ISSN 2320 - 2602
Volume 9 No.7, July 2020
International Journal of Advances in Computer Science and Technology
Available Online at http://www.warse.org/IJACST/static/pdf/file/ijacst09972020.pdf
https://doi.org/10.30534/ijacst/2020/09972020

RESULT PREDICTION SYSTEM USING MACHINE LEARNING

Megha S [1], Nitha Johny[2], Raicy Ann Mammen[3], Sandra M[4], Ms. Neena Joseph[5]
[1]Department of Computer Science, India, smegha4321@gmail.com
[2] Department of Computer Science, India, nithajohny277@gmail.com
[3] Department of Computer Science, India, annraicy@gmail.com
[4] Department of Computer Science, India, sandramprasanth55@gmail.com
[5]Faculty Department of Computer Science, India, neena.joseph@mangalam.in

ABSTRACT

Higher education institutions are often very curious about the success rate of the students throughout their study. Classification and prediction of student's performance in examination are the typical challenges for educators. Machine learning is expected in the near future to provide various venues and effective tools to improve the education in general.The proposed system is an academic result analysing and predicting software using machine learning algorithms. The system is implemented as a UI application. Datasets are processed in the TensorFlow framework. The dataset containing internal exam marks, attendance percentage, timely submission of assignments, previous semester university result etc will be the input. The results can be predicted as whether pass or fail for each subject. SVM algorithms are used for processing the data and more accurate result will be outputted. Front-end and back-end are developed in Python language, where Flask is used in front-end.

Key words: Machine Learning, UI application, Sklearn, Flask.

1. INTRODUCTION

Education is the cornerstone of the society. Hence every educational organizations intend to teach and train in a way that each student can perform well. Inspite of having the uniform platform to learn in schools or colleges, the performance of students varies enormously. This can be ascribed to difference in cognition level, motivation levels and environmental factors. Now with increased computational power it is also simpler to train machines the various nuances of the teaching learning domain. Here we are using machine learning to identify the performance of a student. Machine learning has been successfully applied in the domain of educational data mining (EDM) which includes domains like performance prediction, student modeling, analysis and visualization of student data, recommendation system, student learning curve analysis etc.

Based on the past examination results and the class assessments, it is possible to forecast the future development of the students. It is a challenging and important as it involves the large volume of data in educational databases and the result could impact the future development of a young kid. A good and accurate prediction can bring the benefits and impacts to students, educators and academic institutions. Various types of data mining techniques have been used for performance prediction from decades including Decision Tree, Naive Bayes, K-Nearest Neighbor and Support Vector Machine (SVM). However, with the rise of artificial intelligence and deep learning application, and also using AI engine such as Google TensorFlow for pattern recognition has now been rising its importance.

Artificial intelligence in general, and, machine learning (ML), in particular, has the potential to revolutionize Science-Technology-Engineering-Math (STEM) education. The main venues to improve education using ML algorithms-based techniques include but are not limited to cases such as 1) customizable student learning experience; 2) student path prediction; 3) better organization of learning content, curriculum and learning process; 4) automatic and unbiased grading system; 5) overall feedback on both students' and teachers' performance; 6) suggested learning path; 7) matching students and teachers; 8) educational experiments. The model creates a solid foundation for the conceptual modeling and software engineering of smart learning analytics (SLA) systems that would have and actively use the unique features of smart university, smart education, and smart classroom, including 1) adaptively, 2) sensing, 3) inferring, 4) anticipation, 5) self-learning, and 6) self-organization. In this, we will investigate how to use artificial intelligence and deep learning algorithm for pattern recognition and correlation of assessment results. There are also some more traditional data mining techniques that have been used to predict student's performance.

1.1 Machine Learning Overview

Machine learning is a new approach to learn and analyze complex and huge data. It is based on algorithms that can learn from data without relying on the conventional programming, i.e., rules-based programming. The primary aim is to allow the machines like computers, etc., to learn automatically without human intervention or assistance and adjust actions accordingly [1]. It emerged individually as a scientific discipline in the late 1990s as steady advances in digitization and cheap computing power enabled data. Scientists have stopped building finished models and have plunged into a novel adventure in training computers, through which an unmanageable volume of data and complexity of the big data can be processed and analyzed using the potentiality of machine learning. But, using classical algorithm of machine learning, text is considered as a consequence of keywords; instead, an approach based on semantic analysis mimics the human ability to understand the meaning of a text. Hence, machine learning is an emerging trend in the era of information technology.

1.2 Confusion Matrix Classifier

A confusion matrix is a summary of prediction results on a classification problem. The number of correct and incorrect predictions are summarized with count values and broken down by each class. This is the key to the confusion matrix. The confusion matrix shows the ways in which your classification model is confused when it makes predictions. It gives us insight not only into the errors being made by a classifier but more importantly the types of errors that are being made. The number of correct and incorrect predictions are summarized with count values and broken down by each class. This is the key to the confusion matrix. Out of all positive classes, how much we predicted correctly , it should be as high as possible [2]. This matrix shows the ways in which your classification model is confused when it makes predictions. It gives an insight not only into the errors being made by our classifier but more importantly the types of errors that are being made. It is this breakdown that overcomes the limitation of using classification accuracy alone. Confusion matrix can be calculated by following process: 1) Test a dataset or validate dataset with expected outcome values, 2) Make a prediction for each row in test dataset, 3) Predicts the count from the expected outcomes.

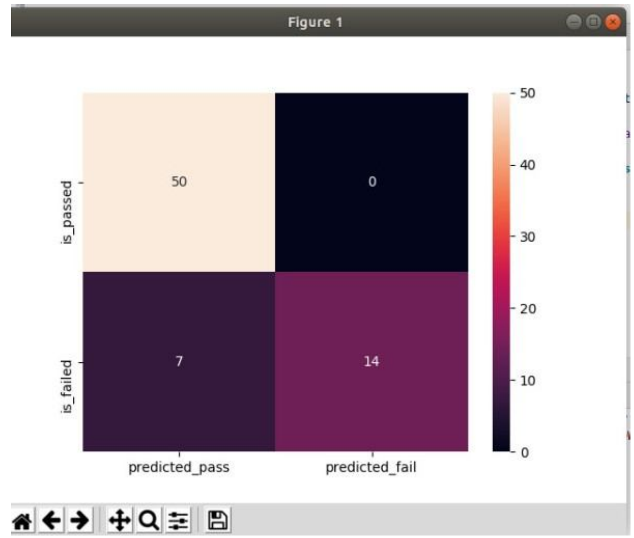


Figure 1: Confusion Matrix of our system

1.3 Support Vector Machine Overview

Support Vector Machine is a supervised machine learning algorithm that can be used for both classification or regression challenges. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data item as a point in n-dimensional space with the value of each feature being the value of a particular coordinate. Here in this, we need not have to add this feature manually to have a hyper-plane, the SVM algorithm has a technique called the kernel trick. The SVM kernel is a function that takes low dimensional input space and transforms it into a higher-dimensional space. It does some extremely complex data transformations, then finds out the process to separate the data based on the output we have provided. The main reason for us to choose SVM as our algorithm as by the reasons, 1) effective in high dimensional space [3], 2) it is effective as in our case we have more number of dimensions than number of samples, 3) memory efficient as there uses subset as training points, and 4) it has a versatile nature.

We are using Python as our system language. As in Python, scikit-learn is a widely used library for implementing machine learning algorithms. SVM is also available in the scikit-learn library and we follow the same structure for using it. We are using this classifier as it works really well with a clear margin of separation, is effective in high dimensional spaces, is effective in cases where the number of dimensions is greater than the number of samples and moreover it uses a subset of training points in the decision function called support vectors, so it is also memory efficient. Here we are using previous year pass out students result as a subset of training to analyze the results of the students which we desired to find out whether a student will pass or not in the upcoming examination.

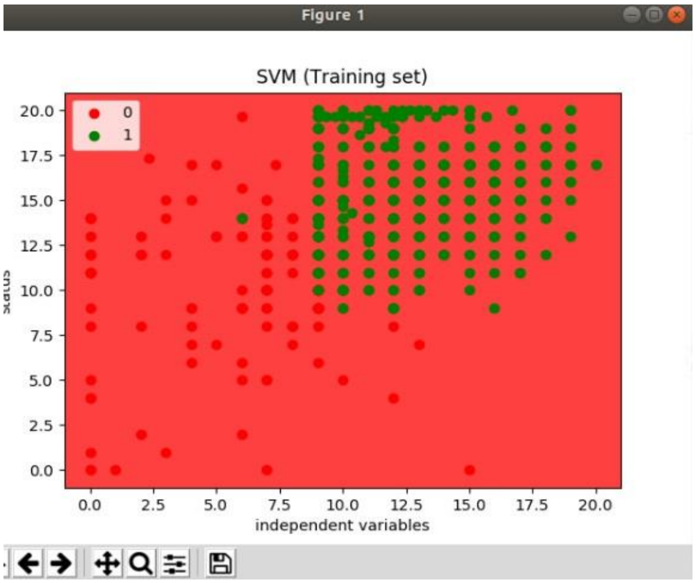


Figure 2: Training set in SVM

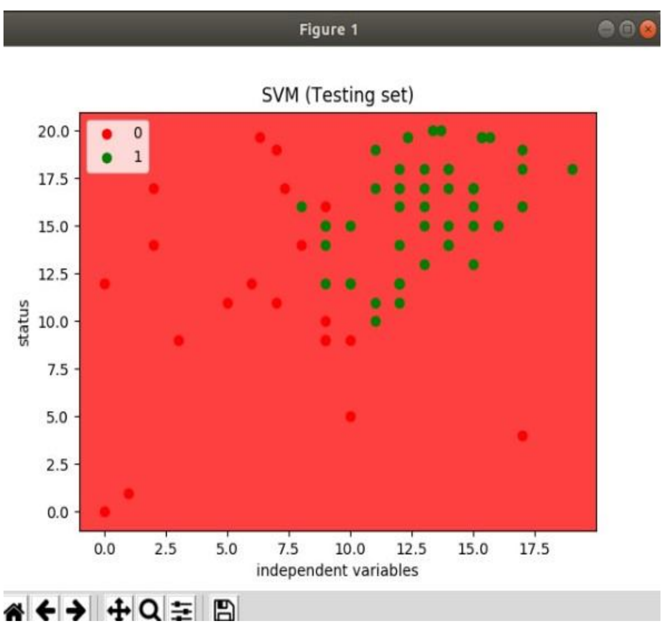


Figure 3: Testing set in SVM

2. LITERATURE SURVEY

In LMS, every action is stored and evaluated as its insight can be obtained into students' behavior online, which can be used to increase the capability of learning as well as teaching. The learning analytics which is the analytics of LMS data is defined to be the collection and reporting measurement of data about the learners and the purposes of understanding and learning their concept and in which environments it occurs. LMS data is used in the field of learning analytics for the prediction of student performance model with the prediction of their grades as well as how many are at the risk of failing a course [4]. This is an important step in learning analytics, as it informs the implementation of interventions, such as personalized feedback.

Studies predicting student success in offline education have typically collected measurements using validated questionnaires, interviews, and observational techniques, with relevant theoretical concepts in mind so that the measurement can be geared towards the concepts that the researcher thinks need to be measured. The use of LMSs allows for tracing and analyzing students' online behavior without the necessity of time-consuming data-collection. However, LMSs provide raw log data that are not concrete measurements of previously outlined theoretical concepts. It is therefore important to understand whether and how these data can be used for learning analytics. Moreover, the question is whether there is actually a single best way to predict students' performance across a diverse set of courses and by evaluating their performance. Studies that have used similar methods and predictors models. Even within an institution that is using the same LMS, such differences have been found in the prediction models of the nine blended courses. Thus, the effects of LMS behavior on student performance might differ per institution or even per course. Indeed, a study using 29 courses (204 offerings, 352 unique students), has found that the variance in students' performance (final grade), was accounted for by individual differences (18 percent) as well as course offerings (22 percent) [4].

In addition, most studies focus on predicting student performance after a course has finished, establishing how well student performance could have been predicted with LMS usage data, but at a point in time where the findings cannot be used for timely intervention anymore. As LMS data provide information during the whole course, it seems useful to determine whether data from only the first weeks of a course are enough for accurate prediction of student performance. In the current study, the authors add to the analysis of the portability of prediction models and the accuracy of timely prediction. First, we provide an overview of the theoretical arguments used in learning analytics and the predictors that have been used in recent studies. The predictive value of these predictors will be examined in 17 blended, undergraduate courses taught at the same institution (Eindhoven University of Technology). This allows us to establish the effects of different types and degrees of LMS usage while controlling at least to some extent for contextual effects. Furthermore, the portability of the prediction models across the 17 courses, i.e., the effect of course, is analyzed. For this, they replicate the study of Ga'sevi'c et al within another institution with a larger sample of more similar courses [5]. Moreover, to ensure the comparability of findings, they only use predictors that are available for all courses. In addition, they analyze whether it is possible to identify students at-risk early on in a course, and to what extent these models can be used to generate targeted interventions.

It is increasingly evident that a large number of college students do not complete the courses on which they enroll, mainly for courses with lower entry requirements. Enrolment numbers to tertiary education are increasing, as is the academic and social diversity which is in the student population. This adds to the challenge of both identifying students at risk of failing and providing appropriate supports and learning environments to enable all students to perform optimally. Learning is a latent variable that is typically measured as academic performance in assessment work and examinations [6]. Factors impacting academic performance have been the focus of the research area for many years. It continues as an active research topic, indicating the inherent difficulty in generating models of learning, particularly in tertiary education. Such education providers collect an increasing volume of data on their students. Therefore, the application of data mining to educational settings is emerging as an evolving and growing research discipline. Learning Analytics (LA) and Educational Data Mining (EDM)aim to better understand students and how they are learning through the use of data analytics on educational data [6]. The objective of this paper is to investigate the accuracy of classification models that identify students at risk of failing in the first by analyzing their academic performance. This study considers both linear and non-linear classification models. Model accuracy which is estimated using cross-validation was compared to model performance when tested on students from a different group. Participating students were from a diverse student population that included mature students, students having disabilities, and students from disadvantaged socio-economic backgrounds. Academic performance was measured as Grade Point Average (GPA), an aggregate score of both semester 1 and semester 2 module grades that were compiled from both the end of term module exams and continuous assessment work.

Student's performance prediction seen to be of dire importance to most academic institutions for higher learning. This may lead to many kinds of research in prediction works that included many students from different backgrounds and academic areas such as MBA students, nursing students and that of Computer background students. Because the predicators variables or the independent values were mostly dealt with demographic profiles. Data collected was mostly from survey forms based on the students' recent background education, residency region, gender and Scholastic Aptitude Test (SAT) scores. Statistical Packages for Social Sciences (SPSS) have been very popular among past researchers that utilize Linear Regression, Data Mining techniques, and Decision Trees. Artificial Neural Network (ANN) came into the picture for student's performance prediction quite recently [7]. These NN models were also considered as demographic backgrounds as inputs to the model. The study presented here with because only considering the Grade Point (GP) of fundamental subjects scored by the students in the first semester as the inputs without considering their former background education or family background. Once the students are accepted into the Program based on merits set by the Faculty of Electrical Engineering UiTM, then everyone should help to be offered to help students to perform in their study before graduation so as to provide them with basic keys into their future lives. Thus such novel effort from academic lecturers is in line with the philosophy of the Universities which states that every individual student has the ability to attain excellence through the transfer of knowledge and the assimilation of moral values so as to they become professional graduates capable for developing knowledge, self, society, and worldwide. The paper describes the development of ANN to predict the performances of bachelor degree engineering students based on the intakes from Matriculation and Diploma level entries. Data of Matriculation and Diploma level intake students were compiled in Excel format which included student reg no, gender, CGPA at semester eight, GP of subjects scored at semester one to semester three for Matriculation students [7]. For the Diploma students, the Grade Points of subjects start with semester three onwards. Such students have a direct entry into the semester three of the Program due to credit exemption for courses in semester one and two. Such data were collected from the Students Information Management System (SIMS) in UiTM developed to store students' academic information until graduation. Such software helps academic advisors to keep track of students' achievement right from the very start of Program.

Many studies have predicted that the future performance of companies for the purposes of making investment decisions. Most of them are based on the qualitative judgments of experts in related industries, who are considered as various financial and firm performance information. However, the qualitative judgments are very highly subjective and are limited in the sense that conclusions come at a significant cost in terms of time and money. With the recent developments in technology, researchers have begun to using machine learning techniques to predict corporate performance. For example, the artificial neural networks, which have relatively good predictive ability in various fields, are widely used in various studies.

However, the models are based on the artificial neural networks often suffer from the problem of overfitting in training data. However, for training a deep neural network takes more time, and the propagation of errors, based on the backpropagation algorithm, back to the input layers can be very difficult. The results of many studies have been shown that prediction models constructed using a support vector machine (SVM), as suggested by Vapnik, have good predictive performance and very fast learning speed [8]. As a result, many researchers have investigated using an SVM machine to predict the corporate performance and stock prices. Recently, the artificial neural network-based for predicting the models have been again attracting attention owing to the development of many parallel processing technology, as well as the algorithms that overcome the limitations of deep neural networks. The algorithm which is used to train a deep neural network is the deep belief network (DBN). A DBN performs the pretraining through unsupervised learning using the restricted Boltzmann machine (RBM) and using fine-tunes the network via supervised learning on training data [8]. In addition, the convolutional neural networks, which are widely used in image processing and voice recognition, demonstrate good performance which is widely used in constructing the classification models in various fields. In order to construct the corporate performance prediction models, the predictors are needs to predict the performance of companies. Most of the corporate performance prediction models use the company's financial performance data and financial indicators as the predictors. However, there is a recent increase in the proportion of technology-intensive firms whose technological capability significantly increases their corporate performance. Thus, in order to predict the corporate performance more accurately, it is most necessary to use both a company's financial information and technical information as predictors. As a result, there are many recent studies that have proposed indicators that show the technological competitiveness of the company. Many of these studies apply for patent data, because they are easier to use in more quantitative analyses, have an internationally uniform structure, and citation information. Among the recent studies predict a company's performance, there are less than construct a prediction model using patent data and the deep learning algorithm. In this, we propose a deep neural network-based performance prediction model that uses the company's financial and technical indicators as predictors. The proposed model includes the unsupervised learning phase, using an RBM, in which the training uses the entire training data set [8]. Then, there is a very fine-tuning phase, which uses the backpropagation algorithm and a relatively up-to-date training data set.

3. SYSTEM ENVIRONMENT AND DESIGN

The objective of this work is to create a web service for the prediction of students results based on their criteria such as their attendance, series exams scores, assignment scores, previous academic scores, etc. It is decided to use Flask web framework for frontend. While using Flask, the front will comprise of HTML, JavaScript, and CSS. Additionally, Python programming tool and Machine Learning Studio are used for the complete execution of the work.

The steps to be followed for the students result prediction are given as follows:

A. Data Set Preparation:

A data set is prepared in the form of csv file to give training to the machine and testing it.

B. Training Model:

Train the machine model based on the data set. The train model has two inputs: 1) SVM algorithm, 2) data provided by the user.

C. Classifier Model:

Here we are using confusion matrix as our classification model on our set of testing data. As this will easily identify any confusion between two classes.

D. Evaluation Model:

This model evaluates the score result and calculate the machine learning parameters: true positive, true negative, false positive, false negative, receiver operating characteristic (ROC), precision, accuracy, recall and F1 score.

E. Web service deployment and publishing

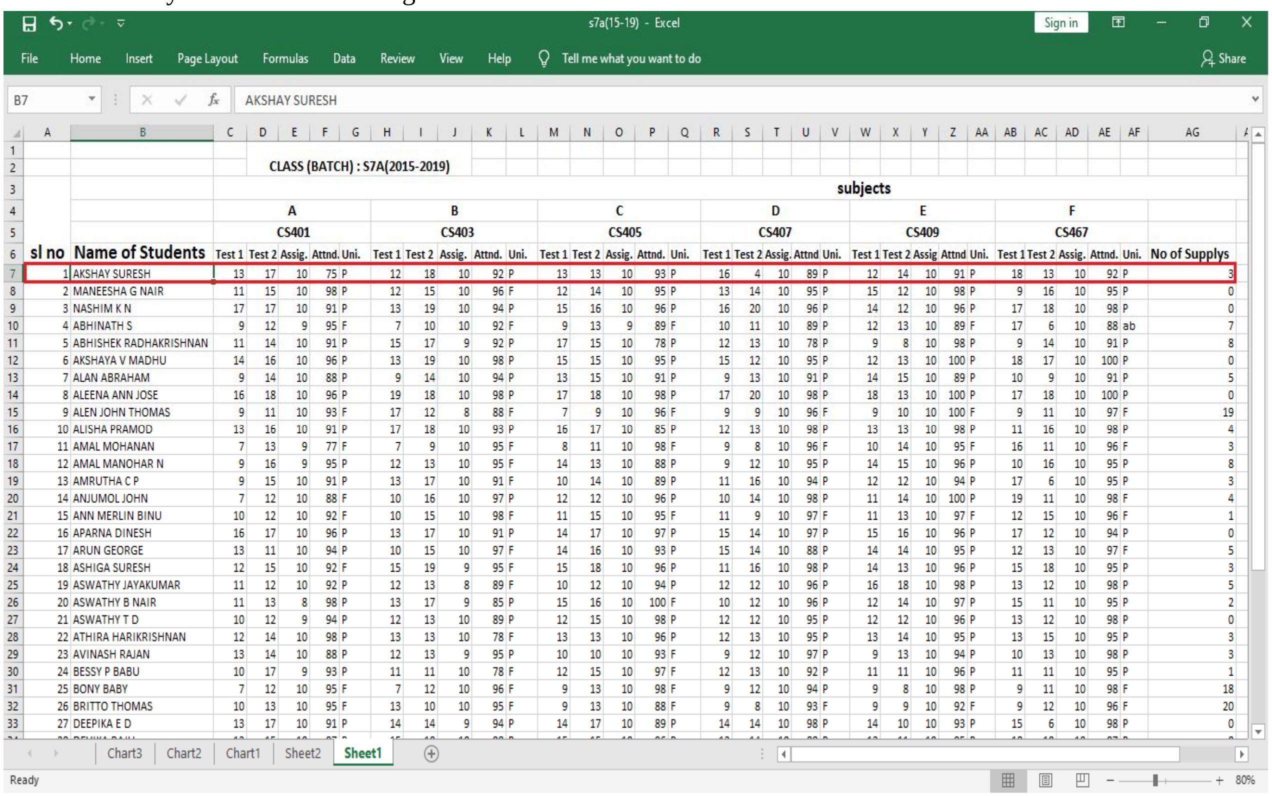
The entire model is deployed in the form of web service so that HOD or class in charge or even other faculty members can access this system.

4. SYSTEM EVALUATION

We are publishing our model as web services that can be easily be consumed by custom apps or business intelligence (BI) tools such as Excel. To develop a predictive analysis model, we can typically use data from one or more resources transform and analyse that data through various data manipulation and statistical functions and generate a set of results.

Let A be true positive, B be false negative, C be false positive and D be false negative. Then accuracy, recall, precision and F1 score can be calculated using the formulae given below:

Accuracy = (A+D)/(A+B+C+D) ------ (1)



CLASS (BATCH) : S7A(2015-2019)

subjects

sl no	Name of Students	A CS401 Test 1	Test 2	Assig.	Attnd.	Uni.	B CS403 Test 1	Test 2	Assig.	Attnd.	Uni.	C CS405 Test 1	Test 2	Assig.	Attnd.	Uni.
1	AKSHAY SURESH	13	17	10	75	P	12	18	10	92	P	13	13	10	93	P
2	MANEESHA G NAIR	11	15	10	98	P	12	15	10	96	F	12	14	10	95	P
3	NASHIM K N	17	17	10	91	P	13	19	10	94	P	15	16	10	96	P
4	ABHINATH S	9	12	9	95	F	7	10	10	92	F	9	13	9	89	F
5	ABHISHEK RADHAKRISHNAN	11	14	10	91	P	15	17	9	92	P	17	15	10	78	P
6	AKSHAYA V MADHU	14	16	10	96	P	13	19	10	98	P	15	15	10	95	P
7	ALAN ABRAHAM	9	14	10	88	P	9	14	10	94	P	13	15	10	91	P
8	ALEENA ANN JOSE	16	18	10	96	P	19	18	10	98	P	17	18	10	98	P
9	ALEN JOHN THOMAS	9	11	10	93	F	17	12	8	88	F	7	9	10	96	F
10	ALISHA PRAMOD	13	16	10	91	P	17	18	10	93	P	16	17	10	85	P
11	AMAL MOHANAN	7	13	9	77	F	7	9	10	95	F	8	11	10	98	F
12	AMAL MANOHAR N	9	16	9	95	P	12	13	10	95	F	14	13	10	88	P
13	AMRUTHA C P	9	15	10	91	P	13	17	10	91	F	10	14	10	89	P
14	ANJUMOL JOHN	7	12	10	88	F	10	16	10	97	P	12	12	10	96	P
15	ANN MERLIN BINU	10	12	10	92	F	10	15	10	98	F	11	15	10	95	F
16	APARNA DINESH	16	17	10	96	P	13	17	10	91	P	14	17	10	97	P
17	ARUN GEORGE	13	11	10	94	P	10	15	10	97	F	14	16	10	93	P
18	ASHIGA SURESH	12	15	10	92	F	15	19	9	95	F	15	18	10	96	P
19	ASWATHY JAYAKUMAR	11	12	10	92	P	12	13	8	89	F	10	12	10	94	P
20	ASWATHY B NAIR	11	13	8	98	P	13	17	9	85	P	15	16	10	100	F
21	ASWATHY T D	10	12	9	94	P	12	13	10	89	P	12	15	10	98	P
22	ATHIRA HARIKRISHNAN	12	14	10	98	P	13	13	10	78	F	13	13	10	96	P
23	AVINASH RAJAN	13	14	10	88	P	12	13	9	95	P	10	10	10	93	F
24	BESSY P BABU	10	17	9	93	P	11	11	10	78	F	12	15	10	97	F
25	BONY BABY	7	12	10	95	F	7	12	10	96	F	9	13	10	98	F
26	BRITTO THOMAS	10	13	10	95	F	13	10	10	95	F	9	13	10	88	F
27	DEEPIKA E D	13	17	10	91	P	14	14	9	94	P	14	17	10	89	P

sl no	Name of Students	D CS407 Test 1	Test 2	Assig.	Attnd	Uni.	E CS409 Test 1	Test 2	Assig	Attnd	Uni.	F CS467 Test 1	Test 2	Assig.	Attnd.	Uni.	No of Supplys
1	AKSHAY SURESH	16	4	10	89	P	12	14	10	91	P	18	13	10	92	P	3
2	MANEESHA G NAIR	13	14	10	95	P	15	12	10	98	P	9	16	10	95	P	0
3	NASHIM K N	16	20	10	96	P	14	12	10	96	P	17	18	10	98	P	0
4	ABHINATH S	10	11	10	89	P	12	13	10	89	F	17	6	10	88	ab	7
5	ABHISHEK RADHAKRISHNAN	12	13	10	78	P	9	8	10	98	P	9	14	10	91	P	8
6	AKSHAYA V MADHU	15	12	10	95	P	12	13	10	100	P	18	17	10	100	P	0
7	ALAN ABRAHAM	9	13	10	91	P	14	15	10	89	P	10	9	10	91	P	5
8	ALEENA ANN JOSE	17	20	10	98	P	18	13	10	100	P	17	18	10	100	P	0
9	ALEN JOHN THOMAS	9	9	10	96	F	9	10	10	100	F	9	11	10	97	F	19
10	ALISHA PRAMOD	12	13	10	98	P	13	13	10	98	P	11	16	10	98	P	4
11	AMAL MOHANAN	9	8	10	96	F	10	14	10	95	F	16	11	10	96	F	3
12	AMAL MANOHAR N	9	12	10	95	P	14	15	10	96	P	10	16	10	95	P	8
13	AMRUTHA C P	11	16	10	94	P	12	12	10	94	P	17	6	10	95	P	3
14	ANJUMOL JOHN	10	14	10	98	P	11	14	10	100	P	19	11	10	98	F	4
15	ANN MERLIN BINU	11	9	10	97	F	11	13	10	97	F	12	15	10	96	F	1
16	APARNA DINESH	15	14	10	97	P	15	16	10	96	P	17	12	10	94	P	0
17	ARUN GEORGE	15	14	10	88	P	14	14	10	95	P	12	13	10	97	F	5
18	ASHIGA SURESH	11	16	10	98	P	14	13	10	96	P	15	18	10	95	P	3
19	ASWATHY JAYAKUMAR	12	12	10	96	P	16	18	10	98	P	13	12	10	98	P	5
20	ASWATHY B NAIR	10	12	10	96	P	12	14	10	97	P	15	11	10	95	P	2
21	ASWATHY T D	12	12	10	95	P	12	12	10	96	P	13	12	10	98	P	0
22	ATHIRA HARIKRISHNAN	12	13	10	95	P	13	14	10	95	P	13	15	10	95	P	3
23	AVINASH RAJAN	9	12	10	97	P	9	13	10	94	P	10	13	10	98	P	3
24	BESSY P BABU	12	13	10	92	P	11	11	10	96	P	11	11	10	95	P	1
25	BONY BABY	9	12	10	94	P	9	8	10	98	P	9	11	10	98	F	18
26	BRITTO THOMAS	9	8	10	93	F	9	9	10	92	F	9	12	10	96	F	20
27	DEEPIKA E D	14	14	10	98	P	14	10	10	93	P	15	6	10	98	P	0

Figure 4: A database sheet that is used as a training set for the system

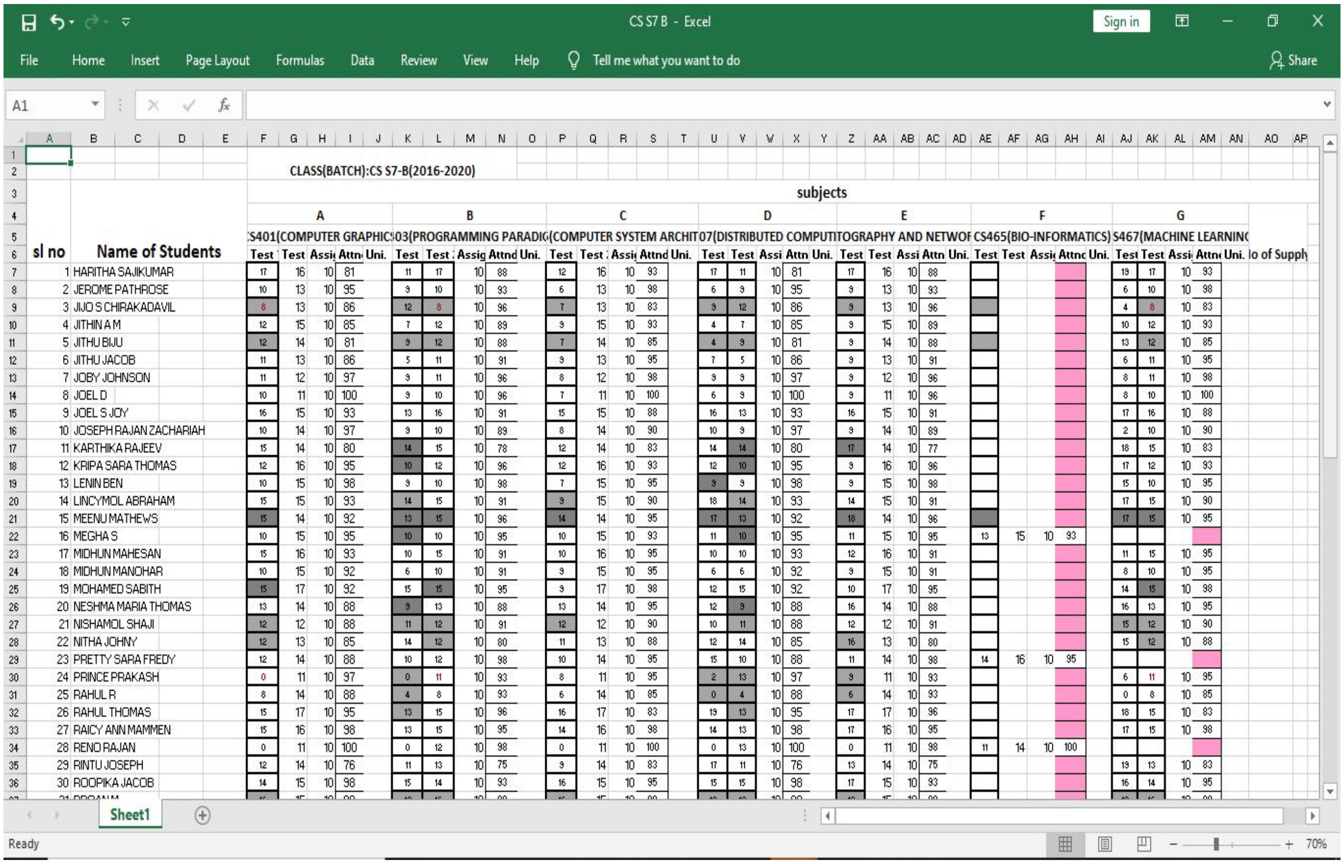


Figure 5: A sample database sheet for student data entry

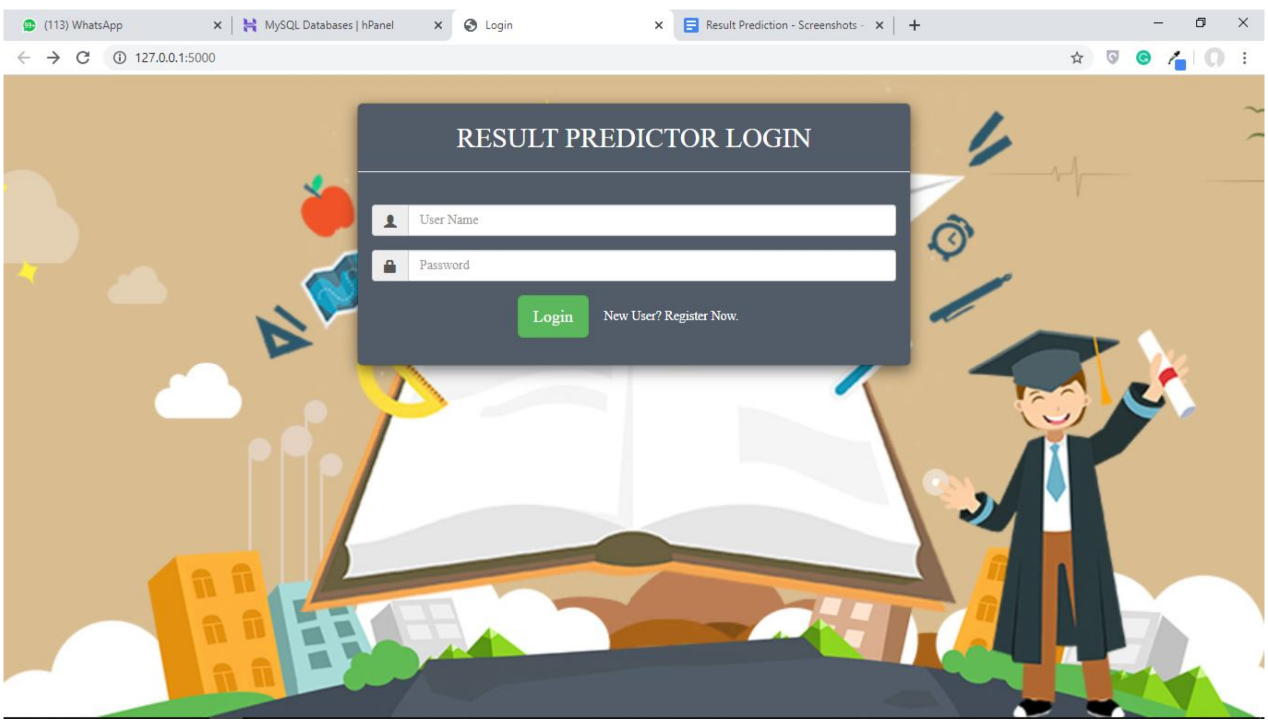


Figure 6: Login page

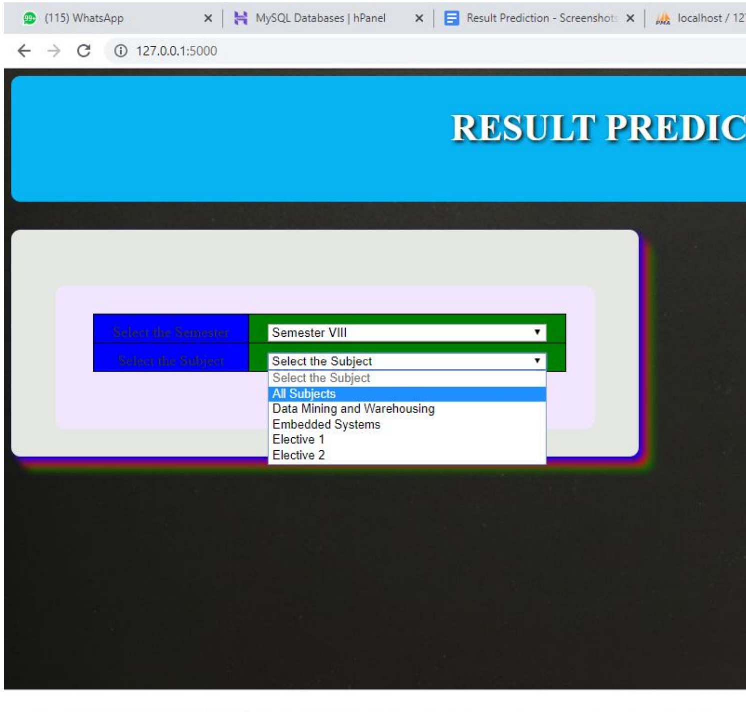


Figure 7: Home page: Admin

Figure 9: Result page- Admin view (subject wise)

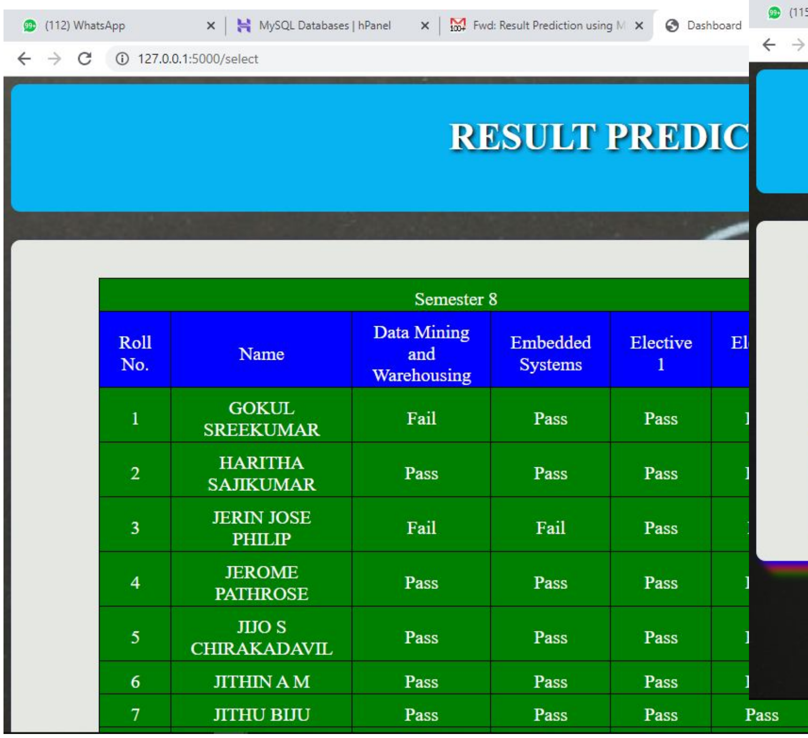


Figure 8: Result page- Admin view (all subjects wise)

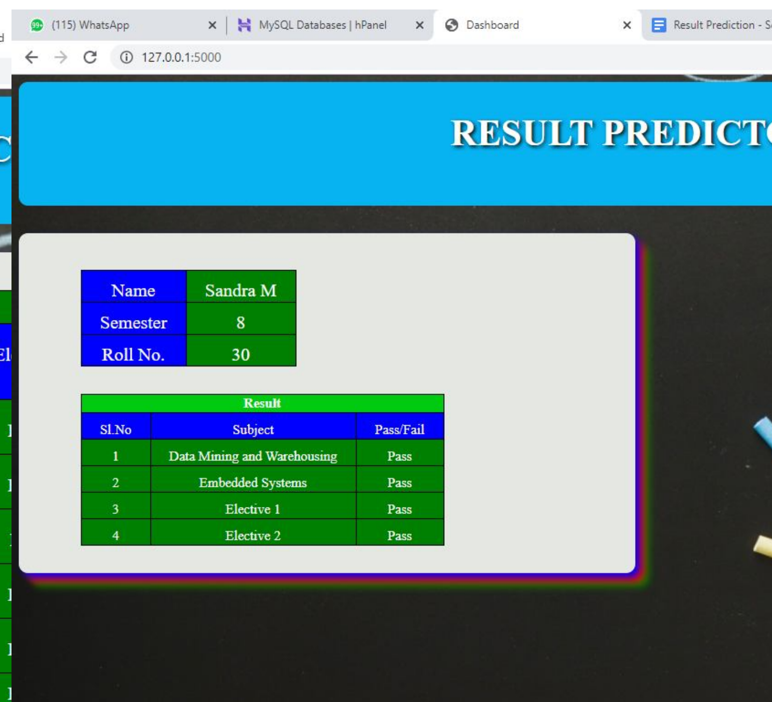


Figure 10: Result page: User

Recall = A/(A+D) ------ (2)

Precision = A/(A+C) ------ (3)

F1 score = (2*Recall*Precision)/(Recall + Precision) --- (4)

Initially a csv file is created for training the system. For this we have taken the database sheet of our pass out seniors. The sheet was created with the fields: 1) Student full name, 2) Subject wise: a) Two series scores, b) Assignment scores, and c) Attendance, and 3) University result. Similarly, the file which need to be get predicted is created in the exact format that of the training set except the university result column which is to be predicted by the system. The pass prediction is done based on the academic criteria which has to be manually entered. The figure 4 shows a sample database sheet which is used to train the system and figure 5 shows a sample database sheet for student data entry.

5. CONCLUSION AND FUTURE SCOPE

The accuracy of the system is given as 0.901 or 90%. The above obtained research outcomes and findings enabled us to make many conclusions. The proposed software shows improvement in prediction the academic results of students in a very efficient manner where the involvement of manual work is very less. The outcomes of formative and summative surveys of students clearly shows students interest in ML-based predictive analytics of student academic performance in a course. This would be helpful for the students, teachers and the overall educational institutions and make awareness on those students who are not met and they can keep their focus on those students who are not eligible. This work tries to improve students' quality of education.

However, to have a deeper insight on the educational data and academic performances, we also need to have more social psychological investigations that are beyond our scope. In the future, we intent to collect and analyze more data, in order to refine our findings and extracts more valuable information to improve by adding more techniques and incorporating more machine learning approaches

ACKNOWLEDGEMENT

We thank the faculty members of department of Computer Science and Engineering, Mangalam College of Engineering for providing us the data required.

REFERENCES

1. https://expertsystem.com/machine-learning-definition/
2. https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62
3. https://scikit-learn.org/stable/modules/svm.html.
4. Rianne Conijn, Chris Snijders, Ad Kleingeld, Uwe Matzat, “Predicting Student Performance from LMS Data: A Comparison of 17 Blended Courses Using Moodle LMS. October 2016IEEETransactions on Learning Technologies PP (99):1-1
5. Amirah Mohamed, Shahiri Wahidah, “A Review on Predicting Student's Performance Using Data Mining Techniques”, Procedia Computer Science, Volume 72, 2015, Pages 414-422, https://www.sciencedirect.com/science/article/pii/S1877050915036182
6. G. Gray, C. Mc Guinness, P. Owende, An application of classification models to predict learner progression in tertiary education, in: Advance Computing Conference (IACC), 2014IEEE International, IEEE, 2014, pp. 549–554.
7. P. M.Arsad, N.Buniyamin, J.-l. A. Manan, A neural network students' performance prediction model (nnsppm), in: Smart Instrumentation, Measurement and Applications (ICSIMA), 2013 IEEE International Conference on, IEEE, 2013, pp. 1–5.
8. Joonhyuck Lee, Dongsik Jang and Sangsung Park, “Deep Learning-Based Corporate Performance Prediction Model Considering Technical Capability” MDP http://www.mdpi.com/2071-1050/9/6/899/pdf.