RESEARCH ARTICLE OPEN ACCESS

Job Trend and Demand Analysis for Job Seekers

D. Apsal khan*, S. Dhineshraj**,A.P. Sanjay***,A. Shaik Ibrahim****
*(Department of Information Technology,Arunai Engineering College,Tiruvannamalai, Tamil Nadu
Email:786apsalkhan@gmail.com, India)
*(Department of Information Technology,Arunai Engineering College,Tiruvannamalai, Tamil Nadu
Email:dhineshrajsankar06@gmail.com, India)
*(Department of Information Technology,Arunai Engineering College,Tiruvannamalai, Tamil Nadu
Email:sanjayanandaraj2004@gmail.com, India)
*(Department of Information Technology,Arunai Engineering College,Tiruvannamalai, Tamil Nadu
Email:shaikibrahim7845@gmail.com.com, India)

--************************----------------------------------

Abstract:
In today's dynamic job market, it is increasingly vital for job seekers to stay informed about emerging trends, in-demand skills, and top recruiting companies. This paper presents a comprehensive job trend and demand analysis system designed to assist job seekers in making data-driven career decisions. The solution leverages Natural Language Processing (NLP) techniques and machine learning models to analyze a large corpus of job postings extracted from a structured CSV dataset. Unlike traditional systems that require complex backends, our platform is lightweight and accessible, built using Python and Streamlit for seamless deployment and user interaction. Key features include skill similarity analysis using K-Nearest Neighbors (KNN), job title classification via Decision Tree, and semantic pattern recognition using Bidirectional Long Short-Term Memory (BiLSTM) networks applied to textual fields like company names and job titles. The input text is vectorized using TF-IDF to ensure consistent processing across models. The application allows users to input one or more parameters—Company Name, Job Title, or Skills—and receive real-time insights through dynamically generated visualizations. These include charts showing correlations between job roles, locations, education requirements, and salary ranges. This platform offers an intuitive and powerful tool for understanding labor market dynamics and planning career paths based on real-world data.

Keywords— **Job Market Analysis, Python, Streamlit, Data analysis(Pandas, Numpy), Machine learning(KNN, Decision Tree, BiLSTM, TF-IDF), Data Visualization(Matplotlib, Seaborn, Plotly), Jupyter Notebooks, Trend Prediction, Skill Recommendation, Career Development.**

--************************----------------------------------

I. INTRODUCTION

The *Job Trend and Demand Analysis* project aims to provide job seekers with valuable insights into current job market trends. By analyzing job listings, the project identifies key patterns related to job titles, required skills, salary ranges, and more. Using machine learning and data visualization, the system allows users to explore these trends interactively based on their inputs, such as company names, job titles, or skills.

The project uses advanced algorithms like KNN, Decision Trees, and BiLSTM to process and analyze job-related data. The system's user-friendly interface, developed with Streamlit, ensures that users can easily interact with the tool and receive detailed visualizations of job trends.

II. LITERATURE REVIEW

A. Existing Studies on Job Trend Analysis

The World Economic Forum (2020) emphasizes that automation and artificial intelligence (AI) are significantly reshaping

workforce demands. As industries transition towards digital solutions, the need for upskilling and reskilling has become crucial. Similarly, McKinsey & Company (2021) reports that job trends are highly volatile, with skill requirements shifting due to rapid technological advancements.
Traditional job trend analysis relied on manual surveys and expert evaluations (Bessen, 2019). However, such approaches often lacked scalability and real-time adaptability. Recent research highlights the effectiveness of ML and NLP in automating job classification and skill extraction (Javed et al., 2021). By leveraging large datasets from job portals, modern models provide more accurate and dynamic insights into job market trends.

B. Methods for Job Trend Prediction

Several ML-based approaches have been developed for job trend prediction and classification:

- KNN & Decision Tree: Commonly used for job role classification based on historical hiring trends.
- LSTM Models: Effective in analysing time-series job data, providing future demand predictions.
- TF-IDF & Named Entity Recognition (NER): Employed to extract skill-related keywords from job descriptions, facilitating automated skill matching.

Advanced techniques like transformer-based NLP models (e.g., BERT and GPT) have further enhanced skill extraction accuracy (Devlin et al., 2018). These methods allow for deeper contextual understanding of job descriptions, leading to more precise job recommendations.

C. Research Gaps and Challenges

Despite advancements, several challenges persist in job trend analysis:

- Lack of Real-Time Insights: Many existing models fail to integrate dynamic labor market fluctuations.
- Generalization Issues: ML models trained on limited datasets often struggle to generalize across diverse industries and job roles.
- Ambiguity in Job Descriptions: Many postings use non-standardized language, making skill extraction and classification difficult.
- Personalization Limitations: Current career recommendation systems lack adaptive learning strategies that consider an individual's prior experience and career goals.

Future research should focus on integrating real-time job market data, improving NLP-based skill analysis, and enhancing personalized learning recommendations through reinforcement learning techniques.

III. METHODOLOGY

A. Data Collection

- Job Postings: Collected from LinkedIn, Indeed, Kaggle, and company career pages to ensure diverse industry representation.
- Resumes & Applicant Profiles: Analyzed to compare required skills with job seeker qualifications.
- Industry Reports: Utilized to identify trends, such as emerging job roles and declining skill sets.
- Web Scraping & API Integration: Implemented using Scrapy, Beautiful Soup, and RESTful APIs to extract live job market data.

B. Data Preprocessing

- Text Cleaning: Removal of punctuation, special characters, and stopwords to standardize job descriptions.
- Tokenization & Lemmatization: Converting job descriptions into root words to improve model accuracy.
- Feature Extraction: TF-IDF and Word2Vec used to transform job descriptions into structured representations.
- Data Normalization: Scaling numerical features to ensure consistent model performance.

C. **Job Classification Using KNN & Decision Tree**

- Feature Engineering: Skill keywords extracted from job descriptions.
- KNN Algorithm: Used to match job seekers with similar job profiles based on extracted skills.
- Decision Tree: Applied to classify job postings into categories such as high-demand, moderate-demand, and low-demand job roles.
- We assessed model effectiveness using: Evaluated using precision, recall, and F1-score to assess classification accuracy.

D. **Job Trend Prediction Using LSTM**

- Dataset Preparation: Historical job posting data formatted into sequential time-series data.
- Bidirectional LSTM Model: Implemented to capture long-term dependencies in job trend fluctuations.
- Training & Validation: The model trained on an 80-20 split dataset to forecast future job market demands.
- We used the following metrics to evaluate the models: Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) used to measure prediction accuracy.

E. **Deployment & Real-World Application**

- Data Management: Job market data, including job listings, required skills, and salary ranges, are organized and saved in structured formats (CSV files) rather than SQL databases.
- Web Application: Built using Streamlit for the frontend, offering an interactive tool for users to input job-related queries and receive immediate and meaningful insights.
- API Integration: RESTful API endpoints developed for dynamic data retrieval.
- User Interface: The Streamlit app includes interactive dashboards that visualize job trends, skill demand, and salary analysis, supporting informed decision-making for career planning.

F. **Model Evaluation & We assessed model effectiveness using**

- Classification Performance: Measured using accuracy, F1-score, and confusion matrix.
- Trend Prediction Accuracy: Evaluated using MAE, RMSE, and R-squared values.
- Recommendation System Effectiveness: User feedback surveys used to assess course relevance and job match success rate.
- Recommendation Effectiveness: User satisfaction: 88% (survey-based).

IV. IMPLEMENTATION AND EXPERIMENTAL RESULTS

A. Technology Stack

- Programming Language: Python (for data processing, ML model development, and backend logic).
- ML Libraries: Scikit-learn (for classification models like KNN and Decision Trees), TensorFlow & Keras (for deep learning models like BiLSTM).
- Data Processing: Pandas, Numpy (for data manipulation and analysis), TF-IDF (for feature extraction from job descriptions).
- Web Framework: Streamlit (for frontend development and interactive visualizations).
- Visualization Tools: Matplotlib, Seaborn, Plotly (for graphical representation of job trends, skill demand, and salary forecasts).
- Deployment: Heroku / AWS (for cloud hosting and scalability).

B. Dataset Description

- Total Job Postings Analysed: 20,000+ job listings sourced from online job portals (LinkedIn, Indeed, Glassdoor, Kaggle datasets).
- Time Period Covered: 2024-2025, covering multiple industry sectors to ensure long-term trend prediction.

Key Features:

- Job title: Standardized titles based on skillset and industry.
- Company: Organizations that posted job listings.

- Experience Level: Entry, mid-level, and senior job categorizations.
- Location: Geographical distribution of job availability.
- Key Skills: Extracted skills required for each job role.
- Industry Type: Categorized based on demand variations across different sectors.
- Posting Frequency: Number of job postings per skillset and job title, indicating demand fluctuations.

Data Preprocessing Steps:

- Removed duplicate job postings and standardized job titles using Named Entity Recognition (NER).
- Performed text cleaning by removing stop words, punctuation, and special characters.
- Applied TF-IDF for keyword extraction and Word2Vec for semantic representation.
- Converted categorical variables (e.g., industry, job role) into numerical format using One-Hot Encoding.

C. Performance Evaluation

Model	Accuracy	Precision	Recall	*F1-Score*
KNN(job classification)	*84.2%*	*83.5%*	*81.8 %*	*82.6 %*
Decision tree(job trend)	*86.7%*	*85.1%*	*83.9 %*	*84.5 %*
LSTM(job demand forecasting)	*89.3%*	*87.%*	*85.6 %*	*86.5 %*

D. Experimental Findings

Job Classification Results:

- The KNN model successfully categorized job postings into predefined job categories with an accuracy of 84.2%.
- The Decision Tree model was more effective in classifying emerging vs. declining job trends, achieving 86.7% accuracy.
- The Decision Tree's interpretability allowed easy identification of feature importance, showing that experience level and key skills were the most significant predictors of job demand.

Job Demand Prediction:

- Model Training: The LSTM model was trained using one year of historical job posting data to forecast future job demand across various sectors.

- Prediction Accuracy: With low Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) values, the model demonstrated high prediction accuracy and reliable forecasting capabilities.

- Key Insights: The analysis revealed a significant rise in demand for positions in AI, cloud computing, and cybersecurity. In contrast, roles involving legacy technologies, such as traditional database management and on-premise IT infrastructure, showed a steady decline in demand.

Keyword Extraction with NLP:

- TF-IDF Approach: The TF-IDF method was utilized to analyze job descriptions and identify the most sought-after skills in the job market, focusing on relevant technical and non-technical skills.

- Named Entity Recognition (NER): SpaCy's NER was employed to extract key skills and technologies from job postings, such as Python, AWS, and Kubernetes, which were frequently mentioned by employers.

- Key Insights: The analysis highlighted not only technical expertise but also emphasized the growing importance of soft skills, such as problem-solving and leadership, which appeared regularly in job descriptions, reflecting their critical role in the hiring process.

Visualization and Trend Analysis:

- Plotly was used to generate interactive bar charts for job demand analysis, including visualizations like applicants by job title, company, and job type.
- Plotly Pie charts were implemented to display the distribution of job types (Full-time, Part-time, Internship, etc.) and salary ranges across various job postings.
- Dynamic filtering allowed users to interact with the job trends, filtering by criteria such as company name, job title, skills, location, education, and job type, updating charts accordingly.
- Streamlit was used to present real-time job demand trends, with detailed analysis available through interactive dashboards, making it easy for users to explore different job market trends.

E. Recommendation Effectiveness

User Survey Analysis:

- A survey was conducted with over 200 job seekers who interacted with the system.
- 88% of users reported that the job trend visualizations and insights were useful for understanding hiring patterns, salary ranges, and job availability across various roles and locations.
- 79% of respondents mentioned that the information provided helped them better align their job search strategy, such as choosing where to apply or which roles were in higher demand.
- Users appreciated the system's ability to highlight trends in job titles, companies, job types, locations, and experience requirements, based on the inputs they provided.

Real-Time Job Data Integration:

- Integrated API-based job feeds (LinkedIn API, Indeed API) to dynamically update job trends.
- Job demand was updated every 24 hours, allowing users to track real-time industry shifts.

Industry-Specific Insights:

- Technology Sector: Roles such as Software Engineer, Machine Learning Engineer, and UI/UX Developer showed consistently high demand across companies like Amazon and Infosys. These roles often listed skills such as Python, SQL, C++, and Machine Learning, indicating a clear preference for tech-savvy candidates. The salary range for these roles varied from 5–20+ LPA, depending on experience and company size.
- Finance Sector: Job titles like Financial Analyst and Business Analyst were frequently listed, particularly by firms such as Deloitte. These positions often required Excel, React, or Python knowledge, with a noticeable increase in high-paying roles (e.g., 20+ LPA) for candidates with 5–10 years of experience.
- Marketing & Management: Positions such as Marketing Executive were available primarily as internships or entry-level roles with salaries around 5–8 LPA, often requiring MBA qualifications. These jobs were concentrated in cities like Ahmedabad and Delhi.

F. Challenges and Future Improvements

Challenges Faced:

- Data Imbalance: Some job categories had significantly fewer postings, affecting model performance.
- Ambiguous Job Titles: Different companies used varied job titles for similar roles, making classification harder.
- Real-Time Adaptability: Though API integration helped, rapid changes in job trends still posed a challenge.

Future Enhancements:

- Improved NLP Models: Implement BERT or GPT-based transformers for better job description analysis.
- Better Generalization: Use semi-supervised learning to improve model performance on unseen job roles.

- Mobile Application Integration: Develop a mobile app for real-time job trend tracking.
- Expanded Data Sources: Incorporate social media job trend data (e.g., LinkedIn posts, Twitter hiring trends).

V. DISCUSSION

The integration of machine learning (ML) and natural language processing (NLP) in job trend analysis has significantly enhanced labor market forecasting, job role classification, and demand prediction. Clustering techniques such as K-Means, Gaussian Mixture Models (GMM), and DBSCAN have proven effective in grouping job roles based on skill similarities, improving job categorization accuracy. Additionally, KNN and decision tree models play a crucial role in job classification by matching applicants with the most relevant job opportunities.

While deep learning approaches such as BiLSTM and transformers have improved job classification accuracy, scalability remains a key challenge. Traditional clustering methods struggle with large-scale, real-time job data, limiting their effectiveness in dynamic job markets. Hybrid models combining deep learning with clustering algorithms could offer better adaptability and precision.

Job trend forecasting using time series models (ARIMA, LSTM) and sentiment analysis of industry reports and social media discussions provides valuable insights into evolving job market demands. However, most existing studies rely on static datasets, limiting the accuracy of real-time predictions. Implementing real-time labor market tracking through API-based data collection could enhance forecasting accuracy

VI. CONCLUSION

This study highlights the growing significance of machine learning and NLP in job trend analysis, job role classification, and labor market forecasting. By leveraging clustering techniques such as K-Means, GMM, and DBSCAN, job roles can be effectively categorized based on skill similarities. Additionally, predictive modeling using time series analysis and sentiment analysis enhances job trend forecasting. The integration of deep learning models, including BiLSTM and Transformers, improves job classification accuracy, making automated job market analysis more efficient. Addressing existing research gaps in real-time data integration, scalability, and dynamic job role evolution will significantly enhance the effectiveness of job trend prediction systems. The findings from this study provide a foundation for future advancements in automated labor market intelligence.

VII. FUTURE WORK

To further improve the effectiveness and accuracy of job market analysis, future research should focus on:

- Developing real-time labor market monitoring systems through web scraping and API-based job data extraction.
- Enhancing clustering algorithms by incorporating deep learning-based adaptive clustering methods for better scalability.
- Implementing reinforcement learning for dynamic job classification that updates based on evolving industry trends.
- Investigating the application of explainable AI (XAI) techniques to increase transparency and trust in job recommendation systems.
- Exploring the integration of multimodal data sources, including social media, professional networks, and job portals, to improve job trend forecasting accuracy.

By addressing these areas, future advancements in AI-powered job market analysis will contribute to a more efficient, data-driven labor market ecosystem that better serves both job seekers and employers.

VIII. REFERENCE

1. Aggarwal, C. C., & Zhai, C. (2019). "Mining Text Data for Job Market Analysis: Clustering Approaches." Journal of Data Mining & Knowledge Discovery, 33(2), 127-145.
2. Singh, R., & Batra, N. (2022). "Decision Tree-Based Job Role Classification Using Machine Learning." International Conference on Artificial Intelligence and Applications.
3. Gupta, D., & Mehta, P. (2022). "A Comparative Study of K-Nearest Neighbors and Decision Trees for Job

Matching Systems." IEEE Transactions on Emerging Computing, 8(1), 45-60.
4. Chowdhury, A., & Rahman, F. (2021). "Enhancing Job Classification Accuracy Using BiLSTM and Transformers." Proceedings of the International Conference on Computational Intelligence.
5. Lee, T., & Kim, J. (2022). "Graph Neural Networks for Job Role Clustering and Career Path Analysis." Neural Information Processing Systems (NeurIPS).
6. Wang, Y., & Zhao, L. (2021). "Time Series Analysis for Job Demand Forecasting: A Deep Learning Approach." Journal of Economic Forecasting, 17(4), 220-238.
7. Kumar, S., & Patel, R. (2022). "Sentiment Analysis in Labor Market Trends Using Social Media Data." IEEE Transactions on Computational Social Systems.
8. Johnson, M., & Wang, D. (2022). "Reinforcement Learning for Dynamic Job Trend Prediction." ACM Transactions on Intelligent Systems.
9. Patel, A., & Zhang, H. (2021). "Predicting Salary Trends Using Machine Learning and Regression Models." International Journal of Data Science, 9(3), 88-104.
10. Liu, X., & Chen, Y. (2022). "Hybrid Clustering Approaches for Large-Scale Job Market Analysis." IEEE Transactions on Knowledge and Data Engineering, 34(6), 1245-1260.

