iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** III **Month of publication:** March 2021

DOI: https://doi.org/10.22214/ijraset.2021.33256

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue III Mar 2021- Available at www.ijraset.com

Job Recommendation System

Bhavya Chawla[1], Naitik Kansara[2], Sakshie Pathak[3], Mr. S. B. Nikam[4]
[1, 2, 3]*Bharati Vidyapeeth Deemed to be University College of Engineering, Pune*
[4]*Assistant professor, Department of Computer Engineering, Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune*

Abstract: Recommendation Systems are omnipresent on the web nowadays. Most websites today are striving to provide quality recommendations to their customers in order to increase and retain their customers. In this paper, we present our approaches to style employment recommendation system for a career based social networking websites. We take a bottom-up approach: we start with deeply understanding and exploring the info and gradually build the smaller bits of the system. We also consider traditional approaches of advice systems like collaborative filtering and discuss its performance. Our experiments show the efficacy of our approaches.
Keywords: Recommendation System, Collaborative Filtering, Content-based recommendation, Jobs recommendation.

I. INTRODUCTION

Rapid advancements in computer technology have led to a tremendous amount of user data. With the increase in the magnitude of data, the requirements and expectations of users have also increased. Companies have started predicting quality results on the basis of data and needs of the customers. This has led to increased customer satisfaction, and this quickly transformed into a multi-fold profit increase for companies. Online shopping websites like Amazon, Flipkart, and Alibaba recommend products, friendship social networking websites like Facebook and Google recommend friends, professional networking websites like LinkedIn and XING recommend jobs, Guidance systems, etc.
We believe the design of any good system is solely based on how deeply the developers understand the system and the data.
To accomplish the best result we have used a bottom-up approach, it works on collaborative filtering using the dataset to get the quality recommendation for the user.
Once we get a deeper perception of the data, we can help users by combining various aspects of a person's resume to the job description provided by the company and accquire the best possible recommendation.

II. DATASETS

Here we provide a brief description of the 2 major datasets used in the making of this model. They are 'Jobs', which contains all the jobs listed on the platform. This is all the essential information a candidate needs to know before they apply for a position.
The 'users' dataset is used to store all the information of the users of the platform looking for a job. The data taken is essential for determining an ideal job for a candidate. We have obtained the datasets from a job recommendation engine program posted on Github that was freely available to download and use.

A. Jobs Dataset

In this dataset, there is a list of all the job listings which have been posted. The attributes present in this dataset are JobID, WindowID, Job Title, Description, Requirements, City, State, Country, Zip Code, Start Date, and End Date.

B. Users Dataset

In this dataset, a list is created of the users that are present on the platform looking for jobs. The various fields present in this field are User ID, Window ID, Split, City, State, Country, Zip Code, Degree Type, Major, Graduation Date, Work History Count, Total Years Experience, Currently Employed, Managed Others, Managed How Many.

III. LIBRARIES AND TECHNOLOGY STACK

A. Matplotlib

This python library is used for creating visualizations in Python. This library was essential as bar graphs, box plots, and pie charts are commonly and preferably used to visualize data. The data is visualized before creating the test and train data sets so that a

uniform distribution of the data can be obtained that is similar to the actual dataset. This is helpful in obtaining similar accuracy percentages for the sample and actual dataset.

B. Seaborn

Seaborn is a statistical graphics library built on top of matplotlib for better visualization of data. The library has dataset oriented plotting functions that operate on data frames and arrays for processing and converting large datasets into plots and graphs. This was used along with matplotlib for data exploration and analysis.

C. Pandas

The data obtained is in the form of a csv file. The data had to be converted into a python data frame using pandas.

D. Numpy

Numpy is a scientific computation library in Python. It is the basic package used to form the basis of libraries like scikit-learn and SciPy. Since we use scikit learn, numpy is an library dependency we need to add for using scikit-learn and SciPy.

E. Ast

The ast module in python helps various applications to process trees of abstract grammar in Python. This library helps in determining what the correct grammar for various syntax looks like. This is a library is used to load the tree module while building the model.

F. Scipy

It is a basic module containing basic scientific, mathematical, and engineering packages. These are essential for developing every machine learning model as they have functions for carrying out basic operations in our model.

G. Scikit-learn

It is one of the most popular and one of the best modules to be used for building any machine learning model. It has simple and effective tools for performing EDA which is a very important part of building the model.

IV. METHODOLOGY

A. Using Jobs Dataset

The jobs dataset is used here to determine the various components that a recruiter wants to tell a candidate before they apply for a job. These parameters are extracted from the job listing of companies that post jobs on the portal. It includes all the essential information needed to match a job to a candidate and to recommend that particular job to the candidate.

B. Using Users Dataset

The data in this database is the information of the users. It is the basis on which the keywords from the job description is matched to a candidate's professional profile. This dataset is used to give us information on the users present in the platform. The keywords from the skills of people from their profiles are extracted and stored in this database and then used to map similarity for the recommendation engine.

C. Exploratory Data Analysis

There are 3 data frames where the attribute split is used. These are 'apps', 'user_history', and 'users'. The aim of the EDA is to distribute the data into train and test dataframes. The data attribute indicates whether the data can be used for training or for testing.
Here, a matrix of these attributes in every dataset is created based on their location information groups. After the data is distributed into train and test sets, we explored the data via bar graphs based on various combinations like country, state, and city.
Data is explored for the test as well as train data. We explore the data in such a way so that data is distributed in a uniform pattern across train and test data.

D. Collaborative Filtering

Collaborative Filtering is one of the most commonly used techniques for building recommendation systems where users are recommended items based on their preferences. It is used by some tech giants like Youtube, Amazon and Netflix.

We have used this technique to match keywords from a job description of a job posting and matched it to a description a professional has used to match the kind of jobs that are recommended to that person.

There are 2 major kinds of classification under Collaborative Filtering.

1) *Item- Item Similarity:* In an item-item similarity, an item is recommended to the user based on an item they have recently shown interest in. Like various brands of mobile phones advertising to a person after they look at a couple of models.
2) *User- User Similarity:* This is based on the idea that a similar set of users might like a similar product. Like a new target of monsoon sandals might target women between 18-15 living in a particular geographical region that is currently experiencing monsoon.

E. Using Cosine Similarity

Cosine_similar computes L2 normalized dot product of vectors. Euclidean normalization projects the dot product as the cosine of the angle between the points denoted by the vectors.

In this experiment, we take the cosine similarity of the 2 matrices created using the datasets. It is used to compare the text keywords in the job description to the skills in the person's job profile.

Working****

V. FUTURE WORK

Using better question sets and availability of more data can enhance the recommendation of jobs not just based on current profiles, but an alternative approach can be used to recommend jobs and job profiles to people using their personality types. This will make the platform more robust and create a separate segment of the target audience of fresh graduates.

The existing model can be enhanced by using facial recognition technologies for recommending jobs. The input keywords used will be the words scraped from the video job requirements put out by the recruiters and the candidates can apply via a video application. With the growing popularity of video interviews, this can be a very useful tool for hiring managers and recruiters using the platform. As well as the candidates can get a better knowledge of the company culture by looking at video job requirements that are posted.

A similar approach can be used for matching the jobs to candidates when there is technology available to successfully extract the proper keywords from the video applications.

VI. CONCLUSION

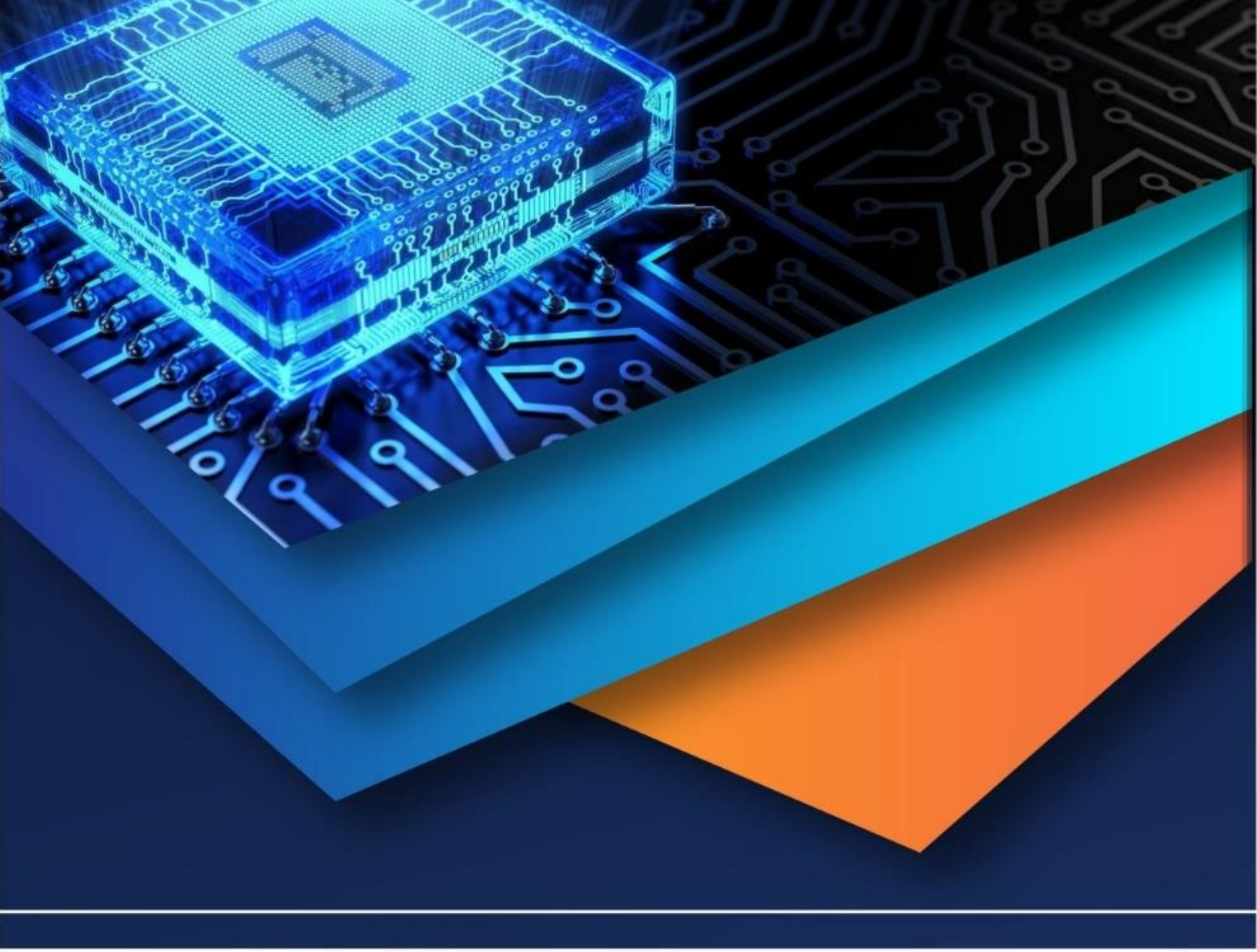
Since a large number of hirings have become online, job recommendation engines have become very essential in recommending the right jobs to the right candidate. A recommendation engine takes preference over a normal system hardcoded to recommend certain things to a person as it can recommend items in a way that is personalized.

For eg: A person living in a certain area, of a certain gender and in a certain age range might ideally use a product. So such products are probably useful for other people fitting that demographic. Hence that item may be recommended to the person.

In the system that we have implemented the items whose similarity we have taken into considerations are just skills. The skills a person has listed on their profile is matched to the skills listed by the employer for the same profile. This can be further customized to include the job locations, years of experiences, industry, and other similar parameters.

REFERENCES

[1] RecSys Challenge 2016 Official Website. Available at: http://2016.recsyschallenge.com/
[2] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen. Collaborative filtering recommender systems. The adaptive web 2007, pp. 291-324. Springer Berlin Heidelberg.
[3] Caret: classification and regression training. Available at: https: //cran.r-project.org/web/packages/caret/index.html
[4] AppliedPredictiveModeling: Functions and Data Sets for 'Applied Predictive Modeling'. Available at: https://cran.r-project.org/web/packages/AppliedPredictiveModeling/index.html
[5] Sonu K. M., Manoj Reddy, A Bottom-Up Approach to Job Recommendation System, Conference: Recommender Systems Challenge, 2016 International ACM



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY