LITERATURE SURVEY

TEAM ID : PNT2022TMID15346
TEAM NAME : TEAM INFINITY
TEAM LEADER : RAGAV BHARADWAJ S
TEAM MEMBERS:

- PRAVEEN KUMAR A (111519104102)
- PRAVEEN KUMAR S (111519104103)
- SABARIVASAN M (111519104117)

PROJECT NAME: Skill / Job Recommender Application
DOMAIN NAME : Cloud Application Development

ABSTRACT:

In today's world, students face an immense of options relating to the number of courses that they may choose from. To make this seemingly massive choice relatively easy to make, many authors have created their own recommender systems to map students to the courses that are best suited for them. However, they are not widely used as they give good results only for the dataset that they consider. In our paper, we have mapped the current students to their alumni based on multiple criteria. Afterwards, unlike other papers that used k-means, we used c-means and fuzzy clustering to arrive at a better solution to predict an elective course for the student. Since all of this is done on a broad actual dataset, the results can be applied anywhere in a real world scenario.

INTRODUCTION:

In the highly competitive environment for jobs, students are often let down by the old curriculum and courses offered by the colleges. Many entry-level employees do not feel that the university courses prepare them for actual job roles. Even if students have knowledge of a particular domain, they are often offered a job role which they never had an interest in or they were never prepared for. Hence it is very important that students develop necessary skills required for various placement profiles by enrolling in online courses. But there is huge diversity of online courses available and there are nearly hundreds of courses for various job roles. Hence it is important for students to get enroll in courses which are suitable to their interests and understanding to save time and eventually help them in the future.

This is where a recommender system can be useful for selecting a suitable track according to a student's understanding of various job domains. This paper proposes course recommendation system that suggests courses by finding students similar to target students and then searching for a pattern in their understanding of the domains and courses that they took. Finally the system recommends courses whose association with the target student's job profile is high.

LITERATURE SURVEY:

In the Paper, Amer Al-Badarenah et al. have proposed a system which recommends student courses based on the courses taken by their peers, having similar interests. The similar interest students are then grouped based on their grades. Clusters are created by using the nearest neighbor algorithm. Their system also predicts grades that the student will achieve if they take up that course. The author has used Manhattan distance to find similarity between two students by calculating the distance between their grades in common subjects and to form clusters. After the clusters are formed, the similarity of the student in interests and the cluster is found from the group by using the n-nearest neighbor algorithm. Course association rules are used in finding recommendations. It is found that in some scenarios the algorithm's performance is low. The prediction depends only on the grades which may not help in knowing the students entire potential.

Sh. Asadi et al talk about using both clustering and fuzzy logic to predict courses. The author uses a clustering algorithm to group similar kinds of students and fuzzy logic to find association rules. PCA is used to decrease the number of components. The clustering is done with K means algorithm and the maximum number of clusters considered in this paper are 5. Student's scores are labeled as low, medium, high. If two regions have equal membership that is if students mark in a subject belongs to two different regions with equal membership then both will be added in a set of rules. This is a drawback.

Feng-Hsu Wanh et al talks about improving the recommendation system of websites and making the user's experience personalized. They have integrated clustering with the associative-mining method. The customers are clustered based on the time-framed navigation session assuming that their preference may change with time. For the selection of a good time frame they have used the Hierarchical Bisecting Medoids Algorithm. The nearest cluster is found and association rule is applied to that cluster. The maximum matching method preserves the sessions and finds maximum pages that match the rules and pages will be displayed whose confidence is beyond a threshold. This paper suggests that this method will be ideal in education system courses recommendation. However, the drawback of this system is that a customer wonaˆt be given recommendations if there are no pages match the association rule.

Huiyi Tan et al. [4] propose a system that can generate recommendations for E-Learning. They collect the data of the user's previous interest and frequently visited pages are considered. Different recommendation models are stored in a separate dataset and the model is randomly selected. Once the model is selected the useful information from the user's history and recommendations are generated based on the algorithm that the model uses and sent back to the E-Learning websites.

EXISTING SYSTEM:

The developed system consist of three modules: college campus recruitment system, keyword based search from online recruitment sites and Android application. In college campus recruitment system student's profiles and company's profiles are collected . Students profile generated by taking informationfrom students through registration and login portal. Company's profile will be generated by the admin from the information and requirement provided by the company to admin. After that profile matching is perform on the students and company's profiles . In second module i.e., keyword based search module students have the provision to search for the companies from various online recruitment sites. Web crawling technique is used for searching through these sites. Students have to put the keyword e.g. C# and web crawler searches for those companies whohave vacancies for C# developers through various online recruitment sites like Naukri.com .

RELATED WORK:

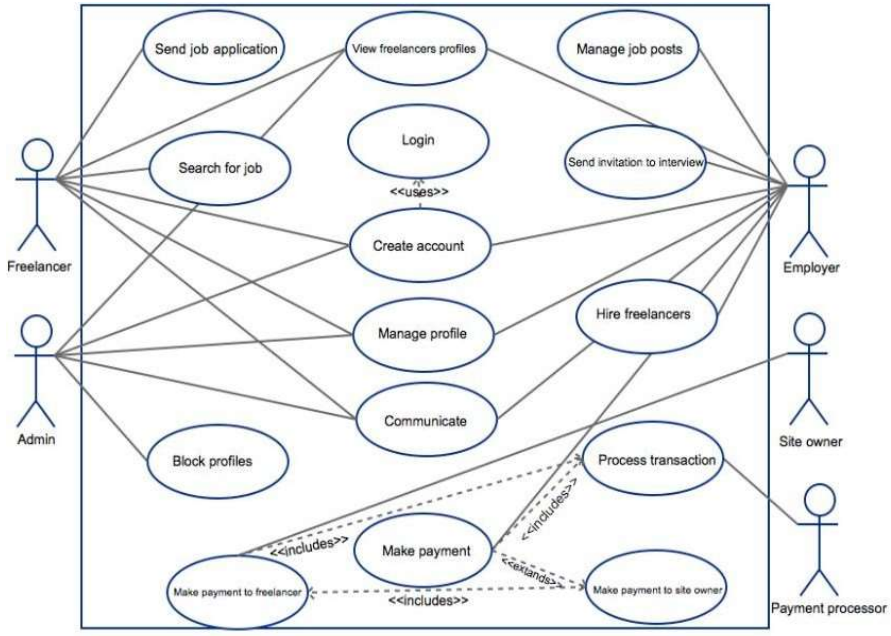
As we have seen, the present-day job seeker is faced with an array of problems before they can find a suitable job for themselves. All existing work is so promising but lacks in some of the other aspects. The need is to eliminate these issues posed by past research and minimize the weaknesses of the systems. The proposed system is designed to go forth with developing a fully functional user interface supporting a job aggregator and recommendation system. Every aspect of the operation is made from scratch and in a customized sort of manner.

Hence, the problem statement devised by us as a building starter for the research is as follows:

1. Developing a hybrid model that aggregates and recommends relevant jobs to the user based on their profile, skills, or interests.
2. Emphasizing quality over quantity and delivering only the most appropriate results to the user. The results were achieved by applying intelligent filters and filtering out great amounts of data using appropriate parameters.
3. Recommending jobs to users of any age and background in real time, based on the popularity of jobs among the other user base. Additionally, allowing users to study job popularity, skill demand, grossing market skills, etc. are discussed.
4. Finally, designing a fully useable and understandable UI for the Recommender System for practical usage.
5. The proposed system consists of the following three major modules, which are completed as part of this research as follows:
6. Data collection and preprocessing followed by the unification of the database.
7. Recommendation of suitable results using a hybrid system of content-based and collaborative filtering. Development of a fully functional user interface in the form of a web application.

USECASE DIAGRAM:

Use Case Diagrams referred as behavior diagram which describes the commutation between actors or participations and set of actions. This is set of actions or use cases will be enclosed by system boundary and can also have relation with each other. Division among tupelos will based on the information gain computed for each attribute.



PROPOSED RECOMMENDATION SYSTEM:

As discussed previously, Recommender System are the system that analyses user preference history and caters them with different options of services related to the requirement. Recommender systems emerged as an independent research area in the mid-1990s(Ricci *et al.*,2011). In recent years, the interest in recommender systems has dramatically increased. In theRecommendation algorithm, it classifies into four types: Content-based filtering, Collaborative filtering, Rule-based, and Hybrid approaches (Mobasher, 2007; Al-Otaibi and Ykhlef, 2012).

Collaborative Filtering (CF): Collaborative Filtering is a technique is based on the human ratings that are given to an item by a user and find similarity between different users who have given similar ratings to an items(Hu and Pu, 2011). The essential operation used here is the memory-based nearest neighbor approach to group users who have asimilar interest. As the volume of data grows gradually, there will be high latency in generating recommendations Mobasher (2007); Herlocker *et al.* (1999). Collaborative filtering has an advantage over content-based filtering techniques, but due to the nature of the hiring process, a job cannot be rated by the user and will not be possible to create a similarity matrix.

Content-based filtering (CBF): These are the most subjective and descriptive based filtering. Content-based filtering can also be called as attribute-based recommender as it uses the explicitly defined property of an item. It is an approach to an information retrieval or machine learning problem. The assumption made in content-based filtering is that user prefers item with similar properties. Content-based filtering recommends items to the user whose properties are similar to the item which the user has previously shown interest. Mobasher (2007) express that drawback of this filtering technique is their tendency to over-specialize in suggesting the item to a user profile as user profiles are relayed on an attribute of the previous item opted by the user. Nevertheless, in the job domain, the job listed in the job board be available only for few days; due to the nature of the domain, the tendency to over-specialize in recommending the same item would not be any problem in the job domain recommender system. In domains like entertainment, user preference are tends to change depending on variousfactors, but In Job domain, the user tends to look for the job where he can use his previous skills. New recommendation of jobs can be made when there is a change in user preference, i.e. if a user thinks to change his/her job domain by updating his new skills andthe job domain if he/she wishes. Another scenario of new recommendation is when new jobs are listed in the database; system would identify the properties of the job listed, such as job domain and skills required for the job and matches with the users with a high similarity score.

Rule-based Filtering (RBF): These filtering techniques depend upon decision rules such as an automatic ormanual decision rule that are manipulated to obtain a recommendation for the user profile. Currently, the E-commerce industry uses a rule-based filtering technique to recommend an item based on the demographic region of a user, purchase history, and other attributes that can be used to profile an user. A drawback in rule-based filtering is user feeds the information to the system. These inputs are utilized as a description of a user profile orcan be considered as a preference of a user, defined by the user. Thus the data acquired is prone to bias. With the age of the user's profile, recommendation tends to hit the saturation and become static Mobasher(2007).

Hybrid filtering (HF): As the title describe, its incorporation of multiple techniques to improve the performance of recommendation. The previously discussed recommendation technique has its weakness and strengths. In order to get a better recommendation and overcome the challenges posed by earlier techniques, this technique is sought after. All of the learning/model-based techniques suffer from cold-start in one or other form. It is a problem related to handling a new user or new item. These and other shortcomings of the CF,CBF, and RBF could be resolved by using hybrid filtering techniques Burke (2007); Jain and Kakkar(2019); Dhameliya and Desai (2019).

The surveys conducted by Burke (2002) and Dhameliya and Desai (2019) have identified different types of hybrid filtering techniques that could be used by integrating CF, CBF, and RBF.

1. Weighted: The similarity score obtained from different recommendation components are couplednumerically to get one better recommendation.
2. Mixed: Recommendations obtained from different recommending techniques are put together andpresented as one recommendation.
3. Switching: choosing one among the recommendation components based on the scenar- ios where it suits best.
4. Feature Combination: Attributes derived from diverse knowledge origins are fused and supplied to a recommendation algorithm.
5. Feature Augmentation: One recommendation technique is used to compute a set of attributes of user or item, which is then part of the input to the next recommendation technique. Two or more recommendation techniques are serialised to get on recommendation.
6. Cascade: Recommending systems are given strict priority, with the lower priority ones breaking ties in the scoring of the higher ones. Here one Recsys technique refines recommendation of another.
7. There had been attempts to develop a recommendation system by several researchers. One such implementation was done by Rafter *et al.* (2000).

   They had devised a hybrid Recsys CASPER for Job finding search engine. They had implemented an automated collaborative filtering module and personalized case retrieval module in their job recommendation system. ACF module utilized user behavior information such as read time and activity onthe page during his time on the system to profile the user. Similarity measure such as the Jaccard index and other clustering algorithms was used for similar grouping user against target user. Their other module PCR finds the similarity between the user's query and jobs in the system. The module computes similarity with a target user's query and jobs from the job case base using different similarity measures. This system has faced sparsity and scalability problems

FUTURE SCOPE:

The previous recommendation system clusters students only on the basis of their grades. They also map a student to only one cluster and do not consider any other possibilities. However, this paper takes into account a student's interests and clusters them accordingly. The target student is associated with all the clusters using the c-means fuzzy algorithm. It finds the clusters in which the students are similar to the target student and group the cluster. The courses taken by the most similar students to the target, from the cluster, are found out and the most common course is recommended.

REFERENCE:

[1] The International Workshop on Artificial Intelligence and Smart City Applications (IWAISCA)August 9-12, 2020, Leuven,Belgium Job Recommendation based on Job Profile Clustering and JobSeeker Behavior

[2] Wang, F.H. and Shao, H.M., 2004. Effective personalized recommenda- tion based on time-framed navigation clustering and association mining.Expert systems with applications, 27(3), pp.365-377.

[3] Al-Badarenah, A. and Alsakran, J., 2016. An automated recommender system for course selection. International Journalof Advanced ComputerScience and Applications, 7(3), pp.166-175.

[4] Asadi, S., Jafari, S. and Shokrollahi, Z., 2019. Developing a Course Recommender by Combining Clustering and Fuzzy Association Rules. Journal of AI and Data mining, 7(2), pp.249-262.

[5] Wang, F.H. and Shao, H.M., 2004. Effective personalized recommenda- tion based on time-framed navigation clustering and association mining.Expert systems with applications, 27(3), pp.365-377.

[6] H. Tan, J. Guo and Y. Li, "E-learning Recommendation System," 2008 International Conference on Computer Science and Software Engineer- ing, Hubei, 2008, pp. 430-433.