An Integrated Framework for Automated Resume Screening Using RoBERTa, Random Forest and Explainable AI

Kevin Frederick Yapiter[1*], Alfin[2], Yoga Hasim[3], Ronsen Purba[4], Mustika Ulina[5]

[1,2,3,4,5]Department of Informatics Engineering, Faculty of Informatics, Universitas Mikroskil, Medan, North Sumatera, Indonesia
E-mail: [1*]21111669@students.mikroskil.ac.id, [2]211110493@students.mikroskil.ac.id, [3]211110781@students.mikroskil.ac.id, [4]ronsen@mikroskil.ac.id, [5]mustika.ulina@mikroskil.ac.id

(Received: 18 Aug 2025, revised: 14 Sep 2025, accepted: 16 Sep 2025)

Abstract

The resume screening process is a critical stage in recruitment, yet conventional methods and traditional applicant tracking systems (ATS) often rely on manual review or keyword matching, resulting in slow, biased, and less objective evaluations. This study proposes an integrated automated screening system that combines RoBERTa for contextual feature extraction, Random Forest for candidate classification, and SHAP-based Explainable AI for interpretable decisions, enhancing transparency, efficiency, and fairness beyond traditional ATS. The dataset consists of real resumes and synthetically generated ones designed to mimic the distribution of real data, with K-means clustering used to establish labeling thresholds. Experimental results show that RoBERTa achieved an F1 Score of 81.08% in feature extraction, while Random Forest reached 96% accuracy in suitability classification. SHAP-based explanations provide insights into feature contributions for each prediction, offering an actionable understanding for recruiters. This integrated framework not only improves the efficiency and fairness of resume screening but also demonstrates a practical application of explainable AI in recruitment.

Keywords: Resume Screening, RoBERTa, Random Forest, Explainable AI, Recruitment

I. INTRODUCTION

The recruitment process plays a critical role in acquiring qualified personnel aligned with organizational needs [1]. Among the key stages in this process is resume screening, which serves to identify the most suitable candidates for specific job positions [2]. Resumes typically contain personal information, educational background, achievements, competencies, skills, and relevant experiences [3]. However, the large volume of submissions in diverse formats often presents significant challenges for recruiters, rendering the screening process time-consuming, prone to human error, and susceptible to bias [4], [5]. To mitigate these challenges, organizations increasingly adopt Applicant Tracking Systems (ATS) to automate and organize candidate selection [6]. While offering certain efficiencies, conventional ATS primarily relies on keyword matching, which often fails to capture the broader context of candidates' experiences and qualifications. Furthermore, such approaches may introduce systemic bias, reducing the accuracy and fairness of candidate evaluations [7]. The situation is exacerbated by the growing number of job seekers [8]. In Indonesia, for instance, the number of job seekers reached 1.82 million in 2023, reflecting a 94.18% increase from the previous year's total of 937,176 [9].

Recent advancements in Machine Learning (ML) and Natural Language Processing (NLP) offer promising solutions for automating resume screening with higher accuracy and efficiency [10]. NLP allows computers to process and interpret human language naturally [11]. One powerful NLP model is RoBERTa (Robustly Optimized BERT Pretraining Approach), an enhanced version of BERT that leverages larger corpora and dynamic masking during pretraining to better capture contextual meaning [12], [13]. Previous research demonstrates that RoBERTa can effectively extract relevant information from resumes, achieving an F1 Score of up to 85.24% [14]. However, most existing works stop at information extraction, without providing a mechanism to evaluate overall candidate suitability or classify candidates automatically. Consequently, recruiters remain responsible for the final decision-making, leaving the process partly manual and inefficient.

To overcome this limitation, this research integrates RoBERTa with the Random Forest algorithm. Random Forest, as part of the ensemble learning paradigm, has demonstrated strong performance in handling high-dimensional text data

while reducing overfitting [15], [16]. Comparative studies highlight its effectiveness in resume classification tasks, where it achieves 91.38% accuracy higher than K-Nearest Neighbors (81.63%), Decision Tree (85.53%), and Support Vector Classifier (90.62%) [17]. By combining RoBERTa's contextual representation with Random Forest's classification strength, this study aims to produce more accurate and robust candidate categorization.

Despite its advantages, Random Forest and similar ML models are often regarded as "black boxes" due to the opacity of their decision-making processes [18]. To enhance transparency and interpretability, Explainable Artificial Intelligence (XAI) has emerged as a framework that enables models to generate human-understandable explanations [19]. XAI plays a critical role in improving trust and accountability in AI-based decision-making [20]. Prior research suggests that XAI can reduce bias in candidate evaluation, particularly regarding gender and age, and enhance recruiter confidence in automated systems [21]. This integration ensures that the automated screening process is not only accurate but also transparent and interpretable for HR practitioners.

In summary, this study proposes an automated resume screening approach that combines RoBERTa for contextual information extraction, Random Forest for candidate classification, and SHAP-based Explainable AI for interpretable predictions. The system addresses limitations of conventional ATS by capturing nuanced information from resumes, providing robust candidate evaluations, and offering insights to support recruiter decisions. Specifically, the research aims to improve both the accuracy of automated classification and the transparency of the screening process, enabling recruiters to understand the reasoning behind each recommendation. By integrating pretrained NLP embeddings with ensemble learning and explainable AI techniques, this approach builds upon existing methods in the literature, contributing to more objective, fair, and interpretable automated resume screening.

II. RESEARCH METHODS

A. Research Flow

This study adopts a multi-stage methodology combining NLP, Machine Learning, and XAI. The research flow, illustrated in Figure 1:

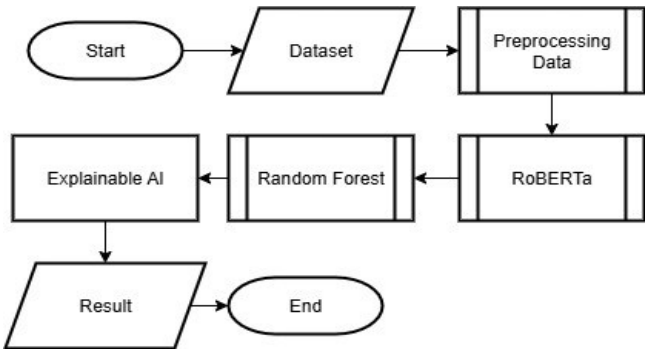


Figure 1. Research Flow

1. Dataset

In this study, a total of 1,095 data records were used, originating from two sources: a public dataset and a synthetic dataset. The public dataset consisted of 595 records obtained from GitHub in JSON file format. It contained collections of resume texts along with labels such as name, skills, work experience, education, and other related information. Additionally, a synthetic dataset comprising 500 records was generated using a Python program. The synthetic dataset was designed to have the same structure as the dataset obtained from GitHub. To construct the synthetic records, candidate attributes such as job titles, technical skills, years of experience, and education were programmatically varied using controlled randomization. This ensured that the generated data preserved realistic resume characteristics while preventing duplication of the original records. The synthetic dataset augmented training data to improve model learning.

2. Preprocessing

At this stage, a series of preprocessing steps were carried out to clean and prepare the resume data for model training (Figure 2). The preprocessing aimed to reduce noise, standardize text formats, and ensure that the data conformed to the required structure for effective modeling, including:

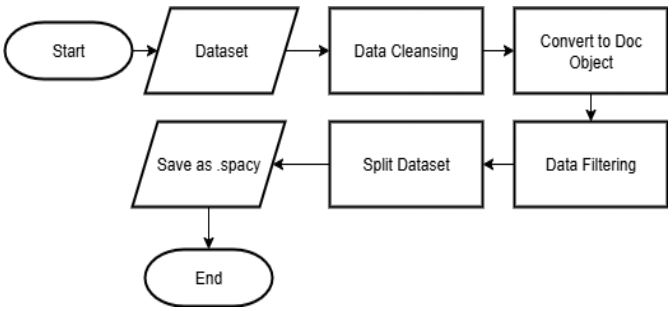


Figure 2. Preprocessing Flow

a. Data Cleansing

Non-ASCII characters and excessive spaces were removed from the text using Python's Regular Expression module, and duplicate entries were eliminated.

Table 1. Cleansing Data

Before	After
Kavya U.\nNetwork Ops	Kavya U. Network Ops

In Table 1 above, it can be seen that extra space has been removed.

b. Convert to Doc Object

Cleaned texts and label lists, excluding irrelevant name and email attributes, were converted into document objects via a two-step process:

- Tokenization

Splitting text and punctuation into tokens and storing them in doc objects.

DOI: 10.34148/teknika.v14i3.1359
TEKNIKA, Volume 14(3), November 2025, pp. 424-432
ISSN 2549-8037, EISSN 2549-8045

Table 2. Tokenization

Before	After
Kavya U.\nNetwork Ops	'Kavya', 'U.', 'Network', 'Ops'

In Table 2 above, it can be observed that the raw text has been systematically tokenized into individual, meaningful units.

- Entity Spans

Identifying important word ranges (entity spans) and storing them in doc.ents.

Table 3. Create Entity Spans

Before	After
Kavya U.\nNetwork Ops	Text: Network Ops Associate, Label: Designation, Start: 2, End: 5

In Table 3 above, it can be observed that entity spans have been accurately identified and extracted from the text for annotation and model training purposes.

c. Data Filtering

In this step, records containing more than 4,000 tokens were removed to maintain data quality and ensure efficient model training.

Table 4. Data Filtering

Before	After
1073 data	1065 data

Table 4 presents the number of records before and after the filtering process.

d. Split Dataset

The doc objects were divided into training and validation sets with an 80:20 ratio, producing 852 training and 213 validation records.

e. Save as.spacy

The final DocBin objects were saved in spacy format for use in RoBERTa model training.

3. RoBERTa

In this stage, information extraction from text was performed using the RoBERTa (Robustly Optimized BERT Pretraining Approach) model, integrated into the Named Entity Recognition (NER) pipeline of spaCy (Figure 3). The workflow consisted of the following steps:

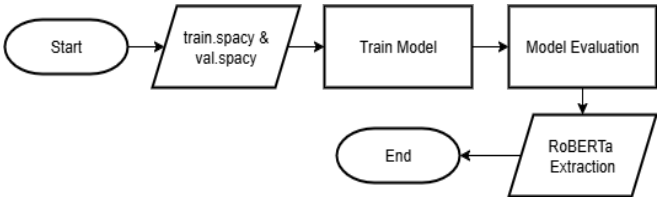


Figure 3. RoBERTa Flow

a. Train.spacy & Val.spacy

These files, generated during preprocessing, contain collections of doc objects, each including resume text and labeled entity spans. Train.spacy was used for model training, while Val.spacy served as validation data to assess model performance.

b. Model Training

RoBERTa underwent fine-tuning using the training dataset to learn patterns in the resumes and predict corresponding labels. The spaCy pipeline was trained for 4,000 steps with an initial learning rate of 4×10^{-5}, and evaluation was conducted every 200 steps.

c. Model Evaluation

The model's performance was evaluated on the validation dataset using precision, recall, and F1-score metrics.

d. Extraction Result

The output of RoBERTa represents key information extracted from resumes.

Table 5. Extraction Result

Text	Extraction
Kavya U. Network Ops	Designation : ["Network Ops"]

Table 5 presents an example of extracted results format.

4. Random Forest

The output from RoBERTa extraction was used as input for the Random Forest model, chosen for its ability to handle complex and non-linear data (Figure 4). The implementation process included the following stages:

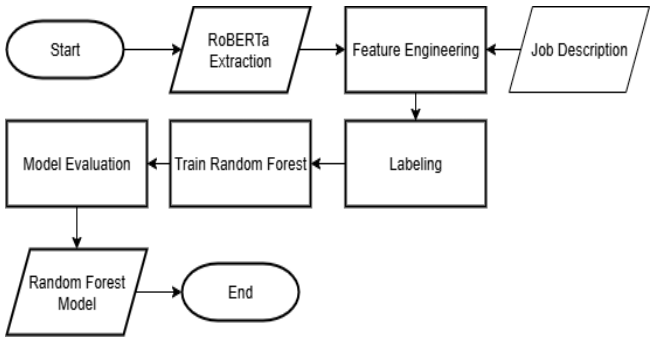


Figure 4. Random Forest Flow

a. RoBERTa Extraction Result

The data presented contains the key information systematically extracted from resumes using the RoBERTa model.

b. Job Description
Data outlining criteria and details of job vacancies, such as position, skills, education, and experience.

c. Feature Engineering
Matching each resume with a corresponding job description. Textual similarity between the two was measured using the SentenceTransformer model (stsb-roberta-base-v2), which converts text into vector. Cosine Similarity then applied to compute similarity scores for different aspects, including position, skills, education, and experience (Formula 1).

$$Cosine\ Similarity\ (A,B) = \frac{A \times B}{||A||\ ||B||} \quad (1)$$

Where A and B are the vector representations of the resume and job description, $A \times B$ are the dot product, and $||A||$ and $||B||$ are the magnitudes of the vectors.

d. Labeling
Since the original dataset had no predefined labels, unsupervised labeling was performed by clustering total similarity scores from feature engineering using the K-Means algorithm to generate target labels.

Table 6. Label and Description

Label	Description
Highly Suitable	Very relevant
Moderately Suitable	Moderately relevant
Not Suitable	Not relevant

Table 6 presents the label categories produced by this process along with their descriptions.

e. Model Training
The Random Forest model was trained using feature engineering results as input features (X) and generated labels as target variables (Y). The dataset was split into 80% training and 20% testing data.

f. Model Evaluation
The model's performance was rigorously assessed on the validation set through multiple evaluation metrics, including a confusion matrix, precision, recall, and F1-score.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Precision measures the proportion of correctly predicted positive instances out of all instances predicted as positive (Formula 2). Here, true positives (TP) are instances correctly predicted as positive, while false positives (FP) are instances incorrectly predicted as positive.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

Recall indicates the proportion of actual positive instances that were correctly indentified by the model (Formula 3). False negatives (FN) are positive instances that the model failed to predict.

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

The F1 Score combines precision and recall into a single metric, providing a balanced measure of the model's accuracy, especially when dealing with imbalanced datasets (Formula 4).

5. Explainable AI
To enhance the interpretability of the classification results, the system incorporated Explainable AI (XAI) using the SHAP (Shapley Additive Explanations) method. In this stage, the Random Forest model and the feature engineering outputs were used as inputs for SHAP. The SHAP analysis provides explanations of predictions by indicating the contribution of each feature to the final prediction.

6. Result
The filtering results contain information such as the classifications produced by the Random Forest model and the decision explanations generated by the Explainable AI component. These explanations highlight the features that contributed to the model's decision-making process.

B. System Design

The system design section describes the functional architecture of the automated resume screening system from the perspective of its users. Figure 5 illustrates the system design.

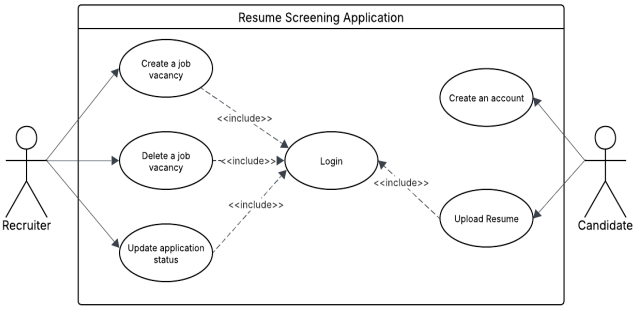


Figure 5. System Design

The actors of the system are:
a. Recruiter
Responsible for creating job vacancies, viewing screening results, and updating candidate application status.
b. Candidate
Responsible for uploading resumes and checking the status of their applications.

The main use cases of the system include:

1. Create a job vacancy
   Allows the recruiter to define and post new job openings.
2. Delete a job vacancy
   Allows the recruiter to remove an existing job posting.
3. Update Application Status
   Enables the recruiter to change the status of a candidate's application based on screening results.
4. Create an Account
   Allows candidate to register and create an account in the system.
5. Upload Resume
   Allows the candidate to submit their resume for automated screening.

III. RESULT AND DISCUSSION

This section presents the outcomes obtained from the developed resume screening system and provides an in-depth analysis of the findings. The discussion highlights the system's performance, interprets the observed results, and examines how the outputs reflect the effectiveness and reliability of the proposed approach.

A. Model Evaluation

The system was evaluated using a validation dataset derived from the original resume corpus, which had undergone preprocessing steps. The dataset was split into training and validation sets with an 80:20 ratio to ensure sufficient data for model learning while retaining enough unseen data for performance assessment.

1. RoBERTa Feature Extraction

RoBERTa was fine-tuned on the training set to extract relevant information from resumes. Validation data was used to evaluate the model performance. Various combinations of learning rate and batch size were tested to select the optimal configuration. Table 7 summarizes the experiments.

Table 7. RoBERTa Experiment Results

learning_rate	batch_size	steps	F1 Score (%)
0.00001	32	4000	77.54
0.00002	32	4000	79.25
0.00003	32	4000	79.67
0.00004	32	4000	80.02
0.00005	32	4000	80.70
0.00006	32	4000	80.22
0.00007	32	4000	79.85
0.00008	32	4000	79.89
0.00001	64	4000	79.36
0.00002	64	4000	80.55
0.00003	64	4000	80.52
0.00004	**64**	**4000**	**81.08**
0.00005	64	4000	80.88
0.00006	64	4000	80.96
0.00007	64	4000	80.73
0.00008	64	4000	80.40

The best model was achieved with learning_rate = 0.00004 and batch_size = 64. The model achieved an average F1 Score of 81.08% across resume attributes.

To further analyze the model's capability in extracting relevant information from resumes, the performance of RoBERTa was evaluated for each resume attribute individually. This evaluation uses Precision, Recall, and F1 Score to measure the accuracy, completeness, and balance of predictions for attributes such as Location, Designation, Years of Experience, Job-Specific Skills, Tech Tools, Soft Skills, Degree, College Name, Graduation Year, and Companies Worked at. The results are summarized in Table 8.

Table 8. RoBERTa Evaluation

Attribute	Precision (%)	Recall (%)	F1 Score (%)
Location	85.91	87.46	86.68
Designation	76.97	71.75	74.27
Years of Experience	78.99	88.95	83.67
Job Specific Skills	84.24	74.44	79.03
Tech Tools	84.57	86.92	85.73
Soft Skills	83.86	86.57	84.62
Degree	70.11	68.16	69.12
College Name	76.38	83.87	79.95
Graduation Year	93.84	84.26	88.79
Companies Worked at	64.01	68.33	66.11
Model Average	**81.66**	**80.53**	**81.08**

Overall, the RoBERTa model achieved an average F1-score of 81.08%. Graduation Year performed best (F1 = 88.79%), while Companies Worked at performed lowest (F1 = 66.11%), reflecting challenges due to attribute variability and limited training data. Other attributes, such as Soft Skills and Tech Tools, also showed strong performance (F1 > 80%), indicating reliable extraction of key resume information.

This analysis emphasizes the importance of examining attribute-level performance, as aggregate metrics alone may obscure weakness in specific resume sections.

2. Random Forest Classification

The embeddings generated by RoBERTa were used to train a Random Forest algorithm, which serves as the second stage of the system. Experiments were conducted varying n_estimators and max_depth to determine the optimal configuration. Results are presented in Table 9.

Table 9. Random Forest Experiment Results

n_estimator	max_depth	Accuracy (%)	F1 Score (%)
100	none	0.962	0.96
500	none	0.957	0.96
100	5	0.934	0.96
100	7	0.953	0.96
100	9	0.953	0.96
100	**11**	**0.962**	**0.96**
100	13	0.962	0.96

The optimal configuration was n_estimator = 100 and max_depth = 11, resulting in an accuracy of 96%.

Random Forest performance was further evaluated using Precision, Recall, and F1. Table 10 shows the results.

Table 10. Random Forest Evaluation

Label	Precision (%)	Recall (%)	F1 Score (%)	Support
Highly Suitable	0.99	0.99	0.99	95
Moderately Suitable	0.94	0.96	0.95	79
Not Suitable	0.95	0.90	0.92	39
Accuracy			0.96	213
Macro Avg	0.96	0.95	0.95	213
Weighted Avg	0.96	0.96	0.96	213

The classifier achieved high performance overall, with an accuracy of 96%.

To provide a more detailed analysis of predictions, a confusion matrix was also generated, illustrating the number of correct and incorrect predictions for each class (Figure 6). Based on the confusion matrix:

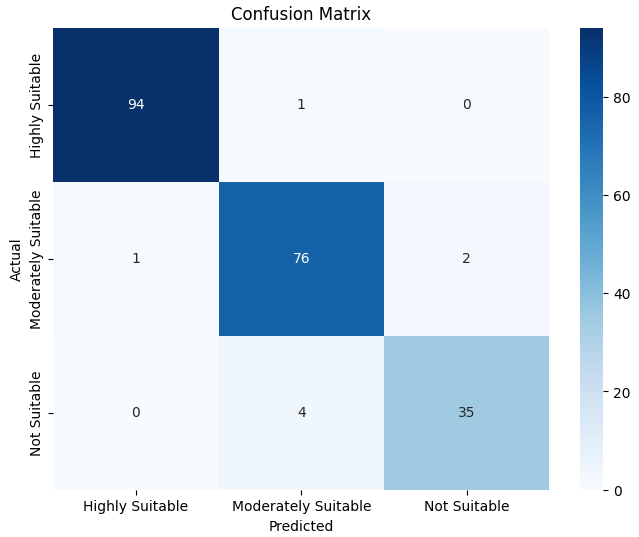


Figure 6. Confusion Matrix of the Random Forest

- Highly Suitable: Out of 95 samples, 94 were correctly predicted (True Positive) and 1 misclassified as Moderately Suitable (False Negative). One Moderately Suitable instance was misclassified as Highly Suitable (False Positive).
- Moderately Suitable: Out of 79 samples, 76 were correctly predicted. Misclassifications included 1 as Highly Suitable and 2 as Not Suitable (False Negatives). Four Not Suitable and 1 Highly Suitable instances were incorrectly predicted as Moderately Suitable (False Positives).
- Not Suitable: Out of 39 samples, 35 were correctly predicted. Misclassifications included 4 as Moderately Suitable (False Negatives) and 2 Moderately Suitable instances predicted as Not Suitable (False Positives).

B. System Evaluation

To assess the practical performance of the proposed resume screening system, a system evaluation was conducted using a set of 17 resumes in conjunction with a single job description. Each resume was manually labeled according to the expected suitability level: Highly Suitable, Moderately Suitable, or Not Suitable. The system's predictions were then compared to these expected results to determine classification accuracy and practical effectiveness. The evaluation results, summarized in Table 11.

Table 11. System Evaluation

Resume	Expected Result	System Prediction	Conclusion
Resume 1	Highly Suitable	Highly Suitable	Classification matches
Resume 2	Moderately Suitable	Moderately Suitable	Classification matches
Resume 3	Not Suitable	Not Suitable	Classification matches
Resume 4	Moderately Suitable	Highly Suitable	Classification differs
Resume 5	Highly Suitable	Highly Suitable	Classification matches
Resume 6	Not Suitable	Not Suitable	Classification matches
Resume 7	Moderately Suitable	Moderately Suitable	Classification matches
Resume 8	Highly Suitable	Highly Suitable	Classification matches
Resume 9	Highly Suitable	Highly Suitable	Classification matches
Resume 10	Highly Suitable	Highly Suitable	Classification matches
Resume 11	Moderately Suitable	Moderately Suitable	Classification matches
Resume 12	Moderately Suitable	Moderately Suitable	Classification matches
Resume 13	Not Suitable	Moderately Suitable	Classification differs
Resume 14	Not Suitable	Not Suitable	Classification matches
Resume 15	Not Suitable	Not Suitable	Classification matches
Resume 16	Not Suitable	Not Suitable	Classification matches
Resume 17	Highly Suitable	Highly Suitable	Classification matches

DOI: 10.34148/teknika.v14i3.1359
TEKNIKA, Volume 14(3), November 2025, pp. 424-432
ISSN 2549-8037, EISSN 2549-8045

Table 11 show that the system correctly classified 15 out of 17 resumes according to the expected suitability levels. Overall, the system demonstrated high agreement with human expectations, indicating that it is capable of providing reliable and interpretable recommendations in real-world resume screening scenarios.

C. System Interface and Result Explanation

The system interface is designed to facilitate seamless user interaction while clearly presenting classification results and key resume attributes.

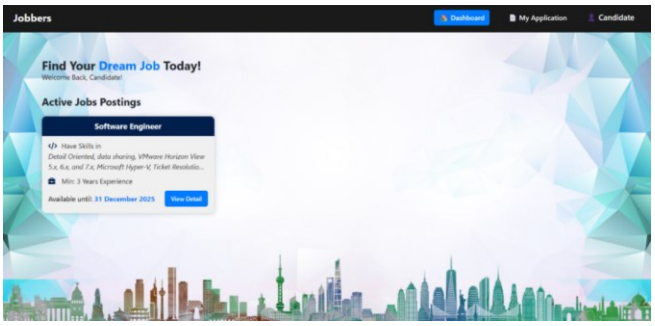


Figure 7. Dashboard Page



Figure 8. Upload Resume Page

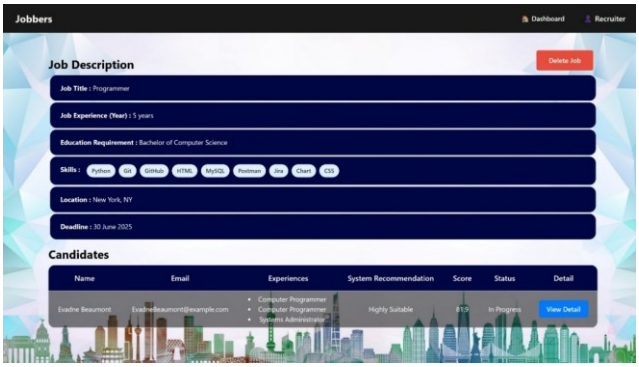


Figure 9. Job Description Page

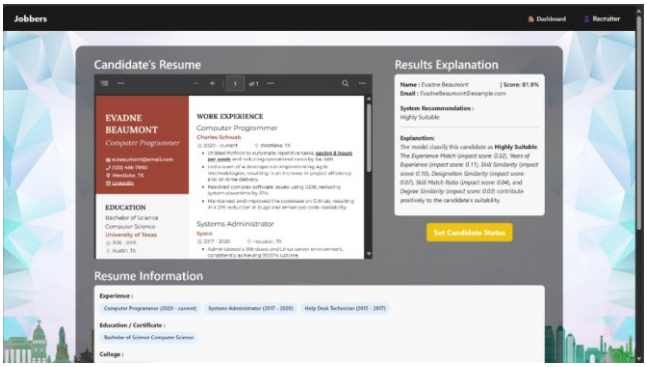


Figure 10. View Resume Page

Figures 7-10 illustrate the main components, including the dashboard, resume upload, job description, and resume viewing pages, which collectively guide users through the workflow from input to result interpretation in a coherent and intuitive manner.

The system displays the extracted attributes from RoBERTa, such as experience, education, college, and more. These attributes are then used by the Random Forest algorithm to generate candidate suitability predictions categorized into Highly Suitable, Moderately Suitable, or Not Suitable. The system also integrates explanation results using SHAP (SHapley Additive exPlanations). This feature highlights the contribution of each attribute to the classification decision, allowing users to understand why a candidate is categorized in a certain way.

D. Discussion

The results demonstrate that combining RoBERTa with a Random Forest algorithm produces a robust and accurate resume screening system. RoBERTa's contextual embeddings capture both explicit and latent information, ensuring strong performance across most resume attributes, while the Random Forest algorithm effectively categorizes candidates into Highly Suitable, Moderately Suitable, and Not Suitable classes. Analysis of the confusion matrix indicates that misclassifications mainly occur between adjacent categories, which is common in nuanced human resources data.

In practical recruitment scenarios, this system addresses key challenges faced by HR practitioners. Traditional manual resume screening is often time-consuming, inconsistent, and prone to bias, especially with large volumes of applicants. By leveraging contextual NLP and interpretable machine learning, the proposed system allows recruiter to quickly identify the most qualified candidates, make fairer decisions, and understand why a candidate was classified in a certain category. For example, borderline candidates can be evaluated with greater clarity, helping recruiters prioritize interviews more efficiently.

Moreover, the integration of Explainable AI through SHAP provides actionable insights for decision-makers. Recruiters can see which resume features most influenced the classification, supporting more transparent and accountable

CC BY SA

hiring practices. Overall, these findings demonstrate that advanced automated screening systems can not only improve efficiency but also enhance fairness and transparency in real-world recruitment processes.

E. Limitations and Future Works

This study has demonstrated that the integration of RoBERTa for contextual feature extraction, Random Forest for classification, and Explainable AI (SHAP) for interpretability provides an effective, reliable, and objective approach for automated resume screening. Nevertheless, several limitations should be carefully considered. First, the dataset employed in this research is relatively limited in size and exhibits noticeable class imbalance, which could potentially affect the generalizability and robustness of the system when applied to more diverse, real-world resumes with varying formats, lengths, and content richness. Second, while the current system achieves high performance metrics, certain resume attributes such as companies previously worked at exhibited lower F1 Scores, indicating challenges in handling attributes with high variability or limited training examples. Third, although Explainable AI enhances interpretability, its impact on actual decision-making and user trust was not formally measured in this study, leaving room for further investigating regarding user experience and practical effectiveness.

Future work could address these limitations by expanding the dataset to include a broader range of resumes, ensuring better representation of minority classes, and potentially applying data augmentation techniques to mitigate sparsity in underrepresented categories. Additionally, exploring alternative or ensemble classification methods, such as gradient boosting, stacking models, or neural network classifiers, may further improve predictive accuracy, robustness, and adaptability of the system.

IV. CONCLUSION

Based on the research methodology and the experimental results, several key conclusions can be drawn regarding the effectiveness of the proposed automated resume screening system. The fine-tuned RoBERTa model exhibits strong capability in extracting essential attributes from resumes while capturing the contextual meaning of words and phrases, effectively overcoming the limitations of traditional keyword-based matching approaches. Evaluation of the model demonstrated an average F1 score of 81.08%, Precision of 81.66%, and Recall of 80.53%, indicating consistently balanced and reliable performance across a diverse set of resume attributes, including job-specific skills, years of experience, education, and professional background.

At the classification stage, the Random Forest model categorized resumes into three suitability levels: Highly Suitable, Moderately Suitable, and Not Suitable, achieving an overall accuracy of 96%. The system is designed to enhance interpretability through SHAP-based explanations, which provide insights into feature contributions for each prediction without claiming quantitative evaluation of the XAI component.

Compared to prior studies focusing solely on information extraction or conventional ATS relying on keyword matching, this approach addresses gaps by combining contextual embeddings with ensemble learning to support more accurate, transparent, and practical candidate evaluation. The findings suggest potential improvements in the efficiency, fairness, and objectivity of initial screening, providing actionable insights for recruiters while building upon existing methodologies in automated resume assessment.

Nevertheless, the study has several limitations, including a relatively small dataset, the use of synthetic data to supplement real resumes, and the absence of a formal evaluation of XAI interpretability. Future research could expand the dataset, explore alternative classification algorithms, and assess the practical benefits of explainable AI in real recruitment settings.

REFERENCES

[1] R. V. K. Bevara, N. R. Mannuru, S. P. Karedla, B. D. Lund, T. Xiao, H. Pasem, et al., "Resume2Vec: Transforming Applicant Tracking Systems with Intelligent Resume Embeddings for Precise Candidate Matching," Electronics (Switzerland), vol. 14, no. 4, Feb. 2025, doi: 10.3390/electronics14040794.

[2] Z. Naveed, B. Nisar, M. Saifullah, and J. I. Baig, "Resume Ranking Using Natural Language Processing," Journal of Computers and Intelligent Systems, vol. 2, no. 1, pp. 61-66, May. 2024.

[3] A. Rawat, S. Malik, S. Rawat, D. Kumar, and P. Kumar, "A Systematic Literature Review (SLR) On The Beginning of Resume Parsing in HR Recruitment Process & SMART Advancements in Chronological Order," Jul. 2021, Research Square. doi: 10.21203/rs.3.rs-570370/v1.

[4] A. K. Sinha, M. A. K. Akhtar, and M. Kumar, "Automated Resume Parsing and Job Domain Prediction using Machine Learning," Indian J Sci Technol, vol. 16, no. 26, pp. 1967–1974, Jul. 2023, doi: 10.17485/IJST/v16i26.880.

[5] Q. Trint and T. T. Dang, "Automatic Process Resume in Talent Pool by Applying Natural Language Processing," International Conference on Logistics and Industrial Engineering, pp. 234-240, Jan. 2021.

[6] N. Novaković, L. Dražeta, "Applicant Tracking System: A Powerful Recruiters' Tool," In Sinteza 2024 – International Scientific Conference on Information Technology, Computer Science, and Data Science, Belgrade, Signidunum University, pp. 240-245, 2024. doi: 10.15308/Sinteza-2024-240-245.

[7] E.Albaroudi, T.Mansouri, and A. Alameer, "A Comprehensive review of AI Techniques for Addressing Algorithmic Bias in Job Hiring," AI, vol. 5, no. 1, pp. 383 404, Feb. 2024, doi: 10.3390/ai5010019.

[8] M. R. Sareddy, "Revolutionizing Recruitment: Integrating AI and Blockchain for Efficient Talent Acquisition," IMPACT: International Journal of Research in Business Management (IMPACT: IJRBM), vol. 10, no. 8, pp. 33–44, Aug. 2022.

[9] Badan Pusat Statistik, "Pencari Kerja Terdaftar, Lowongan Kerja Terdaftar. Dan Penempatan/Pemenuhan Tenaga Kerja Menurut Provinsi dan Jenis Kelamin," bps.go.id, 2023.

[10] P. Tijare, M. Waseem, M. A. Sherani, K. S. K. Sai Krishna, and P. Kavitha, "A smart resume screening tool for customized shortlisting," ITM Web of Conferences, vol. 56, p. 04001, 2023, doi: 10.1051/itmconf/20235604001.

[11] B. Kinge, S. Mandhare, P. Chavan, and S. M. Chaware, "Resume Screening using Machine Learning and NLP: A proposed system," International Journal of Scientific Research in Computer Science, Engineering and Information Technology, vol. 8, no. 2, pp. 253–258, Apr. 2022, doi: 10.32628/cseit228240.

[12] M. A. Aleisa, N. Beloff, and M. White, "Implementing AIRM: a new AI recruiting model for the Saudi Arabia labour market," Journal of Innovation and Entrepreneurship, vol. 12, no. 1, pp. 59, Sep. 2023. doi: 10.1186/s13731-023-00324w.

[13] K. Gupta, R. Jinad, and Q. Liu, "Comparative analysis of NLP models for detecting depression on Twitter," In 2023 International Conference on Communications, Computing and Artificial Intelligence (CCCAI), pp. 23-28, Jun. 2023. doi: 10.1109/CCAI59026.2023.00013.

[14] I. J. A. Biol and C. C. Abalorio, "Utilizing Named Entity Recognition for Web-Based Resume Scoring," International Journal of Engineering Trends and Technology, vol. 72, no. 7, pp. 381-387, 2024. doi: 10.14445/22315381/IJETT-V7217P142.

[15] A. Pina, C. Petersheim, J. Cherian, J. N. Lahey, G. Alexander, and T. Hammond, "Using Machine Learning with Eye-Tracking Data to Predict if a Recruiter Will Approve a Resume," Mach Learn Knowl Extr, vol. 5, no. 3, pp. 713–724, Sep. 2023, doi: 10.3390/make5030038.

[16] H. Al-Quhfa, A. Mothana, A. Aljbri, and J. Song, "Enhancing Talent Recruitment in Business Intelligence Systems: A Comparative Analysis of Machine Learning Models," Analytics, vol. 3, no. 3, pp. 297–317, Jul. 2024, doi: 10.3390/analytics3030017.

[17] B. Surendiran, T. Paturu, H. V. Chirumamilla, and M. N. R. Reddy, "Resume Classification Using ML Techniques," in Proceedings of 2023 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication, IConSCEPT 2023, Institute of Electrical and Electronics Engineers Inc., 2023, doi: 10.1109/IConSCEPT57958.2023.10169907.

[18] M. Aria, C. Cuccurullo, and A. Gnasso, "A comparison among interpretative proposals for Random Forests," Machine Learning with Applications, vol. 6, p. 100094, 2021. doi: 10.1016/J.MLWA.2021.100094.

[19] G. Airlangga, "Decoding Energy Usage Predictions: An Application of XAI Techniques for Enhanced Model Interpretability," Indonesian Journal of Artificial Intelligence and Data Mining, vol. 7, no. 2, p. 275, Apr. 2024, doi: 10.24014/ijaidm.v7i2.29041.

[20] G. Srivastava, R. H. Jhaveri, S. Bhattacharya, S. Pandya, Rajeswari, P. K. R. Maddikunta, et al., "XAI for Cybersecurity: State of the Art, Challenges, Open Issues and Future Directions," ACM Computing Surveys, vol. 1, no. 1, pp. 1-33, Jun. 2022, doi: 10.1145/1122445.1122456.

[21] L. Hofeditz, S. Clausen, A. Rieß, M. Mirbabaie, and S. Stieglitz, "Applying XAI to an AI-based system for candidate management to mitigate bias and discrimination in hiring," Electronic Markets, vol. 32, no. 4, pp. 2207–2233, Dec. 2022, doi: 10.1007/s12525-022-00600-9.