JOB-QUEST: A Job Seeking Assistance Portal Based on User Resume

Chayan Tank*, Dhruv Patel†, Kanti-lal Patel‡, Sonik Sarungale§, and Vivek Singh¶

* † ‡ § ¶ Dept. of Computer science and Engineering
Indraprastha Institute of Information Technology, Delhi, India.
* chayan23030@iiitd.ac.in † dhruv23056@iiitd.ac.in ‡ kantilal23064@iiitd.ac.in
§ sonik23098@iiitd.ac.in ¶ vivek23106@iiitd.ac.in

Abstract—In today's job market, the vast number of daily job postings can overwhelm job seekers, leading to inefficiencies in matching skills and career goals with available opportunities. This challenge also extends to recruiters who struggle to attract suitable applicants for their open positions. To address these issues, we propose a resume-based job ranking system designed to streamline the job search process for both job seekers and recruiters.

Our system focuses on benefiting job seekers by leveraging their resumes to dynamically rank and list relevant job postings from various internet sources. Unlike traditional approaches that primarily assist recruiters in automating resume filtering, our system prioritizes matching job opportunities to an individual's professional profile. By appropriately processing the resume information, we rank job descriptions stored in our database against specific input resumes. The job database comprises a collection of jobs fetched from the Internet, providing a diverse range of opportunities. We use Okapi's BM25 algorithm for similarity matching between resumes and job postings. Our results demonstrate the system's ability to provide users with top-ranked job opportunities tailored to their skills and career aspirations, facilitating a more efficient and effective job application process.

Index Terms—Information Retrieval, Job matching, Okapi BM25 algorithm, Skill-based job recommendations, Job seeker efficiency, Recruiter effectiveness.

I. PROBLEM STATEMENT

The contemporary job market presents a formidable challenge for both job seekers and recruiters due to the overwhelming volume of daily job postings. Job seekers often find it hard and tedious to locate opportunities that align precisely with their skills, experiences, and career aspirations. This struggle results in inefficiencies as they navigate through numerous job listings, often applying to positions that do not match their professional goals. Similarly, recruiters face hurdles in attracting and identifying qualified candidates for their open positions. The lack of an effective job-matching system contributes to prolonged recruitment cycles, resource wastage, and potential mismatches between job requirements and candidate profiles.

The existing job search landscape is characterized by a cumbersome and time-consuming application process for both job seekers and recruiters. Job seekers encounter challenges in filtering through a sea of job listings, leading to frustration and a sense of disconnection from suitable opportunities. On the other hand, recruiters grapple with identifying the right candidates from a vast pool of applicants, resulting in prolonged hiring timelines and increased resource utilization.

The overarching problem revolves around the inefficiency and lack of precision in job matching, impacting both job seekers' ability to secure suitable employment and recruiters' efforts to fill positions effectively. There is a critical need for a robust job ranking and matching system that addresses these challenges by accurately analyzing job seekers' resumes, extracting pertinent information, and providing tailored job recommendations. Such a system would streamline the job search process, enhance recruitment effectiveness, and improve the overall job market ecosystem for all stakeholders.

II. MOTIVATION

In today's rapidly evolving job market especially tech and IT jobs, this process of job searching and candidate selection has become increasingly complex and challenging. Job seekers are faced with a vast array of job postings daily, making it difficult to identify opportunities that align with their specific skills, experiences, and career aspirations. On the other hand, recruiters are inundated with numerous resumes and job applications, leading to time-consuming and resource-intensive candidate screening processes.

This scenario highlights the critical need for innovative solutions that can streamline the job search and recruitment processes, benefiting both job seekers and recruiters alike. A resume-based job ranking system presents a promising approach to address these challenges by leveraging advanced technologies such as machine learning, natural language processing, and semantic analysis.

The motivation behind developing a resume-based job ranking system stems from the following key factors:

- **Enhanced Efficiency**: By automating the resume screening and job matching processes, the system aims to significantly reduce the time and effort required for both job seekers and recruiters. This increased efficiency translates to faster candidate selection and improved job placement rates.
- **Personalized Job Recommendations**: The system's ability to analyze resumes and match them with job descriptions enables the generation of personalized job recom-

mendations tailored to each candidate's skills, location, experiences, and career goals. This personalized approach enhances the likelihood of successful job placements and job satisfaction among candidates.

- **Optimized Recruitment Processes**: For recruiters, the systems should offers a streamlined and systematic approach to candidate evaluation and selection. By ranking candidates based on their resume relevance to job requirements, recruiters can identify top candidates more efficiently, leading to improved recruitment outcomes and reduced hiring costs.
- **Scalability and Adaptability**: A well-designed resume-based job ranking system can be scaled to handle large volumes of resumes and job postings across diverse industries and job roles. Its adaptability ensures that it can cater to the evolving needs of both job seekers and recruiters in a dynamic job market.

III. LITERATURE REVIEW

The field of job retrieval systems using resume-based approaches has seen significant advancements and contributions from various researchers. Tejaswini et al. [6] proposed a machine learning-based resume ranking system that initiates with a candidate assessment through Multiple Choice Questions (MCQs). Upon passing this initial assessment, candidates' resumes undergo a rigorous cosine similarity-based ranking process. This methodology ensures that the most suitable candidates are selected based on their resume content and alignment with job requirements, enhancing the efficiency of candidate selection processes for recruiters.

In a similar vein, Wailthare et al. [7] introduced a system where recruiters post job openings and subsequently accept candidate resumes for evaluation. The system employs similarity-matching techniques to assess the compatibility of candidate profiles with job descriptions. By ranking candidates based on their resumes' relevance to job requirements, this approach streamlines the candidate selection process for recruiters, enabling them to identify top candidates efficiently.

On the classification front, Gopalakrishna et al. [5] focused on resume classification using Natural Language Processing (NLP) techniques and classical Machine Learning (ML) algorithms such as naive Bayes. Their approach aimed to automate the resume filtering process for recruiters, categorizing resumes into predefined classes based on specific criteria. While effective for simplifying the hiring process, this method may lack scalability and adaptability to diverse user groups and job roles.

Jiechieu et al. [4] took a deep learning approach using Convolutional Neural Networks (CNNs) for resume classification, particularly within the Information Technology (IT) domain. Their methodology involved computing word embedding matrices and feeding them into CNNs for classification into predefined classes. While this approach demonstrated efficacy within a specific domain, it may require further enhancements

to incorporate job ranking and comparison mechanisms for broader applicability.

Zaroor et al. [1] introduced the Job Post and Resume Classification System (JRC), which combines section-based segmentation of resumes with an integrated skills knowledge base for efficient online recruitment. JRC's unique approach of extracting skills from segmented resumes and utilizing a knowledge base for classification yielded higher precision and efficiency compared to traditional methods. This system's effectiveness in automated job and resume classification showcases its potential for improving recruitment processes significantly.

Furthermore, Zaroor et al. [3] presented an innovative hybrid approach for matching candidate resumes with job postings. This approach leveraged advanced NLP techniques and regular expressions to extract critical skills from resumes, combining them with integrated skills knowledge bases like DICE or O*NET for classification. Their empirical experiments demonstrated higher precision results in less time compared to conventional methods, marking a substantial advancement in recruitment optimization.

Senthil Kumaran et al. [2] developed 'EXPERT,' an intelligent screening tool that utilizes semantic analysis and ontology mapping for candidate evaluation. EXPERT's semantic analysis and ontology mapping techniques enable comprehensive candidate assessments based on job requirements, achieving over 90% accuracy in CV screening. Future improvements aim to enhance mapping precision and broaden assessment criteria to include additional factors such as social network participation and research involvement, providing a holistic evaluation of candidates.

These diverse methodologies and approaches in the literature contribute significantly to the advancement of resume-based job retrieval systems, showcasing innovations in candidate assessment, job matching accuracy, and recruitment process optimization

IV. NOVELTY

Our approach to resume-based job ranking introduces several novel elements that enhance the traditional methods used in the job search and recruitment process. One of the key innovations is our system's ability to dynamically match job postings from the internet with candidate resumes in real time. Unlike static databases that may quickly become outdated, our dynamic matching ensures that job seekers are presented with the most current and relevant job opportunities available, maximizing their chances of finding suitable employment.

Another aspect of novelty lies in the advanced weighted tokenisation of extracted skill tokens with the help of the large language model Llama and Okapi's BM25. Instead of relying solely on keyword matching, our system goes deeper by extracting and analyzing key information from both resumes and job descriptions. This advanced analysis provides a more comprehensive understanding of candidate skills and

job requirements, leading to more accurate and meaningful job recommendations.

Our system also incorporates a personalized ranking algorithm that takes into account not just the relevance of skills but also factors such as experience level, location, education, Hobbies, career goals, etc. This personalized approach ensures that job matches are tailored to each candidate's unique profile, enhancing the precision and relevance of job recommendations.

Furthermore, the integration of Llama into our system allows it to continuously learn and adapt based on user interactions and feedback. This adaptive learning mechanism improves the system's accuracy and relevance over time, making it more effective in assisting users in their job search and recruiters in candidate selection. It also helps in increasing the robustness of the system and token extraction.
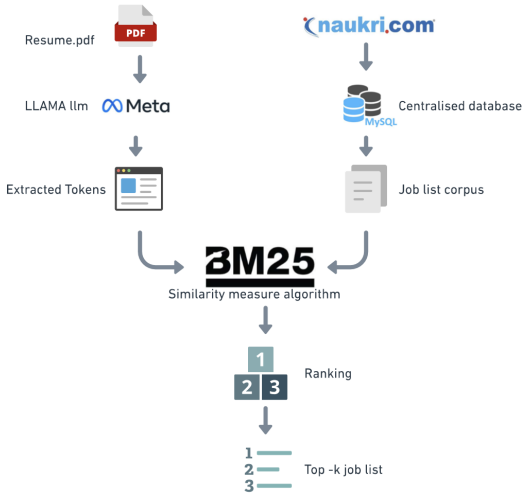

Fig. 1. Flowchart of proposed system.

V. METHEDOLOGY

A. *Data Collection and Preprocessing*

Our methodology begins with the collection of diverse job postings from online sources such as Naukri.com, ensuring a comprehensive representation of industries, job roles, and skill requirements. To standardize and improve processing efficiency, both job postings and resumes undergo preprocessing steps such as text normalization, tokenization, stop-word removal, stemming, etc. This preprocessed weekly job posting data is stored in pickle files for better and more efficient matching against the user resumes using Okapi's BM25.

The User resumes on the other hand would be inputted by the user on the platform, which would first be preprocessed and then with the help of Llama output relevant tokens such as 'primary skills', 'secondary skills', 'latest education', 'past experience', 'soft skills', 'location', 'hobbies', 'personality

uniqueness' etc which would then be matched against the job postings.

B. *Feature Extraction and Representation*

As mentioned in previous section that after preprocessing, we extract relevant features from resumes, including skills, education, experience, certifications, and achievements, using the Llama Large language model by meta. Similarly, job postings are also processed to extract key job requirements such as skills, qualifications, responsibilities, and industry-specific keywords against the user resume tokens. These extracted features are then structured and represented in pickle format suitable for BM25 comparison and matching during the job ranking process.

C. *Semantic Analysis and Matching*

Our methodology employs advanced semantic analysis techniques such as word embeddings and semantic similarity. Semantic similarity measures like cosine similarity on out baseline experimentation and okapi's BM25 for our baseline imrovement are utilized to quantify the similarity between candidate profiles and job descriptions accurately.

D. *Personalized Ranking and Recommendation*

The system integrates a personalized ranking algorithm that goes beyond semantic similarity. It considers additional weighed tokens which represent factors such as experience level, education, career goals, and location preferences to provide tailored job recommendations for each candidate. This personalized approach ensures that job recommendations align closely with the candidate's unique profile, enhancing the accuracy and relevance of the job ranking process.

After this, we display the Top 10 highest-matched job descriptions in our database with the user's resume and also provide the URL to apply straight to the job application page. The matching metric can also be chosen by the user from our menu for matching Through this comprehensive methodology, our resume-based job ranking system aims to revolutionize the job search and recruitment landscape by providing users with personalized, accurate, and relevant job recommendations while optimizing efficiency and effectiveness for recruiters.

VI. DATABASE

Our database serves as a comprehensive repository for job data scraped from Naukri.com on a weekly basis. The data includes job profiles and locations, focusing on prominent job markets like Bangalore, Delhi, Pune, and others. Job profiles encompass a wide range of roles such as ML Engineer, Software Developer, Web Developer, Network Analyst, and more, reflecting the diverse job landscape in the industry.

The structure of our MySQL database is designed to efficiently store and manage this job data. We have tables dedicated to storing job profiles, locations, Job application URLs

and other relevant information extracted from Naukri.com. These tables are interconnected to maintain data integrity and facilitate seamless data retrieval and analysis.

Additionally, we employ a systematic approach to data storage and management. Weekly job data is stored in the MySQL database, allowing for easy access and querying of job information. These data are then preprocessed and stored into a pickle file for better and more efficient matching against our OKapi BM25 similarity matching with weighted tokens extracted from the user resume.

By maintaining a structured and organized database, we ensure efficient data retrieval, analysis, and reporting for our job ranking and matching system. This database serves as a vital component in powering our system's capabilities to provide accurate and tailored job recommendations to users.

VII. CODE

The Algorithms that we are using for enabling this pipeline are mainly similarity-matching algorithms along without token weighting logic. In Our Project Baseline, we have used a cosine similarity measure algorithm for the matching process. The cosine similarity between two vectors $\mathbf{a}$ and $\mathbf{b}$ is given by:

$$\text{similarity}(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|}$$

In the Improved Project Baseline, We have used Okapi's BM25 similarity algorithm to accurately quantify the similarity between candidate profiles and job descriptions.

The BM25 formula is given by:

$$\text{BM25}(q, d) = \sum_{i=1}^{n} \text{idf}(t_i) \cdot \frac{f(t_i, d) \cdot (k_1 + 1)}{f(t_i, d) + k_1 \cdot \left(1 - b + b \cdot \frac{|d|}{\text{avgdl}}\right)} \quad (1)$$

where:

q : Query

d : Document

n : Total number of terms in the query or document

$f(t_i, d)$: Frequency of term t_i in document d

$|d|$: Length of document d

avgdl : Average document length in the corpus

k_1, b : Tuning parameters

VIII. EVALUATION

For evaluation purposes, we employ a standard resume dataset containing resume texts of individuals along with their current job titles. This dataset serves as a benchmark against which we assess the performance of our resume-based job ranking system. The evaluation process involves inputting the resume text of each individual into our portal and matching the resulting job recommendations with the actual job titles from the dataset.

To quantify the effectiveness of our system, we calculate the Precision@10 metric for each sample in the dataset. Precision@10 measures the proportion of relevant job recommendations among the top 10 ranked jobs provided by our system. A higher Precision@10 score indicates a more accurate and relevant job-matching capability.

Furthermore, we aggregate the Precision@10 scores across all samples in the dataset and compute the Mean Average Precision (MAP) as a summary evaluation metric. MAP provides a comprehensive assessment of the overall performance of our system in terms of job recommendation accuracy and relevance across the entire dataset.

This evaluation methodology ensures that we rigorously test and validate the job-matching capabilities of our system against a diverse set of resume profiles and job titles. By calculating Precision@10 and MAP, we obtain quantitative insights into the precision and effectiveness of our job ranking algorithm, enabling us to make data-driven decisions for system optimization and improvement.

Upon comparing this precision@10 among our baseline and final implementation, we have seen a 168% increase in results, which is a considerable margin. Our baseline implementations achieve a mean average precision of 0.25, whereas on the state of the art, we got 0.67 as mean average precision. These mean average precisions are across the precison@10 scores of every job prediction that we got.

Hence the results clearly demonstrate a huge improvement over the baseline system.

IX. CONCLUSION

In conclusion, our proposed solution portal aims to address the challenges faced by both job seekers and recruiters in the contemporary job market, mainly the IT tech industry. The sheer volume of job postings online, coupled with the complexities of job matching and recruitment processes, necessitates innovative solutions to streamline the job search experience and enhance recruitment effectiveness. Through the development of a resume-based job ranking system, we aim to revolutionize how job seekers navigate through job opportunities and how recruiters identify and attract suitable candidates. By leveraging advanced data analysis techniques and a comprehensive job database, our system provides personalized job recommendations tailored to individual skill sets, experience levels, and career aspirations.

The significance of our project lies in its potential to significantly improve the efficiency, accuracy, and overall user personal experience in the job search and recruitment processes. By addressing the inefficiencies and mismatches prevalent in the current job market, our system contributes to a more seamless and productive job market ecosystem. Moving forward, continuous refinement, evaluation, and adaptation of our system will be essential to ensure its ongoing effectiveness and relevance in the dynamic job market landscape. We further aim to enhance user satisfaction, optimizing job matching accuracy, and scaling our system to meet the evolving needs of

job seekers and recruiters alike. In essence, our resume-based job ranking system represents a transformative approach to job search and recruitment, offering tangible benefits to job seekers, recruiters, and the overall job market ecosystem.

X. FUTURE SCOPE

The development of our resume-based job ranking system lays a solid foundation for future enhancements and expansions. As we look ahead, several avenues of future scope and improvement emerge, aiming to further enhance the functionality, usability, and effectiveness of our system. Some key areas for future development include:

- **Integration of Evaluation Tests**: We plan to integrate evaluation tests directly into the portal, allowing job seekers to assess their skills, competencies, and readiness for specific job roles. This feature will not only empower job seekers with valuable insights but also improve the accuracy of job matching by considering test results alongside resume data.
- **Real-time Job Market Insights**: Enhancing the portal with real-time job market insights and trends will provide users with up-to-date information on in-demand skills, emerging job roles, and industry-specific developments. This feature will enable job seekers to make informed career decisions and recruiters to stay abreast of evolving talent requirements.
- **Expand Data Sources**: In addition to naukri.com, expanding our data sources to include other reputable job portals and industry-specific platforms will enrich our job database and broaden the scope of job recommendations. This expansion will ensure a more comprehensive coverage of job opportunities across various sectors and regions.

By focusing on these future scope areas, we aim to continuously enhance the value proposition of our resume-based job ranking system, provide a holistic and integrated job search and recruitment experience, and contribute to a more efficient and effective job market ecosystem.

REFERENCES

[1] Zaroor, Abeer, Mohammed Maree, and Muath Sabha. "JRC: a job post and resume classification system for online recruitment." In 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI), pp. 780-787. IEEE, 2017.

[2] Senthil Kumaran, V., and A. Sankar. "Towards an automated system for intelligent screening of candidates for recruitment using ontology mapping (EXPERT)." International Journal of Metadata, Semantics and Ontologies 8, no. 1 (2013): 56-64.

[3] Zaroor, Abeer, Mohammed Maree, and Muath Sabha. "A hybrid approach to conceptual classification and ranking of resumes and their corresponding job posts." In Intelligent Decision Technologies 2017: Proceedings of the 9th KES International Conference on Intelligent Decision Technologies (KES-IDT 2017)–Part I 9, pp. 107-119. Springer International Publishing, 2018.

[4] Jiechieu, Kameni Florentin Flambeau, and Norbert Tsopze. "Skills prediction based on multi-label resume classification using CNN with model predictions explanation." Neural Computing and Applications 33 (2021): 5069-5087.

[5] Gopalakrishna, Suhas Tangadle, and Vijayaraghavan Vijayaraghavan. "Automated tool for Resume classification using Semantic analysis." International Journal of Artificial Intelligence and Applications (IJAIA) 10, no. 1 (2019).

[6] Tejaswini, K., V. Umadevi, Shashank M. Kadiwal, and Sanjay Revanna. "Design and development of machine learning based resume ranking system." Global Transitions Proceedings 3, no. 2 (2022): 371-375.

[7] Wailthare, Kadambari, Aruna Tamhane, Vedashree Mulik, and Kirti Suryawanshi. "A COSINE SIMILARITY-BASED RESUME SCREENING SYSTEM FOR JOB RECRUITMENT."