

A

# Semester Project-II Report

## On

### “Skill-Based Job Recommendation System”

In partial fulfillment of requirements for the degree of

Bachelor of Technology

In

Computer Science & Engineering (Data Science)

**Submitted By**

1. Kunal Bharat Chaudhari (231106029)
2. Vaishnav Nagraj Patil (231106048)
3. Yashodip Ravindra Chaudhari (231106049)
4. Jayesh Ganesh Borale (241206002)

**Under the Guidance of**  
**Prof. S. P. Salunkhe**



**The Shirpur Education Society's**  
**R. C. Patel Institute of Technology, Shirpur - 425405.**

A

# Semester Project-I Report

## On

### “Skill-Based Job Recommendation System”

In partial fulfillment of requirements for the degree of  
Bachelor of Technology

In

Computer Science & Engineering (Data Science)  
Submitted By

1. Kunal Bharat Chaudhari (231106029)
2. Vaishnav Nagraj Patil (231106048)
3. Yashodip Ravindra Chaudhari (231106049)
4. Jayesh Ganesh Borale (241206002)

Under the Guidance of  
Prof. S.P. Salunkhe.



R. C. PATEL  
INSTITUTE OF TECHNOLOGY  
An Autonomous Institute

**The Shirpur Education Society's  
R. C. Patel Institute of Technology, Shirpur - 425405.**

Computer Science & Engineering (Data Science) 2025-26



**R. C. PATEL**  
**INSTITUTE OF TECHNOLOGY**  
An Autonomous Institute  
The Shirpur Education Society's

## **R. C. Patel Institute of Technology Shirpur, Dist. Dhule (M.S.)**

### **Computer Science & Engineering (Data Science)**

#### **CERTIFICATE**

This is to certify that the Semester Project-I entitled "**Skill-Based Job Recommendation System**" has been carried out by team:

- 1. Kunal Bharat Chaudhari (TYDS-231106029)**
- 2. Vaishnav Nagraj Patil (TYDS-231106048)**
- 3. Yashodip Ravindra Chaudhari (TYDS-231106049)**
- 4. Jayesh Ganesh Borale (TYDS-241206002)**

under the guidance of **Prof. S. P. Salunkhe**, in partial fulfillment of the requirement for the degree of Bachelor of Technology in Department of Computer Science & Engineering (Data Science) (Semester-III) of Dr. Babasaheb Ambedkar Technological University, Lonere during the academic year 2024-25.

**Date:**

**Place: Shirpur**

**Guide**

**Prof. S. P. Salunkhe**

**Semester Project-I Coordinator**

**Prof. S. P. Salunkhe**

**H.O.D.**

**Prof. Dr. U. M. Patil**

**Director**

**Prof. Dr. J. B. Patil**

## ACKNOWLEDGEMENT

We would like to express our sincere gratitude to our project guide, Prof. S. P. Salunkhe, for her valuable guidance, support, and expertise throughout the development of the Skill-Based Job Recommendation System. Her patience, encouragement, and valuable insights have been instrumental in shaping the project and making it a success. We are truly thankful for her mentorship and the time they dedicated to helping us navigate through the challenges and learn from the experience.

We would also like to extend our thanks to the Head of the Department, Prof. Dr. U.M. Patil, for her continuous support and belief in our abilities. Their encouragement and provision of necessary resources created an environment conducive to innovation and growth. We are grateful for her leadership and for fostering a positive learning atmosphere that has contributed to our development.

Furthermore, we are deeply grateful to the Director, Prof. Dr. J.B. Patil, for his visionary leadership and unwavering support. His guidance and confidence in our project gave us the motivation to strive for excellence. We appreciate his belief in our potential and his efforts to create an environment that encourages students to explore their creativity and pursue their ideas.

### Project team:

1. Kunal Bharat Chaudhari
2. Vaishnav Nagraj Patil
3. Yashodip Ravindra Chaudhari
4. Jayesh Ganesh Borale

## **Vision of Institute**

"Fostering technical excellence through ethics, sustainable education, innovation and research"

## **Mission of Institute**

To impart high quality Technical Education through:

1. Innovative and Interactive learning process and high quality, globally recognized instructional programs.
2. Fostering a collaborative scientific temper among students with ethical responsibility towards the society.
3. Preparing students from diverse backgrounds to have aptitude for employment, entrepreneurship and research with a spirit of Professionalism.
4. To contribute to nation's sustainable development.

## **Vision of the Computer Science & Engineering (Data Science)**

To provide cutting-edge Computer Engineering education in Data Science while instilling socio-moral values.

## **Mission of the Computer Science & Engineering (Data Science)**

1. To deliver state-of-the-art, ICT-enabled teaching and learning to achieve excellence in Data Science education.
2. To develop professionally competent Data Science Engineers, meeting evolving industrial and societal needs.
3. To prepare employable professionals with ethical values and a commitment to professional and social responsibility.

## **Program Specific Outcomes (PSOs)**

1. Apply programming concepts, algorithms, and data structures to develop data-driven software and web solutions.
2. Develop data-driven solutions using machine learning, data analysis, and cloud technologies for practical problem-solving.

## PAGE INDEX

<b>Chapter No.</b>	<b>Topic</b>	<b>Page No.</b>
<b>1.</b>	ABSTRACT	
	<b>INTRODUCTION</b>	<b>1-6</b>
	1.1 Background	4
	1.2 Problem Statement	5
1.3 Objective of work	6	
<b>2.</b>	<b>Related Concepts</b>	<b>7-13</b>
	2.1 Working System	8-9
	2.2 Algorithm	10-11
	2.3 Software Requirements	12-13
<b>3.</b>	<b>IMPLEMENTATION</b>	<b>14-16</b>
<b>4.</b>	<b>CONCLUSION</b>	<b>17-18</b>
	<b>Bibliography</b>	

---

## ABSTRACT

---

In the rapidly evolving digital era, online recruitment platforms and job portals have become the primary medium for connecting job seekers with employment opportunities. Although these platforms provide access to a vast number of job postings, they often fail to deliver relevant and personalized recommendations to users. Traditional job recommendation systems mainly depend on keyword matching, job titles, and basic filters such as location and experience. As a result, job seekers frequently receive irrelevant job suggestions, while recruiters struggle to identify candidates whose skills truly match job requirements.

To overcome these challenges, this project presents a **Skill-Based Job Recommendation System**, an intelligent solution that focuses on matching candidate skills with job requirements rather than relying solely on keywords. The proposed system leverages **Natural Language Processing (NLP)**, **Machine Learning**, and **Large Language Models (LLMs)** to analyse unstructured textual data such as resumes and job descriptions. Using advanced text-processing techniques, the system extracts relevant skills, qualifications, and experience from candidate resumes and compares them with job descriptions collected from online job portals.

The system performs preprocessing operations such as tokenization, stop-word removal, stemming, lemmatization, and vectorization using techniques like TF-IDF. Semantic similarity methods, including cosine similarity, are applied to calculate the relevance between candidate skill profiles and job requirements. Based on similarity scores, the system ranks and recommends the most suitable jobs to users in real time, ensuring accurate and meaningful job suggestions.

The implementation of the system is carried out using **Python** as the core programming language, with **Django** used for backend development and **HTML/CSS** for frontend design. **Selenium** is employed for automated job data collection from online portals to ensure updated job listings. The modular architecture of the system ensures scalability, reliability, and ease of maintenance.

Overall, the Skill-Based Job Recommendation System enhances the job search experience by reducing irrelevant recommendations, saving time, and supporting informed career decisions. It also assists recruiters and placement cells in identifying suitable candidates efficiently, demonstrating the practical application of Artificial Intelligence in modern recruitment systems.

## CHAPTER-1

### INTRODUCTION

---

In the present digital age, the process of recruitment and job searching has undergone a significant transformation due to the widespread adoption of online platforms and automated systems. Job portals, professional networking websites, and recruitment applications have become the primary medium through which employers advertise job openings and candidates search for suitable employment opportunities. While these platforms provide easy access to a large number of job postings, they also create a new challenge—information overload. Job seekers are often overwhelmed by thousands of job listings, many of which are irrelevant to their skills and career aspirations. Similarly, recruiters face difficulties in identifying candidates with the most suitable skill sets from a large pool of applicants. This mismatch between job requirements and candidate skills results in inefficiencies, increased recruitment time, and reduced satisfaction for both parties.

Traditional job recommendation systems rely heavily on keyword matching, job titles, and simple filters such as location, experience level, and education. Although these techniques are easy to implement, they fail to capture the true capabilities of a candidate and the actual requirements of a job. For instance, two job roles may have different titles but require similar skill sets, while jobs with the same title may demand entirely different technical competencies. Keyword-based systems cannot understand such semantic relationships, leading to poor recommendation quality. As a result, candidates often miss relevant opportunities, and employers may overlook suitable applicants. This limitation highlights the need for a more intelligent and skill-focused job recommendation approach.

A skill-based job recommendation system aims to overcome these challenges by focusing on the actual skills possessed by a candidate rather than superficial keywords. Skills such as programming languages, tools, frameworks, domain knowledge, and soft skills play a crucial role in determining job suitability. By analyzing resumes and job descriptions as unstructured textual data, advanced computational techniques can be used to extract meaningful information and identify hidden patterns. Natural Language Processing (NLP) enables machines to understand, interpret, and process human language, making it possible to analyze resumes and job postings at a deeper semantic level. When combined with machine learning algorithms, NLP can significantly improve the accuracy and relevance of job recommendations.

In recent years, the integration of Artificial Intelligence (AI) into recruitment systems has gained considerable attention. Machine learning models can learn from historical data, identify trends, and adapt to changing job market demands. Large Language Models (LLMs) further enhance this capability by understanding contextual meaning, synonyms, and relationships between different skills. These technologies enable the development of

intelligent recommendation systems that provide personalized job suggestions tailored to an individual's skill profile. Such systems not only reduce job search time but also improve employability by guiding candidates toward roles that best match their competencies.

The Skill-Based Job Recommendation System proposed in this project is designed to address the limitations of traditional job portals by introducing a data-driven and intelligent approach to job matching. The system analyzes candidate resumes and job descriptions using NLP techniques to extract relevant skills, experience, and qualifications. These extracted features are transformed into numerical representations using vectorization methods, allowing the system to calculate similarity scores between candidates and job roles. Based on these scores, the system ranks and recommends the most suitable job opportunities to users. This approach ensures that recommendations are based on actual skill compatibility rather than simple keyword matches.

The system is implemented using Python as the primary programming language due to its extensive support for data analysis, machine learning, and NLP libraries. Django is used as the backend framework to manage application logic, database interactions, and user requests, while HTML and CSS are used to create a clean and user-friendly interface. Selenium is employed to automate the collection of job data from online portals, ensuring that the system works with updated and relevant job listings. The modular architecture of the system allows easy scalability, making it suitable for future enhancements such as real-time labor market analysis, career path prediction, and integration with mobile platforms. In conclusion, the Skill-Based Job Recommendation System represents an important step toward modernizing the recruitment process by leveraging Artificial Intelligence and data-driven techniques. By focusing on skills rather than keywords, the system improves the quality of job recommendations, enhances user satisfaction, and supports efficient hiring. This project demonstrates the practical application of NLP and machine learning in solving real-world problems and highlights the growing importance of intelligent recommendation systems in today's competitive job market.

## 1.1 Background

The rapid advancement of digital technologies and the widespread adoption of the internet have fundamentally transformed the global employment landscape. Recruitment processes that once relied on newspapers, walk-in interviews, and manual screening have now shifted to online job portals, professional networking platforms, and automated recruitment systems. This digital transformation has enabled organizations to reach a wider pool of candidates and has provided job seekers with instant access to thousands of employment opportunities. However, despite these advantages, the increasing volume of job listings and candidate profiles has introduced new challenges related to efficiency, relevance, and personalization in the job search process.

Modern job portals store massive amounts of unstructured textual data in the form of resumes, job descriptions, cover letters, and company profiles. While this data contains valuable information about candidate skills and job requirements, traditional systems are unable to fully utilize it. Most existing job recommendation mechanisms depend on keyword matching, predefined rules, and basic filters such as job title, experience, and location. These approaches fail to understand the semantic meaning of skills and do not consider the contextual relationship between different technologies, tools, or competencies. As a result, job seekers often receive recommendations that do not align with their actual abilities, and recruiters may overlook suitable candidates whose resumes do not contain exact keyword matches.

At the same time, the nature of employment itself has evolved. Employers are increasingly adopting a skill-based hiring approach rather than focusing solely on degrees or job titles. With the rapid emergence of new technologies, especially in fields such as data science, software development, cloud computing, and artificial intelligence, skill requirements change frequently. This dynamic environment demands intelligent systems that can adapt to evolving job market trends and accurately map candidate skills to relevant roles. Static and rule-based recommendation systems are not capable of handling such complexity, highlighting the need for intelligent, data-driven solutions.

The emergence of Artificial Intelligence (AI), Machine Learning (ML), and Natural Language Processing (NLP) has opened new possibilities for improving recruitment systems. NLP techniques enable machines to process and understand human language, making it possible to analyse unstructured resumes and job descriptions at a deeper level. By applying techniques such as tokenization, lemmatization, and semantic analysis, relevant skills and experiences can be extracted automatically. Machine learning algorithms can then learn patterns from historical recruitment data and predict suitable job matches for candidates. In this context, the Skill-Based Job Recommendation System proposed in this project aims to leverage modern AI techniques to create an efficient, scalable, and user-centric recruitment solution. By combining NLP, machine learning, and automated data collection, the system provides personalized job recommendations that are aligned with real-world skill requirements. This background establishes the motivation for the project and highlights the importance of intelligent job recommendation systems in today's competitive job market.

## 1.2 Problem Statement

The rapid digitalization of recruitment has led to the widespread use of online job portals and professional networking platforms for hiring and job searching. While these platforms provide easy access to numerous job opportunities, they also introduce several challenges related to relevance, accuracy, and efficiency in job matching. Most existing job recommendation systems rely heavily on keyword-based search mechanisms, job titles, and basic filtering options such as location, education, and years of experience. These traditional approaches fail to capture the actual skills and competencies of candidates, resulting in ineffective job recommendations.

One of the major problems with existing systems is their inability to accurately understand and interpret unstructured resume data. Resumes and job descriptions are written in natural language and often use different terminologies to describe similar skills. For example, a candidate skilled in “Python programming” may not be recommended a job that requires “scripting in Python,” even though both represent the same competency. Keyword-based systems treat such terms as different, leading to missed opportunities for both job seekers and employers. This limitation significantly reduces the quality of recommendations and causes frustration among users.

From the recruiter’s perspective, the problem is equally significant. Employers receive a large number of applications for each job posting, many of which do not meet the required skill criteria. Screening these applications manually is time-consuming and prone to human error. Recruiters may overlook suitable candidates whose resumes do not match exact keyword requirements or who use different terminology to describe their skills. This leads to delayed hiring processes and increased operational costs for organizations.

Furthermore, existing job recommendation systems are not well-equipped to handle the dynamic nature of the job market. Skill requirements evolve rapidly due to technological advancements and changing industry demands. Traditional rule-based systems cannot adapt easily to these changes and often rely on outdated data. This results in recommendations that do not reflect current job market trends, further reducing their effectiveness.

In summary, the core problem addressed by this project is the absence of an intelligent, skill-focused job recommendation system capable of accurately analyzing resumes and job descriptions, understanding semantic relationships between skills, and providing relevant job recommendations. Existing systems lack adaptability, accuracy, and deep skill understanding. Therefore, there is a strong need for a data-driven solution that leverages Natural Language Processing and Machine Learning to bridge the gap between job seekers and employers through efficient and skill-based job matching.

### 1.3 Objective of Work

- The primary objective of the Skill-Based Job Recommendation System is to design and develop an intelligent platform that matches job seekers with suitable employment opportunities based on their actual skills and competencies. The system aims to overcome the limitations of traditional keyword-based job portals by focusing on skill relevance and meaningful job matching using data-driven techniques.
- Another key objective of this project is to automatically extract important information from unstructured resumes and job descriptions. Using Natural Language Processing techniques such as tokenization, stop-word removal, and lemmatization, the system identifies relevant skills, qualifications, and experience, thereby reducing manual effort and improving accuracy in skill extraction.
- The project also aims to convert extracted textual information into a structured and comparable format. By applying vectorization techniques like TF-IDF and semantic embeddings, candidate skills and job requirements are represented numerically, allowing similarity measures and machine learning algorithms to calculate accurate job matches.
- Providing personalized job recommendations is another important objective of the system. Each candidate receives job suggestions based on their individual skill profile rather than generic recommendations. This personalization helps job seekers identify opportunities that align closely with their strengths and career goals, improving the overall job search experience.
- From an organizational perspective, the system aims to support recruiters by reducing the time and effort required for candidate screening. By ranking candidates and job roles based on skill similarity, the system helps recruiters identify suitable applicants efficiently, leading to faster and more effective hiring decisions.
- Finally, the project aims to develop a scalable, user-friendly, and practical web-based application using Python, Django, and modern web technologies. The system demonstrates the real-world application of Artificial Intelligence in recruitment and provides a foundation for future enhancements such as advanced recommendation models and real-time labour market analysis.

## CHAPTER-2

### RELATED CONCEPT

---

The Skill-Based Job Recommendation System is based on the concept of intelligent recommendation systems that aim to suggest relevant items to users based on meaningful analysis rather than simple rules. In recruitment, recommendation systems are used to connect job seekers with suitable job roles. Traditional systems rely mainly on keyword matching, job titles, and basic filters, which often fail to capture the real capabilities of candidates. Skill-based recommendation focuses on matching actual competencies such as technical skills, tools, and domain knowledge, making job recommendations more accurate and relevant in today's dynamic job market.

Natural Language Processing (NLP) is a core concept used in this system to analyse unstructured textual data such as resumes and job descriptions. Since these documents are written in natural language and vary in format, NLP techniques like tokenization, stop-word removal, stemming, and lemmatization are applied to clean and preprocess the text. NLP helps the system understand semantic meaning, handle synonyms, and extract important skills and experience details, which are essential for effective job matching.

Another important concept is text representation and similarity measurement. After preprocessing, resumes and job descriptions are converted into numerical vectors using techniques such as TF-IDF and word embeddings. These vector representations allow the system to compare candidate skills with job requirements mathematically. Similarity measures like cosine similarity are then used to calculate how closely a candidate's profile matches a particular job, and jobs are ranked accordingly to provide the most relevant recommendations first.

Machine Learning and personalization further enhance the effectiveness of the system. Machine learning algorithms help identify patterns in skill requirements and job trends, while Large Language Models (LLMs) improve contextual understanding of text. Personalization ensures that each job seeker receives customized job recommendations based on their individual skill profile rather than generic suggestions. Together, these concepts enable the development of an intelligent, scalable, and efficient skill-based job recommendation system.

## 2.1 Working System

### Resume and Job Data Collection:

---

Resume and job description data form the core input of the Skill-Based Job Recommendation System. User resumes are collected through file upload or text input, while job descriptions are gathered from online job portals or datasets. These documents are usually unstructured and written in natural language. Collecting accurate and up-to-date job data ensures that the recommendation system provides relevant and meaningful job suggestions to users.

### Natural Language Processing (NLP):

Natural Language Processing is the core engine behind text understanding in this project. NLP techniques are used to analyze resumes and job descriptions by breaking text into tokens, removing stop words, converting words to their base form using stemming or lemmatization, and handling variations in terminology. NLP enables the system to understand skills, experience, qualifications, and domain knowledge from raw text, making the system intelligent and context aware.

### Skill Extraction:

Skill extraction is a crucial step where relevant technical and non-technical skills are identified from resumes and job descriptions. Skills such as programming languages, tools, frameworks, certifications, and soft skills are extracted using predefined skill lists and NLP-based methods. This process converts unstructured text into meaningful skill information that can be compared across candidates and job roles.

### Text Vectorization:

After extracting skills, the textual data is converted into numerical form so that machine learning algorithms can process it. Techniques such as TF-IDF (Term Frequency–Inverse Document Frequency) or word embeddings are used to represent resumes and job descriptions as vectors. These vectors capture the importance of each skill and allow mathematical comparison between candidate profiles and job requirements.

### **Similarity Calculation and Ranking:**

Once vector representations are created, similar measures such as cosine similarity are applied to calculate how closely a candidate's skills match a job's requirements. Each job is assigned a similarity score based on this comparison. Jobs with higher similarity scores indicate a better match and are ranked higher in the recommendation list. This ranking ensures that users receive the most relevant job suggestions first.

### **Recommendation Generation:**

Based on the similarity scores and ranking, the system generates personalized job recommendations for the user. The top-matching job roles are displayed along with relevant details such as job title, company name, required skills, and match score. This personal recommendation approach helps job seekers quickly identify suitable opportunities without manually searching through numerous listings.

### **Web-Based Interface:**

A simple and user-friendly web-based interface is developed using frameworks such as Django, Flask, or Streamlit. Users can upload resumes, view extracted skills and receive job recommendations in real time. The interface is designed to be clean, responsive, and easy to navigate, making the system accessible even to non-technical users.

## 2.2 Algorithm

---

### **Input Acquisition:**

The algorithm begins by taking input from the user in the form of a resume file or a list of skills entered through the system interface. At the same time, job descriptions are loaded from the system database or collected from online job portals. These inputs serve as the primary data sources for the recommendation process.

### **Text Preprocessing:**

In this step, Natural Language Processing techniques are applied to clean and normalize the input text. This includes converting text to lowercase, removing stop words and special characters, and applying stemming or lemmatization. Preprocessing ensures that irrelevant words are removed and only meaningful information is retained for further analysis.

### **Skill Extraction:**

After preprocessing, relevant skills are extracted from both resumes and job descriptions. This step identifies technical skills, tools, programming languages, certifications, and important keywords using predefined skill dictionaries and NLP-based methods. The extracted skills form the core features used for job matching.

### **Feature Vector Generation:**

The extracted skills and processed text are converted into numerical vectors using vectorization techniques such as TF-IDF or word embeddings. These vectors represent the importance of skills in each document and allow the system to perform mathematical comparison between candidate profiles and job requirements.

### **Similarity Computation:**

Once vector representations are generated, similarity measures such as cosine similarity are applied to compute the similarity score between the candidate's skill vector and each job description vector. Higher similarity scores indicate a stronger match between the candidate and the job role.

### **Job Ranking:**

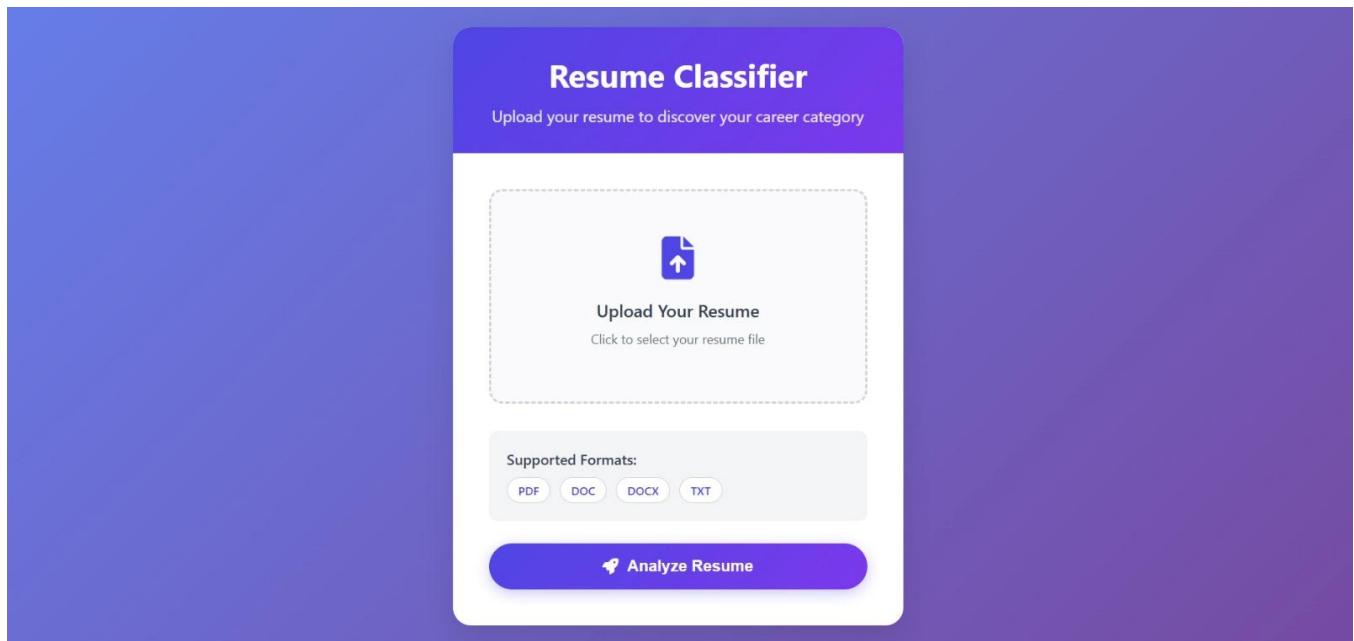
Based on the similarity scores, all job roles are ranked from highest to lowest relevance. Jobs with the highest similarity values are considered the most suitable for the candidate. This ranking mechanism ensures that the best-matching jobs are prioritized in the recommendation list.

### **Recommendation Output:**

Finally, the top-ranked job recommendations are selected and displayed to the user through the web interface. The output includes job details such as job title, company name,

required skills, and match score. This allows users to easily evaluate and apply for suitable job opportunities.

### User Interface:



### Output:

#### SUGGESTED JOBS

**Results: 10 jobs found**

##### Python developer

Tata Consultancy Services

Pune, Chennai, Bengaluru | 6-8 Yrs | Not disclosed | Recently

Role & responsibilities Knowledge in object-relational mapping (ORM) Familiarity with ...

URL: <https://www.naukri.com/job-listings-python-developer-tata-consultancy-services-pune-chennai-bengaluru-6-to-8-years-201225008476>

##### Python developer

PwC Service Delivery Center

Hybrid - Bengaluru | 5-10 Yrs | Not disclosed | Recently

Participate in system design, architecture, and technical decision-making processes. Ot...

URL: <https://www.naukri.com/job-listings-python-developer-pwc-service-delivery-center-bengaluru-5-to-10-years-161225016414>

##### Python Software Developer

Embedsense

Bengaluru(Hebbal) | 0-3 Yrs | 3.5-5.5 Lacs PA | Recently

Required Qualifications . Bachelors degree in computer science, Engineering, or related...

URL: <https://www.naukri.com/job-listings-python-software-developer-embedsense-bengaluru-0-to-3-years-201225012053>

## 2.3 Software requirements:

### ➤ Software Requirements

Software requirements refer to the set of tools, libraries, programming environments, and services required to develop, deploy, and execute the Skill-Based Job Recommendation System. This project utilizes a combination of open-source technologies and web-based platforms to ensure efficient resume processing, skill extraction, and job recommendation. The detailed software components are listed below:

#### 1. Python 3.10+

- Core programming language used for resume parsing, text preprocessing, skill extraction, similarity computation, and recommendation logic.
- Supports essential libraries such as NLTK, spaCy, NumPy, Pandas, and Scikit-learn.
- Provides strong community support and seamless integration with Machine Learning and NLP frameworks.

#### 2. Django / Flask (Web Framework)

- Used to develop the backend of the job recommendation system.
- Handles resume uploads, job data processing, user requests, and recommendation output.
- Provides a secure, scalable, and efficient server-side environment for web-based applications.

#### 3. Natural Language Processing Libraries (NLTK / spaCy)

- Used for processing unstructured resume and job description text.
- Supports:
  - Tokenization
  - Stop-word removal
  - Stemming and Lemmatization
  - Skill and keyword extraction
- Helps the system understand and analyze human language effectively.

#### 4. Machine Learning Libraries (Scikit-learn)

- Used to implement vectorization and similarity-based algorithms.
- Supports :
  - TF-IDF Vectorization
  - Feature extraction and ranking

## 5. Pandas and NumPy

- Pandas is used for data manipulation, cleaning, and handling resume/job datasets.
- NumPy is used for numerical operations and array-based computations.
- Together, they ensure efficient processing of large text datasets.

## 6. Selenium (Web Automation Tool)

- Used for automated collection of job descriptions from online job portals.
- Helps fetch real-time and updated job listings.
- Reduces manual effort in job data gathering.

## 7. Database Management System (SQLite / MySQL)

- Used to store user profiles, resumes, extracted skills, and job descriptions.
- Ensures structured data storage and quick retrieval.
- Supports efficient recommendation generation and system scalability.

## 8. Frontend Technologies (HTML / CSS / JavaScript)

- HTML is used to structure the web pages.
- CSS is used for styling and layout design.
- JavaScript is used for form validation and basic interactivity.
- Provides a clean, responsive, and user-friendly interface.

## 9. Web Browser (Chrome / Edge / Firefox)

- Required to access and test the web-based job recommendation system.
- Ensures smooth rendering of the user interface and compatibility across platforms.

## CHAPTER-3

# IMPLEMENTATION

---

### 3.1 Resume Data Collection and Parsing

The first stage of implementation focuses on collecting and extracting meaningful textual information from candidate resumes. Resumes are uploaded by users in common formats such as PDF or DOC files. A resume parser is used to extract raw text content from these documents. The extracted text includes details such as skills, qualifications, experience, and project descriptions.

Each resume is cleaned and converted into a standardized text format to ensure compatibility with further processing steps. This parsed resume text serves as the primary input for skill extraction and job matching. All extracted resume data is stored in a structured format for further analysis.

### 3.2 Job Description Preprocessing and Skill Vocabulary Creation

Alongside resumes, job descriptions are collected from online job portals or datasets. These job descriptions undergo a series of text preprocessing steps to prepare them for analysis:

- Converting all text to lowercase
- Removing unnecessary characters
- Removing stop words
- Applying stemming or lemmatization
- Normalizing whitespace

After preprocessing, important skills and keywords are extracted from job descriptions. A common skill vocabulary is created by combining skills from resumes and job descriptions. Each unique skill is assigned an index, allowing consistent representation across candidate and job profiles.

### 3.3 Dataset Preparation Using Structured Skill Representation

The resume and job datasets can be large and unstructured. To manage this efficiently, structured datasets are created where each entry contains:

- Extracted resume skills
- Corresponding job description skills

- Candidate identifiers
- Job identifiers

This structured representation allows efficient processing and comparison. Instead of loading all data at once, the system processes records dynamically during similarity computation, ensuring scalability and optimal memory usage.

### **3.4 Skill Matching and Similarity-Based Recommendation Model**

The system uses a similarity-based recommendation approach to match candidate skills with job requirements.

#### **Skill Vector Generation**

Extracted skills from resumes and job descriptions are converted into numerical vectors using TF-IDF vectorization. These vectors represent the importance of each skill in the document.

#### **Similarity Computation**

Cosine similarity is applied between resume vectors and job vectors to calculate similarity scores. Higher similarity values indicate a stronger match between candidate skills and job requirements.

#### **Recommendation Logic**

Based on similarity scores, job roles are ranked in descending order. The top-ranked jobs are considered the most suitable recommendations for the candidate.

### **3.5 Model Execution and Recommendation Generation**

The recommendation engine processes each candidate profile and compares it with all available job descriptions. For every candidate:

- Resume text is preprocessed
- Skills are extracted and vectorized
- Similarity scores are computed
- Job roles are ranked based on relevance

The system continuously improves recommendation accuracy by refining skill extraction and similarity computation. This approach ensures consistent, reliable, and skill-focused job matching.

### **3.6 User Interface Development Using Web Framework**

A user-friendly web interface is developed using Django/Flask and frontend technologies such as HTML and CSS. The interface allows users to interact easily with the system.

#### **Features of the User Interface:**

- Allows users to upload resumes
- Displays extracted skills for review
- Processes resume through the recommendation engine
- Shows recommended job listings with relevance scores
- Offers a clean, simple, and responsive layout

#### **1. Resume Upload Logic**

The uploaded resume is validated and sent to the backend for parsing and preprocessing.

#### **2. Skill Extraction Logic**

The parsed resume text is processed using NLP techniques to extract relevant skills.

#### **3. Recommendation Generation Logic**

The similarity-based model computes job matches and ranks suitable job roles.

#### **4. Result Display Handling**

The recommended jobs are displayed on the web interface with job details and match scores.

This complete workflow ensures smooth interaction between the frontend and backend, providing real-time and accurate job recommendations to users.

## CONCLUSION

---

The Skill-Based Job Recommendation System successfully demonstrates how modern Artificial Intelligence techniques can be applied to improve the recruitment and job search process. By focusing on skill matching rather than traditional keyword-based methods, the system provides more accurate and meaningful job recommendations. The project effectively bridges the gap between job seekers and employers by analyzing resumes and job descriptions using Natural Language Processing and similarity-based machine learning techniques.

The system implementation covers the complete workflow starting from resume collection and preprocessing to skill extraction, vectorization, similarity computation, and final job recommendation. By using NLP techniques such as tokenization, lemmatization, and stop-word removal, the system efficiently processes unstructured text data. The use of TF-IDF vectorization and cosine similarity ensures reliable and consistent comparison between candidate skills and job requirements.

A user-friendly web interface enhances accessibility and usability of the system. Users can upload resumes, view extracted skills, and receive personalized job recommendations in real time. The modular design of the system allows easy maintenance and scalability, making it suitable for future enhancements such as advanced machine learning models, real-time labor market analysis, and integration with professional networking platforms.

Overall, the Skill-Based Job Recommendation System provides an effective, scalable, and intelligent solution for modern recruitment challenges. It reduces manual effort, improves hiring efficiency, and supports informed career decisions. The project highlights the importance of skill-based approaches in today's evolving job market and demonstrates the practical application of Artificial Intelligence in building smart and future-ready recruitment systems.

---

**BIBLIOGRAPHY**

---

1. Jurafsky, D., & Martin, J. H., *Speech and Language Processing*, 3rd Edition, Pearson Education, 2020.
2. Géron, A., *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 2nd Edition, O'Reilly Media, 2019.
3. Pedregosa, F., Varoquaux, G., Gramfort, A., et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, Vol. 12, pp. 2825–2830, 2011.
4. Manning, C. D., Raghavan, P., & Schütze, H., *Introduction to Information Retrieval*, Cambridge University Press, 2008.
5. Mikolov, T., Chen, K., Corrado, G., & Dean, J., "Efficient Estimation of Word Representations in Vector Space," *Proceedings of the International Conference on Learning Representations (ICLR)*, 2013.
6. Kaggle, "Job Recommendation and Resume Dataset," Available at: <https://www.kaggle.com/> (Accessed: 2025).
7. NLTK Documentation, "Natural Language Toolkit," Available at: <https://www.nltk.org/>
8. spaCy Documentation, "Industrial-Strength Natural Language Processing," Available at: <https://spacy.io/>
9. Scikit-learn Documentation, "Text Feature Extraction and Similarity Measures," Available at: <https://scikit-learn.org/>
10. Django Software Foundation, "Django Web Framework Documentation," Available at: <https://docs.djangoproject.com/>
11. Selenium Documentation, "Web Automation Tool," Available at: <https://www.selenium.dev/documentation/>
12. IBM, "Artificial Intelligence in Recruitment and Talent Management," IBM Research White Paper, 2022.
13. LinkedIn Talent Solutions, "Skills-Based Hiring: The Future of Recruitment," Industry Report, 2023.