

Career Prediction using Support Vector Machine Algorithm

Steffi F Andrade, Nishita M Bhatankar, Nadia M D'silva and Prof. Prachi Patil
Department of Information Technology Engineering,
Fr.Conceicao Rodrigues College of Engineering, Bandra

Abstract- Career is one of the important decisions of our life. We all face great challenges in choosing a right career path considering the impact of a right career in our life. It is also essential to select a career based on our best abilities in order to achieve success and contentment. Many systems have been developed over the past years to help students select right career paths by predicting the best career option based on his/her academic factors. Academic factors however is not the only suitable factor, we also need to consider one's cognitive abilities and psychometric factors like speed, learning ability, patience and memory to provide best career results. Therefore, this paper focuses on a system that considers both academic as well as psychometric factors to predict one's career.

Index terms- Career, prediction, classification, ML algorithms, SVM.

I. INTRODUCTION

Machine learning advancements have introduced several systems to solve various complex problems. Career has been a stressful and one of the most important decisions of an individual's life for years. Machine learning technology makes career related decisions easier by developing systems that predict career on the basis of academic and technical abilities of an individual. However, cognitive abilities and psychological abilities should be taken into consideration too. Hence, there is a need for a new system in the market that considers academic as well as psychometric abilities to predict career efficiently. For this purpose, we need to adopt new methods to measure the cognitive abilities of an individual by analyzing the performance of the individual during the course of a game like 8-puzzle game. The system should also be able to answer career related questions. Selection of a classification algorithm with high accuracy rate is also essential to develop a career prediction system providing better results thus, helping individuals make career related decisions wisely.

II. PAPER SUMMARY

The paper [1] on predicting career based on advanced Machine learning algorithms like Support Vector Machine(SVM), Decision tree (Random forest), eXtreme Gradient boosting(XG Boost) compare the classification and prediction outcomes of these

algorithms and determine the accuracy of each. The prediction is based on a well-supported dataset that is divided as training set and testing set. The collected data undergoes pre-processing by performing cleaning, replacing missing values, integrating operations on the raw dataset. The dataset is supervised consisting of attributes based on academic and technical abilities of a student along with a few classes defining the output label. For better results, before applying the training set to the classifier algorithms the data is transformed into numerical format using One-hot encoding. After data transformation the training and test sets are applied to algorithms like SVM, Decision trees and XG boost. The predicted result of the algorithms are compared to determine which algorithm has the best accuracy.

The proposed model [2] is designed to predict a career based on various academic factors of a student. It is also suitable for answering many queries related to career guidance. For example, a student may have certain queries regarding a particular field or the factors he needs to work on for a particular stream, these queries can be effectively answered using the proposed model. There are various levels like primary, secondary and higher level in this model, depending on which the system searches a matching guidance for the student. All the factors are ranged from a minimum to maximum value, which are inputs to this system given by the student. Data mining technique used in this system is Nearest Neighbour (NN). The closest matching score is found from the database using the Euclidean distance formula to suggest a precise career field. Various factors including score in a particular field for a student can also be calculated by comparing his score to the closest matching score in the database.

The main aim [3] is to predict the student's career based on his academic scores and cognitive abilities. 8-puzzle game is used to determine the student's psychological abilities. Psychological factors such as intelligence and planning ability, speed of solving problem, patience and perseverance, learning ability is used to predict the student's career. The student has to play the game several times. Based on the number of times the student plays the game his patience and perseverance is measured. Intelligence of the student is tested by his performance. A relation is established between intelligence score and CGPA of the student. So, in this expert system the student manually enters his CGPA as input and also plays the 8-puzzle game several times. The system determines his psychological

abilities and builds the students cognitive model which helps in predicting the student's career.

The objective [4] is to find an accurate classification algorithm that can be used to predict student's career. For building an expert system for career prediction, first the dataset is collected. Framework is designed which consist of different input parameters, that is passed to the Weka tool. Weka tool is used for pre-processing of the dataset. After pre-processing different classification algorithms are used to classify the dataset into classes. Classification algorithms like FURIA, JRIP, PART, PRISM, ID3, J48 and IB1 are applied on the dataset and their performance is observed. It is observed that the algorithms like ID3, PRISM and PART give 100% accuracy in classifying the dataset into classes. Table is maintained containing details regarding accuracy and errors. Thus, classification can be used to predict a student's career.

The proposed model [5] is used to answer various career related queries using a chatbot. The input can either be in a text format or a voice format. In case of a voice format, it is first changed into a text format and then processed further. The received input or queries are then processed using various modules like the Natural Processing Language (NLP) and Knowledge Base (KB). It then matches the input with the data stored in the database using pattern matching and generates an appropriate output for the query. The output of the query is in audio form, making it more understandable rather than just reading the output.

III. PROPOSED SYSTEM

A) PROBLEM STATEMENT

The systems developed so far are either focusing on factors like academic records and personal interests or only psychological factors. The system will focus on both academic records and psychological factors. Some of the algorithms which we will be using are Euclidean distance, One Hot Encoding, SVM (Support Vector Machine), XG Boost (eXtreme Gradient Boosting). Games like 8 puzzle game, Sudoku are used to determine the psychological factors of the student. The system requests the student to play the 8-puzzle game many times and displays the scores of student. The academic record and psychological factors of the student are compared with the minimum levels required for each career and thus the system predicts the matching career(s) for the student.

The dataset should comprise of academic attributes like scores, certifications, computer knowledge, courses and also psychometric attributes and cognitive abilities like speed of calculation, learning ability, memory, and patience. The different possible games can be Sudoku, tic tac toe and 8 puzzle game.

B) ALGORITHMS

- One Hot encoding- This is a pre-processing method that converts categorical variables into a form (numeric) that could be provided to ML algorithms to do a better job in prediction.
- Support Vector Machine (SVM)- This is a machine learning algorithm that analyses data for classification and regression analysis.
- XG Boost- This is an optimized distributed gradient boosting library designed to be highly efficient, flexible and portable.
- Nearest Neighbour(NN)- This is a machine learning algorithm that measures distance between a pair of samples (p and q). It is based of Euclidean Distance formula.

The process to develop this system includes pre-processing using one hot encoding, classification by SVM, XG boost is used for performance boosting. In order to answer career related queries nearest neighbour technique can be used.

IV. CONCLUSION

We developed a model which considers the psychological factors of the student, apart from only the academic record which also plays a major role while predicting the best career for a student. Though there are many traditional tests like aptitude and reasoning tests, IQ tests etc. that are used to assess the psychological factors of a student, a novel idea is to assess the psychological factors through game playing because playing game is much more interesting for every student instead of attending other general tests. A system is developed that supports 8 puzzle game which is played by the student for specific number of times and the expert system evaluates his performance, assesses the psychological factors, builds the cognitive model of the student and predicts the most suitable careers for the student. A correlation between intelligence score and the student's academic track record is established and this shows that the system is rightly assessing the psychological factors of the student. The number of career(s) can be extended so that the system will be useful for any graduate or a post-graduate student belongs to various disciplines.

REFERENCES

- 1] K. Sripath Roy, K. Roopkaanth, V.Uday Teja, V.Bhavana, J.Priyanka, "Student Career Prediction Using Advanced Machine Learning Techniques", International Journal Of Engineering and Technology, Vol.7, No.2.20, pp.26-29, 2018.

- 2] Gufran Ahmad Ansari, "Career Guidance through Multilevel Expert System Using Data Mining Technique", International Journal of Information Technology and Computer Science(IJITCS), Vol.9, No.8, pp.22-29, 2017.
- 3] V.Chandra Prakash, J.K.R. Sastry, K. Anusha, "Expert System for building cognitive model of a student using 8-puzzle game and for career assessment", International Journal of Engineering and Technology, Vol.7, No.2.27, pp.113-117, 2018.
- 4] M.L.Waghmode, Dr. P.P. Jamsandekar, "Expert system for career selection: A Classifier Model", International Journal Of Advance Research in Computer Science and Management Studies, Vol.4, No.1, pp.174-179, 2016.
- 5] Archana Parab, Siddhesh Palkar, Satish Maurya, Sonal Balpande, "An Intelligent Career Counselling Bot", International Research Journal Of Engineering And Technology(IRJET), Vol.4,No.3, pp.56-72, 2017.
- 6] K. B. Eashwar, R. Venkatesan, "Student Performance Prediction Using SVM", International Journal of Mechanical Engineering and Technology (IJMET), Vol 8, No. 11, pp. 649-662, 2017.

