



# SKILL BASED JOB RECOMMENDATION SYSTEM

Mansi Makhi, Urmila Jagdhane, Pooja Dhumal, Shruti Mane

Bharati Vidyapeeth's College of Engineering for Women Katraj-Pune;

Affiliated to Savitribai Phule Pune University

## Abstract:

Data science's subfield of machine learning focuses on creating algorithms that can learn from and make predictions based on the data. By suggesting jobs based on a candidate's skill set, recommender systems can be extremely useful in helping college graduates realize their career goals. There are many websites today that offer a ton of information on work opportunities, but this chore is very time consuming for students because they must sift through a lot of material to find the appropriate job. And many students are unaware of the types of jobs that would be best for them.

These days, the Technology industries are booming. This study looks at the users resume, compares the degree knowledge, soft skills, hard skills, and projects he has completed, and only then does the system suggest jobs for that user.

**Key words:** Recommender System, Support Vector Machine (SVM), Machine Learning etc.

## I. INTRODUCTION

According to a recent study, most college graduates find it challenging to choose their career field. A lot of engineers are working to make IT the dominant discipline. Thus, they are doing various online courses and aimlessly looking for employment. Many students today are interested in careers in the IT industry, but they are unsure of which field is best for them.

To get around this, Candidates require a job suggestion that considers their skills and suggests a

position that would be a good fit. Designing a system that recognises their skills is the answer. Using a machine learning algorithm, the system advised a job based on the skills of the individual.

Predicting the jobs that would be best for candidates is the main goal of the proposed work. It makes accurate predictions by using machine learning models to look for similarities between job descriptions and competencies. Any candidate can use this application to learn more about their ideal jobs and to develop both their soft skills and their hard skills. They will benefit from it because they won't have to waste time looking for work. Also, they can develop their expertise in that field and advance more quickly.

## PROBLEM DEFINATION

Regarding recommendation systems, the job recommendation system is the most important. The system must adjust to the novel algorithms and techniques that are available. The main objective of this work is to recommend a job for the candidates.

## II. LITERATURE REVIEW

[1] Collaborative filtering is one of the most effective and adequate technique used in recommendation. The fundamental aim of the recommendation is to provide prediction of the different items in which a user would be interested in based on their

preferences. Recommendation systems based on collaborative filtering techniques are able to provide approximately accurate prediction when there is enough data. User based collaborative filtering techniques have been very powerful and success in the past to recommend the items based on user's preferences. But there are also some certain challenges such as scalability and sparsity of data which increases as the number of users and items increases. In a large website, it is difficult to find the interested information in a certain time. But the recommendation system filters out information and items that are best suitable for us. Although there are different recommendation approaches, yet collaborative filtering technique is very popular because of the effectiveness. In this work, movie recommender system has been described, which basically uses item-based technique of collaborative filtering to provide the recommendations of items, which is dynamic and will learn from the positive feedback Background of recommender system.[2] In this paper, it is proposed that a dynamic user profile-based job recommender system. To address the challenge that the job applicants do not update the user profile in a timely manner, we update and extend the user profile dynamically based on the historical applied jobs and behaviors of job applicants. In particular, the statistical results of basic features in the applied jobs are used to update the job applicants. In addition, feature selection is employed in the text information of jobs that applied by the job applicant for extending the feature. Then a hybrid recommendation algorithm is employed according to the characteristics of user profiles for achieving the dynamic recommendation. [3] The recommendation system uses prior obtained information about the user to present user interested data. Personalized

(DJSP) used to be discussed separately isolated from process planning data given by CAD/CAM file. However, both the job shop scheduling and process planning are interdependent problems. Moreover, the new smart manufacturing systems requires more awareness between the available job floor data and machine availability. From that perspective, the integrated Job shop - process planning gained new scale of attention in recent studies to cope with smart factories. Within the used techniques phase, Evolution based algorithms proven to be a good

results aim to provide relevant information to the user based on the user's basic information or activity with the system. The user's basic information can be modeled into a user profile using ontology. Ontology is the systematic representation of various entities in a domain and the relationships between them.. The system collects basic information and models into a user profile. The dynamic aspects such as favorite jobs list and recently viewed jobs are then used as a source of data for the system. The recommendation algorithm works on the input given to present the list of relevant jobs to the user. [4] This paper presents the development of an online-based continuous assessment repository system designed to cater the needs for both lecturers and students. Conventional methods of notifying continuous assessment marks, e.g., test or assignment marks to students include displaying the students' marks on the bulletin boards or merely by returning the marked tests or assignments to students. While these are common practices, some lecturers are doing a better job by taking advantages of the Internet, i.e., by uploading their students' carry marks onto the student portal or by notifying students about their results via emails. The proposed system offers useful features for lecturers to upload their students' continuous-assessment marks onto an online database. This system provides an interface for the students to view their ongoing assessment marks throughout the semester via a web browser. Being an online-based system, it provides 24 hours access for students via any types of electronic devices with data connection that are able to run a web-browser. This system offers additional advantages when compared with other online systems, where it does not only display the raw marks obtained by students, but also the analysis of the marks and

solution for each separate problem, developing the same belief for the integration process.

### III. ALGORITHM

#### A. Support Vector Machine (SVM)

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. It is used for Classification problems in Machine Learning.

SVM chooses the extreme points/vectors that help in creating the hyperplane. These cases are called as support vectors, and the algorithm is termed as Support Vector Machine.

targeted final exam mark based on currently achieved marks. Having such a system will benefit not only the students but also for the faculty to progressively monitor students' performance throughout the semester. [5] Job Shop Scheduling Problem (JSSP) and its relatives; Flexible Job Shop Problem (FJSP) and Dynamic Job Shop Problem

## IV. PROJECT REQUIREMENTS

External interface requirement

### User Interface

Application Based On Job Recommendation System using Machine Learning.

### Hardware Interfaces:

1. Hardware : intel core
2. Speed : 2.80 GHz
3. RAM : 8GB
4. Hard Disk : 500 GB
5. Key Board: Standard Windows Keyboard

### Software Interfaces

1. Operating System: Windows 10(64 Bit)
2. IDE: Spyder
3. Programming Language : python version 3.7,3.8

## V.PROPOSED METHODOLOGY

1. Data Collection: Collect data on job descriptions and candidate skill sets. This data can be sourced from job portals, company websites, or by scraping data from other online sources.

2. Data Pre-processing: Pre-process the collected data by removing any irrelevant or duplicate data, and converting the data into a format that can be used for machine learning models

3. Feature Extraction: Select the most important features from the pre-processed data using techniques like PCA, LSA, or LDA. These features will be used as inputs for the machine learning model.

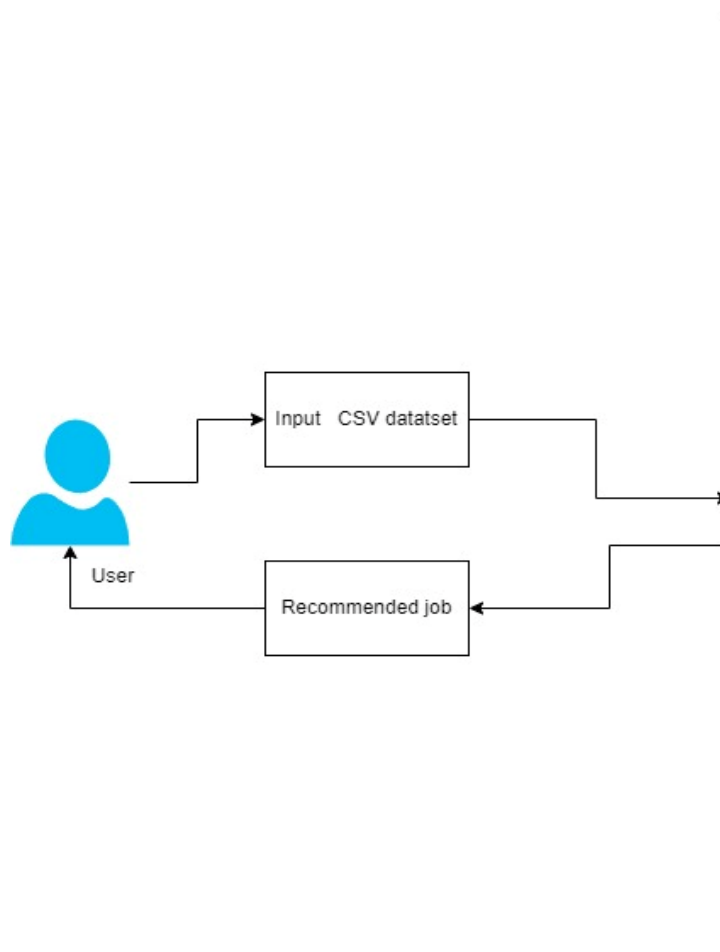
4. Train/Test Split: Split the pre-processed and feature-selected data into training and testing datasets. The training dataset will be used to train the SVM model, and the testing dataset will be used to evaluate the performance of the model.

5. SVM Model Training: Train an SVM model on the training dataset using the selected features. SVM is a supervised machine learning algorithm that can classify data into different categories based on the input features. The SVM algorithm seeks to maximize the margin between different classes of data points.

6. Model Evaluation: Evaluate the performance of the trained SVM model using the testing dataset. Metrics like accuracy, precision, recall, and F1 score can be used to evaluate the performance of the model.

7. Deployment: Deploy the trained SVM model as a skill-based job recommendation system. The system will take a candidate's skill set as input, and recommend relevant job listings based on the input.

## VI. SYSTEM ARCHIECTURE



The system approaches are classified into the following main four categories: Recommender System, Data Pre-Processing, Feature Extraction, Support

Vector Machine (SVM) Training Model and Testing. The detailed descriptions of different techniques are presented in the following paragraphs.

## 1. Recommender system

A recommender system is an algorithmic approach that suggests items or products to users based on their past behavior, preferences, and interests. The system analyzes user data such as purchase history, browsing behavior, ratings, and reviews to make personalized recommendations.

Recommender systems use different techniques to generate recommendations. Some of the most commonly used techniques include collaborative filtering, content-based filtering, and hybrid recommender systems.

## 2. Data pre-processing

By preprocessing data, we make it easier to interpret and use. This process eliminates inconsistencies or

duplicates in data, which can otherwise negatively affect a model's accuracy. Data preprocessing also ensures that there aren't any incorrect or missing values due to human error or bugs. In short, employing data preprocessing techniques makes the database completer and more accurate.

There are 3 types used in data pre-processing

1. Data Cleaning
2. Data Transformation
3. Data Reduction

## 3. Feature extraction

Feature extraction refers to the process of transforming raw data into numerical features that can be processed while preserving the information in the original data set. It gives better results than



applying machine learning directly to the raw data. Manual feature extraction requires identifying and describing the features that are relevant for a given problem and implementing a way to extract those features. Many machine learning practitioners believe that properly optimized feature extraction is the key to effective model construction.

#### 4. Support vector machines (svm)

Support Vector Machines (SVM) is a machine learning algorithm that can be used for classification or regression analysis. In this response, I will focus on SVM for classification. The basic idea of SVM is to find a hyperplane that separates the data into different classes. The hyperplane is selected such that it maximizes the margin between the two classes. The margin is the distance between the hyperplane and the closest data points from each class. SVM can be used for both linear and non-linear classification problems.

#### 5. Creating testing model

Creating a training module in machine learning involves designing and developing a set of materials and resources that will teach individuals how to build and train machine learning models. The training module should be structured in a way that is easy to understand and follow, and should cover key concepts and techniques in machine learning.

The following steps can be followed to create a training module in machine learning:

- 1) Define the learning objective
- 2) Identify the target audience
- 3) Develop the content
- 4) Determine the delivery method
- 5) Test and refine the module
- 6) Evaluate the module

Overall, creating a training module in machine learning requires careful planning, development, and evaluation to ensure that it meets the needs of the target audience and effectively teaches them the skills they need to build and train machine learning models.

#### 6. Testing

Testing in machine learning is an essential part of the development process as it allows for evaluating the performance of a model and assessing its ability to generalize to new, unseen data. In machine learning, testing involves splitting the dataset into training and testing sets. The training set is used to train the model, while the testing set is used to evaluate the performance of the model on new, unseen data. The testing set should be representative of the data that the model will encounter in the real world, and should be distinct from the training set to avoid overfitting.

### VII. FUTURE SCOPE

The main goal of this project is to provide skill-based jobs as per student/user inputs. For future work, we would like to examine and implement special ranking aggregation techniques for our ranking aggregations collaborative filtering frameworks. We'll apply these machine learning algorithms to modify content-based and collaborative techniques to obtain accurate results.

### VIII. CONCLUSION

To conclude about our system, we have made analysis of different research papers and algorithm implemented in it about recommendation systems. In our proposed system we have improvised and modified the recommendation systems. This job

Recommendation System has considered many parameters like marks, experience, skill etc. This framework facilitates the understanding of job recommendation process as well as it allows the use of a variety of recommendation methods according to the preferences of the job recommender system designer.

Future directions of our work will focus on performing a more exhaustive evaluation considering a greater number of methods and data as well as a comprehensive evaluation of the impact of each professional skill of a job seeker on the received job recommendation. Machine learning has been improvising the recommendation systems, also it brings more possibilities to improve performance of recommendation system.

## IX. 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-Kaun 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] 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.
- [6] L. Zahrotun, "Comparison jaccard similarity, cosine similarity and combined both of the data clustering with shared nearest neighbour method," Computer Engineering and Applications Journal. vol. 5. Pp. 11-18, 2016, doi:10.18495/comengapp.v5i1.160, 2016.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] Leah G. Rodriguez, Enrico P. Chavez, "Feature Selection for Job Matching Application using Profile Matching Model", IEEE 4th International Conference on Computer and Communication Systems (ICCCS), pp. 2-4 , Feb. 2019.