	43	99

Table 5.3.3 Confusion Matrix of NB

A	B	C
114	12	1
38	120	53
3	31	108

Table 5.3.4 Confusion Matrix of ANN (MLP)

A	B	C
124	2	1
3	205	3
0	2	140

Table 5.4 is tabulated with results of K-Means clustering consisting of attributes, full data, cluster 0 and cluster 1 which is grouped with information that belongs to students' performance dataset.

Table 5.4: K-Means Clustering

Sl.No	Attribute	Full Data	Cluster 0	Cluster 1
1	Gender	M	F	M
2	Nationality	KW	Jordan	KW
3	Place of Date	Kuwait	Jordan	Kuwait
4	Stage ID	Middle School	Middle School	Middle School
5	Grade ID	G-02	G-02	G-02
6	Topics	IT	French	IT
7	Semester	F	S	F
8	Relation	Father	Mom	Father
9	Raised hands	46.775	55.8037	39.1992
10	Visited Resources	54.7979	67.5023	44.1379
11	Announcements View	37.9188	46.8037	30.4636
12	Discussion	43.2833	42.8402	43.6552
13	Parents Answering Survey	Yes	Yes	Yes
14	Parent School Satisfaction	Good	Good	Good
15	Number of Days Absent	< 7	< 7	< 7

The scattered data points of student performance information is plotted in figure 5.2. Fig. 5.2 shows the traces of low, medium and high score values.

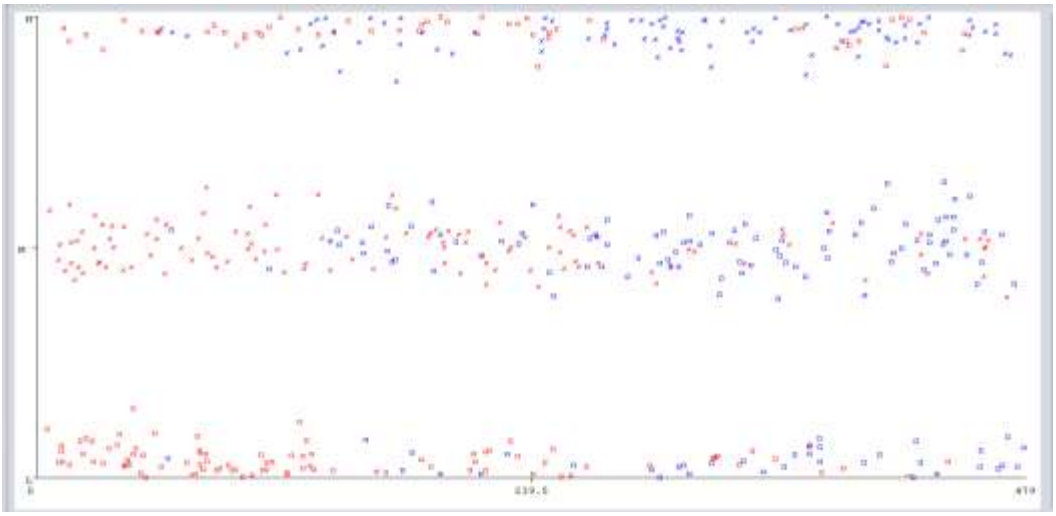

Fig. 5.2: KNN Cluster information of Target Label Score

Table 5.5 is tabulated with clustered instances such as cluster 0 with 219 data points and cluster 1 with 261 data points. Table 5.6 shows the clustered class values.

Table 5.5: Clustered Instances

Cluster 0	219
Cluster 1	261

Table 5.6: Classes to Cluster

Cluster 0 (High)	**Cluster 1 (Middle)**
36	91
95	116
88	54

Table 5.7 is tabulated with DB-SCAN clustering algorithm where value of ε is 0.9. Identified minimum points (min points) as 6, number of clusters found to be 1 and algorithm time elapsed is found to be 0.6 and its corresponding illustration is shown in Fig. 3.

Table 5.7: DB-SCAN Clustering

Clustered Data Objects	480
Number of Attributes	15
Epsilon	0.9
Min Points	6
Number of generated clusters	1
Elapsed Time	0.6

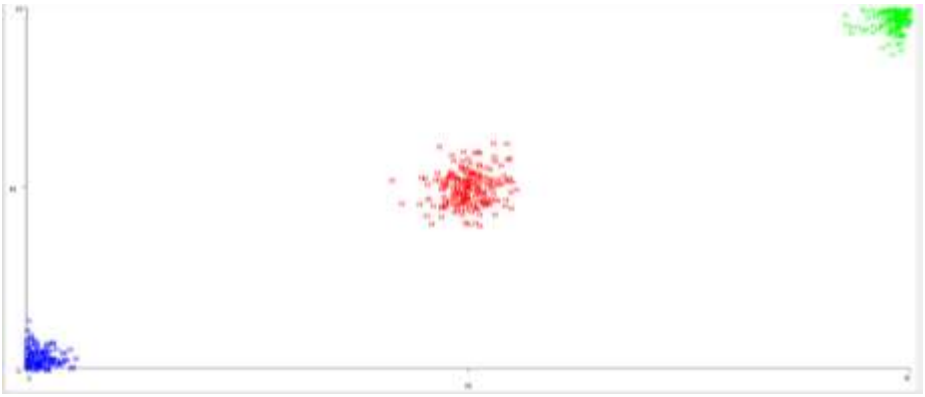

Fig. 5.3 : DBSCAN Cluster information of Target Label Score

Table 5.8, 5.9 and 5.10 are tabulated with clustered instances of agglomerative (or) hierarchical clustering algorithm and the same is illustrated as figure 5.4.

Table 5.8: Clustered Instances

0	7 (100%)

Table 5.9: Clustered Instances

0	479 (100%)
1	1 (0%)

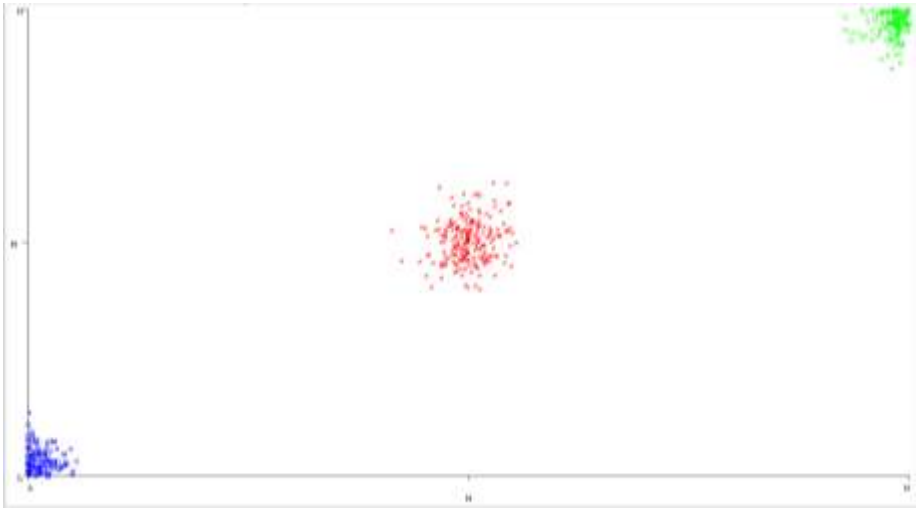

Fig. 5.4 : Hierarchical clustering information of Target Label Score

Table 10: Cluster Information

Cluster 0 (Middle)	Cluster 1 (High)
127	0
211	0
141	1

VI. Conclusion and Future Work

Educational sector is a big boon for the society and it is utmost important to strengthen the university admission system before pre-admission of a student by defining certain basic eligibility criteria. In this research work, prediction system is proposed in which student academic details and personal records are thoroughly studied and to predict forthcoming semester results in order to maintain consistent grades throughout the end semester. The proposed prediction system is deployed with supervised and unsupervised learning algorithms for predicting the performance of students in forthcoming semesters. It is determined that higher prediction accuracy as 98% of student performance is obtained in supervised machine learning classification algorithms rather than unsupervised clustering algorithms. Future research direction is to propose hybrid prediction system to analyze students' academic performance in both physical and blended learning environments.

References

[1]. Heinesen, E., "Admission to higher education programmes and student educational outcomes and earnings–Evidence from Denmark", Economics of Education Review, Vol. 63, pp. 1-19, 2018.
[2]. Chango, W., Cerezo, R., & Romero, C, "Multi-source and multimodal data fusion for predicting academic performance in blended learning university courses", Computers & Electrical Engineering, Vol. 89, pp. 106908, 2021.
[3]. Gonzalez-Nucamendi, A., Noguez, J., Neri, L., Robledo-Rella, V., García-Castelán, R. M. G., & Escobar-Castillejos, D, "The prediction of academic performance using engineering student's profiles", Computers & Electrical Engineering, Vol. 93, pp. 107288, 2021.
[4]. Helal, S., Li, J., Liu, L., Ebrahimie, E., Dawson, S., Murray, D. J., & Long, Q, "Predicting academic performance by considering student heterogeneity", Knowledge-Based Systems, vol. 161, pp. 134-146, 2018.
[5]. Adekitan, A. I., & Salau, O, "The impact of engineering students' performance in the first three years on their graduation result using educational data mining", Heliyon, Vol. 5, No. 2, 2019.
[6]. H. Kishan Das Menon, V. Janardhan, "Machine learning approaches in education", Materials Today Proceedings, Vol. 43, No. 6, pp. 3470 – 3480, 2021.
[7]. https://archive.ics.uci.edu/ml/index.php
[8]. Ahmed, N. S., & Sadiq, M. H, "Clarify of the random forest algorithm in an educational field", IEEE International conference on advanced science and engineering, pp. 179-184, 2018.
[9]. Pallathadka, H., Wenda, A., Ramirez-Asís, E., Asís-López, M., Flores-Albornoz, J., & Phasinam, K, "Classification and prediction of student performance data using various machine learning algorithms", Materials Today: Proceedings, 2021.
[10]. Rodríguez-Hernández, C. F., Musso, M., Kyndt, E., & Cascallar, E, "Artificial neural networks in academic performance prediction: Systematic implementation and predictor evaluation", Computers and Education: Artificial Intelligence, Vol. 2, pp. 100018, 2021.
[11]. Maheswari, K., Priya, A., Balamurugan, A., & Ramkumar, S, "Analyzing student performance factors using KNN algorithm", Materials Today: Proceedings, 2021.
[12]. Lau, W. W, "Effects of social media usage and social media multitasking on the academic performance of university students", Computers in human behavior, Vol. 68, pp. 286-291, 2017.
[13]. Dabhade, P., Agarwal, R., Alameen, K. P., Fathima, A. T., Sridharan, R., & Gopakumar, G, "Educational data mining for predicting students' academic performance using machine learning algorithms", Materials Today: Proceedings, Vol. 47, pp. 5260-5267,2021.
[14]. Wakefield, J., & Frawley, J. K, "How does students' general academic achievement moderate the implications of social networking on specific levels of learning performance?", Computers & Education, Vol. 144, pp. 103694, 2020.

[15]. Dabhade, P., Agarwal, R., Alameen, K. P., Fathima, A. T., Sridharan, R., & Gopakumar, G, "Educational data mining for predicting students' academic performance using machine learning algorithms", Materials Today: Proceedings, Vol. 47, pp. 5260-5267, 2021.
[16]. Rodríguez-Hernández, C. F., Musso, M., Kyndt, E., & Cascallar, E, "Artificial neural networks in academic performance prediction: Systematic implementation and predictor evaluation", Computers and Education: Artificial Intelligence, Vol. 2, pp. 100018, 2021.
[17]. Pallathadka, H., Wenda, A., Ramirez-Asís, E., Asís-López, M., Flores-Albornoz, J., & Phasinam, K, "Classification and prediction of student performance data using various machine learning algorithms", Materials Today: Proceedings, 2021.
[18]. Farhana, S, "Classification of Academic Performance for University Research Evaluation by Implementing Modified Naive Bayes Algorithm", Procedia Computer Science, Vol. 194, pp. 224-228, 2021.
[19]. Mak, B., & Munakata, T, "Rule extraction from expert heuristics: A comparative study of rough sets with neural networks and ID3", European journal of operational research, Vol. 136, No. 1, pp. 212-229, 2002.
[20]. Rahman, S. R., Islam, M. A., Akash, P. P., Parvin, M., Moon, N. N., & Nur, F. N, "Effects of co-curricular activities on student's academic performance by machine learning", Current Research in Behavioral Sciences, Vol. 2, pp. 100057, 2021
[21]. Matzavela, V., & Alepis, E, "Decision tree learning through a predictive model for student academic performance in intelligent m-learning environments", Computers and Education: Artificial Intelligence, Vol. 2, No. 100035, 2021
[22]. Tarik, A., Aissa, H., & Yousef, F, "Artificial intelligence and machine learning to predict student performance during the COVID-19", Procedia Computer Science, Vol.184, pp. 835-840, 2021.
[23]. Xu, X., Wang, J., Peng, H., & Wu, R, "Prediction of academic performance associated with internet usage behaviors using machine learning algorithms", Computers in Human Behavior, Vol. 98, pp. 166-173, 2019.
[24]. Zeineddine, Hassan, Udo Braendle, and Assaad Farah, "Enhancing prediction of student success: Automated machine learning approach", Computers & Electrical Engineering, Vol. 89, 106903, 2021.
[25]. Maheswari, K., Priya, A., Balamurugan, A., & Ramkumar, S, "Analyzing student performance factors using KNN algorithm", Materials Today: Proceedings, 2021.
[26]. Waheed, H., Hassan, S. U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R, "Predicting academic performance of students from VLE big data using deep learning models", Computers in Human behavior, Vol. 104, pp. 106189, 2020.
[27]. Atlam, E. S., Ewis, A., Abd El-Raouf, M. M., Ghoneim, O., & Gad, I, "A new approach in identifying the psychological impact of COVID-19 on university student's academic performance", Alexandria Engineering Journal, Vol. 61, No.7, pp. 5223-5233, 2022.
[28]. Van Rossum, "Python Programming language", USENIX annual technical conference, Vol. 41, No. 1, pp. 1-36, 2007.
[29]. Daw, S., & Basak, R. "Machine learning applications using Waikato environment for knowledge analysis", IEEE Fourth International Conference on Computing Methodologies and Communication, pp. 346-351, March 2020.