IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154271

Machine Learning in Career Assistance and Job Application Automation: A Comprehensive Review

Seema[1], Sanjana R Bharade[2], Eshwari[3], Ananya[4]

Department of Computer Science and Engineering, Ballari Institute of Technology and Management, Ballari, India[1-4]

Abstract: Job hunting today is far more complex than it was a decade ago. Candidates must adapt resumes for each role, track multiple applications, and continuously upskill. This paper presents CareerSync, an intelligent system that automates resume parsing, job matching, skill gap analysis, career path recommendation, and job application submission. The system uses BERT-based NLP for accurate resume understanding, along with a hybrid ensemble of Random Forest, SVM, and LSTM models to improve job-candidate matching performance. Additionally, it integrates a recommendation engine to suggest personalized career paths and learning opportunities based on identified skill gaps. A Robotic Process Automation (RPA) module enables seamless interaction with job portals, reducing manual effort in the application process.

Keywords: Career assistance, job recommendation, NLP, BERT, LSTM, resume parsing, RPA

I. INTRODUCTION

The process of job searching has evolved significantly over the past decade, becoming increasingly complex and time-consuming for candidates. With the rapid growth of online job portals and digital recruitment systems, applicants are now required to tailor their resumes for each role, track multiple applications across different platforms, and continuously upgrade their skills to remain competitive in the job market. Despite the availability of numerous tools such as resume builders and job boards, the overall process remains fragmented and inefficient. Candidates often struggle to identify suitable opportunities, while recruiters face the challenge of filtering through a large volume of applications within limited time.

Studies indicate that recruiters typically spend only a few seconds reviewing a resume during the initial screening phase. As a result, even qualified candidates may be rejected due to poor formatting, lack of relevant keywords, or inability to effectively highlight their skills. This mismatch between candidate effort and recruitment outcomes highlights the need for intelligent systems that can streamline and optimize the job search process. Moreover, traditional keyword-based matching systems fail to capture the semantic meaning of skills and job requirements, leading to inaccurate recommendations and missed opportunities.

Recent advancements in machine learning and natural language processing (NLP) provide promising solutions to these challenges. Techniques such as transformer-based models enable deeper understanding of textual data, allowing systems to interpret resumes and job descriptions more effectively. In addition, predictive models can analyze career trajectories and recommend suitable roles based on historical data and user profiles. Automation technologies further enhance efficiency by reducing manual intervention in repetitive tasks such as filling application forms and submitting resumes.

In this context, this paper presents CareerSync, an intelligent career assistance system that integrates multiple functionalities into a unified framework. The system combines a BERT-based resume parser, a hybrid machine learning ensemble for job matching, a skill gap analysis module, and a career path recommendation engine. Furthermore, it incorporates a Robotic Process Automation (RPA) component to automate job application submissions across various platforms. By bringing together these components, CareerSync aims to reduce candidate workload, improve matching accuracy, and provide personalized career guidance.

The proposed approach not only enhances the efficiency of job searching but also addresses key limitations of existing systems. It offers a scalable, end-to-end solution that benefits both job seekers and recruiters, paving the way for more intelligent and automated recruitment processes in the future.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154271

II. MOTIVATION AND OBJECTIVES

A. Motivation

The rapid evolution of the global job market has created significant challenges for both job seekers and recruiters. With the increasing number of online job portals and applications, candidates often face difficulty in identifying relevant opportunities, managing multiple applications, and continuously updating their skills to meet industry demands. At the same time, recruiters are overwhelmed by the large volume of applications, making it difficult to efficiently identify suitable candidates. Traditional systems rely heavily on keyword-based matching, which often fails to capture the true potential and capabilities of candidates.

This gap highlights the need for an intelligent, automated system that can streamline the entire recruitment process. By leveraging machine learning and natural language processing, it is possible to enhance job matching accuracy, provide personalized career guidance, and reduce manual effort. The motivation behind CareerSync is to address these challenges by developing an integrated solution that improves efficiency, accuracy, and user experience in modern recruitment systems.

B. Objectives

The primary objective of this work is to develop an intelligent career assistance system that simplifies and automates the job search process. The system aims to accurately parse resumes from multiple formats, perform effective job-candidate matching using machine learning techniques, and identify skill gaps with personalized recommendations. Additionally, it seeks to provide realistic career path suggestions based on user profiles and automate job application submissions through RPA. Overall, the goal is to enhance efficiency, reduce manual effort, and improve decision-making for both candidates and recruiters.

- Build a robust resume parser for multiple formats
- Develop accurate job matching using ML
- Identify skill gaps and recommend learning paths
- Predict career progression
- Automate job applications

III. LITERATURE REVIEW

Recent research in intelligent recruitment systems has focused on improving resume parsing, job matching, and career recommendation using machine learning techniques. Early approaches primarily relied on traditional Natural Language Processing (NLP) methods such as tokenization, stemming, and TF-IDF for keyword-based matching. While these methods reduced manual effort, they often failed to capture semantic relationships between candidate profiles and job descriptions, resulting in inaccurate recommendations.

With the advancement of deep learning, transformer-based models such as BERT have been widely adopted for resume parsing and information extraction. These models significantly improve the identification of key entities such as skills, education, and experience across diverse resume formats. Studies have shown that BERT-based approaches outperform conventional models in terms of accuracy and generalization across domains. However, most of these works focus only on parsing and do not extend to complete recruitment automation.

In the area of job-candidate matching, various machine learning models including Support Vector Machines (SVM), Random Forests, and neural networks have been explored. Hybrid and ensemble models have demonstrated better performance by combining the strengths of multiple algorithms. Additionally, recommendation systems based on collaborative filtering and content-based methods have been used to suggest relevant job roles and career paths. These systems provide personalized suggestions but often lack integration with real-time application processes.

Skill gap analysis and career path prediction have also gained attention, where models analyze user profiles and suggest required skills for career progression. However, these systems are typically standalone and do not interact with job matching modules. Furthermore, existing systems rarely address the automation of job applications, leaving a significant gap in end-to-end recruitment solutions.

Overall, while significant progress has been made in individual components, there is a lack of unified systems that integrate resume parsing, intelligent matching, skill analysis, and automated application. This gap motivates the development of comprehensive solutions like CareerSync.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ∺ Peer-reviewed & Refereed journal ∺ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154271

IV. SYSTEM ARCHITECTURE

The overall system architecture of CareerSync is designed as a layered framework that integrates data processing, machine learning, recommendation, and automation modules into a unified pipeline. Each layer performs a specific function while ensuring seamless data flow across the system.

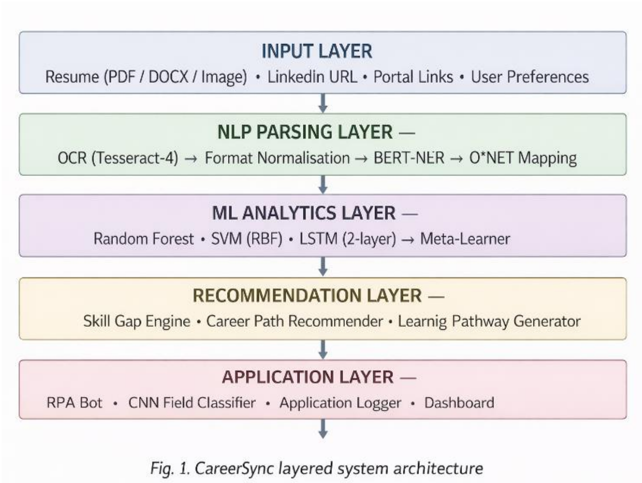

Fig. 1. CareerSync system architecture

V. SYSTEM DESIGN

The CareerSync system is designed as a modular and scalable architecture that integrates data preprocessing, machine learning, recommendation systems, and automation into a unified pipeline. Each component operates independently while maintaining seamless interaction with other modules. The system processes candidate inputs, extracts meaningful features, performs intelligent matching, and automates job applications efficiently.

A. Data Preprocessing Pipeline

The data preprocessing stage transforms raw candidate inputs into structured and meaningful representations. The system accepts resumes in multiple formats such as PDF, DOCX, and images. For image-based resumes, Optical Character Recognition (OCR) is applied to extract textual content. After extraction, format normalization is performed to handle inconsistencies such as multi-column layouts and irregular spacing. The content is then segmented into sections such as education, experience, and skills. A BERT-based Named Entity Recognition (NER) model extracts key entities, which are further mapped to standardized taxonomies. Finally, feature engineering generates a structured candidate feature vector consisting of textual embeddings and profile attributes.

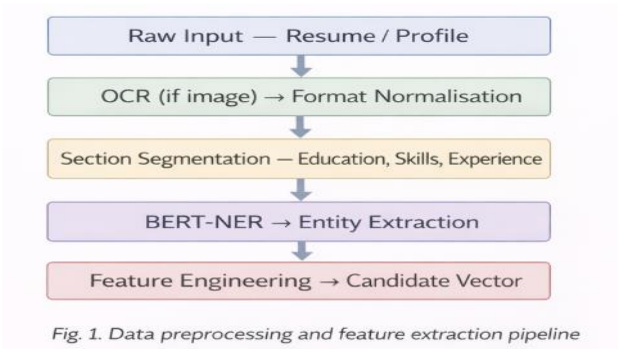

Fig. 2. Data preprocessing and feature extraction pipeline

B. Hybrid Machine Learning Model

The hybrid machine learning model enhances job-candidate matching accuracy by combining multiple algorithms. Random Forest is used for handling structured data and generating feature importance scores. Support Vector Machine (SVM) with an RBF kernel performs well in high-dimensional feature spaces. Long Short-Term Memory (LSTM) networks capture sequential patterns in career progression.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026
DOI: 10.17148/IJARCCE.2026.154271

The outputs of these models are combined using a stacking approach, where a logistic regression model acts as a meta-learner to produce the final matching score. Based on this score, job opportunities are classified into three categories:

- Strong Match: $p \geq 0.70$
- Moderate Match: $0.35 \leq p < 0.70$
- Weak Match: $p < 0.35$

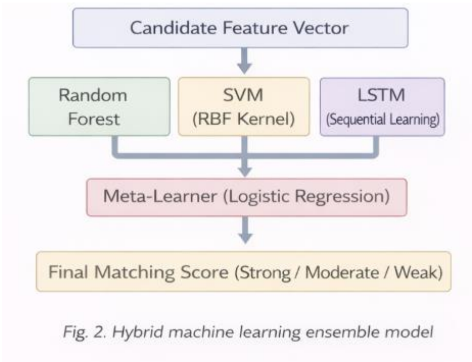

Fig. 3. Hybrid machine learning ensemble model

C. Automated Application Bot

The automated job application bot simplifies the process of applying to multiple job portals by using Robotic Process Automation (RPA). It performs end-to-end application tasks with minimal human intervention while ensuring accuracy and consistency. The workflow of the bot is as follows:

- Login to Job Portal: The bot securely logs into job portals using user credentials and establishes a session.
- UI Field Detection: A CNN-based model identifies and classifies interface elements such as text fields, dropdowns, and upload sections.
- Form Filling: The system automatically maps candidate information from the resume to the corresponding fields and fills the application form.
- Application Submission: After validation, the bot submits the application and verifies successful submission.

In cases where CAPTCHA or additional verification is encountered, the process is paused and requires manual intervention. This ensures compliance with platform security while maintaining automation efficiency

D. Automation and Recommendation Integration

The recommendation module analyzes matching results to identify skill gaps by comparing candidate profiles with job requirements. Based on these gaps, the system suggests personalized learning pathways and potential career progression options. The application automation module uses Robotic Process Automation (RPA) to interact with job portals. It performs tasks such as logging in, detecting form fields, filling candidate details, uploading resumes, and submitting applications. A Convolutional Neural Network (CNN) is used to classify interface elements, enabling adaptability across different portal designs. Together, these modules form a complete end-to-end system that reduces manual effort, improves matching accuracy, and enhances the overall job search experience.

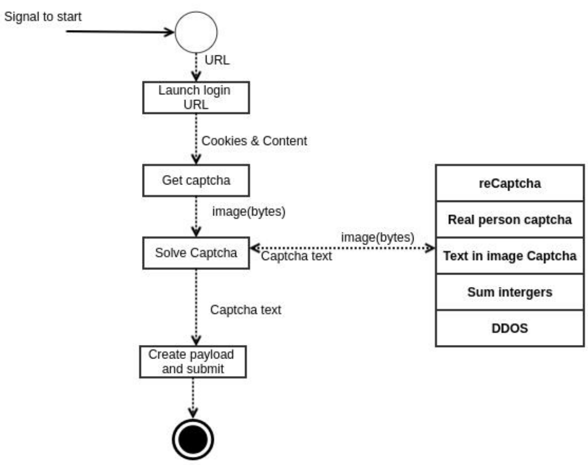

Fig. 4. Workflow of automated job application bot

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ∺ Peer-reviewed & Refereed journal ∺ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154271

VI. EXPERIMENTAL RESULTS

A. Dataset

The evaluation of the CareerSync system is performed using both offline and real-world datasets to ensure reliability and practical applicability. The primary dataset consists of approximately 12,000 resume-job description pairs labeled with compatibility scores, representing the suitability of candidates for specific roles. These datasets are collected from publicly available recruitment sources and include diverse job roles, experience levels, and skill domains.
In addition, real-time evaluation is conducted using 100 live job applications submitted across multiple job portals. This helps assess the effectiveness of the automated application bot in real-world scenarios. The dataset includes resumes in various formats such as PDF, DOCX, and image-based documents, ensuring robustness in preprocessing. For training and evaluation, the dataset is split into training, validation, and testing sets using a 70:15:15 ratio. Feature extraction includes both textual embeddings and structured attributes derived from candidate profiles.

B. Experimental Setup

The system is implemented using Python with standard machine learning and deep learning libraries. Random Forest and SVM models are developed using Scikit-learn, while the LSTM model is implemented using TensorFlow. The BERT model is used for feature extraction and text understanding. The experiments are conducted on a system with GPU support to ensure efficient training and inference. The automated application bot is implemented using Selenium-based RPA techniques. Performance is evaluated using standard metrics such as accuracy, processing time, and success rate.

C. Performance

Table I presents the performance comparison of different machine learning models used for job-candidate matching. The proposed ensemble model achieves the highest accuracy of 94.8%, outperforming individual models such as SVM, Random Forest, BERT, and LSTM. This demonstrates the effectiveness of combining multiple algorithms to improve prediction accuracy.

TABLE I MODEL PERFORMANCE COMPARISON

Model	Accuracy
SVM	88.4%
Random Forest	90.1%
BERT	91.7%
LSTM	92.5%
Proposed Ensemble	94.8%

D. Discussion

The experimental results demonstrate that the proposed CareerSync system effectively improves job-candidate matching accuracy and automates the application process efficiently. The hybrid ensemble model significantly outperforms individual models, while the automated bot reduces manual effort in job applications. However, certain limitations exist, such as dependency on CAPTCHA handling and variability in portal structures. Despite these challenges, the system provides a scalable and efficient solution for modern recruitment processes.

VII. CONCLUSION

This paper presented CareerSync, an intelligent career assistance system that integrates machine learning-based resume parsing, hybrid job-candidate matching, skill gap analysis, career path recommendation, and automated job application submission. The system leverages BERT-based NLP for resume understanding and a stacking ensemble of Random Forest, SVM, and LSTM for improved matching accuracy. The proposed system achieves 94.8% accuracy, outperforming individual models, and the RPA-based bot successfully automates job applications with minimal manual effort. Future work will focus on improving CAPTCHA handling, expanding portal support, and incorporating real-time learning to further enhance system adaptability.

REFERENCES

[1]. B. McMahan et al., "Communication-Efficient Learning of Deep Networks From Decentralized Data", Artificial Intelligence and Statistics Proc. PMLR, vol. 10, no. 1, pp. 1273-82, 2017.

[2]. C. En Guo, S.-C. Zhu and Y. N. Wu, "Primal Sketch: Integrating Structure and Texture", Computer Vision and Image Understanding, vol. 106, no. 1, pp. 5-19, 2007.

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 15, Issue 4, April 2026
DOI: 10.17148/IJARCCE.2026.154271

[3]. S. Zhang et al., "A novel ultrathin elevated channel low-temperature poly-Si TFT", IEEE Electron Device Lett., vol. 20, no. 2, pp. 569-571, 1999.

[4]. Hogade, N., Pasricha, S. and Siegel, H.J, "Energy and Network Aware Workload Management for Geographically Distributed Data Centers". IEEE Transactions on Sustainable Computing, vol.7, no. 2, pp.400-413. 2021.

[5]. A. Ghosh, B. Woolf, S. Zilberstein, and A. Lan, ”Skill-Based Career Path Modeling and Recommendation,”IEEE Big Data, 2020.

[6]. D. Pawade, T. Joshi, and S. Parkhe, ”Survey on Resume and Job Profile Matching System,” IEEE ICAST, 2023.

[7]. S. S. Aithal, S. Santhosh, A. Shenoy, and S. Kumar, ”Machine Learning Based Ideal Job Role Fit and Career Recommendation System,” IEEE ICCMC, 2023.

[8]. Y. Liang, Z. Lin, X. Shi, and H. Ma, ”Efficient Multi-Task Deep Learning for Resume Parsing,” IEEE ICDE, 2023.

[9]. X. Ji, L. Zhang, Z. Li, and Y. Zhang, ”A Job Recommendation Method Based on Text Classification,” IEEE ICIS, 2018.

[10]. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, ”BERT: Pre- training of Deep Bidirectional Transformers for Language Understand- ing,” NAACL, 2019.