Job Recommendation Systems
INFO 7390
Advance Data Science

Instructor: Ram HariHaran

Team Members
Aniruddha Tambe - 2101113
Sangram Vuppala - 2958963
Shubhankar Salvi - 1541699

May 2022

Introduction

Every industry has been greatly impacted by the rise of digital communication and the expansion of the internet. The hiring process is one such domain, in which a job seeker applies for a job by creating a profile on a job portal and detailing all of his or her work preferences. Each user's work preferences can be collected, and job recommendations can be made based on their choices.

A recommendation system is a method of determining a user's preferences based on their online behavior, previous purchases, or system history. Recommendation systems were used in e-commerce and online news, but now we are attempting to use them in the employment process.

Problem Statement

Can it be possible to develop an efficient recommender system for job seekers that recommends jobs based on the user's skill set and job domain?

Solution

We have developed recommendation systems using Collaborative Filtering, Content based Filtering and Model based filtering using two datasets, which are sourced from Stack Overflow for User Data and Dice.com for job postings in the IT Sector in the USA.

Collaborative Filtering:

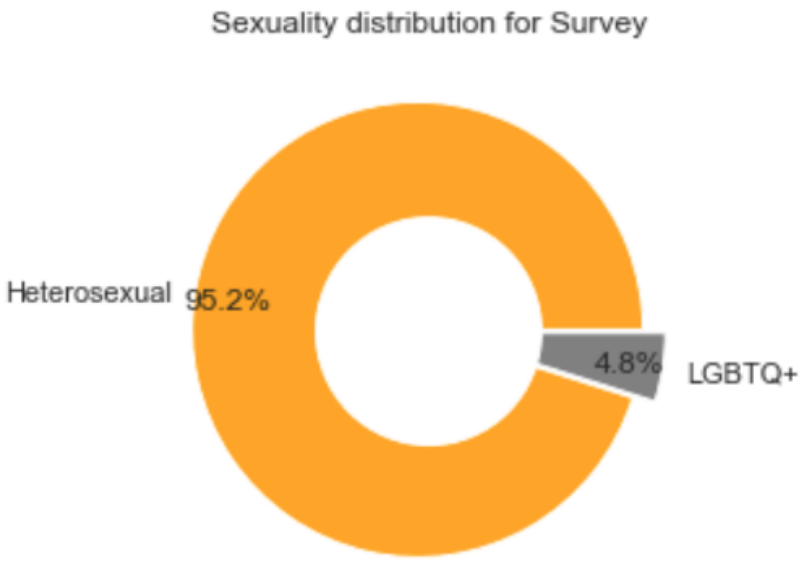
Collaborative Filtering is a technique that uses a user's human evaluations to uncover similarities between various users who have given comparable ratings to the same item. The memory-based nearest neighbor strategy is used to group people who share a common interest. As the volume of data expands, there will be a significant increase in the time it takes to generate recommendations.

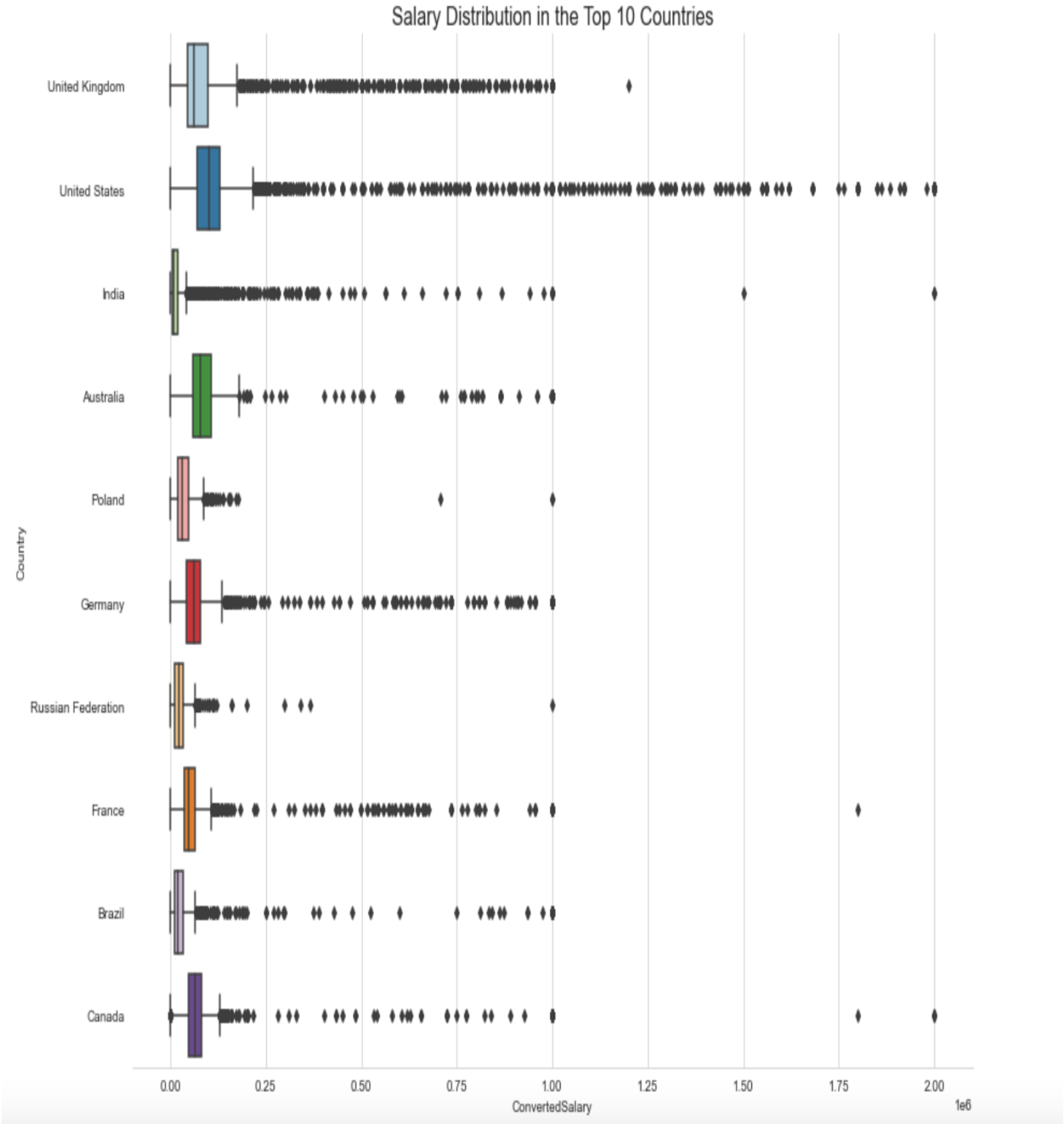
We have built dictionaries for Skills and then built user - user similarity matrix and then used collaborative filtering to recommend jobs to a job seeker based on the user skills.

Content Based Filtering:

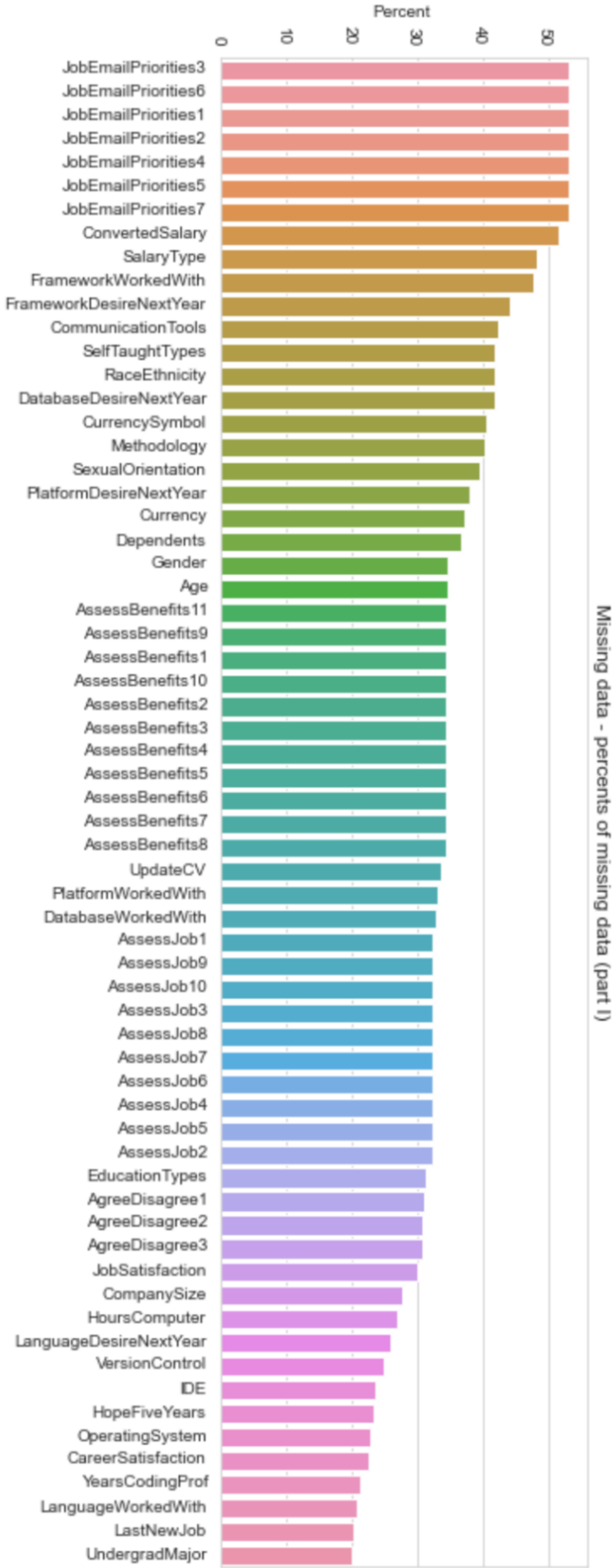
These are the most subjective and descriptive based filtering. Content-based filtering can also be called as attribute-based recommender as it uses the explicitly defined property of an item. It is an approach to an information retrieval or machine learning problem. The assumption made in content-based filtering is that the user prefers items with similar properties. Content-based filtering recommends items to the user whose properties are similar to the item which the user has previously shown interest.

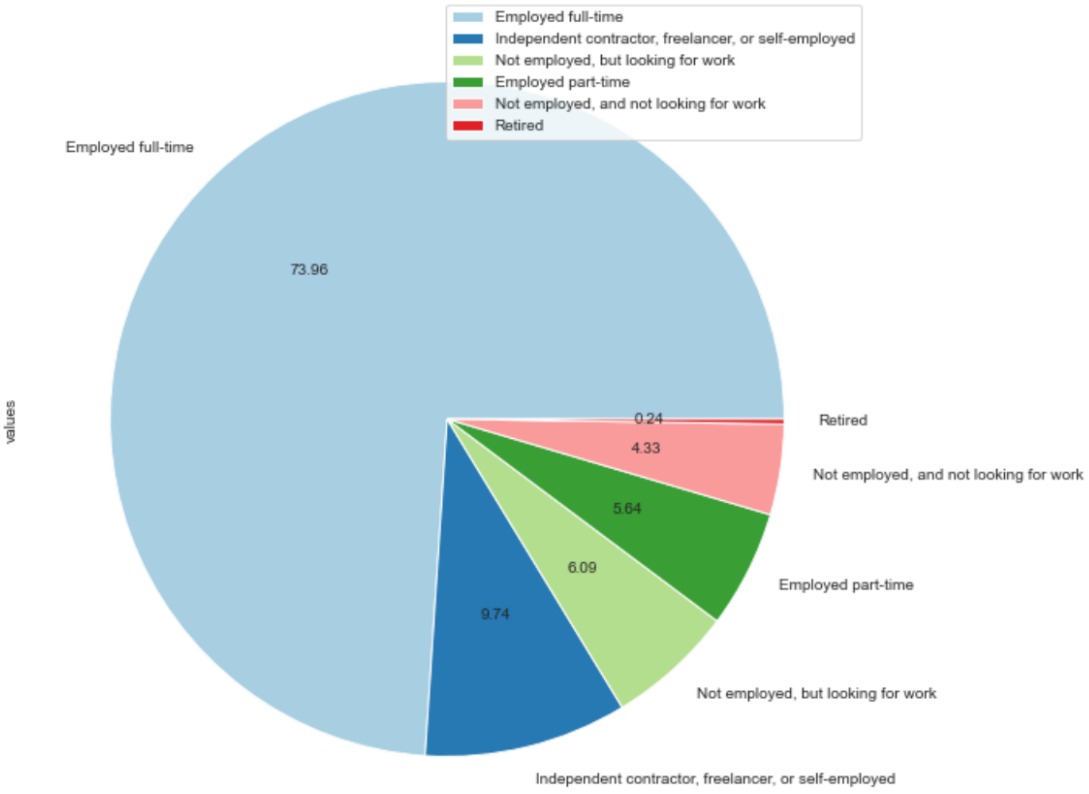
Visualizations



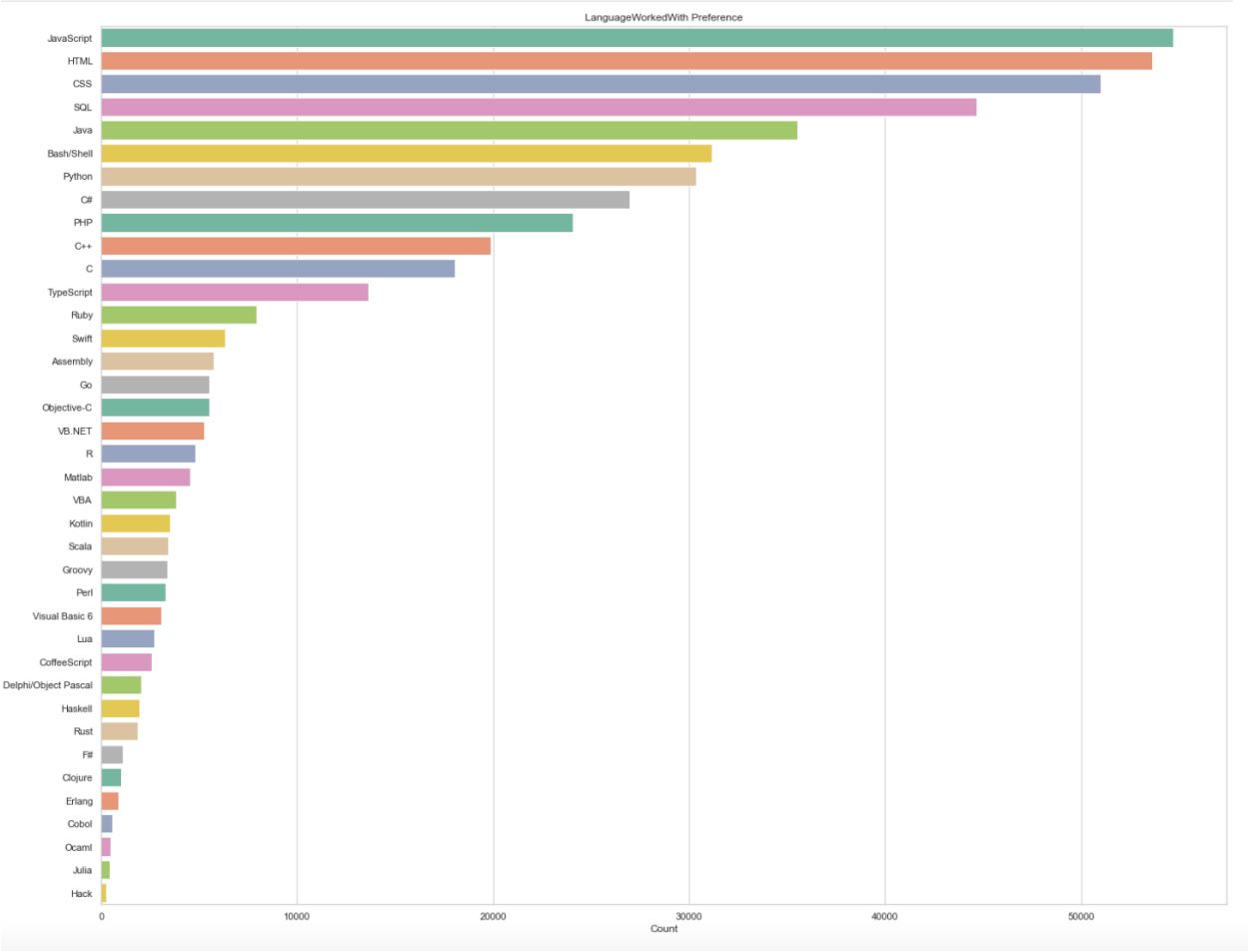

Salary Distribution in the Top 10 Countries
United Kingdom
United States
India
Australia
Poland
Germany
Russian Federation
France
Brazil
Canada
Country
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
ConvertedSalary
1e6


Missing data - percents of missing data (part II)
Percent
0
1
2
3
4
5
6
7
DevType
YearsCoding
FormalEducation
Student
Employment
Country
OpenSource
Respondent


Percent
0
10
20
30
40
50
JobEmailPriorities3
JobEmailPriorities6
JobEmailPriorities1
JobEmailPriorities2
JobEmailPriorities4
JobEmailPriorities5
JobEmailPriorities7
ConvertedSalary
SalaryType
FrameworkWorkedWith
FrameworkDesireNextYear
CommunicationTools
SelfTaughtTypes
RaceEthnicity
DatabaseDesireNextYear
CurrencySymbol
Methodology
SexualOrientation
PlatformDesireNextYear
Currency
Dependents
Gender
Age
AssessBenefits11
AssessBenefits9
AssessBenefits1
AssessBenefits10
AssessBenefits2
AssessBenefits3
AssessBenefits4
AssessBenefits5
AssessBenefits6
AssessBenefits7
AssessBenefits8
UpdateCV
PlatformWorkedWith
DatabaseWorkedWith
AssessJob1
AssessJob9
AssessJob10
AssessJob3
AssessJob8
AssessJob7
AssessJob6
AssessJob4
AssessJob5
AssessJob2
EducationTypes
AgreeDisagree1
AgreeDisagree2
AgreeDisagree3
JobSatisfaction
CompanySize
HoursComputer
LanguageDesireNextYear
VersionControl
IDE
HopeFiveYears
OperatingSystem
CareerSatisfaction
YearsCodingProf
LanguageWorkedWith
LastNewJob
UndergradMajor
Missing data - percents of missing data (part I)



EMPLOYMENT STATUS OF DEVELOPERS



Languages worked with preference

Future Scope

The future scope of this project can be to develop a full fledged application which can be deployed on Heroku and AWS Cloud platforms where users can input their skills and their current status of employment to help generate better recommendations.

Conclusion

We conclude that a job recommendation system using collaborative filtering is better than content based filtering with analysis of job description to recommend a job based on the user's skills and preferences.