PREDICTIVE MODEL TO REDUCE STUDENT DROPOUTS IN M.TECH USING DECISION TREE

Project Report

Submitted in partial fulfillment of the requirements for the award of degree of

Bachelor of Technology

In

COMPUTER SCIENCE AND ENGINEERING

By

M. AKHILESWARI	(17K61A0565)
S. NAGARAJU	(17K61A0596)
K. SAI BHARGAVI	(17K61A0559)
Y. NAVYA	(17K61A05B5)

Under the esteemed guidance of
Mrs. A.N.V.K.Swarupa, M.Tech
Assistant Professor , CSE.

Department of Computer Science and Engineering
SASI INSTITUTE OF TECHNOLOGY &ENGINEERING
(Approved by AICTE, New Delhi, Permanently Affiliated to JNTUK, Kakinada and SBTET-Hyderabad, Accredited by NAAC with 'A' Grade and NBA, Ranked as "A" Grade by Govt. of A.P., Recognised by UGC 2(f) & 12(B))
Kadakatla, TADEPALLIGUDEM– 534 101

ACADEMIC YEAR 2020-2021

VISION AND MISSION OF INSTITUTE

VISION

Confect as a premier institute for professional education by creating technocrats who can address the society's needs through inventions and innovations.

MISSION

1. Partake in the national growth of technological, industrial, industrial area with Social responsibilities.
2. Provide an environment that promotes productive research.
3. Meet stakeholder's expectations through continued and sustained quality improvements.

VISION AND MISSION OF DEPARTMENT

VISION

Build a learning environment that enhances creativity, social awareness and leadership skills for total personality development.

MISSION

The Computer Science and Engineering department's consistent effort is to provide the learner with an exposure to emerging technology by providing hands-on experience making them creative and research oriented professionals with values, Leadership qualities and zeal to serve the society.

PROGRAM OUTCOMES (POs)

Students in the Computer Science and Engineering program should, at the time of their graduation be in possession of:

PO1.Engineering Knowledge: Apply knowledge of mathematics, science, engineering fundamentals and an engineering specialization to the solution of complex engineering problems.

PO2.Problem Analysis: Identify, formulate, research literature and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences and engineering sciences.

PO3.Design/ Development of Solutions: Design solutions for complex engineering problems and design system components or processes that meet specified needs with appropriate consideration for public health and safety, cultural, societal and environmental considerations.

PO4. Conduct investigations of complex problems using research based knowledge and research methods including design of experiments, analysis and interpretation of data and synthesis of information to provide valid conclusions.

PO5.Modern Tool Usage: Create, select and apply appropriate techniques, resources and modern engineering and IT tools including prediction and modelling to complex engineering activities with an under- standing of the limitations.

PO6.The Engineer and Society: Apply reasoning informed by contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to professional engineering practice.

PO7.Environment and Sustainability: Understand the impact of professional engineering solutions in societal and environmental contexts and demonstrate knowledge of and need for sustainable development.

PO8.Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of engineering practice.

PO9. Individual and Team Work: Function effectively as an individual, and as a member or leader in diverse teams and in multidisciplinary settings.

PO10.Communication: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as being able to comprehend and write effective reports and design documentation, make effective presentations and give and receive clear instructions.

PO12.Project Management and Finance: Demonstrate knowledge and understanding of engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

PROGRAM SPECIFIC OUTCOMES (PSOs)

PSO1.Mobile & Web Application Development: Ability to develop mobile & web applications using J2EE, Android and J2ME.

PSO2.Cloud Services: To deploy virtualized and cloud based services in the organization.

PROGRAM EDUCATIONAL OBJECTIVES (PEOs)

PEO1: Graduates will be able to analyze, design and develop advanced computer application to provide solution to the complex problems.

PEO2: Graduates are well trained, confident, research oriented and industry ready professionals who are intellectual, ethical and socially committed.

PEO3: Graduates will have the technical, communication skills and character that will prepare them for technical and leadership roles.

COURSE OUTCOMES (COs)

CO1. Develop problem formation and design skills for engineering and real world problems.

CO2. Collect and Generate ideas through literature survey on current research areas which

help to analyze and present to impart knowledge in different fields.

CO3. Impart knowledge on software & hardware to meet industry perspective needs and

standards.

CO4. Create interest to carry out research on innovative ideas as a lifelong learning.

CO5. Ability to work with team, and enrich presentation and communication skills.

CO6. Create a platform that makes students employable.

EXPECTED OUTCOMES

PROGRAM OUTCOMES (POs)

PO1: Engineering Knowledge

PO2: Problem Analysis

PO3: Design/Development of Solutions

PO4: Conduct investigation of complex problems

PO5: Modern Tool Usage

PO6: The Engineer and Society

PO7: Environment and Sustainability

PO8: Ethics

PO9: Individual Team Work

PO10: Communication

PO12: Project Management and Finance

PROGRAM SPECIFIC OUTCOME (PSOs)

PSO1: **Mobile & Web Application Development**

SASI INSTITUTE OF TECHNOLOGY & ENGINEERING

(Approved by AICTE, New Delhi, Affiliated to JNTUK, Kakinada and SBTET-HyderabadAccredited by NAAC with 'A' Grade and NBA)
Kadakatla, TADEPALLIGUDEM – 534 101, WG-Dst, A P, Ph: 08818 244989

Department of Computer Science and Engineering

Accredited by NAAC with "A" Grade,
Recognised by UGC under section 2(f) &12(B)
Approved by AICTE - New Delhi
Permanently Affiliated to JNTUK, SBTET,
Ranked as "A" Grade by Govt. of A.P.,

CERTIFICATE

*This is to certify that the project work entitled "**Predictive Model to Reduce Student Dropouts in M.Tech using Decision Tree.**" is being submitted by **M.Akhileswari (17K61A0565), S.Nagaraju (17K61A0596), K. Sai Bhargavi (17K61A0559), Y. Navya (17K61A05B5)** in partial fulfilment for the award of the degree of **BACHELOR OF TECHNOLOGY**, in **Computer Science and Engineering** to Jawaharlal Nehru Technological University, Kakinada during the academic year 2020 to 2021 is a record of bonafide work carried out by them under my/our guidance and supervision. The results presented in this thesis have been verified and are found to be satisfactory. The results embodied in this thesis have not been submitted to any other University or Institute for the award of any other degree or diploma.*

Project Supervisor	**Head of Department**
Mrs. A.N.V.K.Swarupa	**Dr. M.V.S.S. Nagendranath**
Assistant Professor	**Associate Professor & HOD**
Depart of CSE	**Depart of CSE**

External Examiner

sasi INSTITUTE OF TECHNOLOGY & ENGINEERING
autonomous
Accredited by NAAC with "A" Grade,
Recognised by UGC under section 2(f) &12(B)
Approved by AICTE - New Delhi
Permanently Affiliated to JNTUK, SBTET,
Ranked as "A" Grade by Govt. of A.P.,

DECLARATION BY THE CANDIDATES

We **M.Akhileswari (17K61A0565), S.Nagaraju (17K61A0596), K. Sai Bhargavi (17K61A0559), Y. Navya (15K61A0588),** hereby declare the project report entitled "**Predictive Model to Reduce Student Dropouts in M.Tech by using Decision Tree.**" under esteemed supervision of **Mrs.A.N.V.K.Swarupa**, is submitted in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology in Computer Science & Engineering. This is a record of work carried out by us and the results embodied in this project have not been reproduced or copied from any source. The results embodied in this project report have not been submitted to any other University or Institute for the award of any other degree or diploma.

M. Akhileswari (17K61A0565)

S. Nagaraju (17K61A0596)

K. Sai Bhargavi (17K61A0559)

Y. Navya (17K61A05B5)

ABSTRACT

Currently, there is a rapid growth and improvement within the educational zone. But there is an increase in academic dropouts in M.Tech due to various reasons like ambiguity in domain selection, lack of student's knowledge in selected domains , lack of guidance, lack of motivation , Income. So, to overcome these problems we are developing a predictive model to predict the rate of academic dropouts based on the domain they are interested in and the income of their parents and the marks that they gained in the B. Tech.The prediction of dropout rate helps the student to take care about his/her career by declining their concern of dropout in M.Tech. The prediction of dropout rate can be done by using data mining technique Decision tree.

ACKNOWLEDGEMENT

We pay obeisance to the dynamic Chairman of Sasi Educational Society, **Mr. Sri Burugupalli Venu Gopala Krishna**, for his inspiring presence, which has always been the principle driving force behind all over endeavors.

First of all, we would like to thank **Mr. M. Narendra Krishna**, Vice-Chairman, Sasi Institute of Technology & Engineering, for his everlasting support.

We would like to thank **Dr. K. Bhanu Prasad**, Director, Sasi Institute of Technology & Engineering, for his support.

We would like to thank **Dr. J. Annand Chadu Lal**, Principal, Sasi Institute of Technology & Engineering, for his support.

We would like to thank **Dr. M.V.S.S. Nagendranath**, Associate Professor and Head of the Department, Computer Science and Engineering, for providing us with invaluable feedback on our work, which allowed us to constantly improve.

We are grateful to our project coordinators **Dr. M. Kandan** and thanks to all teaching and non teaching staff members those who contributed for the successful completion of our project work.

We would like to thank our Guide **Mrs. A.N.V.K. Swarupa**, Asst Professor, Sasi Institute of Technology & Engineering for her constant encouragement, monitoring and guidance throughout the submission of the project. She motivated us whenever we encountered an obstacle along the way.

With gratitude,

M. Akhileswari(17K61A0565)

S. Nagaraju (17K61A0596)

K.Sai Bhargavi (17K61A0559)

Y. Navya (17K61A05B5)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER-1

INTRODUCTION

PREAMBLE

Data mining is a technique where we predict the information from the large data sets that are used for the user. It is a process of identifying the different patterns and parameters to predict the outcomes. The process of identification of useful data that brings a data-driven decision from huge data sets. It depends on effective data collection and warehousing and computer processing.

Data mining models used to identify the academic and non-academic data uses two classification techniques that acquire the quality of data for the classification task, which can be useful for predicting the loss of academic status due to low academic performance. In data mining, we extract the information that is useful for the user, but when we discuss the data analytics by using that information we make a conclusion for the information, data analytics is used to increase efficiency which is used to predict the conclusions from the massive information.

Data analytics is also using quantitative and qualitative approaches for the hidden data which is more useful to the user. In data analytics, we extract the data and clean and analyze the various behavioural patterns. Data analytics can play a key role in improving the business as it is used to gather hidden sights, generate reports, perform market analysis and improve business requirements. Nowadays all organizations are collecting data that is useful to their customers, markets, and business processes.

Classification is a data mining activity that assigns items to a group to identify categories or classes. The goal of Classification is to accurately predict the casting phase of each case in the data. Decision Tree Mining is a type of data mining technique used to build classification models. Create classification models in the form of a tree, such as its name. Decision tree analysis is a predictive modelling tool that can be used across many areas. In the Decision tree, there are two main entities: Decision nodes, leaves. Decision node means where data is split whereas leaves are where we get the outcome.

1.1 PROBLEM STATEMENT

If a student is in his final year of the Bachelor Degree, if he wants to do M.Tech after his B.Tech then it is very difficult for him to select which domain he needs to choose. He is in a dilemma to choose the correct domain. If he selects a particular domain, after joining that domain if his knowledge is very less in that domain then there might be a chance to drop out from that course. So the student's time and money are wasted.

From the previous research, we can observe that if we do prediction after a student joins the course then we can't solve the entire problem. Only we can save some people from dropouts. So here we need an early prediction system that predicts a student's interest before he chooses the course. Then definitely he can't drop out due to lack of knowledge in choosing a domain. Then we can save a cent percent of people from the dropout due to this reason.

1.2 OBJECTIVE OF THE PROJECT

Here nowadays no of students dropout from higher education (M.Tech) are increasing rapidly. The students drop out due to various reasons such as ambiguity in selecting a domain, lack of knowledge in the selected domain, and the lack of facilities in the selected college. Our agenda is to sort out the above problems and provide the early prediction before joining the course.

The proposed system aims to provide a better prediction of higher education (M.Tech admission) through data mining. Here the system provides the prediction of the chance of student dropout in that particular domain. Here we can take the parameters from the student and then we give those parameters to the trained decision tree algorithm and then we can provide the prediction to the student. Then there might be a chance for the student to dropout.

1.3 METHODOLOGY

We have developed an application based on a decision tree in data mining. First we collect the data from the real time users. Here we can create our own dataset. Then we build the decision tree model. And we can train the model with our dataset. After training our model, we just try to check its accuracy level. For that we are testing our model with the testing dataset and note down the accuracy.

After creation of the model, we build the user interface for interacting with the model. Here first we collect the data from the users from GUI and then we give that data to the decision tree algorithm to evaluate our model. Here when we start our model it is simultaneously train the model in the background. Finally we got the output from our decision tree algorithm. Here we can show the output to the user with the help of the web page. And here we provide the reason of dropping also.

1.4 SIGNIFICANCE OF WORK

Presently, there's rapid growth and improvement within the educational zone. This growth emerges from the current technologies and the technique that is directed toward enhancing student performance. But nowadays a big mission is for an educational institution to pick out the dropout college students at the beginning.

The significance of this project is to decline the academic dropouts in Higher Education (M.Tech admissions) by conducting an online examination using data mining algorithms to assign colleges based on allocated domains.

1.5 ORGANISATION OF THE REPORT

We have divided the proposed work into eight chapters. In the first chapter, we will give a brief introduction about the project which consists of the project area, scope of work, problem statements, the objective of the project, methodology, significance of work.

In the second chapter, we will discuss literature review which gives the Information about various projects related to our project, their proposed methodologies, merits, and demerits. In the third chapter, we will discuss System requirements like Hardware Tools, Communication interfaces, Software requirements/Hardware requirements. In the fourth chapter, we will discuss System design including System Architecture, Data Flow diagrams, and UML diagrams.

In the fifth chapter, we will discuss the Implementation and Proposed approach of the project. In the sixth chapter, we will discuss all types of testing related to the project. In the seventh chapter, we will discuss the Experimental results, Project outcome, and Comparison with the existing approaches. In the eighth chapter, we will discuss the Conclusion and Future Enhancement of the project using your approach.

1.6 CONCLUSION

Therefore, with the help of our application, we can greatly decline the academic dropouts in Higher Education (M.Tech admissions). Our application can effectively lower the dropout rate by conducting the online examination to assign colleges based on allocated domains using clustering technique and helps the student to take care of his/her career.

CHAPTER-2

LITERATURE SURVEY

2.1 PREAMBLE

Educational Data Mining refers to a set of special techniques for analyzing and extracting useful information from student data. In recent years the importance of this discipline has grown exponentially, seen, among others, by comprehensive literature. In many European educational systems, government subsidies are one of the most important (source: European University Association - EUA, 2018) to universities' economic stability. Mostly, in the Italian study system, the community support schema has changed rapidly in the file of 20 years ago, and the most complex method, linked to several different parameters, is used to convert the value of the income that an institution receives a year.

One such barrier is a measure of success, that is, the difference between the number of students who decide to take a course and the number of students who graduate, as long as one finds the right model for accuracy, ease of implementation, and data availability. For this paper, we will explain dropping out of school, the issue of non-completion of studies, and official resignation. A student can be automatically classified as an expulsion if he or she has not paid the full fee year, or more, or because you have notified the institution of the intention to leave that course and/or each institution.

2.2 RELATED TOPIC OF LITERATURE REVIEW

2.2.1Creating a Recommender System to Support Higher Education Students in the Subject Enrolment Decision:

In this paper [1], the challenge is though the quantity of college students has been increasing but not the graduating rate. So, they proposed a Decision Support System and focused on increasing graduation rates by constructing a Recommender System to assist students with their selection of subjects. Created a Decision Support System Based on a Recommender System using a dataset with few instances and imbalanced frequencies in the class label that can assist students in selecting the

subjects best suited to them.The Decision Support System deals with both prediction and ambitions at the student's decision making but doesn't provide an effective follow-up of the Student's academic path in comparison with the system recommendations.

2.2.2Predicting Achievement of Students in Smart Campus:

In this paper [2], there are a lot of information systems on the campus, thousands of educational data get generated day by day which may include ineffective data that leads to crises in predicting student achievement. So, they designed a student achievement predicting framework, which contains data processing and student achievement predicting. In data processing, data extraction, data cleaning, and feature extraction are explained. With the help of these data in a data warehouse, they proposed a layer-supervised multi-layer perception (MLP)-based method to predict the achievement of students.By comparing with SVM, Naive Bayes, logistic regression, and MLP algorithms, layer-supervised multi-layer perception (MLP)-based methods get better performance. Supervisors are fed to each corresponding hidden layer of MLP to improve the overall performance of student achievement prediction but didn't pay more attention to weblogs details to find the behavior pattern of students.

2.2.3 Identification of Drop out Students Using Educational Data Mining:

In this paper [3], identically, the wide variety of better knowledgeable persons in a rustic environment can contribute to the improvement of the country. But, this wide variety decreases due to the dropout of college students at an early level of schooling. Moreover, if a student can't continue or drop out, the sources of a state are attenuated. A big mission foreign educational institution is to pick out the dropout college students at the beginning. Several processes had been mentioned in educational data mining to discover the rate of dropout college students. Following this line in this paper, a threshold-based approach has been proposed to identify dropout college students that outperform the prevailing methods. To extract the critical functions, one only needs the characteristic values and their corresponding information gain. From the Extracted functions the threshold cost may be calculated. Overall performance has been proven in two stages: for authentic datasets and after detecting outliers. Each for the authentic datasets and the datasets after detecting

outliers this proposed method works higher. The imbalance of the datasets isn't always considered and removal of the imbalance can't be explained by the proposed technique.

2.2.4 Comparing Different Resampling Methods in Predicting Students' Performance Using Machine Learning Techniques:

In this paper [4], in today's international, because of the advancement of the era, predicting the students' performance is a few of the most useful and essential research topics. It is a fact that predicting the student's performance has come to be a severe task due to the imbalanced datasets in this subject, and there is not any assessment among one-of-a-kind resampling techniques. This paper tries to evaluate diverse resampling strategies which include borderline smote, random over sampler, SMOTE, SMOTE-ENN, SVM- SMOTE, And SMOTE-Tomek to address the problem of the imbalanced fact at the same time as predicting students' overall performance the usage of two specific datasets.The effects of the Friedman take a look at, which is statistical importance, take a look at, confirm that the SVM-smote is greater efficient than the opposite resampling strategies. Moreover, the random forest classifier has completed the great result among all other models while using SVM-smote as a resampling approach.

2.2.5 Clustering Algorithms in an Educational Context: An Automatic Comparative Approach:

In this paper [5], regardless of a growing consensus regarding the importance of nicely identifying the maximum appropriate clustering method for a given problem, a shocking quantity of tutorial research, which includes both academic data mining (EDM) and getting to know analytics (L.A.), neglects this essential project. This shortcoming may want to in many cases harm the prediction electricity of each EDM Primarily based tactics. To cope with such problems, this work proposes an assessment approach that mechanically compares several clustering methods using more than one internal and external performance measure on nine actual-global educational datasets of different sizes, created from the University of Tartu's Moodle gadget, to supply two-way clustering.The proposed method employs a multiple criteria selection-making approach (i.e., TOPSIS) to rank the maximum appropriate techniques for every dataset. Our consequences monitor that the proposed approach

can mechanically examine the overall performance of the clustering strategies and therefore, endorse the maximum appropriate approach for each dataset. The proposed method can't be applied to larger datasets.

2.2.6 An Early Feedback prediction System for Learners At-risk within a First-year Higher Education Course:

In this paper [6], identifying at-risk students as soon as possible is a challenge in educational institutions. Decreasing the time lag between identification and real at-risk state may significantly reduce the risk of failure or challenge. They proposed a 'Predictive model' based only on student's grades specifically trained for each course. An early warning system has been developed focusing on dashboards visualizing for attack holders (i.e., students and teachers), and an early feedback prediction system to intervene in the case of at-risk identification. The merits are early detection of potentially at-risk students, better guidance of students with visualization dashboard and feedback, and increment of the interaction with at-risk students. We need to improve the accuracy of the model by enhancing some other features, and here we predict the student's performance after joining the course.

2.2.7 Early prediction of dropout and final exam performance in an online statistics course:

In this paper [7], Higher education students who either do not complete the courses they enrolled in or interrupt their studies indefinitely remain a major concern for practitioners and researchers. They Proposed two procedures for identifying at-risk students (dropouts-prone and non-achievers) early on in an online university statistics course. Here they identify dropout-prone and fail-prone students before the halfway point of the semester, through tree-based classification models. It provides a simple and interpretable procedure to identify dropout-prone and fail-prone students before the halfway point of the semester. It has low classification performance values of which teachers need to evaluate, and the second limitation is associated with the validation set.

2.2.8 Predicting At-Risk Students at Different Percentages of Course Length for Early Intervention Using Machine Learning Models:

IN this paper [9], Online learning platforms have many advantages as well as they face several challenges such as student's lack of interest, high dropouts,

low engagement, student's self-regulated behaviour, and compelling students to take responsibility for setting their own goals. They proposed a predictive model that analyzes the problems faced by at-risk students, subsequently, facilitating instructors for timely intervention to persuade students to increase their study engagements and improve their study performance. The predictive model was trained and tested using ML and DL algorithms. Predicting models can facilitate instructors to make timely inventions and persuade at-risk students to improve their study performance. In the future, we can also add a feedback system. Here we predict after the student joins the course.

2.2.9 An Integrated Framework with Feature Selection for Dropout prediction in Massive Open Online Courses:

In this paper [10], Massive open online courses have flourished in recent years, which is conducive to the redistribution of high-quality educational resources globally. However, the high dropout rate in the course of operation has seriously affected its development. They proposed an integrated framework with feature selection (FSPred) to predict dropouts in MOOCs, which includes feature generation, feature selection, and dropout prediction. After finding the affected features of the student dropout, then they take certain measures to solve that problem. Hereafter joining the course only we can find the students performance and interest, it is the disadvantage.

2.2.10 Usage of Machine Learning for Strategic Decision Making at Higher Educational Institutions:

In this paper [10], Decisions made at the strategic level of Higher Educational Institutions affect policies, strategies, and actions that the institutions make as a whole. The problems are: The decision process takes longer, the whole picture is not involved along with all data necessary, and small academic impact is produced by the decision, among others. In this paper, they used three supervised classification algorithms which are Decision Trees, Random Forests, and Logistic Regression, among random forests, which gives the best results. Having these students recognized early can allow HEIs governance to reduce the student's dropout rates, retention rates. We can also include analysis of other effectiveness

metrics such as F-measure, etc.,

2.2.11 A Model to Predict Low Academic Performance at a Specific Enrolment Using Data Mining:

In this paper [11], they have used two data mining models to predict the loss of academic performance not only by using the socio-economic data but also the academic records. Here they have used the Decision tree and Bayesian classifier to build the academic models. They also used the C4.5 algorithm to manage the discrete and continuous attributes. They also studied the academic success of the academic performance at a given course to predict the failure at completing the course and the failure in passing the course. To do this they have used the students' past and present performance in the class or year or throughout the course. The admission data sets that can be classified into three categories are Initial Academic Information. Demographic, Socio-Economic and Academic Potential.

2.2.12 Predicting Students' GPA And Developing Intervention Strategies Based on Self-Regulatory Learning Behaviour:

In this paper [12], they have used different techniques to identify the factors that influence and predict the overall student behaviour that differs in accuracy, complexity, and sample size requirements. Based on the several samples that have been collected they have constructed a maximum-weight first-order dependency tree(MWDT). They also found the error classifier by using the datasets and processed them based on the data they have taken.

2.2.13 Predicting Student Performance and its Influence Factors Using Hybrid Regression and Multi-Label Classification:

In this paper [13], they have studied the different student performance at different courses and they identified the key factors that lead to the achieved performance and those identification gives new ideas to the program leads to increase the ability to the success rate of the students. Firstly they have used the hybrid regression model to optimize the accuracy of the predictions of the student's performance. Then after they have used the multi-label classifier to find the qualitative values of the factors that influence the student's academic performance.

2.2.14 Improving Predicting Modelling for At-Risk Student Identification: A Multi-Stage Approach:

In this paper[14], they have studied the research of the students' performance prediction and they have classified their studies to higher education-related research and the lack of success rate in the prediction due to the lack of information they have taken and the research gaps also lead to the failure of predictions. Here, they have introduced a learning management system (LMS) to identify the students' performance during the online learning platforms. The system can track the behaviours of the students during the courses and their learning behaviours.

2.2.15 Using Data Mining Techniques to Predict the Student Performance to Support Decision Making in the University Admission Systems:

In this paper[15], they have studied the data sets related to admissions. They have done this based on data mining techniques to support the educational institutes to analyze their educational databases to predict student performance. And also support them in decision making to set efficient admission criteria. They have done a wide range of research by using data mining techniques in the medical field also. They have used Artificial Neural Network (ANN), Decision Tree, Naive Bayes, and Support Vector Machine (SVM).

2.2.16 Predicting Instructor Performance Using Data Mining Techniques in Higher Education:

In this paper [16],the challenge is to enable online instructors to develop corresponding instructional via course design or student-teacher communication.So,they have used decision trees and linear discriminant analysis, support vector machines, and artificial neural networks are used to build classifier models. The results indicate that the best model has high practical value to capture at-risk students while the course is still in progress.In general, research in educational mining focuses on modelling student's performance instead of instructor performance. One of the common tools to evaluate an instructor's performance in this course evaluation questionnaire is to evaluate based on student's perceptions.

2.2.17 Educational Data Mining for Tutorial Support in Higher Education: A Web-Based Tool Case Study in Engineering Degrees:

In this paper [17], Nowadays it also supports the recording of the resulting retention-oriented interventions for the further analysis is challenging .So,they focused on the analysis of student's performance, in terms of observable scores of the completion of their students.Among all the retention actions, phone calls seem to be more effective than written interactions.The analysis of the impact of the student's explanatory variable and early performance in the graduation probability can lead to a better understanding of the causes of dropout. Real data from engineering degrees of EU Higher Education institutions show the potential of the tool for managing higher education and validate its applicability in the real scenario.

2.2.18 From Lab to Production: Lessons Learnt and Real-Life Challenge of an Early Student- Dropout Prevention System:

In this paper [18], The challenge is to prevent students from abandoning the university by means of retention actions centred at the maximum at-chance Students, seeking to maximize the effectiveness of institutional efforts on this path.So,this paper presents the work done to support student dropout risk prevention in a real online e-learning environment. This generated the predictive models based on the C5.0 algorithm using data from more than 11,000 students collected over five years. The white-box predictive models used in production furnished reasonably appropriate results, very near those received in the laboratory. Then they developed SPA (Sistema de Predicción de Abandono, dropout prediction system in Spanish), an early warning system that uses these models to generate stats, early dropout risk prediction and dynamic periodically updated ones. This used warning systems, prediction methods, and student dropout.

2.2.19 Identifying At-Risk Students for Early Interventions-A Time -Series Clustering Approach:

In this paper [19], The challenge is to perceive at-hazard online students in advance, greater frequency, and with greater accuracy the usage of Time-series clustering.So,they proposed a best-performing model which can start to capture at-risk students from week 10. In addition, the four phases in the student learning process detected the holiday effect and illustrate at-risk student’s behaviours before and after a long holiday break. In this, they used clustering, classifications, LMS; predictive modelling.The findings additionally allow online instructors to broaden corresponding instructional interventions via route layout or student-instructor communications.

2.2.20 A Surrogate-Assisted Multi-objective Evolutionary Algorithm for Large-Scale Task-Oriented Pattern Mining:

In this paper [20], the problem was based on the framework of our previous multi-objective evolutionary algorithm for task-oriented pattern mining; the proposed algorithm estimates the objective values of most solutions using an ensemble of surrogates instead of the real objective functions, thereby highly improving the efficiency of the algorithm. Experimental outcomes on three assignment-oriented programs indicate that the proposed set of rules has higher efficiency than the latest algorithms. In this, they used data-driven optimization, surrogate, and multi-objective optimization.

Table 2.3: Comparison of Student drop-out Literature Survey

S.No	Author and Year	Proposed Approach	Merits	Demerits
1	Antonio Jesus Fernandez Garcia, 2020.	Proposed Decision Support System which specializes in maximizing graduation rates by using building a recommender gadget to assist students with their selection of subjects	It not only deals with predictions but also ambitions at Assisting college students and college acquire better outcomes by way of assisting them for their decision-making.	It doesn't assist follow-up strategy that lets in the recommender gadget to be updated according to with Changes within the diploma direction's business enterprise.
2	ShaojieQu, 2018.	Proposed a layer-supervised multi-layer perceptron (MLP)-based technique to predict the	As compared with SVM, Naive Bayes, logistic regression, and MLP, layer-supervised multi-	It didn't pay more attention to weblogs details to find the behaviorpattern of students.

		achievement of students.	layer perception techniques get better performance.	
3	Nafisa Tasnim, 2019.	Proposed threshold-based approach to discover dropout students that outperforms the prevailing methods.	The threshold-based approach is used for both authentic datasets and the datasets after detecting outliers.	This approach doesn't work for imbalanced datasets.
4	Ramin Ghorbani, 2020.	The random forest classifier has completed the quality result among all different models during the use of SVM-smote as a resampling technique to predict student performance.	The consequences of the Friedman test, which is of statistical importance take a look at, to verify that the SVM-Smote is more efficient than the other resampling techniques. The overall performance of classifiers is stepped forward by the use of balanced datasets.	It didn't use the Hybrid classifiers for having a better evaluation and also reaching higher results in the prediction of student performance.
5	DanialHoosh yar, 2020.	TOPSIS technique to rank the results produced using	This method no longer simply locates the maximum	The proposed automatic The approach couldn't be applied

		more than one inner and external overall performance measure when carried out to several clustering algorithms. This approach might help to automatically endorse the best clustering methods for each dataset.	suitable clustering method for every dataset however additionally investigates the possible impact of the use of distinctive sizes and sorts of educational datasets on the performance of clustering algorithms in two-manner clustering.	to larger educational datasets and unable to reduce the effect of noise and outliers in large datasets automatically.
6	David Baners, M.Elena Rodriguez and Montse Serra, 2019.	Proposed a predictive model, An early Warning System along with an early feedback system.	Early detection of potentially at-risk students, better guidance for students, and increment of interaction with students.	We need to improve the accuracy of the system by enhancing some other features, and we predict a student's performance after joining the course.
7	Josep Figueroa-Canas, and Teresa Sancho-Vinuesa, 2020.	Proposed a tree-based classification model.	It provides a simple and interpretable procedure to identify dropout-prone and fail-prone students	It has low classification performance values, and we calculate the performance of the students after

			before the halfway point of the semester.	joining the course.
8	Muhammad Adnan, AsabHabib, Jawad Ashraf, ShafaqMuaasdiq, Arsalan Ali Raza, Muhammad Abib, Maryam Bashir, and Sana Ullah Khan, 2021.	Proposed a Predictive model which uses Machine Learning algorithms and compared using accuracy, precision, recall, support, and f-score.	It helps instructors in identifying the at-risk early in the course for timely intervention thus avoiding student dropouts.	Need to examine the activity-wise significance, and need to build a feedback system.
9	Lin Qiu, Yanshen Liu, and Yi Liu, 2018.	Feature selection Prediction Framework (FSPred).	They find the affected people and then assess them to save them from dropout from the program.	They predicted the students who have already joined the course only.
10	Yuri Nieto, Vicente Gacia-Diaz, Carlos Montenegro, Claudio Camilo Gonzalez, and Ruben	Proposed a Supervised classification algorithm Decision Trees algorithm.	By recognizing the students they can reduce the students' dropout rates by assessing them to some mentors.	We need to include analysis of other effectiveness metrics such as F-measure, etc.,`

	Gonzalez Crespo, 2019.			
11	Camilo Ernesto Lopez Guarin, 2017.	Two Data mining models.	They predicted the student's information for only first and second-year students.	They have predicted only the graduated students' information.
12	Amin Zollanvari, 2018.	Constructed a predictive model.	They have proposed the Maximum-Weight Decision tree to predict the GPA.	They have predicted only based on the grade point bases.
13	Abdullah Alshanqiti, 2020.	They predicted student performance by combining the three dynamically weighted techniques.	They have used the fuzzy rules, lasso linear regression, matrix factorization.	By using the prediction techniques the complexity of the increase decreases the success rate of the student predictions.
14	Jui-Long Hung, 2019.	These studies provide the two-stage approach.	They evaluated the effectiveness based on the approach but the complementary datasets.	The approaches they used resulted in high unbalanced data.
15	Hanan Abdullah Mengash, 2020.	They have created an approach to support the educational institutes to	By providing an approach to identify the student's academic	They have only predicted the reasons for the success rate of the students'

		identify the student information.	performance and their success rate.	performance.
16	Mustafa Agaoglu, 2016.	Seven classification models are generated: two the usage of decision tree algorithms (c5.Zero, and cart), one the usage of SVM, three the use of ANNS, and one using DA.	C5.0 classifier takes care of most of the decisions mechanically using fairly affordable defaults.	It didn't explain about the subject of the course are more important to students than instructors' behavior in evaluating the instructor Performance.
17	Ramon Vilanova, 2020	Proposed a web-based software tool that is focused on the analysis of students' Performance, in terms of the observable scores and of the completion of their studies.	It is conceived to provide easy upload and processing of the academic data, different perspectives on the information, and interactivity for self-guided exploration.	It didn't include further information about classroom attendance and results at the course level, obtained from learning management systems, to the analysis.
18	Alvaro Ortigosa, 2019.	Proposed SPA (Sistema de Predicciœn de Abandono, dropout	It provided a reasonably Good separabilityof classes, close to	Good results achieved in the laboratory may not be possible in the long term in a real

		prediction system in Spanish), an early warning system that uses these models to generate static early dropout-risk predictions and dynamic - periodically Updated ones.	the performance of the Random Forest models tested in the lab: addressing all the students with dropout risk scores greater than 25% would have covered 60% of real dropouts and 20% of persistent Students.	production system due to Sustainability issues.
19	Jui-Long Hung,2017.	Distance learning, engineering education, higher education, The final course provides useful input for future developments and a more effective student-focused	A decision tree is a decision support tool that uses a tree-shaped graph or model for classification. It is a supervised learning method.	F students might be a little behind or less engaged than other students in the class. The long holiday break catalyzes to trigger the at-risk situation.
20	FuchenDuan, 2019.	In this, the author is based on Mining methods and Algorithms, Time- Series Analysis, LMS, Predicate modeling data.	Educational data mining can be used for classifying and predicting students' performance, dropouts as well as teachers'	"Restriction of processing" right: functionality to stop the processing of any individual's data by the system and the consequent cessation of derived

			performance.	information production

2.4 CONCLUSION

Currently, there is rapid growth and improvement within an educational zone. But there is still an increase of academic dropouts day by day due to various reasons like ambiguity in domain selection, lack of knowledge in selected domains, etc. So, the above problems need to be overcome.In this chapter, the existing technologies used for reducing academic dropouts were present. The existing systems have their limitations and need to be solved. All the existing systems need to be improved to get a better prediction system to decline academic dropouts.

CHAPTER-3

SYSTEM REQUIREMENTS

PREAMBLE

The previous chapter describes the literature review related to the prediction of academic dropouts and various approaches and techniques of detecting fake reviews. We have discussed the various research papers and journals with their merits, demerits and future direction in chapter 2. In this chapter we describe the different kinds of tools used and different requirements to develop the proposed system. We provide the details about tools used like front end tools, back end tools, hardware tools and different system requirements like hardware and software requirements of the system.

3.1 TOOLS USED

3.1.1 FRONTEND TOOLS

HTML:

Html stands for a hypertext markup language. Its miles used to layout internet pages using a mark up language. Hypertext defines the hyperlink among the web pages. A mark up language is used to outline the textual content record inside the tag which defines the structure of web pages.Html is a markup language that is used by the browser to manipulate text, photos, and different content material to display it in the required format.

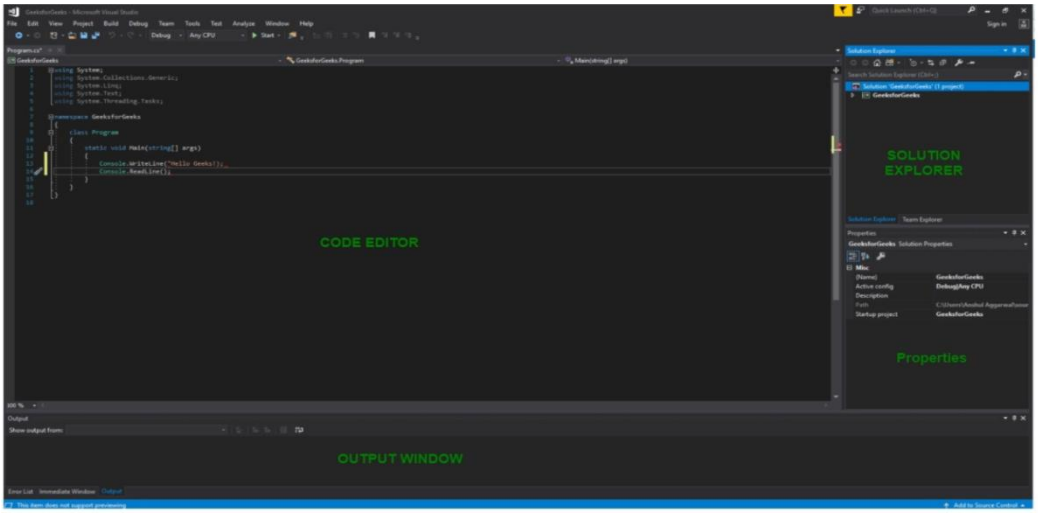
```
1 <html>
2 <head>
3     <title></title>
4 </head>
5 <body>
6
7 </body>
8 </html>
9
```

CSS:

CSS (cascading style sheets) is a stylesheet language used to layout the webpage to make it attractive. The reason for the usage of CSS is to simplify the process of making net pages presentable. CSS permits you to use patterns on internet pages. More importantly, CSS permits you to try this independent of the HTML that makes up every net web page.

VISUAL STUDIO:

Visible Studio is an incorporated improvement environment (ide) developed by Microsoft to expand GUI(graphical person interface), console, web programs, web apps, cellular apps, cloud, and internet services, etc. With the help of this idea, you can create controlled code in addition to local code. It makes use of the various structures of Microsoft software improvement software programs like windows keep, Microsoft Silverlight, and home windows API, and so on. It is not a language-precise idea you could use this to put in writing code in C#, C++, VB(Visual Basic), python, JavaScript, and lots of extra languages. It gives a guide for 36 extraordinary programming languages. It is to be had for windows in addition to for macOS.



Components of Visual Studio:

Code Editor: where the user will write code.

Output Window: right here the visible studio shows the outputs, compiler warnings, error messages, and debugging facts.

Solution Explorer: it suggests the files on which the consumer is presently working.

Properties: It'll give additional statistics and context about the chosen components of the contemporary challenge.

3.2 BACKEND TOOLS

JAVASCRIPT:

JavaScript is the world's most famous lightweight, interpreted compiled programming language. It's also known as a scripting language for internet pages. It is well-known for the improvement of net pages; many non-browser environments additionally use

it. JavaScript may be used for client-facet trends in addition to server-side trends.

JavaScript can be delivered on your html record in two methods:Innerjs: we are able to add java Script immediately to our html record by writing the code in the <script> tag. The <script> tag can either be positioned within the <head> or the <body> tag in keeping with the requirement.Outsidejs: we can write javascript code in other record having an extension .Js and then hyperlink this file in the <head> tag of the html document wherein we need to feature this code.

```
2  <html lang="en">
3
4  <head>
5    <title> Adding JavaScript to HTML</title
       >
6    </head>
7    <script language ="javascript" type
       ="text/javascript">
8
9    </script>
10 </body>
11
```

PHP:

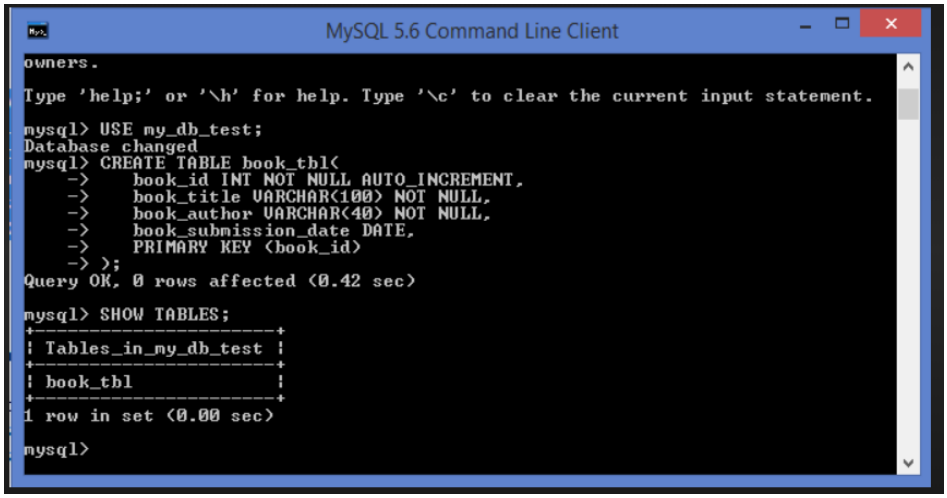
The term Hypertext Preprocessor is an acronym for Hypertext Preprocessor: Hypertext Preprocessor. Hypertext Preprocessor is a server-facet scripting language designed especially for net development. PHP may be effortlessly embedded in HTML files and HTML codes also can be written in a PHP record. The component that differentiates Hypertext Preprocessor from a consumer-side language like HTML is, PHP codes are executed at the server while HTML codes are directly rendered on the browser.

```
<html>
<body>
<?php
echo "Hello World";
?>
</body>
</html>
```

MYSQL:

Mysql is a quick, smooth-to-use RDBMS getting used for plenty of small and massive groups. Mysql is advanced, advertised, and supported by MySQL ab, which is a Swedish organization. Mysql is becoming so popular due to many correct reasons:

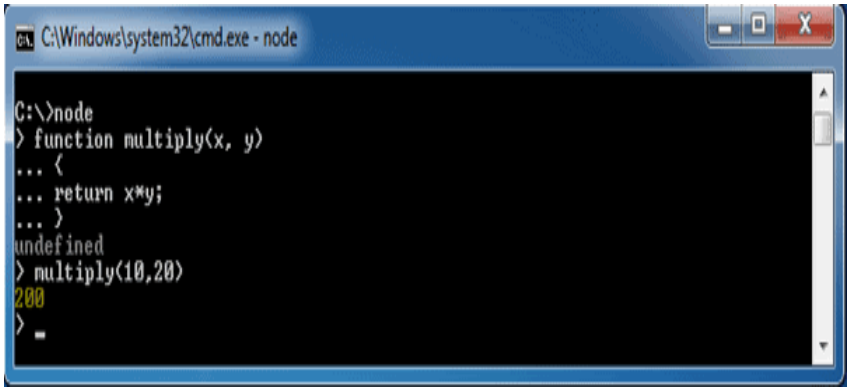
- Mysql works on many running systems and with many languages such as Hypertext Preprocessor, perl, c, c++, java, and many others.
- Mysql works in no time and works nicely in spite of massive statistics sets.
- Mysql may be very friendly to personal home pages, the most favored language for internet development.
- Mysql supports massive databases, up to 50 million rows or more in a desk. The default document size limit for a desk is 4gb, but you could boom this (in case your working device can manage it) to a theoretical limit of eight million terabytes (tb).



Node.js:

Node.js is an open-source runtime environment which is built on the chrome V8 JavaScript engine . Node.js can be used to build different types of applications such as command-line application ,web application etc , where it is mainly used to build network programs like web servers for example PHP, Java etc .

- ❖ It uses JavaScript for server-side applications.
- ❖ Lightweight Framework that includes minimum modules .other modules can be included as per the need of the application.
- ❖ By default it is asynchronous so it can perform faster than any other frameworks.
- ❖ Cross-platform framework that can run on windows, Linux .



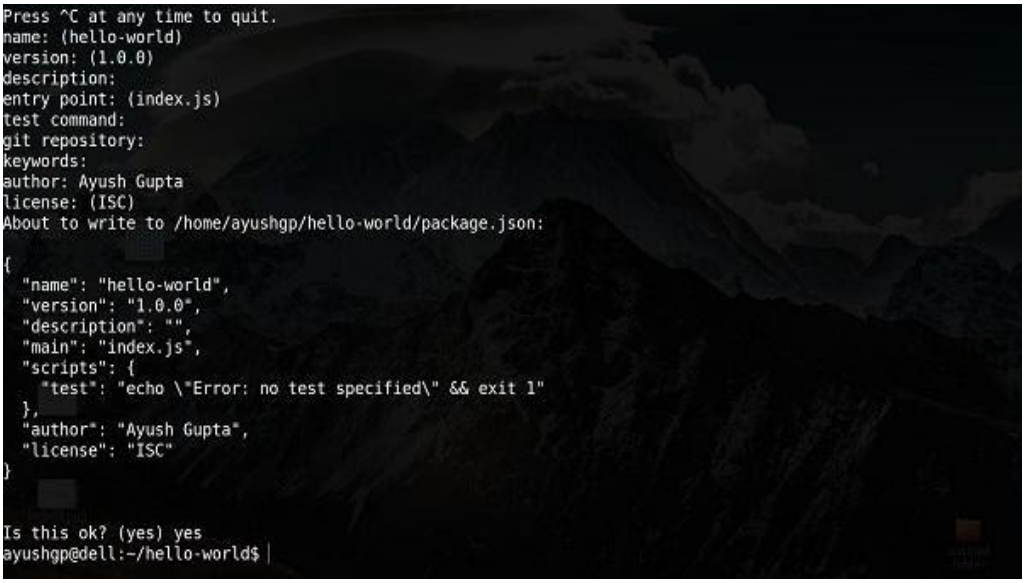
Express.js:

Express is a flexible and minimal Node.js web application that provides a set of features for the mobile and web applications .It provides a minimal interface for our application. It is flexible as there are numerous modules available on npm ,which can be directly plugged into Express .

- PUG is the most used language to build the HTML pages in Express.

- MongoDB is an open source database designed for the ease of Development and Scaling and also used to store the data.

- Mongoose is a client API for Node .js which makes it easy to access the data from our Express Application .



3.3 HARDWARE REQUIREMENTS

The hardware requirements required to develop this project are discussed here. We have applied our experiments on a system with processor Intel or compatible Pentium class 3 processor or higher and with memory or RAM at least 4GB. The system type is 64 bit operating system (OS), x64-based processor and the hard disk with 4GB. We have used windows (7/10) as our operating system. We have used these hardware requirements to develop the project using HTML, CSS, JavaScript, PHP and MySql. We have given the summary of software and hardware requirements in the next section of this chapter.

3.4 COMMUNICATION INTERFACES

The communication among the I/O gadgets and the processor of the computer device has applied the use of an interface unit. In a pc machine data is transferred from an input tool to the processor and from the processor to an output device.

In our project,the browser acts as a communication interface which provides interaction between user and system. So, we are considering Google Chrome or Internet Explorer as browsers of our project.

3.5 SOFTWARE/HARDWARE REQUIREMENTS

- HTML,CSS, JavaScript, PHP and MYSQL
- Computer: Intel Core i3 processor
- Memory (RAM): 4GB
- OS: Windows 10
- Editor: Visual Studio
- Browser: Google Chrome / Internet Explorer
- System type: 64-bit Operating System

3.6 CONCLUSION

This chapter introduced the software and hardware requirements of the system. The above requirements are required for the successful implementation of our project.For the implementation of an application of online test we have chosen HTML, CSS, JavaScript, PHP, MYSQL and also using decision tree algorithms in data mining for better implementation and accurate results. We have used a system with the 64-bit windows operating system and an Intel core Processor device. The hardware requirements are the basic ones that are required for the execution of the application.

CHAPTER-4

SYSTEM DESIGN

INTRODUCTION

The previous chapter describes the different requirements and tools used for developing the proposed system. It also consists of the hardware requirements like processor, memory and software requirements like Node JS, Visual Studio and frontend and backend tools. This chapter deals with the design of the proposed system. It includes the system architecture, data flow diagram and various UML diagrams to represent the design of the system.

4.1 SYSTEM ARCHITECTURE DIAGRAM

The System architecture is the conceptual model that defines the shape, conduct, and greater perspectives of a gadget. An architecture description is a formal description and illustration of a system, prepared in a way that supports reasoning about the structures and behaviour of the device. A system architecture can consist of a device and the sub-structures advanced, so one can work together to implement the general machine.

The system consist several modules, so it works as seen below. First it started, after starting the system we collect the data from the real world and store it on the database. So from there it will take the dataset. After taking the set we simply try to train the model in the following way:

First it separates the data into parts based on the class label, and then it will calculate the information gain for each and every node. After calculating the information gain it will select the best information gain value as the root node. Then we will take the remaining data and we will separate based on the data values. And after again we calculate the information gain for the each attribute and try to take the best information gain to pick the nest node. This process continues until all the attributes are finished on the tree. And finally we can have a decision tree. And we complete our training part of the model.

Then we collect the data from the user by using graphical user interface and give that data to the model to find out the results. And then the Decision Tree model gives the output. And we collect this output and displayed to the user via graphical user interface. This is the architecture of our system.

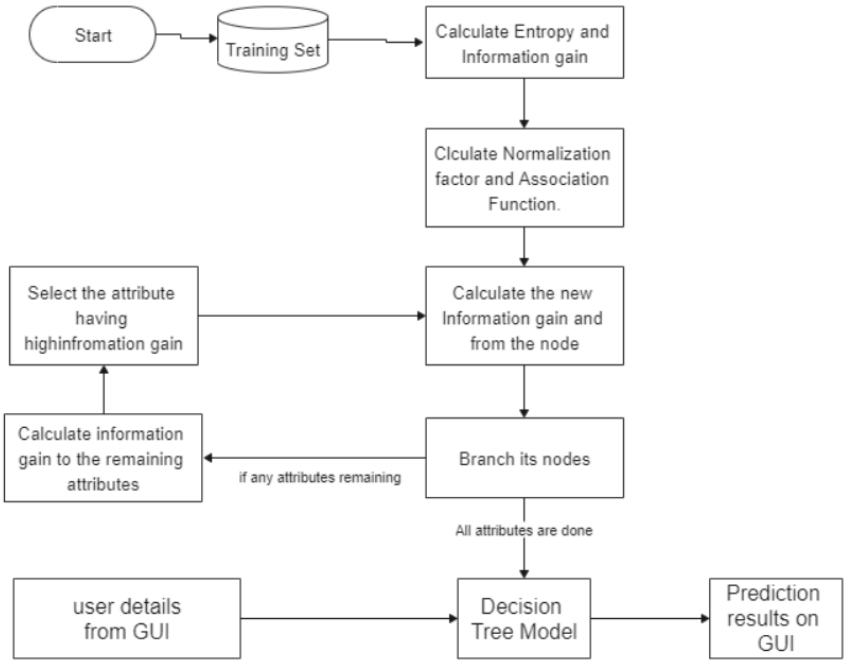


Fig. 4.1: Architecture Diagram

4.2 DATA FLOW DIAGRAM

A data flow diagram (DFD) is a way of representing a guide of records of a method or a machine. The data flow diagram additionally affords statistics approximately the outputs and inputs of each entity and the process itself. Specific operations primarily based on the statistics may be represented by means of a flowchart. An information float diagram is a way of representing a waft of statistics of a process or a machine. It also presents the internal and external entities of the gadget.

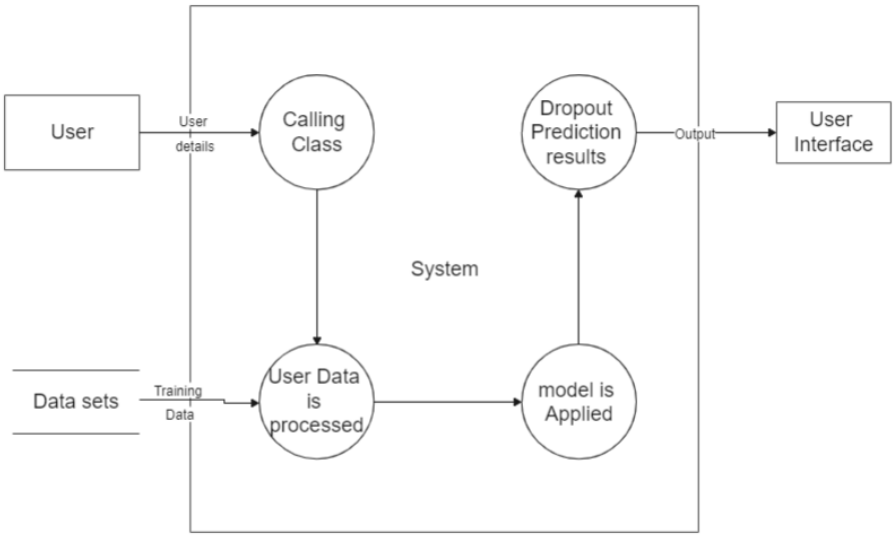


Fig: 4.2 Data Flow Diagram

In the Data Flow Diagram (DFD), the square boxes represent the entities, circles represent the process and named arrows represent the direction of flow of information. We have considered the user as an entity, and it is shown in the sure box in the figure. The processes represented in the data flow diagram are calling class, User data processing, modeland results.

User initially enters into the system and he gives required details to the system. Then it go to the calling function module. From there it will validate those details and from there the data is go to the data processing phase. Here it will collect the data from the student's dataset and processing it here. Then the data is transfer to the model. If it is a training data then the machine is get trained. Or if it is the testing data then it will evaluate the model. And it produces some output and it sends that to the output dashboard. From there we display the results to the user via graphical user interface.

4.3 USE CASE DIAGRAM

The below use case diagram shows actors and use cases for predictive models to reduce academic dropouts by using decision trees.

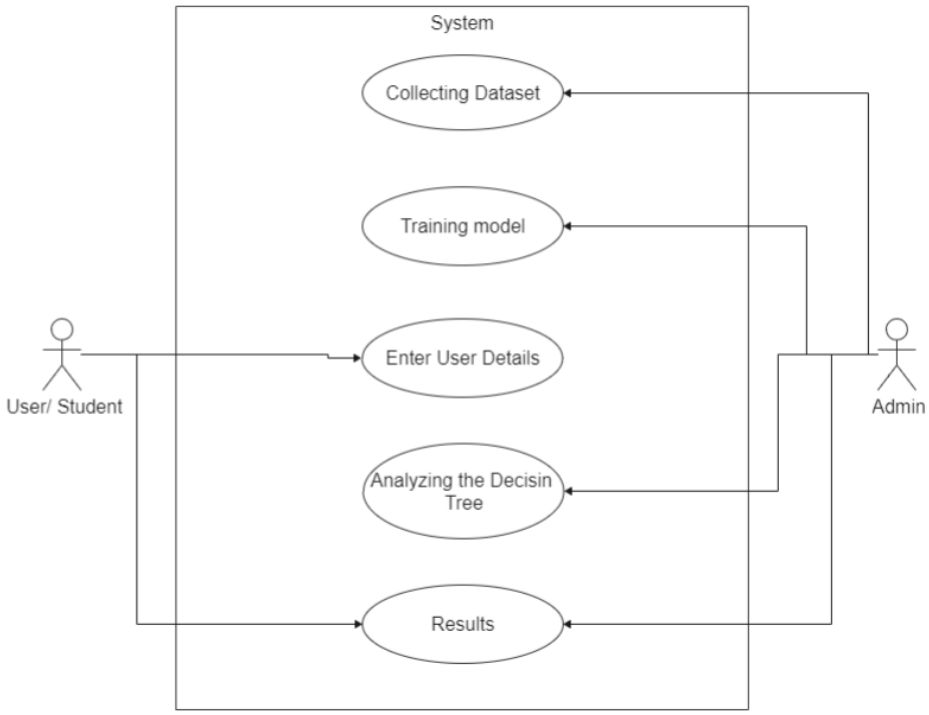


Fig: 4.3 Use case Diagram

Actor

An actor inside the UML "specifies a position played by way of a user or another machine that interacts with the difficulty".

Use case

A use case diagram at its simplest is an illustration of a person's interaction with the gadget that indicates the relationship among the user and the one-of-a-kind use instances wherein the user is involved. The use cases are represented by way of both circles or ellipses.

User

In this use case diagram, the user is the only person who is attending the online test and then evaluates the test and displays the resulting domain with a related list of colleges.

4.4 CONCLUSION

The design of the system is explained in chapter 4. This chapter provides information about the system design. Here, the system architecture diagram and data flow diagram of our proposed system has been discussed. This chapter deals with the design of the proposed system. It includes the system architecture, data flow diagram and various UML diagrams to represent the design of the system. This project has only system architecture, use case diagrams and a data flow diagram. The next chapter describes the implementation details like modules or steps in implementation and proposed approach, techniques and equations.

Chapter-5

IMPLEMENTATION

PREAMBLE

The previous chapter describes the design of the proposed system. It deals with the design phase consisting of system architecture, data flow and various UML diagrams. The next step after designing is the implementation of the proposed approach. This chapter deals with the implementation of the proposed approach. This chapter consists of the proposed approach and various steps required implementing the proposed approach. This chapter gives a brief view about the implementation of the system using various steps and proposed approaches.

5.1 PROPOSED APPROACH

In the proposed approach to decline academic dropouts in M.Tech , we develop a web application based on the decision tree algorithm. Here initially we take the students dataset which is having the full information about the students who passed on the MTech programme. Then we build a system which is having the three modules which are given below. Here first we build the web page which is used to take the student details and then we send that details to the already trained algorithm. Then that model is processing this data by using its knowledge and then it will produce some output that is the prediction of the student information that to the chance of dropout percentage. And we show these results on the dashboard for user reference.

5.2 PROPOSED METHODOLOGY MODULE

The following modules are proposed in our methodology:

- Calling Class
- Decision Tree
- Binary Node

We split our project into above three modules. And each module is having its own working and they are connected to each other, we can see it as follows:

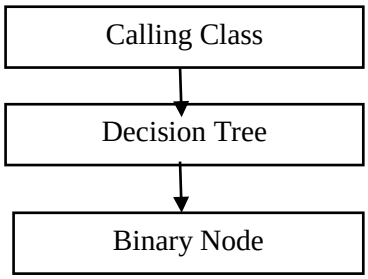


Fig 5.2: System module division

Calling class module includes the information collection from the user interface and it is having another work that it will take care about the data which is coming from the user is valid or not we will check here. Then finally if the data is valid then simply we will collect that data from the user and we give it to the decision tree algorithm.

Decision Tree module includes lot of tasks and it is the heart of our project. Here we will do building, training and testing the data and all are done here.

Binary node, since the system is constructed in an object oriented trend, considering all the involvement as objects. Hence, representation of each particular value, classes of the each attribute in form of node is necessary in order to create the decision tree.

5.3 CALLING CLASS

Calling class module includes the information collection from the user interface. Here we build a web page by using the html and node js and then we can collect the data from the user. It is having another work that it will take care about the data which is coming from the user is valid or not we will check here. Then finally if the data is valid then simply we will collect that data from the user and we give it to the decision tree algorithm.

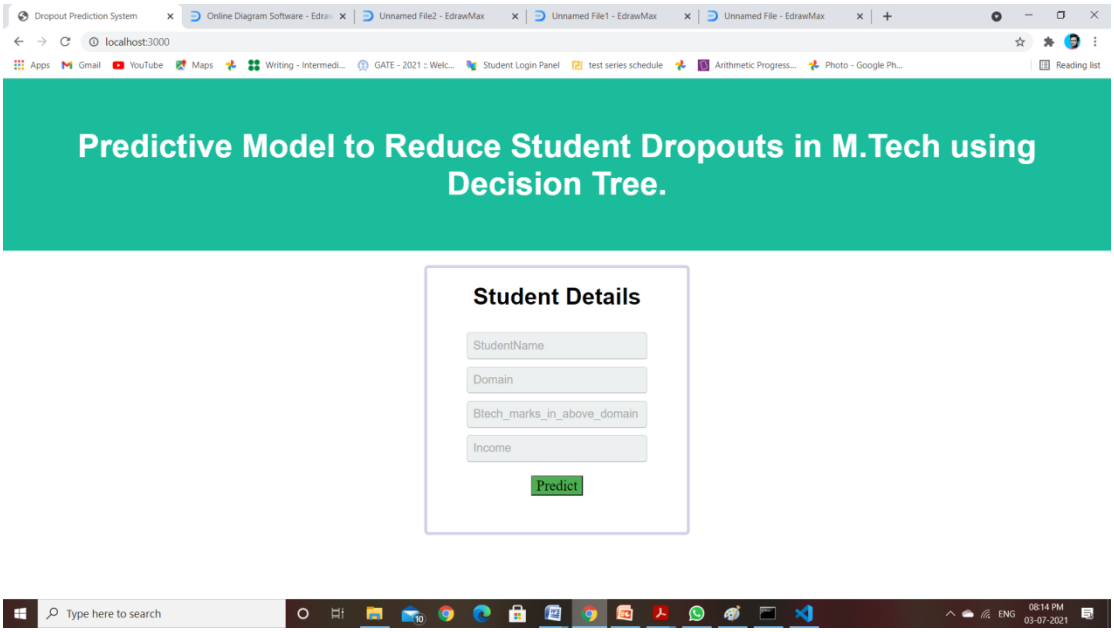


Fig 5.3: Student Details Collection.

5.4 DECISION TREE MODULE

Decision tree module consists of number of tasks. It is the heart of the project. In this module we take the details from the calling function and save them. First we collect the dataset from the real world facts. Then we can do the following process: First it separates the data into parts based on the class label, and then it will calculate the information gain for each and every node. After calculating the information gain it will select the best information gain value as the root node. Then we will take the remaining data and we will separate based on the data values. And after again we calculate the information gain for the each attribute and try to take the best information gain to pick the nest node. This process continues until all the attributes are finished on the tree. And finally we can have a decision tree. And we complete our training part of the model. Then we store this model for our future purpose.

5.4.1 DECISION TREE ALGORITHM

Decision tree algorithm is a supervised learning technique that can be used for classification problems. It is a tree structured classifier, where internal nodes represent the features of the dataset, branches represent the decision rules and each leaf node represents the outcome. In the decision tree, there are two nodes, which are the Decision node and Leaf node. Decision nodes are used to make any decision and

have multiple branches, whereas leaf nodes are the output of those decisions and do not contain any further branches. The decisions or the test are performed on the basis of features of the given dataset. It is a graphical representation for getting all the possible solutions to a problem/decision based on the given conditions.

It is called a decision tree because, similar to a tree, it starts with the root node, which expands on further branches and constructs a tree-like structure. In order to build a tree, we use the CART Algorithm (Classification and Regression Tree algorithm).A decision tree simply asks a question, and based on the answer(Yes/No), it further splits the tree into sub trees.

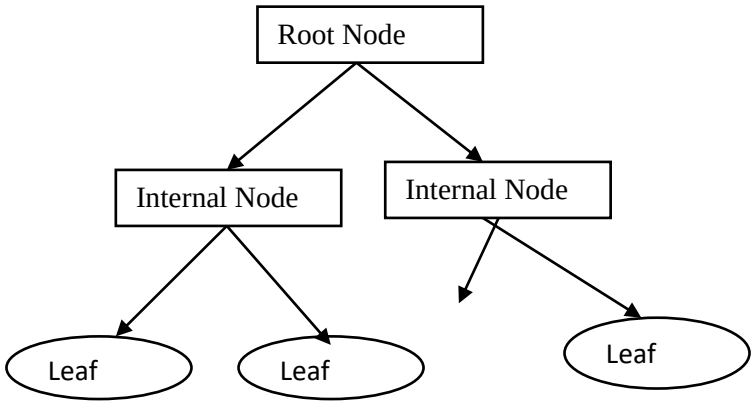


Fig 5.4.1: Decision Tree Structure

In a decision tree, for predicting the class of the given dataset, the algorithm starts from the root node of the tree. This algorithm compares the values of the root attribute with the record attribute and, based on the comparison, follows the branch and jumps to the next node.

Algorithm:

- Begin the tree with the root node, says S, which contains the complete dataset.
- Find the best attribute in the dataset using Attribute Selection Measure (ASM).
- Divide the S into subsets that contain possible values for the best attributes.
- Generate the decision tree node, which contains the best attribute.
- Recursively make new decision trees using the subsets of the dataset created in

step-3. Continue the process until we reach the leaf node.

5.4.2 ATTRIBUTE SELECTION MEASURE

While implementing the decision tree, the main issue arises that how to select the best attribute for the root node and for the sub-nodes. So, to solve such problems there is a technique which is called Attribute selection measure. By this measurement, we can easily select the best attribute for the nodes of the tree. There are two popular techniques for ASM, which are:

- Information Gain
- Gini Index

1. Information Gain:

Information gain is the measurement of changes in entropy after the segmentation of the dataset based on the attribute. It calculates how much information a feature provides us about a class. According to the information gained, we split the node and build the decision tree. A decision tree algorithm always tries to maximize the value of Information gain, and a node having the highest information gain is split first.

Information gain= Entropy(S)-[(Weighted Avg) Entropy(each feature)]*

Entropy is a metric to measure the impurity in a given attribute.

Entropy(S) = -p(yes) log2 P(yes)- P(no) log2 P(no)
Where,

- *S=Total number of samples*
- *P(yes)=probability of yes*
- *P(no)=probability of no*

2. Gini Index:

Gini index is a measure of impurity or purity used while creating a decision tree in the CART algorithm. An attribute with the low Gini index should be preferred as compared to the high gini index. It only creates binary splits.

Gini Index = 1-∑jPj2

Here in this project we make use of this decision tree algorithm to find the results. Here for the process of Attribute selection measure we make use of the Information gain and the entropy values.

5.5 BINARY NODE

Binary node, since the system is constructed in an object oriented trend, considering all the involvement as objects. Hence, representation of each particular value,classes of the each attribute in form of node is necessary in order to create the decision tree. Here we are taking the information from the calling class and then we try to give that information to the model and the it predict something and we take that as output. We call this module as the binary node because here we do all these process by using the binary node decision tree. And finally we print this output to the user via user interaction page (GUI).

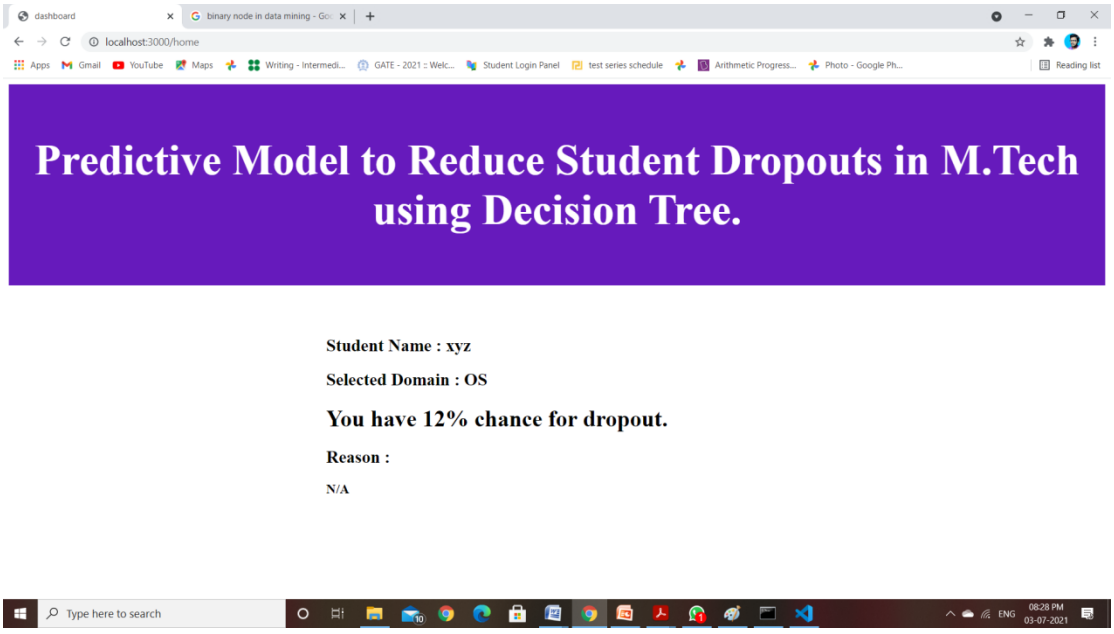


Fig. 5.5: Results Page.

5.6 CONCLUSION

The implementation of the system is explained in Chapter 5. This chapter describes the implementation of the system. We have discussed the proposed methodology and steps of implementation. This chapter consists of the proposed approach and various steps required implementing the proposed approach. This chapter gives a brief view about the implementation of a system using various steps and proposed approaches. Here, we provided a detailed explanation of the steps in the proposed approach and how they are implemented. We have discussed various techniques, methods and equations for each step of the implementation. We gave a brief description about each step in the process, its purpose and also the packages, equations, diagrams required for the implementation in each step clearly. The next chapter discussed types of testing and test cases of the project.

CHAPTER-6

TESTING

PREAMBLE

This chapter completely deals with testing the model. By analyzing all the previous chapters' clear explanation about optical character recognition, problem statement, aim of the project, related reference papers, system requirements, system design, methodology and modules related to the project. Whereas in this chapter it completely deals with testing phase related to model.

6.1 SOFTWARE TESTING

Testing is the process of evaluating a system or its component(s) with the intent to find whether it satisfies the specified requirements or not. Testing is executing a system in order to identify any gaps, errors, or missing requirements in contrary to the actual requirements. Software testing is the process used to assess the quality of computer Software. Software testing is an empirical technical investigation conducted to provide stakeholder with information about the quality of the product or service under test, with respect to the context in which it is intended to operate. This Includes, but is not limited to, the process of executing a program or application with the intent of finding software bugs. Quality is not an absolute; it is value to some person. What that in mind, testing furnishes a criticism or comparison that compares the state and behavior of the product against specification. An Important point is that compares the state and behavior of the product against a Specification. An important point is that software testing should be distinguished from the separate discipline of Software Quality Assurance, which encompasses all business process areas, not just testing.

6.1.1 Verification

Verification is the process of evaluating work-products of a development phase to determine whether they meet the specified requirements. Verification ensures that the product is built according to the requirements and design specifications. It also answers to the question, are we building the product, right?

Verification Testing

Verification testing can be best demonstrated using V-Model. The artifacts such as test Plans, requirement specification, design, code and test cases are evaluated.

6.1.2 Validation

The process of evaluating software during the development process or at the end of the development process to determine whether it satisfies specified business requirements. Validation Testing ensures that the product actually meets the client needs. It can also be defined as to demonstrate that the product fulfils its intended use when deployed on appropriate environment. It answers to the question, Are we building the right product?

6.2 TYPES OF TESTING

Software testing methods are traditionally divided into black box testing and white box testing. These two approaches are used to describe the point of view that a test engineer takes when designing test cases.

Black box testing

Black box testing is a testing approach which is used to test the software without the knowledge of the internal structure of program or application. It is also known as data driven box testing, data and functional testing. Testing is based on external expectations; internal behavior of the application is unknown. This type of testing is ideal for higher levels of testing like System Testing, Acceptance testing. Programming knowledge is not needed to perform Black Box testing. Implementation knowledge is not requiring doing Black Box testing. The main objective of this testing is to check what functionality of the system under test.

White box testing

White box testing is used when the tester has access to the internal data structures, code and algorithms. White box testing methods include creating tests to cause all statements in the program to be executed at least once. Other examples of white box testing are mutation testing and fault injection methods. The main objective of White Box testing is done to check the quality of the code.

Regression testing

After modifying software, either for a change in functionality or to fix defects, a regression test re-runs previously passing tests on the modified software to ensure that the modifications

6.3 TEST CASES

Test Case 1:

Input: Provide data that has high income and high domain marks.

Output: Processing the given data and predicts the dropout chance.

Validation: Accept.

Test Case 2:

Input: Provide data that has low income and high domain marks.

Output: Processing the given data and predicts the dropout chance.

Validation: Accept.

Test Case 3:

Input: Provide data that has low income and low domain marks

Output: Processing the given data and predicts the dropout chance.

Validation: Accept

Test Case 4:

Input: Provide data that has high income and low domain marks

Output: Processing the given data and predicts the dropout chance.

Validation: Accept

Test Case 5:

Input: Provide data of Low income and domain marks greater than 100.

Output: cannot process it due to invalid details of data.

Validation: Not Validated

6.4 CONCLUSION

This chapter gives a detailed explanation of testing techniques related to the project. The technique we have used is one of the most important techniques used in Data mining to test the model i.e testing model performance. The next chapter is related to results and comparison of the proposed approach and results with the existing methods and results. It also give comparison chart graphs and tables related to the comparison. The next chapter gives the explanation of the experiment.

Chapter 7

EXPERIMENTAL RESULTS

7.1 Results

Here in the system trained with the data. When we enter the student name, interested domain, marks got on that domain in the graduation and student's income then when student click on predict button then it is going to call the decision tree algorithm on the background. That algorithm is already trained with real time data. Then the algorithm will take the values and processing the input and produce the output. Output is the chance of student dropout percentage.

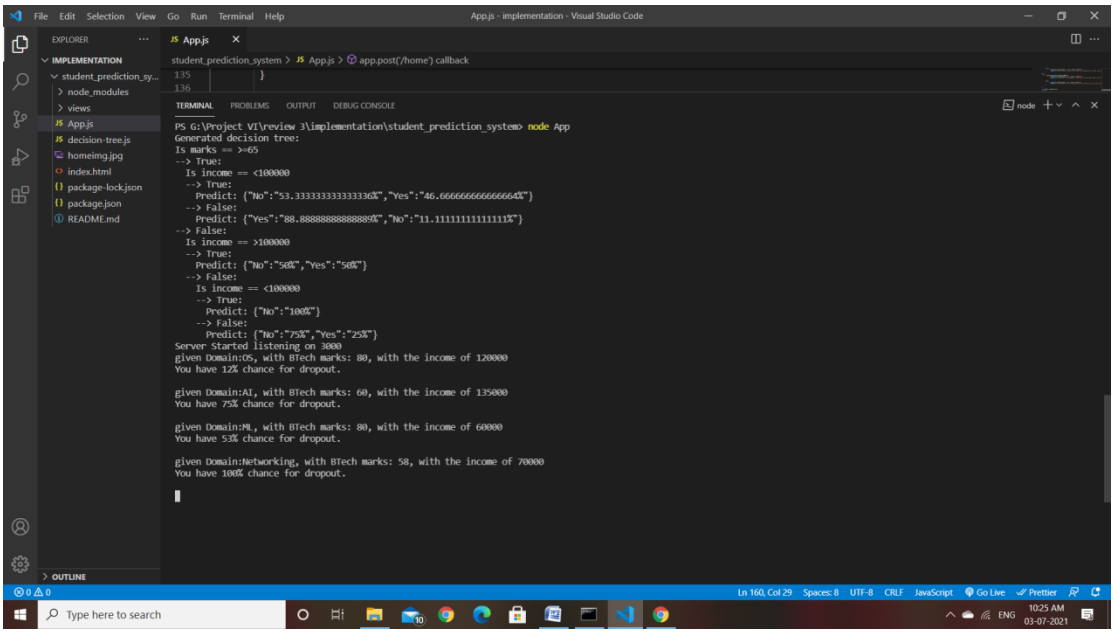


Fig. 7.1: Output of the proposed model.

7.2 Result Comparison

Here we compare our results with the existing system results. When we test our with test data then we got 85% accuracy. In the existing system, it is having the 74% to 77% of accuracy. So here we improved the accuracy of the model.

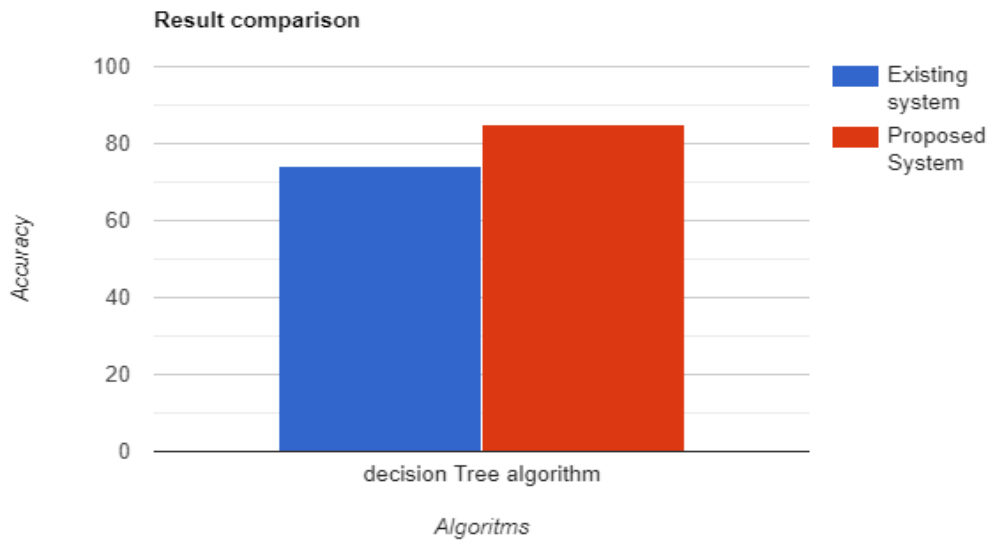


Fig. 7.2: Result comparison.

7.3 Applications

Our system is used to reduce the student dropouts from the higher education. Students get to know that, he can complete his academics after joining into the higher education or not. This is the main application of this system.

CONCLUSION AND FUTURE ENHANCEMENT

8.1 Conclusion

The Obtaining results suggest that the usage of the decision tree algorithm to predict the student drop outs is gives more accuracy. The performance is found to be good in decision tree. In this system we are taking four parameters form the student, among those four parameters we give two parameters to the decision tree algorithm. And our algorithm works fine when we provide those two parameters.

Here first we start our system, and then it is go to the user interaction page. And in the background the decision tree algorithm is trained with the training data. Simultaneously we can collect the data from the student/user from the graphical user interface. Then we pass these data to processing. After processing the decision tree will provide a result by using its knowledge. After we provide this output to the user along with the reason. Here the system works fine when we provide valid inputs, otherwise there may be a chance of crash of algorithm.

8.2 Future Enhancement

Here e are taking the student graduation marks only to predict the dropout rate. So, here we can also add an examination interface to the system, and based on the examination results and the time log of the examination as the additional parameters to the decision tree algorithm. Then we have a chance of increase in the accuracy. And we can also add the authentication interface to system to provide the security. If we are adding all those things to the system then it is going to be more efficient.

REFERENCES:

[1] Antonio Jesus Fernandez Garcia, Roberto Rodriguez- Echeverria, Juan Carlos Preciado Jose Maria ConejeroManzano, and Fernando Sanchez-Figueroa ." Creating a Recommender System to Support Higher Education Students in the Subject Enrollment Decision", IEEE Journal, Vol.08, pp.189069-189088, 2020.

[2] Shaojie qu1, Kan li2, Shuhui Zhang1, And Yongchao Wang3." Predicting Achievement of Students in Smart Campus", IEEE, Journal, Vol.06, pp:60264-60273, 2018.

[3] NafisaTasnim, Mahit Kumar Paul, "Identification of Drop Out Students Using Educational Data Mining", IEEE, Conference, 2019.

[4] RaminGhorbani and RouzbehGhousi, "Comparing Different Resampling Methods in Predicting Students' Performance Using Machine Learning Techniques", IEEE, Journal, Vol.08,pp:67899-67911, 2020.

[5] DanialHooshyar 1, Yeongwook Yang 1, Margus Pedaste1, And Yueh-Min Huang 2, "Clustering Algorithms in an Educational Context: An Automatic Comparative Approach", IEEE, Journal, Vol.08,pp:146994-147014, 2020.

[6]David Baners, M. Elena Rodriguez and Montse Serra, "An Early Feedback Prediction System for learners At-risk within a First-year Higher Education Course", IEEE Transactions, vol:12, pp:249-263, 2019.

[7] Josep Figueroa- Canas and Teresa Sancho-Vinuesa, "Early prediction of dropout and final exam performance in an online statistics course", IEEE Transactions, vol: 15, pp: 86-94, 2020.

[8] Muhammad Adnan, AsabHabib, Jawad Ashraf, ShafaqMuaasdiq, Arsalan Ali Raza, Muhammad Abib, Maryam Bashir, and Sana Ullah Khan, "Predicting At-Risk Students at Different Percentages of Course Length for Early Intervention Using Machine Learning Models", IEEE Access, vol:9, pp: 7519-7539, 2021.

[9] Lin Qiu, Yanshen Liu, and Yi Liu, "An Integrated Framework with Feature Selection for Dropout prediction in Massive Open Online Courses", IEEE Access, vol:6, pp:71474-71484, 2018.

[10] Yuri Nieto, Vicente Gacia-Diaz, Carlos Montenegro, Claudio Camilo Gonzalez, and Ruben Gonzalez Crespo, “Usage of Machine Learning for Strategic Decision Making at Higher Educational Institutions”, IEEE Access, vol:7, pp: 75007-75017, 2019.

[11] Camilo Ernesto Lopez Guarian, Elizabeth Leon Guzman, Fabio A.Gonzalez." A Model to Predict Low Academic Performance are a Specific Enrollment Using Data mIning”, IEEE REVISTA, Vol:10,pp:119-125,2015.

[12] Amin Zollanvari, RefilkCaglarKizilirmak, YauHeeKho, DanielHernandez-Torrano," Predicting Student GPA and Developing Intervention Strategies Based on Self-Regulatory Learning Behaviours ", IEEE ACCESS, VOL:5, PP:23792-23802,2017.

[13] Abdullah Alshanqiti, and Abdallahnamoun," Predicting Student Performance and Its Influential Factors Using Hybrid Regression and Multi-Label Classification", IEEE ACCESS, Vol-8,pp:203827-203844,2020.

[14] Jui-long Hung, Brett E.Shelton, Juan Yang and Xu Du, "Improving Predictive Modelling For At-Risk Student Identification. A multi-stage Approach", IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES, Vol:12,pp:148-157,2019.

[15] Hanna Abdullah Mengash, "Using Data Mining Techniques to Predict Student Performance to Support Decision Making in University Admission Systems", IEEE ACCESS, Vol:8,pp:55462-55470,2020.

[16] Mustafa Agaoglu, “Predicting Instructor Performance Using Data Mining Techniques in Higher Education”. IEEE Journal, Vol:04, pp.2379-2387,2016.

[17] Miguel Angel Prada, Manuel Dominguez, Jose Lopez Vicario, Paulo AlexandreVaraAlves, Marian Barbu, Michal Podpora, Umberto Spagnolini, Maria J. VarandaPereire, Ramon Vilanova. “Educational Data Mining for Tutorial Support in Higher Education: A Web-Based Tool Case study in Engineering Degrees”. IEEE journal, Vol:08, pp.212818-212836, 2020.

[18] Jui-Long Hung, Morgan C Wang, “Identify At-Risk students for Early interventions-A Time Series Clustering Approach”. IEEE, 2017.

[19] Rosa M. Carro, AlvaroOrtigosa, “From Lab to Production: Lessons Learnt and Real Life Challenges of an Early Student-Dropout Prevention System” IEEE, 2019.

[20] Ye Tian, Shangshang Yang, Lei Zhang, FunchenDuan, Xingyi Zhang. "A Surrogate-Assisted Multi-objective Evolutionary Algorithm for Large-Scale Task-Oriented Pattern Mining". IEEE journal, Vol:03,2019.

Appendix A

SOURCE CODE

Decision tree implementation:

```
classDecisionTree {
constructor(classPropertyName, featurePropertyNames, trainingData) {
    this._classPropertyName = classPropertyName;
    this._featurePropertyNames = featurePropertyNames;
this.decisionTree = this._buildTree(trainingData);
  }
  _buildTree(rows) {
const split = this._findBestSplit(rows);
if (split.bestGain === 0) {
return new Leaf(this._classCounts(rows));
    }
const partition = this._partition(rows, split.bestQuestion);
consttrueBranch = this._buildTree(partition.trueRows);
constfalseBranch = this._buildTree(partition.falseRows);
return new DecisionNode(split.bestQuestion, trueBranch, falseBranch);
  }
class that occurred in the row
```

```
  _classCounts(rows) {
constclassCounts = {};
constclassPropertyName = this._classPropertyName;
rows.forEach(row => {
const classifier = row[classPropertyName];
if (!classCounts[classifier]) {
classCounts[classifier] = 0;
      }
classCounts[classifier] += 1;
    });
returnclassCounts;
  }
  _partition(rows, question) {
const [trueRows, falseRows] = [[], []];
rows.forEach(row => {
if (question.match(row)) {
trueRows.push(row);
      } else {
falseRows.push(row);
      }
```

```
    });
return {trueRows, falseRows};
  }
  _gini(rows) {
constclassCounts = this._classCounts(rows);
let impurity = 1;
Object.keys(classCounts).forEach(key => {
constclassProbability = classCounts[key] / parseFloat(rows.length);
impurity -= Math.pow(classProbability, 2);
    });
return impurity;
  }
  _infoGain(left, right, current_uncertainty) {
const weight = left.length / (left.length + right.length);
return (
current_uncertainty -
weight * this._gini(left) -
      (1 - weight) * this._gini(right)
    );
  }
```

```
 _findBestSplit(rows) {
letbestGain = 0;
letbestQuestion = null;
constcurrent_uncertainty = this._gini(rows);


   this._featurePropertyNames.forEach(column => {
constuniqueValues = new Set();
     // get all unique values in the column
rows.forEach(row => {
uniqueValues.add(row[column]);
     });
for (let value of uniqueValues) {
const question = new Question(column, value);
       // split dataset
const partition = this._partition(rows, question);
       // skip the split if it does not divide the dataset
if (!partition.trueRows.length || !partition.falseRows.length) {
continue;
       }
```

```
        // information gain from split
const gain = this._infoGain(
partition.trueRows,
partition.falseRows,
current_uncertainty,
        );
if (gain >= bestGain) {
          [bestGain, bestQuestion] = [gain, question];
        }
      }
    });
return {bestGain, bestQuestion};
  }
  _printTree(node, indent = '') {
if (node instanceof Leaf) {
console.log(
        `${indent}Predict:
${JSON.stringify(node.percentageProbabilities())}`,
      );
return;
    }
```

```
console.log(`${indent}${node.question.toString()}`);
console.log(`${indent}--> True:`);
    this._printTree(node.trueBranch, indent + '  ');
console.log(`${indent}--> False:`);
    this._printTree(node.falseBranch, indent + '  ');
  }

  _treeToStringArray(node, stringArray, indent = '') {
if (node instanceof Leaf) {
stringArray.push(
        `${indent}Predict:
${JSON.stringify(node.percentageProbabilities())}`,
      );
returnstringArray;
    }

    // print the question at this node
stringArray.push(`${indent}${node.question.toString()}`);
stringArray.push(`${indent}--> True:`);
    this._treeToStringArray(node.trueBranch, stringArray, indent + '      ');
stringArray.push(`${indent}--> False:`);
```

```
    this._treeToStringArray(node.falseBranch, stringArray, indent + '      ');
  }
  _classify(row, node) {
if (node instanceof Leaf) {
returnnode.percentageProbabilities();
    }
if (node.question.match(row)) {
returnthis._classify(row, node.trueBranch);
    } else {
returnthis._classify(row, node.falseBranch);
    }
  }
classify(row) {
returnthis._classify(row, this.decisionTree);
  }
printTree() {
    this._printTree(this.decisionTree);
  }
toString() {
consttreeStringArray = [];
```

```
    this._treeToStringArray(this.decisionTree, treeStringArray);
returntreeStringArray.reduce((acc, string) => {
return `${acc} \n ${string}`;
    }, '');
  }
}

constructor(column, value) {
this.column = column;
this.value = value;
  }
match(example) {
constexampleValue = example[this.column];
if (isNumeric(exampleValue)) {
returnexampleValue>= this.value;
    } else {
returnexampleValue === this.value;
    }
  }
toString() {
```

```
let condition = '==';
if (isNumeric(this.value)) {
condition = '>=';
    }
return `Is ${this.column} ${condition} ${this.value}`;
  }
}
class Leaf {
constructor(classCounts) {
this.predictions = classCounts;
  }
percentageProbabilities() {
const total = Object.keys(this.predictions).reduce((acc, key) => {
returnacc + this.predictions[key];
    }, 0);
const probabilities = {};
Object.keys(this.predictions).forEach(key => {
probabilities[key] = `${(this.predictions[key] / total) * 100}%`;
    });
return probabilities;
```

```
  }
}
classDecisionNode {
constructor(question, trueBranch, falseBranch) {
this.question = question;
this.trueBranch = trueBranch;



this.falseBranch = falseBranch;
  }
}
functionisNumeric(value) {
returntypeof value === 'number';
}
module.exports = DecisionTree;
```

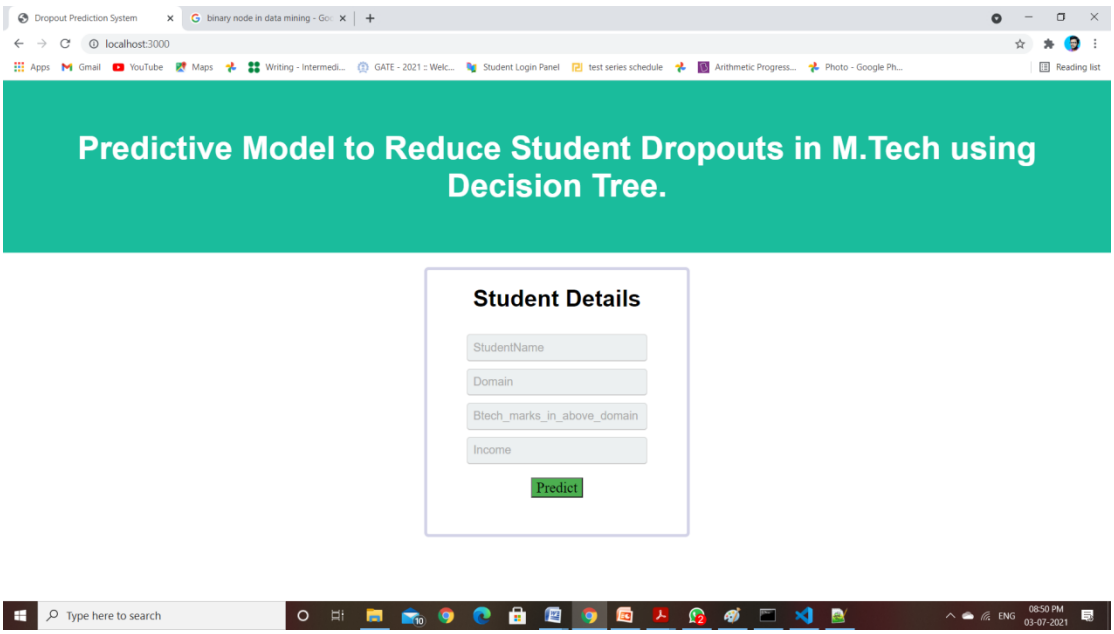
Appendix B

SCREENSHOTS

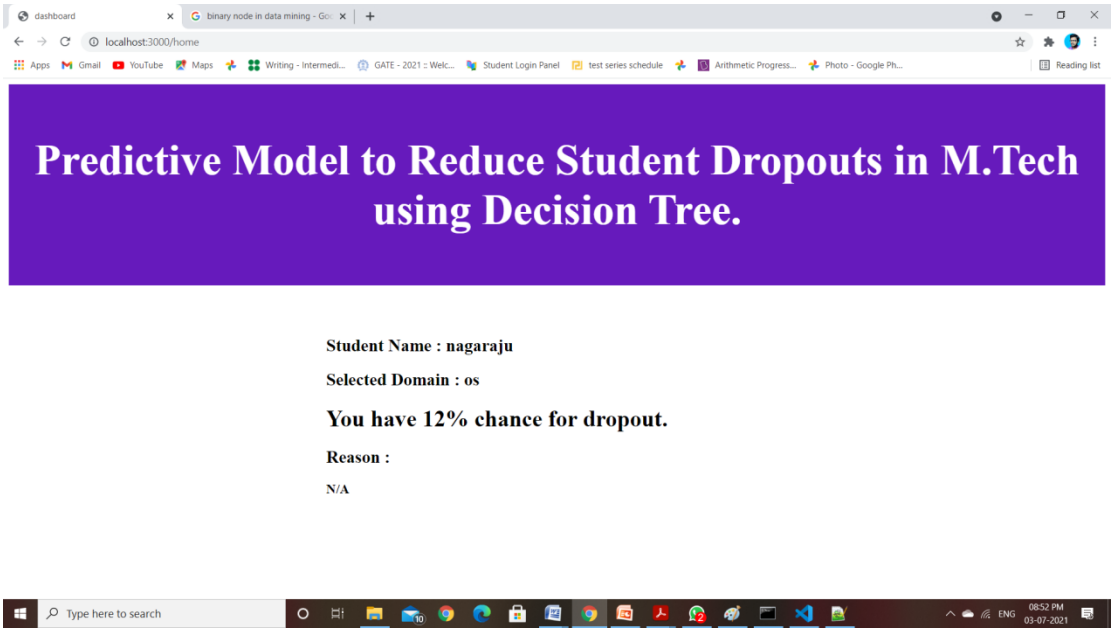
Here we trained the model by giving our dataset to the algorithm.

```
PS G:\Project VI\review 3\implementation\student_prediction_system> Node App
Generated decision tree:
Is marks == >=65
--> True:
  Is income == <100000
  --> True:
    Predict: {"No":"53.333333333333336%","Yes":"46.666666666666664%"}
  --> False:
    Predict: {"Yes":"88.88888888888889%","No":"11.11111111111111%"}
--> False:
  Is income == >100000
  --> True:
    Predict: {"No":"50%","Yes":"50%"}
  --> False:
    Is income == <100000
    --> True:
      Predict: {"No":"100%"}
    --> False:
      Predict: {"No":"75%","Yes":"25%"}
Server Started listening on 3000
```

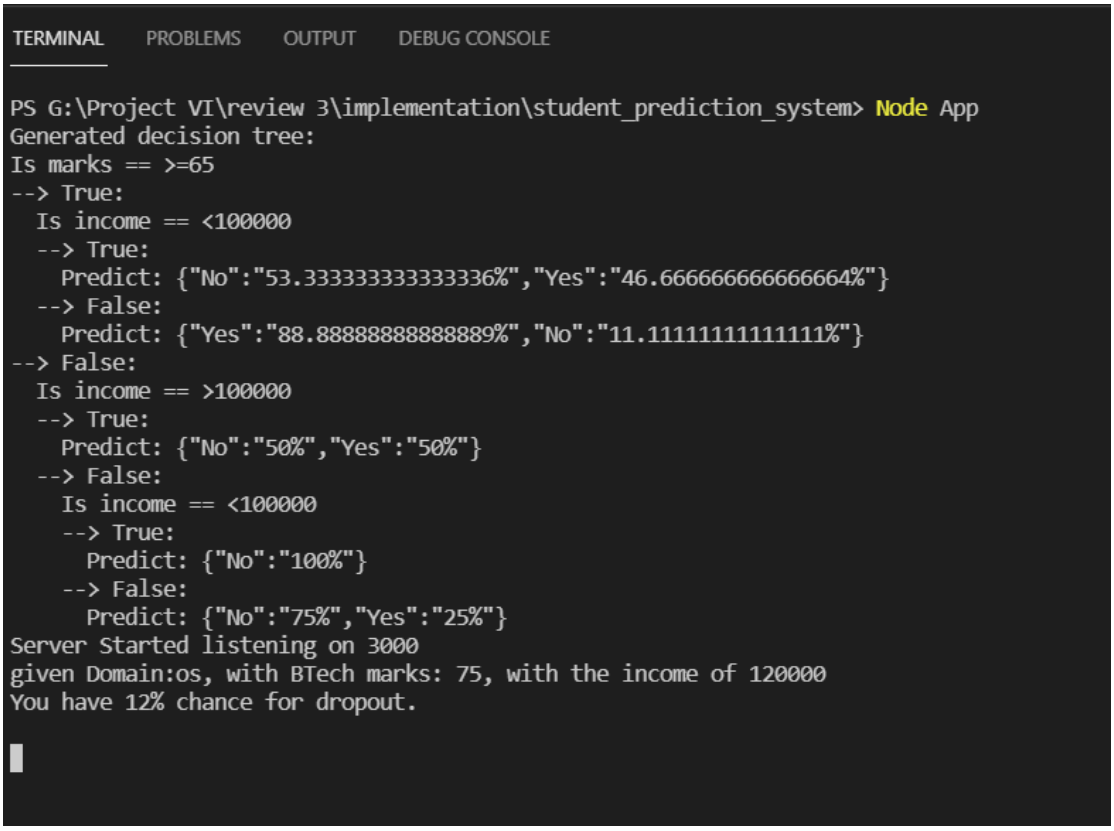
TRANING MODEL



DATA COLLECTING FROM THE USER



OUTPUT DASHBOARD



RESULTS ON CONSOL

Appendix C

STUDENT CONTRIBUTION

S. No	Topic		Project Associates			
			M. Akhileswari	S. Nagaraju	K. SaiBhargavi	Y. Navya
			17K61A0565	17K61A0596	17K61A0559	17K61A05B5
1	Literature survey		✓	✓	✓	✓
2	Problem Formulation				✓	✓
3	Requirements identification		✓	✓		
4	Design	Data flow	✓	✓	✓	
		System architecture		✓		✓
		Use Case Diagram	✓		✓	✓
5	Implementation		✓	✓	✓	✓
6	Testing		✓	✓	✓	✓
7	Documentation		✓	✓	✓	✓

APPENDIX D

POs, PSOs, PEOs and Cos Relevance with Project

PROGRAM OUTCOMES (POs):

POs	Program Outcome	Relevance
PO1	**Engineering Knowledge:** Apply knowledge of mathematics, science, engineering fundamentals and an engineering specialization to the solution of complex engineering problems.	With the use of engineering knowledge of React Js Language, and Data Mining techniques.
PO2	**Problem Analysis:** Identify, formulate, research literature and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences and engineering sciences.	In Literature review specified in(Chapter 2) identified problems in existing system and we came with a solution (refer to appendix-B for screenshots)
PO3	**Design/ Development of Solutions:** Design solutions for complex engineering problems and design system components or processes that meet specified needs with appropriate consideration for public health and safety, cultural, societal and environmental considerations.	Architectural diagram, use case diagram and data flow diagram (refer to chapter 4) is used for designing the proposed system. Refer sample code (refer appendix- A), complete implemented source code is available.
PO4	**Conduct investigations of complex problems** using research based knowledge and research methods including design of experiments, analysis and interpretation of data and synthesis of information to provide valid conclusions.	We have conducted a specificresearch including analysis of data and identified the problems and provided information to provide conclusions (refer chapter- 2).
PO5	**Modern Tool Usage:**Create, select and apply appropriate techniques, resources and modern engineering and IT tools including prediction and modelling to complex engineering activities with an under- standing of the limitations.	We have taken IT tools like Node js and Express js (refer to chapter 3) for implementation purpose.

PO6	**The Engineer and Society:** Apply reasoning informed by contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to professional engineering practice.	The professional engineering practice helped to produce efficient and accurate results to improve effectiveness in the system (refer to chapter 5).
PO7	**Environment and Sustainability:** Understand the impact of professional engineering solutions in societal and environmental contexts and demonstrate knowledge of and need for sustainable development.	By providing accurate result leads to sustainable development and future enhancement.
PO8	**Ethics:**Apply ethical principles and commit to professional ethics and responsibilities and norms of engineering practice.	To find the accuracy rate of the system can be proposed in this system (refer chapter 5).
PO9	**Individual and Team Work:**Function effectively as an individual, and as a member or leader in diverse teams and in multidisciplinary settings.	Provide the solutions for the activities refer to page no are carried as an individual as well as a team refer to appendix-C.
PO10	**Communication:**Communicate effectively on complex engineering activities with the engineering community and with society at large, such as being able to comprehend and write effective reports and design documentation, make effective presentations and give and receive clear instructions.	Communicated individually aswell as a team regarding project, our communication enhanced a lot in terms of developing a project. By giving presentation infront of evaluators and guides we enhanced our communication skills refer to project department Evolution committee assessment file.
PO11	**Life-long Learning:**Recognize the need for and have the preparation and ability to engage in independent and life- long learning in the broadest context of technological change.	----------

PO12	**Project Management and Finance:**Demonstrate knowledge and understanding of engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.	Demonstrates knowledge and awareness to engineering students(refer to chapter 5).

PROGRAM SPECIFIC OUTCOME (PSOs):

PSO	Program Specific Outcome	Relevance
PSO1	**Mobile and Web Application Development**: Ability to develop mobile and web applications using J2EE, Android and J2ME	Ability to develop online application.

PROGRAM EDUCATIONAL OUTCOMES (PEOs):

PEO	Program Educational Outcome	Relevance
PEO1	Graduates are well trained, confident, research oriented and industry ready professionals who are intellectual, ethical and socially committed	By analysing different techniques and problems related to the project, we designed the system and provided an accurate solution to the problems.
PEO2	Graduates will have the technical, communication skills and character that will prepare them for technical roles	Following the ethical principles while working on project we are ethically and socially committed. We are well trained and intellectual as we worked as a team for our project and enhanced our confidence and research-oriented skills
PEO3	Graduates will be able to analyze, design and	By working on the

	develop advanced computer applications to provide solutions to real world problems	project implementation and presenting in front of evaluators and guides we have enhanced skills required for technical and leadership roles.

COURSE OUTCOME (COs):

CO	Course Outcomes	POs, PEOs, PSOs Mapped
CO1	Develop problem formulation and design skills for engineering and real world problems	PO2, PO3.
CO2	Collect and generate ideas through survey on current research areas which help to analyse and present to impart knowledge in different fields	PO2, PO4, PO10.
CO3	Import knowledge on software and hardware to meet industry perspective needs and standards	PO1, PO6, PO7, PO8.
CO4	Create interest to carry out research on innovative ideas as a lifelong learning	PO4
CO5	Ability to work with team and enrich presentation and communication skills	PO1, PO3, PO9, PO10
CO6	Create a platform that everyone make use of it	PO3, PO7, PO8, PO9, PO10.