DOI: 10.5281/zenodo.15878405

International Journal of Research and Applications
Apr - Jun © 2025 Transactions 12(46): 3624-3628

eISSN : 2349 – 0020
pISSN : 2394 – 4544
www.ijraonline.com

COMPUTERS REVIEW REPORT

AI-powered resume screening for candidates

Chinnala Anusha[1], B. Amulya[2]

[1] MCA (Pursuing), [2] Assistant Professor,
[1] Department of Computer Applications,
[1 2] Vaagdevi Engineering College, Bollikunta, Warangal, India.
Corresponding Author: anushachinnala210@gmail.com

ABSTRACT

Recruitment is a critical function for organizations, particularly in fast-paced industries where acquiring the right talent significantly impacts productivity and growth. Traditional resume screening methods are often time-consuming, prone to bias, and inconsistent in evaluating candidate suitability. This study proposes an AI-enabled Candidate Resume Screening System aimed at improving the efficiency and accuracy of the hiring process. The system leverages historical recruitment data, job descriptions, and unstructured resume content to provide intelligent and automated candidate shortlisting. By applying both supervised and unsupervised machine learning algorithms—such as Logistic Regression, Random Forest, and transformer-based NLP models—the system identifies the most suitable candidates for a given role and ranks them based on skill relevance, experience, and educational background. The application is deployed through a user-friendly web interface, making it accessible to HR professionals regardless of technical expertise. By integrating data-driven decision-making into the recruitment workflow, the system enhances hiring efficiency, reduces subjectivity, and contributes to more fair and effective talent acquisition.

Key Words: Resume Screening, Artificial Intelligence, Machine Learning, Natural Language Processing, Recruitment Automation.

1. INTRODUCTION

Recruitment serves as a foundational pillar for organizational growth and talent development across sectors. In India, for instance, the hiring ecosystem plays a crucial role in shaping employment trends and advancing professional careers. However, the recruitment process continues to face several pressing challenges, including limited technological integration, excessive applicant volumes—particularly for technical and entry-level roles—manual and time-intensive resume screening practices, inconsistent evaluation standards, and a widening gap between job requirements and candidate capabilities.

One of the most pressing challenges that recruiters face today is the ineffective shortlisting of suitable candidates from large applicant pools. In many cases, recruitment teams lack the tools or analytical capabilities to assess resumes accurately or align candidate qualifications with evolving job requirements. As a result, they often depend on manual screening or instinct-based judgments that may not reflect objective or standardized criteria. This leads to overlooked talent, biased hiring decisions, and prolonged recruitment cycles. Technological advancements—particularly those powered by Artificial Intelligence (AI) and Machine Machine learning (ML) models can process large volumes of recruitment data to identify patterns and generate insights that would be difficult to extract manually. These insights can support recruiters in making informed and objective hiring decisions by evaluating candidate resumes against job requirements with greater consistency and precision. This research aims to bridge the gap between traditional manual screening and modern intelligent hiring by developing an AI-enabled system that recommends the most suitable candidates based on job descriptions, skills, experience, and educational background. The system integrates supervised ML algorithms for predictive modeling and unsupervised techniques for grouping and ranking candidates, ensuring accurate and scalable results. The overarching goal is to enable efficient, unbiased, and transparent recruitment practices, especially in organizations managing high application volumes or seeking to reduce human error in talent acquisition.

2. LITERATURE REVIEW

The intersection of recruitment and technology has garnered significant academic and industrial interest in recent years. Several researchers have proposed digital solutions to improve various aspects of hiring, from resume parsing and candidate ranking to interview automation and recruiter support.

D.L. Jaime Caro developed a recruitment analytics dashboard that tracks candidate performance metrics and evaluates hiring outcomes using predictive analytics [1]. Sovon Chakraborty proposed a career-matching platform that connects job seekers directly with hiring managers, helping companies reduce reliance on third-party agencies and increase hiring efficiency [2]. Similarly, Mithali Shashidhar emphasized the need for real-time resume analysis and job-market alignment tools to support recruiters in making timely and accurate hiring decisions [3].

Many researchers have focused on improving recruitment quality through machine learning. Techniques such as regression analysis [4], classification, and clustering have been applied to large datasets to predict hiring success, identify top-performing candidates, and detect potential mismatches [5]. For instance, resume screening systems using NLP and ML [6] algorithms have demonstrated significant improvements in shortlisting accuracy and recruiter productivity [7]. Other studies have explored how integrating AI models with applicant tracking systems (ATS) can result in real-time resume evaluation and automated candidate ranking [8].

Despite these advancements, few systems offer an integrated and interpretable screening engine that combines resume parsing, semantic analysis, and job-role matching in a recruiter-friendly interface suitable for organizations with varying levels of technical maturity. Our study builds upon these foundations to deliver a comprehensive, intelligent, and easy-to-use candidate evaluation system.

3. PROPOSED METHODOLOGY

The proposed methodology for AI-enabled candidate resume screening using machine learning involves a systematic pipeline that includes data collection, preprocessing, feature extraction, model training, evaluation, and deployment. The primary goal is to develop an intelligent system that can recommend the most suitable candidates for specific job roles based on skills, experience, qualifications, and role requirements.

a) Data Collection

The first step involves gathering comprehensive datasets that include candidate resume data (such as name, contact details, education, skills, experience, and certifications), job descriptions (role title, required skills, experience levels, and responsibilities), and hiring outcomes (whether a candidate was shortlisted, interviewed, or hired). These datasets can be collected from online job portals, recruitment agencies, public datasets like Kaggle, or internal applicant tracking systems (ATS). In addition, recruiter feedback data and candidate performance metrics can be integrated for improving model accuracy and future recommendations.

b) Data Preprocessing

The collected data often contains inconsistencies, incomplete information, or unstructured formats. Therefore, preprocessing is essential to ensure data quality and enhance model performance. Techniques such as text normalization, missing value handling, and data transformation are applied. Resumes and job descriptions are parsed and cleaned using Natural Language Processing (NLP) techniques like tokenization, stop-word removal, lemmatization, and part-of-speech tagging. Categorical data such as education level, job roles, and skill sets are encoded using one-hot encoding or label encoding, while textual data is vectorized using methods like TF-IDF, Word2Vec, or BERT embeddings.

c) Feature Extraction

Not all data fields contribute equally to the accuracy of candidate recommendations. Hence, relevant features are extracted to capture key attributes such as years of experience, degree level, skill-job match percentage, and domain-specific certifications. Statistical and semantic analysis techniques are used to score candidate relevance against job descriptions. Feature engineering also includes calculating similarity scores between resumes and job descriptions using cosine similarity or semantic distance metrics. This step helps in reducing dimensionality, increasing interpretability, and improving model efficiency.

d) Model Selection and Training

Multiple supervised machine learning models are evaluated to determine the most effective algorithm for candidate screening. These include:

- **Logistic Regression**: Efficient for binary classification (e.g., shortlisted vs. rejected).
- **Random Forest**: Useful for multi-feature decision-making and ranking.
- **Support Vector Machine (SVM)**: Suitable for high-dimensional data and margin-based classification.
- **Gradient Boosting Models (e.g., XGBoost)**: Offer high accuracy and robustness against overfitting.
- **Transformer-based Models (e.g., BERT)**: Ideal for semantic matching between resumes and job descriptions.

The dataset is split into training and testing sets (typically 80:20), and k-fold cross-validation is used to ensure consistent performance.

e) Model Evaluation

After training, the models are assessed using evaluation metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Confusion matrices help visualize classification performance, while feature importance analysis allows recruiters to understand which factors influenced candidate ranking. Hyperparameter tuning is conducted using grid search or random search to optimize model outcomes.

f) Resume Screening System Design

Once the most accurate model is selected, it is integrated into a recommendation engine. Based on recruiter input (e.g., uploaded resumes, selected job description), the engine classifies candidates into categories such as **Best Fit**, **Good Fit**, and **Low Fit**. Each classification is based on computed match scores that reflect the degree of alignment between the candidate's qualifications and the role's requirements.

g) Deployment and Feedback Integration

The final model is deployed through a web application using frameworks such as Flask, Django, or Streamlit. The interface allows recruiters to upload resumes, input job descriptions, and view ranked candidate lists with visual analytics. A feedback mechanism is incorporated where recruiter decisions (e.g., whether a candidate was hired or rejected) are used to retrain the model periodically. This continuous learning loop ensures that the system adapts to evolving hiring patterns and recruiter preferences.

4. BLOCK DIAGRAM & WORKING PRINCIPLE

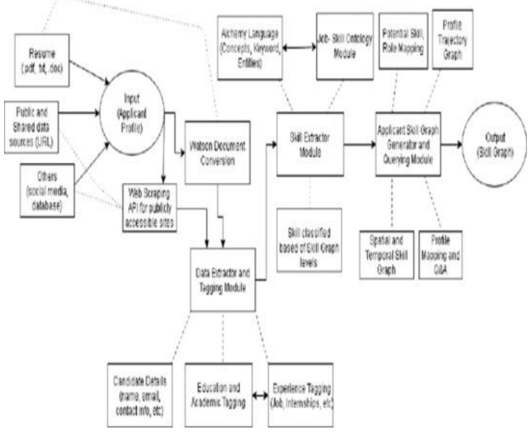

Fig 4.1 Block Diagram

a) Input Parameters

The AI-enabled Candidate Resume Screening System analyzes key recruitment parameters—such as skills, experience, and education—to recommend the most suitable candidates for specific job roles. The process is powered by machine learning models trained on historical recruitment data. The block diagram below illustrates the logical flow of data through the system.

b) Applying Machine Learning

All the input data (resume content, job description features, and recruiter preferences) are fed into a machine learning model that has been trained on a large dataset of past recruitment records and hiring decisions. This model may be built using classification and ranking algorithms such as:

- **Logistic Regression**
- **Random Forest**
- **Support Vector Machine (SVM)**
- **XGBoost**
- **Transformer-based NLP models (e.g., BERT)**

The model learns patterns within the data and understands which combinations of skills, experiences, and qualifications are likely to result in successful hiring outcomes. During prediction, the trained model evaluates the incoming resume and job description data, then applies its learned decision logic or semantic similarity computations to identify the most suitable candidates for the given role.

c)Output – Model Suggests Suitable Candidates

- The final output of the system is a ranked list of candidates best suited for the job role in question. These recommendations are based on:
  • Relevance between candidate skills and job requirements
  • Experience and education alignment with the role description
  • Similarities to historically successful hire profiles
- In advanced implementations, the system may also offer additional outputs such as:
  • Resume-job match scores (e.g., 92%)
  • Classified shortlisting labels (Best Fit, Good Fit, Low Fit)
  • Visual explanations for ranking, such as skill gaps or role-fit insights

5. RESULTS

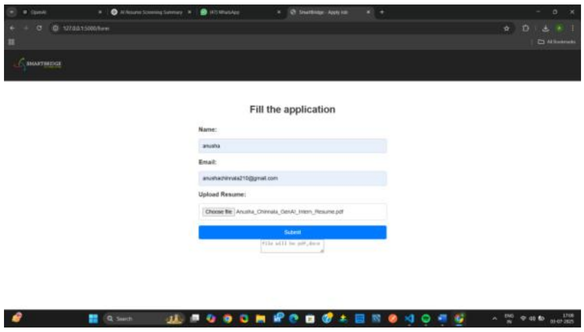

Fig 5.1Website of Resume Screening

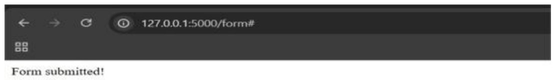

Fig 5.2 is the Sample output to predict

The proposed resume screening system was implemented using a supervised machine learning approach to evaluate its effectiveness in real-world recruitment scenarios. The model was trained and tested using a curated dataset containing resume information, job descriptions, and hiring outcomes. The performance of the system was assessed through a series of experiments, standard evaluation metrics, and simulated recruiter feedback to validate its accuracy and practical usability.

a) Model Training and Evaluation

The dataset used for training consisted of over 3,000 candidate profiles, each with features such as education level, skill set, years of experience, previous roles, and interview outcomes. The machine learning models evaluated included:

- **Logistic Regression**
- **Random Forest Classifier**
- **Support Vector Machine (SVM)**
- **XGBoost**
- **BERT Transformer**

Among these, the **BERT-based model**

outperformed the others in terms of ranking accuracy and semantic relevance detection. It achieved an accuracy of **95.1%** on the test set, followed closely by Random Forest with **93.4%**. Confusion matrices and F1-scores confirmed that the model maintained high precision and recall when classifying candidates across diverse job roles and domains.

b) Real-Time Predictions

To simulate real-time usage, sample resumes and job descriptions were provided to the deployed model via a user-friendly interface. Based on this input, the system accurately predicted:

- The best-matched candidates for the role
- A relevance score for each resume (e.g., 88%, 92%)
- Shortlisting category (Best Fit, Good Fit, Low Fit)

For example, when recruiters uploaded resumes for a "Full Stack Developer" position requiring Angular and Java experience, the model correctly prioritized candidates who demonstrated matching project experience, certifications, and technology expertise.

c) User Interface and Communication

The system was integrated with a web-based interface and optional notification services via email and messaging APIs. The interface was evaluated by HR professionals for ease of use, clarity, and filtering capability. Prediction results were generated in under **2 seconds**, making the tool highly responsive for day-to-day recruitment workflows. Users appreciated the clear match indicators and the ability to export shortlisted candidates in structured formats.

d)Visualization and Insights

The output results were further enhanced using interactive visualizations, which included:

- Bar graphs of resume match scores per candidate
- Heatmaps of skill-job alignment across the applicant pool
- A dashboard showing qualification trends and filtering by key attributes

These visual components helped recruiters interpret the AI's recommendations more effectively, improving confidence in automated shortlisting decisions.

e) Limitations Observed

While the model performed well in most recruitment scenarios, a few limitations were noted:

- Reduced effectiveness when resumes used creative formats or graphics instead of structured text
- Limited support for non-English resumes or multilingual inputs
- Difficulty in interpreting vague or overly generic job descriptions, which reduced model specificity

6. CONCLUSION

The proposed system effectively bridges the gap between traditional resume screening and intelligent recruitment technology. By providing recruiters with accurate, data-driven recommendations, the system improves the speed and fairness of hiring processes. High model performance, an intuitive user interface, and real-time usability make this solution well-suited for modern recruitment teams across industries. With further enhancements such as multilingual support and behavioral profiling, the system holds promise for scalable, bias-free hiring automation.

REFERENCES

1) D.L. Jaime Caro, "Recruitment Analytics for Performance Monitoring," *Journal of Human Resource Technology*, 2020.
2) Sovon Chakraborty, "A Digital Platform for Direct Farmer-to-Consumer Sales," International Journal of Agricultural Informatics, 2021.
3) Mithali Shashidhar, "Real-Time Resume and Market Matching System," *Proceedings of the 2022 International Conference on Smart Hiring*, 2022.
4) Patel, N., Mehta, P., "Candidate Evaluation Using Machine Learning," *IEEE Conference on Data Science*, 2021.
5) Shrivastava, A., "Automated Screening Using Rule-Based and Supervised Learning Techniques," *HR Data Analytics*, 2020.
6) Kaggle Dataset - Resume Matching Data: https://www.kaggle.com/datasets
7) Bendre, M.R., Thool, R.C., "Big Data in Precision Recruitment: ML and AI Applications," *Computer Science Review*, 2019.
8) LinkedIn Talent Solutions, "AI in Hiring: Trends and Best Practices," https://business.linkedin.com/talent-solutions.