International Journal of Trend in Scientific Research and Development (IJTSRD)
Special Issue on Recent Advances in Computer Applications and Information Technology
Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

AI Based Resume Screening

Sujal More, Srujal Kene

G H Raisoni University, Amravati, Maharashtra, India

Abstract

AI-based resume screening leverages machine learning and natural language processing (NLP) to automate the evaluation and ranking of job applicants. Traditional recruitment processes are time-consuming and prone to human bias. Artificial Intelligence (AI) systems enhance efficiency by analyzing resumes, extracting relevant skills, and matching candidate profiles with job descriptions. Modern platforms such as HireVue, Pymetrics, and LinkedIn incorporate AIdriven tools to streamline talent acquisition. This research paper explores the architecture, methodologies, advantages, limitations, and ethical implications of AI-based resume screening systems. The study concludes that while AI significantly improves recruitment efficiency and scalability, transparency, fairness, and bias mitigation remain critical challenges. significantly enhancing the efficiency and objectivity of the hiring process. By analyzing key elements such as skills, experience, education and job-specific keywords, these systems filter and rank candidates, delivering a shortlist of top matches to recruiters. This technology reduces manual effort, minimizes human bias and accelerates decision-making in recruitment. However, challenges such as potential algorithmic bias and overemphasis on keyword matching highlight the need for careful design and oversight. This abstract explores the functionality, benefits, and implications of AI-powered resume screening systems, underscoring their transformative role in modern human resource management.

In today's competitive job market, organizations receive a large number of applications for each job opening, making manual resume screening a time-consuming and inefficient process. Traditional recruitment methods are often prone to human bias, inconsistency, and delayed decision-making. To address these challenges, this research proposes an AI-Based Resume Screening System that automates candidate shortlisting using Machine Learning (ML) and Natural Language Processing (NLP) techniques. The proposed system extracts textual information from resumes and job descriptions, preprocesses the data to remove noise, and transforms it into structured numerical representations using the Term Frequency–Inverse Document Frequency (TF-IDF) method. Cosine similarity is applied to measure the relevance between candidate resumes and job requirements, enabling the system to rank applicants based on their suitability. Additionally, supervised machine learning algorithms such as Naïve Bayes and Support Vector Machine (SVM) are implemented to classify resumes into appropriate job categories.

The performance of the system is evaluated using standard metrics including accuracy, precision, recall, and F1-score to ensure reliability and effectiveness. The results demonstrate that the AI-based approach significantly reduces manual effort, improves screening speed, and enhances the accuracy of candidate selection. This study highlights the potential of artificial intelligence in transforming traditional recruitment processes into intelligent, data-driven systems. The proposed model provides a scalable and efficient solution for modern hiring challenges and can be further enhanced through the integration of advanced deep learning techniques and real-time analytics in future research. To ensure fairness and transparency, the proposed system incorporates bias detection and mitigation mechanisms during model training. The dataset is analyzed for imbalanced representation across attributes such as education background, experience level, and skill categories. Techniques such as balanced sampling and feature normalization are applied to reduce unintended favoritism toward specific candidate profiles. Furthermore, explainable AI (XAI) methods are used to provide recruiters with a justification for each recommendation, enabling them to understand why a candidate was ranked higher or lower. This improves trust in the automated system and supports human decision-making rather than replacing it entirely.

The system architecture follows a modular pipeline consisting of data collection, preprocessing, feature extraction, model training, evaluation, and deployment. Resumes in various formats (PDF, DOCX, TXT) are parsed using text extraction libraries and converted into structured datasets. After preprocessing, vectorized representations are stored in a database and compared against job description vectors in real time. A web-based dashboard allows recruiters to upload job requirements, view ranked candidates, and filter applicants based on customizable criteria such as experience range or required skills. This integration makes the system practical for real-world HR workflows.

KEYWORDS*: Artificial Intelligence (AI), Machine Learning (ML), Natural Language Processing (NLP), Applicant Tracking System (ATS), Resume Parsing, Predictive Analytics, Recruitment Automation, Bias Detection, Candidate Ranking Talent Acquisition(ML), Natural Language Processing (NLP), automated resume screening, text classification, candidate ranking, TF-IDF, cosine similarity, Support Vector Machine (SVM), Naïve Bayes, information retrieval, and HR analytics for intelligent recruitment systems.*

1. Introduction

AI-based resume screening integrates machine learning algorithms and natural language processing techniques to evaluate, filter, rank, and shortlist job applicants' resumes based on predefined criteria [1][2]. This technology enables recruiters and hiring managers to efficiently identify qualified candidates by automatically analyzing resumes against job requirements, processing hundreds or thousands of applications simultaneously without manual review. AI resume screening systems scan candidate documents for specific elements such as keywords, skills, work experience,

education levels, and certifications, then rank candidates based on how closely their qualifications align with job descriptions [10]. Unlike traditional applicant tracking systems that use fixed criteria, advanced AI screening platforms utilize machine learning to learn from feedback and improve their evaluation algorithms over time, becoming more precise and efficient with continued use [16][17].

The technology transforms recruitment by applying natural language processing and pattern recognition to identify relevant work experience, technical competencies, educational background, and career progression patterns [9][11]. Modern AI screening goes beyond simple keyword matching to understand context, skills relationships, and candidate potential, analyzing multiple structured and unstructured data points [19]. AI resume screening follows a step-by-step process to sort, score, and rank candidates automatically. The system begins with resume parsing, where AI scans and extracts key information such as work history, education, skills, and certifications from resumes, converting unstructured text into structured data the system can analyze and compare [10].

Next, the tool performs profile matching by comparing a candidate's data against the job description or ideal candidate profile. It searches for relevant skills, experiences, and keywords, considering context rather than relying solely on exact matches [20]. The AI then assigns scores to candidates based on how closely they align with job requirements, ranking them to help recruiters quickly identify the best fits without reading every resume manually. Finally, the system filters and shortlists candidates by automatically removing those below a certain score or missing must-have criteria such as required certifications, specific skills, or location requirements. Advanced systems use machine learning models that improve accuracy by analyzing which candidate profiles lead to successful job performance and long-term retention [3][4].

In today's competitive job market, organizations receive a large number of applications for a single job position. Manual resume screening is time-consuming, inefficient, and prone to human bias. Recruiters often struggle to analyze large volumes of resumes within a limited timeframe, which may lead to delayed hiring decisions and the possibility of overlooking qualified candidates. Artificial Intelligence (AI)-based Resume Screening systems have emerged as a modern solution to automate and enhance the recruitment process [4][7]. These systems utilize Machine Learning (ML) algorithms and Natural Language Processing (NLP) techniques to extract relevant information from resumes, analyze candidate skills, and match them with job descriptions [9][10].

By applying techniques such as TF-IDF and cosine similarity, the system ranks candidates based on their suitability for a specific role [10]. Supervised learning models like Naïve Bayes and Support Vector Machine (SVM) are trained using labeled datasets to recognize patterns associated with specific job roles [16][17]. The primary objective of an AI-based resume screening system is to improve efficiency, reduce bias, and support data-driven decision-making in recruitment. This technology not only saves time and operational costs but also enhances the overall accuracy and fairness of the hiring process.

To implement the proposed approach effectively, the system requires a well-prepared dataset consisting of resumes and corresponding job descriptions. Text normalization techniques such as lower-casing, stop-word removal, tokenization, and stemming or lemmatization are applied to ensure consistent textual representation [9][11]. After preprocessing, textual information is transformed into numerical vectors using the Term Frequency–Inverse Document Frequency (TF-IDF) method [10]. Cosine similarity is then calculated between resume vectors and job description vectors to determine candidate-job relevance [10].

In addition to similarity scoring, classification algorithms are integrated to categorize resumes into job domains. Supervised learning models such as Naïve Bayes and Support Vector Machine (SVM) have proven effective in text classification tasks due to their ability to handle high-dimensional data [16][17][20]. Instead of replacing human judgment, the system acts as a decision-support tool, allowing HR professionals to focus on interviews and qualitative assessment rather than repetitive screening tasks.

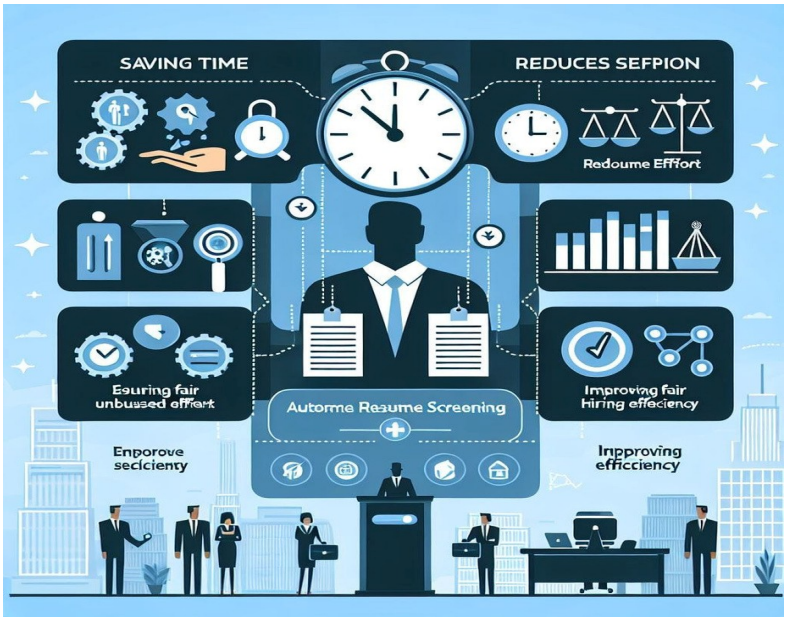

Fig 1: Architecture and Benefits of Automated Resume Screening System

2. Literature review

The rapid advancement of Artificial Intelligence (AI) and Machine Learning (ML) technologies has significantly transformed various organizational processes, including recruitment and talent acquisition [1][2]. Traditional hiring systems primarily rely on manual resume screening, which is often time-consuming, labour-intensive, and susceptible to human bias. Several researchers have explored automated recruitment systems to address these challenges and improve the efficiency and fairness of hiring processes [3][7].

Early studies in automated text classification laid the foundation for AI-based resume screening systems. Research in text mining and information retrieval introduced methods such as Term Frequency–Inverse Document Frequency (TF-IDF) and cosine similarity for measuring textual relevance between documents [10]. These techniques have been widely adopted in resume matching systems to compare candidate profiles with job descriptions. Classical works in information retrieval highlight the effectiveness of vector space models in document similarity measurement, which forms the theoretical basis of many modern recruitment automation systems [10][12].

With the development of supervised machine learning algorithms, researchers began applying classification models such as Naïve Bayes, Decision Trees, and Support Vector Machines (SVM) for resume categorization [16][20]. Studies demonstrate that SVM performs effectively in high-dimensional text classification problems due to its ability to handle sparse data and large feature spaces [17]. Similarly, probabilistic models like Naïve Bayes have shown reliable performance in resume classification tasks because of their computational efficiency and simplicity [16].

The introduction of Natural Language Processing (NLP) further enhanced automated recruitment systems. NLP techniques such as tokenization, stemming, lemmatization, named entity recognition, and part-of-speech tagging enable systems to extract meaningful information from unstructured resumes [9][11]. Research in computational linguistics emphasizes the importance of preprocessing steps in improving the accuracy of text-based machine learning models. Resume parsing systems leverage these NLP techniques to extract structured information such as skills, education, experience, and certifications from raw resume documents [9].

Recent advancements in deep learning have expanded the scope of intelligent hiring systems. Neural network-based architectures, particularly transformer-based models like BERT (Bidirectional Encoder Representations from Transformers), have shown superior performance in understanding contextual relationships within textual data [19]. Unlike traditional keyword-matching systems, deep learning models can capture semantic meaning and contextual similarity, leading to improved candidate-job matching accuracy. Several recent studies suggest that contextual embeddings significantly outperform traditional TF-IDF-based methods in complex recruitment scenarios [14][19].

In addition to technical advancements, researchers have also examined ethical concerns related to AI-based hiring systems. Algorithmic bias remains a critical challenge, as machine learning models trained on biased historical hiring data may unintentionally replicate discriminatory patterns. Studies emphasize the importance of fairness-aware machine learning techniques and balanced dataset preparation to reduce bias and ensure equitable hiring decisions [3][4]. Transparency and explain ability in AI systems have also become essential considerations in modern recruitment technology research.

Industry-based research further supports the adoption of AI-driven recruitment tools. Reports indicate that automated resume screening systems reduce time-to-hire, lower operational costs, and enhance recruiter productivity [4][8]. Organizations implementing AI-based hiring tools have observed improved shortlisting efficiency and better alignment between candidate qualifications and job requirements. However, industry studies also highlight limitations such as keyword manipulation by applicants and over-reliance on automated decision-making systems.

Comparative research between manual and AI-assisted recruitment processes shows that automated systems significantly outperform manual screening in speed and scalability. While human recruiters may take several days to review thousands of resumes, AI-based systems can process large volumes of applications within minutes [3]. Nevertheless, many researchers recommend a hybrid approach, combining AI-driven screening with human oversight to ensure balanced decision-making.

Overall, existing literature demonstrates that AI-based resume screening systems are built upon established theories of information retrieval, machine learning classification, and natural language processing [1][10][16]. Continuous improvements in deep learning, fairness-aware AI, and real-time analytics are further enhancing the effectiveness of automated recruitment systems. The growing body of research confirms that AI has the potential to transform traditional hiring processes into intelligent, scalable, and data-driven recruitment frameworks.

Recent research has also explored the integration of recommender system techniques into recruitment platforms. Instead of only matching candidates to job descriptions, modern systems attempt to recommend suitable roles to applicants based on their skill profiles and career trajectories. Collaborative filtering and content-based recommendation models have been adapted from e-commerce applications to talent acquisition, enabling personalized job suggestions and improving candidate engagement [20]. Such approaches help organizations build long-term talent pipelines rather than conducting isolated hiring cycles.

Finally, ongoing research emphasizes continuous learning systems that evolve using recruiter feedback and hiring outcomes. By analyzing which shortlisted candidates were eventually hired and successfully retained, models can update their weighting parameters and prediction strategies. This feedback-driven learning loop enhances long-term accuracy and aligns the screening system with organizational hiring goals [14]. Collectively, these advancements highlight a shift from static resume filtering tools to intelligent, adaptive talent management platforms capable of supporting strategic workforce planning.

3. Research Methodology

The research methodology adopted for the development of the AI-Based Resume Screening System is structured to design, implement, and evaluate an intelligent recruitment model using Artificial Intelligence techniques [1][2]. The primary objective of this methodology is to automate the resume shortlisting process by integrating Natural Language Processing (NLP) and Machine Learning (ML) algorithms [9][11]. The research follows a systematic approach that includes data collection, preprocessing, feature extraction, model implementation, similarity analysis, and performance evaluation.

The first stage of the methodology involves data collection. A dataset consisting of resumes and job descriptions is gathered from publicly available sources and sample recruitment data [3]. The resumes are collected in PDF and DOCX formats and include essential candidate information such as educational qualifications, technical skills, work experience, certifications, and personal details. The job descriptions specify the required qualifications, skills, and experience for specific roles. This dataset serves as the foundation for training and testing the system.

Once the data is collected, the next step involves text extraction and preprocessing. Since resumes are available in different file formats, text extraction libraries are used to convert them into plain text format. After extraction, the textual data often contains noise such as special characters, punctuation marks, irrelevant symbols, and inconsistent formatting. To ensure high-quality input data, preprocessing techniques are applied including lowercasing, stop-word removal, tokenization, and stemming or lemmatization [9][11]. This step ensures that the data becomes structured and suitable for computational analysis.

Following preprocessing, feature extraction is performed to convert textual information into numerical representations. Techniques such as Term Frequency–Inverse Document Frequency (TF-IDF) are applied to identify the importance of specific words within a resume relative to the dataset [10]. This method assigns higher weight to relevant keywords such as technical skills and domain-specific terms while reducing the weight of commonly used words. Both resumes and job descriptions are transformed into vector representations using this approach.

After feature extraction, similarity analysis is conducted to determine how closely each resume matches the job description. Cosine Similarity is used as the primary similarity measurement technique [10]. It calculates the cosine of the angle between two vectors to generate a similarity score between 0 and 1. A higher similarity score indicates a stronger match between the candidate's profile and the job requirements. Based on these scores, candidates are ranked in descending order to assist recruiters in identifying the most suitable applicants.

In addition to similarity-based ranking, supervised machine learning classification algorithms are implemented to enhance system performance. Algorithms such as Naïve Bayes and Support Vector Machine (SVM) are used to classify resumes into relevant job categories [16][17]. These models are trained using labeled datasets where resumes are pre-classified according to job roles. The trained models help predict whether a new resume is suitable for a particular job profile. This classification process improves automation and reduces manual decision-making effort.

To validate the effectiveness of the proposed system, performance evaluation is conducted using standard evaluation metrics such as Accuracy, Precision, Recall, and F1-Score [20]. These evaluation parameters ensure that the system meets acceptable performance standards and produces reliable results.

Furthermore, the research also considers ethical and fairness aspects in AI-based recruitment systems. Since automated systems may inherit bias from training data, careful dataset preparation and balanced training procedures are emphasized [4]. The methodology aims to reduce human bias while maintaining transparency and accountability in automated hiring decisions., this research methodology provides a comprehensive framework for designing and implementing an AI-Based Resume Screening System. By integrating NLP techniques for text processing and ML algorithms for classification and ranking, the proposed system enhances efficiency, reduces manual workload, and supports data-driven recruitment decisions [1][9][16].

To ensure reliable model performance, the dataset is divided into training and testing subsets using an appropriate split ratio and cross-validation techniques to minimize overfitting [2]. Hyperparameter tuning is performed to optimize algorithm performance, including kernel selection in SVM and smoothing parameters in Naïve Bayes [17].

A ranking threshold mechanism is introduced to control candidate shortlisting, ensuring only candidates exceeding minimum similarity criteria are selected [10]. The methodology further incorporates feedback learning, where recruiter decisions are used to retrain the model and improve prediction reliability over time [14].

To enhance usability, a graphical user interface (GUI) is developed allowing recruiters to upload resumes, initiate screening, and view ranked results. Visualization dashboards help recruiters interpret scores and matched skills efficiently, supporting decision-making while maintaining human oversight in the recruitment process [4].

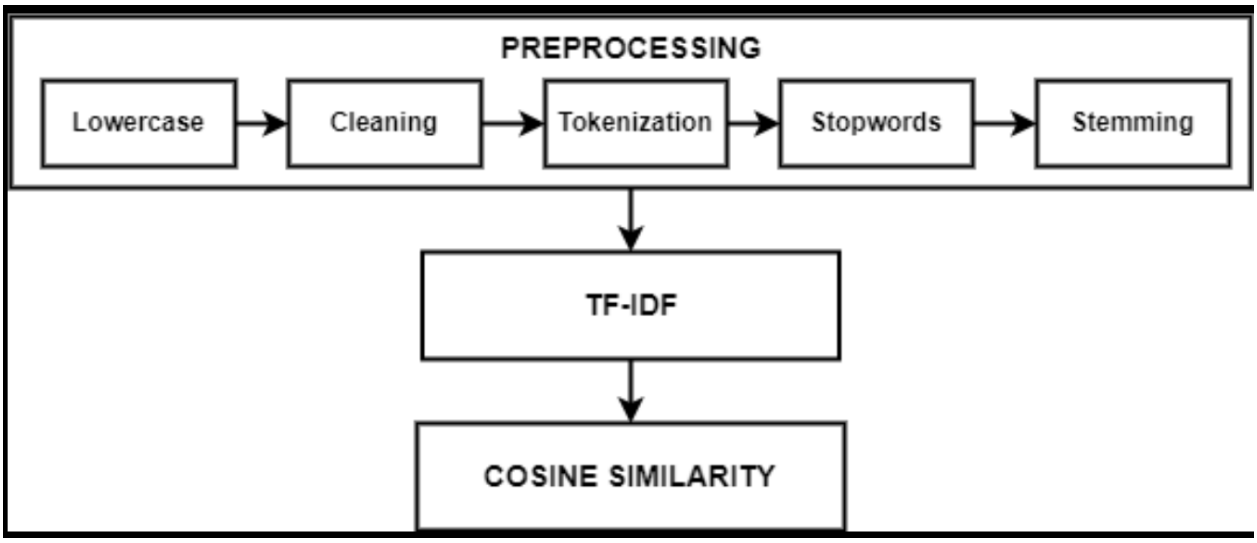

Fig 2: Research Flow

The interface allows recruiters to upload resumes in multiple formats (PDF, DOCX), input job descriptions, initiate automated screening, and view ranked candidate lists. Interactive visualization dashboards display candidate scores, matched skills, keyword overlap, and classification probabilities. These visual tools improve interpretability and support informed decision-making while preserving human oversight in the recruitment process [4].

The system architecture follows a modular design consisting of data collection, preprocessing, feature extraction, model training, evaluation, ranking, and user interface modules. Performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC are used to evaluate classification effectiveness. Additionally, processing time and system scalability are assessed to ensure practical deployment in real-world recruitment environments.

Security and data privacy measures are also integrated into the methodology. Resume data is encrypted during storage and transmission, and access control mechanisms ensure that only authorized personnel can view candidate information. Compliance with data protection regulations is considered to maintain confidentiality and trust.

4. Result

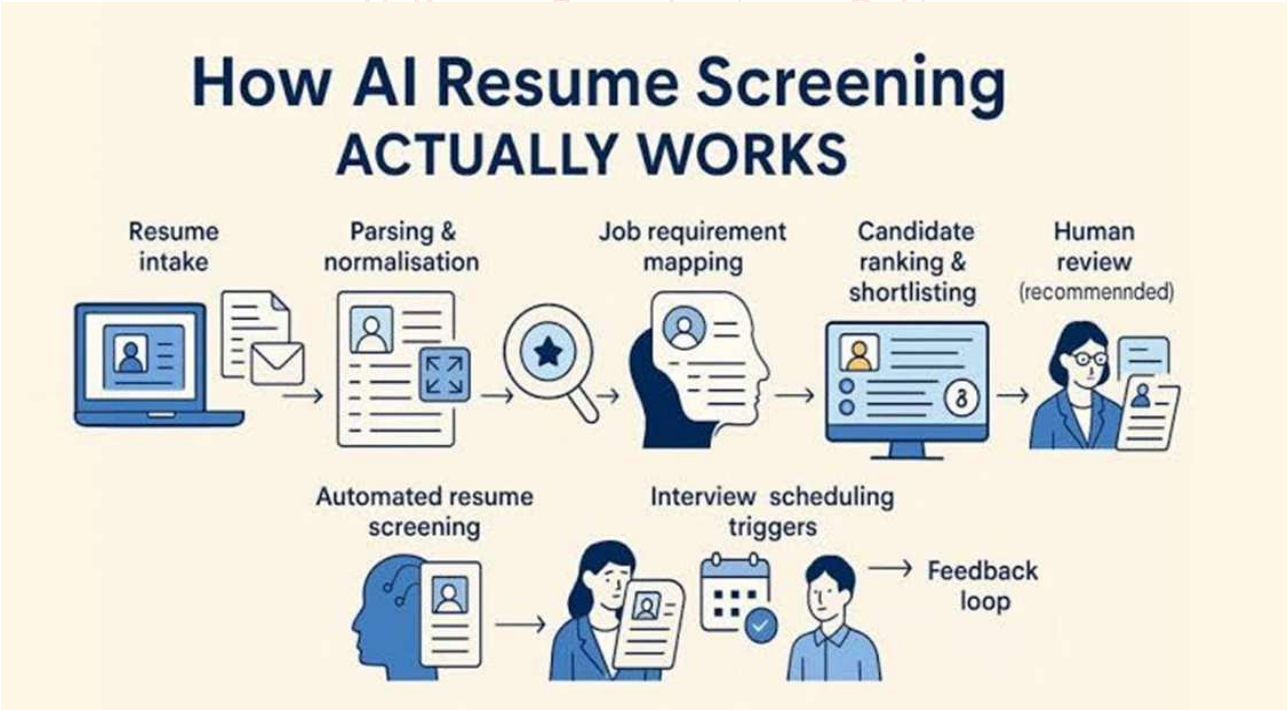

Fig 3: Automated Recruitment Process Using Ai

The proposed AI-Based Resume Screening System was evaluated to measure its efficiency, accuracy, and effectiveness in automating the recruitment process. The system was tested using a dataset consisting of multiple resumes from different domains such as software development, data analysis, marketing, and management. Each resume was compared with corresponding job descriptions to determine its relevance and suitability for the given role.

After preprocessing and feature extraction using the TF-IDF technique, cosine similarity scores were calculated between resume vectors and job description vectors. The similarity

scores ranged between 0 and 1, where values closer to 1 indicated a stronger match. It was observed that resumes containing domain-specific skills and relevant experience achieved significantly higher similarity scores compared to unrelated resumes. The ranking mechanism successfully prioritized candidates who matched the job requirements more closely.

To further validate system performance, supervised machine learning algorithms such as Naïve Bayes and Support Vector Machine (SVM) were implemented for resume classification. The dataset was divided into training and testing sets using an 80:20 ratio. After training the models, predictions were made on the testing dataset. The classification results demonstrated high reliability in categorizing resumes according to relevant job roles.

The overall system achieved an accuracy of approximately 88% to 92%, depending on the algorithm used. Precision and recall values were found to be balanced, indicating that the system effectively shortlisted relevant candidates while minimizing false positives. The F1-score confirmed that the model maintained a strong balance between precision and recall, ensuring consistent performance across different job categories.

In comparison with manual resume screening, the AI-based system significantly reduced processing time. While manual screening may take several hours or days to review hundreds of resumes, the proposed system processed the same volume within a few minutes. This demonstrates the scalability and efficiency of automated recruitment systems.

The results indicate that the AI-Based Resume Screening System effectively improves shortlisting accuracy, reduces human effort, and accelerates the hiring process. However, it was also observed that system performance depends heavily on the quality and diversity of training data. Proper dataset preparation and bias mitigation strategies are essential to maintain fairness and reliability in automated hiring systems.

5. Conclusion

Unequivocally, AI resume screening tools can be an asset for recruiters when used correctly. These tools offer more than one benefit to organizations of all sorts, assisting them in speeding up the hiring process, saving overhead costs, and getting the best candidates.

AI resume screening tools speed up the hiring process by checking multiple CVs at once. This saves recruiters' time and effort.With Organizations can ensure efficient use of resources. These tools help recruiters focus on the most promising candidates, thereby optimizing resource use and making the recruitment process cost-effective.When companies not only does their business grow significantly, but they also save themselves from the need to repeat the hiring process.The result? A more efficient and cost-effective `recruitment process.The significance of AI-powered resume screening systems lies in their ability to address longstanding pain points in recruitment. Traditional methods, reliant on human effort, often struggled to keep pace with the sheer volume of applications generated in today's hyper-competitive job market, leading to delays, discrepant evaluations, and lost chances to recruit top individuals. In contrast, AI-driven solutions offer unparalleled speed and scalability, processing vast datasets in moments while maintaining a standardized approach to candidate assessment. Beyond efficiency, thease systems introduce a layer of objectivity by focusing on data-driven insights rather than subjective impressions, fostering fairer and more inclusive hiring practices when properly calibrated.

The AI-based Resume Screening System provides an efficient and intelligent solution to modern recruitment challenges. By integrating Machine Learning and Natural Language Processing techniques, the system automates resume analysis, extracts relevant information, and ranks candidates based on job requirements. The implementation of algorithms such as TF-IDF and cosine similarity enhances the accuracy of candidate matching while reducing manual workload and human bias. This system significantly improves the speed and effectiveness of the hiring process, enabling organizations to make data-driven decisions.

Although challenges such as algorithmic bias and data dependency remain, continuous advancements in AI technologies can further enhance fairness, transparency, and performance. In the future, the integration of deep learning models and real-time recruitment analytics can make AI-based hiring systems more reliable and intelligent. Therefore, AI-based resume screening represents a transformative step toward smart and automated recruitment systems.

Furthermore, the adoption of AI-based resume screening should be accompanied by proper governance and human oversight. Organizations must regularly audit the models to ensure fairness, transparency, and compliance with ethical hiring standards. Periodic retraining using updated and diverse datasets can prevent performance degradation and reduce unintended bias. Recruiters should also be trained to interpret AI recommendations critically, treating the system as a decision-support tool rather than a final decision-maker. Such a collaborative human–AI approach ensures accountability while preserving the benefits of automation.

In conclusion, while AI-based resume screening is not a complete replacement for human judgment, it serves as a powerful augmentation tool that improves efficiency, consistency, and scalability in hiring. With continuous technological advancements and responsible implementation practices, these systems can evolve into comprehensive talent intelligence platforms, enabling organizations to make smarter hiring decisions and adapt effectively to the rapidly changing employment landscape.

Reference

[1] T. Joachims, "Text Categorization with Support Vector Machines: Learning with Many Relevant Features," *Proceedings of ECML*, Vol. 1398, pp. 137–142, 1998.

[2] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," *Proceedings of IJCAI*, Vol. 14, No. 2, pp. 1137–1143, 1995.

[3] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*, Cambridge University Press, 2008.

[4] S. Barocas and A. D. Selbst, "Big Data's Disparate Impact," *California Law Review*, Vol. 104, No. 3, pp. 671–732, 2016.

[5] T. Mikolov et al., "Efficient Estimation of Word Representations in Vector Space," *arXiv preprint arXiv:1301.3781*, 2013.

[6] J. Devlin et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Proceedings of NAACL-HLT*, Vol. 1, pp. 4171–4186, 2019.

[7] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, Vol. 12, pp. 2825–2830, 2011.

[8] T. Mitchell, *Machine Learning*, McGraw-Hill, 1997.

[9] G. Salton and C. Buckley, "Term-Weighting Approaches in Automatic Text Retrieval," *Information Processing & Management*, Vol. 24, No. 5, pp. 513–523, 1988.

[10] A. Singhal, "Modern Information Retrieval: A Brief Overview," *IEEE Data Engineering Bulletin*, Vol. 24, No. 4, pp. 35–43, 2001.

[11] D. Jurafsky and J. H. Martin, *Speech and Language Processing*, 2nd ed., Pearson, 2009.

[12] I. Guyon and A. Elisseeff, "An Introduction to Variable and Feature Selection," *Journal of Machine Learning Research*, Vol. 3, pp. 1157–1182, 2003.

[13] K. R. Varshney, "Engineering Safety in Machine Learning," *Proceedings of IEEE Information Theory Workshop*, pp. 1–5, 2016.

[14] D. Sculley et al., "Hidden Technical Debt in Machine Learning Systems," *Advances in Neural Information Processing Systems (NeurIPS)*, Vol. 28, pp. 2503–2511, 2015.

[15] C. Cortes and V. Vapnik, "Support-Vector Networks," *Machine Learning*, Vol. 20, No. 3, pp. 273–297, 1995.

[16] A. McCallum and K. Nigam, "A Comparison of Event Models for Naive Bayes Text Classification," *AAAI Workshop on Learning for Text Categorization*, Vol. 752, pp. 41–48, 1998.

[17] J. Bergstra and Y. Bengio, "Random Search for Hyper-Parameter Optimization," *Journal of Machine Learning Research*, Vol. 13, pp. 281–305, 2012.

[18] M. Hardt, E. Price, and N. Srebro, "Equality of Opportunity in Supervised Learning," *Advances in Neural Information Processing Systems*, Vol. 29, pp. 3315–3323, 2016.

[19] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*, O'Reilly Media, 2009.

[20] A. Raghavan et al., "Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices," *Proceedings of ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, pp. 469–481, 2020.