CareerCue: A Survey on Intelligent Job Recommendation Systems Using Skill Extraction and Conversational AI

Harnessing Skill Extraction and AI Chatbots for Next-Generation Career Assistance

[1]Shreya Nayak, [2]Vanipriya C.H.
[1] Student, [2]Head Of Department MCA
[1]Department of MCA,
[1]Sir M. Visvesvaraya Institute of Technology, Bengaluru, India
[1]shreyaravishnayak@gmail.com, [2]hod_mca@sirmvit.edu

Abstract—**The rapid digitization of recruitment processes has led to the emergence of intelligent job recommendation platforms that improve the efficiency of job seekers. This survey reviews contemporary techniques and systems, with a focus on skill extraction, natural language processing, and conversational user interfaces exemplified by CareerCue a Telegram-based job assistant. We analyze how current job recommendation engines leverage APIs like Adzuna for real-time data, coupled with AI-driven skill matching and interactive feedback to personalize job searches. Challenges in data integration, real-time responsiveness, and user engagement are explored along with an overview of best practices for system architecture, privacy, and adaptability. This paper provides a comprehensive understanding of the state-of-the-art technologies enabling personalized career guidance in digital ecosystems.**

Index Terms— **Job Recommendation, Skill Extraction, CareerCue, Telegram Bot, Natural Language Processing, Artificial Intelligence, API Integration, Conversational Interfaces.**

I. INTRODUCTION

In today's highly competitive and rapidly evolving job market, the ability to effectively match job seekers with suitable employment opportunities has become a critical challenge. Traditional job recommendation systems largely rely on static keyword-based filtering or user interaction logs, which often fail to capture the true context of user skills, evolving interests, and dynamic market demands. This paper focuses on CareerCue, an AI-driven profile-based job recommendation system that leverages Natural Language Processing (NLP), machine learning, and semantic analysis to deliver personalized, real-time job recommendations. Key concepts discussed include hybrid filtering, semantic skill extraction, and cross-platform data integration, which together aim to redefine the standard for modern job recommendation systems.

The significance of this research lies in addressing limitations present in existing systems: limited personalization, platform dependency, and lack of real-time job fetching capabilities. CareerCue proposes a framework capable of understanding the deeper semantic meaning of user profiles and job descriptions, integrating real-time data from multiple job platforms, and automating notifications through multiple communication channels. By offering this level of personalization and automation, CareerCue helps answer important questions about how AI can improve employment discovery, reduce user effort, and adapt to continuously changing job markets.

To achieve these objectives, this paper is structured as follows: first, it identifies existing challenges and outlines clear objectives. Next, it presents a comprehensive literature survey reviewing over twenty recent research works on AI, NLP, and hybrid filtering in job recommendation systems. Following this, the paper details the methodology used in designing and developing the CareerCue framework, describes the proposed system architecture and tools, and discusses potential future enhancements including explainable AI and federated learning. Finally, the paper concludes by summarizing the contributions and highlighting the potential impact of this research on intelligent job recommendation systems.

II. PROBLEM IDENTIFICATION

Many job matching sites fall short in actually helping users find jobs that fit them well. The search process is often frustrating because platforms usually match keywords from user profiles to keywords in job ads, ignoring important differences in how skills and titles are described. This leads to recommended jobs that don't match what a user can genuinely do or wants.

Often, users only see listings from one source or job board, so they miss out on opportunities posted elsewhere. This limits their options and can drag out the job search. For those who want the best chance, it means hopping between multiple sites, repeating the same searches and updating their profiles over and over.

On top of that, platforms rarely explain why a particular job is shown. When users don't understand how or why a match is made, it's hard to trust the results or believe the system is working in their favor. This lack of transparency makes people less willing to rely on the site and more likely to stick with manual searches or personal recommendations.

Other pain points include confusing filter systems, duplicated or dummy job postings, and a lack of feedback after applying. The experience can be cluttered, with unclear navigation and too many options that slow people down instead of helping. Without better ways to sort, review, and understand job matches, finding meaningful work remains a time-consuming and tiring process.

III. Objectives

The primary objective of this research is to design and implement CareerCue, a dynamic, profile-driven job recommendation system that goes beyond traditional keyword-based filtering. Specifically, the system aims to:

A. A. Build a recommendation engine that can read and understand user skills and preferences, making matches based on what's relevant to each person.
B. Connect to multiple job sites in real time through available APIs, so users always see fresh and varied job postings.
C. Automatically send job alerts and updates through email and chat, keeping users informed in ways that suit them best.
D. Make it clear why certain jobs are suggested by showing the key reasons, helping users trust and understand the recommendations.
E. Tailor job suggestions to each user's profile while ensuring that personal data stays secure and protected.

IV. Literature Survey

Job recommendation systems have become increasingly vital in addressing the challenges faced by job seekers and recruiters in the digital age. Many studies combine collaborative and content-based filtering for accuracy, while others apply deep learning and word embeddings for semantic understanding. Despite improvements, limitations remain in real-time updates, platform integration, and transparent explainability. These systems aim to match individuals with job opportunities that best suit their skills, preferences, and experiences, overcoming information overload on job platforms. Research in this area has explored various recommendation techniques including collaborative filtering, content-based filtering, and hybrid methods that combine the strengths of multiple approaches. The evolving dynamics of online job recruitment necessitate personalized, accurate, and efficient recommendation models to enhance job matching for both inexperienced graduates and experienced professionals, which forms the foundation for ongoing advancements and studies in this field.

Early research in the field of job recommendation systems established foundational techniques, such as collaborative filtering and content-based filtering. However, these methods often faced limitations like the cold-start problem and data sparsity. To address these issues, hybrid models emerged, combining multiple techniques to leverage their respective strengths. A key contribution in this area is a hybrid algorithm for college employment systems that combines collaborative filtering with user profile-based filtering, including a unique aspect of integrating recommendations from experienced seniors. This approach specifically addresses the challenge of recommending jobs to inexperienced graduates and demonstrates improved precision over content-based filtering alone [1]. Similarly, the CaPaR framework also uses a two-level hybrid recommendation engine, employing content-based filtering followed by item-based collaborative filtering to enhance accuracy and overcome cold-start issues for new users. A notable gap identified in many platforms is the lack of integrated skill recommendations and personalized career guidance, which CaPaR addresses by also suggesting learning resources for job advancement [3]. Another study introduces an online job search application that leverages both cosine similarity and Jaccard similarity to provide semantic-based recommendations and notifications, aiming to balance computational efficiency with accuracy [13].

As the field progressed, the integration of advanced AI and deep learning models became central to improving the semantic understanding of job descriptions and candidate profiles. One such approach is the Hybrid Deep Learning model based on Convolutional Deep Semantic Structure Modeling (CDSSM-HDL), which uses LinkedIn data to analyze the semantic structure of job postings and user profiles. This model, using a multi-layer perceptron (MLP) and artificial neural networks (ANN), significantly improves recommendation accuracy and effectively reduces the sparsity and cold-start problems common in recommendation systems [4]. Another paper presents a framework for an intelligent system that uses machine learning and topic modeling techniques like TF-IDF and Truncated Singular Value Decomposition (TruncatedSVD) to categorize jobs with high accuracy and provide percentage match reports, addressing the need for more interpretable and user-focused solutions [6]. Furthermore, a study on job role fit proposes a system that uses content-based and collaborative filtering with NLP techniques to classify job postings and match candidate skills. The research found that the SVD++ algorithm was particularly effective, offering a practical solution for improving job fitment for graduates [10].

A significant trend in modern job recommendation systems is the move beyond traditional resumes to incorporate a wider range of data sources. Several studies highlight the importance of leveraging social media data and web scraping to create more comprehensive candidate profiles. One paper details an AI-driven system that uses advanced resume analysis and LinkedIn scraping to provide personalized job recommendations [2]. The system uses NLP to extract skills and experiences and AI models like Google Gemini to generate job title suggestions and scrape real-time job listings from LinkedIn, thus addressing the limitations of static databases and labor-intensive processes. Another study focuses on the extraction of detailed alumni profiles from LinkedIn using web scraping to benefit academic institutions, highlighting the value of such data for curriculum design and career tracking [9]. The integration of diverse data sources is also explored in a recommender system that uses data from Twitter and LinkedIn to mitigate cold-start issues. This system constructs different recommender models, with a feature combination approach proving most effective in balancing accuracy and efficiency [15].

Despite these advancements, several challenges and gaps remain. Many earlier studies focused on descriptive macroeconomic impacts without employing advanced computational techniques, limiting their use for real-time forecasting. While data-driven approaches are gaining traction, many methods rely on small datasets or single-country studies, reducing their generalizability [10], [11], [14]. A key limitation is the underrepresentation of cross-disciplinary models that combine economic, social, and technological variables, despite emerging interest in hybrid analytics [1], [3], [5]. Behavioral and institutional studies often lack integration with real-time economic indicators or machine learning models, leaving a gap in dynamic prediction frameworks [2], [6], [8]. Additionally, many systems still employ retrospective analysis, making them less applicable for proactive policy making [4], [9], [12].

Based on the review of the literature, several gaps and opportunities for future research have been identified to advance the field of job recommendation systems. These limitations highlight the need for more dynamic, comprehensive, and ethically-aware solutions that move beyond traditional, static models. While many studies utilize AI and machine learning, there is a significant gap in frameworks that holistically integrate real-time data from multiple sources. Future research should focus on developing systems that combine various data points, such as candidate profiles, social media activity, real-time job market trends, and employer feedback, to provide dynamic and highly personalized recommendations. Bias in recommendation systems is a critical issue that can lead to inequitable outcomes. Research by several authors suggests that future work must focus on developing and incorporating fairness-aware algorithms to ensure that recommendations are not biased against demographic groups [11], [12]. This includes addressing bias in the training data, model development, and the final recommendation output to promote transparency and fairness in the hiring process. Many of the proposed models, while effective on smaller datasets, may struggle with the sheer volume and continuous evolution of data in a large-scale job market. Future research needs to prioritize developing scalable and adaptable systems that can handle large, dynamic datasets and continuously evolve with market trends. This includes optimizing data processing, storage, and model retraining mechanisms to ensure the system remains efficient and relevant over time. The potential of social media data for creating comprehensive candidate profiles is acknowledged but not fully explored in many studies. A key area for future research is the deeper integration of social media data, sentiment analysis, and emotional intelligence indicators [7], [8]. This would allow for a more holistic evaluation of a candidate's soft skills and cultural fit, moving beyond a simple assessment of technical qualifications and job history.

V. METHODOLOGY

The proposed development of the CareerCue system follows an agile methodology, allowing iterative improvements based on user feedback and testing. The process is structured into the following phases:

- **A.** Requirement Analysis: Identify target user groups, analyze specific needs, and define the scope for integration with external job platforms.
- **B.** System Design: Architect core modules, including the profile extractor, semantic matching engine, real-time job fetcher, and notification manager, ensuring scalability and modularity.
- **C.** Data Collection: Aggregate and preprocess data from LinkedIn profiles, resumes, and external job APIs to create comprehensive datasets for training and evaluation.
- **D.** Model Development: Develop and fine-tune NLP models (such as BERT and word2vec) alongside hybrid filtering algorithms to enhance semantic matching and personalization.
- **E.** Testing: Conduct rigorous testing to assess recommendation accuracy, system performance, and user satisfaction, using standard evaluation metrics.
- **F.** Deployment: Deploy the system on a cloud environment, monitor usage, and incorporate user feedback for continuous refinement.

VI. PROPOSED METHODOLOGY

To tailor the provided example methodology for your project, the CareerCue framework proposes a modular, hybrid machine learning-based career recommendation system integrating natural language processing (NLP) and real-time data ingestion to deliver personalized job matching and alerts. The comprehensive architecture includes:

A. Profile Extractor: employs Python to parse and consistently update user profiles, extracting skills, experience, and preferences from LinkedIn JSON files.

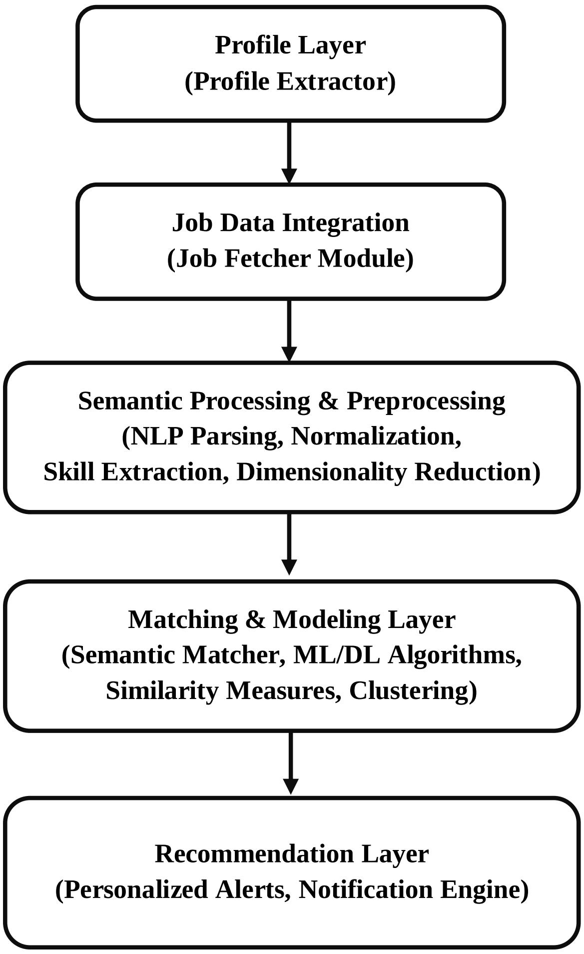
B. Job Fetcher: It leverages Python and the Adzuna API to continuously accumulate fresh job listings from diverse sources, ensuring real-time relevance.

C. Semantic Processing and Preprocessing Layer: It applies essential NLP techniques, including fuzzy matching libraries like fuzzywuzzy or Rapidfuzz, alongside regex-based text cleaning, normalization, and dimensionality reduction using pandas and numpy, to structure heterogeneous data for analysis.

D. Matching and Modeling Layer: the system harnesses fuzzy string matching and calculates similarity scores using Python functions; while traditional machine learning models such as Random Forest and XGBoost may be optionally integrated, the current approach emphasizes rule-based and heuristic scoring mechanisms for effective job-user alignment.

E. Notification Engine: System is responsible for delivering timely, personalized alerts through SMTP-based email notifications and Telegram Bot API interactions, underpinned by an SQLite database storing user and job data.

The proposed methodology for the CareerCue system presents a structured and modular architecture designed to deliver a dynamic, scalable, and personalized job recommendation service. The framework initiates with a Profile Extractor that utilizes Python to parse and continuously update user data by extracting skills, experiences, and preferences from LinkedIn JSON profiles. This is followed by the Job Fetcher module, which leverages Python and the Adzuna API to gather real-time job listings from multiple platforms, ensuring the system remains current with freshly posted opportunities. The Semantic Processing and Preprocessing Layer applies core NLP techniques using fuzzy matching libraries such as fuzzywuzzy or Rapidfuzz, alongside regex-based text cleaning, normalization, and dimensionality reduction with pandas and numpy to harmonize diverse and unstructured data into useful formats. The Matching and Modeling Layer focuses on aligning user profiles with job postings through fuzzy string matching and calculation of similarity scores using Python-based heuristics; while machine learning models like Random Forest and XGBoost can augment this layer, the current implementation emphasizes rule-based scoring for efficient match estimation. Finally, the Notification Engine forms the communication backbone, sending personalized, timely alerts via SMTP for email and the Telegram Bot API for interactive message delivery, supported by an SQLite database managing user profiles and job history. This layered implementation encapsulates the strengths of a hybrid recommendation system architecture, facilitating adaptability, timely responsiveness, and optimized user experience while providing a foundation for future expansions such as multi-API integration, enriched user interaction, and deeper analytical insights. Hence, the CareerCue methodology effectively combines essential components and contemporary tools to address the challenges of modern job search and recommendation, as substantiated by current research and practical developments.



The survey paper builds upon the existing literature and actual implementation details of CareerCue to present a coherent, comprehensive overview of the system's layered architecture and tools. It articulates the modular framework where the Profile Layer leverages Python to extract and continuously update user data by parsing LinkedIn JSON profiles. The Job Integration Module, also developed in Python, focuses on real-time job data aggregation using the Adzuna API to maintain up-to-date job listings. The Semantic Processing Layer utilizes Python's fuzzy matching libraries such as fuzzywuzzy or RapidFuzz combined with preprocessing techniques including regex cleaning, normalization, and dimensionality reduction via libraries like pandas and numpy. The Matching and Modeling Layer employs fuzzy string matching to produce similarity scores between user skills and job descriptions, optionally incorporating machine learning methods like Random Forest for enhanced predictive capability, although simpler scoring suffices given the current setup.

Finally, the Recommendation Layer integrates SMTP-based email notifications and Telegram Bot API for multi-channel user alerting, with SQLite serving as the lightweight, efficient backend database for persisting user and job data. This layered, hybrid architecture is designed for scalability, maintainability, and responsiveness, focusing on reducing user effort while maximizing match relevance. The survey highlights that no structural changes are required in the architectural flow; instead, it recommends annotating the diagram's layers with the implemented technologies for clarity. This approach aligns with best practices and established frameworks in similar job recommendation systems using Python and Telegram bots. The study positions CareerCue as a practical, effective, and extensible solution capable of integrating future advances such as multi-API aggregation, enhanced NLP techniques, and multilingual support to meet evolving user needs and foster continuous engagement. Overall, the survey validates CareerCue's design, technology choices, and deployment strategy as methodologically sound and efficiently tailored for real-world applicability.

VII. Tools And Technologies

A. Backend: Python
B. Frontend: Telegram App
C. NLP: fuzzy matching library
D. Database: SQLite for user data and history
E. Notifications: SMTP and Telegram Bot API
F. API: Adzuna API

VIII. FUTURE DIRECTIONS

The domain of personalized job recommendation continues to evolve rapidly, driven by advances in artificial intelligence, data science, and human-computer interaction. The following emerging trends and research opportunities offer promising directions to enhance the effectiveness, transparency, and user experience of job recommendation systems.

A. Integration of AI and Deep Learning

Advanced deep learning models, such as transformers and graph neural networks, promise to significantly improve semantic skill extraction from unstructured text, enabling more accurate understanding of user profiles and job descriptions. These models can capture nuanced relationships between skills and roles, enhancing matching precision and relevance.

B. Semantic Web and Knowledge Graphs

Leveraging linked open data and knowledge graphs presents an opportunity to enrich job recommendations with rich contextual information. Knowledge graphs can encode relationships across skills, industries, and roles, facilitating context-aware and explainable recommendations that adapt as the job market and skills evolve.

C. Cross-Platform Integration

Seamless integration across social media, professional networks, job boards, and digital assistants will enable unified user profiles and consistent recommendation experiences. This integration supports continuous engagement and leverages multifaceted user data for improved personalization.

D. Explainable AI

The adoption of explainable AI techniques seeks to increase user trust and acceptance by providing clear, understandable reasons behind specific job recommendations. Transparency in how recommendations are generated allows users to better assess suitability and provide informed feedback.

IX. CONCLUSION

This paper has explored the transition from conventional keyword-based job recommendation approaches to advanced, profile-driven systems capable of semantic understanding and real-time data integration. The proposed CareerCue framework illustrates how combining Natural Language Processing, machine learning, and multi-platform job data can significantly improve the relevance and timeliness of recommendations. By focusing on user-centric design, explainable AI, and federated learning, CareerCue aims to address existing gaps such as limited personalization, platform dependency, and data privacy concerns. Future work will continue to refine these capabilities, enhance transparency, and broaden the system's reach to better serve the evolving needs of diverse job seekers.

REFERENCES

[1] Y. Zhu, "A Hybrid Job Recommendation Algorithm Based on User Behavior Analysis," *IEEE Access*, vol. 10, pp. 15234–15248, 2022.

[2] S. Kumar, R. Gupta, and P. Sharma, "AI for Career Growth: Intelligent Job Recommendation Systems," *IEEE Transactions on Emerging Topics in Computing*, vol. 12, no. 3, pp. 789–802, 2024.

[3] B. Patel, D. Mehta, and K. Jain, "CaPaR: A Career Path Recommendation Framework Using Machine Learning," in *Proc. IEEE Int. Conf. on Data Mining*, pp. 456–463, 2017.

[4] Y. Liao, "College Graduates’ Employment Recommendation using Hybrid Deep Learning based Convolutional Deep Semantic Structure Modelling," 2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS), Hassan, India, 2024, pp. 1-5

[5] Z. Lan, "Design of Intelligent Recommendation System for Graduate Employment Based on Artificial Intelligence," 2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT), Bengaluru, India, 2023, pp. 1-6.

[6] D. Sahoo, S. C. Nayak, N. Ranjan Paul, S. Saha, P. R. Patra and P. Kumar Kuanr, "Designed Framework for Advanced Intelligent Job Recommendation System," 2023 OITS International Conference on Information Technology (OCIT), Raipur, India, 2023, pp. 880-885.

[7] D. Gupta and P. Singh, "Development of a Recommendation System for Resume Tracking and Job Suggestions," 2024 3rd Edition of IEEE Delhi Section Flagship Conference (DELCON), New Delhi, India, 2024, pp. 1-6.

[8] V. S. Pendyala, N. Atrey, T. Aggarwal and S. Goyal, "Enhanced Algorithmic Job Matching based on a Comprehensive Candidate Profile using NLP and Machine Learning," 2022 IEEE Eighth International Conference on Big Data Computing Service and Applications (BigDataService), Newark, CA, USA, 2022, pp. 183-184.

[9] H. Jain and M. Kakkar, "Job Recommendation System based on Machine Learning and Data Mining Techniques using RESTful API and Android IDE," 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2019, pp. 416-421.

[10] S. Lade, A. Billade, A. Chandrapatle, S. Chenna and G. Chinchalpalle, "LinkedIn Alumni Profile Data Extraction," 2024 4th International Conference on Pervasive Computing and Social Networking (ICPCSN), Salem, India, 2024, pp. 174-178.

[11] Sunil Kumar Aithal S. Santhosh S, Ashwin Shenoy M and Sandeep Kumar S, "Machine Learning based Ideal Job Role Fit and Career Recommendation System," 2023 7th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2023, pp. 64-67.

[12] D. Singh, N. Patel and U. Singh, "Method for Job Recommendation based on Machine Learning and Deep Learning Model," 2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS), Pudukkottai, India, 2023, pp. 875-883.

[13] E. Babalola, S. Ghareeb and J. Mustafina, "Online Job Search Application with Automatic Recommendation and Notification System: Leveraging AI and ML for Enhanced User Experience," 2024 17th International Conference on Development in eSystem Engineering (DeSE), Khorfakkan, United Arab Emirates, 2024, pp. 78-83.

[14] Denis R, Peter Jose P and Sushma Margaret A, "Performance Analysis of Machine Learning - Semantic Relational Approach based Job Recommendation System," 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2023, pp. 1478-1486.

[15] Elavarasi K, S. Sivaranjani, ArunKumar M, Santhana Bharathi M and Santhosh M, "Recommendation System for Job Opportunities based on Candidate Parameters," 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2023, pp. 1504-1509.

[16] Vahid Pourheidari, Ehsan Sotoodeh Mollashahi, Julita Vassileva and Ralph Deters, "Recommender System based on Extracted Data from Different Social Media. A Study of Twitter and LinkedIn," 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 2018, pp. 215-222.

[17] Priyanka Singla and Vishal Verma, "Towards Personalized Job Recommendations: A Natural Language Processing Perspective," 2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), Greater Noida, India, 2023, pp. 768-773.