A Comprehensive Framework for Online Job Portals for Job Recommendation Strategies Using Machine Learning Techniques

Kamal Upreti, Shikha Mittal, Prakash Divakaran, Prashant Vats, Manpreet Bajwa, and Sandeep Singh

Abstract The employment market in today's modern society is growing increasingly active, which makes choosing a clear opportunity for yourself a difficult endeavor, particularly for newcomers who are unfamiliar with the numerous possible professions. As a result, the need for employment recommendation systems has been steadily increasing. Many systems employ suggestions to provide consumers with personalized solutions. By examining job recommendation articles, we are taking into account various machine learning algorithms as well as models provided in this study. The information in the student's résumé is compared to the specifications of the job opportunities. Users' abilities, knowledge, past previous employment, demographic data, as well as other necessary details are extracted from recommendation apps. The applicant is presented with fresh positions that are unrelated to the one being sought based on the extraction of information. We discovered that by using content-based filtering to unsupervised based on deep learning classification methods such as SVM, KNN, and randomized forest, the random forest approach

K. Upreti (✉)
Department of Computer Science and Engineering, Dr. Akhilesh Das Gupta Institute of Technology and Management, New Delhi, India
e-mail: kamalupreti1989@gmail.com

S. Mittal
Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India
e-mail: shikhs.mittal@gmail.com

P. Divakaran
Department of Business Administration, Himalayan University, Itanagar, Arunachal Pradesh, India
e-mail: prakashtek@gmail.com

P. Vats · M. Bajwa · S. Singh
Department of Computer Science and Engineering, SGT University, Gurugram, Haryana, India
e-mail: prashantvats12345@gmail.com

M. Bajwa
e-mail: Manpreet_feat@sgtuniversity.org

S. Singh
e-mail: sandeep_feat@sgtuniversity.org

M. Tuba et al. (eds.), *ICT Infrastructure and Computing*, Lecture Notes in Networks and Systems 520, https://doi.org/10.1007/978-981-19-5331-6_74

delivers the highest outcomes for our applications. Python is used to construct the recommendation engine.

Keywords Job recommendation · Collaborating filtering · Random forest · K-nearest neighbor · SVM · Machine learning · First section

1 Introduction

Presently, everything we look for on the computer returns a recommendation. All of this is done through an intelligence filtration process that predicts a user's preferences for certain products to give personalized recommendations. Several businesses are reaping huge profits as a result of this fascinating strategy, which is also beneficial to an individual's development. We looked into several articles that are testing their efficacy in a variety of ways [1]. We feel that recommending the most relevant job opportunities to newcomers who are seeking appropriate open positions depending on particular talents and background is incredibly helpful. This method can make job seeking easier for them, which is a major issue with job posting websites that just provide responsibilities and job descriptions to their visitors [2]. We look at a particular recommendation systems sector, and job suggestions, and offer a novel strategy and suitable classification method for this area [3]. This domain can easily handle millions of objects, such as user credentials and job posts, and also even additional information in the form of user engagement with these items. Furthermore, many E-commerce webpages, the most general implementation of existing algorithms, use information retrieval methodologies without contemplating the user's curriculum vitae and item's characteristics in this particular instance, that is, educators' resumes and specifics of trying to recruit data, thus according Zhang et al. [4]. Numerous E-commerce webpages, the broadest sense implementation of classification techniques, use information retrieval methodologies without contemplating the user's curriculum vitae and item's characteristics in this case, that is, educators' resumes and information of attempting to recruit relevant data as a result, we suggested an enhanced item-based cooperative thresholding technique [5]. The goal of this study is to provide an accurate method of online resume seeking advice utilizing a few well-known categorization algorithms. Our next paper will consolidate more research on populating users' preference matrices with implicit behavior. The act of identifying, comprehending, and classifying concepts, and things into current groups or "sub-populations" is known as the classification method [6]. Machine learning techniques categorize future information into groups using which was before training examples and a range of techniques. This is relevant to textual classification tools, which is our study subject; for instance, sentiment classification, where a well-trained model delivers extremely reliable data [7]. Moreover, we explore what recommendations categories we use and also why, as well as which classification techniques provide the best results for our system, in this document. There are five major components to this study [8]. The second section discusses the related research of recommender systems. The architecture and techniques used to choose the preferable one are described in Sect. 3. Section 4 summarizes our findings and includes system evaluations. Following that, Sect. 5 gives an overview of the work that will be done in the foreseeable.

2 Related Work

Employment recommendation systems (ERS) are now being explored from a variety of angles, and a variety of technological techniques have been used to construct them. Hong et al. [6] examined the differences between conventional recommendation engines (RS) and ERS. Al-Otaibi et al. [5] did a review of several recommender systems, whereas Hong et al. [6] conducted a comparative study between conventional recommendation engine (RS) and ERS. The profile page analysis and preservation of the user's contributions are done through the application of various suggestion algorithms by the recommender system. Information screening, content-based filtering, experience and understanding filtering, and mixed suggestion systems are examples of recommendation engines that use many strategies to arrive at more reliable recommendations; however, this is a much more comprehensive approach. Alghieth et al. [7] presented a compilation of all other recommender system and their requirements, including Caspar, persistent, prospective, eRecruiter, iHR+, and a neural model for employment opportunity, to name a few. A recommender system is a simple method that produces efficient outcomes [9]. Since our systematic review needs to take into account fundamental strategies of suggestion.

1. Minimum requirements engineering effort is needed.
2. People and objects are identifiers with no inner structure or attributes that provide satisfactory outcomes in the vast majority of cases.
3. In most cases, it delivers acceptable outcomes.

There are two forms of recommender systems (CRF): consumer CF [10] and piece of information CF [11].

1. Recommender systems (CRF): identify comparable viewers who have rated in the same way as the user account in the previous and then use their recommendations on other products to anticipate what the user of the system would enjoy [12].
2. Component CF: Instead of utilizing commonalities in consumers' rating behavior to forecast inclinations, artifact component CF looks for parallels in product rating patterns [13]. In most cases, a piece of information approach is preferred over a consumer one. Product Cfm makes use of commonalities in product-based detection. Object approaches are excellent for off-line computation because identifying compared to manufacturing is easier than creating similar people, and object qualities are much more stable that user preferences [14]. Both are shown in Fig. 1.

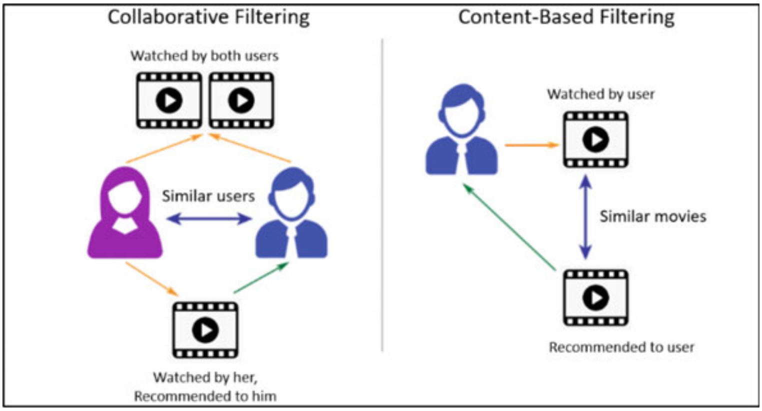

Fig. 1 Two types of job recommendation system

3 Proposed Methodology

This section gives an overview of the major machine learning approaches (ML). Deep learning is divided into three types: unsupervised, supervised, and evolutionary reinforcement computation. Classifier, on the other hand, is a method that may be employed by any ML kind. Data may be classified in two ways: organized and unregulated [15]. Classification is the process of categorizing data into a set of categories. The basic purpose of a categorization challenge is to determine which group or class to new evidence will belong [16]. The following are a few examples of machine learning techniques that are used to assess predictive performance.

3.1 Closest K Neighbor (K)

Closest K neighbor (KNN) is one of artificial learning's very basic but crucial categorization methods. Recognition systems, data gathering, and penetration testing are just a few of the applications; it discovers in the directed learning area. KNN arose from studies conducted for the military [17]. In 1951, Fix and Hodge, two officers from the United States Air Force School of Aviation Medicine, published a technical paper inventing the clustering techniques. This method uses prior data (also known as training examples) to arrange locations into categories based on a characteristic. The integrator table is extensively used in real-world issues since it has no basic assumptions about the distribution of data. To provide job suggestions, the nearest neighbor strategy is employed to calculate the shortest route between training and testing data. KNN is a deep learning approach that uses the people with cognitive of the top-k-nearest neighbors to locate clusters of similar users based on shared abilities and generate predictions [18]. To find the correct *K*, run the KNN algorithm multiple

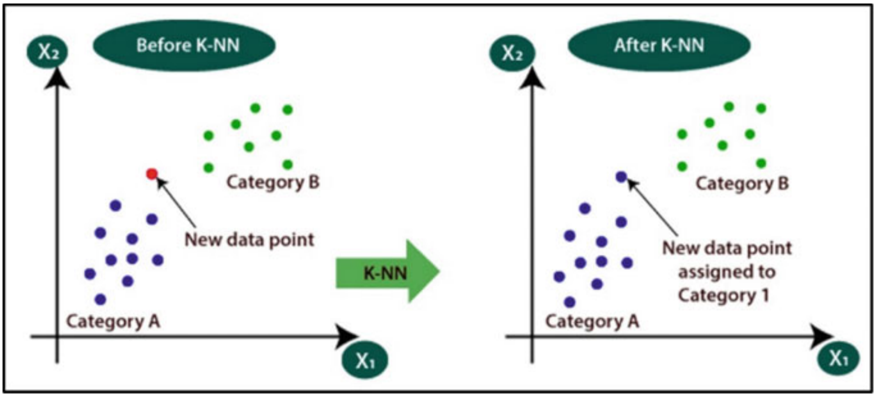

Fig. 2 Nearest *K* neighbor clustering approach

times with values of *k* and choose the one with the fewest mistakes. For instance, we may provide awareness regarding in a matrix with one row for every item (job opportunity) and columns for every user (job applicant), subsequently, sort by job criteria and add a new row for overall skills number [19]. When we integrate the abilities information with the overall abilities count, we have precisely what we need to determine which occupations are comparable and screen those out from everyone else as shown in Fig. 2.

The KNN method, as presented by Suharyadi and Kusnadi [15], achieved a satisfying result of 80%, respectively, for his trials, which is acceptable enough. To get adequate data, their employment suggestion method incorporates data cleansing, extraction, selection, and reduction. Tunak et al. [12] found that because KNN does not require any training, it is quicker and has fewer parameters to tweak. For a satisfactory outcome, adequate scaling is required. With its low computational, KNN may accomplish a lot of tasks that other classification techniques cannot.

3.2 Vector Classifier Machines (SVM) for Reinforcement

Vector support networks (SVM) is a categorization approach that uses the geometrical concept of hyperplane to segregate sequence data. The SVM is a simple method that may be used to create a hyperplane that separates and split data into sections with only a single set of data in each sector if the data has previously been classed. The SVM is extremely efficient and precise for categorization. This approach is used to reduce categorization mistakes [20]. The bigger the margins in the hyperplane-divided section, the greater the outcomes. The hyperplane is chosen after considering every one of the options and selecting the best one. SVM utilizes the notion of architectural risk aversion, according to Min and Han [14], by generating an optimum hyperplane that separates in the concealed subspace and utilizing nonlinear computing to

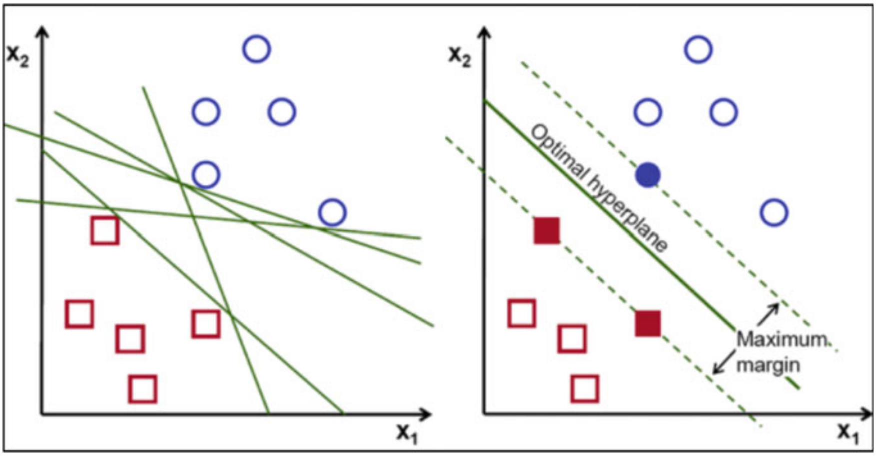

Fig. 3 Vector classifier machines (SVM) for reinforcement

discover a single answer. The goal is to determine the hyperplane that separates the relevant employment offerings from the non-related advancement opportunities the best. Whenever new content is loaded, we try to track down the hyperplane that best separates the information. SVM provides several benefits, according to Martinez-Gil et al. [13]. (1) The SVM contains a servicing and maintenance that prevents over-fitting in terms of geospatial margin. (2) SVM is characterized by an optimization process whereby a variety of effective systems already exist. Vector classifier machines (SVM) for reinforcement are shown in Fig. 3.

3.3 Random Forest of Chance

The term "randomized" is used because we use randomly generated processes: bootstrap and feature subset picking. To create various data sets, we employ random selection with repetitions, a technique known as bootstrap, and the databases we produce are known as bootstrap sample sets of data. Also, we select a set of attributes from each tree and utilize them exclusively for retraining; preferably, the length of our set of features which we use for further assessment is sqrt (total set of optimized features) or log (total set of generic features) [21, 22]. Likewise, the term forest is utilized because we are interpreting the intended outcomes using many decision trees. Appropriate for large data sets if applicability is not an issue, as these contain many trees structure, making interpretation more challenging. In a randomized forest, we send a piece of data thru every tree and record the forecast produced by such a tree, after which we collect all of the guesses and use the consensus voting to get our final prediction. Integration is the process of merging findings from several models. In an essence, confusion matrix is a mix of Lean startup and aggregating, referred to as

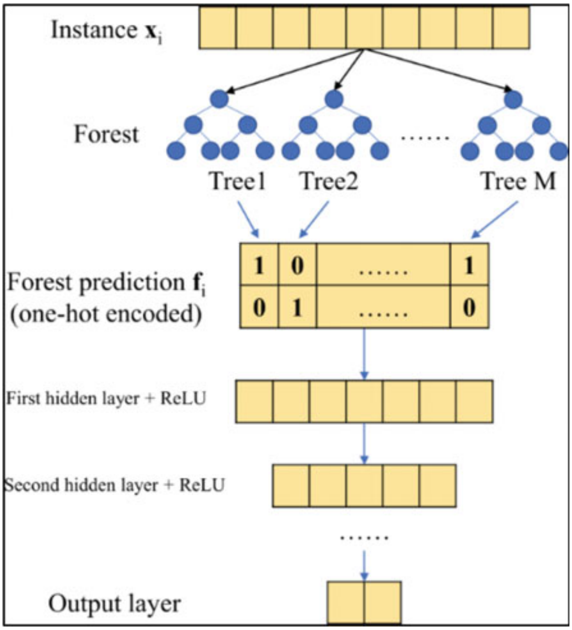

Fig. 4 Random forest optimization

bagging. This approach prevents fitting problem by combining these two procedures. The number of selected trees utilized in randomized forest has a direct relationship with efficiency. According to the classification algorithm in the research paper by Parida et al. [11], the efficiency of utilizing the supplied data set was 93.45% as shown in Fig. 4 to represent the random forest optimization.

4 Possible Recommendations Based on Machine Learning for Online Job Hiring

The recruitment process has a unique subtlety; hence, AI will never again be capable of replacing recruiters. However, it will increase the speed and quality with which recruiters can execute their jobs, which has financial and productivity ramifications for any recruiting company. AI can assist in the following areas:

4.1 Reduced Time to Recruit

Screening people for certain jobs may be a time-consuming process, therefore, AI helps with all of this important measure. Up to 45% of a recruiter's time has been spent on mechanical duties such as entering information into an application monitoring system (ATS) or simply examining applications, time that could be best invested in the human element of creating applicant connections, and optimizing the recruiting process [23].

4.2 Enhancing Recruitment Performance

Selecting the "exactly correct" applicant for each post is critical, as well as the prince of darkness is in the specifics. AI is critical in quickly looking through a large data set to connect talents or traits to jobs. That when a recruiter utilizes the automation tool, the quicker AI learns what such an accurate model appears like to enhance the performance of potential candidates [22].

4.3 Recruiting Idle Applicants

AI can help with the acquisition of "potential recruits," or people who are not proactively searching for work. Because the depth of talent for all of this category is so large (it's very nearly everybody who is not searching for work), AI could swiftly zero in on exceptional prospects, if they are searching or otherwise.

5 Conclusion

The research establishment is becoming increasingly interested in recommendations and categorization systems. The conclusions of the most closely connected publications to the study suggested in this piece are represented in this part. We determined that KNN requires greater computing effort than SVM, while SVM also consumes a long time than RF, particularly in nonlinear categorization, following researching such techniques and related findings. When numerous characteristics are included in an employment recommendation engine, KNN fails, although SVM and RF both keep improving whenever the variable numbers are high. SVM is a more result-oriented strategy, whereas RF operates in a much more understandable format, understanding the interconnections between both the parameter settings. Finally, we have shown that RF and SVM can be used to create automated job suggestion algorithms. If we utilize evolutionary algorithms, we can improve SVM. As the combination of multiple trees grows, the RF becomes more exact.

6 Future Work

The proposed methodology recommends us based on a large quantity of data. Rather than providing an immense amount of information, it suggests the precise thing. The success of the recommendation system across a large quantity of data has boosted its appeal in current history. Based on our research, we believe that we do some client personality psychology utilizing temporal features from our perception and engagement statistics. Furthermore, in the future employment, we may explore advanced algorithms such as a combination of RF and SVM, as well as a combination of spam detection and cooperation, to make the recommending engine more interactive.

References

1. Varshney S, Vats P, Choudhary S, Singh D (2022) A blockchain-based framework for IoT based secure identity management. In: 2nd international conference on innovative practices in technology and management (ICIPTM), pp 227–234. https://doi.org/10.1109/ICIPTM54933.2022.9753887
2. Chauhan K, Gupta K, Vats P, Mandot M (2022) A comparative study of various wire-less network optimization techniques. In: Joshi A, Mahmud M, Ragel RG, Thakur NV (eds) Information and communication technology for competitive strategies (ICTCS 2020). Lecture notes in networks and systems, vol 191. Springer, Singapore. https://doi.org/10.1007/978-981-16-0739-4_61
3. Kaur R, Vats P, Mandot M, Biswas SS, Garg R (2021) Literature survey for IoT-based smart home automation: a comparative analysis. In: 9th international conference on reliability, Infocom technologies and optimization (trends and future directions) (ICRITO), pp 1–6. https://doi.org/10.1109/ICRITO51393.2021.9596421
4. Zhang Y, Yang C, Niu Z (2014) A research of job recommendation system based on collaborative filtering. In: Seventh international symposium on computational intelligence and design
5. Al-Otaibi ST, Ykhlef M (2012) A survey of job recommender systems. Int J Phys Sci 7(29):5127–5142
6. Hong W, Zheng S, Wang H (2013) A job recommender system based on user clustering. J Comput 8(8)
7. Alghieth M, Shargabi AA (2019) A map-based job recommender model. Int J Adv Comput Sci Appl 10(9)
8. Qureshi A, Batra S (2022) A review of machine learning (ML) in the internet of medical things (IOMT) in the construction of a smart healthcare structure. J Algebraic Stat 13(2)
9. Schafer JB, Frankowski D, Herlocker J et al (2007) Collaborative filtering recommender systems. In: The adaptive web. Springer Berlin Heidelberg, pp 291–324
10. Sarwar B, Karypis G, Konstan J et al (2001) Item-based collaborative filtering recommendation algorithms. In: Proceedings of the 10th international conference on World Wide Web. ACM, pp 285–295
11. Parida B, Patra PK, Mohanty SP (2020) Use of social network for recommending job by applying machine learning techniques
12. Taunk K, Verma S, Swetapadma A (2019) A brief review of nearest neighbor algorithm for learning and classification. In: Proceedings of the international conference on intelligent computing and control systems (ICICCS)
13. Martinez-Gil J, Freudenthaler B, Natschläger T (2018) Recommendation of job offers using random forests and support vector machines. Conference paper

14. Min S-H, Han I (2005) Recommender systems using support vector machines. Conference Paper in lecture notes in computer science
15. Suharyadi J, Kusnadi A Design and development of job recommendation system based on two dominants on psycho test results using KNN algorithm. ISSN 2355-0082
16. Kamal Upreti, Mohammad Shahnawaz Nasir, Mohammad 12. Shabbir Alam, Ankit Verma, A.K. Sharma (2021) Analyzing real time performance in Vigil Net using wireless sensor network. Mater Today: Proc. https://www.sciencedirect.com/science/article/pii/S2214785321005812. ISSN 2214-7853. https://doi.org/10.1016/j.matpr.2021.01.490
17. Upreti K, Sharma AK, Vargis B, Sidhu RS (2020) An efficient approach for generating IRIS codes for optimally recognizing IRIS using multi objective genetic algorithm. Mater Today: Proc. https://www.sciencedirect.com/science/article/pii/S2214785320376525. ISSN 2214-7853. https://doi.org/10.1016/j.matpr.2020.10.085
18. Upreti K, Vargis BK, Jain R, Upadhyaya M (2021) Analytical study on performance of cloud computing with respect to data security. In: 5th international conference on intelligent computing and control systems (ICICCS), pp 96–101. https://doi.org/10.1109/ICICCS51141.2021.9432268
19. Bedi P, Upreti K, Rajawat AS, Shaw RN, Ghosh A (2021) Impact analysis of industry 4.0 on realtime smart production planning and supply chain management. In: IEEE 4th international conference on computing, power and communication technologies (GUCON), pp 1–6. https://doi.org/10.1109/GUCON50781.2021.9573563
20. Juneja N, Upreti K (2017) An introduction to few soft computing techniques to predict software quality. In: 2nd international conference on telecommunication and networks (TEL-NET), pp 1–6. https://doi.org/10.1109/TEL-NET.2017.8343581
21. Sharma A, Singh UK, Upreti K, Yadav DS (2021) An investigation of security risk and taxonomy of cloud computing environment. In: 2nd international conference on smart electronics and communication (ICOSEC), pp 1056–1063. https://doi.org/10.1109/ICOSEC51865.2021.9591954
22. Alanya-Beltran J et al (2022) Machine learning-based intelligent wireless communication system for solving real-world security issues. Secur Commun Netw 2022
23. Upreti K, Kumar N, Alam MS, Verma A, Nandan M, Gupta AK (2021) Machine learning-based congestion control routing strategy for healthcare IoT enabled wireless sensor networks. In: Fourth international conference on electrical, computer and communication technologies (ICECCT), pp 1–6. https://doi.org/10.1109/ICECCT52121.2021.9616864