International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 25, No. 5, October, 2020, pp. 559-568
Journal homepage: http://iieta.org/journals/isi

An Effective Recommendation System to Forecast the Best Educational Program Using Machine Learning Classification Algorithms

Joy Dhar[1*], Asoke Kumar Jodder[2]

[1] Department of NSQF, Hatgobindapur M. C. High School, East Bardhaman 713407, West Bengal, India
[2] Hatgobindapur M. C. High School, East Bardhaman 713407, West Bengal, India

Corresponding Author Email: joy.dhar@hatgobindapurschool.co.in

https://doi.org/10.18280/isi.250502

Received: 10 July 2020
Accepted: 20 September 2020

Keywords:
automated recommendation system, students' academic performance, machine learning techniques, machine learning classification algorithms, educational data mining

ABSTRACT

After passing the 10th class, every student is eager to know which educational program will be the best for their higher education to match their career goal. Sometimes, they are very much confused to decide the best path for their higher education, and they need help to determine the best suitable academic program to develop their careers and achieve their goal. So, we introduce an effective recommendation system to forecast each student's best educational program for their career development. This proposed research is accomplished by utilizing machine learning (ML) approaches to forecast every student's best academic path based on their past academic performances and recommend them the best suitable academic program for their higher studies. Class 10th standard passing student data are supplied to this automated system, and a correlation-based feature selection approach is applied to extract the relevant features for each academic program. This study utilizes multiple ML algorithms to provide the best results and forecast each student's academic performance and select the best model based on their performance for each educational program. Hence, the best-selected model and related features are involved in the recommendation process to provide the best suitable academic path for achieving every student's career goals.

1. INTRODUCTION

Consider the fact that every student has a dream to be a successful person. However, some while they have faced some problems after completing their examination of class 10th standard. Therefore, they cannot understand which academic path is best for them to match their career goal. At this stage, they are following other people like their friends, relatives. Thus, there is a big chance for selecting an educational program through their parent choice or putting pressure on each student by society. Otherwise, they have a lack of knowledge to select a suitable program. Therefore, they even join any academic program to pursue higher education and develop their careers. At that time, they even cannot understand relating subjects of the selected program. Therefore, they may face an unfortunate result. Sometimes, some of them are called off their studies or may join another program or do small businesses or work in tiny offices. If they acquired help earlier, then there may be a huge probability for such students to become a successful person. To solve this problem and help such students, we proposed this research by which every student will get enormous help for their bright future. Effective recommendation methodology provides the best facility to every student by guiding them. Thus, they can select the best path for their higher education.

In India, after passing class 10th standard, there are different kinds of academic programs available for studying class 12th standard. Such programs are a science-based Intermediate program, a commerce-based intermediate program, an arts/humanities-based intermediate program, a polytechnic-based diploma engineering program, Industrial Training Institute (ITI) based program, a paramedical-based program like Diploma in Medical Laboratory Training (DMLT). Students are allowed to join any of these above-said programs after passing class 10th standard successfully.

Generally, students are selected for such programs based on their examination result of class 10th standard. However, in case of admission in polytechnic for the diploma engineering program, every student needs to pass the Joint Entrance Exam for Polytechnic (JEXPO). The problem of forecasting the individual student's academic performance based on their past educational records and providing the most significant features by recommending the best suitable educational program is expressed as a multiclass multilabel classification problem. Vast amounts of data are collected from the Hatgobindapur M. C. High School, the government-aided school situated in East Bardhhaman in West Bengal, India. The last nine-year student passing records of class 10th standard are collected from this school and supplied to our proposed system. This dataset is split into six sub-datasets, which belong to several above-said academic programs based on the class 12th standard. This study is applied to forecast the accurate result based on different ensemble-based machine learning (ML) classification models along-with the other ML classification models such as Random Forest (RF) [1], XGBoost (XGB), Gradient Boosting (GB), the Gaussian Naive Bayes (GNB) [1], Logistic Regression (LR) [1], CatBoost (CATB), Decision Tree (DT) [1], LightGBM (LGBM), and K-Nearest Neighbor (KNN) [1].

Generally, this paper addresses the following questions:

1. How does the proposed system show the performance of each model for the individual educational program?
2. Which classification model is best suited for forecasting the academic result of each student?
3. What are the relationships between different 10th class subjects and different educational programs at class 12th standard, and which subject straight affects a particular educational program?
4. Which feature or set of features are belonging to each particular program of class 12th standard?

This proposed system chooses the best model from the above-said models and automatically selects the set of features for forecasting each student's academic performance and thus provides a recommendation to each student by forecasting the best suitable academic program.

However, vast amounts of data are gathered from different educational organizations for developing related research work. Educational Data Mining (EDM) technique is most frequently used for the relevant research work. Different approaches have been taken regarding relevant educational research. Most of the earlier literature relies on predicting the students' academic performance. However, some researchers suggest different approaches to recommend the best course or select the best subjects for achieving each student's career goal.

However, past researchers utilized different machine algorithms: Quadratic Discriminant Analysis (QDA), Random Forest (RF), Gaussian Naive Bayes (GNB) Support Vector Machines (SVM), AdaBoost, K-Nearest Neighbors (KNN), to forecast and recommend the best course or select the best subjects for achieving each student's career goal [2, 3].

Regarding the prediction of selecting the best academic program based on the recommendation system is the best suitable approach for any educational organization. Every student needs such a recommendation system through which they will be able to choose the best suitable academic program based on their past academic records. Concerning educational purposes, this is the most challenging problem because it needs multiple methods to solve such problems, such as predicting the students' academic performance based upon their past educational records and then recommending them by providing the suitable academic program(s). Regarding the prediction of developing such a recommendation system, a branch of research is given. Ezz and Elshenawy develop an adaptive recommendation system for predicting student performance and recommend the best department based on their preparatory program [2]. Whereas, Rovira et al. [3] introduced the prediction for student dropout intension, forecasting student grades based on the relevant courses and personalized course recommendations.

Regarding the intelligent recommendation system, Goga et al. [4] predicted each student's first-year academic performances and then recommended the required actions for enhancement. Kurniadi et al. [5] introduce another literature based on an intelligent recommendation system for forecasting the student performance and career's interest and then predicting the recommendation for selecting subjects. Thai-Nghe et al. [6] use recommender system techniques to forecast the performance of the student.

After analyzing most of the previous literature, it revealed that most of the studies do not solve the multiclass multilabel classification problem for forecasting the individual student's academic performance and provides recommendation through predicting the best academic path for each student. Thus, this paper fills a research gap. However, literature [2] that only related to this proposed research. They developed a model for forecasting each student's educational performance based on preparatory courses' past performance [2]. By which such a system predicts suitable a department in the faculty of engineering for further studies. Thus, these findings have limitations and restrictions regarding their scope and hence significantly less usable. Perhaps our proposed research can be widely used as their scope is not limited and restricted in any manner. Apart from these, the prediction accuracy of such published literature is very much less as compared to our proposed research. Thus, these findings differ from the finding of our proposed research.

The remaining part of this study is composed below. Segment 2 illustrates the research methodology of this proposed system. Segment 3 exhibits the experiment and results through which a detailed analysis of data is performed for finding the solutions based on various classification algorithms with the help of different performance evaluation metrics. In the end, the conclusion is expressed in Segment 4.

2. RESEARCH METHODOLOGY

This segment describes an innovative methodology of our proposed system that is exhibited in Figure 1. We offer an innovative automated method through which individual students get information about the best educational program to meet their careers' goal. Such a system predicts the best suitable academic program for individual students among the six different programs of class 12th based on the students' educational performance of class 10th standard. This proposed system automatically splits the whole dataset into six sub-datasets for each educational program. Thus, such sub-datasets become binary classification based datasets for each educational program for the class 12th standard. Then further splits the datasets into training and test dataset for each program. After that, most related features are extracted from each dataset's original features for the individual program, and the resampling technique is applied to relinquish the imbalanced nature of the dataset. The different machine learning algorithms (i.e., models) are used for prediction and obtain the best result. In the end, collecting the classification result of each program, and then it is stored in the recommendation storage to provide the best educational program for each student.

2.1 Stage-1

There are two sections available for this proposed architecture in this stage, such as collecting data and data preprocessing.

2.1.1 Collecting data

This study collects the passing students' data of class 10th standard from Hatgobindapur M.C. High School. There are 2996 samples of students' data who were appearing for class 10th exam between 2010 to 2018 available for this proposed research. There are 592 invalid data available in this proposed research. Such invalid data are registered student(s) who did not appear in the examination, and some registered students after passing the 10th exam did not join in higher education. Removing such invalid data from the dataset and generating a new dataset. Such a new dataset comprises 2404 samples of valid student data. The dataset consists of those students' records who are passed out all the learning year beginning from class 10th to the class 12th standard. The target group of students has studied class 12th standard for 2 to 3 years [2]. Here the duration of different science-based intermediate programs, Commerce-based intermediate programs, arts/humanities-based intermediate programs, ITI based programs, and various paramedical-based programs are up to 2 years, while polytechnic-based diploma engineering programs are up to 3 years.

The dataset comprises two central parts, which are specified in the following:

Class 10th standard relevant data. It comprises all actions such as final theory exam marks and practical marks for all related subjects (Mathematics, Bengali, Physical Science, Life science, English, History, Geography, and Informational Technology). It also contains the final academic status and obtained marks of passing class 10th standard for each student.

The final year's academic status and obtained marks. It comprises the students' final year academic status, obtained marks of all subjects for each student or obtained marks, and prior academic status for those students who are still studying the class 12th standard program.

2.1.2 Data preprocessing

In this data preprocessing stage, already the collected raw data is supplied into this stage. Data is preprocessed through different data mining techniques such as Data cleaning techniques and Class encoding with feature discretization and binarization technique.

Data cleaning technique. In this proposed system, there are different unrelated data available in the dataset. Such data are related to the students' personal information, family-based information, students' social and demographic-based information. Such data does not affect the proposed system. That is why dropping tools are applied to remove such unrelated data from the dataset. In this research, there are several numbers of noisy data and missing data available in the dataset [7, 8]. We have used several strategies to resolve the difficulty of missing data and noisy data from the dataset, such as fill the missing data using the fill method and using the binning technique to solve the problem of noisy data [7, 8]. Apart from the above-said issues, we have found another problem, namely outliers. To remove the outliers from the dataset, we use the outlier detection technique using standard z-score cut-off values.

Class encoding with feature discretization and binarization technique. The proposed system uses final year academic status and marks as an essential feature to represent student performance. Based on earlier-stated features, a minimum criterion to pass the relevant program is that every student should obtain a minimum of 30% marks from total marks after then such students will get final academic status based on the above-said obtained marks. Thus, final-year academic status is selected to demonstrate the students' educational performance, which will be 1 for those who have received at least 30% marks for passing the relevant academic program and 0 for those who do not pass such educational programs. A feature discretization technique is applied to the final academic status and transforms the numerical values into two categorical values:

(1) Pass (for those whose final academic status is equal to or greater than 30%)

(2) Fail (for those whose final academic status is less than 30%)

After performing the feature discretization technique, a final academic status feature that holds only categorical data is then transformed into binary data using the binarization technique. After performing the binarization technique on the feature (final academic status) then estimates the average/mean value of all features relating to obtaining marks for all subjects as below:

$$\text{Final Academic Status } = \begin{cases} 1, & avg \geq 30\ [Pass] \\ 0, & avg < 30\ [Fail] \end{cases}$$

where, avg = average/mean value of all features relating to obtaining marks for all subjects.

After completing the above-said operations on the raw data, a multi-program preprocessed dataset DS is generated.

2.2 Stage-2

Such dataset DS is then split into each educational program, $\mathrm{DS} = \{\mathrm{DS}_1, \mathrm{DS}_2, \mathrm{DS}_3, \cdots, \mathrm{DS}_\mathrm{n}\}$ where n is the number of educational programs (programs) of the actual dataset DS as shown in Table 1. In this stage, the different machine learning techniques are used, such as partitioning the dataset, Correlation-based Feature Selection (CFS) technique, resampling techniques using random oversampling, and model development evaluation metrics, and the best model selection using performance evaluation metrics. In this stage, for each educational program DS_i, the following steps are evaluated:

Table 1. Split the dataset and generate a new dataset for each educational program

The name of the educational program	Total no. of students registered
Science-based intermediate program	543
Commerce based intermediate program	421
Arts/Humanities based intermediate program	855
Polytechnic based diploma engineering program	302
ITI based program	211
Paramedical based program	72

2.2.1 Partitioning the dataset

Each educational program DS_i is split into a training dataset (TR_i) and test dataset (TST_i) for *i*th educational programs. Such training dataset TR_i comprises m features, $TR_{i,m}$={$TR_{i,1}$, $TR_{i,2}$, $TR_{i,3}$,…, $TR_{i,m}$} and *k* dependent classes are supplied as input dataset P_i for ith educational programs.

2.2.2 Correlation-based Feature Selection (CFS) technique

A correlation matrix $\mathrm{C}^{\mathrm{i}}_{\mathrm{m,k}}$ between the feature *m* of input dataset P_i for *i*th educational programs and the dependent classes *k* for *i*th educational programs is estimated using the "Kendall's Tau correlation coefficient."

A correlation matrix $\mathrm{C}^{\mathrm{i}}_{\mathrm{m,k}}$ for P_i and a correlation threshold value T_i is used to check all the highly correlated features $F_{i,\,m}$ with the correlational matrix $\mathrm{C}^{\mathrm{i}}_{\mathrm{m,k}}$. Attributes $A_{i,m}$ in the correlation matrix, $\mathrm{C}^{\mathrm{i}}_{\mathrm{m,k}}$ is used to remove highly correlated features $\mathrm{F}_{\mathrm{i,\,m}}$ from input dataset P_i for ith educational programs to generate a different dataset $\mathrm{P}^{*}_{\mathrm{i}}$ with m significant features.

Thus, for every m significant features of input dataset P_i for *i*th educational programs:

$$P_{i,m} = P_{i,m} - A_{i,m} \quad (1)$$

where, $C^{i}_{m,k} > T_i$ (for *i*th educational programs, m features of input dataset P_i is subtracted with the attributes of highly correlated features $A_{i,m,}$ and the result is stored in input dataset P_i with m feature set where the correlation matrix's value $C^{i}_{m,k}$ is higher than the correlation threshold value T_i. Thus, it removes highly correlated features $F_{i,m}$ from the input dataset P_i.). Hence,

$$P_i^{*} = \forall \ P_{i,m} \quad (2)$$

where, $C^{i}_{m,k} \leq T_i$ (for each dataset *i*, all m features of input dataset P_i of each educational program with a correlation matrix $C^{i}_{m,k}$ is less than or equal to the correlation threshold value T_i will be selected and thus various datasets P_i^* are generated with significant features).

2.2.3 Resampling technique

A newly generated dataset P_i^* for each educational program is supplied as input to this phase. In this phase, the newly generated dataset P_i^* for each educational program may be imbalanced. The proposed system needs a resampling technique to solve such a problem. Such a resampling technique helps rebalance the class distribution for the used imbalanced dataset P_i^* for each educational program. Thus, we get new datasets P_i^R for each program. In this phase, the random oversampling technique is used to rebalance such input datasets in this proposed architecture.

$$P_i^R \leftarrow P_i^* \quad (3)$$

where, P_i^R= new datasets after the resampling technique, and P_i^*= generated datasets after the CFS technique.

2.2.4 The model development with performance evaluation metrics

10-fold cross-validation mechanism is used for this proposed system [9, 10]. The cross-validation technique helps train our newly generated dataset P_i^R of each program and perform a cross-validation strategy with new independent data to represent how accurate they are [9, 10]. In this phase, various machine learning models are applied for training and validation purposes through a 10-fold cross-validation technique [11]. The dataset P_i^R is used for training and validation purposes through a 10-fold cross-validation technique [9, 10, 12].

This cross-validation technique is used to train and validate this proposed system through training dataset P_i^R for each program with the set of q models [9, 10, 12],

$$M = \{M_1^{LGBM}, M_2^{GB}, M_3^{XGB}, M_4^{CATB}, M_5^{RF}, M_6^{KNN}, M_7^{GNB}, M_8^{DT}, M_9^{LR}, \cdots, M_q\}$$

where, M_1^{LGBM}= LightGBM model, M_2^{GB}= Gradient Boosting model, M_3^{XGB}= XGBoost model, M_4^{CATB}= CatBoost model, M_5^{RF} = Random Forest model, M_6^{KNN}= KNN model, M_7^{GBN}= Gaussian Naive Bayes model, M_8^{DT} = Decision Tree model, M_9^{LR} = Logistic Regression classification model and M_q= set of q models, and the performance of the set of q models,

$$PE = \{PE_1^{LGBM}, PE_2^{GB}, PE_3^{XGB}, PE_4^{CATB}, PE_5^{RF}, PE_6^{KNN}, PE_7^{GNB}, PE_8^{DT}, PE_9^{LR}, \cdots, PE_q\}$$

where, PE_1^{LGBM}= LightGBM model, PE_2^{GB}= Gradient Boosting model, PE_3^{XGB}= XGBoost model, PE_4^{CATB}= CatBoost model, PE_5^{RF} = Random Forest model, PE_6^{KNN} = KNN model, PE_7^{GBN} = Gaussian Naive Bayes model, PE_8^{DT}= Decision Tree model, PE_9^{LR} = Logistic Regression classification model and PE_q = performance of the set of q models.

Regarding each model's performance, different performance evaluation metrics such as F-measure, Cohen's Kappa, ROC-AUC, and Log loss are applied to the set of q models.

2.2.5 The best model selection using performance evaluation metrics

In this phase, each model's performance PE_q is applied to choose the best suitable model M_B from the set of q models. The proposed architecture uses the formula, which is specified below, for choosing the best model M_B.

$$M_B = M_q \text{ where } PE_B = \arg\max \sum_{q=1}^{n} PE_q \quad (4)$$

where, M_B= the best model selected from the set of q models in terms of the best performance PE_B for each educational program, the best performance PE_B selected in terms of checking different performance evaluation metrics such as F-measure, Cohen's Kappa, ROC-AUC, Log loss, and PE_q= set of performance for the models M_q.

2.2.6 Performance verification using test datasets and performance evaluation metrics

The best adaptive model $M_{B,}$ and the selected most reliable feature set P_i^R for each educational program, is further used to classify testing dataset TST_i for each educational program using different performance evaluation metrics for each set of q models. After performing the above-said operations, each educational program's classification results are supplied to recommendation storage, which is used to recommend the best suitable educational program for class 10th standard passing students.

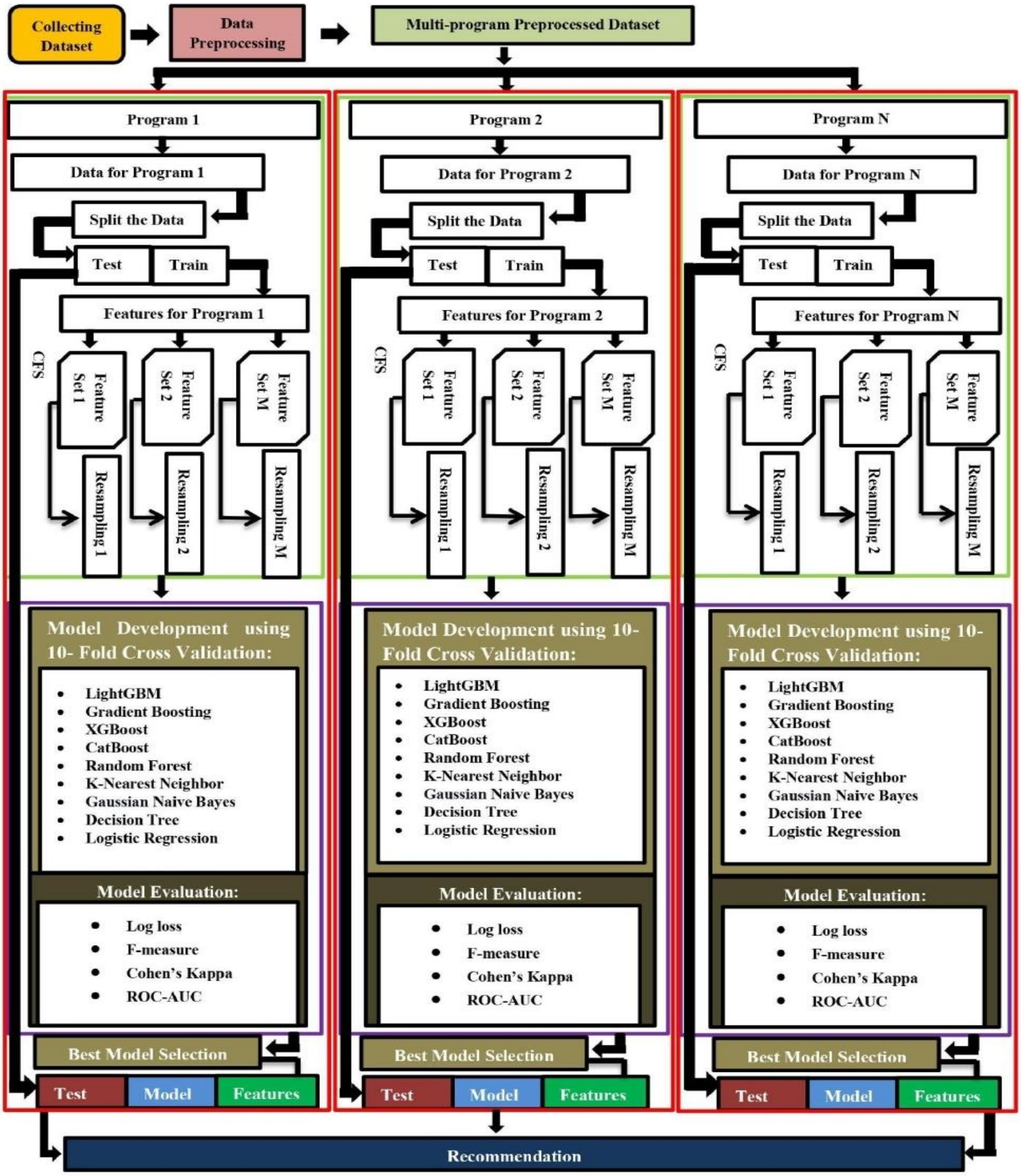


Figure 1. The architecture of an effective automated recommendation system

3. EXPERIMENT AND RESULTS

In this segment, the proposed research is applied to forecast the best educational program through an automated recommendation system. Class 10th standard passing student data is supplied to this proposed system. Through which such a system automatically forecasts student academic performance and recommending the best suitable academic path. Class 12th standard students' past performance, such as their class 10th results, are supplied as input to this proposed system. In this proposed system, the final year academic status data may be either pass or fail. After performing different operations on such data, as we explained earlier, the automated system's performance needs to measure. For measuring the performance of each model for each program, this study needs different performance evaluation metrics such as F-measure, Cohen's Kappa, and ROC-AUC score. Sometimes, the performance value of different models for each educational program may be equal. Therefore, this proposed system could not measure the best model according to their performances for each program. Thus, this proposed research needs another performance evaluation metric, namely, Log loss, to solve the above-said problem.

(1) F-measure: Precision and Recall metrics are utilized to evaluate the F-measure performance evaluation metric [13]. F-measure can be determined as:

$$\mathrm{F-measure\ (F)} = 2\times\frac{\mathrm{Precision\ (PR)}\times\mathrm{Recall\ (RE)}}{\mathrm{Precision\ (PR)}+\mathrm{Recall\ (RE)}}$$

$$\text{where Precision (PR)} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP) + False Positive (FP)}}$$

$$\text{and Recall (RE)} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP) + False Negative (FN)}}$$

(2) Cohen's Kappa: The expected agreement and the observed agreement are used to measure Cohen's Kappa performance evaluation metric [14]. Cohen's Kappa can be estimated as:

$$\mathrm{Cohen's\ Kappa\ (CK)} = \frac{\mathrm{P_{OA}} - \mathrm{P_{EA}}}{1-\mathrm{P_{EA}}} = 1 - \frac{1-\mathrm{P_{OA}}}{1-\mathrm{P_{EA}}}$$

where, $\mathrm{P_{OA}}$ = observed agreement and $\mathrm{P_{EA}}$ = expected agreement.

(3) ROC-AUC: The area under the Receiver Operating Characteristic (ROC) curve helps determine the ROC-AUC performance assessment metric [15].

(4) Log loss: Actual test data of y and the predicted value are applied to compute log loss. Log loss can be estimated as:

$$Log\ Loss\ (L) = -\frac{1}{N}\sum_{i=1}^{N}[A\times \mathrm{P} + (1-A)\times \mathrm{Q}]$$

where, A = y_i^{TST}, P = $\log(y_i^{PRED})$, Q = $\log(1-\ y_i^{PRED})$, $\mathrm{y^{TST}}$ = actual test data of y and $\mathrm{y^{PRED}}$ = predicted value.

The resulting performances of the individual models for each educational program are presented in Table 2, and the resulting performances of each educational program with the help of the best model and relevant features are displayed in Table 3. Those above-said tables also show the F-measure, Cohen's Kappa, ROC-AUC, and the Log loss value for each educational program's model and the resulting features related to the dependent classes. The acquired outcomes from Table 2 and Table 3 can determine the best explanation for the earlier-mentioned questions.

In respect of the first question, the answer exhibits the data presented in Table 2. Such a table represents the performance of each model based on the individual academic program.

In respect of the second question, the answer exhibits that the data represented in Table 3. Table 3, this table summarizes the outcome of the best model for each educational program. Table 3 reveals that no particular model is the best model that supplies the best suitable performance for all earlier stated educational programs. From Table 3, we can see that the best model for the science-based intermediate program goes to LightGBM with an F-measure of 100%, Cohen's Kappa of 100%, ROC-AUC of 100%, and Log loss of 0.02%. Whereas in respect of the commerce-based intermediate program, the best model is CatBoost with an F-measure of 98%, Cohen's Kappa of 96%, ROC-AUC of 98%, and Log loss of 6%. In the context of the arts/humanities-based intermediate program, LightGBM signifies as the best model with an F-measure of 100%, Cohen's Kappa of 100%, ROC-AUC of 100%, and Log loss of 0.3%. CatBoost is the best model for the polytechnic-based diploma engineering program with an F-measure of 98%, Cohen's Kappa of 96%, ROC-AUC of 98%, and Log loss of 8%. In the case of ITI based program, Random Forest (RF) signifies as the best model with an F-measure of 97%, Cohen's Kappa of 94%, ROC-AUC of 97%, and Log loss of 18%. Apart from the above-said programs, we have another educational program, namely a paramedical-based program in which Random Forest (RF) represents the best model with an F-measure of 90%, Cohen's Kappa of 80%, ROC-AUC of 90%, and log loss of 27%.

Concerning the third question, the response represented the data displayed in Table 3, through which the suggested system obtains the relevant features for each educational program of class 12th standard. In the science-based intermediate program, seven relevant features, such as Physical Science, Bengali, Life Science, English, Mathematics, Geography, and Information Technology, are contemplated. In other educational programs, this study obtains Bengali, English, Mathematics, and Geography as the four relevant features are contemplated for the commerce-based intermediate program. This paper obtains Bengali, English, History, Geography, and Life Science as the five relevant features are pondered for the arts/humanities-based intermediate program. While those intermediate programs are in huge demand, at this point, another academic program, namely a polytechnic based diploma engineering program, provides healthy competition to the above-said programs. This automated system obtains Mathematics, Physical science, English, Life Science, Geography, and Information Technology as six relevant features are contemplated for the polytechnic based diploma engineering program. In respect of ITI based program, the proposed system obtains three relevant features such as Mathematics, Physical Science, and Information Technology are pondered. There are only two relevant features concerning the paramedical-based program: Life Science and Physical Science, which are recognized by the proposed system.

From Figure 2 to Figure 5, different individual program models exhibit its performance through different performance evaluation metrics such as an F-measure, Cohen's Kappa, ROC-AUC, and Log loss. In this proposed research, forecast the students' educational performance with the ranges of F-measure and ROC-AUC from 90% for the paramedical-based program, 97% for the ITI based program, 98% for both of the commerce-based intermediate and polytechnic based diploma engineering program, and 100% for the science-based and arts/humanities-based intermediate program. The ranges of Cohen's Kappa from 80% for the paramedical-based program, 94% for the ITI based program, 96% for both commerce-based intermediate and polytechnic based program, and 100% for the science-based and arts/humanities-based intermediate programs which are used for forecasting students' academic performance. The proposed system is also used another performance evaluation metric to recognize the best model of the individual educational program. Such evaluation metric is Log loss, which is applied to forecast the best academic performance of each student with the ranges from 27% for the paramedical-based program, 18% for the ITI based program, 8% for the polytechnic based program, 6% for the commerce-based intermediate program, 0.3% for the arts/humanities-based intermediate program and 0.02% for the science-based intermediate program. Thus, this proposed research provides averages of different performance evaluation metrics: 97.16% for F-measure, 94.33% for Cohen's Kappa, and 97.16% for ROC-AUC along-with 9.88% for Log loss for all educational programs. By analyzing the above-said evaluation metrics from Figure 2 to Figure 5, the highest performance is achieved by the science-based and arts/humanities-based intermediate program with an F-measure of 100%, Cohen's Kappa of 100%, a ROC-AUC of 100%, and a Log loss of 0.02% and 0.3% for the science-based and arts/humanities-based program respectively. The paramedical-based program achieves the lowest performance with an F-measure of 90%, Cohen's Kappa of 80%, a ROC-AUC of 90%, and a Log loss of 27%.

Concerning the fourth question, all the above-said results are examined to recognize the correlation between the extracted relevant features and forecast classes (Fail/Pass) for each educational program. This proposed research needs to apply a visualization technique such as a radar chart to analyze the result. Analysis of results for each academic program can be exhibited using the radar chart. Figure 6 shows the best related and unrelated features for both science-based and arts/humanities-based intermediate program. Figure 6 shows the mean of individual features for both fail students and pass students. Figure 6A displays that Math, Physical Science, Life Science, and Information Technology are the healthiest subjects for success in the Science-based intermediate program of the class 12th standard, and Figure 6B exhibits that History, Geography, and Bengali are the most crucial subjects for success in the Arts/Humanities-based intermediate program of class 12th standard. Such results are matched up with produced best-related features from the proposed system.

Table 2. Performance of different models based on each educational program

	Model used	F-measure	Cohen's Kappa	ROC-AUC	Log loss
Science-Based Intermediate Program	**LGBM**	**1.0**	**1.0**	**1.0**	**0.0002**
	GB	1.0	1.0	1.0	0.0008
	XGBoost	1.0	1.0	1.0	0.009
	CatBoost	1.0	1.0	1.0	0.01
	KNN	0.98	0.96	0.98	0.07
	RF	1.0	1.0	1.0	0.03
	DT	0.97	0.94	0.97	0.26
	LR	0.99	0.98	0.99	0.07
	GNB	0.96	0.92	0.96	0.09
Commerce Based Intermediate Program	LGBM	0.97	0.94	0.97	0.07
	GB	0.97	0.94	0.97	0.13
	XGBoost	0.96	0.92	0.96	0.19
	CatBoost	**0.98**	**0.96**	**0.98**	**0.06**
	KNN	0.96	0.92	0.96	0.09
	RF	0.96	0.92	0.96	0.1
	DT	0.88	0.76	0.88	0.52
	LR	0.89	0.78	0.89	0.21
	GNB	0.89	0.78	0.89	0.23
Arts/Humanities Based Intermediate Program	**LGBM**	**1.0**	**1.0**	**1.0**	**0.003**
	GB	1.0	1.0	1.0	0.02
	XGBoost	1.0	1.0	1.0	0.008
	CatBoost	1.0	1.0	1.0	0.006
	KNN	0.96	0.92	0.96	0.06
	RF	1.0	1.0	1.0	0.01
	DT	0.99	0.98	0.99	0.34
	LR	0.99	0.98	0.99	0.05
	GNB	0.98	0.96	0.98	0.04
Polytechnic Based Diploma Engineering Program	LGBM	0.97	0.94	0.97	0.17
	GB	0.96	0.92	0.96	0.20
	XGBoost	0.97	0.94	0.97	0.12
	CatBoost	**0.98**	**0.96**	**0.98**	**0.08**
	KNN	0.97	0.94	0.97	0.39
	RF	0.97	0.94	0.97	0.11
	DT	0.95	0.90	0.95	0.46
	LR	0.94	0.88	0.94	0.16
	GNB	0.94	0.88	0.94	0.20
ITI Based Program	LGBM	0.97	0.94	0.97	0.38
	GB	0.97	0.94	0.97	0.31
	XGBoost	0.95	0.90	0.95	0.21
	CatBoost	0.97	0.94	0.97	0.19
	KNN	0.96	0.92	0.96	0.65
	RF	**0.97**	**0.94**	**0.97**	**0.18**
	DT	0.95	0.90	0.95	0.82
	LR	0.86	0.70	0.85	0.30
	GNB	0.84	0.66	0.83	0.35
Paramedical Based Program	LGBM	0.85	0.71	0.85	0.32
	GB	0.89	0.78	0.89	0.34
	XGBoost	0.88	0.76	0.88	0.37
	CatBoost	0.90	0.80	0.90	0.28
	KNN	0.89	0.78	0.89	0.69
	RF	**0.90**	**0.80**	**0.90**	**0.27**
	DT	0.77	0.53	0.76	0.47
	LR	0.75	0.50	0.75	0.51
	GNB	0.76	0.53	0.76	0.51

Table 3. Performance of each educational program with the best model and relevant features

Program	Best Model	F-measure	Cohen's Kappa	ROC-AUC	Log loss	Threshold value (T)	Number of Features	Related Features
Science-Based Intermediate Program	LGBM	1.0	1.0	1.0	0.0002	0.1	7	Physical Science, Bengali, Life Science, Mathematics, English, Geography and Information Technology
Commerce Based Intermediate Program	CatBoost	0.98	0.96	0.98	0.06	0.2	4	Bengali, English, Mathematics and Geography
Arts/Humanities Based Intermediate Program	LGBM	1.0	1.0	1.0	0.003	0.1	5	Bengali, English, History, Geography and Life Science
Polytechnic Based Diploma Engineering Program	CatBoost	0.98	0.96	0.98	0.08	0.15	6	Mathematics, Physical Science, English, Life Science, Geography and Information Technology
ITI Based Program	RF	0.97	0.94	0.97	0.18	0.25	3	Mathematics, Physical Science, and Information Technology
Paramedical Based Program	RF	0.90	0.80	0.90	0.27	0.3	2	Life Science and Physical Science

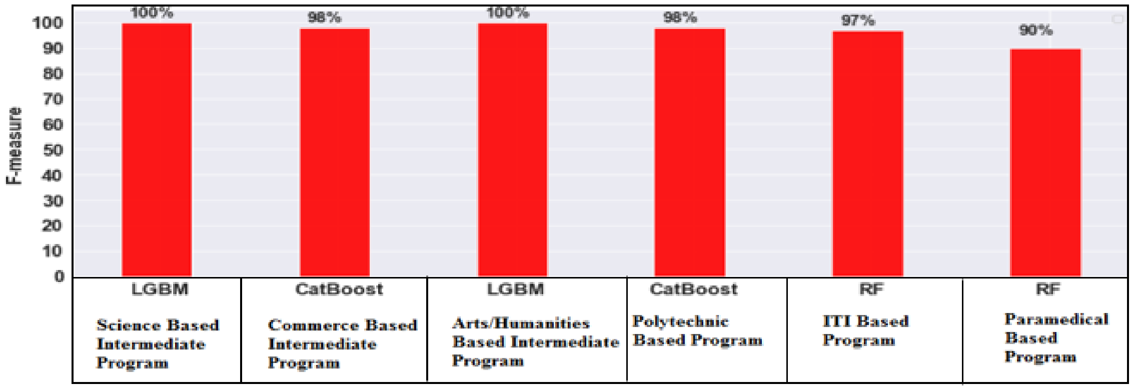


Figure 2. Values of F-measure for the proposed research

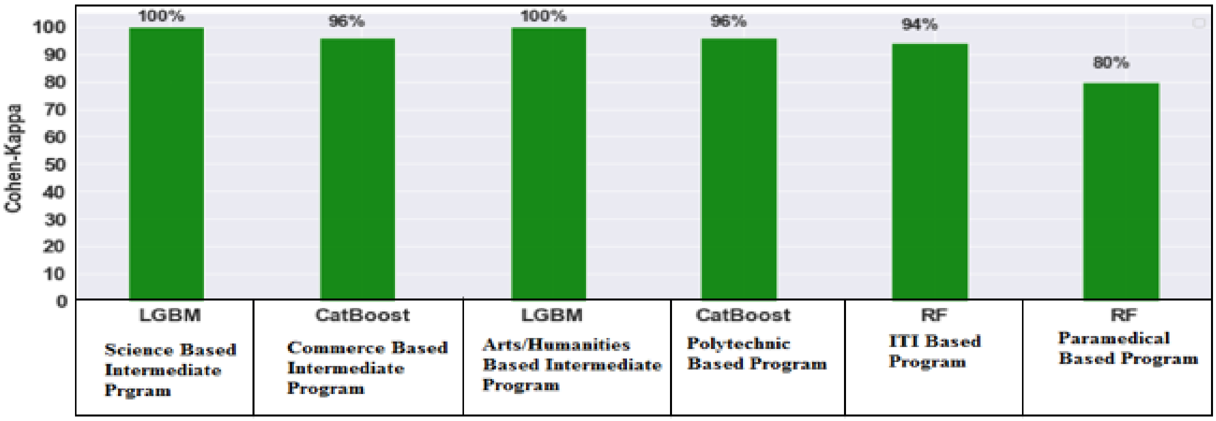


Figure 3. Values of Cohen's Kappa for the proposed research

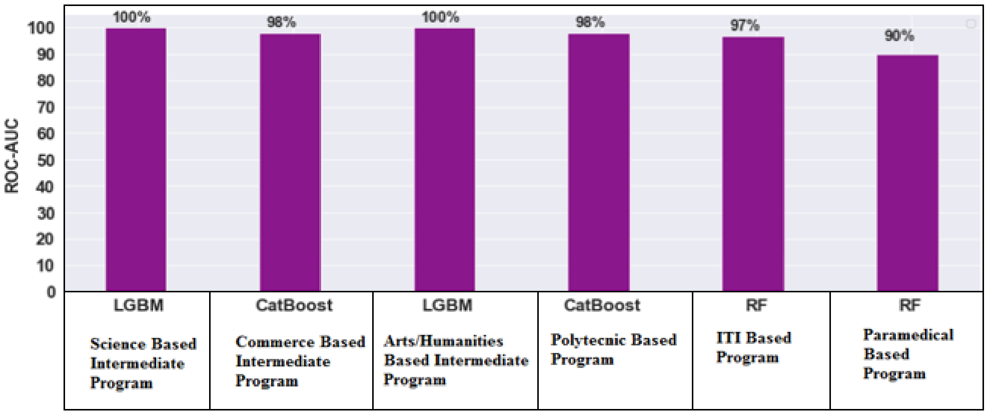


Figure 4. Values of ROC-AUC for the proposed research

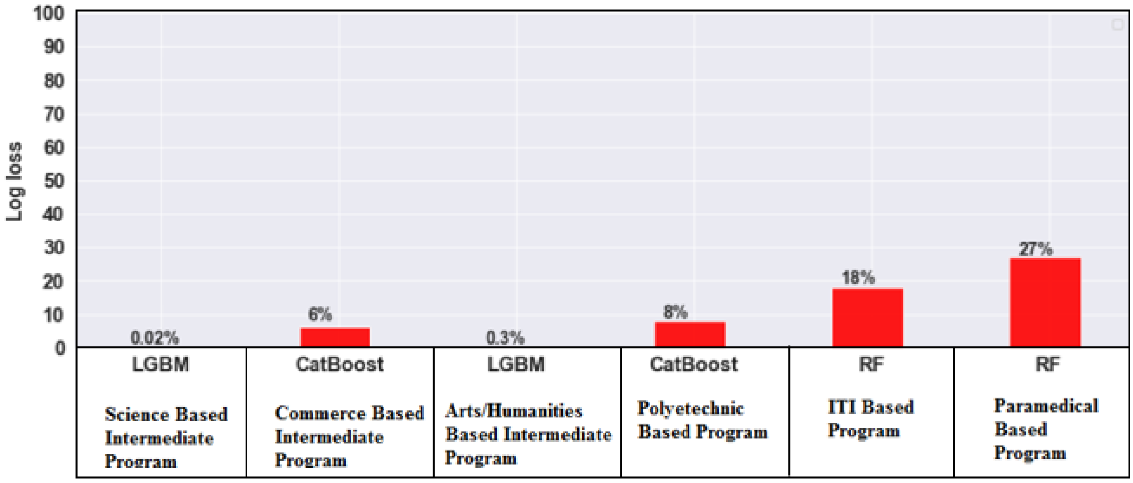
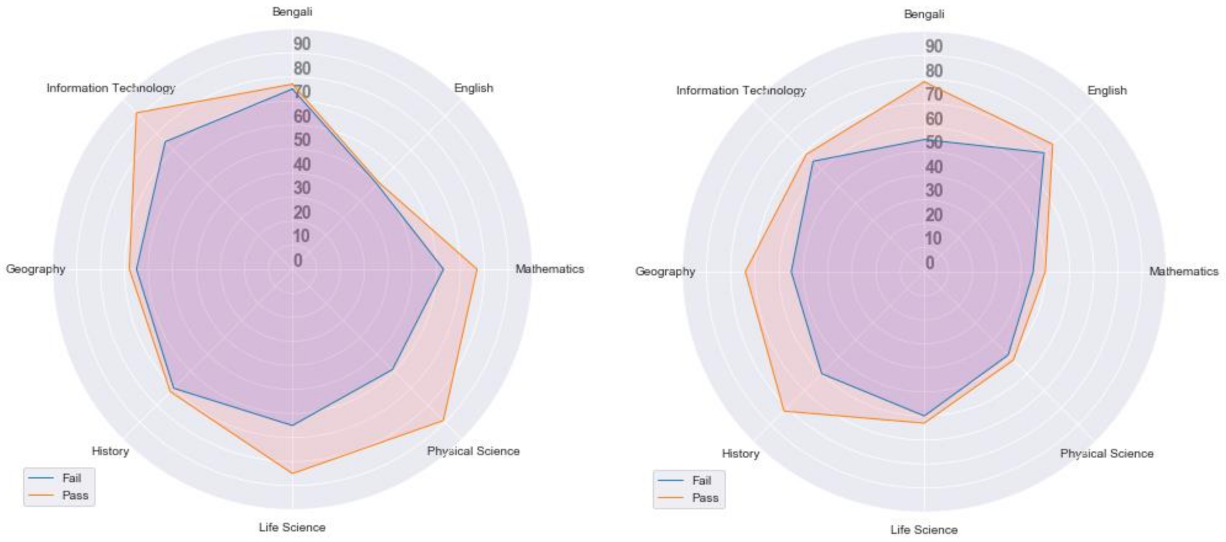


Figure 5. Values of Log loss for the proposed research



(A) Science Based Intermediate Program

(B) Arts/Humanities Based Intermediate

Figure 6. Radar chart for two sample educational programs used in this proposed research

4. CONCLUSION

After completing the 10th standard examination, most students do not understand which academic program can fulfill their dreams or which program is the most suitable for their further studies. Sometimes, they have taken the wrong decision, and this decision to make their life not so well. Hence, this paper presented an effective recommendation system that provides the best suitable academic path for each student's higher education. This study forecasts the best suitable educational program for each student to pursue the 12th standard class based on their past educational performance of the class 10th standard. This proposed research collects raw data of passing students' records of the class 10th standard from Hatgobindapur M.C. High School. Such raw data are then preprocessed and generate a new dataset after utilizing several machine learning techniques. This suggested research utilizes several machine learning algorithms to predict the most reliable academic path for pursuing the class 12th standard. The suggested research discovers effectively the most robust set of features and the most reliable machine learning classification algorithms for each academic program based on class 12th standard. The LightGBM algorithm is received as the best classification model for the science-based and arts/humanities-based intermediate program in terms of the F-measure value: 1.0, Cohen's Kappa value: 1.0, and ROC-AUC value: 1.0. The LightGBM algorithm is also gained as the best classification model for the Log loss values: 0.0002 and 0.003 for the science-based and arts/humanities-based intermediate program. While the CatBoost algorithm achieves the reward as the best classification model for the commerce-based intermediate program and the polytechnic based diploma engineering program in terms of the F-measure value: 0.98, Cohen's Kappa value: 0.96, and ROC-AUC value: 0.98. Moreover, the CatBoost algorithm is also awarded as the best classification model for the Log loss values: 0.06 and 0.08 for the commerce-based intermediate program and polytechnic based diploma engineering program.

In contrast, the Random Forest algorithm receives the reward as the best classification model for the ITI based program and paramedical-based program in terms of the F-measure values: 0.97 and 0.90, Cohen's Kappa values: 0.94 and 0.80, and ROC-AUC values: 0.97 and 0.90, respectively. Furthermore, the Random Forest algorithm is also awarded as the best classification model for the Log loss values: 0.18 and 0.27 for the ITI based and paramedical-based programs, respectively. Therefore, all the utilized best classification models provide averages of different performance evaluation metrics, in terms of F-measures value: 97.16%, Cohen's Kappa value: 94.33%, ROC-AUC value: 97.16%, and the Log loss value: 9.88% for all educational programs. Hence, this proposed research recommends the best suitable educational program for each student for their higher education. This proposed research's future work requires more datasets based on the educational program to solve other relevant academic problems and further verify the generated predictions. For this reason, different test datasets (relevant academic datasets) require to apply to this proposed system through which the proposed system will become more powerful.

REFERENCES

[1] Sonule, A.R., Kalla, M., Jain, A., Chouhan, D.S. (2020). Unsw-Nb15 dataset and machine learning based intrusion detection systems. International Journal of Engineering and Advanced Technology Regular Issue, 9(3): 2638-2648. https://doi.org/10.35940/ijeat.c5809.029320

[2] Ezz, M., Elshenawy, A. (2019). Adaptive recommendation system using machine learning algorithms for predicting student's best academic program. Education and Information Technologies, 25: 2733-2746. https://doi.org/10.1007/s10639-019-10049-7

[3] Rovira, S., Puertas, E., Igual, L. (2017). Data-driven system to predict academic grades and dropout. Plos One, 12(2): e0171207. https://doi.org/10.1371/journal.pone.0171207

[4] Goga, M., Kuyoro, S., Goga, N. (2015). A recommender for improving the student academic performance. Procedia - Social and Behavioral Sciences, 180: 1481-1488. https://doi.org/10.1016/j.sbspro.2015.02.296

[5] Kurniadi, D., Abdurachman, E., Warnars, H.L.H.S., Suparta, W. (2019). A proposed framework in an intelligent recommender system for the college student. Journal of Physics: Conference Series, 1402(6): 066100. https://doi.org/10.1088/1742-6596/1402/6/066100

[6] Thai-Nghe, N., Drumond, L., Krohn-Grimberghe, A., Schmidt-Thieme, L. (2010). Recommender system for predicting student performance. Procedia Computer Science, 1(2): 2811-2819. https://doi.org/10.1016/j.procs.2010.08.006

[7] Anjum, N., Cavallaro, A. (2008). Multifeature object trajectory clustering for video analysis. IEEE Transactions on Circuits and Systems for Video Technology, 18(11): 1555-1564. https://doi.org/10.1109/tcsvt.2008.2005603

[8] Anjum, N., Cavallaro, A. (2008). Multifeature object trajectory clustering for video analysis. IEEE Transactions on Circuits and Systems for Video Technology, 18(11): 1555-1564. https://doi.org/10.1109/TCSVT.2008.2005603

[9] Singh, S., Kumar, R. (2020). Histopathological image analysis for breast cancer detection using cubic SVM. 2020 7th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India. https://doi.org/10.1109/spin48934.2020.9071218

[10] Curtis, J.R., Yang, S., Chen, L., Park, G.S., Bitman, B., Wang, B., Navarro-Millan, I., Kavanaugh, A. (2011). Predicting low disease activity and remission using early treatment response to antitumour necrosis factor therapy in patients with rheumatoid arthritis: Exploratory analyses from the TEMPO trial. Annals of the Rheumatic Diseases, 71(2): 206-212. https://doi.org/10.1136/ard.2011.153551

[11] Chakraborty, S., Shaikh, S.H., Chakrabarti, A., Ghosh, R. (2020). A hybrid quantum feature selection algorithm using a quantum inspired graph theoretic approach. Applied Intelligence, 50: 1775-1793. https://doi.org/10.1007/s10489-019-01604-3

[12] Thomas, L., Kumar Manoj, M.V., Annappa, B. (2019). Clinical decision support system for early disease detection and management. Pre-Screening Systems for Early Disease Prediction, Detection, and Prevention Advances in Medical Diagnosis, Treatment, and Care, 108-155. https://doi.org/10.4018/978-1-5225-7131-5.ch005

[13] Barolli, L., Takizawa, M., Xhafa, F., Enokido, T. (2019). Web, artificial intelligence and network applications. Proceedings of the Workshops of the 33rd International Conference on Advanced Information Networking and Applications (WAINA-2019). Springer. https://doi.org/10.1007/978-3-030-15035-8

[14] Delgado, R., Tibau, X. (2019). Why Cohen's kappa should be avoided as performance measure in classification. PLOS ONE, 14(9): e0222916. https://doi.org/10.1371/journal.pone.0222916

[15] Kroncke, B.M., Duran, A.M., Mendenhall, J.L., Meiler, J., Blume, J.D., Sanders, C.R. (2016). Documentation of an imperative to improve methods for predicting membrane protein stability. Biochemistry, 55(36): 5002-5009. https://doi.org/10.1021/acs.biochem.6b00537