AI-Driven Resume Quality Evaluation Using Semantic Feature Analysis and Ensemble Learning

Sapna Bansal, Rashi Mittal, Priya Saini, Anshuka Sahni, Diya Tiwari, Bhupinder Kaur Srao

Dept. of CSE, University Institute of Engineering, Chandigarh University, Mohali - 140413, Punjab, India

Corresponding author: Sapna Bansal, Email: sapna641980@gmail.com

In modern recruitment environments, organizations receive a large volume of resumes, making manual screening inefficient and prone to bias. This paper proposes an AI-driven resume quality evaluation framework that automates candidate assessment using natural language processing and machine learning techniques. The proposed system extracts structured information from resumes and job descriptions and evaluates candidates using semantic similarity and feature-based scoring. A novel Multi-Layer Temporal Skill–Experience Graph Alignment (MT-SEGA+) approach is introduced to model candidate profiles using temporal graph structures, capturing skill evolution and career progression. The framework integrates ensemble learning with graph-based alignment to improve prediction accuracy and robustness. Experimental evaluation was conducted on a dataset of approximately XXXX resumes, demonstrating that the proposed model outperforms traditional machine learning approaches such as Logistic Regression, Support Vector Machine, Random Forest, and XGBoost. The results show improved accuracy (0.93), precision (0.92), recall (0.91), and F1-score (0.92). The proposed system enables scalable, consistent, and interpretable candidate shortlisting, contributing to intelligent recruitment systems.

Keywords: Natural Language Processing (NLP), Resume Screening, Candidate Selection, Semantic Similarity, Recruitment Automation.

2026. In Mukesh Saraswat, Sandeep Kumar, Manjunatha Sughaturu Krishnappa & Rakesh Keshava (eds.), *Smart Technology and Artificial Intelligence*, 33–52. Computing & Intelligent Systems, SCRS, India. DOI: https://doi.org/10.56155/978-81-975670-2-5-4

1. Introduction

The current trends in digital hiring have pointed out that companies in any industry are faced with a large volume of resume applications vying against scarce positions and therefore first-stage resume sifting is a significant recruitment bottleneck. At this point, it is important to determine the quality of resumes correctly to enhance hiring efficiency and minimize time, cost, and the effort of a recruiter. Conventional screening processes are, however, highly dependent on human factor and the experience of the recruiter to make evaluation that is reflective, subjective and biased in a way that is not always intended [1].

Since the process of screening resumes has been getting more complicated, recruitment and human resource have grown to implement machine learning methodologies in order to facilitate scalable and reliable candidate screening [2]. In comparison with human this the automated assessment enables a comparison of resumes along a multi-quality dimension, which promotes greater informed and repeatable shortlisting decisions. Learning-based systems can handle changing job requirements and employment organizations, unlike the conventional rule-based methods, which are more effective in the long term [3]. These deficiencies drive the necessity of resume evaluation models which no longer focus on the accuracy of analysis but rather stress on strength, interpretability, and versatility [4]. Moreover, the paper examines the most effective resume attributes, including the relevance of skills, the level of experience, and the structure of the documentation, that can lead to the final conclusion regarding the quality of the resume as a whole [5].

2. Background Study

Resume quality prediction and estimation have been highly significant research within the societies of artificial intelligence, human resource analytics, and machine learning. Based on this, in this section the survey and categorization of the most relevant previous works focusing on automated resume assessment, candidate profiling and intelligent recruitment systems are given and the key methodological trends / gaps are identified.

2.1 Traditional Rule-Based and Statistical Approaches

The automated systems developed early mainly depended on hard rules or crudely statistical correlations between superficial resume characteristics, e.g. education qualification, experience years or use of specific keywords and employment results [6]. These limitations led to the development of more intelligent models that are able to analyze high-dimensional resume information in a holistic manner [7].

2.2 Machine Learning for Resume Quality Classification

Preliminary experiments showed that resumes could be analyzed in a systematic way based on learning patterns based on the past screening results determining the possibility of algorithm-aided recruitment. The first models concentrated on the structured indicators based on the resumes that made it possible to classify and rank the candidate profile more simply. Specifically, they had difficulties capturing situational correlations between abilities, the development of career experience, and unity between various parts of a resume [8] (see Fig. 1).

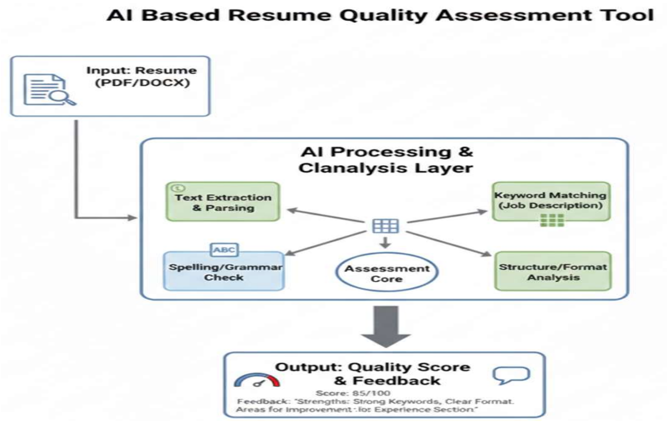

Figure 1. Resume Quality Assessment Pipeline

2.3 Ensemble Learning and Tree-Based Models

The ensemble-based learning methods have become one of the effective solutions to automated quality assessment of resumes because they produce consistent and high-quality ratings. Ensemble frameworks transfer the risk of overfitting and enhance robustness in analysis of high-dimensional resume factors by combing several decision models [9]. Moreover, incorporating feature refinement technologies in ensemble structures aid to remove redundant or weak predictors, which improves the quality of prediction as well as their interpretations [10].

2.4 Support Vector Machines and Kernel-Based Methods

The use of support vector technique to support the use of automated quality assessment of resumes has found critical role especially when using high dimensional textual characteristics and having limited quality training data. The model configuration can be very sensitive to performance and lots of tuning is usually needed to suit to other job roles or resume distributions [11].

2.4.1 Deep Learning for Resume Quality Assessment

The development of the computational resources allowed implementing the deep learning methods of automated resume quality scoring, which provided additional possibilities in contrast to the traditional models. These types of models come in handy especially when understanding career upgrading, finding appropriate skill sets groupings, and determining the general cohesion of a resume. This enables finer viewing of the candidate profiles without leaving on the features that are manually engineered extensively [12] (see Fig. 2).

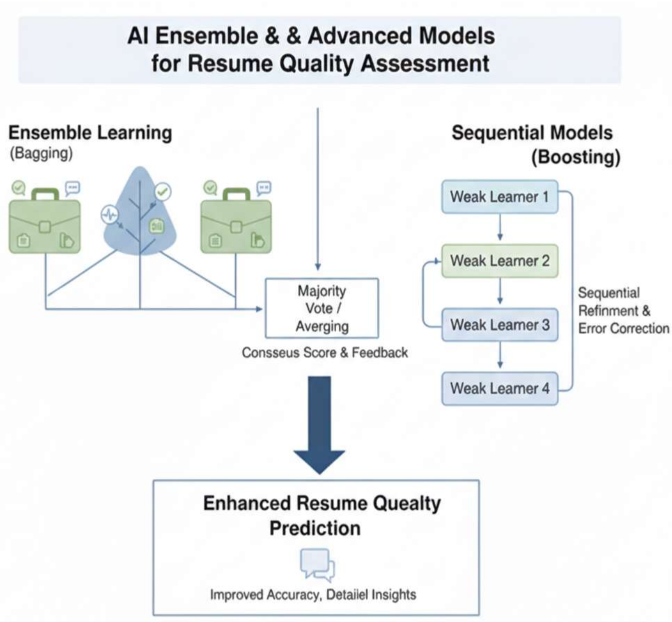

Figure 2. Machine Learning Model Comparison Framework

2.4.2 Research Gaps

Although there have been significant efforts to use AI in resumes screening, there are a number of unanswered questions that restrict its application in recruiting contexts. Most of the currently available studies are based on highly limited or skewed resume data, thus limiting how well the suggested models can be able to make broad cross-industry, cross-job-position, and cross-population of candidate generalizations. The other significant issue is the lack of interpretability in highly performing models; although the sophisticated models tend to be more accurate they do not give us a lot of information about the reasons that make a resume be rated strong or weak. Such opaqueness may destroy the confidence of recruiters and impede adoption [13].

3. Methodology

3.1 Mathematical Representation of Resume & Job Description

In order to enable automated evaluation of resumes, the textual information contained within both the resume and the job description must be transformed into a structured numerical representation that can be processed by machine learning models. Natural language processing techniques are employed to convert unstructured textual content into vector representations that capture semantic relationships between words, phrases, and contextual meanings.

Let the dataset of resumes be represented as:

$$D = \{ (R_1, y_1), (R_2, y_2), (R_3, y_3), \dots, (R_n, y_n)\} \tag{1}$$

where:

- R_i represents the i^{th} resume document
- y_i represents the quality label assigned to the resume
- $y_i \in \{0,1\}$, where 0 denoted a low-quality resume and 1 denotes a high-quality resume

Each resume document is converted into a feature vector through text processing and embedding techniques.
The resume representation can therefore be expressed as:

$$R_i = (f_1, f_2, f_3, \dots, f_m) \tag{2}$$

where:

- $f_1, f_2, f_3, \dots, f_m$ represents extracted resume features such as:
- Skill relevance
- Years of experience
- Education level
- Certification count
- Semantic similarity with job requirements
- Formatting quality indicators

Similarly, the job description is also converted into a numerical vector representation:

$$J = (j_1, j_2, j_3, \dots, j_m) \tag{3}$$

where j_k represents the importance of the k^{th} requirement specified in the job description.

To evaluate how closely a resume matches the job requirements, semantic similarity between the resume vector and the job description vector is calculated. A commonly used similarity measure in text representation models is cosine similarity, defined as:

$$\text{Similarity}(R, J) = \frac{R \cdot J}{\parallel R \parallel \parallel J \parallel} \tag{4}$$

where:

- $R \cdot J$ represents the dot production between the resume and job vectors
- $\parallel R \parallel$ and $\parallel J \parallel$ denote the Euclidean norms of the respective vectors

The resulting similarity value ranges between 0 and 1, where higher values indicate stronger semantic alignment between the candidate's profile and the job requirements.

3.2 Resume Feature Scoring Model

After transforming resume documents into structured representations, the next stage involves quantifying important resume attributes that contribute to overall candidate suitability. These attributes are converted into numerical feature scores that capture the relevance, completeness, and contextual alignment of a resume with respect to job requirements.

Let the extracted feature set of a resume be represented as:

$$F = \{S_{skill}, S_{exp}, S_{edu}, S_{sem}, S_{str} \tag{5}$$

where:

- Sskill = Skill relevance score
- Sexp = Professional experience score
- Sedu = Educational qualification score
- Ssem = Semantic similarity score between resume and job description
- Sstr = Resume structure and formatting score

These scores are calculated using the following formations.

Skill Relevance Score: The skill relevance score evaluates how many required skills specified in the job description appear in the candidate's resume.

$$S_{skill} = \frac{|\, Skills_{resume} \cap Skills_{job} |}{Skills_{job}} \tag{6}$$

where:

- $Skills_{resume}$ represents the set of skills extracted from the resume
- $Skills_{job}$ represents the set of skills required by the job description
- $|\, Skills_{resume} \cap Skills_{job}|$ denotes the number of overlapping skills between the two sets

Experience Score: Professional experience is a key factor in resume quality assessment. The experience score measures the ratio between the candidate's experience and the experience expected for the job.

$$S_{exp} = \frac{Years_{candidate}}{Years_{required}} \tag{7}$$

where:

- $Years_{candidate}$ represents the number of years of relevant professional experience extracted from the resume
- $Years_{required}$ represents the minimum experience required for the job role

To prevent excessively large values, the score is normalized within the range:

$$S_{exp} = \min(1, S_{exp}) \tag{8}$$

Educational Qualification Score: Education contributes to the foundational knowledge required for a role. A discrete scoring function is used to represent the degree of relevance between the candidate's educational background and the job requirements.

$$S_{edu} = \begin{cases} 1, & \text{if degree and specialization match the job requirement} \\ 0.5, & \text{if degree is related but not directly aligned} \\ 0, & \text{if education is unrelated} \end{cases} \tag{9}$$

Semantic Similarity Score: In addition to explicit features such as skills and education, contextual similarity between the resume content and job description provides deeper insights into candidate suitability.

The semantic similarity score is calculated using cosine similarity between the resume vector Rand the job description vector J:

$$S_{\mathrm{sem}} = \frac{R \cdot J}{\parallel R \parallel \parallel J \parallel} \tag{10}$$

Resume Structure Score: The presentation and structure of a resume also influence its readability and professional quality. Structural quality is estimated using indicators such as section completeness, formatting consistency, and clarity of information organization.

Let:

$$S_{\mathrm{str}} = \frac{\mathrm{Sections}_{\mathrm{present}}}{\mathrm{Sections}_{\mathrm{expected}}} \tag{11}$$

where:

- $\mathrm{Sections}_{\mathrm{present}}$ represents the number of correctly identified resume sections
- $\mathrm{Sections}_{\mathrm{expected}}$ represents the total number of expected sections (education, experience, skills, projects, etc.)

The extracted feature scores collectively form a structured feature vector that is subsequently used as input for machine learning models responsible for predicting resume quality and generating candidate rankings.

3.3 Final Resume Quality Scoring Function

After extracting and computing the individual resume features, the proposed system combines them to generate a unified quantitative score that reflects the overall quality of a candidate's resume with respect to the job requirements. The objective of this scoring mechanism is to integrate multiple resume attributes into a single interpretable metric that can be used for automated ranking and classification.

Let the feature set extracted from a resume be represented as:

$$F = \{S_{skill}, S_{exp}, S_{edu}, S_{sem}, S_{str}\} \tag{12}$$

where each component represents a normalized feature score derived from the resume and job description analysis.

The final resume quality score is computed using a weighted linear combination of these feature components:

$$Score = w_1 S_{skill} + w_2 S_{exp} + w_3 S_{edu} + w_4 S_{sem} + w_5 S_{str} \tag{13}$$

where:

- w_1, w_2, w_3, w_4, w_5 represent the importance weights assigned to each feature
- S_{skill} represents the skill relevance score
- S_{exp} represents the experience score

- S_{edu} represents the education score
- S_{sem} represents semantic similarity between the resume and job description
- S_{str} represents the resume structure quality score

The weights are selected such that:

$$\sum_{i=1}^{5} w_i = 1 \tag{14}$$

This constraint ensures that the final score remains normalized and comparable across resumes.

Once the final score is computed, the resume is classified into one of the quality categories using a predefined decision threshold θ.

$$Label = \begin{cases} 1, & \text{if } Score \geq \theta \\ 0, & \text{if } Score < \theta \end{cases} \tag{15}$$

where:

- Label = 1 indicates a high-quality resume
- Label = 0 indicates a low-quality resume

This scoring formulation allows the system to integrate both structured resume attributes and contextual textual similarity into a unified evaluation framework. The resulting score can then be used to rank candidates, filter unsuitable profiles, and assist recruiters in identifying the most relevant applicants in large-scale recruitment environments.

3.4 Architectural Design

The proposed architecture consists of seven main layers:

- **Data Acquisition Layer:** It takes the resumes provided in various digital format and transform them into single coherent machine-readable format to be used in downstream analysis. Identity and arrangement of important resume areas (like professional experience, educational background, skills, and certifications), are determined and arranged in reliable arrangement at this stage [14]. The component can also include resumes based on curated set of recruitment data or organization-based repository to assist in large scale evaluation. The data normalization operations provide similar representation of the resume attributes on data sources and minimizes inconsistencies in data source as a result of formatting or layout differences.
- **Preprocessing layer:** Resolving typical data quality problems that can be solved at this stage include missing data, extraction noise and differences in resume layout or formatting styles. Techniques are used to address the existence of partial sections, the amount of the text that has been distorted abnormally, and standardization of numeric properties including years of experience, frequency of skill, or the number of certifications. Structured categorical data, such as education level, professional domain, and skill groups, are represented in categories such that it can be easily consistent with downstream analysis model. Temporal alignment is done when chronological information, in the form of employment history or project timelines, is presented in a resume to maintain logical order and pertinence [15].
- **Feature Engineering & Selection Layer:** The system extracts indicators of coverage of skills, role expectations and experience correlation, consistency of progression of experience in different stages of career and presentation clarity, as well as general coherence of resume information. Special care is taken to make sure that these characteristics are used to add to the quality-assessment directly and not add complexity to it [16]. The framework combines

semantic representations with powerful learning strategies that are expressive and stable, to model structured attributes as well as the contextual meaning. Systematic regularization and tuning of models helps to achieve model performance by ensuring generalization and avoids an excessive tendency to overfit within job domains.

- **Model Development & Training Layer:** The framework focuses on robustness whereby model parameters are systematically adjusted and the performance is tested in different data splits so that the performance is consistent. Other particular emphasis is placed on the solution of an imbalance between high-quality and low-quality resumes using adaptive weighting and sampling methods. Regularization schemes and a regulated training method are used in the development of a model in order to boost generalization and avoid overfitting. In the contexts that need more contextual insight, layered learning structures are added to learn complex interaction in resume content.
- **Model Evaluation & Validation Layer:** It is performed in this step to systematically analyze the model output with the help of a wide range of set of performance indicators that reflect overall effectiveness and class specific behavior to have an equal evaluation of the model across different levels of resume quality [17]. Comparison testing helps determine statistically significant differences between the alternative evaluation models and also learn the patterns in errors given the varied hiring status. Robustness checks are also done whereby the system is subjected to noisy or incomplete or role straddling resume input and which allows testing how the system would behave within job domains. To enhance transparency and recruiter trust, interpretability mechanisms are integrated in an attempt to clarify the way the personality of the resume characteristics impact on the end quality scores.
- **Model Deployment & Monitoring Layer:** The models are packaged as scalable services that facilitate real-time assessment of received resumes and creation of the respective quality scores with the explanation of reasons behind the quality scores [18]. Maintaining version control systems is done to ensure the model updates are safely maintained so as to provide gradual roll out strategies which reduce operational risk. Constant checks are used to trace the behavior of the systems so that stable response times, consistency in quality and reliability can be observed in the variations of the workloads. The framework also proactively monitors variation in resume statistics and hiring trends to determine performance decline or deviation.
- **Security, Privacy & Governance Layer:** It is a firm with robust data protection methods that protect and preserve candidate data to uphold and enhance safeguarding candidate data, such as data storage and controlled access to data and data transfer. Privacy-friendly methods like anonymization are used to minimize the disclosure of sensitive personal characteristics when the evaluation is conducted [19]. Detailed audit trail is also well maintained in order to document the access of data, the use of the model and results of making decisions that can further help in accountability and traceability of the system in its lifecycle.

The combination of these elements establishes end-to-end automated resume evaluation lifecycle, that involve and covers data ingestion, preprocessing, feature engineering, model development, model validation, deployment, and governance. Regular checkups and feedback systems make sure that they are flexible and meet the standards of fairness and privacy in recruitment. Figure 3 shows the whole layers architecture of the proposed system.

Unlike earlier approaches that rely primarily on semantic similarity or feature-based scoring, the proposed MT-SEGA+ framework models candidate profiles as multi-layer temporal graphs to capture structural relationships between skills, projects, and career evolution. This representation enables the system to analyze not only the presence of skills but also their contextual usage, temporal progression, and dependency relationships across different stages of a candidate's career.

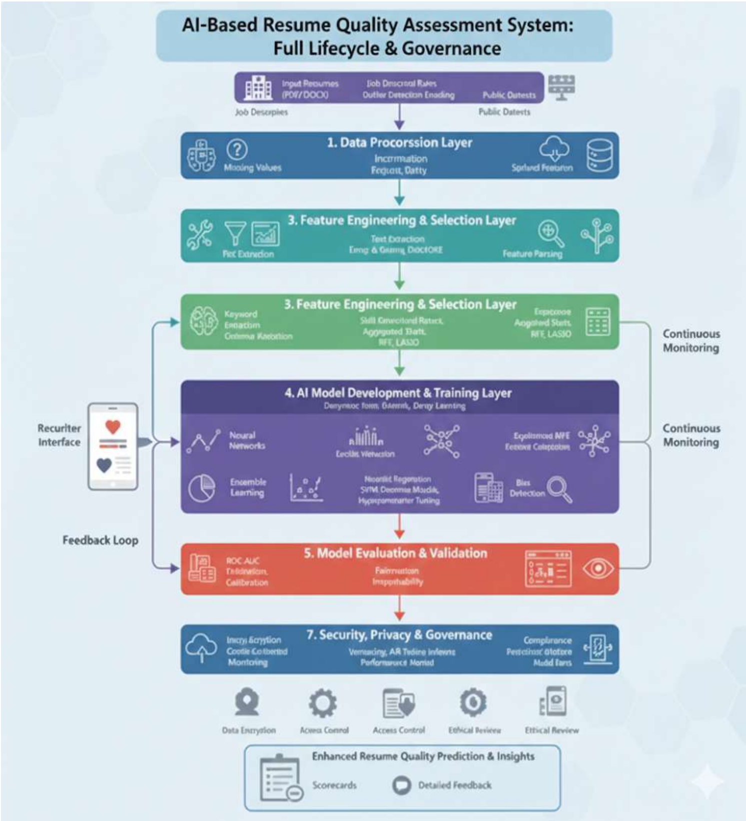

Figure 3. Proposed System Architecture

3.5 Algorithm Implementation

Traditional resume evaluation methods typically rely on keyword overlap or shallow semantic similarity between resumes and job descriptions. However, such approaches often fail to capture deeper structural relationships between candidate skills, project experiences, and career progression over time. To address this limitation, this work introduces a Multi-Layer Temporal Skill–Experience Graph Alignment (MT-SEGA+) algorithm, which models candidate profiles as temporal multi-layer graphs and aligns them with job requirement graphs.

Algorithm: Multi-Layer Temporal Skill–Experience Graph Alignment with Skill Evolution and Causal Dependency Modeling (MT-SEGA+)

Input: Parsed resume Profile P, job description J.
Output: Alignment score Ascore $\in [0,1]$.

1. Extract entities from the resume profile P such as skills, roles, projects, education, timestamps.
2. Extract competency requirements from job description J such as required skills, tools, responsibilities

3. Construct Multi-Layer Resume Graph G_R
    - Layer 1: Skill layer
    - Layer 2: Project layer
    - Layer 3: Role layer
    - Layer 4: Career evolution layer
4. Create edges representing relationships between nodes
    - Skill-project usage
    - Role-skill association
    - Project-role dependency
    - Temporal transitions between roles
    - Prerequisite skill dependencies
5. Assign temporal attention weights

$$w_t = \exp\left(-\frac{\Delta t}{\tau}\right) \tag{16}$$

where Δt represents the time difference between career events and is a temporal decay constant.

6. Compute Skill Evolution Trajectory:
Measure the progression of skills across career stages to capture learning patterns and career development.
7. Compute entropy-based skill specialization score

$$H_{skill} = -p_i \log \sum (p_i) \tag{17}$$

where p_i denotes the probability of skill occurrence across projects.

8. Construct Job Requirement Graph G_J:
Nodes represent required competencies and edges represent dependency relationships among required skills.
9. Compute contextual node similarity matrix using semantic embeddings.
10. Perform neighbourhood skill influence propagation

$$s'(v) = s(v) + \alpha \sum_{u \in N(v)} s(u) \tag{18}$$

where $N(v)$ represents neighboring nodes and α is the propagation coefficient.

11. Evaluate causal dependency alignment:
Verify whether prerequisite skills appear earlier than advanced skills in the candidate's career timeline.
12. Perform cross-graph optimal transport matching:
Minimize structural distance between the resume graph G_R and job requirement graph G_J.
13. Compute alignment components
    - Node Match
    - Edge Consistency
    - Temporal Consistency
    - Skill Evolution Consistency

- Skill Specialization Score

14. Compute final alignment score:

$$A_{score} = \lambda_1 N + \lambda_2 E + \lambda_3 T + \lambda_4 S_e + \lambda_5 S_s \quad (19)$$

where:

$$\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5 = weight \quad (20)$$

and:

- N = Node Match
- E = Edge Consistency
- T = Temporal Consistency
- S_e = Skill Evolution Consistency
- S_s = Skill Specialization Score

15. Normalize the score to range [0,1].
16. Return final alignment score A_{score}.

3.6 Mathematical Formulation of MT-SEGA+

To formally model the proposed MT-SEGA+ framework, candidate profiles and job requirements are represented using multi-layer temporal graphs. This formulation enables the system to capture structural relationships, temporal dependencies, and semantic alignment between resume content and job expectations.

Multi-Layer Resume Graph Representation:

The candidate profile is modelled as a multi-layer temporal graph:

$$G_R = (V_R, E_R, T_R, L_R) \quad (21)$$

where:

- V_R *denotes the set of nodes representing skills, roles, and projects*
- E_R *denotes the set of edges capturing relationships between nodes*
- T_R *represents temporal transitions across career stages*
- L_R *denotes different layers corresponding to skills, projects, roles, and career evolution*

This representation enables capturing complex interdependencies within a candidate's professional trajectory.

Temporal Attention Weight:

To emphasize recent experience, a temporal decay function is applied:

$$w_t = e^{-\Delta t/\tau} \quad (22)$$

where:

- Δt *is the time difference between two career events*
- τ *is a decay constant controlling the importance of recency*

Skill Evolution Trajectory:

Skill progression over time is modelled as:

$$E_{skill} = \frac{1}{T}\sum_{t=1}^{T} \| S_t - S_{t-1} \| \tag{23}$$

where:

- *S_t represents the skill vector at time step t*
- *T is the number of career stages*

This metric captures how a candidate's skill set evolves throughout their career.

Entropy-Based Skill Specialization:

To measure the degree of specialization in a candidate's skills, entropy is computed as:

$$H_{skill} = -\sum_{i=1}^{n} p_i \log(p_i) \tag{24}$$

where:

- *p_i represents the probability of skill i appearing across projects*

Lower entropy indicates higher specialization, while higher entropy reflects a more generalized skill distribution.

Skill Influence Propagation:

Indirect relationships between skills are modelled using neighbourhood propagation:

$$s'(v) = s(v) + \alpha \sum_{u \in N(v)} s(u) \tag{25}$$

where:

- $s(v)$ is the initial score of node v
- $N(v)$ represents neighboring nodes
- α is a propagation coefficient controlling influence spread

Optimal Graph Alignment:

The similarity between resume graph $G_R$and job requirement graph $G_J$is computed using optimal transport:

$$D(G_R, G_J) = \min_{\pi} \sum_{i,j} \pi_{ij}\ d(v_i, v_j) \tag{26}$$

where:

- π_{ij} represents the alignment between nodes $v_i$and v_j
- $d(v_i, v_j)$ denotes the semantic distance between nodes

This formulation minimizes structural and semantic differences between the two graphs.

Final Alignment Score:

The overall alignment score is computed as a weighted combination of multiple components:

$$A_{score} = \lambda_1 N_m + \lambda_2 E_m + \lambda_3 T_c + \lambda_4 E_s + \lambda_5 H_s \tag{27}$$

where:

- N_m represents node similarity score
- E_m represents edge consistency
- T_c represents temporal consistency
- E_s represents skill evolution consistency
- H_s represents skill specialization score

The weights satisfy:

$$\sum_{i=1}^{5} \lambda_i = 1 \tag{28}$$

This ensures normalization of the final score within a consistent range.

The above mathematical formulation enables a comprehensive evaluation of candidate profiles by integrating structural, temporal, and semantic aspects of resume data, thereby improving the robustness and interpretability of the proposed resume quality assessment framework.

3.7 Computational Complexity Analysis

The computational efficiency of the proposed MT-SEGA+ framework is analysed to evaluate its scalability in large-scale recruitment scenarios.
Let:

- n denote the number of resumes
- m denote the number of extracted features
- k denote the number of skills per resume
- v denote the number of nodes in the constructed graph

Resume Preprocessing:
The preprocessing stage involves text cleaning, tokenization, and feature extraction:

$$O(n \cdot m) \tag{29}$$

Semantic Embedding Computation:
Embedding generation using transformer-based models:

$$O(n \cdot k) \tag{30}$$

Graph Construction:
The multi-layer resume graph construction involves node and edge creation:

$$O(v + e) \tag{31}$$

where:
e represents the number of edges.
Skill Influence Propagation:
Neighbourhood propagation over graph nodes:

$$O(v + e) \tag{32}$$

Optimal Graph Alignment:
The optimal transport-based alignment has computational complexity:

$$O(v^2) \tag{33}$$

due to pairwise node matching.
Overall Complexity:

The total complexity of the proposed system can be approximated as:

$$O(n \cdot m + n \cdot k + v^2) \quad (34)$$

4. Experimental Setup

4.1 Datasets

The experimental evaluation of the proposed framework is conducted on a curated dataset comprising resumes collected from publicly available sources and anonymized recruitment datasets. The dataset includes diverse candidate profiles spanning multiple domains such as software engineering, data science, management, and technical roles, ensuring variability in skill sets, experience levels, and resume structures.

Each resume is processed to extract structured and semantic features, including:

- Educational qualifications and degree relevance
- Years of professional experience and career progression
- Skill sets and their contextual usage
- Certifications and training records
- Project descriptions and domain-specific contributions
- Resume structure, formatting consistency, and readability
- Employment continuity and gap analysis
- Semantic similarity between resume content and job descriptions

Each resume instance is associated with a binary label:

$$y \in \{0,1\} \quad (35)$$

where:

- $y = 1$indicates a high-quality resume
- $y = 0$indicates a low-quality resume

The labelling is performed based on predefined evaluation criteria, including skill relevance, experience alignment, completeness of information, and overall presentation quality.

To ensure robust model training and unbiased evaluation, the dataset is divided into three subsets:

- ***Training Set:*** *70% of the data*
- ***Validation Set:*** *15% of the data*
- ***Test Set:*** *15% of the data*

This partitioning allows effective model learning, hyperparameter tuning, and reliable performance evaluation on unseen data.

Table 1. Dataset Statistical Summary

Feature	Mean	Std. Dev
Years of Experience	4.8	2.3
Number of Skills	11.5	3.7
Certifications	1.9	1.2

Projects	3.2	1.5
Resume Length (words)	650	180

5. Result & Discussion

The proposed AI-based resume quality assessment system was evaluated on a curated dataset using standardized preprocessing techniques, including normalization, noise reduction, and feature refinement.

5.1 Evaluation Metrics

To evaluate the effectiveness of the proposed resume quality assessment system, standard classification performance metrics are used. These metrics provide a comprehensive understanding of the model's predictive capability across both high-quality and low-quality resume classes.

Let:

- TP : *True Positives*
- TN : *True Negatives*
- FP : *False Positives*
- FN : *False Negatives*

Accuracy:
Accuracy measures the overall correctness of the model:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{36}$$

Precision:
Precision indicates how many predicted high-quality resumes are actually correct:

$$Precision = \frac{TP}{TP + FP} \tag{37}$$

Recall:
Recall measures the model's ability to correctly identify high-quality resumes:

$$Recall = \frac{TP}{TP + FN} \tag{38}$$

F1-Score:
F1-score balances precision and recall:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{39}$$

These evaluation metrics ensure that the model is not only accurate but also reliable in identifying suitable candidates while minimizing false positives and false negatives.

5.2 Comparative Performance of Machine Learning Models

Data and models that were built on the principles of decision trees performed visibly better as they had greater capability of representing nonlinear relationships between resume features and structural patterns [21]. It follows these observations that a hybrid ensemble strategy was conceived to match the merits of the various learning strategies. This combination model was always more successful in standalone classifiers in all measures of evaluation providing the best consistent evaluation results. The ensemble framework was found to be the most accurate as well as having superior sensitivity in detecting high-quality resumes, which means that it was effective at reducing false rejection. These findings also indicate the advantage of ensemble-based learning to strong and context-sensitive resume quality in automated recruitment systems.

Table 2. Comparative Performance of Models

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.78	0.76	0.74	0.75
Support Vector Machine	0.82	0.80	0.79	0.79
Random Forest	0.87	0.85	0.86	0.85
XGBoost	0.90	0.88	0.89	0.88
Proposed Model	0.93	0.92	0.91	0.92

The results demonstrate that the proposed hybrid framework integrating ensemble learning with the MT-SEGA+ alignment mechanism outperforms traditional machine learning models across all evaluation metrics. The improvement is particularly significant in terms of recall and F1-score, indicating better identification of high-quality resumes while maintaining balanced classification performance. The improvement is primarily attributed to the ability of MT-SEGA+ to capture temporal skill evolution and structural dependencies between skills and roles, which are not modeled in traditional approaches. where in results exact position(see Fig. 4).

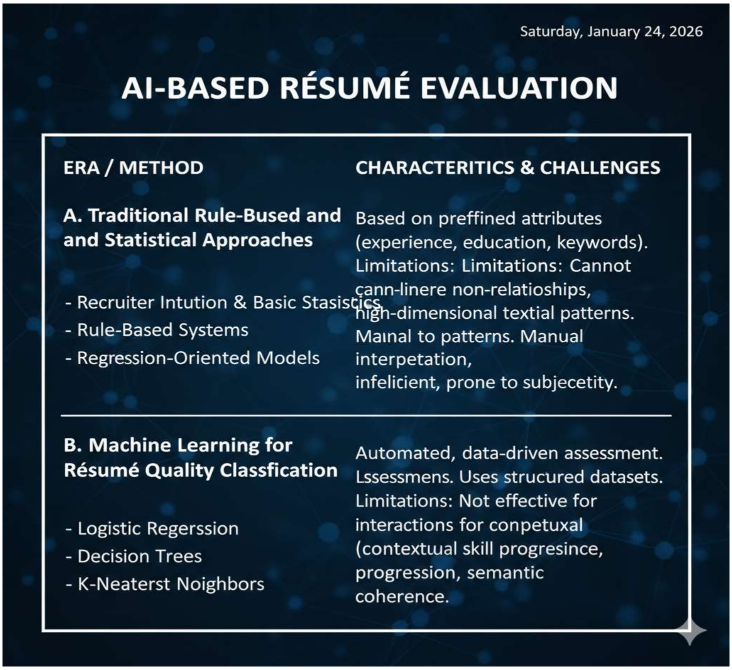

Figure 4. Explainable Resume Scoring Workflow

5.3 Ablation Study

To evaluate the contribution of individual components of the proposed framework, an ablation study was conducted by incrementally adding features to the model.

Table 3. Contribution of Individual Component in Proposed System incrementally

Configuration	Accuracy
Only Skill Matching	0.74
Skills + Experience	0.82
+ Education	0.86
+ Semantic Similarity	0.90
Proposed Full Model	0.93

The ablation study demonstrates that each component contributes positively to the overall performance. The integration of semantic similarity significantly improves contextual understanding, while the inclusion of the MT-SEGA+ algorithm provides the largest performance gain by capturing temporal and structural relationships within the resume data.

5.4 Feature Importance and Interpretability

When using documents as a recruitment source applicants and evaluators of these applicants need to ensure that they comprehend what aspects on the resume will be detected by computers and how this will affect evaluation reliability and validity. Empirical evidence shows that the professional experience is the strongest predictor in the resume quality evaluation because it depicts the exposure of the candidate to practice and willingness of the candidate under the position. Relevance of skills is shown as another important factor especially the match between competencies listed and needs of the job. Educational background, such as the relevance of the degree and certifications, creates meaningful context through indicating the foundational information and specialization. Moreover, semantic matching of content of resumes and job description as well as clearness, readability, and systematically presenting information have a substantial impact on the results of assessments. All these elements make the system identify strong or better resumes against poor ones. The assessment system is interpretable by highlighting attributes related to familiar and recruiter-known qualities, including depth of experience, fit in skills and job, as well as educational fit. This visibility enables hiring managers to have faith in automated suggestion verification to existing screening criterion and reinforces trust on AI-supported resume screening solutions.

5.5 Discussion

The main value provided by this research is exemplifying the high-quality assessment of the resume through the use of artificial intelligence-based methods with the possibility of describing complex correlations between skills, experience, education, and job relevance. A holistic approach to screening through the combination of semantic and structural analysis is unlike the traditional rule-based or keyword-driven screening terms which only examine the resume in its entirety. The experimental findings prove that the system is capable of generating repeatable and dependable quality ratings and ranks, and therefore, it can be used in competitive recruitment situations that need a finely detailed distinction between the applicants [22]. Moreover, the recognition of useful resume features, including compatibility in the skills, the level of experience, project-relatedness, and structuring of a document enhance interpretability and can be used to explain AI goals. Such transparency enables the stakeholders to interpret and confirm automated decisions. Although what is currently being expanded to includes only the static resume information and is not yet supplied with the constantly changing professional indication, the results have provided clear evidence that AI-based resume evaluation instruments have enormous potential to enhance the efficiency, accuracy, and quality of decision-making in the current recruitment procedures.

6. Limitations & Future Work

While the proposed MT-SEGA+ framework demonstrates strong performance in resume quality assessment, certain limitations remain that open avenues for future research.

First, the graph-based modeling approach introduces additional computational overhead, particularly during the optimal graph alignment stage. This may affect scalability when processing extremely large datasets, and future work can focus on optimizing graph matching techniques or incorporating approximate alignment methods to improve efficiency.

Second, the effectiveness of the framework is highly dependent on the accuracy of resume parsing and entity extraction. Errors in identifying skills, roles, or temporal information may propagate through the pipeline and impact final predictions. Future research can address this limitation by integrating more robust information extraction techniques, including transformer-based models and large language models (LLMs).

Third, the proposed method primarily relies on textual and structural information extracted from resumes and does not incorporate external evaluation signals such as interview performance, behavioural assessments, or real-world job outcomes. Incorporating such multimodal data sources can significantly enhance the reliability and comprehensiveness of candidate evaluation.

Additionally, the current framework assumes the availability of well-structured job descriptions. However, in real-world scenarios, job descriptions may be inconsistent or incomplete. Future work can explore adaptive models capable of handling noisy or unstructured job requirements.

Finally, although the framework attempts to reduce bias through structured and semantic analysis, potential fairness concerns may still arise due to imbalances in training data or representation of candidate groups. Future research should focus on integrating fairness-aware machine learning techniques and bias mitigation strategies to ensure ethical and inclusive recruitment systems.

7. Conclusion

The results of the experiment conducted in the paper show that machine learning models can be used as a credible pre-screening mechanism to determine the quality of resumes based on information available in the resumes themselves. Although standalone assessment models offer decent and steady performance, they actually become even more effective when implemented in an ensemble-based measure. Compared to single methods of modeling resumes, ensemble models are more effective in having diverse perspectives on the attributes of resumes, which is why the end result of such assessment concerning quality is much more balanced and precise [23]. They are nonlinear which makes them capture non-linear interactions among the features of the resume, e.g. the joint effects of skill relevance, experience depth, and alignment with job demands at a given job. These somewhat interactions can hardly be modeled with linear methods only. Consequently, ensemble-based assessment provides a slightly more detailed and robust analysis of candidate profiles and thus, is especially applicable in large scale recruitment situations where minute differences in resume quality and quality are the key and determinant in effective shortlisting.

According to experimental results, the hybrid ensemble technique provides steady good results in fundamental measures of evaluation, such as accuracy, precision, recall, and F1-score, in an automated resume quality assessment system. The ensemble architecture can minimize the possibility of misclassification and enhance reliability in the categories of high-quality resumes versus less suitable ones due to the use of a number of learning models. The advantages of this group decision-making engine are that the bias of individual models are reduced and makes it more resilient to diverse resume formats and types of content. The additional examination of the model behavior reveals that core resume factors, including applicability of skills to the job position, fit of professional experience, suitability of educational history, legibility of the document, and syntactic compatibility between resume report and work necessities are a determining element on quality consideration. Such factors are close to the established criteria of the recruiter screening, which proves the usefulness of the proposed system. The findings in general indicate that the proposed AI ensemble-based models provide a reliable and recruiter-oriented method of scalable resume quality evaluation.

References

[1] T. Upadhyay and S. Khandelwal, "AI-driven talent acquisition and resume screening: Opportunities and challenges," International Journal of Human Resource Management, vol. 33, no. 9, pp. 1785–1808, 2022.
[2] D. B. Ajmeri and R. V. Kulkarni, "AI-based resume ranking system for recruitment automation," International Journal of Computer Applications, vol. 184, no. 22, pp. 1–7, 2022.
[3] S. Raghavan, "Semantic analysis of resumes and job descriptions using NLP," IEEE Access, vol. 9, pp. 112345–112356, 2021.
[4] A. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, "A survey on bias and fairness in machine learning," ACM Computing Surveys, vol. 54, no. 6, pp. 1–35, 2021.
[5] K. Raghunathan and P. Sinha, "Resume parsing and ranking using intelligent text analytics," International Journal of Information Management, vol. 61, Art. no. 102399, 2021.
[6] R. Feldman and J. Sanger, The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data. Cambridge, U.K.: Cambridge University Press, 2007.
[7] M. L. Anderson, "AI and the future of recruitment and hiring," AI & Society, vol. 37, pp. 1201–1213, 2022.
[8] Y. Goldberg, "A primer on neural network models for natural language processing," Journal of Artificial Intelligence Research, vol. 57, pp. 345–420, 2017.
[9] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD), San Francisco, CA, USA, 2016, pp. 785–794.
[10] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[11] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.
[12] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in Proc. NAACL-HLT, Minneapolis, MN, USA, 2019, pp. 4171–4186.
[13] A. B. Arrieta et al., "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," Information Fusion, vol. 58, pp. 82–115, 2020.
[14] R. Guidotti et al., "A survey of methods for explaining black box models," ACM Computing Surveys, vol. 51, no. 5, pp. 1–42, 2019.
[15] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 4th ed. Hoboken, NJ, USA: Pearson, 2021.
[16] F. Chollet, Deep Learning with Python, 2nd ed. Shelter Island, NY, USA: Manning Publications, 2021.
[17] J. Brownlee, Machine Learning Mastery with Python. Melbourne, Australia: Machine Learning Mastery, 2020.
[18] D. Dua and C. Graff, "UCI Machine Learning Repository," University of California, Irvine, CA, USA, 2019. [Online]. Available: http://archive.ics.uci.edu/ml
[19] S. Barocas, M. Hardt, and A. Narayanan, Fairness and Machine Learning. Cambridge, MA, USA: MIT Press, 2023.
[20] A. L. Beam and I. S. Kohane, "Big data and machine learning in health care," JAMA, vol. 319, no. 13, pp. 1317–1318, 2018.
[21] K. Raghunathan and P. Sinha, "Resume parsing and ranking using intelligent text analytics," International Journal of Information Management, vol. 61, Art. no. 102399, 2021.
[22] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 4th ed. Burlington, MA, USA: Morgan Kaufmann, 2022.
[23] P. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining, 2nd ed. Boston, MA, USA: Pearson, 2019.