Recommendation System using Content Filtering

A case study for College Campus Placement

Ms. Nishigandha Karbhari
Department of Computer Engineering
Shree L.R. Tiwari College of engineering
Mira Road, Mumbai, India
nishi_karbhari@yahoo.com

Prof. Asmita Deshmukh
Department of Computer Engineering
K.C. College of Engineering
Thane, India.
asmitadeshmukh7@gmail.com

Dr. Vinayak D. Shinde
Department of Computer Engineering
Shree L.R. Tiwari College of engineering
Mira Road, Mumbai, India
vdshinde@gmail.com

Abstract—**Recommender system known as information gathering system aims at creating an algorithm which, keeps in consideration the diverse needs and varying level of competence. It offers better opportunities in project development cycle under requirement phase and design phase. Social media and E-commerce market has tapped in the recommender system to boost its growth by providing with precise results. It provides with either service or product recommendation using the information gathered in the software engineering process. It is broadly divided in three categories which are Collaborative, Content-based and Hybrid recommendation approach. This paper presents a model to generate recommendations based on marks of student. It discovers best solutions which would have otherwise remained hidden. The case study performed on the recommender system implementation in college campus will result a recommendation in placement of students (employee) to companies (employer) as per their requirements in shortest possible time. It can be expected as a situation where we have tried to achieve the results while keeping in mind the requirements of employer and employee.**

Keywords— Recommendation system, Content-Based Filtering, Explicit System, Machine Learning, Placement System.

I. INTRODUCTION

Today Recommender systems (RS) have carved a niche for themselves by making a powerful impact on an individual. Social communication sites like Google, Twitter, LinkedIn, and Netflix owe a lot of their success and user friendliness to RS [1].

RS have widespread helping to find new products or services, such as books, music, transportation or even people, based on information about the user and the recommended item. Not just in decision making rather play an imperative part to settle on choices which helps to pick up benefits by connecting the best alternative as suggestion. Machine learning (ML) an area of artificial intelligence (AI) was traced as cognitive science yet was named and actualized in mid-1990s [3]. Here are wide varieties in the calculation being utilized for this recommendation to either service or any product.

The process relies on software engineering process [3]. Information is gathered in different structures and this same information is utilized for preparing or processing to create suggestions.

Recommendations are done based input data provided which can be:

A. Collaborative recommendation approach

Collaborative approach filters the information using the available information in the data clusters [10]. For instance if an individual needs to know movies to watch, then the individual with comparative preferring information will be gathered looked at and recommendation will be created. This will be completely on the premise others information gathered. The approach is collaborating others data for recommendation.

B. Content-based recommendation approach

Content-based approach compares between the users profile and the content of the item. This approach provides results with high relevance, the recommendations are transparent, and technique is easier for implementation or changes. The users profile is kept as reference with the outcomes to be generated both include some buzz keywords which are used as recommendation keywords [10]. This deals with comparable ideas of single individual.

C. Hybrid recommendation approach

Hybrid approach is combination of multiple recommendation system [10]. It consolidates both collaborative and content-based approach to generate the desired outcomes. This is generally required when there are multiple cluster with and without information are being processed where generating an outcome is mandatory.

II. PROBLEM DEFINATION

Indeed, even in today's times of automation/ computerization the greater part of the procedures involved in training and placement of students is generally manual. Making this procedure completely computerized will not just reduce time period in determination but will also add clinical accuracy to the assignment process. In addition, it might likewise additionally enhance the execution of the framework, as has been the situation with the vast majority of the frameworks which have been automated. Certifications, paper publications, paper presentation, courses, seminars, short term training programs attended by student will discover its way in resume of understudies.

This is also done presently but what makes difference is the proficiency with which it is done. This information maintenance and data usage is prerequisite for an efficient job selection procedure. [3] The current framework is a tedious framework. The process is difficult as the records of students is increasing which prompts to the limitation of the current framework. In current training and placement system, all the work is done manually, this also risks the increase in error. This may lead to non-allocation of best suited candidate for the job. The communication between candidate and administration makes present framework tedious. They submit their resume every year, leaving them solidified in time. The list is produced for each company and student need to check all the company with all its job role. The problem on excel sheet is we cannot retrieve the required data of the student or cannot find query for retrieving such criteria in existing system. There are many problems that occur in retrieving particular data. Finding student is done manually based on the criteria of the company system should find the eligible student form the data available. The current framework will check entire database and then locate a particular candidate for a particular job. This is a time consuming process. The process additionally includes checking number of understudies found, generate report of students, shortlist them and so on which takes more time. For generation of excel sheet the administration has to check entire resume and enter data into database so there is a possibility of wrong section. All this can be taken care of utilizing a automated framework and the system can be made more efficient utilizing the content based recommendation system. The Choice based system proposed in this paper depends on the preferences and interest of the candidate. The current framework lacks in this which leads to constant job change. It is required to provide the desired job role while the placement procedure. In Recommender system it furnishes with the choice based job in the result recommendation.

III. PROPOSED SYSTEM

The steps involved in applying algorithm to find results are:

A. Preprocessing

Incorporates preparing of the framework. There will be diverse sorts of information being put away in the database. Like marks acquired, courses, certification, publication presentation, short term training program, hands-on sessions and so on.
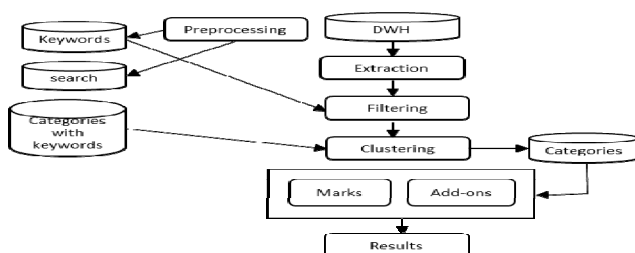

Fig 1: Proposed Model

For good analysis we need to gather buzz words identified with technical names. It is basic at top notch to first-rate the substantial keywords. Keywords are defined as a sequence of one or more words and provide a compact description of a content [5].

B. Extraction

In the Extraction we include Data extraction from database. A fundamental token is allotted to user and can be utilized to seek any information. Information such as student's details, mark sheet and add-ons will be extracted for processing. This will in any case exclude a job role list. However, keywords related to the search can be identified [8].

C. Filtering

Now comes the important aspect, separating wheat from the husk. Once the entire data is gathered, we need to convert it into information. As the extracted information is stored in a separate database known as knowledge source, which is retrieved for further tasks. Filtering includes tokenization and cleaning functions. Filtering is done based on the list of keywords, which are analyzed or mined by examination from the training corpus. Techniques are used to find out the linking data and table of databases respectively in the preprocessing phase for providing assistance in the process of filtering the contents of message [2].

D. Clustering

Clustering partitions the informational index into groups or equality classes. We have clustered users into various categories based on the requirements of result generation. For analysis of the outcomes, we have made use of clustering calculations [8].

E. Identification of Users

Identify the applicants who are keen on data identified with a specific class. The process utilizes selection of category to identify the interested users from the posts. The procedure includes Extraction of users who are linked or associated to the various keywords related to the search have to be shown inclination towards similar jobs. From that point we set up the database of users in each category [7].

F. Content Based filtering in Recommendation system

Next step is, Content-based filtering, also referred to as cognitive filtering, recommends items in view of a correlation between the content of the items and user profile. The content of each item is represented as a set of descriptors or terms, typically the words that occur in a document. The user profile is represented with the same terms and built up by analyzing the content of items which have been seen by the user [10].

IV. RESULTS AND DISCUSSION

We conclude the paper with an account of the remaining significant challenges for the field of content-based information retrieval. A significant number of the issues that stay as difficulties have tended most in th current frameworks, however it will probably be some time before these difficult

challenges are resolved. We attempted to carry out more extensive experimentation with various item collection.

Fig 2: Adding new students

The figure 2 demonstrates adding student's records. The content based system uses either content provided or extracted to perform the operation. Here in including new record it performs a three phase filtering process. The primary channel is the understudies details, then it checks with the understudies mark sheet and then add-ons are included the add-ons are optional but it is an important section for producing results.

Fig 3: Subject Report

The figure 3 incorporates Subjects in mark sheet from first year to final year. To generate results this information is fundamental. Marks obtained by students is linked for recommending the respective job role. The second criteria to generate result is add-ons from figure 2. These are courses, short term training program, certification, paper publication and soon. Figure 4 produces the final results. The outcomes are displayed only after filtering marks and job role selection. With marks as input searching is initiated to check with the list of students qualifying the requirement. Further to show the suggestions actual recommendations each job role is associated with set of subjects and add-ons. That is compared with individual subject and add-ons linked to the job.

Final results are generated including the marks obtained by the student in individual subject and add-ons of the student. The placement can then be performed by the outcomes created. This framework produces the job role recommendation according to the interest of the understudy.

Looking at the current understudy framework it will convey more stability to the placement system.

Fig 4: Linking Mark Sheet with publication, presentation, course to generate the results

V. CONCLUSION

Recommender systems is a technology used everywhere in ecommerce these days. Providing the results on social networks it has graphically increased the users as targeted by the networking owners. This paper presents a method to combine this and consider the diverse needs with varying level of competence. Categorizing students based on their credentials thereafter, it discovers best solutions to generate recommendations for placement based on the marks and various other factors included in the profile of the student. Using these soft computing techniques the student can be referred to the job profile which is not used as reference for the placement otherwise.

REFERENCES

[1] Z. L., L. W. and S. G. , "Recommendations on Social Network Sites," *IEEE,* 2009.

[2] Adomavicius, G., Tuzhilin and A., "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," *IEEE Transactions on Knowledge and Data Engineering,* vol. 17, no. 6, pp. 734 - 749, 2005.

[3] M. H. Henrik, "Two notes on machine "Learning," vol. 2, pp. 364-379, 1959..

[4] G. a. A. T. Adomavicius, "Context-aware recommender systems," *Springer US,* pp. 217-253, 2011.

[5] K. V. , N. M. , X. O. , M. W. , H. D. and B. I. D. Erik , "Context-Aware Recommender Systems for Learning: A Survey and Future Challenges," *IEEE,* vol. 5, no. 4, pp. 318-335, 2012.

[6] H. G. M. . J. L. S. N. Jeff, "Friends Recommendations in Social Network using Genetic Algorithm and Network," *IEEE,* no. 11, pp. 2207-2214, 2016.

[7] X. Xie, "Potential Friend Recommendation in Online Social Network," *IEEE/ACM International Conference on Green Computing and Communications & 2010 IEEE/ACM International Conference on Cyber, Physical and Social Computing,2010.*

[8] A. A. C. G. Karuna C.Gull, "A clustering technique to rise up the marketing tactics by looking out thekey users," *IEEE International Advance Computing Conference,* 2014.

[9] A. V. D. Tofik R. Kacchi, "Friend Recommendation System based on Lifestyles of Users," *IEEE,* 2016 .

[10] A. D. a. V. S. N. Karbhari, "Recommender System in Machine Learning," *ICEMTE,* vol. 5, no. 3, pp. 62-65, 2017.