

# An Efficient Bi-Encoder-Based Skill Classification and Employer Notification Model for Smart Job Market Analysis

**Jangam Akshaya**, department of Computer Science and Engineering, GNITC, 22-5B5,  
22wj1a05b5gniindia.org@gmail.com

**K.Varshith Kumar** , department of Computer Science and Engineering, GNITC, 22-5C2,  
22wj1a05c2@gniindia.org

**K.Hemanth Kumar** , department of Computer Science and Engineering, GNITC, 22-5D3,  
22wj1a05d3@gniindia.org

**P.Sunil**, Assistant Professor, department of Computer Science and Engineering, GNITC,  
psunil.csegnitc@gniindia.org

\*\*\*

**Abstract-** The modern labour market requires efficient analysis of unstructured job advertisements to identify required skills and job roles. However, many existing labour market analysis systems rely on large language models (LLMs), which are computationally expensive and difficult to deploy in real-time environments. To address this limitation, this project proposes an AI-enabled Job Portal System that integrates lightweight natural language processing techniques for intelligent recruitment analysis. The system utilizes a bi-encoder-based contrastive learning approach inspired by ConTeXT-match to extract relevant skills from job advertisements and map them to standardized taxonomies such as ESCO. In addition, JobBERT V2-based job title normalization ensures consistent classification of diverse job titles. The platform allows applicants to register their educational and skill profiles, which are automatically matched with job requirements. A Gmail-based notification system provides real-time updates regarding application status. The proposed system improves recruitment transparency, computational efficiency, and scalability for smart labour market analysis.

**Key Words:** Job Portal, Skill Extraction, ConTeXT-match, Job Title Normalization, Labour Market Analysis, Gmail Notification System

## 1.INTRODUCTION

The modern labour market is rapidly evolving due to technological advancements and changing workforce demands. Job advertisements provide valuable information about job roles and required skills, making them an important source for labour market analysis. However, these job descriptions are typically written in unstructured text and often use different terminology for similar skills and occupations, making large-scale analysis difficult. Recent research has applied natural

language processing (NLP) and machine learning techniques to automatically extract skills and normalize job titles from job advertisements. While large language models (LLMs) can achieve high accuracy, they require significant computational resources and are not always suitable for real-time applications. To address these challenges, this project proposes an AI-enabled Job Portal System that efficiently extracts skills from job advertisements, matches them with applicant profiles, and provides real-time Gmail-based notifications to improve transparency and recruitment efficiency

## 2. LITERATURE REVIEW

Labour market analysis and recruitment systems have increasingly relied on data-driven techniques to understand job trends, required skills, and workforce demands. With the rapid growth of online job portals, large volumes of job advertisements are generated daily, containing valuable information about occupations and required competencies. However, these job descriptions are often written in unstructured text and use different terminology to describe similar roles and skills, making automated analysis challenging. Recent advancements in natural language processing (NLP) and machine learning have significantly improved the ability to extract meaningful insights from job advertisements. Napat Laosaengpha et al. (2024) proposed a framework for learning job title representations from job descriptions using a job description aggregation network. Their approach analyzes complete job descriptions rather than relying only on extracted skills, allowing the model to capture contextual relationships between job titles and responsibilities. The proposed method improves job title classification accuracy and provides a more comprehensive representation of job roles. Ibrahim Rahhal et al. (2024) presented a detailed survey on the application of data science techniques in labour market

analysis. The authors reviewed numerous studies that utilize machine learning, natural language processing, and data mining techniques to analyze job market trends, extract skills, and support decision-making in workforce planning. Their work highlights how automated analysis of job advertisements can provide valuable insights into evolving industry skill demands. Karel D'Oosterlinck et al. (2024) introduced an innovative framework called Infer–Retrieve–Rank for solving extreme multi-label classification problems using language models. The proposed approach combines retrieval-based methods with language models to efficiently assign relevant labels in datasets containing thousands of possible classes. This framework demonstrates improved classification performance while reducing the complexity associated with traditional approaches. Mike Zhang et al. (2024) proposed the Nearest Neighbor Occupational Skill Extraction (NNOSE) method for extracting occupational skills from job advertisements. The system leverages retrieval-based techniques to identify similar skills across multiple datasets, improving the prediction of rare or infrequently occurring skills. Experimental results showed significant performance improvements, particularly in cross-dataset evaluation scenarios. Boyang Shang et al. (2023) introduced a diversity subsampling method designed to select representative samples from large datasets while preserving the distribution of the original data. This approach improves training efficiency for machine learning models by ensuring that selected samples provide a comprehensive representation of the dataset. Their method demonstrated improved performance compared to traditional subsampling techniques in large-scale machine learning tasks. Jens-Joris Decorte et al. (2023) focused on the problem of extreme multi-label skill extraction using large language models. The authors proposed a contrastive learning approach combined with synthetic dataset generation to improve the performance of skill extraction systems. Their method achieved significant improvements in precision and recall across multiple benchmark datasets, demonstrating the effectiveness of contrastive learning for large-scale labour market analysis. Collectively, these studies highlight the growing importance of machine learning and natural language processing techniques in analyzing job advertisements, extracting relevant skills, and supporting intelligent recruitment systems.

### 3. RELATED WORK

Several studies have explored the use of machine learning and natural language processing techniques for labour market analysis and intelligent recruitment systems. Early

research primarily focused on extracting skills from job advertisements using traditional natural language processing methods and rule-based systems. However, these approaches often struggled to handle the large diversity of job descriptions and the ambiguity present in unstructured textual data. With the advancement of deep learning and large-scale language models, more sophisticated approaches have been proposed to improve the accuracy and scalability of skill extraction systems. Decorte et al. (2023) proposed an approach for extreme multi-label skill extraction using large language models and contrastive learning strategies. Their work demonstrated that generating synthetic labeled datasets can significantly improve the performance of skill extraction models when dealing with thousands of possible skill labels. Similarly, D'Oosterlinck et al. (2024) introduced the Infer–Retrieve–Rank framework for solving extreme multi-label classification problems using language models and retrieval-based techniques. This framework allows efficient identification of relevant labels from large taxonomies and has shown competitive performance across several benchmark datasets. Zhang et al. (2024) proposed the Nearest Neighbor Occupational Skill Extraction (NNOSE) method, which improves the prediction of rare skills by retrieving similar skills from external datasets. Their approach demonstrates improved performance in cross-dataset evaluation and provides better generalization for unseen skill categories. In another study, Laosaengpha et al. (2024) introduced a job description aggregation network for learning job title representations directly from job descriptions. This method improves job title normalization by capturing the semantic relationship between job titles and their associated descriptions. Rahhal et al. (2024) conducted a comprehensive survey on the application of data science techniques in job market analysis, highlighting how machine learning and natural language processing methods can be used to analyze labour market trends, skill demands, and workforce development strategies. Their study emphasizes the importance of automated systems for processing large volumes of job advertisement data. Despite these advancements, many existing systems rely heavily on computationally expensive large language models, which may not be suitable for real-time applications or large-scale deployment. Therefore, there is a growing need for lightweight and efficient models that can accurately extract skills and normalize job titles while maintaining scalability. The proposed system addresses these challenges by integrating a lightweight bi-encoder architecture inspired by ConTeXT-match with job title normalization and a real-time notification system

to enhance recruitment efficiency and labour market analysis.

#### 4. PROPOSED METHODOLOGY

The proposed system introduces an AI-enabled Job Portal designed to improve labour market analysis and recruitment efficiency by integrating advanced natural language processing techniques with a scalable web-based platform. The primary objective of the system is to automatically extract relevant skills from job advertisements, normalize job titles, and match them with suitable applicant profiles. Job advertisements typically contain unstructured textual information that describes job responsibilities, qualifications, and required skills. However, the use of diverse terminology and inconsistent formatting across job postings makes it difficult to perform large-scale automated analysis. To address this challenge, the proposed system utilizes a lightweight skill extraction framework inspired by the ConTeXT-match model, which applies contrastive learning and token-level attention to identify relevant skills within job descriptions.

The system follows a multi-stage architecture consisting of job data upload, skill extraction, job title normalization, applicant profile management, and notification services. In the first stage, employers upload job advertisements through a web interface or file upload mechanism. These job descriptions are stored in a structured database for further processing. In the second stage, the Skill Extraction Engine processes the textual content of the job advertisement. The text is first tokenized and converted into numerical embeddings using a transformer-based encoder. The system then applies a contrastive learning strategy that compares job description sentences with a predefined taxonomy of skills, such as those defined in the ESCO framework. By mapping textual information to standardized skill representations, the system ensures consistent identification of relevant competencies across different job advertisements.

Once the skills are extracted, the next stage performs job title normalization using a model inspired by JobBERT V2. This component analyzes raw job titles and contextual information from job descriptions to map them to standardized occupational titles. This normalization process reduces redundancy and improves consistency across job listings, enabling more accurate analysis and matching of job opportunities. In the next stage, applicants register on the platform and create profiles containing their educational background, experience, and

skill sets. The system compares the extracted job requirements with applicant profiles to identify suitable matches and generate job recommendations. Finally, the platform integrates a Gmail-based notification service that provides real-time updates to applicants regarding the status of their applications, such as under review, shortlisted, or rejected. This automated communication mechanism improves transparency and user engagement while eliminating the need for frequent manual checks on the portal. By combining efficient machine learning techniques, scalable architecture, and automated notification mechanisms, the proposed system provides a practical solution for intelligent recruitment and labour market analysis.

#### 5. RESULTS AND DISCUSSION

The proposed AI-enabled Job Portal system was successfully developed and tested to evaluate its effectiveness in job-skill matching and recruitment automation. The system integrates job posting, applicant registration, resume analysis, skill matching, and automated email notification modules to provide a complete recruitment solution.

##### Home Page

Initially, the home page of the system provides an intuitive interface where users can navigate to different modules such as admin, employer, login, and registration. The interface displays key statistics related to job listings, companies, and job seekers to provide users with a quick overview of the platform. The system is designed with a modern user interface to improve usability and user engagement.

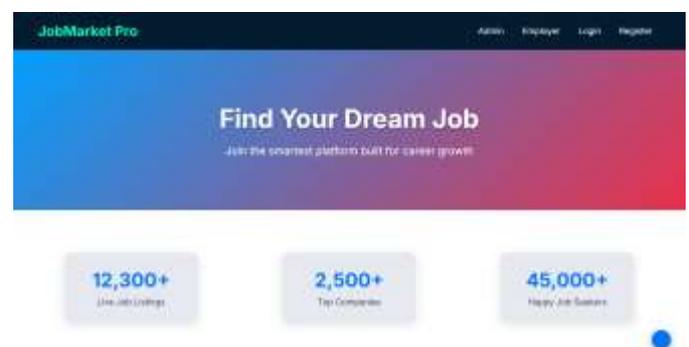


Fig 1: Home page of the JobMarket AI portal

##### Registration Page

Applicants can create an account by providing personal information, educational qualifications, and uploading their resume in PDF format. The registration module

stores applicant details securely in the database and prepares the profile for further job matching analysis.



Fig 2: Applicant registration interface

### Employee login Page

After successful registration, users can log into the system through the employee login interface. The login module verifies user credentials and grants access to the job application and resume analysis modules.



Fig 3: Employee login interface

### Employer Login Page

The Employer Login Interface allows registered employers to securely access the job portal by entering their email and password. After successful authentication, employers can access their dashboard to post job listings, manage existing job advertisements, and review

applications from job seekers. The interface is designed to be simple and user-friendly, enabling efficient management of recruitment activities within the platform.

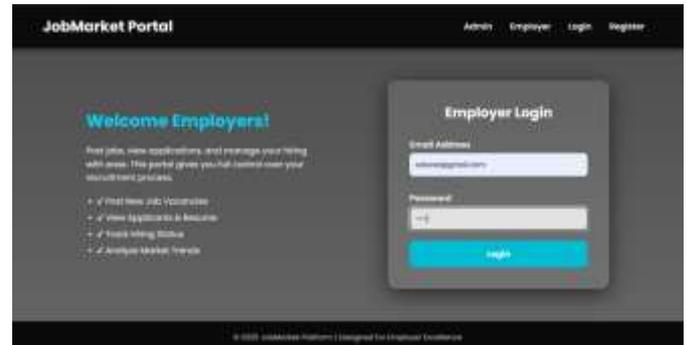


Fig 4: Employer login interface

### Employer Portal

Employers can post job opportunities through the portal. The system displays available job listings including job title, description, required skills, and priority skills. Applicants can browse these job listings and apply directly through the platform.

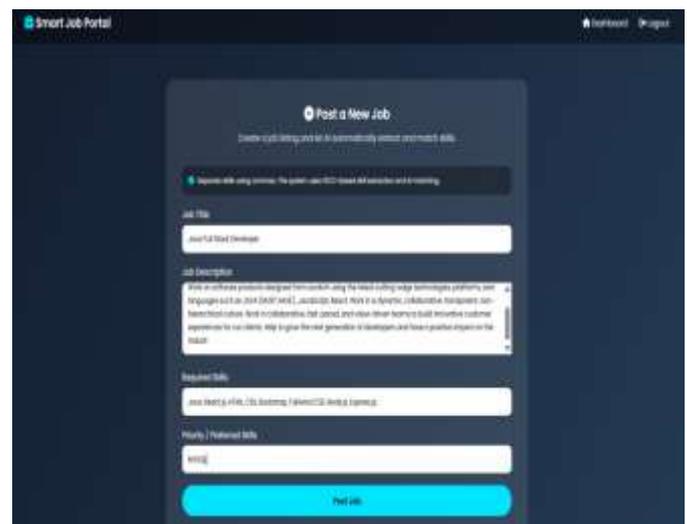


Fig 5: Job listings and skill requirements

### Employee Job Application Interface

The Employee Job Application Interface allows job seekers to apply for available job positions through the portal. In this module, applicants enter their basic details such as name and email address and upload their resume in PDF format. After submitting the resume, the system analyzes the applicant's skills and compares them with the job requirements. The interface also provides application insights through visual charts that display the application status and skill matching results, helping applicants understand their eligibility for the selected job role.

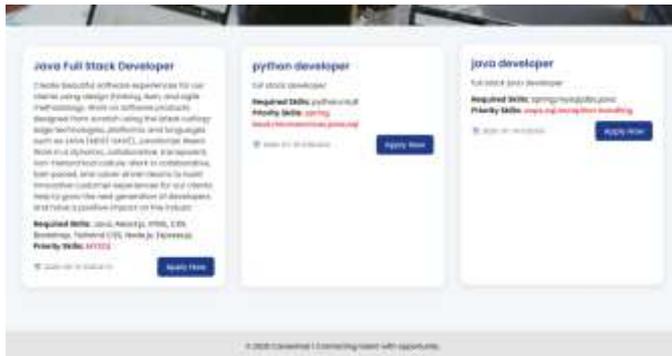


Fig 6: Job List

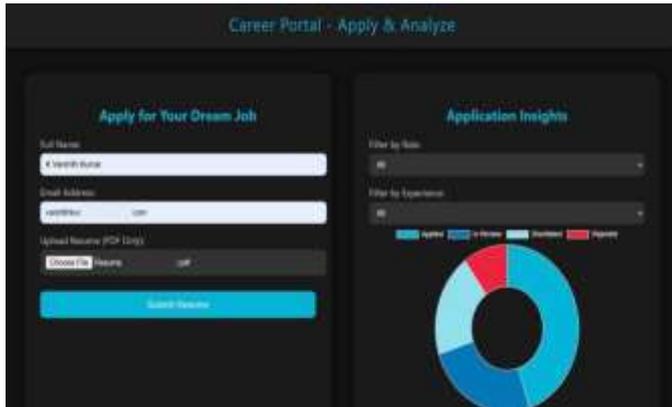


Fig 7: Applying for a Job



Fig 8: Resume match analysis showing required skills match and priority skills match

### Automated Gmail Notification

Once the resume analysis is completed, the system automatically sends a notification email to the applicant informing them about the job match results. This Gmail-based notification system improves communication between the platform and applicants and ensures real-time updates about application status.



Fig 9: Automated Gmail notification sent to the applicant

## 6. CONCLUSION

This project presented the design and implementation of an AI-enabled Job Portal system aimed at improving the efficiency and transparency of the recruitment process. The system integrates multiple modules including employer job posting, applicant registration, resume upload, skill extraction, resume analysis, and automated notification services. By utilizing intelligent resume analysis and skill matching techniques, the platform is able to compare applicant skills with job requirements and identify suitable candidates for specific job roles. This reduces the manual effort required in traditional recruitment processes and helps employers quickly identify potential candidates based on relevant skills.

The developed system also provides an interactive and user-friendly interface for both employers and job seekers. Applicants can easily register, upload their resumes, and apply for jobs directly through the portal. The resume analysis module processes the uploaded resumes and identifies matching and missing skills with respect to the selected job role. The results are presented using graphical charts and detailed reports, allowing applicants to clearly understand their suitability for the job. This feature also helps candidates identify areas where they need to improve their skills in order to increase their chances of selection.

In addition, the integration of Gmail-based notifications enhances communication between the platform and applicants by automatically sending updates regarding job match results and application status. This real-time notification mechanism improves user engagement and ensures that applicants remain informed throughout the recruitment process. The experimental results demonstrate that the system successfully performs job listing management, resume analysis, skill matching, and automated communication in an efficient and reliable manner.

Overall, the proposed AI-enabled Job Portal provides a practical solution for modern recruitment platforms by combining automation, intelligent skill analysis, and real-time notifications. The system improves recruitment efficiency, enhances transparency in the hiring process, and offers a scalable framework that can be further expanded to support more advanced AI-driven recruitment functionalities in the future.

## 7. FUTURE SCOPE

In the future, the system can incorporate deep learning-based natural language processing models to improve the accuracy of skill extraction from resumes and job descriptions. Additional features such as real-time job recommendations, interview scheduling, and automated resume ranking can also be implemented to further streamline the recruitment process. The platform can be expanded to support mobile applications, allowing users to access job opportunities and receive notifications more conveniently. Integration with professional networking platforms and external job databases could further increase the availability of job listings. Moreover, advanced analytics dashboards can be developed for employers to monitor recruitment trends and candidate statistics. These improvements would enhance the scalability, intelligence, and usability of the system, making it more effective for modern recruitment and labour market analysis.

## REFERENCES

- [1] I. Rahhal, I. Kassou, and M. Ghogho, “Data science for job market analysis: A survey on applications and techniques,” *Expert Systems with Applications*, vol. 251, 2024.
- [2] W. V. Cunningham and P. Villaseñor, “Employer voices, employer demands, and implications for public skills development policy,” *World Bank Research Observer*, vol. 31, no. 1, pp. 102–134, 2016.
- [3] I. Khaouja, I. Kassou, and M. Ghogho, “A survey on skill identification from online job ads,” *IEEE Access*, vol. 9, pp. 118134–118153, 2021.
- [4] B. Clavié and G. Soulié, “Large language models as zero-shot ESCO skills matchers,” *RecSys in HR Workshop*, 2023.
- [5] K. D’Oosterlinck et al., “In-context learning for extreme multi-label classification,” 2024.
- [6] P. Safikhani et al., “Automated occupation coding with hierarchical features,” *Discover Artificial Intelligence*, vol. 3, 2023.
- [7] J. J. Decorte et al., “JobBERT: Understanding job titles through skills,” *ECML-PKDD Workshop*, 2021.
- [8] M. Zhao et al., “SKILL: A system for skill identification and normalization,” *AAAI Conference on Artificial Intelligence*, 2015.
- [9] S. Jia et al., “Representation of job-skill in artificial intelligence with knowledge graph analysis,” *IEEE ISPCE-CN*, 2018.
- [10] L. Sayfullina et al., “Learning representations for soft skill matching,” *Springer*, 2018.
- [11] M. Zhang et al., “SkillSpan: Hard and soft skill extraction from job postings,” *NAACL-HLT*, 2022.
- [12] M. Zhang et al., “NNOSE: Nearest neighbor occupational skill extraction,” *EACL*, 2024.
- [13] A. Herandi et al., “Skill-LLM: Repurposing general-purpose LLMs for skill extraction,” 2024.
- [14] J. J. Decorte et al., “Design of negative sampling strategies for distantly supervised skill extraction,” 2022.
- [15] A. Gnehm et al., “Fine-grained extraction and classification of skill requirements in job advertisements,” *NLP+CSS Workshop*, 2022.
- [16] J. J. Decorte et al., “Extreme multi-label skill extraction training using large language models,” *AI4HR Workshop*, 2023.
- [17] T. Jiang et al., “LightXML: Transformer with dynamic negative sampling for extreme multi-label text classification,” *AAAI Conference*, 2021.
- [18] K. Dahiya et al., “DeepXML: A deep extreme multi-label learning framework applied to short text documents,” *ACM WSDM*, 2021.
- [19] J. Zhang et al., “Fast multi-resolution transformer fine-tuning for extreme multi-label text classification,” *NeurIPS*, 2021.
- [20] R. Hadsell, S. Chopra, and Y. LeCun, “Dimensionality reduction by learning an invariant mapping,” *IEEE CVPR*, 2006.

[21] A. van den Oord, Y. Li, and O. Vinyals, "Representation learning with contrastive predictive coding," 2018.

[22] T. Chen et al., "A simple framework for contrastive learning of visual representations," ICML, 2020.

[23] A. Radford et al., "Learning transferable visual models from natural language supervision," ICML, 2021.

[24] F. Javed et al., "Carotene: A job title classification system for the online recruitment domain," IEEE Big Data Computing Service and Applications, 2015.

[25] N. Laosaengpha et al., "Learning job title representation from job description aggregation network," ACL Findings, 2024.

[26] K. Song, X. Tan, T. Qin, and T. Y. Liu, "MPNet: Masked and permuted pre-training for language understanding," NeurIPS, 2020.

[27] A. Vaswani et al., "Attention is all you need," Advances in Neural Information Processing Systems, 2017.