**SKILL AND JOB RECOMMENDER – LITERATURE SURVEY**

The amount of information available online rose in tandem with the Internet's rapid expansion, driving a greater need to improve users' capacity for information management. This stimulates a great deal of interest in particular study areas and technological advancements that can help manage this information overload. Information retrieval and information filtering are the two most crucial domains. Information filtering works to help users get rid of undesirable information, whereas information retrieval deals with automatically matching users' information. Recommender systems, which have their roots in cognitive science, approximation theory, information retrieval, forecasting theories, as well as management science and consumer choice modelling in marketing, are the most recent technology to combat information overload (Adomavicius and Tuzhilin, 2005). The recommender systems used a number of information sources that are related to users and objects to decide the interested items for a certain user.

Widely used in many applications, recommender systems make recommendations for goods, services, and information to potential clients. In order to improve customer services, boost sales, and shorten customer search times, many e-commerce platforms integrate recommender systems (Schafer et al., 1999).For instance, a wide range of businesses, including books (Mooney and Roy, 2000), news articles (Das et al., 2007), and the online book retailer Amazon.com (Linden et al., 2003). Additionally, Microsoft offers customers a variety of recommendations, including products that may be downloaded for free and bug fixes (Shani and Gunawardana, 2011). All of these businesses have established effective commercial recommender systems, which have boosted client fidelity and increased web sales. Additionally, a lot of software providers offer independent generic recommendation technologies.

**Term Frequency-Inverse Document Frequency (TF-IDF)**

As a statistical tool for assessing a word's significance in a corpus of documents, TF-IDF gives words weights. This significance is inversely correlated with the frequency of the term in the corpus and inversely proportional to the number of times it appears in the document. When used on smaller text datasets, such those frequently seen in job descriptions, this method performs less well than when used on larger ones for subject identification. TF-IDF has been used to deal with recommendations, though, and the results are intriguing.

**Word2vec**

A general predictive model called Word2vec is used to learn word vector representations. The distributional semantics and co-occurrence statistics are captured by these vector representations, also known as word embeddings.

CBOW and Skip-gram are two Word2vec models that we can utilise to get word embeddings.

Model of the Continuous Bag of Words (CBOW). Based on the n words before and n words following the target word, this model predicts the target word. For instance, consider the following:

Sit amet lorem ipsum dolor

Lorem, ipsum, sit, and amet are the n = 2 words before and after it that CBOW will use to predict the word. These terms are called the context of the target word and their quantity is a parameter of the model.

RecSys 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 the Recommendation algorithm, it classifies into four types: Content-based filtering, Collaborative filtering, Rule-based, and Hybrid approaches.

**Collaborative Filtering (CF)** is a technique that uses human ratings given to items by users to compare the ratings of various users who have given comparable ratings to those same items (Hu and Pu, 2011). In order to group users who share a common interest, the memory-based nearest neighbour method is the key operation performed here. There will be significant lag when creating recommendations as the volume of data gradually increases.

Herlocker et al. (2007); Mobasher (2007) (1999). Although content-based filtering techniques have an advantage over collaborative filtering, due to the nature of the hiring process, a job cannot be rated by the user and cannot be compared using a similarity matrix.

The most subjective and descriptive kind of filtering are **content-based filters (CBF)**. As it utilises attributes, content-based filtering is also known as attribute-based recommender the element's specifically stated characteristic. It is a method for retrieving information or difficulty with machine learning. In content-based filtering, it's assumed that users favour item with comparable qualities. The user is given recommendations for products depending on the content. the user has already expressed interest in the item's attributes. Mobasher According to (2007), a downside of this filtering method is their propensity to over-specialize in recommending the product to a user profile since user profiles are based on a characteristic of the earlier item that the user selected. 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 various factors, 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 and the job domain if he/she wishes.

**Rule-based Filtering (RBF):** These filtering techniques depend upon decision rules such as an automatic or manual 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 or can 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 staticMobasher (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 coupled numerically to get one better recommendation.

2. Mixed: Recommendations obtained from different recommending techniques are put together and presented as one recommendation.

3. Switching: choosing one among the recommendation components based on the scenarios 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.

Several researchers had made an effort to create a recommendation system.  Rafter and colleagues carried out one such implementation (2000). A hybrid Recsys had been created. CASPER is a search engine for jobs. They had put in place an automatic collaboration system. Their job recommendation system includes a filtering module and a customised case retrieval module. ACF module made use of user behaviour data like page activity and read time. while he was using the system, in order to profile the user. For comparable grouping users versus target users, clustering methods like the Jaccard index and other similarity measures were used. Their second module, PCR, determines whether the user's query and system jobs are similar. Using several similarity metrics, the module determines how closely a target user's query and jobs from the job case base are related. Problems with scalability and sparsity have plagued this system.

**REFERENCES**

**[1]** Adomavicius G, Tuzhilin A (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. IEEE Trans. Knowl. Data Eng. 17(6):734-749.

**[2]** Barbieri N, Costa G, Manco G, Ortale R (2011). Modeling Item Selection and Relevance for Accurate Recommendations: A Bayesian Approach. In Proceedings of the fifth ACM conference on Recommender systems (RecSys '11), Chicago, Illinois, USA, ACM pp. 21-28.

**[3]** Chen PC (2009). A Fuzzy Multiple Criteria Decision Making Model in Employee Recruitment. IJCSNS Int. J. Comput. Sci. Netw. Secur. 9(7):113-117.

**[4]** Fazel-Zarandi M, Fox MS (2010). Semantic Matchmaking for Job Recruitment: An Ontolgy Based Hybrid Approach. In Proceedings of the 3rd International Workshop on Service Matchmaking and Resource Retrieval in the Semantic Web at the 8th International Semantic Web Conference. Washington D.C., USA, CEUR-WS.org.

**[5]** Ghazanfar MA, Pr¨ugel-Bennett A (2010). Building Switching Hybrid Recommender System Using Machine Learning Classifiers and Collaborative Filtering. IAENG Int. J. Comput. Sci. 37:3.

**[6]** Laumer S, Eckhardt A, Weitzel T (2010). Electronic Human Resources Management In An E-business Environment. J. Electron. Commer. Res. 11(4):240-250.

**[7]** Lee DH, Brusilovsky P (2007). Fighting Information Overflow with Personalized Comprehensive Information Access: A Proactive Job Recommender. In Proceedings of the 3rd Conference on Autonomic and Autonomous System, Athens, Greece, IEEE pp. 21-26.

**[8]** Mooney RJ, Roy L (2000). Content-Based Book Recommending Using Learning for Text Categorization. In Proceedings of DL ‟00: Proceedings of the Fifth ACM Conference on Digital Libraries, New York, NY, ACM pp. 195-204.

**[9]** Rafter R, Smyth B (2001). Passive Profiling from Server Logs in an Online Recruitment Environment. In Proceedings of the IJCAI Workshop on Intelligent Techniques for Web Personalization (ITWP 2001), Seattle, Washington, USA pp. 35-41.

**[10]** Singh A, Catherine R, Visweswariah K (2010). PROSPECT: A System for Screening Candidates for Recruitment. In Proceedings of 19th ACM International Conference on Information and Knowledge Management (CIKM'10), Toronto, Ontario, Canada, ACM pp. 659- 668.