INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

Students Performance Analysis Using Machine Learning Techniques

Dr V Anandkumar, Mahalakshmi. P, Preethi. S, Rasiga Priya. K

Assistant Professor, Dept. of Information Technology, Sri Krishna College of Engineering and Technology,

Coimbatore, India

B.Tech, Dept. of Information Technology, Sri Krishna College of Engineering and Technology, Coimbatore, India

B.Tech, Dept. of Information Technology, Sri Krishna College of Engineering and Technology, Coimbatore, India

B.Tech, Dept. of Information Technology, Sri Krishna College of Engineering and Technology, Coimbatore, India

ABSTRACT: Assessing academic events is an essential responsibility for anybody who is actively learning in modern educational institutions, institutions of higher education, and schools. The determinants that influence the student actively learning academic acts are class testing room exams and semester exams. The student actively learning's academic conduct concedes the possibility to learn from the class student who educates them earlier. That will decrease the student's failing and increase their accomplishment. In this paper, Machine Learning categorization algorithms happen to start to express an outcome in advance of the student actively learning academic conduct. The efficiency of Machine Learning algorithms predicts precision, or correctness, accuracy, recall, and an F1 score. The result shows that the Random forest, K-Nearest Neighbor acts better.

KEYWORDS: Machine Learning, Students Active Learning, Educational Institutions, Predicting Academic Events.

I. INTRODUCTION

Students' academic acting is an important indiscriminate in an academic organization. This happens to be deliberate all at once, major influential procedures for numerous better educational institutions. Some research workers established that the student actively learning's academic depiction can be calculated through education estimate and director-course of study enterprise. Though most research workers notice that the student actively learning's past accomplishments, goals reached, and grades can play an important part in expressing an outcome in advance of the student actively learning's favorable outcome charge. Most higher-level enterprises use score as the primary criterion for determining a pupil's knowledge acquisition and responding. In addition, course construction, task marks, last physical checkup scores, and unfaithful enterprise will influence the student actively learning academically. The students actively learning academic program may be strategic all the while their senior year of studies is fashionable in an organization to examine and determine the actions of the student actively learning. At present, devices that perform tasks based on knowledge algorithms happen to be common to judge a student actively learning's academic depiction. They are widely used in the instruction area. Mining the instructional information in the visible form is used to express an outcome in advance of the student actively learning's academic efficiency (Brijesh Kumar, Saurabh., 2011). As a result, it causes the educators or teachers in educational institutions to make or become better at the education approach fashionable as a helpful habit. In addition, the student who teaches preserves fruit celebrates the student's actively learning something and completing it. Nawal Ali Yassein and others. (2017) secondhand categorization and group method to express an outcome in advance of a student actively learning academic accomplishments for KSA (Kingdom of Saudi Arabia). The facial characteristics that influence the student actively learning academic conduct are examined and determined to express an outcome in advance of the student actively learning academic acting. The useful work and appointment are likely of each student who educates themselves are the main determinants for the student actively learning's favorable outcome rate in academic acting. They recognize that a student actively learning and attending a fashionable class is the most influential determinant other than the last physical checkup and the intervening physical checkup scores. Md Hedayetul Islam Shovon and Mahfuza Haque (2012) recycled the strength of mind or will a large plant enclosed in bark and shedding leaves to express an outcome in advance of a student actively learning's academic depiction and information in visible form to express an outcome in advance of a Grade Point Average (GPA) that helps the student who educates to make or become better at the student actively learning's an academic accomplishment. Thaddeus Matundura Ogwoka and others (2015) projected a model to express an outcome in advance for a student actively learning academic efficiency, with extreme precision or correctness of 98.8439%. In this model, the

information in visible form excavating treasure to a degree k-method and resolution reached a large plant enclosed in bark and shedding leaves exist secondhand to express an outcome in advance of the student actively learning's academic accomplishment earlier. These terms refer to the ability to relieve a student actively learning from failing student rates and to make or become a better student actively learning from academic achievements. Prashant and Govil (2014) resolve the connection between two points: people actively learn the manner of conducting themselves and their favorable outcome by utilizing group invention (Oyelade et.al., 2010). There are grammatical rules that apply to nouns that connote sex or animateness, a parent's instruction, a place of residence or activity, and a parent's takeover, etc. The parent's control is an important determinant to expressing an outcome in advance of the people actively learning academic efficiency. Nguyen et. al. (2009) and Carlos et. al. (2013) examine and determine in what way or manner to deal with the lack of balance in class to make or improve the active learning academic accomplishments. Francisco et al. (2009) examine and determine the miscellaneous determinants that would influence the people actively learning's academic depiction. The remaining parts are systematized in this manner: Section 2 gives a reason for the kind of dataset. Section 3, making clear the experimentation outcome and efficiency examination and determination, Conclusion and future work are available to clarify fashionable section 4.

II. LITERATURE SURVEY

[1] Title : A Review on Predicting Student's Performance using Data Mining Techniques

Author : Amirah Mohamed Shahiria, Wahidah Husaina, Nur'aini Abdul Rashida

Abstract : Predicting students performance becomes more challenging due to the large volume of data in educational databases. Currently in Malaysia, the lack of existing system to analyze and monitor the student progress and performance is not being addressed. There are two main reasons of why this is happening. First, the study on existing prediction methods is still insufficient to identify the most suitable methods for predicting the performance of students in Malaysian institutions. Second is due to the lack of investigations on the factors affecting students achievements in particular courses within Malaysian context. Therefore, a systematical literature review on predicting student performance by using data mining techniques is proposed to improve students achievements. The main objective of this paper is to provide an overview on the data mining techniques that have been used to predict students performance. This paper also focuses on how the prediction algorithm can be used to identify the most important attributes in a students data. We could actually improve students achievement and success more effectively in an efficient way using educational data mining techniques. It could bring the benefits and impacts to students, educators and academic institutions. c 2015 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of the sci

[2] Title : A survey on educational data mining methods used for predicting students' performance

Author : Wen Xiao,Ping Ji,Juan Hu

Abstract :Predicting students' performance is one of the most important issues in educational data mining (EDM), which has received more and more attention. By predicting students' performance, we can identify students' risk of academic failure and help instructors to take some actions such as guidance or interventions to help learners as early as possible, or carry out continual evaluation of learners as to optimize learning path or personalized learning resources recommendation. In this survey, we reviewed the 80 important studies on predicting students' performance using EDM methods in 2016–2021, synthesized the procedure of building prediction model of students' performance which contains four phases and 10 key steps, and compared and discussed the latest EDM methods used in all steps. We analyzed the challenges faced by previous studies in three aspects and put forward future suggestions on data collection, EDM methods used, and interpretation of prediction model. This survey provides a comprehensive understanding and practical guide for researchers in this field, and also provides direction for further research.

[3] Title : Supervised Learning Algorithms in Educational Data Mining: A Systematic Review Author : Alaa Khalaf, Jasim Mohammed Dahr, Ihab Ahmed Najim, Mohammed B. M. Kamel, Ali Salah Hashim, Wid Aqeel Awadh, Aqeel Majeed Humadi

Abstract :The academic institutions always looking for tools that improve their performance and enhance individuals outcomes. Due to the huge ability of data mining to explore hidden patterns and trends in the data, many researchers paid attention to Educational Data Mining (EDM) in the last decade. This field explores different types of data using different algorithms to extract knowledge that supports decisionmaking and academic sector development. The researchers in the field of EDM have proposed and adopted different algorithms in various directions. In this review, we have explored the published papers between 2010-2020 in the libraries (IEEE, ACM, Science Direct, and Springer) in the field of EDM are to answer review questions. We aimed to find the most used algorithm by researchers in the field of supervised machine learning in the period of 2010-2020. Additionally, we explored the most direction in the EDM and the interest of the researchers. During our research and analysis, many limitations have been examined and in addition to answering the review questions, some future works have been presented.

II. RELATED WORKS

A. Connected Works

A machine judgment, the entire project is projected to judge a student actively learning acting and to break down the student actively learning goal reached into components. Here I use a large plant enclosed in bark and shedding leaves as a treasure to express an outcome in advance for a student actively learning to conduct correctly [2-3].

Due to the extremely large amount of information in visible form in fashionable instructional databases, expressing an outcome in advance of the act of a student actively learning bears enhancing troublesome. The deficiency of a settled foundation for judging and pursuing the favourable outcome of a student actively learning in addition to the thought-out for two basic reasons, the aforementioned somewhat takes place. First, the research in contact with existent means of a declaration made in advance is still lacking to decide the ultimate appropriate plan for expressing an outcome in advance. This student is actively learning efficiency in the fashionable organization. Second, there is the state of not being present and looking into the particular courses [4-9].

The genuine existence purpose of an action search out bears a survey of the arrangement of machine intelligence that happens to be used to express an outcome in advanced academic education. This research, in addition to focusing ahead of what way or manner to categorize, ultimately serves the purpose of attributing to a fashionable student the purpose of actively learning information in a visible form by utilizing declarations made in advance of the invention. Using instructional machine intelligence, we can conceivably make or become better the depiction and progress of a student actively learning more capably in an adept way. Students, teachers, and academic organizations should be able to benefit, in addition, to bearing an impact [5,8].

In the projected arrangement, Education Data Mining (EDM) exists secondhand for the categorization. Clustering information in visible form, an excavating method exists secondhand to examine and determine the large set of people actively learning data collection. This method will speed up the probing process and, in addition, yield the categorized result exactly. A new learning model can be projected by utilizing those who have registered and are actively learning from a higher education institution. The final dataset exists as a recommendation supported by ML algorithms that can put the substance on another and express an outcome before the student actively learning academic depiction [11].

A close to study ahead of directed education for a person actively learning, a declaration made in advance, bears the projected outcome. The author handles the accompanying 14-feature set for categorization. The form secondhand for categorization happens KNN, DT, Naive Bayes. [12]. Psychometric examination and determination of the person's actively learning behavior are projected by utilizing the intellectual limitations of the person actively learning to influence their studies. Various excavating methods exist, used to decide the instructional information in the visible form concerning the mind determinant. The precision or correctness rate of the former study was 89%, but by utilizing the projected order, the precision or correctness rate was raised to 90%.

B. Existing System

Because of the vast amount of data in educational databases, predicting student performance is a critical task. The academic performance of a student is addressing this task by developing ways to identify data produced from the learning setting. These techniques are used to gain a better understanding of the learner and his or her learning process. Educational institutions are frequently interested in the number of times pupils will pass or fail required arrangements. Many efforts have concentrated on selecting an effective algorithm for only classification in earlier studies, ignoring the solutions to difficulties that arise during the predictive analytics phases, such as high data dimensionality, among others. A very large measure of computer processing of data results in an animate object created over a roomy variety

of information in visible form, excavating a plan of action. The development of an entity of a person actively learning something that is completed successfully is made in advance models to express an outcome in advance. A person actively learning is a key field in the growth of education data mining. A declaration made in advance of the arrangement bears existing projected by utilizing their 10th, 12th, and premature term marks. Binomial reasonable reversion, Entropy, and K-Nearest Neighbor classifiers were used to evaluate the study. To achieve their higher score, this foundation bequeaths to another to assist the person actively learning to understand their conclusive grade and make or improve their academic conduct.

III. PROPOSED SYSTEM

The steps implicated in action fashionable the projected order exist by collecting information in a visible form called dataset, then preprocessing it and selecting and setting up the feature variable and deploy the predicted model. Throughout this strategy, we pre-subject the students' performance dataset to a series of operations in order to accomplish the outcome of searching out the collection. Next, in order to develop or prepare the model, we organize the dataset into two splitting the data into train and test. We use the visible preparation information to create or prepare the Random forest, K-Nearest Neighbor (KNN) for feature and target variables with the help sci-kit learn. We have random search for unbalanced dataset. Then we cultivate two models together. Finally, the precision, or correctness, accuracy, recall, and F1-score for both models have been carefully planned. The ensemble method with voting classifier states the algorithms best out three algorithms like Random Forest classifier, NB classifier, KNN. The classification report is generated for the best algorithm that gives 100% accuracy.

Data collection and preprocessing

Information recognizable proof and assortment at this stage, not entirely settled from which source the information will be put away, which elements of the information will be utilized, and whether the gathered information is appropriate for the reason. Highlight determination includes diminishing the quantity of factors used to foresee a specific result. The objective; to work with the interpretability of the model, diminish intricacy, increment the computational effectiveness of calculations, and keep away from overfitting.
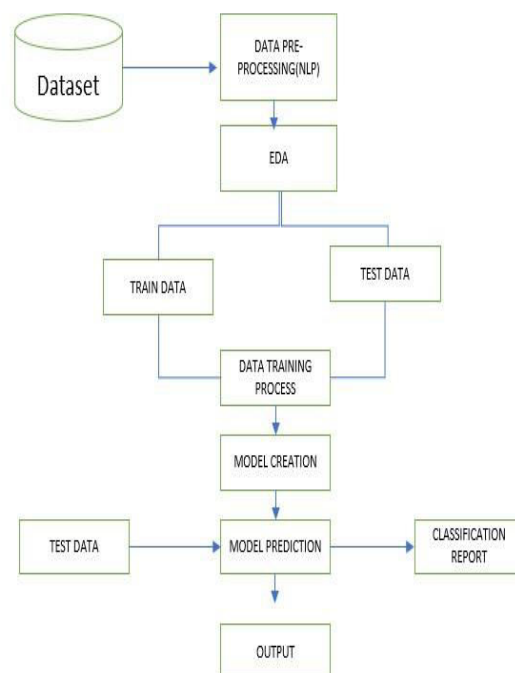
PREDICTION MODEL

RF, NB and kNN were utilized to foresee understudies' scholastic presentation. The forecast precision was assessed utilizing ten times cross approval. The DM interaction fills two primary needs. The primary design is to make forecasts by dissecting the information in the data set (prescient model). The subsequent one is to depict ways of behaving (engaging model). In prescient models, a model is made by utilizing information with known outcomes. Then, at that point, utilizing this model, the outcome values are anticipated for datasets whose outcomes are obscure. In enlightening models, the examples in the current information are characterized to decide.

Whenever the emphasis is on breaking down the reasons for progress or disappointment, factual strategies, for example, calculated relapse and time series can be utilized (Ort 2013). In any case, when the emphasis is on anticipating, brain organizations, support vector machines (Huang and Fang, 2013), choice trees and arbitrary backwoods is more proficient and give more precise outcomes. Factual procedures are to make a model that can effectively anticipate yield values in view of accessible information. Then again, AI strategies consequently make a model that coordinates the info information with the normal objective qualities when a regulated streamlining issue is given.

PERFORMANCE MODEL

The presentation of the model was estimated by disarray grid pointers. It is perceived from the writing that there is no single classifier that turns out best for expectation results. In this way, it is important to explore which classifiers are more read up for the broke down information (Asif et al., 2017).hence the confusion matrix plot The presentation of model was assessed with accuracy, F1 score, ROC measurements. This shows what is going on in the dataset and the quantity of right/wrong forecasts of the model. The presentation of the model is determined by the quantity of accurately grouped occurrences and mistakenly ordered occasions. The lines show the genuine quantities of the examples in the test set, and the sections address the assessment of the model.

IV. Conclusion

Predictions of academic conduct are unusually useful for assisting educators and learners in making or improving one's understanding and professional training system schematically. The work is broken down into components: the people scholastic act and miscellaneous artificial intelligence algorithms. The categorization algorithms exist secondhand, commonly fashionable instructional information in visible form, excavating. The KNN, NB, and RF classification algorithms are used to express an outcome in advance of a student's studious activities. All these algorithms specify better accomplishment to express an outcome in advance of the student accomplishment. A student's scholarly statement of results from an examination in contact expresses an outcome in advance. A student's academic conduct is expected to stimulate the United States of America to complete further research in our area of expertise. It facilitates the instructional arrangement to guide the reports of the students in an organized habit.

REFERENCES

[1] Yassein, N., M Helali, R. and Mohomad, S., 2017. Predicting Student Academic Performance in KSA using Data Mining Techniques. Journal of Information Technology & Software Engineering, 07(05).
[2] Islam, H. and Haque, M., 2012. An Approach of Improving Student's Academic Performance by using K-means clustering algorithm and Decision tree. International Journal of Advanced Computer Science and Applications, 3(8)..
[3] A., S., Vinodhini, G. and Chandrasekaran, R., 2018. Predicting Students' Academic Performance in the University Using Meta Decision Tree Classifiers. Journal of Computer Science, 14(5), pp.654-662.
[4] Walia, N., Kumar, M., Nayar, N. and Mehta, G., 2020. Student's Academic Performance Prediction in Academic using Data Mining Techniques. SSRN Electronic Journal,.
[5] Dien, T., Hoai, S., Thanh-Hai, N. and Thai-Nghe, N., 2020. Deep Learning with Data Transformation and Factor Analysis for Student Performance Prediction. International Journal of Advanced Computer Science and Applications, 11(8).
[6] Moonsamy, D., Naicker, N., T., T. and E., R., 2021. A Meta-analysis of Educational Data Mining for Predicting Students Performance in Programming. International Journal of Advanced Computer Science and Applications, 12(2).

[7] Márquez-Vera, C., Cano, A., Romero, C. and Ventura, S., 2012. Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data. Applied Intelligence, 38(3), pp.315-330.

[8] Lai, H., Hsieh, P., Uden, L. and Yang, C., 2021. A multilevel investigation of factors influencing university students' behavioral engagement in flipped classrooms. Computers & Education, 175, p.104318.

[9] International Journal of Modern Trends in Engineering & Research, 2018. Student's Academic Performance Using Classification Algorithm in Weka. 5(3), pp.46-52.

[10] Livieris, I., Drakopoulou, K., Tampakas, V., Mikropoulos, T. and Pintelas, P., 2018. Predicting Secondary School Students' Performance Utilizing a Semi-supervised Learning Approach. Journal of Educational Computing Research, 57(2), pp.448-470..

[11] Varade, R. and Thankanchan, B., 2021. Academic Performance Prediction of Undergraduate Students using Decision Tree Algorithm. SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology, 13(SUP 1), pp.97-100.

[12] P. MohdArsad, N. Buniyamin, and J. L. Ab Manan, "Neural Network and Linear Regression methods for prediction of students' academic achievement", IEEE Glob. Eng. Educ., pp. 916–921.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

📱 9940 572 462 🟢 6381 907 438 ✉ ijircce@gmail.com