

College Recommendation System

Nilesh Dhikale¹, Apurva Dudhane², Himanshu Srivastava³, Umesh Gholap⁴, Prof L.A. Deshpande⁵

¹ *Engineering Student, Computer Science, Vishwakarma Institute Of Information Technology, Maharashtra, India.*

² *Engineering Student, Computer Science, Vishwakarma Institute Of Information Technology, Maharashtra, India.*

³ *Engineering Student, Computer Science, Vishwakarma Institute Of Information Technology, Maharashtra, India.*

⁴ *Engineering Student, Computer Science, Vishwakarma Institute Of Information Technology, Maharashtra, India.*

⁵ *Engineering Professor, Computer Science, Vishwakarma Institute Of Information Technology, Maharashtra, India.*

ABSTRACT

Educational organizations are one of the important parts of our society and playing a vital role for growth and development of any nation. For that getting appropriate college is of foremost importance. Hence proposing a system which involves data analysis and data mining techniques. Recommendation system is a part of data mining techniques which we have used. It includes machine learning and data mining techniques for filtering data and presenting the required information. Our system is made to target students, parents and educationalist who search and need counseling for getting admission in engineering colleges. This paper aims at presenting a framework of system which can recommend best engineering colleges on merit of students.

Keyword : - Recommendation system, Android, Decision tree(J48) algorithm, Data mining, Eclipse ADT, WEKA .

1. INTRODUCTION

Data Mining techniques are of great importance in present education and business system. A data mining task can be specified in the form of a data mining query, which is input to the data mining system. A data mining query is defined in terms of data mining task primitives. These primitives allow the user to interactively communicate with the data mining system and examine the findings from different angles or depths. A Decision tree approach is proposed which may be taken as an important basis of selection of college during any course program. It is a tedious job searching the best deserving college due to lack of articulate information about them. It also consumes lot of time. Recommender Systems (RSs) are software tools and techniques providing suggestions for items to be of use to a user[8]. The suggestions relate to various decision-making processes. We are proposing a system where student can shortlist colleges based on their merit, interest, fees and locality.

2. MOTIVATION

In present scenario different websites are needed to be browsed to collect articulate information regarding colleges and lot of time is consumed in searching the best deserving college. Along with giving an insight to college administration for overall rating, cutoff of their colleges, overall admission intake and preferences of students, so that such analysis can be further used by any college system for their advertisement and publicity. Also to avoid spending time and money on Counselor and stressful research related to finding cut-off of each college.

3. ALGORITHMIC COMPARISON

We have used WEKA tool for comparing algorithms based on inputs. Giving 160 entries as input which includes cut-off of 5 colleges namely COEP, VIT, PICT, VIIT and SCOE. Considering 6 attributes for classification including college, branch, gender, university, and caste and merit number. After giving training data we passed testing data for classification. College's names were taken as class labels for classification. And therefore found that Naive Bayes gave an accuracy of 50 percent that means it classified 5 out of 10 colleges correctly. While for J48 the accuracy was found out to be 80 percent.[4]

4. DATA ANALYSIS

After studying and referring papers on recommendation system it was found that Naive Bayes and decision tree algorithms are most appropriate to solve the such problem. The algorithm, which after comparing with WEKA tool gave higher accuracy is used.[5] This system would be helpful for students minimizing their time in searching colleges. The co-curricular rating is collected from a Google Form which was circulated to students of other Engineering Colleges. The Interest fields of the Google form includes attributes like infrastructure, cultural and technical activities, sports, NSS, Edc and other attributes include faculty, hostel, placement and fees. Based on the rating given by the students of various colleges we have averaged their ratings to get a mean value. A threshold value was set in order to recommend best college. The cut-off of colleges according to their categories ,branches and the branch wise intake of colleges was collected from the Directorate of Technical Education website and its published book. The information regarding colleges was gathered from DTE website.

5. RULE EXTRACTION FROM DECISION TREE

Rules are a good way of representing information or bits of knowledge. A rule-based classifier uses a set of IF-THEN rules for classification. An IF-THEN rule is an expression of the form.[6]
IF condition THEN conclusion.

The IF-part (or left-hand side) of a rule is known as the rule antecedent or precondition. The THEN-part (or right-hand side) is the rule consequent. In the rule antecedent, the condition consists of one or more attribute tests that are logically ANDed. The rules consequent contains a class prediction. To extract rules from a decision tree, one rule is created for each path from the root to a leaf node. Each splitting criterion along a given path is logically ANDed to form the rule antecedent (IF part). The leaf node holds the class prediction, forming the rule consequent (THEN part). Extracting classification rules from a decision tree. The decision tree can be converted to classification IF-THEN rules by tracing the path from the root node to each leaf node in the tree. The rules extracted are

R1: IF percentile ≥ 90 AND interest= NSS THEN likely to get VIIT = yes
R2: IF percentile ≥ 87 AND college= VIT THEN likely to get Industrial branch= yes
R3: IF distance = In between 1-25 kms. THEN easy transportation available = yes
R4: IF percentile ≤ 35 AND category= Open THEN gets Civil branch = no

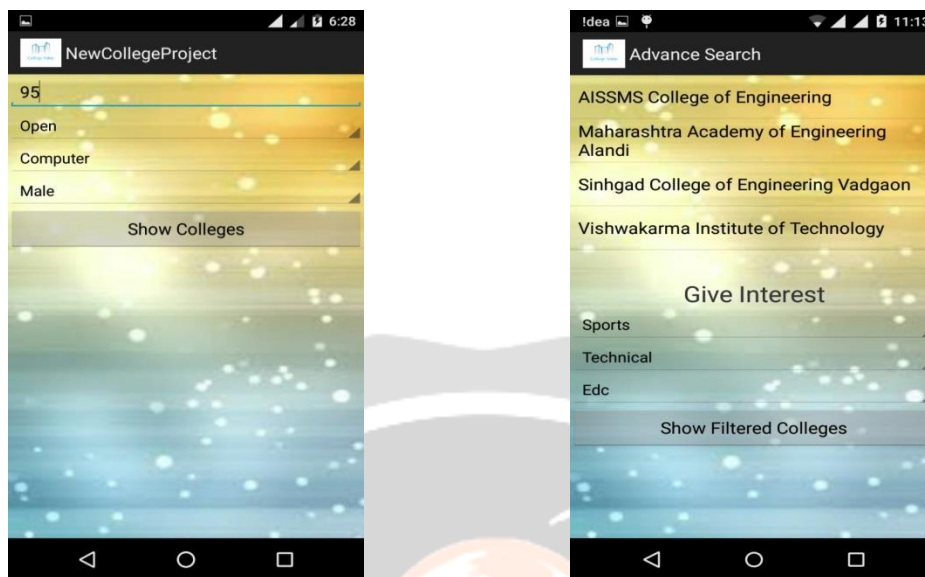
If the condition in a rule antecedent holds true for a given tuple, we say that the rule antecedent is satisfied and that the rule covers the tuple.

6. EXPERIMENTAL SETUP

Our system consists of five modules which describe the various aspects for recommending colleges which includes college details, cut-offs, branches and their comparison.

A. Our Application comprises of five modules

1) Advanced Search: Have Collected interest rating from a survey using Google form. Interest fields includes attributes like infrastructure ,cultural and technical activities, sports, NSS, EDC and other attributes include faculty, hostel, placement and fees. Based on the feedback given by the students of various colleges , have averaged their ratings to get a mean value. In this module user will be asked to give his candidature details and his interests in co-curricular and extracurricular activities. According to his merit and his interests ,colleges will be shortlisted based on their merit, interest, fees and locality.



2) College Search: In this module user will select college to view its details. After selecting college details of college which include their intake, cuttoffs, establish year, fees and contact details will be displayed.



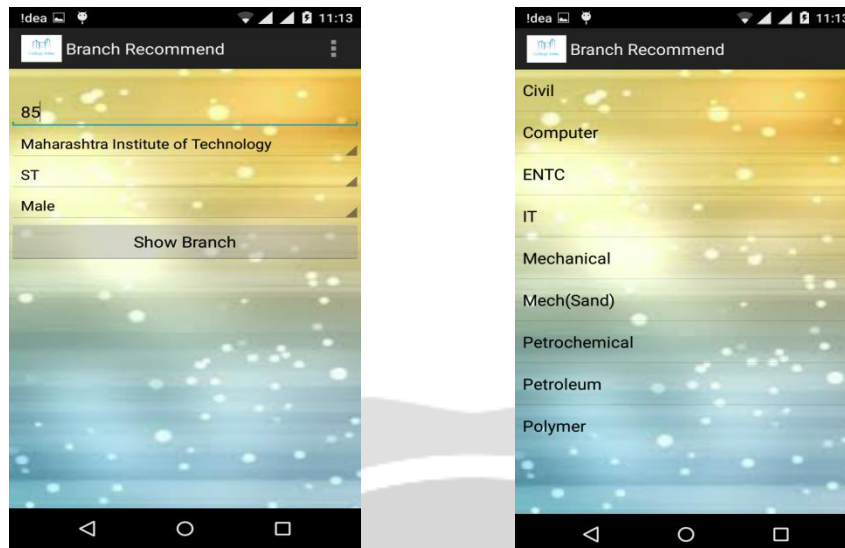
3) Locality Search: In this module the distance between colleges and few peak locations using Google map is taken. User will be asked for a peak location nearer to him and enter distance radius he wants to search. Accordingly all the nearby colleges will be displayed.



4) Comparison: In this module user will be given a choice to select colleges he wants to compare. By this he will get a clearer idea for the distinguished college as comparison will be provided by four filters based on percentile, interests, branch intake and fees.



5) Branch Search: In this module user will be asked to give his candidature details and the college name. According to his merit a list of branches will be recommended with the student is most likely to get according to merit in that college.



7. CONCLUSION

A data mining query language can be designed to incorporate the primitives, allowing users to flexibly interact with data mining systems. Rules are a good way of representing information of knowledge. A rule-based classifier uses a set of IF-THEN rules for classification. These rules can be extracted directly from the training data. Extracting classification rules from a decision tree can be done. The decision tree can be converted to classification IF-THEN rules by tracing the path from the root node to each leaf node in the tree. Thus Query based approach and Rule based classification is used to solve user general varied query

8. REFERENCES

- [1]. F.O. Isinkaye, Y.O. Folajimi, B.A. Ojokoh, Recommendation systems: Principles, methods and evaluation, Received 13 March 2015; revised 8 June 2015; accepted 30 June 2015.
- [2] Sonali Agarwal, G. N. Pandey, and M. D. Tiwari, Data Mining in Education: Data Classification and Decision Tree Approach International Journal of e-Education, e-Business, e-Management and e-Learning, Vol.2, No. 2, April 2012
- [3] Parneet Kaur, Manpreet Singh, Gurpreet Singh Josan, Lakshmi M. Gadhikar, Deepa Vincent, Lavanya Mohan and Megha V. Chaudhari Classification and prediction based data mining algorithms to predict slow learners in education sector rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015)
- [4] Anshul Goyal and Rajni Mehta, Performance Comparison of Naïve Bayes and J48 Classification Algorithms, International Journal of Applied Engineering Research, ISSN 0973-4562 Vol.7 No.11 (2012).
- [5] Trilok Chand Sharma, Manoj Jain, WEKA Approach for Comparative Study of Classification Algorithm, International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 4, April 2013.
- [6] Data Mining : Concepts and Techniques, "Han, Jiawei.; Kamber, Micheline."
- [7] Data Mining: Know It All, "Soumen Chakrabarti; Earl Cox; Eibe Frank; Ralf Hartmut Gting; Jiawei Han; Xia Jiang; Micheline Kamber; Sam S. Lightstone; Thomas P. Nadeau; Richard E Neapolitan; Dorian Pyle; Mamdouh Refaat; Markus Schneider; Toby J. Teorey; Ian H. Witten".
- [8] Recommender Systems Handbook, "Francesco Ricci Lior Rokach Bracha Shapira Paul B. Kantor".
- [9] Institutes Having Special Remark, Directorate of Technical Education, Maharashtra State, Mumbai List of Institutes participating in Centralised Admission Process for the Academic Year 2015-16.
- [10] Cut-Off 2014-2015 Book