A Survey on Large Language Models for Job Recommendations

Sarang Shelke, Aditya Shelke, Vedant Lodha & Yashraj Gaikwad
Department of Artificial Intelligence and Data Science
PVG's COET, Pune, India

Abstract—The rise of Large Language Models (LLMs) has transformed job recommendation systems by moving beyond traditional keyword-based matching and collaborative filtering to a more context-aware and intelligent approach. Leveraging deep contextual understanding and external knowledge, LLMs can analyze job descriptions and candidate profiles with greater accuracy, enabling more personalized and meaningful job rec-ommendations. Techniques such as fine-tuning and prompt engi-neering enhance their ability to establish complex relationships between candidates and job roles, improving overall matching quality. However, despite their potential, LLM-powered job recommendation systems face challenges such as bias, explain-ability, and scalability. Many existing models function as "black boxes," offering little transparency in how recommendations are generated, making it difficult for job seekers to interpret AI-driven career guidance. Furthermore, conventional methods are often restricted to retrieving and ranking job postings rather than acting as dynamic career assistants that proactively provide tailored suggestions. This survey categorizes existing LLM-based job recommendation approaches into discriminative and genera-tive paradigms, analyzing key methodologies, comparing different architectures, and identifying their strengths and limitations. We also discuss recent advancements, address ongoing challenges, and explore future research directions to enhance fairness, transparency, and adaptability in AI-driven hiring solutions.

Index Terms—Job Recommendation, Large Language Models, Context-Aware Matching, NLP, AI in Recruitment.

I. INTRODUCTION

The rapid growth of online recruitment platforms has fundamentally reshaped how job seekers connect with po-tential employers. As digital transformation accelerates, the significance of Artificial Intelligence (AI) is increasing in improving job search efficiency and recommendation accuracy. The Post-COVID era saw a massive shift toward online job platforms, with the global online recruitment market valued at USD 31.37 billion in 2023, expected to reach USD 62.29 billion by 2032 [1]. These platforms rely on recommendation systems to match candidates with job opportunities based on their skills, preferences, and career history. However, despite advancements in algorithmic job matching, current systems still suffer from fundamental limitations.

Most traditional job recommendation models rely on keyword-based matching techniques, which often fail to cap-ture the underlying contextual relationships between job roles and candidate profiles. For example, job postings for "Software Developer" and "Programmer" may not always be linked, despite their strong similarities in required skills and responsi-bilities. Similarly, contextual nuances are often overlooked—a marketing head in a tech company may require vastly different skills than one in a non-tech industry, yet recommendation engines may not distinguish these differences [5]. Additionally, job seekers frequently receive an overwhelming number of recommendations, making it difficult to find truly relevant opportunities. Searching for a software developer position on LinkedIn might return thousands of job listings, requiring significant manual effort to filter the most suitable ones. Furthermore, personalization remains a major challenge in mainstream job-matching systems, as many platforms struggle to account for factors like company culture, remote work preferences, and career aspirations.

The integration of Large Language Models (LLMs) in job recommendation systems aims to address these challenges by enhancing contextual understanding, improving personal-ization, and offering explainable AI-driven recommendations. Unlike traditional methods, LLMs leverage deep learning architectures to analyze job descriptions and resumes beyond simple keyword matching, capturing the semantic meaning of skills, experience, and career trajectories [6]. This survey ex-plores the role of LLMs in modern job-matching systems, cate-gorizing existing approaches into discriminative and generative models, analyzing their performance, and identifying future directions to improve fairness, transparency, and scalability in AI-powered hiring solutions.

II. RELATED WORK

Recent advancements in the design of job search engines focus on automating the scraping of job listings, matching users, and generating job descriptions. Rao et al. (2024) [2] propose a system that leverages frameworks such as Flask for designing web interfaces and automation libraries such as Selenium and BeautifulSoup for scraping websites for job listings. These techniques parallel the underlying data acquisition methods in modern job-matching systems, which utilize multi-threaded scraping to maximize throughput.

The performance of multi-threaded processes has been well documented. A comparative study by de Sa et al. (2012) [3] examines multi-threaded application performance under various scenarios, quantifying the gains and identifying bot-tlenecks when executing concurrent tasks across multi-core processors. These findings have direct implications for the design of scalable job recommendation systems.

The legal landscape for web scraping, especially of plat-forms like LinkedIn, has evolved significantly. Recent work by Padilha et al. (2024) [4] examines the technical, legal, and ethical hurdles of scraping LinkedIn data for research purposes, providing guidelines for responsible data acquisition. These studies underscore the need to balance the benefits of data-driven research with compliance to privacy laws and ethical norms.

Research by Zhang et al. (2022) [5] evaluates the application of large language models (LLMs) to job matching and skill alignment. The authors compare GPT-based, BERT-based, and custom Transformer models in classifying and ranking can-didate profiles against job postings, demonstrating improved semantic understanding and context handling for matching tasks. Similarly, Wu et al. (2023) [6] provide a comprehensive survey on LLMs for recommendation systems, highlighting their potential in contextual job matching.

III. BACKGROUND AND FUNDAMENTALS A. Traditional Job Recommendation Methods

Early job recommendation systems primarily relied on sim-ple rule-based and keyword-driven approaches. While effective for basic filtering, these methods often fail to understand the nuanced relationships between job roles and candidate profiles [5]. Below, we explore some of the most common traditional methods:

1) Keyword-Based Matching: Keyword-based approaches, such as TF-IDF and BM25, match job descriptions and re-sumes based on the presence of predefined terms. However, these systems struggle with synonyms and do not account for the broader context of job roles.

2) Rule-Based Systems: Rule-based job matching involves manually defining filters based on job title, location, experi-ence level, and other predefined criteria. While these systems provide structured recommendations, they lack adaptability and personalization.

3) Collaborative Filtering: Collaborative filtering recom-mends jobs based on user behavior patterns. These systems use past applications, clicks, and interactions to generate recommendations but often fail when dealing with new job seekers with limited history.

4) Graph-Based Approaches: Graph-based models repre-sent job seekers, recruiters, and job postings as interconnected nodes. These methods enhance personalization by leveraging relationships and common patterns in job transitions [6].

B. Introduction to Large Language Models (LLMs)

Large Language Models (LLMs) represent a significant advancement in Natural Language Processing (NLP). Unlike traditional models, which rely on manual feature engineering, LLMs use deep learning techniques to extract semantic mean-ing from job descriptions and resumes [6].

1) Key Innovations in LLMs: LLMs are built on transformer architectures, which introduce self-attention mechanisms to capture long-range dependencies in text. This enables them to:

- Understand job descriptions beyond keyword matching. • Identify hidden patterns in candidate profiles.
- Generate human-like recommendations based on job seeker preferences.

C. Discriminative vs. Generative LLMs

LLM-based recommendation systems can be broadly clas-sified into two categories [6]:

1) Discriminative LLMs (DLLM4Rec): Discriminative models focus on classification tasks, such as ranking job listings based on relevance to a candidate's profile. These models use embeddings to evaluate similarity scores between job descriptions and resumes [7].

2) Generative LLMs (GLLM4Rec): Generative models go beyond ranking by creating personalized job recommenda-tions. These models can:

- Generate custom job descriptions tailored to a candidate's skills.
- Provide conversational AI recommendations through chat interfaces.
- Assist in resume optimization by suggesting improved phrasing [8].

LLMs' ability to bridge the gap between structured job matching and contextual understanding makes them a trans-formative force in the recruitment industry. The next sections of this paper will explore state-of-the-art LLM applications, compare different model architectures, and highlight existing challenges in AI-driven job recommendations.

IV. SURVEY OF LLM-BASED JOB RECOMMENDATION SYSTEMS

A. Overview of LLMs in Job Recommendations

The integration of Large Language Models (LLMs) in job recommendation systems has transformed the way job seek-ers find relevant opportunities. Unlike traditional AI models, which rely heavily on keyword-based filtering, LLMs use deep contextual understanding to analyze job descriptions and candidate profiles [6]. Recent research has demonstrated that these models can improve job matching accuracy, automate resume screening, and enhance candidate-job alignment. This section categorizes and reviews existing LLM-based job rec-ommendation systems based on their underlying architectures.

B. BERT-Based Models for Job Matching

BERT-based models fall under the category of discrimina-tive LLMs (DLLM4Rec) [6], where the primary goal is to classify and rank job postings based on their relevance to a candidate's profile. Some notable approaches include:

- JobBERT [7]: A fine-tuned BERT model designed for job recommendation by embedding job descriptions and resumes into a shared vector space for semantic matching.
- BERT for Skill Extraction [7]: Models that leverage Named Entity Recognition (NER) to extract candidate skills and compare them with job requirements.

C. GPT-Based Models for Personalized Job Recommendations

Generative models, such as GPT-4 and ChatGPT [8], offer a more dynamic approach to job recommendations by generating personalized career suggestions. These models can:

- Generate job descriptions tailored to a candidate's skills and experience [10].

- Act as AI-powered career advisors, providing recommen-dations based on historical job market trends [9].

For example, some systems use GPT-4 to generate career paths and highlight potential skill gaps for job seekers.

D. Hybrid LLM Models (Graph + LLMs for Job Matching)

Recent research has explored combining LLMs with Graph Neural Networks (GNNs) to enhance job recommendation accuracy. Hybrid models:

- Utilize knowledge graphs to represent job seekers and job postings as nodes in a structured database.
- Improve explainability by leveraging structured relation-ships in job market data [6].

Graph-enhanced LLMs are particularly useful in ranking job recommendations more effectively while ensuring personal-ized results.

E. Comparative Analysis of LLM-Based Job Recommendation Models

TABLE I
COMPARISON OF LLM ARCHITECTURES IN JOB RECOMMENDATION SYSTEMS

Model	Type	Key Features
BERT	Discriminative	Semantic Job Matching, Skill Ex-traction, Resume Screening [7]
GPT-4.0	Generative	Contextual Career Guidance, Con-versational AI, Resume Optimiza-tion [8]
Claude 3.5 Sonnet	Generative	Advanced Reasoning, Bias Reduc-tion, AI Career Coaching
LLaMA 3.1	Generative	Lightweight, Efficient Inference, Low-Latency Applications
Gemini 2.0 Flash	Generative	Multimodal Processing, Speed Op-timization, Visual Job Matching
Mistral 7B	Hybrid	Open-Source, Fine-tuning, Custom Job Matching

While specific studies on job recommendation applications for models like Claude 3.5 Sonnet, LLaMA 3.1, Gemini 2.0 Flash, and Mistral 7B are limited, their advanced natural language processing capabilities make them suitable for such tasks, as demonstrated by recent research on LLMs in this domain [12], [13].

F. Applications of LLM-Based Job Recommendation Systems

Large Language Models (LLMs) have transformed job recommendation systems by enabling intelligent, personalized, and scalable solutions. Their applications span across various domains, including recruitment platforms, career guidance, and automated hiring.

1. AI-Powered Resume Screening

LLMs automate resume screening by extracting key skills, experience, and qualifications, significantly reducing the work-load for recruiters [5].
Example: AI-driven Applicant Tracking Systems (ATS) use models like BERT and GPT-4 to match resumes and job descriptions.

2. Personalized Job Recommendations

Unlike traditional keyword-based matching, LLMs leverage deep contextual understanding to recommend jobs tailored to a candidate's profile [6].
Example: LinkedIn and Indeed employ AI-powered recom-mendation engines to enhance job matching accuracy [14], [15].

3. AI-Powered Career Counseling

Conversational AI models serve as virtual career advisors, pro-viding insights into career progression and skill development [9].
Example: AI chatbots assist job seekers by analyzing market trends and suggesting optimal career paths.

4. Automated Job Description Generation

LLMs generate well-structured job descriptions that attract the right candidates by incorporating industry-specific language and essential skill requirements [10].
Example: Companies use AI to create detailed job postings tailored to specific hiring needs.

5. Fraud Detection in Job Applications

LLMs assist in identifying fraudulent job postings and de-tecting fake resumes by analyzing linguistic patterns and inconsistencies [11].
Example: AI-based fraud detection systems flag suspicious job listings, reducing scams in online job markets.

6. Skill Gap Analysis and Workforce Upskilling

Skill gaps in candidates can be identified by analyzing job trends using LLMS that recommend necessary training pro-grams to enhance employability [16], [17].
Example: Educational platforms like Coursera use AI to sug-gest relevant courses for users.

G. Challenges in LLM-Based Job Recommendation Systems

Despite their advantages, Large Language Models (LLMs) in job recommendation systems face several challenges that impact their effectiveness, fairness, and scalability. Addressing these issues is necessary to ensure reliable and ethical AI-driven hiring solutions.

1. Bias in AI Recommendations

LLMs may reinforce existing biases present in training data, leading to gender, racial, or occupational disparities in job recommendations. If not properly managed, AI systems can disproportionately suggest certain roles to specific demograph-ics, limiting diversity and inclusion in hiring processes.

2. Explainability and Transparency Issues

One major limitation of LLM-based job matching is the lack of interpretability. These models often function as black-box systems, making it difficult for job seekers and recruiters to understand why a particular job was recommended. This reduces trust in AI-driven hiring and raises concerns about accountability.

3. Scalability and Computational Costs

Deploying and maintaining LLMs at scale requires significant computational resources. Generating real-time job recommen-dations involves high processing power, making it costly for companies to integrate LLMs into their platforms. Efficient model compression and inference optimization techniques are required to address this challenge.

4. Data Privacy and Security Concerns
LLMs handle sensitive user data, including resumes, employment history, and personal preferences. Ensuring and implementing strong data protection checks is necessary to safeguard user privacy. Additionally, implementing robust encryption and anonymization techniques is necessary to prevent data breaches.

5. Mismatch Between AI Predictions and Real-World Hiring Trends
AI-driven job recommendations do not always align with real-world industry requirements. LLMs rely on historical data, which may not accurately reflect emerging job trends or evolving skill requirements. Continuous model retraining and integration of real-time labor market insights are essential to maintain relevance.

Tackling these issues can help in improving the accuracy, fairness, and adoption of LLM-based job recommendation systems. Future research should focus on mitigating bias, enhancing model explainability, optimizing resource efficiency, and strengthening data privacy measures.

H. Future Research Directions

To further improve the effectiveness and adoption of LLM-based job recommendation systems, future research must address several key areas:

1. Mitigating Bias in AI Recommendations
Future studies should explore fairness-aware AI models to minimize bias in job matching. Developing diverse training datasets and incorporating bias-mitigation techniques can lead to more equitable job recommendations.

2. Enhancing Explainability and Transparency Integrating Explainable AI (XAI) methods can help make LLM-based job recommendations more interpretable. By pro-viding justification for each recommendation, these models can build trust among users and recruiters.

3. Optimizing Computational Efficiency
Current LLMs require significant computational resources, limiting their scalability. Research into model quantization, distillation, and efficient inference techniques can make real-time job recommendations more cost-effective.

4. Improving Real-Time Labor Market Adaptation
Traditional job recommendation models struggle to adapt to rapid changes in job market demands. Future models should leverage reinforcement learning and dynamic retraining tech-niques to ensure job recommendations remain relevant.

5. Strengthening Data Privacy and Security
Techniques such as federated learning and advanced encryp-tion methods should be explored to enhance privacy. Ensuring compliance with global data protection laws will be crucial in increasing user trust.

Advancing these research directions will enhance the reliability, fairness, and effectiveness of LLM-driven job recommendation systems, making them more adaptable to the evolving job market.

V. CONCLUSION

Large Language Models (LLMs) have significantly en-hanced job recommendation systems by enabling contextual understanding, personalized recommendations, and automation in recruitment processes. This paper explored various appli-cations of LLMs in job recommendations, highlighting their ability to streamline resume screening, improve job matching accuracy, and assist in career guidance.

However, LLM-based job recommendations still face challenges related to bias, explainability, computational efficiency, and privacy concerns. Addressing these limitations is crucial to ensure fair and scalable AI-driven hiring systems. Future research should focus on mitigating biases, improving model interpretability, enhancing efficiency, and ensuring compliance with data protection standards.

As AI-powered recruitment continues to evolve, the responsible and ethical use of LLMs will be key to shaping the future of job matching. By leveraging advancements in AI while addressing existing challenges, LLM-driven job recommendation systems have the potential to revolutionize the hiring landscape and make it an easier place to search for jobs.

REFERENCES

[1] SkyQuest Technology Consulting, "Online Recruitment Market Report," SkyQuest, 2024. [Online]. Available: https://www.skyquestt.com/report/ online-recruitment-market. [Accessed: Mar. 20, 2025].

[2] P. P. Rao, K. Radhakrishnan, A. N. R. Ayshwarian, and P. Revathi, "Job searching using web scraping," International Research Journal of Mod-ernization in Engineering Technology and Science, vol. 6, no. 4, pp. 1–6, Apr. 2024. [Online]. Available: https://www.irjmets.com/uploadedfiles/ paper/issue_4 april_2024/54878/final/fin_irjmets1714674060.pdf.

[3] A. de Sa, M. Pereira, P. Moura, and L. Peixoto, "A comparative study of multithreaded applications performance in different scenarios," Revista de Sistemas de Informaçao da FSMA, no. 9, pp. 45–53, 2012.

[4] A. Padilha and J. Mena-Chalco, "Navigating technical, legal, and ethical hurdles to scraping LinkedIn data for academic research," Liinc em Revista, vol. 20, no. 1, p. e7041, 2024. [Online]. Available: https: //doi.org/10.18617/bm06ge67.

[5] Y. Zhang, H. Liu, and X. Wang, "Automated resume parsing and job matching using BERT," in Proc. IEEE Int. Conf. Artificial Intelligence and Employment, 2022, pp. 123–130.

[6] L. Wu, Z. Zheng, Z. Qiu, H. Wang, H. Gu, T. Shen, C. Qin, C. Zhu, H. Zhu, Q. Liu, H. Xiong, and E. Chen, "A survey on large language models for recommendation," arXiv preprint arXiv:2305.19860, May 2023. [Online]. Available: https://arxiv.org/abs/2305.19860.

[7] J.-J. Decorte, J. Van Hautte, T. Demeester, and C. Develder, "JobBERT: Understanding job titles through skills," arXiv preprint arXiv:2109.09605, Sep. 2021. [Online]. Available: https://arxiv.org/abs/ 2109.09605.

[8] Y. Zhang, Y. Jin, L. Chen, and T. Yang, "Navigating user experience of ChatGPT-based conversational recommender systems: The effects of prompt guidance and recommendation domain," arXiv preprint arXiv:2405.13560, May 2024. [Online]. Available: https://arxiv.org/abs/ 2405.13560.

[9] T. Brown, R. Patel, and L. Gomez, "AI-driven career guidance: The role of large language models in employment counseling," International Journal of AI and Human Resources, pp. 205–219, 2023.

[10] K. Lee and A. Kumar, "Automated job description

generation using generative AI," IEEE Transactions on Computational Intelligence, vol. 12, no. 3, pp. 89–102, 2023.

[11] R. Gupta, P. Shah, and Y. Lin, "Fraud detection in online job portals using transformer-based models," in Proc. 2022 IEEE Conf. AI in Security, 2022, pp. 312–320.

[12] Z. Zheng et al., "Generative job recommendations with large language model," arXiv preprint arXiv:2307.02157, Jul. 2023. [Online]. Available: https://arxiv.org/abs/2307.02157. [Accessed: Apr. 11, 2025].

[13] N. Li, B. Kang, and T. De Bie, "LLM4Jobs: Unsupervised occupation extraction and standardization leveraging large language models," arXiv preprint arXiv:2309.09708, Sep. 2023. [Online]. Available: https://arxiv. org/abs/2309.09708. [Accessed: Apr. 11, 2025].