TEST Engineering & Management

January-February 2020
ISSN: 0193-4120 Page No. 10372 - 10375

Data Mining Techniques for Suggesting Student Career

[1]A. Abdul Latheef Nishar, [2]P. Sri Ramya
[1]UG Scholar, [2]Associate Professor (SG),
Department of Computer Science and Engineering, Saveetha School of Engineering,
Saveetha Institute of Medical and Technical Sciences, Chennai
[1]Abdulnizar899@gmail.com, [2]sriramyap@saveetha.com

Article Info
Volume 82
Page Number: 10372 - 10375
Publication Issue:
January-February 2020

Article History
Article Received*: 18 May 2019*
Revised*: 14 July 2019*
Accepted*: 22 December 2019*
Publication*: 19 February 2020*

Abstract
One of the most important decision to be taken in his/ her life is deciding their career and the path they choose to follow. In the recent times, most people often feel to reconsider the fact that whether they have chosen the wrong career and they tend to re-evaluate it. But we can help such kind of people, who are confused about their career and their future. In order to solve this problem, an enormous amount of data related to the students like aptitude, personalities should be collected and then mapped respectively. For gathering the aptitude skills of the students an Intuitive Career System is created where it makes the students to undergo certain tests and also few other questions regarding their personal background which helps us to analyze further. And for determining their personalities various social media platforms are used here with the help of Facebook Graph API. After collecting the required details, those information are used in the model designed for predicting the career of the particular student.

Keywords: *Clustering, Data Mining, Decision Tree, Data Balancing Algorithm, K- Cross*

1. Introduction

The process of choosing a path of career for an individual can help him/her to fix certain goals and come up with strategies to ensure safety of their future. The vital key of selecting a certain path or career for an individual involves an enormous amount of cooperation during the process of testing their skills related to talents, interests and their abilities. In the current world, students are having various fields and domains to work or study in, which easily confuses them to choose a particular goal or a career.

In recent days, it has been noticed that there is a gradual rise in the mobility of jobs from various background, this shows that due to the improper knowledge regarding decision making for their careers. This has increased the need of a counseling system for one's career.

In this project, a well-known algorithm known as Educational Data Mining is implemented. It uses machine learning, mining of data, and statistics on the data sets related to education. One of the algorithm being

preprocessed in EDM is clustering, that is used for machine learning technology and pattern recognition. Most of the objects are same when they are in the same cluster, but they'll vary if the objects from different clusters. When the clustering is implemented in educational system, it is known as Educational Data Clustering. EDM is being used to study online

January-February 2020
ISSN: 0193-4120 Page No. 10372 - 10375

courses which use a clustering type that is non-parametric. Various existing systems use only aptitude data or personality of the candidates for deciding their careers, but it has to be noticed that the candidate's background should also play a major role in deciding the career for them.

2. Literature Survey

In order to perform a literature survey, an intensive study has been performed by collecting various papers and articles from various resources. It has been noticed that each author has used different kind of techniques based on their respective aspects. Those research works were implemented and certain results have been obtained without any problems. In most of the articles[1], they have collected the data from different kinds of social media platforms like LinkedIn. They have predicted the new job and its details based on the details like the candidate's previously worked company's size, role of the candidate, salary being paid, etc. And also various inference models like Bayes net, KNN, Decision tree and SVM were used for the purpose of implementation.

In the research work performed by Gulati[2], it can be seen that they have used details like the count of dropout students from a particular courses. The data for this research was collected from a respective college and also from its students, the data was collected with the help of a well-known Microsoft tool, Excel Sheets. After collecting all the required data, a famous data mining tool known as SMOTE was used for data mining, for dealing with certain situations like over-sampling. Later, a traditional technique like data partitioning was used for data processing the data. The whole data set was processed under three cases. Initially all the attributes were taken into account and respective algorithm was performed on it. In the second[3], all those attributes were used to perform Attribute Selection Algorithm and they were ranked by using that algorithm, and later the top 10 were skimmed off and kept for further process. Finally[4], the algorithm applied was Data Balancing Algorithm, for selecting the required attribute. The results have been compared by using two classification algorithms, rule based and decision tree. In another research work, some data of around 200 students were collected, where each candidate's data had 13 attribute which played a vital role in deciding the career. In order to collect the data, the candidates were made to answer a set of questionnaire. For obtaining the data for data mining[5], pattern mode has been implemented and analyzed. For validating and training the data collected, a validating technique known as K-cross was used. In this implementation process, initially they looked into the base cases. Later, the obtained information is analyzed and calculated to identify the best way of splitting up the data.

January-February 2020
ISSN: 0193-4120 Page No. 10372 - 10375

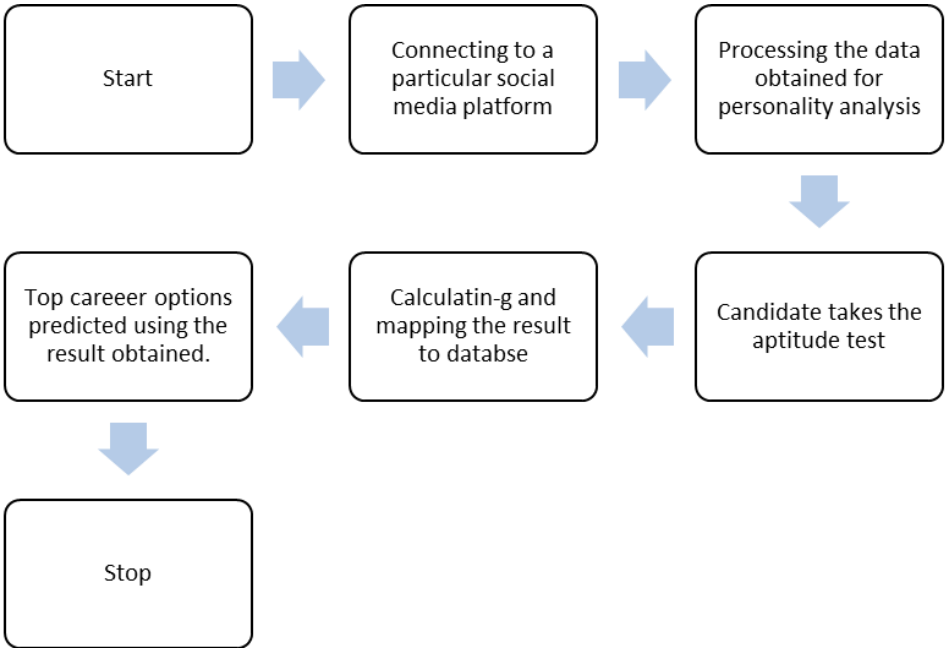

Figure 1: The Flow of Intuitive Career System

Recently, Mr. Kotsiantis performed a study[6], where they investigated in the distance learning of techniques related to machine learning and its features, especially for the dropout students in the institute. This study enabled them to provide a clear path for various data mining using data related to education. They were the first one to use the machine learning methods in the educational field and its data. Initially, the neural network has been used for predicting the performance of the students and later fitting the respective candidate into the predicted cluster, with the help of K-Means algorithm. This method helped the researchers to predict a candidate's talent at the initial stage of their academics itself.

3. Different Data Mining Technique Used

The Data mining technique used in this process are,

1) Classification- It is used to predict the class of objects in which its class label is unknown. Its objective to find a derived model used to describe and distinguishes data classes or concepts.

2) Prediction – Used to predicting the missing numerical data values rather than data labels.

3) Outlier Analysis – May be defined as data objects that don't comply with the general behavior.

4) Evolution Analysis – Refers to the description and model regularities.

4. Result

By performing such an intensive literature survey, enormous amount of useful information related to the project being implemented has been gathered. And also the cons of the existing system has been identified and further studies are being done to make sure that those cons are prevented and further updated. And also it being made sure that the system ensures that it consumes less time, work and cost, so that each and every person can be benefitted and lead a perfect career in the future.

5. Conclusion

This project work will play a key role in all the candidates who have fallen in a dilemma regarding choosing their career and a path. The

January-February 2020
ISSN: 0193-4120 Page No. 10372 - 10375

process of data mining using the educational related data primary focus is to analyze the educational system. This system helps those candidates to decide their career with ease. They will get an opportunity to showcase their talent in which ever they excel. It can be implemented in most of the educational institutions and can guide their students well. Further studies and research work is being performed for enhancing the project work and provide a much accurate result for the candidates in various aspects of their life, and helping them to lead a peaceful and happy life.

References

[1] CG. Kesavaraj and S. Sukumaran, "A study on classification techniquesin data mining", in Computing, Commun and Networking Technol.(ICCCNT) 2013 4th Int. Conf., Jul. 2013.

[2] R. Campagni, D. Merlini, R. Sprugnoli and M. C. Verri, "Data mining models for student careers", in Expert System with Appl. , vol. 42, no. 13,pp. 5508-5521, Aug. 2015.

[3] H. Gulati, "Predictive analytics using data mining technique", in Computing for Sustainable Global Develop. (INDIACom), 2015 2nd Int.Conf., Mar. 2015.

[4] J.S. Cramer, "The Origins of Logistic Regression", Tinbergen Inst., no.2002-119/4, Jan. 2003.

[5] C. Romero and S. Ventura, "Educational data mining: A review of the state of the art", in IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., vol.40, no. 6, pp. 601618, Nov. 2010.

[6] O. R. Zaane and J. Luo, "Web usage mining for a better Web-based learning environment", in Proc. Conf. Adv. Technol. Edu., 2001,pp.6064.

[7] C. Anuradha, T. Velmurugan "A Data Mining based Survey on Student PerformanceEvaluation System." 2014 IEEE International Conference on Computational Intelligence andComputing Research, 978-1-4799-3975-6/14 ©2014 IEEE

[8] R. Sumitha, E.S. Vinothkumar "Prediction of Students Outcome Using Data MiningTechniques"International Journal of Scientific Engineering and Applied Science (IJSEAS) –Volume-2, Issue- 6,June 2016

[9] Zhibing Liu, Huixia Wang, HuiZan "Design and implementation of student informationmanagement system." 2010 International symposium on intelligence information processing andtrusted computing. 978-0-7695- 4196-9/10 IEEE

[10] S.R.Bharamagoudar, Geeta R.B., S.G.Totad, "Web Based Student Information ManagementSystem",International Journal of Advanced Research in Computer and Communication engineering Vol.2, Issue 6, June 2013.