SURVEY BASED ON STUDENTS PERFORMANCE PREDICTION USING VARIOUS ONLINE TECHNIQUES

[1]Aparna Baiju, [2]Athira Babu M, [3]Bincy Baby, [4]Sisira P Sasi, [5]Neha Beegam P E
[1 2 3 4] Student, [5] Assistant professor
[1 2 3 4 5] Computer Science and Engineering, APJ Abdul Kalam Technological University
[1 2 3 4 5] Ilahia College Of Engineering And Technology, Muvattupuzha, Kerala, India

Abstract: Predicting academic performance is an important task for students in university, college, and school, etc. Class quizzes, assignments, lab examinations, midterm, and final exams are all aspects that influence a student's academic achievement. The student's academic performance should be communicated to the class teacher ahead of time to reduce dropout rates and improve performance. Educators face a variety of obstacles, including classifying students and predicting their exam success. Various traditional data mining methods such as decision trees and association rules were used to perform classification. Via learning action, it is important to examine and comprehend the progression and advances in the standard of online education.

Index Terms **- Educational Data mining, Conventional neural network, recurrent neural network, Dropout, Peer learning, Prediction modeling,**

I. INTRODUCTION

Student's instruction overall performance is an essential part of an educational organization. This is taken into consideration as one of the crucial measures for plenty of advanced universities. Some researchers said that the student's instructional overall performance may be measured thus studying evaluation and co-curriculum sports. Though, the bulk of researchers have cited that the students beyond performances, achievements, and grades can play a crucial function to are expecting the student's fulfillment rate. Predominantly, the maximum of the better-stage establishments uses grade as the primary degree to evaluate student's overall performance. Also, direction structure, task marks, very last examination scores, and extracurricular sports will affect the student's instructional overall performance. The student's instructional software may be properly through their sophomore duration of the research in an organization to research the overall performance of college students. At present, device studying algorithms are maximum famous to assess student's instructional overall performance that has been drastically carried out within side the schooling sector. Extracurricular sports as drop out prediction issue in better schooling the use of the choice tree. The gadget will assist the scholars to make suitable desires and a higher future. Student know-how extraction is primarily based totally on their hobby and educational overall performance the version has especially portions. In the First portion, gather the facts from college students who may be finished their secondary schooling. Collecting facts from college students they may be doing their better schooling. By these facts set we're growing an education version and offers prediction.

II. RELATED WORKS

The machine will assist the scholars to make appropriate desires and a higher future. Student expertise extraction is primarily based totally on their hobby and educational overall performance the version has in particular portions. In the First portion, accumulate the facts from college students they're finished their secondary education. Collecting facts from college students they're doing their better education. By this fact set, we're growing a schooling version and offers prediction.

Tomas Hasbun, Alexandra Araya, and Jorge Educational Data [1] *"Extracurricular activities as dropout prediction factor in higher education using decision tree"* Mining can help students who are at risk of dropping out

- The software used was the r studio package part for the cart algorithm.
- Better prediction estimation in dropout students than last year.

According to RianneConijn, Chris Snijders, Ad Klein geld, Use Matzoth [2*] "predicting student performance from LMS data* "With the adoption of Learning Management Systems (LMSs) in academic establishments, plenty of facts has to turn out to be had describing college students' online behavior.

- Use LMS data for predictive modelling
- Use predictive analytical techniques

According to Zhang Hongyan, Jiao Baocong, Jing Li, LuoZhuobi [3] *"Research on using grey system theory to predict the development scale of education"* Based on the facts of the dimensions of schooling in Xinjiang, this paper makes use of the gray machine concept to set up GM (1, 1).

- Use Grey system theory and grey prediction model.
- Result of prediction of the scale of education.

According to GulcheraShadmanova [4] *"Upgrading higher education through open online courses"* The article discusses the significance of introducing open online publications into better schooling to make it greater powerful and accessible.

- The ability to study at multiple top universities at the same time.

- Make informed choices from a wide range of real knowledge opportunities.

According to SayibdjanMirzaev [5] *"Predicting learning behavior online course learners using hybrid deep learning model"* As tendencies in instructional generation hold to advance, the strategies wherein the guides are introduced and obtained with the aid of using freshmen' developed from board coaching to online guides.

- Study the feasibility of implementing a hybrid deep learning model for predicting Online Course Learners' behavior.
- Use Grey system theory and grey prediction model.

Result of prediction of the scale of education

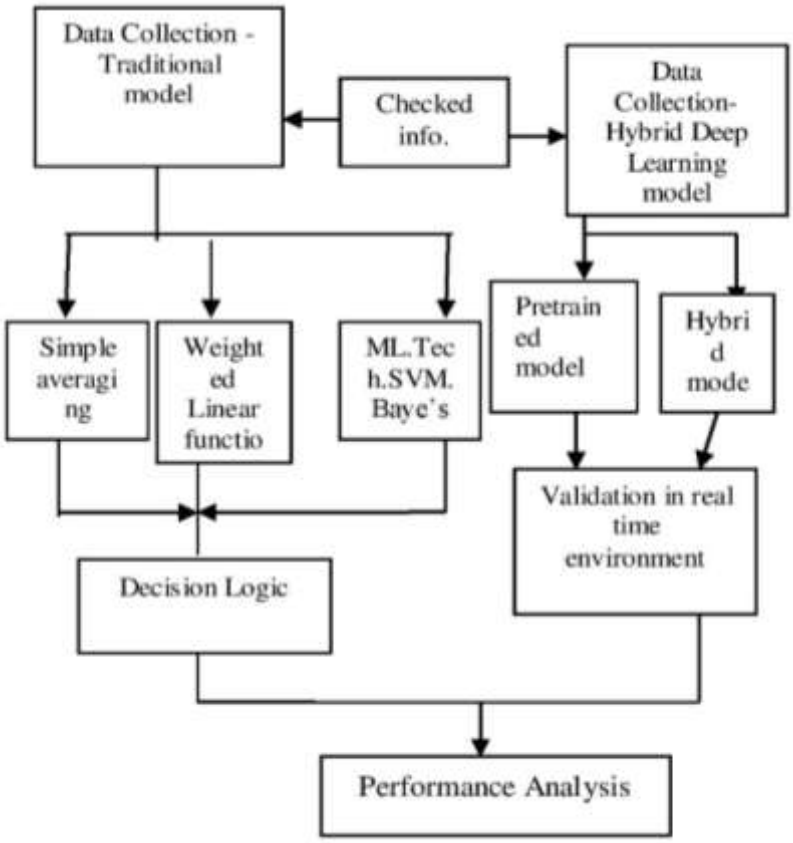

Figure 1: Workflow diagram

According to MinNie1, ZhaohuiXiong 1[6*]"Carrier choice prediction based on campus big data-mining the potential behavior of college students* "Career desire has a pivotal position in university Students existence planning thousand college students.

- To predict students' career choices, use Cluster Centres based on the XGBOOST (ACCBOX) model.
- Career choices are based on their professional skills.

According to Engr.Sana Bhutto, Dr. Isma Farah Siddiqui [7] *"Predicting students' academic performance through supervised machine learning"* There are many supervised and unsupervised forms of gadget studying tactics which can be used to extract hidden statistics and courting among statistics if you want to subsequently, enables decision-makers withinside the destiny to take right interventions.

- Classification to predict the student's future performance, which can be used to categorize students based on their academic records.

According to Rajesh Kumar T, Vamsidhar T, Harika B, Madan Kumar T, NissyRthe [8] *"Student performance prediction using data mining techniques"* fulfillment of an educational organization may be measured in phrases of exceptional of schooling gives to its college students.

- Use ID3 algorithm is based on the recursive process.
- System to predict the fourth-year result of third-year students based on their current and previous Performance.

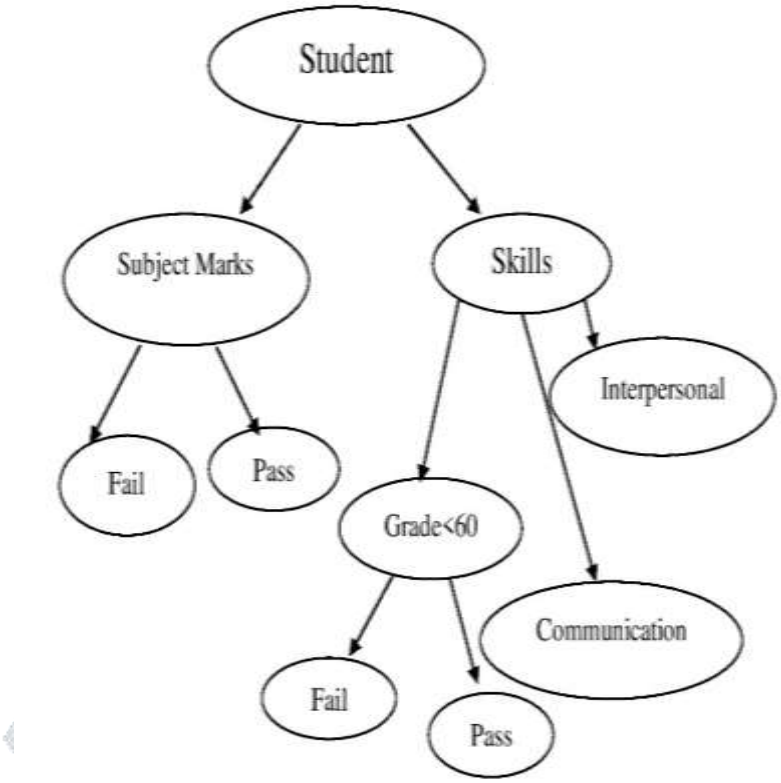

Figure 2: Decision Tree for a scholar's overall performance.

According to Juan L. Rastrollo- Guerrero, Juan A. Gómez- Pulido [9] *"Analyzing and predicting student's performance using machine learning: a review* "Predicting student's performance is one of the foremost important topics for learning contexts like schools and universities since it helps to style effective mechanisms that improve academic results and avoid dropout, among other things.

- By analyzing and processing all data carefully can give us useful information about students.
- The authors used the SVM Algorithm the most, and it produced the most accurate predictions.

According to SonglakSakulwichitsintu [10] *"A peer learning framework for enhancing Students learning experiences in online environments"* online peer learning is increasingly important for education instructions.

- Students' level of accountability when working with others.
- Student's communicative competencies in sharing learning experiences.

"*Improved data mining approaches can be used to improve student performance prediction in educational systems,"* according to Karthik Ryan [11]. The goal of this study was to enhance the prediction of student performance based on personal and academic variables.

- Improving student performance prediction based on their personal and academic performance characteristics.
- Make a study to improve student performance based on data mining make a study on to improve students.

According to Wil.ton.W.T.Fok', Yeung', K.Y. [12] *"Prediction model for Students' Future Development by Deep Learning and Tensor Flow Artificial Intelligence Engine,"* it also features a python-based framework like Tensor Flow.

- Experimenting with Tensor Flow artificial intelligence
- Engine for classifying students' performance and forecasting their future university degree program is studied.
- The prediction result is good enough to provide appropriate recommendations to students.

III. PROPOSED SYSTEM

The system focuses on two sections or categories. Since the first section aims at students who all are completed their plus two education, providing suggestions for higher studies according to their interest and marks. Apart from this, the second section dealing with the students who all are doing their higher studies. Here the datasets are the internal marks obtained from their academics. The collected data have to be pre-processed for further processing. Thus, analyzing their marks in each semester and predict their weaker side. Also, suggest whether they have to concentrate more on the practical side of the theoretical side. It helps the students to become the best Professionals.

The following are the important modules in our proposed work.

- Data Pre-processing
- Learning model
- Model creation
- Analysis and Testing
- Prediction

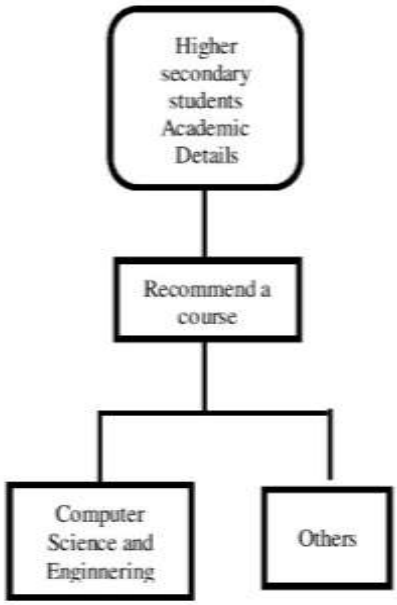

Figure 3: Proposed system Architecture section 1

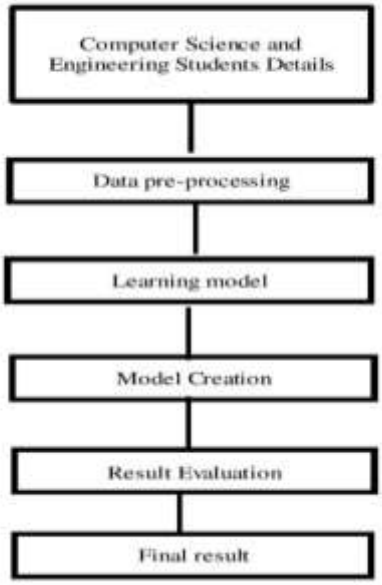

Figure 4: Proposed system Architecture section 2

Data Pre-processing

Here we are incorporating a data pre-processing module to clean and organize the raw data to make it suitable for building and training Machine Learning models. Data pre-processing is the procedure for preparing raw data for use in a machine learning model. It is the first and most important step in the development of a machine learning model. When creating a machine learning project, it's not always a case that we encounter clean and formatted data. And while doing any operation with data, it's mandatory to wash it and put it in a formatted way. So, for this, we use data pre-processing tasks. Real-world data is often incomplete, unreliable, deficient in specific behaviors or trends, and rife with errors. The data is gathered from the students it is collected in raw format which is not feasible for the analysis. So, data pre-processing is a required task for cleaning the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model.

Learning model

Modeling in machine learning is an iterative phase where a data scientist continually trains and tests machine learning models to discover the best one for the given task. During this stage, we are train numerous models to define which one of them provides the most accurate predictions. After pre-processed the collected data and split it into three subsets we can proceed with model training. This technique entails feeding the algorithm with training data. The most prevalent model training approaches are supervised and unsupervised learning.

Model Creation

Modeling in machine learning is an unvaried section wherever regularly train and check machine learning models to get the most effective one for the given task. Typically, Model Creation is simplified by leveraging existing models that are available from a variety of online sources. Most ML/DL frameworks such as PyTorch or Tensor Flow, have pre-trained models that are leveraged to speed up the model implementation stage.

Validation / Result Evaluation

The data set is tested and analyzed using algorithms. Data that has been explicitly defined for use in experiments, usually of a computer programmer, is known as test data. Test data could also be recorded for re-use or used once then forgotten. Test data can be created manually, by using data generation tools or it can be retrieved from an existing production environment. During the prediction process, all algorithms are compared for accuracy. The most important attributes in the dataset are selected and using

that attribute result obtained. At this stage, we should have a trained model and are ready to conduct evaluation techniques on its performance. For evaluation, we utilize a partition of the refined data, usually mentioned because of the 'test data'. During the training of the model, the test data was not seen.

Prediction

The performance of a rule after it has been trained on a historical information set and applied to new data when predicting the probability of a selected outcome is referred to as prediction. The rule can generate probable values for AN unknown variable for every record within the new knowledge, permitting the model builder to spot what that worth can be. By mistreatment of the current knowledge, we will predict a district for college students, and supported by the prediction, we will direct the coed throughout their studies.

IV. CONCLUSION

Predicting academic performance is an important task for the students in university, college, and school, etc. Class quizzes, assignments, lab examinations, midterms, and final exams are all aspects that influence a student's academic achievement. The student's academic performance should be communicated to the class teacher ahead of time in order to reduce student dropout and improve performance. Educators have many obstacles, including classifying students and predicting their exam success. Machine learning algorithms are currently the most popular method of assessing a student's academic achievement, and they have been widely used in the education sector. Using a decision tree, extracurricular activities were found to be a dropout predictor in higher education. The method will assist students in making informed decisions that will lead to a more prosperous future. Student knowledge extraction based on their interest and academic performance the model has mainly two portions. In the First portion, collect the data from students who have completed their secondary education. Collecting data from students who are doing their higher education. By this data set, we are creating a training model and give a prediction.

V. ACKNOWLEDGMENT

Apart from the efforts, the success of this project's preliminary report depends largely on the encouragement and guidelines of many others. We would like to take this opportunity to thank everyone who contributed to the successful completion of this project. We would like to show our heartfelt gratitude towards Prof. Dr. ABDUL GAFUR M, Principal, Ilahia College of Engineering and Technology for permitting us to work on this project. Also, we would like to show our greatest gratitude towards our head of the Department of Computer Science & Engineering Dr. Arun E and project guide Ms. Neha Beegam P E for their valuable advice and guidance. Finally, we express our gratitude and thanks to all our teachers and other faculty members of the Department of Computer Science & Engineering, for their sincere and friendly cooperation in completing this project.

.

REFERENCES

[1] Engr. Sana Bhutto, Dr. Isma Farah Siddiqui, Jamshoro, Pakistan, Dr. Qasim Ali Arain, Dr. Isma Farah Siddiqui (2020), Predicting Students' Academic Performance Through Supervised Machine Learning, Department of Software Engineering Mehran University of Engineering and Technology.

[2] Juan L. Rastrollo- guerrero, Juan A. Gómez- Pulido * And Arturo Durán- domínguezEscuelaPolítécnica, (2020) Universidad De Extremadura, Analyzing and Predicting Students Performance by Means of Machine Learning: A Review 10003 Cáceres, Spain; Juanluisrg@unex.Es (J.L.R.-G.); Arduran@unex.Es (A.D.A Peer Learning Framework for Enhancing Students' Learning Experiences in Online Environments.

[3] GulcheraShadmanova (2019) Upgrading Higher Education through Open Online Courses Department of Information Technologies, Tashkent Institute of Irrigation and Agricultural Mechanization Engineers Tashkent, Uzbekistan, 2019 International Conference on Information Science and Communications Technologies

[4] RianneConijn, Chris Snijders, Ad Kleingeld, UweMatzot (2019), Predicting Student Performance from LMS Data: A Comparison of 17 Blended Courses Using Moodle LMS, IEEE Transactions on Learning Technologies (Volume: 10, Issue: 1, Jan.-March 1 2019).

[5] Wil.ton. W.T. Fok', Y.S He', H.H.AuYeung', K.Y.Law', KH Cheung', YY.Ai', P. Ho (2018), Prediction Model for Students' Future Development by Deep Learning, Tensor the University of HongKong.4th International Conference on Information Management (ICIM).

[6] SayibdjanMirzaev, S. Mohanavalli; B. Bharathi (2018), Predicting Learning Behavior Online Course Learners Using Hybrid Deep Learning Model, Department of Information Technologies, Department, IEEE 6th International Conference on MOOCs, Innovation, and Technology in Education (MITE)

[7] SonglakSakulwichitsintuPaul Turner (2018), A Peer Learning Framework for Enhancing Students' Learning Experiences in Online Environments, School of Engineering and ICT University of Tasmania Hobart, Australia, volume: 2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)

[8] Rajesh Kumar T, Vamsidhar T, Harika B, Madan Kumar T, Nissy R (2017), Student Performance Prediction Using Data Mining Techniques, International Journal of Advanced Research in Computer and Communication Engineering ISO 3297:2007 Certified Vol. 6, Issue 3, March 2017department of computer engineering KoneruLakshmaiah education Foundations, 2018 Third International Conference on Informatics and Computing (ICIC0).

[9] Sri Krishna, Dr. K. Karthikeyan, P. Kavipriya Assistant Professor, On Improving Student Performance Prediction in Education Systems Using Enhanced Data Mining Techniques, Volume 7, Issue 5, May 2017, ISSN: 2277 128X International Journal of Advanced Research in Computer Science and Software Engineering

[10] MinNie1, ZhaohuiXiong 1, RuiyangZhong 2, Wei Deng 3,4, * And Guowu Yang (2016), Carrier Choice Prediction Based on Campus Big data-mining The Potential Behaviors of College Students,2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT).

[11] TomasHasbun, Alexandra Araya, And Jorge Villalon1(2016), Extra-Curricular Activities as Drop Out Prediction Factor in Higher Education Using Decision Tree, School of Engineering and Science Adolfo Ibáñez University Santiago, Chile.

[12]Zhang Hongyan, Jiao Baocong, Jing Li, LuoZhuobi (2007), Research on Using Grey System Theory to Predict, The Development Scale of Education,2007 IEEE International Conference on Grey Systems and Intelligent Services.