SKILL AND JOB RECOMMENDER

APPLICATION

INTRODUCTION:

A recent report claims that most college graduates have difficulty in choosing their domain in their job. Many engineers are trying to shift the domain from their field to IT. So, they are doing some courses online and randomly searching for a job. Nowadays, IT fields are the target of many students but they don't know which domain is fit for them. To avoid this situation candidates, need a Job recommendation that analyses the skills to recommend a suitable job for the candidate. The solution is to design a system that reads a resume and their skills. The resumes are going through pre-processing to make the design more efficient. For pre-processing top words and porter stemmer, Porter Stemmer will make every word their root word, and stop words will remove every meaningless word. This makes the system more efficient. Using tf-idf vectorizer for both resume and job description. Then compare the skills in the resume and description.

LITERATURE SURVEY:

Existing works are mainly found for the company to select a candidate who is fit for their vacancy. There are many experiments for calculating the four recommendation algorithm but with a different distance formula namely the Minkowski distance. And some others are tried a different recommender system like collaborative which only helps when there are more data to relate. That won't help for a person who is searching that which job is the correct choice for him/her. R.J. Mooney and L. Roy used Content-Based Book Recommending where the content-based recommendation helps for a cold start. And some authors also say that a content-based recommender is best when they researched a comparison study of job recommendations . A recommender system is not only the main part of accurate prediction. There are some other things like vectorizing the words and then similarity functions. Authors like Shouning Qu, and Li-Ping Jing aid that for text mining, tf-idf is the best approach for text feature selection. Ravali Boorugu has researched NLP and tried various text summarization techniques. Some papers also say about similarity detection with many languages. Jeevamol Joy and Renumol V G discussed which similarity is the best one for a content-based recommending system. They finally concluded that cosine similarity is the best similarity for content based recommended system. Cosine similarity is not only used for recommender systems but is used to find the similarity functions between two sentences or two paragraphs. Mohammad Alobed has tried "A Comparative Analysis of Euclidean, Jaccard, and Cosine Similarity Measure and Arabic Wordnet for Automated Arabic Essay Scoring", and L. Zahrotum also compared Jaccardidean and cosine similarity. They both said that Cosine similarity with all stemming types has the lowest error compared with the Jaccard and Euclidean similarity. There is already a system that worked with both tf-idf and cosine similarity recommendations. It is used for patient support forums. Tanya V. Yadalam, Vaishnavi, M.Gowda, and Vanditha researched those career recommendations content-based filtering which was mostly like my project but inside it, they mostly discussed security, transparency for the data, and the framework. Most works are just built for the companies or for the

purpose of making money from the people by giving some irrelevant choices. Many were using collaborative recommendation, which recommends the many searched jobs or the jobs which were chosen by some other. It only works if the system deals with more number of resumes which seems it can only be used by the companies. Some systems are asking to log in and some were asked to buy subscriptions. Logging in makes you redirect some spam mails. In many papers, they have been solved through content recommender which is not enough. A literature paper had done research on content recommender system, vectorizer, and cosine similarity in a row but in that the author doesn't think about the implementation process and only concentrated more on securing the data.

REFERENCES:

[1] R. J. Mooney and L. Roy, "Content-Based Book Recommending Using Learning for Text Categorization," in In Proceedings of DL '00: Proceedings of the Fifth ACM Conference on Digital Libraries, New York, NY, pp. 13-20, 2000

[2] Li-Ping Jing, Hou-Kuan Huang, Hong-Bo Shi, "Improved feature selection approach TFIDF in text mining", International Conference on Machine Learning and Cybernetics, pp. 944-946, 2002, doi:10.1109/icmlc.2002.1174522.

[3] Shouning Qu ,Sujuan Wang,Yan Zou, " Improvement of Text Feature Selection Method Based on TFIDF", International Seminar on Future Information Technology and Management Engineering, pp. 79-81, 2008, doi:10.1109/fitme.2008.25.

[4] I. A. Braga, "Evaluation of stopwords removal on the statistical approach for automatic term extraction," Seventh Brazilian Symposium in Information and Human Language Technology, pp. 142-149, 2009.

[5] Nikolaos D. Almalis, Prof. George A. Tsihrintzis, Nikolaos Karagiannis, Aggeliki D. Strati, "A new content- www.ijcrt.org © 2022 IJCRT | Volume 10, Issue 8 August 2022 | ISSN: 2320-2882 IJCRT2208099 International Journal of Creative Research Thoughts (IJCRT) www.ijcrt.org a784 based job recommendation algorithm for job seeking and recruiting", 6th International Conference on Information, Intelligence, Systems and Applications (IISA), pp. 1-7, 2015, doi:10.1109/iisa.2015.7388018.

[6] Mohammad Alodadi and Vandana P. Janeja, " Similarity in Patient Support Forums Using TF-IDF and Cosine Similarity Metrics", International Conference on Healthcare Informatics, pp. 521-522, 2015, doi:10.1109/ichi.2015.99.

[7] L. Zahrotun, "Comparison jaccard similarity, cosine similarity and combined both of the data clustering with shared nearest neighbor method," Computer Engineering and Applications Journal. vol. 5. Pp. 11-18, 2016, doi:10.18495/comengapp.v5i1.160, 2016.

[8] Peng Yi, Cheng Yang ,Chen Li, Yingya Zhang, "A Job Recommendation Method Optimized by Position Descriptions and Resume Information", IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), pp. 762 -764, March 2017, doi:10.1109/rteict.2017.8256590.

[9] Minh-Luan Tran, Anh-Tuyen Nguyen, Quoc-Dung Nguyen, Tin Huynh, "A comparison study for job recommendation", International Conference on Information and Communications (ICIC), pp. 199-204, 2017, doi:10.1109/infoc.2017.8001667.

[10] Gokul P.P, Akhil BK, Shiva Kumar K.M, "Sentence similarity detection in Malayalam language using cosine similarity", 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), pp. 221-225, 2017, doi:10.1109/rteict.2017.8256590.

[11] H. Li, Y. Ge, H. Zhu, H. Xiong, and H. Zhao, "Prospecting the career development of talents: A survival analysis perspective," in *Proc. ACM SIGKDD International Conference on Knowledge Discovery and DataMining*, 2017, pp. 917–925.

[12] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[13] S. Roweis and Z. Ghahramani, "A unifying review of linear Gaussian models," *Neural Computation*, vol. 11, no. 2, pp. 305–345, 1999.

[14] M. Fraccaro, S. K. Sønderby, U. Paquet, and O. Winther, "Sequential neural models with stochastic layers," in *Proc. Advances in Neural Information Processing Systems*, 2016, pp. 2199–2207.

[13] R. Krishnan, U. Shalit, and D. Sontag, "Structured inference networks for nonlinear state space models," in *Proc. AAAI Conference on Artificial Intelligence*, 2017.