RECIPROCAL JOB RECOMMENDER SYSTEM

KARAMJOT KAUR   RITIKA MOHANTY   STUTI AGGARWAL

RJRS

ISSUES ADDRESSED SO FAR

- Retrieve the information from recruiters and append it to the database
- Matching resumes to job profiles based on skills required for the job
- Using various recommendation techniques to implement reciprocal recommendation system

NEW LEARNINGS

- Collaborative Filtering

- Linear Kernels

Steps

01
Upload job description & appending job details to job dataset

02
Preprocessing the job description

03
Recommending resumes for a job

04
Matching jobseeker X to other jobseekers based on skills

05
Integrating the skills of other jobseekers

06
Recommending jobs to jobseeker X similar to the recommendations for other

UPLOAD

UPLOAD JOB DETAILS

```python
job_details = {}

job_details['company_name'] = input("Enter the company name: ")
job_details['job_title']  = input("Enter the job title: ")
job_details['job_des'] = input("Enter the job description: ")
job_details['job_location'] = input("Enter the job location: ")
job_details['exp_req']= input("Enter the experience required from candidate: ")
job_details['edu_req']= input("Enter the educational qualifications required: ")
job_details['cert_req']= input("Enter the exra certifications required (optional): ")
```

```
Enter the company name: Deloitte
Enter the job title: Data Analyst
Enter the job description: You will collaborate with case teams to gather requirements, specify, design, develop, deliver and support analytic solutions
Enter the job location: Hyderabad
Enter the experience required from candidate: 2 years
Enter the educational qualifications required: BCA/Bsc./B.Tech in Computer Science and related field
Enter the exra certifications required (optional): None
```

```
company_name  :  Smart Food Safe Solutions Inc
job_title  :  Software Testing Internship
job_des  :  Selected intern's day-to-day responsibilities include: 1. Reviewing software requirements and preparing test cases 2. Execute test cases (manual or automated)
job_location  :  Bengaluru
exp_req  :  0
edu_req  :  students or freshers can apply who: are available for full time (in-office) internship have relevant skills and interests
cert_req  :  -
{'company_name': 'company_name',
 'job_title': 'job_title',
 'job_des': 'job_des',
 'job_location': 'job_location',
 'exp_req': 'exp_req',
 'edu_req': 'edu_req',
 'cert_req': 'cert_req'}
```

```python
jobs =pd.read_csv('/content/drive/MyDrive/Colab/AI_proj.ample/jobs_dataset.csv')
jobs.tail()
```

	company_name	job_title	job_des	job_location	exp_req	edu_req
30	Redington India Ltd.	Product Technical Specialist	The Product Technical Specialist is responsibl...	New Delhi	Experience of around 0 to 1 years as a presale...	Graduate in any discipline Engineering with C...
31	Cyient	Developer Associate - FULL STACK DEVELOPER	A full stack developer requires skills such as...	Bengaluru	NaN	NaN
32	Cvent	Software Testing Engineer	To ensure success as a software test engineer,...	West Delhi	Experience working with popular operating syst...	Bachelor's degree in computer science, softwar...
33	Userfacet	Software Engineer	We are looking for a skilled software engineer...	Bangalore	5 to 7 years of software engineering or softwa...	A bachelor's degree in computer science, softw...
34	Dell	Information Security Operations Analyst	NaN	Delhi	three or more of experience as Network Intrusi...	Bachelor's Degree or equivalent in Computer/El...

PREPROCESS

```
resumes['test']=resumes['skills'].apply(string_process)
resumes['test']
```

```
0      engineering java tensorflow research python o...
1      digital marketing excel photoshop python mark...
2      engineering java database documentation softw...
3      php computer science java website presentatio...
4                              english email c
5      data analytics seaborn matplotlib pandas pyth...
6      mysql purchasing windows tableau linux schedu...
7      design certification engineering process emai...
8      policies updates mysql engineering forecastin...
9      java email html technical newspaper english c...
10     salesforce finance mis seo website design crm...
11     php training mysql java engineering proposal ...
12     mysql vendors engineering distribution querie...
13     excel mysql engineering analytics sales pytho...
14     c html technical javascript english administr...
15     mechanical engineering reporting research tec...
16     scripting java oracle windows sql english css...
17     sports engineering scrum payments erp windows...
18     ai content mysql c oracle mining banking net ...
19     engineering windows java social media hotel c...
20     policies engagement agile engineering plan sc...
Name: test, dtype: object
```

NGRAMS on Job Description

```python
data_lower = []
for s in test_job['job_des']:
    data_lower.append(s.lower())

test_job_des = ""
test_job_des = test_job_des.join(data_lower)
j=1
for i in test_job_des:
  if j%40:
    print(i, end='')
  else:
    print(i)
  j = j+1
```

selected intern's day-to-day responsibil
ities include: 1. reviewing software req
uirements and preparing test cases 2. ex
ecute test cases (manual or automated) a
nd analyze results 3. evaluate product c
ode according to specifications 4. repor
t bugs and errors to development teams.
skills req: basic knowledge of database/
sql;knowledge of any scripting language
like vbscript, javascript, c#, shell/per
l scripting

<class 'dict'>
95
{'selected': 73,
 'intern': 36,
 'day': 23,
 'to': 89,
 'responsibilities': 65,
 'include': 34,
 'reviewing': 69,
 'software': 79,
 'requirements': 63,
 'and': 4,
 'preparing': 55,
 'test': 87,
 'cases': 16,
 'execute': 32,
 'manual': 46,
 'or': 51,
 'automated': 10,
 'analyze': 2,
 'results': 67,
 'evaluate': 30,
 'product': 57,
 'code': 19,
 'according': 0,
 'specifications': 81,
 'report': 59,
 'bugs': 14,
 'errors': 28,
 'development': 26,

TF-MATRIX of Token Counts

```python
tf_count = v.transform(resumes['test'])
tf_ud = tf_count.toarray()
print(type(tf_count))
print(tf_ud.shape)
print(tf_ud)
```

```
<class 'scipy.sparse.csr.csr_matrix'>
(21, 95)
[[0 0 0 ... 0 0 0]
 [0 0 0 ... 0 0 0]
 [0 0 0 ... 0 0 0]
 ...
 [0 0 0 ... 0 0 0]
 [0 0 0 ... 0 0 0]
 [0 0 0 ... 0 0 0]]
```

MATCH RESUMES to JOBS

```
nbrs = NearestNeighbors(n_neighbors=3, n_jobs=-1).fit(tf_count)
```

```
def getNearestN(query):
    q_tf_count = v.transform(query)
    distances, indices = nbrs.kneighbors(q_tf_count)
    return distances, indices
```

```
distances, indices = getNearestN([test_job_des])
```
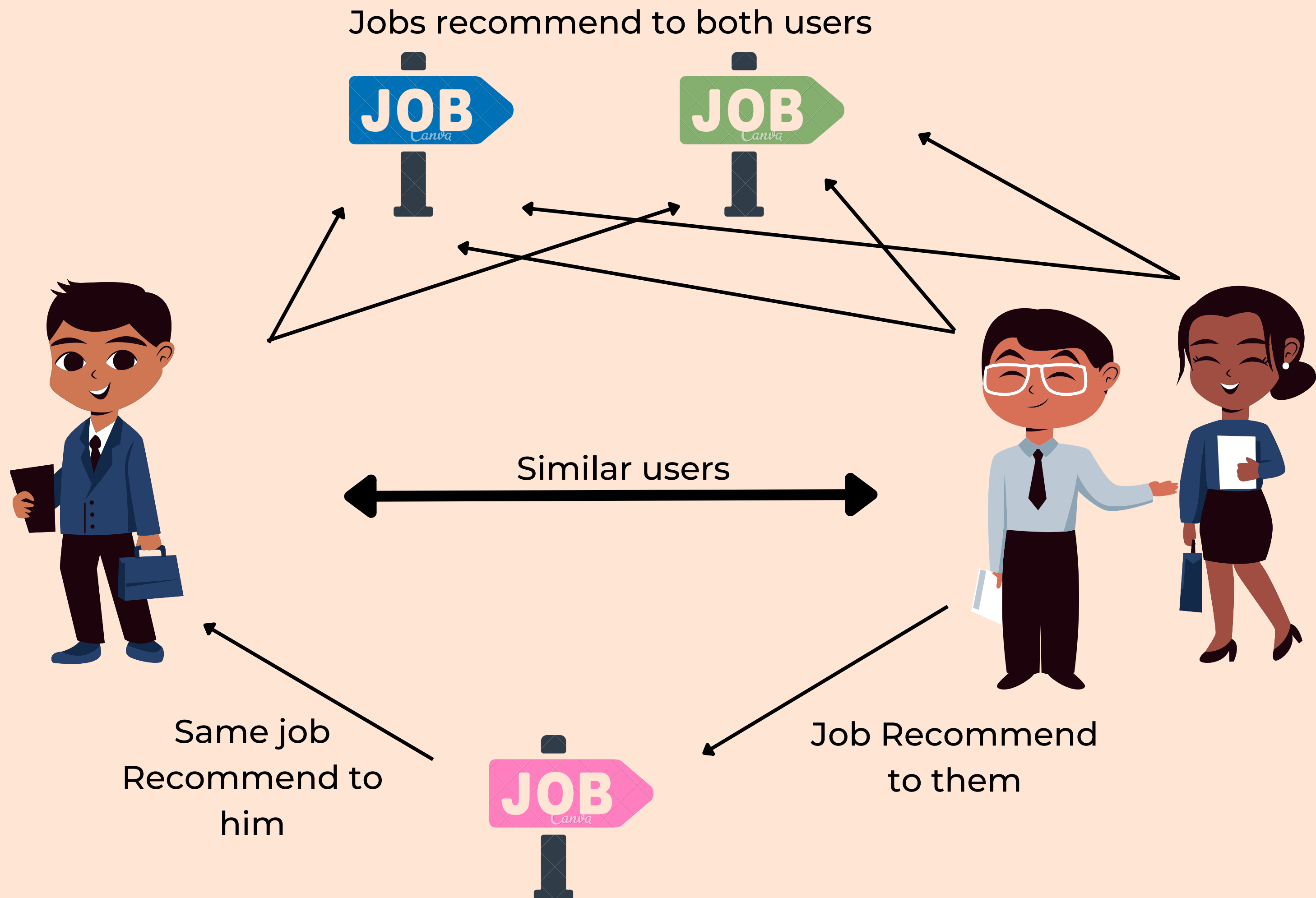
```
len(distances[0])
```

3

```
print("Distances          ", "Indices")
for i in range(0,len(distances[0])):
    print(distances[0][i],"    ", indices[0][i])
```

```
Distances              Indices
11.045361017187261        17
11.180339887498949        19
11.445523142259598        16
```

RECOMMENDING TOP 3 RESUMES

	name	email	mobile_number	skills	college_name	degree	designation	experience	company_names	no_of_pages	total_experience	test
17	Lavanya	lavanya.adhinarayanan@gmail.com	988403855	['Sports', 'Engineering', 'Scrum', 'Payments',...	Dayal Singh College, University of Delhi	B.Sc. Hons	NaN	['*', '• Worked with Insoft.com Private Ltd, C...	NaN	9.0	0.00	sports engineering scrum payments erp windows...
20	Aneet	aneetatomy23@gmail.com	8129945395	['Engineering', 'Windows', 'Java', 'Social med...	NSIT, New Delhi	B.Tech (Computer Science)	NaN	['PageMe Consultancy [Digital Marketing Analys...	NaN	2.0	0.25	engineering windows java social media hotel c...
16	Sireesha	karakondavenkat@gmail.com	9182382214	['Scripting', 'Java', 'Oracle', 'Windows', 'Sq...	IIT Madras	NaN	NaN	NaN	NaN	2.0	0.00	scripting java oracle windows sql english css...

COLLABORATIVE RECOMMENDATION SYSTEM

Jobs recommend to both users

JOB

JOB

JOB

Similar users

Same job Recommend to him

Job Recommend to them

RESUME TO RESUME MATCHING

```
[153] # CREATING A BAG OF WORDS

      from sklearn.feature_extraction.text import TfidfVectorizer

      # Convert a collection of raw documents to a matrix of TF-IDF features.
      # tf-idf means term-frequency times inverse document-frequency
      # tf-idf for a term t of a document d in a document set is tf-idf(t, d) = tf(t, d) * idf(t)
      # idf is computed as idf(t) = log [ n / df(t) ] + 1

      tdif = TfidfVectorizer(stop_words='english')
      # print(type(tdif))
      tdif_matrix = tdif.fit_transform(resumes['test'])
      tdif_matrix.shape

      (21, 167)
```

The **sklearn.metrics.pairwise** submodule implements utilities to evaluate pairwise distances or affinity of sets of samples.

Kernels are measures of similarity, i.e. s(a, b) > s(a, c) if objects a and b are considered "more similar" than objects a and c. A kernel must also be positive semi-definite.

The function linear_kernel computes the linear kernel, that is, a special case of polynomial_kernel with degree=1 and coef0=0 (homogeneous). If x and y are column vectors, their linear kernel is:

$$k(x, y) = x^\top y$$

```python
from sklearn.metrics.pairwise import linear_kernel

# considers not only the similarity between vectors under the same dimension,
# but also across dimensions. When used in machine learning algorithms,
# this allows to account for feature interaction.


cosine_sim = linear_kernel(tdif_matrix, tdif_matrix)


#  tf-idf functionality in sklearn.feature_extraction.text can produce
# normalized vectors, in which case cosine_similarity is equivalent to
# | linear_kernel, only slower


indices = pd.Series(resumes.index, index=resumes['name']).drop_duplicates()
indices
```

```
name
Lakshay Yadav            0
STUTI                    1
KARAMJOT KAUR            2
RITIKA MOHANTY           3
Email Contact           4
Akash Shetty            5
ANURAG CHAUHAN          6
R.Vimalraj EDUCATION    7
```

RECOMMENDING SIMILAR RESUMES

```python
def get_recommendations(title,cosine_sim=cosine_sim):
    idx = indices[title]
    # print(idx)
    sim_scores = list(enumerate(cosine_sim[idx]))
    # print(sim_scores)
    sim_scores = sorted(sim_scores, key=lambda x:x[1], reverse=True)
    # we don't how this 'key=lambda x:x[1]' works
    # print(sim_scores)
    sim_scores = sim_scores[1:4]

    tech_indices = [i[0] for i in sim_scores]
    return resumes.iloc[tech_indices]
```

```python
sim_res = get_recommendations('Lakshay Yadav')
sim_res
```

	name	email	mobile_number	skills	college_name	degree	designation	experience	company_names	no_of_pages	total_experience	test
2	KARAMJOT KAUR	karamjotk03@gmail.com	9654226396	['Engineering', 'Java', 'Database', 'Documenta...	NaN	NaN	NaN	NaN	NaN	2.0	0.00	engineering java database documentation softw...
10	BL Karnataka	Varunbl.vrun@gmail.com	1767605	['Salesforce', 'Finance', 'Mis', 'Seo', 'Websi...	NaN	NaN	['Bot Developer', 'Data Processing Analyst']	['Data Processing Analyst', 'NielsenIQ', 'Oct ...	['Microsoft']	2.0	1.42	salesforce finance mis seo website design crm...
11	Ansari	ansarialiasgar40@gmail.com	9779811193	['Php', 'Training', 'Mysql', 'Java', 'Engineer...	NaN	[BE in Computer Science and Engineering']	['Playing Cricket']	['15-Sep-', '2021 - 19-', 'Oct-2021']	NaN	1.0	0.00	php training mysql java engineering proposal ...

INTEGRATING SKILLS OF SIMILAR JOBSEEKERS

```
[168] sim_res['test']

        2       engineering java database documentation softw...
        10      salesforce finance mis seo website design crm...
        11      php training mysql java engineering proposal ...
        Name: test, dtype: object


[172] sim_skills = ""
      for res in sim_res['test']:
        sim_skills = sim_skills +" "+ res

      sim_skills


      '  engineering java database documentation software engineering sql communicati
      ebsite design crm computer science data analytics reporting excel research migr
      ecurity   php training mysql java engineering proposal technical computer scien
```

MATCHING JOBS TO JOBSEEKERS WITH SIMILAR SKILLS

```python
nbrs = NearestNeighbors(n_neighbors=5, n_jobs=-1).fit(tf_count)
```

```python
def getNearestN(query):
    q_tf_count = v.transform(query)
    distances, indices = nbrs.kneighbors(q_tf_count)
    return distances, indices
```

```python
distances, indices = getNearestN([jobseeker_skills])
```

```python
print("Distances          ", "Indices")
for i in range(1,len(distances[0])):
    print(distances[0][i],"    ", indices[0][i])
```

```
Distances          Indices
1.0        0
1.0        14
1.0        28
1.4142135623730951      19
```

TOP 5 JOBS FOR JOBSEEKER BASED ON SIMILAR JOBSEEKERS

TOP N JOBS RECOMMENDED TO LAKSHAY YADAV USING COLLABORATIVE FILTERING

```
[180] type(indices)
      for i in indices:
          rec_jobs = jobs[['job_title', 'company_name','job_location']].iloc[i]
      rec_jobs
```

	job_title	company_name	job_location
25	Associate Developer (Python with Machine Learn...	Capgemini	Pune
0	Software Testing Internship	Smart Food Safe Solutions Inc	Bengaluru
14	NLP / Machine Learning	Oracle Financial Services Software	Pune
28	Software Testing Engineer	Hexaware Technologies	Navi Mumbai
19	Microsoft Technology Project Manager	AD Infosystem.LLC	Delhi

FUTURE WORK

1. Checking and improving the accuracy of the model.

2. Building the user interface for the project.

3. Create a hybrid recommendation model using different recommendation technologies.

REFERENCES

- https://scikit-learn.org/stable/modules/metrics.html#:~:text=This%20is%20called%20cosine%20similarity,represented%20as%20tf%2Didf%20vectors

- https://www.analyticsvidhya.com/blog/2022/02/introduction-to-collaborative-filtering/