

Reciprocal Job Recommendation System

Make a match that fulfils the mutual needs of both users

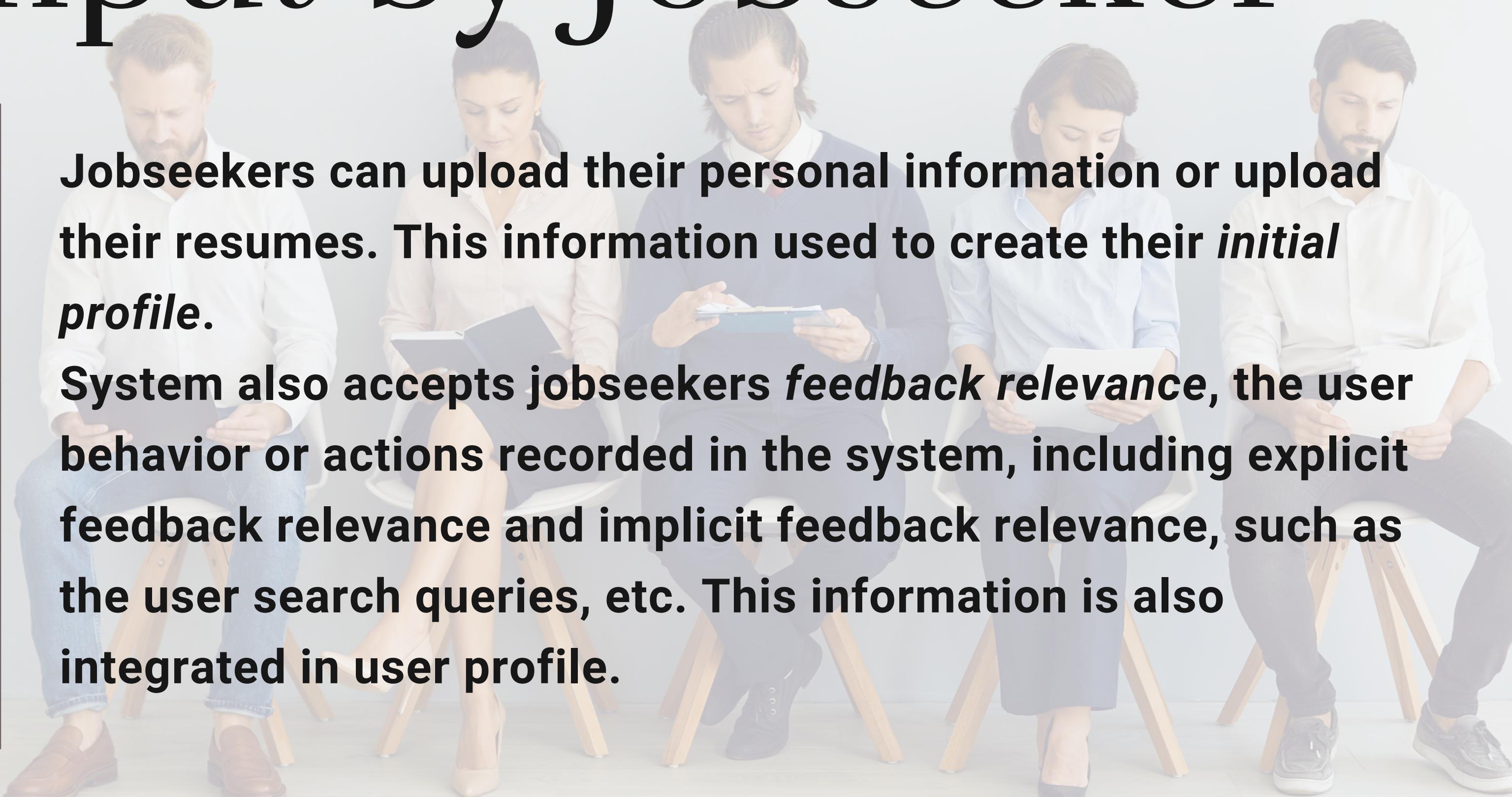
Karamjot Kaur
Ritika Mohanty
Stuti Aggarwal



Problem Statement

Reciprocal Job-Candidate Recommender System should be capable of retrieving a list of talented candidates that meet the requirement of a recruiter as well as should offer job advertisements in a sequence for candidates based on the preference information provided by the users.

Input by Jobseeker



Jobseekers can upload their personal information or upload their resumes. This information used to create their *initial profile*.

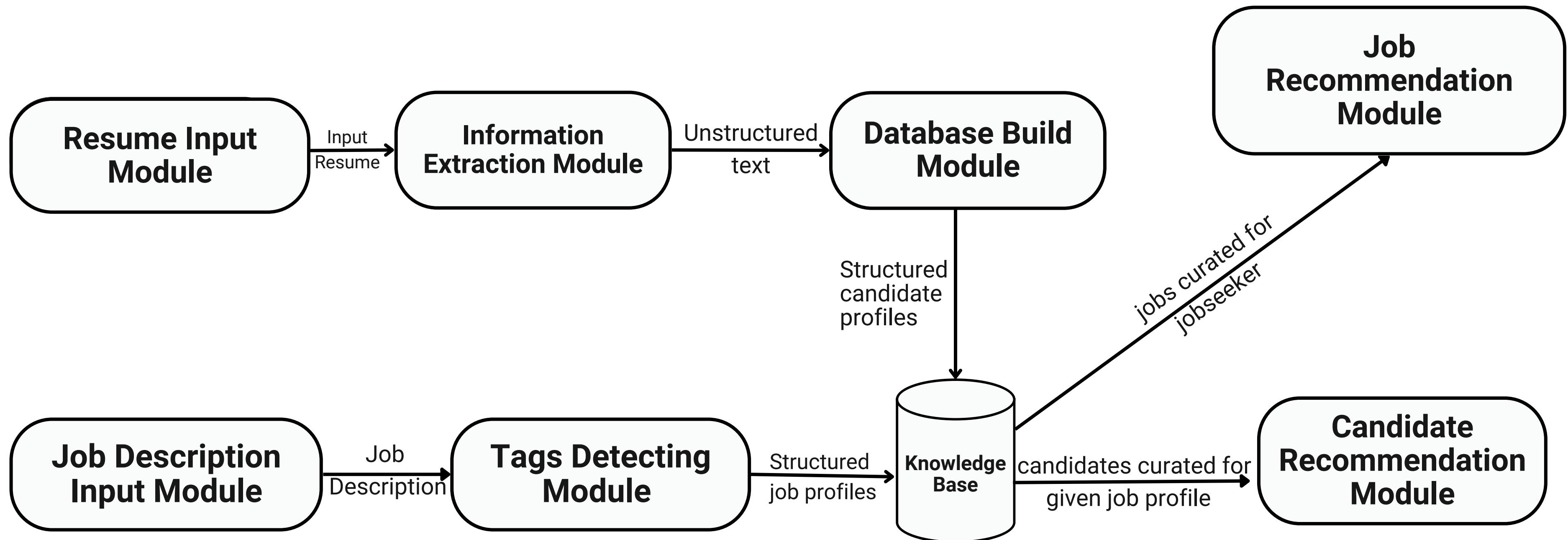
System also accepts jobseekers *feedback relevance*, the user behavior or actions recorded in the system, including explicit feedback relevance and implicit feedback relevance, such as the user search queries, etc. This information is also integrated in user profile.

Input by Recruiter

Recruiter must enter the job description of the position they are offering, which may include:

- **Position Offered:** Position offered for the job
- **job location:** After getting the job, where the candidate will be working
- **Employer/Company:** Company that is offering the job
- **Sector:** major discipline in which the job is offered
- **Experience:** Required experience for the job
- **Education:** Required education for the job
- **Certificate:** Required certificates for the job

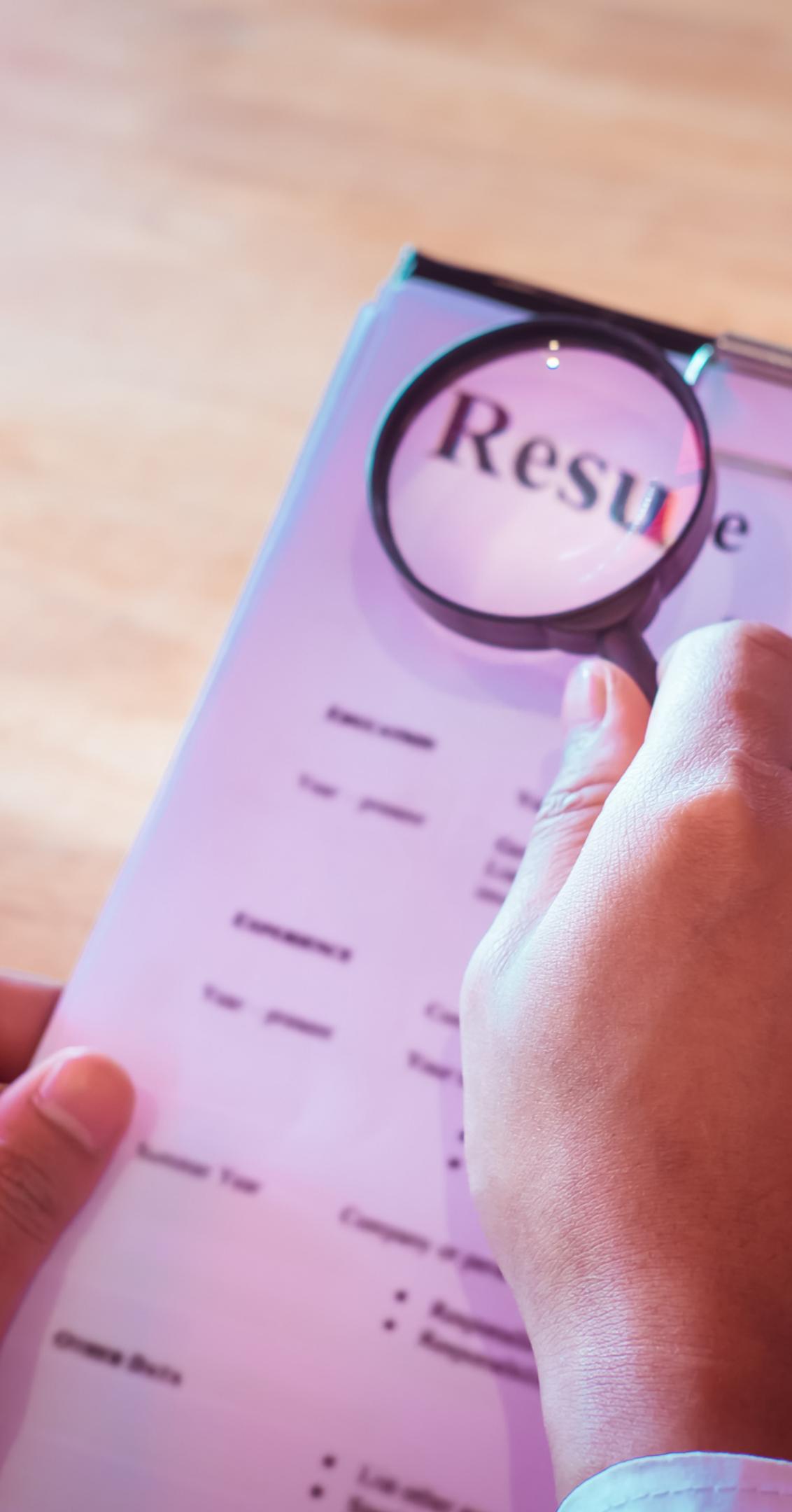
WORK FLOW DIAGRAM



To build the system, we must address these issues

How to extract the information of jobseekers which can effectively contribute in matching jobs and jobseekers?

Which recommendation technology is used in the job recommender system based on user profile?



Resume Information Extraction with Cascaded Hybrid Model

Resume shares a document-level hierarchical contextual structure where the related information units usually occur in the same textual block, and text blocks of different information categories usually occur in a relatively fixed order.

Structure of cascaded hybrid model

First Layer is composed of consecutive general information blocks such as Personal Information, Education etc.

Then within each general information block, detailed information pieces can be found, e.g., in Personal Information block, detailed information such as Name, Address, Email etc. can be further extracted.

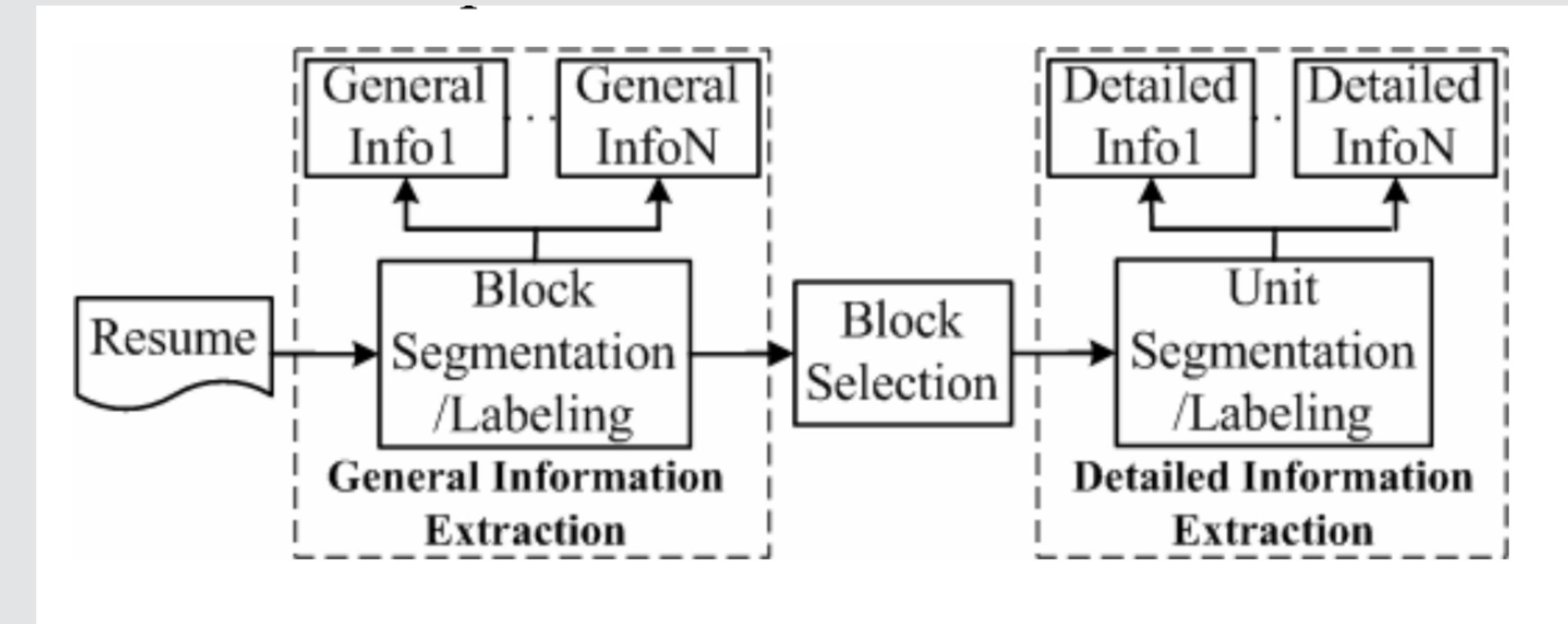


Image source Association for Computational Linguistics, Association for Computational Linguistics (2005)

Experiment Results show that with the cascaded model, the precision is greatly improved compared with the flat model with identical IE method. Although there is some loss in recall, the average F-score is still largely improved in the cascaded model.

Adam Wang (Male)

XXXX Company of Bejing,
Beijing City, 100007
1364-110-XXX
wangXXX@hotmail.com

Education Background

From Sept. 2000 to Apr. 2003, I got master degree from University of XXX in computer software engineering.

From Sept. 1996 to July. 2000, I got bachelor degree from School of XXX and major in computer science and technology.

Experience

From March 2003 to now, working on Human Face Recognition System in XXXX Company of Beijing

From June 2001 to March 2003, working on Content-Based Intelligent Image Retrieval System in Research Center of XXX Company

From Sept. 2000 to May 2001, working on Intelligent Highway Distress Detection System in National Lab. Of XXX University

Interests

Reading, music, and jogging

<Personal Information>

```
<Name>Adam Wang</Name> <Gender>Male</Gender>
<Address>XXXX Company of Bejing, Beijing City </Address>
<Zip code>100007</Zip code>
<Mobile>1364-110-XXX</Mobile>
<Email>wangXXX@hotmail.com</Email>
</Personal Information>
```

<Education>

```
<Graduation School> University of XXX</Graduation School>
<Major>Computer Software Engineering</Major>
<Degree>Master</Degree>
<Graduation School>School of XXX</Graduation School>
<Major>Computer Science and Technology</Major>
<Degree>Bachelor</Degree>
</Education>
```

<Research Experience>

```
From March 2003 to now, working on Human Face Recognition
System in XXXX Company of Beijing
From June 2001 to March 2003, working on Content-Based
Intelligent Image Retrieval System in Research Center of
XXX Company
From Sept. 2000 to May 2001, working on Intelligent
Highway Distress Detection System in National Lab. Of XXX
University
</Research Experience>
```

```
<Interests>Reading, music, and jogging</Interests>
```

Figure 1. Example of a resume and the extracted information.

Resumé Parsing

LAKSHAY YADAV

+91 7982993455 | www.lakshayyadav.com | lakshay.mcs21.du@gmail.com | in Lakshay Yadav | Lakshay Yadav

Education

Department of Computer Science, University of Delhi Master's in Computer Science (8.8 CGPA)	Dec. 2021 – present Vishwavidyalaya, Delhi
Atma Ram Sanatan Dharma College, University of Delhi Bachelor's in Computer Science (8.9 CGPA)	Jul. 2018 – Jun. 2021 Dhaua Kuan, Delhi

Certification & Training

Data Structures & Algorithms in Python Coding Ninjas	Sep. 2020 – Jan. 2021 online
Summer training on Ethical Hacking Internshala	May 2019 – Jul. 2019 online
Certified in Cyber Security DICC	Jun. 2018 – Aug. 2018 offline

Projects

Web Based Face Recognition Application TensorFlow, Scikit-learn, Face recognition, Flask, Html, CSS	Feb 2022
• It is a login mechanism which uses face recognition and it is trained over a set of real and fake faces using neural networks.	
• Features: It can distinguish between real and fake faces.	
Virtual Painter python,openCV,AR (Augmented Reality)	Sep 2021
• A canvas based platform on which we can draw just by motion of the hand.	
• Human gestures are perceived via sight, and computer vision is used to research different gestures.	
Airbnb Clone React.js	May 2022
• It is build upon - React/Tailwind CSS/ MapBox integration - Next.js/ Calendar Search - Next.js/ SSR / SSG/Date-picker/ Maps / API.	
• In this webapp we can search any location to book room and it has integrated map showing location of places found in that region.	
Cybersecurity Projects Python, scapy module, netfilterqueue	Jun 2020
• Python based cybersecurity projects. The projects are as follows:	
1. Backdoor and Listener Using these scripts we can get an unauthorised access to systems.	
2. Keylogger Using this script we can record all the keystrokes that happened on a system.	
3. Replace Download Using this script we can change specific ongoing downloading file on a system.	
4. Packet Sniffer Using this script we can read incoming and outgoing network data.	

Skills

- **Languages:** C++, Python, Html, CSS, SQL, Java Script
- **Cyber Security:** Network Security, Website Security, Social Engineering
- **Theoretical:** Data Structures and Algorithm, Data Base Management System, Object Oriented Programming
- **Libraries/Framework:** React.js, Next.js, openCV, Scapy, Face Recognition

“
LakshayYadav(cid:211)+917982993455(cid:140)www.lakshayyadav.comlakshay.mcs21.du@gmail.comflLaks
hayYadav(cid:135)LakshayYadavEducationDepartmentofComputerScience,UniversityofDelhiDec.2021–
presentMaster'sinComputerScience(8.8CGPA)Vishwavidyalaya,DelhiAtmaRamSanatanDharmaCollege,
UniversityofDelhiJul.2018–
Jun.2021Bachelor'sinComputerScience(8.9CGPA)DaulaKuan,DelhiCertification&TrainingDataStructure
s&AlgorithmsinPythonSep.2020–
Jan.2021CodingNinjasonlineSummertrainingonEthicalHackingMay2019–
Jul.2019InternshalaonlineCertifiedinCyberSecurityJun.2018–
Aug.2018DICCofflineProjectsWebBasedFaceRecognitionApplication|TensorFlow,Scikit-
learn,Facerecognition,Flask,Html,CSSFeb2022•Itisaloginmechanismwhichusesfacerecognitionanditistra
inedoverasetoffrealandfakefacesusingneuralnetworks.●Features:Itcan distinguishesbetweenrealandfakefac
es.VirtualPainter|python,openCV,AR(AugmentedReality)Sep2021•Acanvasbasedplatformonwhichwecan
drawjustbymotionofthehand.●Humangesturesareperceivedviasight, andcomputervisionisusedtoresearc
hdifferentgestures.AirbnbClone|React.jsMay2022•Itisbuildupon-
React/TailwindCSS/MapBoxintegration-Next.js/CalendarSearch-Next.js/SSR/SSG/Date-
picker/Maps/API.●Inthiswebappwecansearchanylocationtobookroomandithasintegratedmapshowingloc
ationofplacesfoundinthatregion.CybersecurityProjects|Python,scapymodule,netfilterqueueJun2020•Pyt
honbasedcybersecurityprojects.Theprojectsareasfollows:1.BackdoorandListenerUsingthesescritswecan
getanunauthorisedaccesstosystems.2.KeyloggerUsingthisscriptwecanrecordallthekeystrokesthathappen
edenasystem.3.ReplaceDownloadUsingthisscriptwecanchangespecificongoingdownloadingfileonasyste
m.4.PacketSnifferUsingthisscriptwecanreadincomingandoutgoingnetworkdata.Skills●Languages:C++,Pyt
hon,Html,CSS,SQL,JavaScript●CyberSecurity:NetworkSecurity,WebsiteSecurity,SocialEngineering●Theo
retical:DataStructuresandAlgorithm,DataBaseManagementSystem,ObjectOrientedProgramming●Librari
es/Framework:React.js,Next.js,openCV,Scapy,FaceRecognitionLakshayYadav(cid:211)+917982993455(cid:140)
www.lakshayyadav.comlakshay.mcs21.du@gmail.comflLakshayYadav(cid:135)LakshayYadavEducatio
nDepartmentofComputerScience,UniversityofDelhiDec.2021–
picker/Maps/API.●Inthiswebappwecansearchanylocationtobookroomandithasintegratedmapshowingloc
ationofplacesfoundinthatregion.CybersecurityProjects|Python,scapymodule,netfilterqueueJun2020•Pyt
honbasedcybersecurityprojects.Theprojectsareasfollows:1.BackdoorandListenerUsingthesescritswecan
getanunauthorisedaccesstosystems.2.KeyloggerUsingthisscriptwecanrecordallthekeystrokesthathappen
edenasystem.3.ReplaceDownloadUsingthisscriptwecanchangespecificongoingdownloadingfileonasyste
m.4.PacketSnifferUsingthisscriptwecanreadincomingandoutgoingnetworkdata.Skills●Languages:C++,Pyt
hon,Html,CSS,SQL,JavaScript●CyberSecurity:NetworkSecurity,WebsiteSecurity,SocialEngineering●Theo
retical:DataStructuresandAlgorithm,DataBaseManagementSystem,ObjectOrientedProgramming●Librari
es/Framework:React.js,Next.js,openCV,Scapy,FaceRecognition”

PDF to Text

Text Extraction

```
import pdfplumber
str1 = ""
with pdfplumber.open("/content/resume.pdf") as pdf:
    for pgs in pdf.pages:
        for chs in pgs.chars:
            print(chs['text'], end='')
            str1 += chs['text']

print(str1)
```

RITIKA MOHANTY ritikamohanty38@gmail.com

540643304

INTRODUCTION An avid learner, a trainee mobile-app developer who e

◀

```
print(str2)
```

LakshayYadav(cid:211)+917982993455(cid:149)www.lakshayyadav.comlakshay.mcs21.du@gmail.comLakshayYadav(cid:135)Laks

1

<p>RITIKA MOHANTY</p>					
<p>ritikamohanty38@gmail.com</p>	<p>9540643304</p>				
<h2>INTRODUCTION</h2> <hr/>					
<p>An avid learner, a trainee mobile-app developer who eminently believes in catering to the users for commercial viability and overall success</p>					
<h2>EDUCATION</h2> <hr/>					
<p>Pursuing B.Sc. (Hons.) in Computer Science (10/10 CGPA (Semester Grade Point Avg.))</p>					
<p>Current Semester: 5th (Final Year)</p>					
<p>University: Delhi University, New Delhi</p>					
<p>Affiliate College: Shyama Prasad Mukherjee College for Women, New Delhi</p>					
<p>July 2019 – May 2022 (Expected)</p>					
<p>Senior Secondary (Class XII) (89.5%)</p>					
<p>Vishal Bharti Public School, New Delhi</p>					
<p>Affiliated to CBSE Board</p>					
<p>2018 –2019</p>					
<p>Secondary (Class X) (10/10 CGPA (Cumulative Grade Point Avg.))</p>					
<p>Vishal Bharti Public School, New Delhi</p>					
<p>Affiliated to CBSE Board</p>					
<p>2016 –2017</p>					
<h2>SKILL SET</h2> <hr/>					
<table border="1"><tr><td>Languages</td><td>C++, Java, PHP, Python</td></tr><tr><td>Tools</td><td>Android Studio, MySQL</td></tr></table>	Languages	C++, Java, PHP, Python	Tools	Android Studio, MySQL	
Languages	C++, Java, PHP, Python				
Tools	Android Studio, MySQL				
<h2>ADDITIONAL COURSES</h2> <hr/>					
<table border="1"><tr><td>Machine Learning (Self-Paced) By: Stanford University Platform: Coursera</td><td>In progress</td></tr><tr><td>Artificial Intelligence By: Tech Saksham</td><td>In progress</td></tr></table>	Machine Learning (Self-Paced) By: Stanford University Platform: Coursera	In progress	Artificial Intelligence By: Tech Saksham	In progress	
Machine Learning (Self-Paced) By: Stanford University Platform: Coursera	In progress				
Artificial Intelligence By: Tech Saksham	In progress				
<h2>PERSONAL PROJECTS</h2> <hr/>					
<p>WorkLog</p>					
<ul style="list-style-type: none">- Employee Attendance android application which marks the user present through QR code scanning and GPS Tracking (using Java, XML & PHP-for backend)- Enhanced user experience visual data representation of monthly records via a customized calendar					
<p>SugarSpirit</p>					
<ul style="list-style-type: none">- Confectionary shopping and delivery android application (using Java, XML & PHP-for backend)- Improved feasibility by integrating online payment gateways					
<p>Farm2Home</p>					
<ul style="list-style-type: none">- Fresh produce vendor website with admin panel (using HTML, PHP)					
<h2>WORK EXPERIENCE</h2> <hr/>					
<p>Part-time Content Writer, Weblink.In</p>					
<p>May 2019 – Feb 2020</p>					
<ul style="list-style-type: none">- Researched and wrote 30+ articles on a wide range of subject according to client requirements- Improved on content presentation skills under mentor supervision					

LAKSHAY YADAV	
• +91 7982993455	• www.lakshayyadav.com
lakshay.mcs21.du@gmail.com	
in Lakshay Yadav	Lakshay Yadav
Education	
Department of Computer Science, University of Delhi <i>Master's in Computer Science (8.8 CGPA)</i>	Dec. 2021 – present Vishwavidyalaya, Delhi
Atma Ram Sanatan Dharma College, University of Delhi <i>Bachelor's in Computer Science (8.9 CGPA)</i>	Jul. 2018 – Jun. 2021 Dhauka Kuan, Delhi
Certification & Training	
Data Structures & Algorithms in Python 🔗 <i>Coding Ninjas</i>	Sep. 2020 – Jan. 2021 online
Summer training on Ethical Hacking 🔗 <i>Internshala</i>	May 2019 – Jul. 2019 online
Certified in Cyber Security 🔗 <i>DIGC</i>	Jun. 2018 – Aug. 2018 offline
Projects	
Web Based Face Recognition Application 🔗 TensorFlow, Scikit-learn, Face recognition, Flask, Html, CSS	
• It is a login mechanism which uses face recognition and it is trained over a set of real and fake faces using neural networks.	
• Features: It can distinguish between real and fake faces.	
Virtual Painter 🔗 python,openCV,AIR (Augmented Reality)	
• A canvas based platform on which we can draw just by motion of the hand.	
• Human gestures are perceived via sight, and computer vision is used to research different gestures.	
Airbnb Clone 🔗 React.js	
• It is build upon - React/Tailwind CSS / MapBox integration - Next.js / Calendar Search - Next.js / SSR / SSG/Date-picker/ Maps / API.	
• In this weapp we can search any location to book room and it has integrated map showing location of places found in that region.	
Cybersecurity Projects 🔗 Python, scapy module, netfilterqueue	
• Python based cybersecurity projects. The projects are as follows:	
1. Backdoor and Listener	Using these scripts we can get an unauthorised access to systems.
2. Keylogger	Using this script we can record all the keystrokes that happened on a system.
3. Replace Download	Using this script we can change specific ongoing downloading file on a system.
4. Packet Sniffer	Using this script we can read incoming and outgoing network data.
Jun 2020	

Skills

- **Languages:** C++, Python, Html, CSS, SQL, Java Script
- **Cyber Security:** Network Security, Website Security, Social Engineering
- **Theoretical:** Data Structures and Algorithm, Data Base Management System, Object Oriented Programming
- **Libraries/Framework:** React.js, Next.js, openCV, Scapy, Face Recognition

Tokenization

It's the process of breaking a stream of textual data into words, terms, sentences, symbols, or some other meaningful elements called tokens.

Tokenization can separate sentences, words, characters, or subwords. When we split the text into sentences, we call it sentence tokenization. For words, we call it word tokenization.

Example of Tokenization

```
↳ ['RITIKA',
    'MOHANTY',
    'ritikamohanty38@gmail.com',
    '',
    '9540643304',
    '',
    'INTRODUCTION',
    'An',
    'avid',
    'learner',
    '',
    'a',
    'trainee',
    'mobile',
    '-',
    'app',
    'developer',
    'who',
    'eminently',
    'believes',
    'in',
    'catering',
    'to',
    'the',
    'users',
    'for',
    'commercial',
    'viability',
    'and',
    'overall',
    'success',
    '',
    'EDUCATION',
    'Pursuing',
    'B.Sc',
    '',
    '(',
    'Hons',
    '',
    ')',
    'in',
    'Computer',
    'Science',
    ''
```

```
doc= nlp(str2)
resume_tokens = [token.text for token in doc]
resume_tokens

['LakshayYadav(cid:211)+917982993455(cid:140)www.lakshayyadav.comlakshay.mcs21.du@gmail.comLakshayYadavEdu',
 '',
 'UniversityofDelhiDec.2021',
 '-',
 'presentMaster'sinComputerScience(8.8CGPA)Vishwavidyalaya',
 '',
 'DelhiAtmaRamSanatanDharmaCollege',
 '',
 'UniversityofDelhiJul.2018',
 '-',
 'Jun.2021Bachelor'sinComputerScience(8.9CGPA)DhulaKuan',
 '',
 'DelhiCertification&TrainingDataStructures&AlgorithmsinPythonSep.2020',
 '-',
 'Jan.2021CodingNinjasonlineSummertrainingonEthicalHackingMay2019',
 '-',
 'Jul.2019InternshalaonlineCertifiedinCyberSecurityJun.2018',
 '-',
 'Aug.2018DICOfflineProjectsWebBasedFaceRecognitionApplication|TensorFlow',
 '',
 'Scikit',
 '-',
 'learn',
 '',
 'Facerecognition',
 '',
 'Flask',
 '',
 'Html',
 '',
 'CSSFeb2022•It is a login mechanism which uses facerecognition and it is trained over a set of real and fake faces using neural networks. •Features',
 '',
 'It can distinguish between real and fake faces',
 '',
 'VirtualPainter|python',
 '',
 'openCV',
 '',
 'AR(AugmentedReality)Sep2021•A canvas based platform on which we can draw just by motion of the hand. •Human gestures are perceived via sight',
 '',
 'and computer vision is used to research different gestures',
 '',
 'AirbnbClone|React.jsMay2022•It is build upon',
 '-',
 'React',
 '/',
 'TailwindCSS',
 '/',
 'MapBox integration',
 '-',
 'Next.js',
 '/',
 'CalendarSearch',
 '-',
 'Next.js',
 '/',
 'SSR',
```

Name entry recognition

```
#name entry recoinization
docc2=nlp(strin2)
for ent in docc2.ents:
    print(ent.text, ent.start_char, ent.end_char, ent.
          type)

→ RITIKA 0 6 ORG
    Computer Science 256 272 ORG
    10/10 274 279 CARDINAL
    Current Semester: 314 331 ORG
    5th 332 335 ORDINAL
    Delhi University 361 377 ORG
    New Delhi Affiliate College 379 406 ORG
    New Delhi 451 460 GPE
    July 2019 461 470 DATE
    May 2022 473 481 DATE
    XII 518 521 ORG
    89.5% 524 529 PERCENT
    Vishal Bharti Public School 531 558 ORG
    New Delhi 560 569 GPE
    CBSE Board 584 594 ORG
```

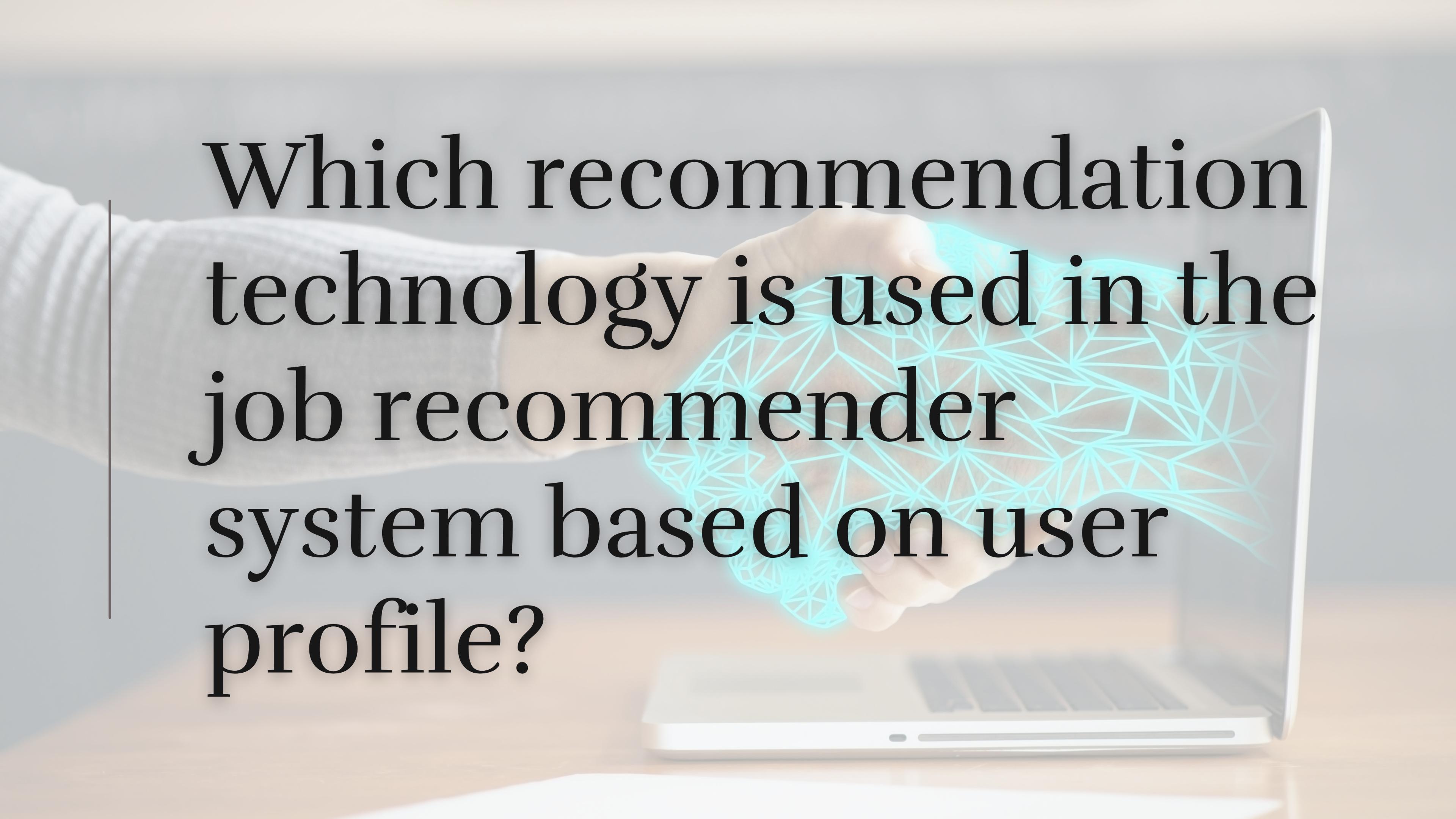
```
----- ----- ----- -----
2018 -2019 595 605 DATE
10/10 629 634 CARDINAL
Vishal Bharti Public School 671 698 ORG
New Delhi 700 709 GPE
CBSE Board 724 734 ORG
Languages C++ 760 773 ORG
Java 775 779 PERSON
PHP 781 784 PERSON
Python Tools Android Studio 786 813 ORG
Stanford University Platform 877 905 ORG
GPS Tracking 1110 1122 ORG
Java 1130 1134 PERSON
XML & PHP-for 1136 1149 ORG
monthly 1216 1223 DATE
SugarSpirit 1258 1269 ORG
Java, XML & PHP-for 1336 1355 ORG
HTML 1492 1496 ORG
PHP 1498 1501 ORG
Content Writer 1532 1546 ORG
Weblink 1548 1555 GPE
```

Hidden Markov Model

Among the methods in IE, Hidden Markov modelling has been widely used (Freitag and McCallum, 1999; Borkar et al., 2001). As a state-based model, HMMs are good at extracting information fields that hold a strong order of sequence.

For the first pass, since there exists a strong sequence among blocks, a HMM model is applied to segment a resume and each block is labelled with a category of general information.

Which recommendation technology is used in the job recommender system based on user profile?



Content-Based Recommendation (CBR)

- Analyzing the keywords and terms found in item profile and another item profile to calculate how they are similar to each other.
- relies on the user individual behavior by tracking the user interesting items to recommend items that have similar content.
- has a ramp-up problem for new users also called user **cold start problem** where new users have no enough interactions to get suitable recommendations.
- Another problem is **stability-plasticity problem**. It is the converse of the user cold start problem. Stability-plasticity happens when user extensively use the system and rate a lot of items.

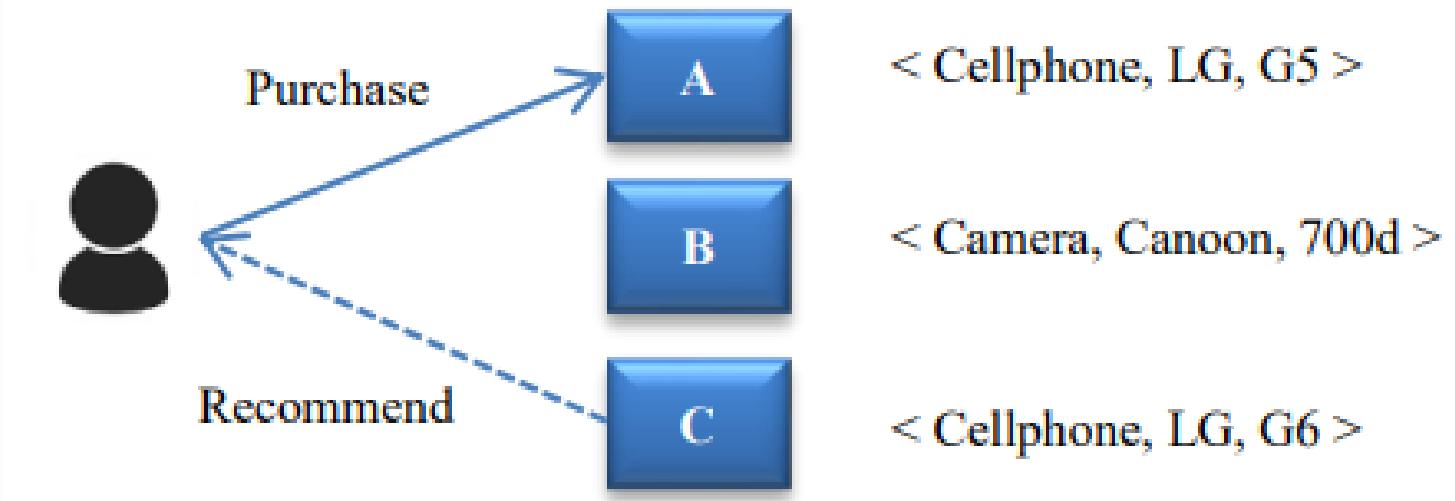


Image Source: Data-Driven Information Filtering Framework for
Dynamically Hybrid Job Recommendation

Collaborative Filtering Recommendation (CFR)

- Fundamental argument: if users X and Y rate n items similarly or have similar behaviors (e.g. buying, watching, listening), hence they will rate or act on other items similarly
- relies on the crowd behavior of all users and each individual behavior
- CFR is better than CBR in the diversity of recommendation.
- User-based CFR faces inaccurate recommendation for new users who have little interaction with the system, aka **cold-start problem**

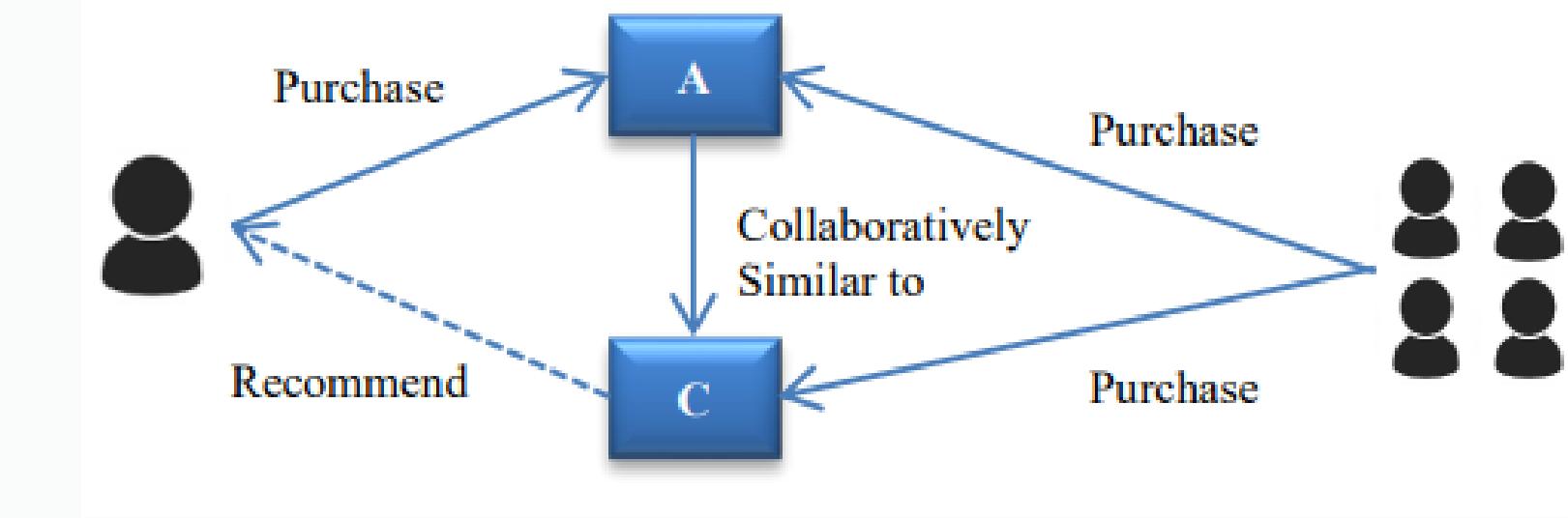


Image Source: Data-Driven Information Filtering Framework for Dynamically Hybrid Job Recommendation

Demographic-Based Recommendation (DBR)

- depends mainly on the demographic data of users (e.g. age, gender). It works on clustering users into groups. Then it starts to handle all users among the same group equally.
- relies on creating stereotypes of users and categorizing them based on their demographic data
- will not be a highly personalized recommendation

Knowledge-based Recommendation (KBR)

- depends on patterns and rules extracted from exploiting deep knowledge of the interesting items for users.
- designed basically for complex domains where customers need to specify their preferences explicitly.
- will overcome the user cold start problem but it is not easy to be used as the main recommender technique. Because it is hard to gain the proper deep knowledge, all needed patterns and rules to fit all user needs.

- Each technique has its limitations so the optimal solution is to combine two or more of those approaches to overcome all the stated shortcomings. Therefore, an hybrid approach we combine different techniques for best results.
- **CASPER (Case-Based Profiling for Electronic Recruitment) project**
 - CASPER system was designed to run as a complementary service to the original Job Finder search engine
 - CASPER ACF is a content free system that starts tracking the user from the moment that they login to the web site. It combines two server side components: a User Profiling System that constructs a user profile of each user's preferences, gathered by monitoring their behaviour within the JobFinder site, and a queryless Automated Collaborative Filtering engine, which uses this profile information to generate personalized recommendations to a user based on what similar users like.

References

- Job Recommender Systems: A Survey Zheng Siting, Hong Wenxing*, Zhang Ning, Yang Fan School of Information Science and Technology Xiamen University (2012)
- Resume Information Extraction with Cascaded Hybrid Model. Kun Yu, Gang Guan, Ming Zhou (2005)
- Data-Driven Information Filtering Framework for Dynamically Hybrid Job Recommendation Islam A. Heggo and Nashwa Abdelbaki (2021)
- Personalised Retrieval for Online Recruitment Services Rachael Rafter, Keith Bradley, Barry Smyth (2000)