Prediction and Analysis of Key Performance Indicators (Kpi) For Students using Data Science

Mrs. G. Chandra Praba
Dept.Of cse, Kings college of engineering,
Punalkulam, Pudukottai, Tamil Nadu.

Mrs. K. Abhirami
Dept. Of cse, Kings college of engineering,
Punalkulam, Pudukottai, Tamil Nadu.

Mrs. R. Suganthalakshmi
Dept.Of cse, Kings college of engineering,
Punalkulam, Pudukottai, Tamil Nadu.

ABSTRACT-Many organizations need future analysis data to overcome the pitfalls or the improvements to be made, by using analytical tools.Predicting student's performance is the mostintricate task due to the large volume of data in student's databases. The performance of a student is predicted mainly by considering their academic details. Better anticipation of student's success in higher academic institutions is one approach to attain top level of quality in education system. This can be used to concentrate more on the students who needs a bit higher attention and to train them for reaching better in academics.Students' data can be evaluated with the help of various techniques. Data science is the most prevalent techniques to evaluate students' performance and is widely usedin educationalsector.The main objective of this paper is to identify the key performance indicators that affect the results (success and failure) of the student in the course and to analyze the various classification models and identify a high accuracy prediction model to predict the result.

Keywords: KPI, Student's Performance

I.INTRODUCTION

In this modern era, all has become digitized and because of that numerous data are retrieved from various sources. Making the most useful information from these data has become a tedious and a very important task. Especially, in the educational sector, more data are collected. Using these data to forecast the future result of the student makes sense in collecting those data. Data science can help more objectively, evaluate the candidates androot out inefficiencies and biases to predict the future using machine learning algorithms. Key performance indicator helps in measuring the performance of the student and is essential to use in data science for good communication. Student's performance measurement is a necessary part in educational institutions. There are a lot of algorithms to calculate the student's performance. It can be obtained by measuring the academic performance like assessment and co-curriculum activities. The existing methodology about the grades in school being the measure of student's success. The most of higher educational institutions uses the final grades to evaluate overall performance of students. Finally prediction are based on course structure, assessment mark, and also extracurricular activities.

The evaluation is important to maintain student's performances and the effectiveness of learning process. By investigating the students' performance a strategic program can be well planned during their period of studies in an institution for better outcome. Currently, there are many techniques being proposed to evaluate performance of students. Data science has been widely applied in educational field.

The proposed system also helps the learner, institutions and faculties to work according to the learning criteria of the students. Actually data science help in the education sector. To understand, analyse and then find the difference between different prediction methods of data Science in education.

II. RELATED WORKS

In [1], educational data science is widespread nowadays due to increase in e- resources, usage of online tools for education and Internet. Lots of research is taking place to make best of education tools and technologies. The usage of these techniques is to predict or analyse the students' performance and improve the students who are falling below satisfactory grades, an artificial neural network classifier model was built which can be beneficial for both students and teachers to discover knowledge from huge data present in educational sector. Student's behavioural features were considered with other featuresandamodelwasproposedbasedon data science techniques which yielded 22.1% high accuracy after removing behavioural features. Further by employing ensemble methods there was found 25.8%increase in accuracy. Academic data set consisting of 473 instances, and found that 70% accuracy was yielded by Bayesian classifier. The naive Bayes classifiers, KNN were used to categorize student'sdropouts. 87% and 79.7% accuracy was yielded by K- nearest neighbours and decision trees applying 10-foldcross-validation.

In [2], it represents the data science techniques used for analysing pupil performance. Educational institutions contain an enormous amount of academic database containing student details. These student databases along with other attributes are taken into consideration like family background, family income, etc. It will helps to identifying promising students and by providing a chance to refine those students who likely get low marks. To prepare a structure which will investigate the performance from their previous performancesusingconceptsofDatascience under Classification? Classification Algorithms like Decision Tree, Naive Bayes and Support Vector Machine (SVM) can

Special Issue - 2021

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
ICRADL - 2021 Conference Proceedings

help us for predicting and analysing the student's performance.

This prediction helps to parents and teachers to keep track of student's and provide them a required counselling. It is actually trying to enhance student's acquirement and success more effectively in a way using educational data mining techniques. It can bring the benefits & influence of learner, teachers and educational institutions. Results of this examination can act as policy improvement technique in higher education.

In **[3]**, in educational data mining field, making a prediction about student academic performanceisusuallydone.Objectiveofthis paper to provide great knowledge and understandingofthedataminingtechniques. Data mining methods into considerationlike classification, clustering association rule mining and regression analysis. In almost every research paper, the only classification algorithm is taken into consideration for predicting student progress and set of activities based on the previous actions. There are so many classification techniques available for prediction but it is taking into consideration only decision tree,Naïve Bayes, Support Vector Machine (SVM), Artificial Neural Networks (ANN), K-Nearest Neighbour, SMO, Linear Regression,Random Forest, Random Tree, Tree, LADTree.

In Decision tree algorithm the maximum and minimum accuracy for expecting student's educational performance are 99.9% and 66.8% respectively. To find the maximum prediction accuracy used by the combination of student's attribute like examination, additional events, end of first sessionresult.

In Naive Bayes algorithm, the maximum and minimum accuracy for predicting student's academic data sets are 100% and 64.3% respectively. The different combination of student's attribute like Gender, Age, Marital Status, Occupation, and Job associated with the computer, Bachelor, Another master, Computer literacy.

In rule-based algorithm, the maximum and minimum accuracy was yielded by the student's academic performance are 95.7% and 55.5% respectively. To find the maximum prediction accuracy used a combination of student's attribute like family, PEP, EES, end of first session.

In K-Nearest Neighbour algorithm the maximum and minimum accuracy for predicting student's academic performance are 100% and 78% respectively.

In Artificial Neural Networks (ANN) the maximum and minimum accuracy for predicting student's academic performance are 90.8% and 67.3% respectively. To find the maximum prediction accuracy used a combination of student's attribute like term examination in their first-year course.

In **[4], o**ne of the important specifics in educational institution is the rapid growth of learning data. The main goal of institution is to progress the education worth. Prediction of student's performance in institution is one manner to extent the good quality in education system. Educational institution staff should identify students who are likely to fail in exams. But, it is hard to identify them in early stage because of large number of enrolled students.

Student's academic performance can be affected by the several factors such as personal, social, demographic data etc. Data mining in education is called educational data mining, EDM. Educational data mining describes a research field with the application of data mining, machine learning and statistics to create information from educational sets e.g., universities and tutoring systems. In addition, these methods lack the capability to reveal useful hidden information. Online exam is being launched because of need for adestination.

In [**5**], **t**he researcher has developed a prototype for performance analysis based on KPI. This exiting system used the Delphi method for analysing the student activities based on their academic performance. A key performance indicator or KPI is an important tool for management control system that obtains valuable feedback for planning and control. In the education sector, there are no set parameters of performance evaluations, it differs from institute to institute. The parameters differ at the primary level in school education depending upon the pattern of education followed by the institutions. In the junior and high level of education, the limitations change based on the type and quality of educationinstructed.

For example, the performance appraisal system followed in an engineering college will vary from those followed by a college importing medical education.

In [**6**], it intensively discusses the important factors on predicting student's performance. There are two main factors in predicting student's performances, which are attributes and prediction methods. A graphical representation for list of common characteristics and list of methods used to calculating the student's performance.First step will be focused on the important attributes used to predicting the student performance and second step will be focused on the prediction methods used in predicting students' performance. There are several algorithms under classification task that have been applied to predict student's performance. Among the algorithms used are Decision tree, Artificial Neural Networks, Naive Bayes, K-Nearest Neighbor and Support Vector Machine. In the proposed system neural networks are used educational data science. The advantages of this neural network are the ability to detect all possible interaction between the predictor variable. Neural networks are used to classify the different studentdata.

In **[7],** performing correlation analysis and using different feature selection methods, a set of factors that impact the result of students have been derived. The prediction model built using Bayes classifier displayed by highest accuracy of 84.4% and can predict unseen data. The researchers can take measures to improve students if they know that higher secondary marks (HSC), medium of instruction and contribute to a students' success rate in their particular course. Such as assignment marks and attendance and include the type of subject namely practical or theoreticalexams.

In this existing system are evaluated and applied different feature selection techniques to identify the most probable factors contributing to the success of students. When most influential variables are known by researchers

can proceed with the prediction model and buildingprocess.

In **[8],** the goal of higher education institutions is to offer good quality education to students. One way to achieve highest level of quality in higher education organization is by discovering knowledge for prediction about enrolment of students in a particular course.

Algorithm used:
- DecisionTree
- NeuralNetwork
- NaïveBayes
- K – NearestAlgorithm

The classification task is used to estimate student's performance and as there are various methodologies that are used for data classification, the decision tree method.

It helps to describe the performance of the student's in end of semester. It is used to avoid the drop outs from college and special attention to the students who need special consideration and permit trainers to provide appropriate counselling.

III. SYSTEM DESIGN

The proposed system is used to predict and motivate the students in enhanced learning. Online entry analysis is used to complement study and learning experience. To identify the weak students early and intervene and to analyse the students attribute of social economical situation, home or community factors, teaching environment features, students perception, students attitude, students ability, past academic performance these are all captured and predict the students in future performance.

The proposed system uses Naive Bayes algorithms, Support vector machine and neural networks algorithms to predict and analyse the student performance using data science. This system uses two different attributes to measure the achievements and that is student grade and student valid information.

There are different kinds of techniques available in data science to predict and analysis the students' performance in future. Education data science to find the hidden information from a huge database of education setting. It also helps the trainers, managing institute and faculties. Actually data science help in the different field of educators sectors. Data science is the process of data mining.

The proposed system track student performance on various categories such as extra circular activities, co-circular activities based on the historical performance and also their practical and theoretical performance. The algorithm can be tested for each of these subsequent steps and then they are compared.

Staff can keep track of student's activities by analysing their performance in various fields. They can predict which students lack in which field and according to that they can provide the training to students. The proposed system are uses the machine learning algorithms to analyse the student performance and system monitor their each set of activities or performance and delivered the guidelines to students or faculties in various steps to develop their skills by their own. These performance are monitored by system step by step of their activities being done.

IV. CONCLUSION

Predicting student's performance is commonly useful to help the learners to improving their learning skills. This paper has been reviewed by previous studies on predicting the students' performance with various analytical methods. Most of the researchers have used cumulative grade point average (CGPA) and internal assessment as data sets. Although for prediction techniques, the classification method is frequently used in educational data science part. Under the classification techniques, Neural Network and Decision Tree are the two methods highly used bythe researchers for predicting student's performance. It will help the educational system to monitor the students' performance in systematicway.

The performance of neuralnetworks increases with increase in data set. It will actually help to increase our education system to check the systematic performance of thestudents.

The Meta analysis on student's performance has stimulated us to transmit the further examination activities to be applied in educational institutions. Educational scheme can take the help of this model to review the students' performance in a suitable behaviours.

REFERENCES:

[1] "Student's Performance Prediction using Deep Learning and Data MiningMethods"InternationalJournalISSN: 2277-3878, Volume-8, Issue-1S4, and June 2019.

[2] "Predicting Student Performanceusingdataminingtechniques" International Journal Volume 119 No. 12 2018.

[3] "Literature Survey on Student's Performance Prediction in Education using Data Mining Techniques" International Conference November2017.

[4] "Predicting and Analysis of Student Performance Using Decision Tree Technique" International Journal volume-5, Issue 4, April2017.

[5] Arora, Amity, and Sukhbir Kaur. "Performance Assessment Model for Management Educators Based on KRA/KPI." International Conference Management March. Vol. 23.2015.10.

[6] "A Review on Predicting Student's Performance using Data Mining Techniques" International Conference2015.

[7] International JournalVolume 25– No.9, July 2011 "Identifying Key Performance Indicators and Predicting the Result from StudentData".

[8] International journal vol. 2, No. 6, 2011 "Mining Educational Data to Analyse StudentsPerformance".