International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 02 | Feb 2021 www.irjet.net p-ISSN: 2395-0072

Prediction of Suitable Career for Students using Machine Learning

Samruddhi Deshpande[1], Pooja Gupta[2], Nikita Singh[3], Deepali Kadam[4]

[1]BE student, Information Technology, Datta Meghe College of Engineering, Navi Mumbai, India
[2]BE student, Information Technology, Datta Meghe College of Engineering, Navi Mumbai, India
[3]BE student, Information Technology, Datta Meghe College of Engineering, Navi Mumbai, India
[4]Asst. Professor, Information Technology, Datta Meghe College of Engineering, Navi Mumbai, India

---

Abstract - *This project is aimed at helping students to choose a suitable career by taking into consideration a student's mental, physical, psychological aspects and academic performances and some other important aspects of the student's life. Students can test their knowledge in their field of interest by taking a quiz. This system is an android as well as a web application and it will be user-friendly. This system will be beneficial for students and colleges. This system will help a lot to students as most of the times students tend to confuse while choosing a career. Students are unaware of their strengths and weaknesses and this results in choosing a random career. To avoid this, the system will analyse students by considering important aspects and results of the quiz and will predict the results using Machine Learning accordingly.*

Key Words: **Prediction of Career, Performance, Quiz, Confusion, Strengths, Weaknesses, Machine Learning.**

1.INTRODUCTION

Nowadays, students are very much confused while choosing a career. While choosing a career a lot of factors have to be taken into consideration. One has to self-analyse himself/herself while choosing a career. The students fail to know their strengths and choose their career randomly which leads to frustration and demoralization. Hence, this system will help in choosing suitable careers for the students. This system is an android as well as web application which will provide facilities to students to check what career is suitable for them. All these things will be predicted by the system by taking all the required data from students. This system will help a lot of students as it will focus on suitable career prediction for students and provide quizzes in different fields of career to students. Using this one system students will be able to know their strengths and suitable career for them and accordingly students can plan their schedules in order to achieve their goal and for their betterment.

1.1 Proposed System

The proposed system predicts the suitable career for students of 10th and 12th standard by taking related data as an input and also gives an option to solve quiz in the field of interest of student so that student can check his/her knowledge in his/her field of interest. It is android as well as web application which will allow students to use the system on desktop as well as mobiles. The prediction will be done by Naïve Bayes Algorithm and a dataset containing different attributes like personal details, curriculum details, background details, interests, etc. will be considered while predicting suitable career for students. This application will help students by giving proper guidance about their career according to their interest and skills.

1.2 Naïve Bayes Algorithm

Naïve Bayes Algorithm is a probabilistic classification technique based on Bayes' Theorem. Bayes' theorem is also known as Bayes' Rule or Bayes' law, which is used to determine the probability of a hypothesis with prior knowledge. It depends on the conditional probability. The formula for Bayes' theorem is given as: P(A|B) = P (B|A) * P(A)/P(B), where, P(A|B) is Posterior probability**:** Probability of hypothesis A on the observed event B. P(B|A) is Likelihood probability**:** Probability of the evidence given that the probability of a hypothesis is true. P(A) is Prior Probability**:** Probability of hypothesis before observing the evidence. P(B) is Marginal Probability**:** Probability of Evidence. It is a supervised learning algorithm. It is not a single algorithm but a family of algorithms where all of them share a common principle, i.e., every pair of features being classified is independent of each other. The steps to implement Naïve Bayes algorithm are: 1. Data Pre-processing 2. Fitting Naïve Bayes to the training data set 3. Predicting the test result 4. Testing accuracy 5. Visualizing the test result. The prediction was done using different algorithms like Decision Tree, SVM, Naïve Bayes Algorithm. Naïve Bayes algorithm performed well amongst all the algorithms and gave highest accuracy among all. As the Naïve Bayes Algorithm considers every feature to contribute independently to the probability of getting a prediction and it requires all the features to be equally important, it is the suitable choice for this problem statement.

2. Approach

The system considers all the important aspects of students' life by considering their physical and mental conditions, academic performance, family background, family support, interests, talent, etc. It takes all these parameters as input from the student and applies the Machine Learning model and predicts suitable career based on this information. The system also intends to allow students to test their knowledge in their field of interest and provides option of solving quizzes in various fields of career and also, allows students to check their Emotional Quotient and Social Quotient. It shows appropriate message based on the marks obtained by the students in the quiz. A dataset containing features like physical and mental conditions, academic performance, family background, family support, interests, talent, etc. is used for training the model. Firstly, three models, using SVM, Decision Tree and Naïve Bayes Algorithm respectively, were trained. The first step of implementation was data collection. After collecting the data set, the next step that was implemented was data pre-processing. In data pre-processing, the data set was made ready for training by removing duplicates, correcting errors, dealing with missing values, normalizing the data set, making the required data type conversions. LabelEncoder was used to normalize the labels and to convert non-numerical labels into numerical labels. The aim of pre-processing is to get clean and processed data which will give maximum accurate outputs. After pre-processing, the data set was split into training data set and testing data set. The training data set was 70% of the total data set and testing data set was 30% of the total data set. The next process after splitting the data set was to choose a model and train the model. In training, the data set is fitted into the classifier and predictions are made. The libraries which were used while training the model were csv, pandas, sklearn, joblib, etc. After getting the predictions, accuracy given by SVM, Decision Tree, Naïve Bayes Algorithm was calculated. The maximum accuracy was given by Naïve Bayes Algorithm. Hence, the system is developed by training the model using Naïve Bayes Algorithm.

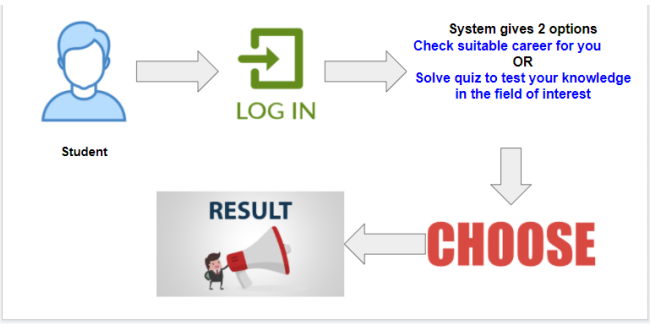

Fig – 1: Approach Diagram

Student signs up/in and after that students are asked the standard in which he/she studied and passed recently. After choosing the standard, students are provided with two options, one is to check suitable career and another is to solve quiz. After choosing the option to check suitable career students will be asked to fill a form which will ask about their health, academic performance, family background, interests, and questions related to different fields of career and on answering those questions, a suitable career will be predicted. If student chooses to solve a quiz, he/she will be shown his/her score in the quiz and an appropriate message will be displayed based on the score obtained by the student.

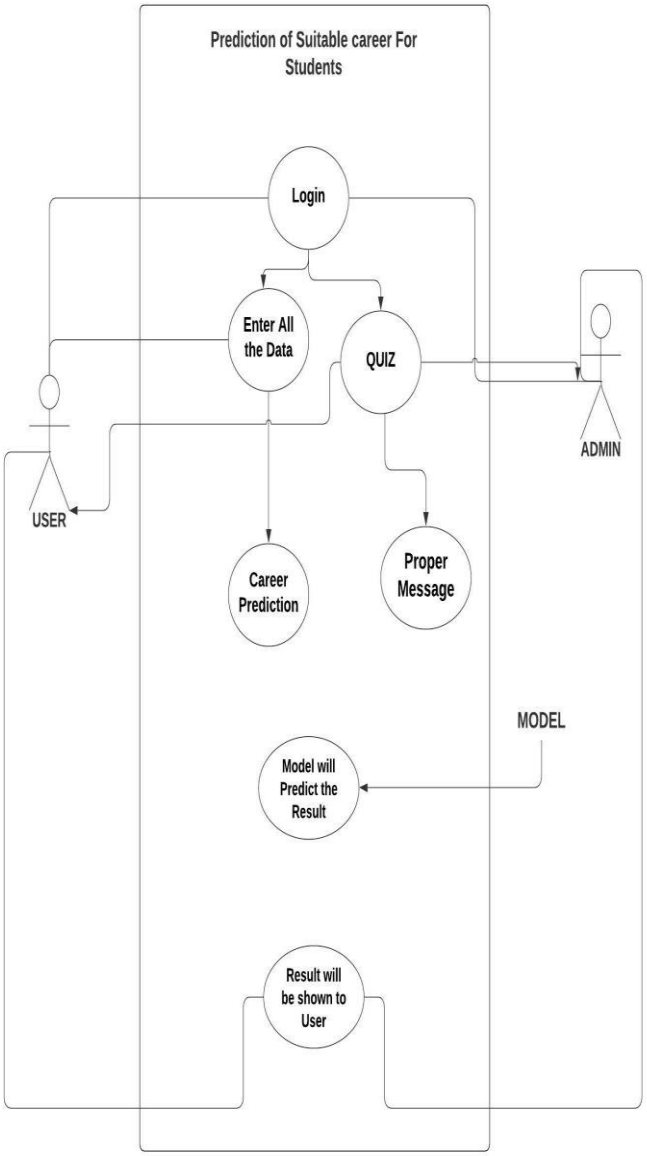

Fig – 2: Use Case Diagram

3. CONCLUSIONS

We have developed a website and mobile application which will help 10th and 12th grade students to know their suitable career according to their strengths and weaknesses. The other module helps the students to take a quiz in their field of interest and displays the score of the quiz which helps the students to analyze their knowledge in that field. The features used in this model are helpful in predicting the suitable career for students. This system will help students in choosing their career and reduce confusion among the students while choosing a career. In future, more powerful system can be developed where maximum fields of career can be covered and technical, logical, memory-based, skill-based tests in the maximum fields of career can be provided by the system in order to give more accurate and beneficial results.

ACKNOWLEDGEMENT

Working on the project on "Prediction of Suitable Career for Students using Machine Learning" was a source of immense knowledge to us. However, this would not have been possible without the equal support of all the members. We would like to express our sincere gratitude to Asst. Prof. Deepali Kadam for her guidance and valuable support. We would like to extend a sincere thanks to her.

REFERENCES

[1] Rucha Hemant Rangnekar, Khyati Pradeep Suratwala, Sanjana Krishna, Sudhir Dhage, "Career Prediction Model Using Data Mining And Linear Classification", Department of Computer Engineering, Sardar Patel Institute of Technology Mumbai, India, 2018.

[2] Md. Yeasin Arafath, Mohd. Saifuzzaman, Sumaiya Ahmed, Syed Akhter Hossain, "Predicting Career Using Data Mining", Daffodil International University Dhaka, Bangladesh, 2018.

[3] Ahmad Slim, Gregory L. Heileman, Jarred Kozlick and Chaouki T. Abdallah, "Predicting Student Success Based on Prior Performance", Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, 87131, USA, 2014.

[4] Pratiyush Guleria, Niveditta Thakur, Manu Sood, "Predicting Student Performance Using Decision Tree Classifiers and Information Gain ", Department of Computer Science, Himachal Pradesh University, Shimla Himachal Pradesh, India, 2014.R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press, 2014.

[5] Fergie Joanda Kaunang, Reymon Rotikan, "Students' Academic Performance Prediction using Data Mining ", Computer Science Department, Universitas Klabat Minahasa, Utara, Indonesia, 2018.

[6] John M. Mativo, Shaobo Huang, "Prediction of students' academic performance: Adapt a methodology of predictive modeling for a small sample size", Career and Information Studies & College of Engineering, University of Georgia, Athens, Georgia, USA, 2014.

[7] Krina Parmar, Dineshkumar Vaghela, Priyanka Sharma, "Performance prediction of students using distributed Data mining", Parul Institute Of Technology, Vadodara, India, 2015.

[8] Tismy Devasia, Vinushree T P, Vinayak Hegde. "Prediction of students performance using Educational Data Mining". Department of Computer Science. Amrita Vishwa Vidyapeetham University, Mysuru Campus, Karnataka, India, 2016.

[9] Lokesh S Katore, Bhakti S Ratnaparkhi, Jayant S Umale, "Novel professional career prediction and recommendation method for individual through analytics on personal traits using C4.5 algorithm". Dept of Computer Engineering. Pimpri Chinchwad College of, Engineering Pune, India, 2015.

[10] Hanife GOKER, Halil Ibrahim BULBUL, Erdl IRMAK, "The Estimation of student's Academic Success by Data Mining Methods", The Institute of Information Gazi University, Ankara, TURKEY, 2013.

[11] Roshani Ade, P. R. Deshmukh, "An incremental ensemble of classifiers as a technique for prediction of student's career choice", Dept. of Computer Science and Engg, Sipna College of Engineering and Technology, Amravati University, Maharashtra, India, 2014.

[12] Utomo Puiianto,Erwina Nurul Azizah,Ayuningtyas Suci Damayanti, "Naive Bayes using to predict students' academic performance at faculty of literature". Electrical Engineering Department. Universitas Negeri Malang, Malang, East Java, Indonesia, 2017.

BIOGRAPHIES

Samruddhi Deshpande
is final year student, Information Technology, Datta Meghe College of Engineering, Navi Mumbai, India

Pooja Gupta
is final year student, Information Technology, Datta Meghe College of Engineering, Navi Mumbai, India

Nikita Singh
is final year student, Information Technology, Datta Meghe College of Engineering, Navi Mumbai, India

Deepali Kadam
is Asst. Professor, Information Technology, Datta Meghe College of Engineering, Navi Mumbai, India