electronics

MDPI

Article

Resume2Vec: Transforming Applicant Tracking Systems with Intelligent Resume Embeddings for Precise Candidate Matching

Ravi Varma Kumar Bevara [1,*], Nishith Reddy Mannuru [1], Sai Pranathi Karedla [2], Brady Lund [1], Ting Xiao [2,3,*], Harshitha Pasem [1], Sri Chandra Dronavalli [1] and Siddhanth Rupeshkumar [2]

1 Department of Information Science, University of North Texas, Denton, TX 76205, USA; nishithreddymannuru@my.unt.edu (N.R.M.); brady.lund@unt.edu (B.L.); harshithapasem@my.unt.edu (H.P.); srichandradronavalli@my.unt.edu (S.C.D.)

2 Department of Computer Science and Engineering, University of North Texas, Denton, TX 76205, USA; saipranathikaredla@my.unt.edu (S.P.K.); siddhanthrupeshkumar@my.unt.edu (S.R.)

3 The Anuradha and Vikas Sinha Department of Data Science, University of North Texas, Denton, TX 76205, USA

* Correspondence: ravivarmakumarbevara@my.unt.edu (R.V.K.B.); ting.xiao@unt.edu (T.X.)

Abstract: Conventional Applicant Tracking Systems (ATSs) encounter considerable constraints in accurately aligning resumes with job descriptions (JD), especially in handling unstructured data and intricate qualifications. We provide Resume2Vec, an innovative method that utilizes transformer-based deep learning models, including encoders (BERT, RoBERTa, and DistilBERT) and decoders (GPT, Gemini, and Llama), to create embeddings for resumes and job descriptions, employing cosine similarity for evaluation. Our methodology integrates quantitative analysis via embedding-based evaluation with qualitative human assessment across several professional fields. Experimental findings indicate that Resume2Vec outperformed conventional ATS systems, achieving enhancements of up to 15.85% in Normalized Discounted Cumulative Gain (nDCG) and 15.94% in Ranked Biased Overlap (RBO) scores, especially within the mechanical engineering and health and fitness domains. Although conventional the ATS exhibited slightly superior nDCG scores in operations management and software testing, Resume2Vec consistently displayed a more robust alignment with human preferences across the majority of domains, as indicated by the RBO metrics. This research demonstrates that Resume2Vec is a powerful and scalable method for matching resumes to job descriptions, effectively overcoming the shortcomings of traditional systems, while preserving a high alignment with human evaluation criteria. The results indicate considerable promise for transformer-based methodologies in enhancing recruiting technology, facilitating more efficient and precise candidate selection procedures.

Keywords: resume screening; applicant tracking system (ATS); large language models (LLMs); transformer models; embeddings; BERT; GPT; Llama; job description; recruitment

Academic Editors: Young-Koo Lee, Jinseok Chae and Joonho Kwon

Received: 22 January 2025
Revised: 14 February 2025
Accepted: 17 February 2025
Published: 18 February 2025

Citation: Bevara, R.V.K.; Mannuru, N.R.; Karedla, S.P.; Lund, B.; Xiao, T.; Pasem, H.; Dronavalli, S.C.; Rupeshkumar, S. Resume2Vec: Transforming Applicant Tracking Systems with Intelligent Resume Embeddings for Precise Candidate Matching. *Electronics* **2025**, *14*, 794. https://doi.org/10.3390/electronics14040794

1. Introduction

In today's rapidly evolving job market, organizations face significant challenges in identifying and recruiting top talent within a complex digital landscape. The digital transformation of recruitment has led to an exponential increase in job applications, making traditional, manual resume screening both inefficient and unreliable. With talent pools growing exponentially, companies face the dual demands of maintaining competitive hiring practices and ensuring that the candidates selected possess the precise skills necessary for specialized roles. This challenge is amplified during economic downturns, where efficiency and strategic hiring are essential to ensuring candidates align closely with organizational needs. As organizations seek ways to streamline hiring processes, without sacrificing candidate quality, there is an urgent need for intelligent systems capable of performing nuanced assessments in a time-efficient manner. Traditional recruitment methods, which often rely on manual processes, are struggling to keep pace, as they lack the capacity to evaluate the nuanced qualifications required for today's specialized roles, while processing vast amounts of unstructured data.

To address these challenges, Applicant Tracking Systems (ATSs) have been widely adopted to streamline recruitment processes by automating and organizing candidate selection. While ATS platforms provide a degree of efficiency, their traditional reliance on keyword matching to filter candidates limits their effectiveness, as this method fails to capture context and often overlooks qualified candidates who do not use exact keywords. The keyword-centric model further risks introducing systemic biases, as certain terms may favor specific groups or backgrounds, which can inadvertently undermine diversity and inclusion efforts within organizations. As a result, many ATS solutions miss out on understanding the broader context of candidates' experiences and skills, limiting their ability to accurately assess suitability for specialized roles. Format dependency and difficulty handling varying resume structures further exacerbate these issues, leading to potentially inaccurate assessments of candidate suitability and biases in the selection process [1].

The integration of advanced natural language processing (NLP) techniques, especially transformer-based models, presents a compelling solution to enhance ATS capabilities. These technologies, including encoders (BERT, RoBERTa, and DistilBERT) and decoders (GPT, Gemini, and Llama), enable ATSs to move beyond keyword-based screening by generating neural embeddings that capture the semantic relationships and context within resumes and job descriptions. Embedding techniques allow systems to consider the full scope of a candidate's qualifications and experiences, recognizing similar competencies expressed in varied terms or across different contexts. This shift offers organizations a more holistic view of a candidate's qualifications, supporting recruitment processes that are more precise, inclusive, and scalable, even with large volumes of applications [2]. Unlike traditional methods, embeddings can handle unstructured data and adapt to various resume formats, addressing critical limitations in conventional ATS approaches [3].

This paper introduces Resume2Vec, an innovative framework designed to transform applicant tracking systems using intelligent resume embeddings. The name Resume2Vec, short for 'Resume-to-Vector', reflects its core function of converting textual resume data into vectorized embeddings for improved semantic matching. Resume2Vec leverages transformer models to analyze and match resumes with job descriptions more accurately, allowing organizations to optimize talent-acquisition processes by enhancing matching accuracy, reducing bias, and ensuring operational efficiency. Resume2Vec's embedding-based approach goes beyond the surface-level text by identifying the underlying intent and the relevance of candidates' experiences to the requirements of specific roles. By evaluating the effectiveness of different transformer architectures, such as encoders (BERT, RoBERTa, and DistilBERT) and decoders (GPT, Gemini, and Llama), Resume2Vec demonstrates improvements over conventional ATS in terms of nuanced candidate evaluation and performance scalability, offering a robust and dynamic solution for modern recruitment challenges. Through comprehensive experimentation, our research highlights the advantages of embedding-based approaches, showcasing their consistency and adaptability across large-scale applications, making Resume2Vec a significant advancement in the field of recruitment technology.

To guide our investigation, this study sought to address the following research questions:

- How does Resume2Vec compare to traditional ATSs in terms of ranking accuracy and alignment with human evaluations?
- How can Resume2Vec be effectively integrated into existing recruitment systems to enhance hiring efficiency and reduce bias?

These research questions align with our goal of evaluating Resume2Vec's effectiveness in improving candidate-job matching, ensuring fairer hiring practices, and optimizing recruitment workflows. In the subsequent sections, we present a detailed review of related work, followed by our methodology, experimental results, and discussion of findings.

2. Related Work

Applicant Tracking Systems (ATSs) are utilized by recruiters to oversee job applications. These systems vary in their functionalities, supporting various elements of recruiting, including application tracking, candidate assessment, and engagement during the hiring process [4].

The automation of resume screening and candidate matching has gained critical importance as organizations face an overwhelming volume of job applications. The introduction of various advanced natural language processing (NLP) techniques has demonstrated potential in addressing these recruitment challenges. However, each approach brings distinct contributions and limitations, creating an opportunity for our proposed framework, Resume2Vec, to build on these advancements and address their shortcomings in scalability, contextual accuracy, and bias mitigation.

A competence-level classification model was developed using context-aware transformers to categorize resumes by experience levels for Clinical Research Coordinator roles [5]. This model effectively sorts candidates based on expertise but remains confined to a single, narrowly defined job role, limiting its application across diverse categories. Resume2Vec extends beyond this limitation by leveraging transformers to capture the semantic context within resumes and job descriptions, making it adaptable to a broader range of domains and job roles.

To address recruitment biases, several researchers have explored techniques to 'degender' resumes by removing gender indicators, thereby reducing bias in automated screening [6]. While this approach reduces explicit gender biases, it compromises candidate information by excessively obfuscating resumes. In contrast, Resume2Vec maintains the richness of candidate data by integrating embedding techniques that capture nuanced professional skills, while incorporating bias mitigation strategies without sacrificing contextual depth.

An automated screening solution using Sentence-BERT (S-BERT) was introduced to rank resumes based on semantic similarity to job descriptions [7]. Although S-BERT reduces dependency on keywords, it struggles with large datasets due to the computational overhead. Resume2Vec optimizes transformer architectures to handle extensive datasets efficiently, allowing for scalable deployment across larger organizations.

A resume shortlisting method using DistilBERT and YOLOv5, which focuses on efficiently parsing specific skills from resumes, was proposed in another study [8]. While their model improves resume parsing accuracy, it primarily extracts isolated skills, limiting its capability to assess holistic candidate qualifications. Resume2Vec addresses this by embedding the entire work history, qualifications, and experience of candidates, offering a more comprehensive assessment for accurate job matching.

The ResumeAtlas framework, developed using BERT and Gemini, demonstrated has high accuracy in categorizing resumes across 43 job categories [9]. However, it lacks a direct matching mechanism with job descriptions, reducing its real-world application in recruitment. Resume2Vec bridges this gap by combining resume embeddings with job-specific requirements, ensuring a precise alignment of candidates' competencies with role demands.

The evolution of ATS capabilities has been further illustrated by studies employing SVM and XGBoost for resume screening and ranking based on company requirements [10]. While effective with structured data, their reliance on uniform resume formats limits adaptability with unstructured resumes. Resume2Vec overcomes this constraint by utilizing transformers capable of processing both structured and unstructured resumes, enhancing versatility across industries.

In examining AI's impact on ATSs, researchers have outlined advancements in parsing, ranking, and bias mitigation, but highlighted the limitations of keyword-based approaches [11]. Resume2Vec advances beyond keyword dependency by leveraging semantic embeddings, providing a holistic assessment of candidates' qualifications, and ensuring compatibility with varied resume formats.

A Siamese Sentence-BERT model, conSultantBERT, was introduced to emphasize cross-lingual matching for multilingual data compatibility [12]. Although it addresses language diversity, it is limited by supervised embeddings without industry-wide scalability. Resume2Vec's embedding techniques enable consistent, large-scale performance across various industries and languages, facilitating broad applicability.

In tackling data sparsity, synthetic resumes were generated using a combination of real and synthetic data to improve BERT and feedforward neural network classification accuracy [13]. While this approach primarily benefits training, Resume2Vec's optimized embeddings demonstrate robust performance in actual candidate–job matching, reducing dependency on synthetic data.

Keyword-based ranking using cosine similarity has been shown to be efficient, but it fails to capture the complete context within resumes [14]. Resume2Vec addresses this shortcoming by using embeddings that reflect a candidate's full qualification spectrum, providing a more contextually aware selection process.

For models with sequential dependencies, a Long short-term memory (LSTM)-based resume screening model was proposed, enhancing accuracy but struggling with large datasets due to high computational costs [15]. Resume2Vec uses transformer-based architectures, which handle long sequences more efficiently, making it suitable for processing high volumes of resumes.

A human-in-the-loop system leveraging ESCOXLM-R+ for multilingual resume matching was introduced, combining manual corrections by recruiters with fine-tuned LLMs [16]. While innovative, the dependency on human input introduces inefficiencies and potential biases. Resume2Vec automates the process, leveraging intelligent embeddings that generalize across diverse datasets and requirements, eliminating the need for manual corrections.

A stacked model using K-Nearest Neighbors (kNN), Support Vector Classifier(SVC), and eXtreme Gradient Boosting (XGBoost) demonstrated improved resume classification and ranking performance but relied heavily on feature engineering [17]. In contrast, Resume2Vec employs transformer-based contextualized embeddings that dynamically encode semantic relationships, outperforming models dependent on manual feature extraction.

A classification system based on knowledge-base-assisted matching and conceptual segmentation was also developed, effectively routing resumes to occupational categories [18]. However, its reliance on predefined knowledge bases like O*NET introduced high runtime complexity and error-prone segmentation. Resume2Vec eliminates this dependency by mapping resumes and job descriptions directly into a shared vector space, enabling faster and more precise classification.

CNNs with Glove embeddings have been used to classify resumes into hierarchical categories [19]. While CNNs offer hierarchical representations, they lack the contextual awareness provided by transformer models, leading to suboptimal performance when handling complex or ambiguous resume content. Resume2Vec's transformer-based embeddings encapsulate contextual and sequential nuances, delivering superior classification granularity.

Finally, studies have highlighted the ethical concerns and risks of bias perpetuation in LLMs used for candidate screening, underscoring the need for fairness in AI-driven recruitment [20]. Resume2Vec incorporates fairness-driven embedding techniques, leveraging balanced training datasets and fine-tuning to minimize bias.

Frameworks such as those emphasizing summarization and grading of resumes offer efficiency but lack the depth of semantic comparison between resumes and job descriptions [21]. Resume2Vec, however, uses tailored embeddings for alignment tasks, ensuring comprehensive matching and precise ranking.

These existing studies highlight advancements in automated resume screening and candidate matching but reveal significant limitations in handling unstructured data, scaling for large datasets, and providing unbiased candidate evaluations. Resume2Vec addresses these issues by integrating neural embeddings for nuanced semantic analysis, enabling efficient, large-scale deployment while supporting unbiased and precise candidate selection. This framework represents a meaningful advancement in recruitment technology, offering enhanced context, inclusivity, and scalability essential for meeting modern recruitment demands.

Building on these advancements, this study examines the research hypothesis that transformer-based embeddings (BERT, RoBERTa, DistilBERT, GPT-4.0, Gemini, and LLaMA) in Resume2Vec capture candidate-job relevance more effectively than traditional ATS keyword-based filtering, leading to rankings that better align with human evaluations in resume ranking.

3. Methodology

This study employed a comprehensive mixed-methods approach to analyze the mapping of resumes to job descriptions (JDs). The methodology consisted of two phases: Phase 1 focused on qualitative human evaluations, while Phase 2 assessed the performance of traditional applicant tracking systems (ATS) and the proposed deep learning-based methods. Phase 1 established a foundational understanding of human decision-making in resume evaluations, which was later compared with quantifiable scoring mechanisms in Phase 2.

The overall architecture of the proposed system is depicted in Figure 1. The pipeline begins with the collection of resumes from secondary sources, including Kaggle and job descriptions through web scraping. Human evaluators performed manual assessments to establish a baseline ranking for resumes across various job descriptions. Simultaneously, traditional ATS methods processed the resumes to generate relevance scores.

The enhanced method, powered by transformer-based models such as BERT, RoBERTa, DistilBERT, GPT-4.0, Gemini, and Llama, converts both resumes and job descriptions into embeddings. These embeddings were computed to capture nuanced relationships between resumes and job descriptions. Prior research has demonstrated that leveraging semantic embeddings improves job–candidate matching by capturing contextual relationships beyond explicit keyword matching, resulting in better alignment with recruiter preferences [22]. Our approach builds upon these findings by utilizing multiple transformer architectures to enhance the robustness of automated resume screening. Results from a traditional ATS and the enhanced method were compared through metrics such as Normalized Discounted Cumulative Gain (nDCG) and Ranked Biased Overlap (RBO) to measure the quality of rankings. The architecture also facilitated a comparative analysis to identify the optimal resumes for each job description, as shown in the results section.

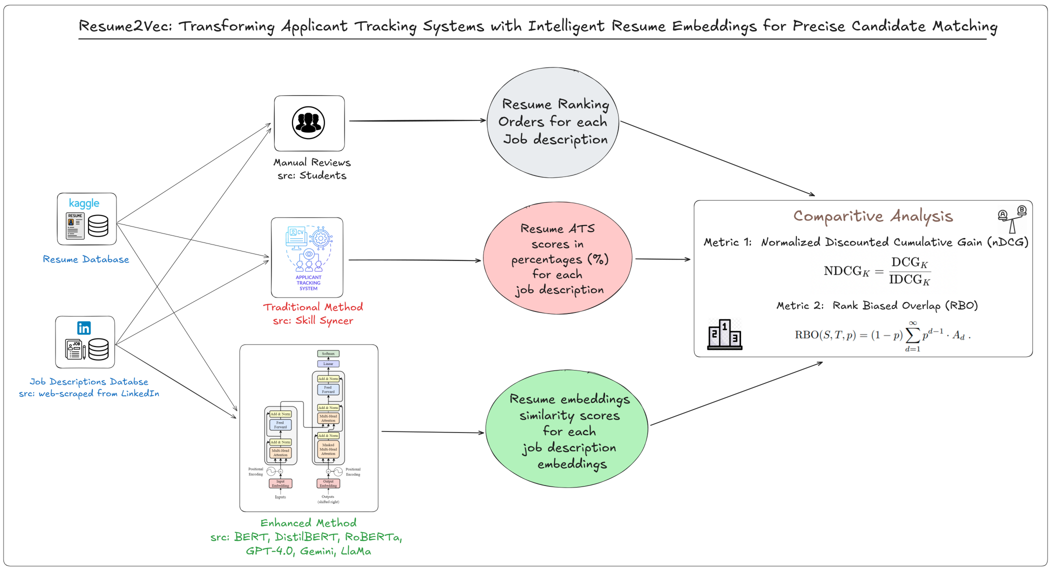

Figure 1. Architecture of the proposed system for resume–JD mapping.

3.1. Data Collection

The dataset for this study was compiled from two primary sources. Job descriptions (JDs) were obtained via web scraping and resumes were acquired from publicly accessible platforms and secondary data sources. The data gathering process was meticulously crafted to guarantee diversity across professional sectors, pertinence to actual recruitment scenarios, and coherence with the study's aims.

Job descriptions were collected over a 5-month period from August to December 2023 and were acquired through automated web scraping methods. A total of 2500 job descriptions were extracted, covering 10 distinct professional domains. LinkedIn, a prominent professional networking platform, was chosen as the major source because of its vast array of job postings across many industries and locations. To extract job descriptions, we employed a combination of Selenium and BeautifulSoup, leveraging their respective strengths in handling dynamic and static web content. Selenium was used to navigate job listing websites that rely on JavaScript to render content, allowing us to interact with web elements such as pagination, dropdown filters, and scrolling mechanisms, to ensure the retrieval of complete listings. Once the desired job postings had been fully loaded in the browser, BeautifulSoup was then utilized to parse the HTML content and extract structured information, including job titles, descriptions, company names, employment types, and locations. This hybrid approach was necessary due to the dynamic nature of modern job boards, where data are not always readily available through static HTML scraping alone. The scraping method focused on particular job categories, like Health and Fitness, Operations Management, and Data Science, while also covering listings for onsite, remote, and hybrid work environments. The essential information gathered for each job posting comprised the job title, firm name, comprehensive job descriptions, employment type, location, posting date, and a link to the original listing. This systematic data collection guaranteed that the information encompassed a wide range of positions and industry-specific needs.

Resumes were obtained from Kaggle's publicly available datasets, last updated in December 2023, yielding a dataset of 15,000 anonymized resumes. These resumes span multiple industries and experience levels. Kaggle was chosen for its varied and publicly accessible datasets, which comprise anonymized resumes featuring comprehensive qualifications, professional experiences, and skill sets. The incorporation of resumes from Kaggle yielded a representative dataset for evaluation, as it exhibited diverse levels of proficiency across several disciplines. To maintain the integrity and usability of the data, stringent cleaning and validation protocols were implemented. Resumes with missing critical fields (e.g., work history, education, or skills) were excluded from the final dataset. Additionally, to mitigate bias, names and gender indicators were removed before analysis. Duplicate posts and incomplete items were eliminated from job descriptions. Particular emphasis was placed on preserving formatting consistency and ensuring that the extracted information appropriately reflected the designated job categories. To ensure quality, entries with missing or incomplete fields were omitted from the dataset for resumes.

This methodology for data collection, including scraping techniques and secondary data platforms, yielded a high-quality dataset that facilitates the efficient assessment of transformer-based approaches for matching resumes to job descriptions.

3.2. Data Preprocessing

To ensure the quality and consistency of the dataset, a multi-step text preprocessing pipeline was implemented. This process was designed to clean, standardize, and prepare the text data for downstream analysis. The key steps in the cleaning process are outlined below:

1. Removal of HTML Tags: HTML elements present in the scraped job descriptions were removed to retain only the core textual content.
2. Text Standardization: Text was converted to lowercase to ensure uniformity across the dataset.
3. Stopword Removal: Common stopwords were removed using the Natural Language Toolkit (NLTK) [23]. Additionally, domain-specific stopwords were included, to improve relevance.
4. Special Characters and Non-ASCII Removal: Non-ASCII characters and special symbols were filtered out to maintain compatibility and readability.
5. Handling Numbers: While irrelevant digits (e.g., phone numbers) were removed, meaningful numeric data such as percentages, monetary values, and years were preserved.
6. Email Address Removal: To maintain privacy and reduce noise, email addresses were stripped from the data.
7. Punctuation Removal: The punctuation was removed using tokenization techniques, to focus on alphanumeric content.

This pipeline ensured that both resumes and job descriptions were effectively cleaned and standardized, addressing common issues such as unstructured formatting, irrelevant content, and noise. The preprocessing steps were essential for preparing the data for subsequent qualitative and quantitative analyses, enabling accurate and reliable evaluation.

3.3. Phase 1: Qualitative Human Evaluation

In the first phase, a structured qualitative ranking methodology was employed to capture the nuanced decision-making processes inherent in resume evaluations across five distinct professional domains. Participants were recruited from a diverse pool of undergraduate, graduate, and professional students at a major, U.S.-based Research University. This selection was motivated by their unique perspectives as both job seekers and future hiring managers, as well as their diverse academic backgrounds. Students in professional programs brought valuable industry experience and current knowledge of emerging field requirements, contributing to a balanced and comprehensive evaluation framework [24].

To ensure domain-specific expertise in the resume evaluations, students were selected based on their academic backgrounds and job search experience within corresponding fields. For instance, graduate students in Mechanical Engineering were chosen to evaluate resumes for Mechanical Engineering positions, as they possess firsthand knowledge of the skills and qualifications required in the job market. Similarly, students from Health and

Fitness, Data Science, Software Testing, and Operations Management were recruited based on their active job search status and familiarity with industry expectations. This approach ensured that the evaluators had a strong understanding of job description requirements and relevant candidate qualifications.

Forty participants were recruited across all domains in the dataset, which originally consisted of ten categories. For the qualitative evaluation, five specific domains, including Health and Fitness, Data Science, Software Testing, Mechanical Engineering, and Operations Management, were selected. These domains were chosen to ensure the evaluation process reflected consistent expertise and technical specificity, while maintaining a manageable scope. Each of the selected domains represented distinct skill requirements, enabling a focused and meaningful analysis of resume evaluation methodologies. Ten anonymized resumes per domain were assessed against two carefully crafted job descriptions, designed to reflect contemporary industry standards and requirements.

The evaluation process was conducted via Qualtrics and utilized a forced-choice ranking system, where participants ranked resumes from 1 to 10. This method emphasized primary criteria, including but not limited to relevance of professional experience, alignment of skills with job requirements, and overall candidate suitability. The forced-choice approach has been demonstrated to mitigate response biases and produce more discriminating results in recruitment contexts [25]. Individual rankings were aggregated using the Borda Count method, which assigns scores based on rank positions and produces a collective ordering. The Borda score $B(c)$ for a candidate c was calculated as

$$B(c) = \sum_{i}(n - r_i(c)) \tag{1}$$

where n is the total number of candidates (10), and $r_i(c)$ represents the rank assigned by participant i. Aggregated rankings were validated through Kendall's coefficient of concordance (W), which ranged from 0.68 to 0.82 across domains, indicating substantial agreement among evaluators. To ensure thoughtful and deliberate evaluations, participants were required to justify their top-five rankings. Optional comment fields captured additional insights, such as the interpretation of technical skills versus practical experience. These qualitative evaluations served as the human benchmark against which the automated methods in Phase 2 were compared.

3.4. Phase 2: Quantitative Evaluation and Model Selection

For this phase, six transformer models were utilized to generate embeddings for resumes and job descriptions (JDs): BERT, RoBERTa, DistilBERT, GPT-4.0, Gemini, and Llama. These models were chosen for their proven capability in capturing nuanced relationships in textual data, with encoder models focusing on extracting contextual embeddings and decoder models enhancing generative capabilities. The inclusion of encoder–decoder models, such as GPT-4.0 and Gemini, ensured that contextual understanding and sequence-level nuances were preserved, as discussed in [26]. The choice of these models was further motivated by their ability to handle both structured and unstructured data, critical for robust resume–JD matching.

Cosine similarity was selected as the primary metric for comparing embeddings, due to its effectiveness in measuring the angular distance between vectors. This metric is particularly suited for high-dimensional embeddings, as it emphasizes the direction of vectors rather than their magnitude, making it robust to variations in embedding scales. Mathematically, cosine similarity between two vectors $\mathbf{u}$ and $\mathbf{v}$ is defined as

$$\text{Cosine Similarity} = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} \tag{2}$$

where $\mathbf{u} \cdot \mathbf{v}$ represents the dot product of the vectors, while $\|\mathbf{u}\|$ and $\|\mathbf{v}\|$ denote their magnitudes. This metric was chosen due to its wide applicability in text similarity tasks and its prior use in bias analysis studies [27].

The preprocessing steps involved cleaning the dataset and converting the textual data into embeddings using the six selected transformer models. Scatter plots were generated for both resumes and JDs, to visualize the embeddings and assess their ability to capture contextual nuances and domain-specific information.

As shown in Figures 2 and 3, the embeddings exhibited clustering patterns based on domain similarities, indicating that the models effectively captured contextual relationships.

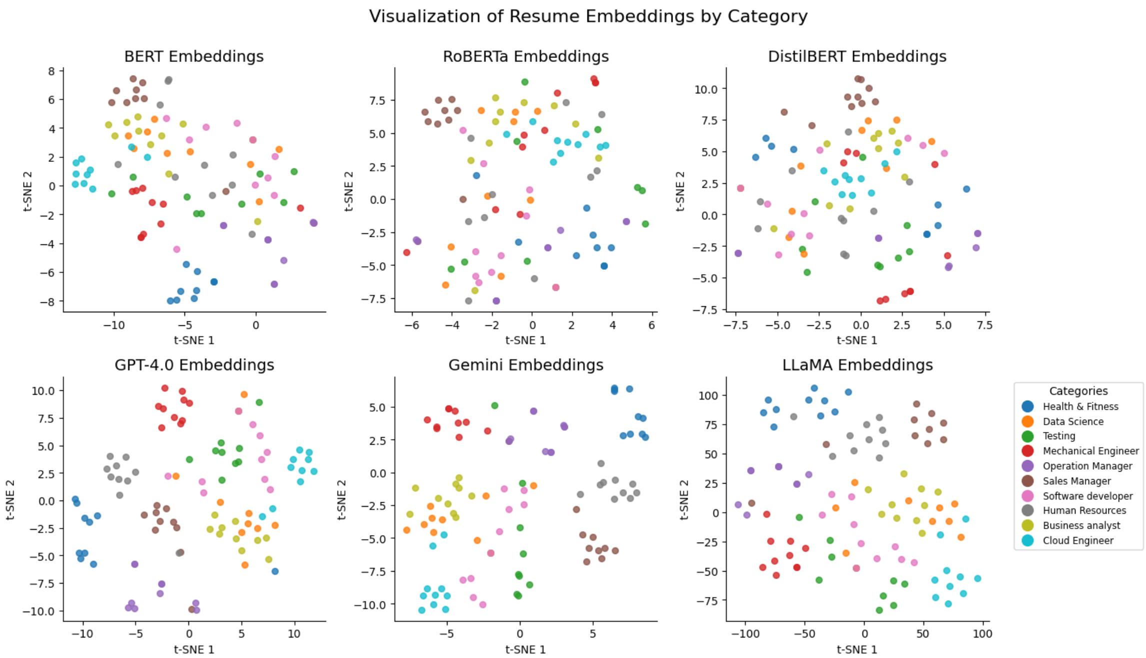

Figure 2. Scatter plot of resume embeddings across domains.

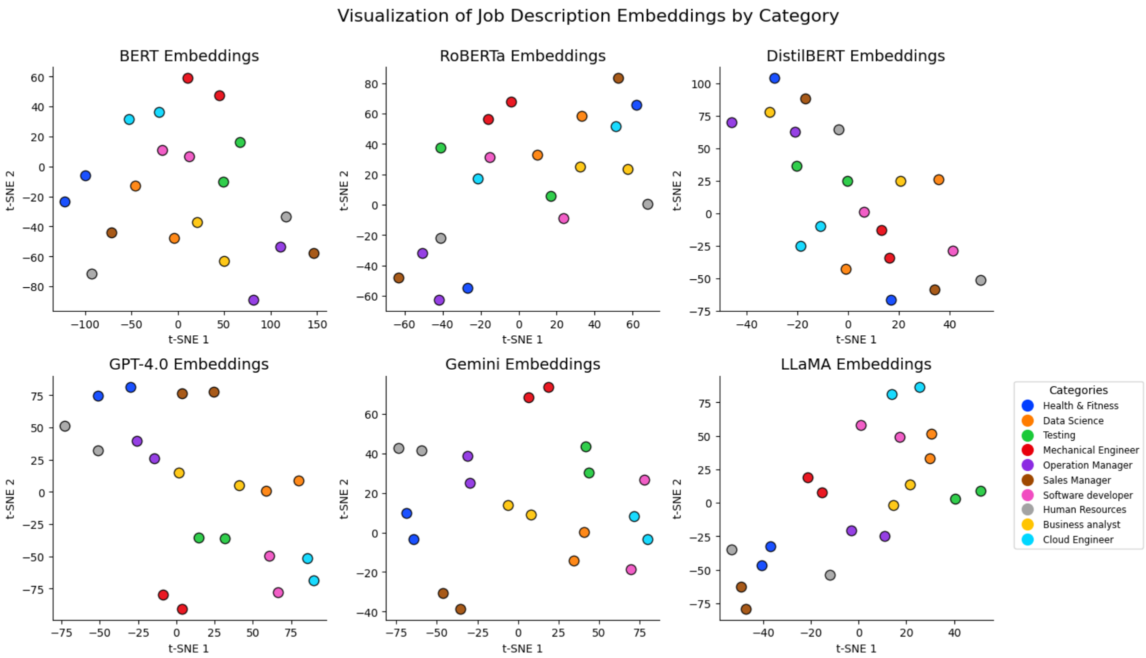

Figure 3. Scatter plot of job description embeddings across domains.

To evaluate the performance of the different transformer models, a downstream classification task was conducted, inspired by the Census2Vec framework [28]. This task involved predicting the category of resumes based on embeddings, with three embedding options considered: resume embeddings, JD embeddings, and stacked embeddings. Resume embeddings were chosen due to their higher data availability for testing. Accuracy was used as the evaluation metric, and the results indicate that while Llama demonstrated the highest accuracy across multiple classification algorithms, achieving the best performance in Random Forest (95.5%), Gradient Boosting (95.5%), and Logistic Regression (95.5%), Gemini performed competitively, with a notable accuracy of 90.5% using SVM. These results suggest that Llama provided the most consistent and optimal performance across the different classification methods, as shown in Table 1.

Table 1. Accuracy comparison of transformer models across classification algorithms. Bolded values indicate the highest accuracy achieved in each column for the corresponding classification algorithm.

Model	SVM	Random Forest	Gradient Boosting	Logistic Regression
BERT	78.5%	94.5%	91.3%	95.2%
RoBERTa	70.5%	93.5%	92.5%	94.5%
DistilBERT	83.0%	94.5%	89.5%	**95.5%**
GPT-4.0	90.5%	91.0%	55.0%	89.5%
Gemini	**91.0%**	94.0%	73.5%	90.5%
Llama	83.5%	**95.5%**	**95.5%**	**95.5%**

To validate the results, data augmentation was performed, and embeddings were analyzed with and without Principal Component Analysis (PCA). As shown in Figures 4 and 5, PCA improved the performance of some models, but Llama consistently delivered the best results across all scenarios.

Using the Llama model, human-evaluated data were also converted to embeddings, and cosine similarity was computed to rank the resumes. The final ranking order was compared across three scenarios: human evaluation, traditional ATS, and the proposed method. Metrics such as Normalized Discounted Cumulative Gain (nDCG) and Ranked Biased Overlap (RBO) were employed to evaluate the performance.

The nDCG metric, calculated as

$$\mathrm{nDCG} = \frac{\mathrm{DCG}}{\mathrm{IDCG}}, \quad \text{where } \mathrm{DCG} = \sum_{i=1}^{p} \frac{2^{\mathrm{rel}_i} - 1}{\log_2(i+1)} \tag{3}$$

measured the quality of ranked lists. Similarly, RBO, a ranking similarity measure, provided further insights into alignment with human rankings. The results, as shown in Tables 2 and 3, demonstrated that the proposed method outperformed the traditional ATS in most domains for this study.

Table 2. Comparison of nDCG scores across domains.

Domain	nDCG (ATS)	nDCG (Proposed)
Data Science	0.88	**0.89**
Health and Fitness	0.83	**0.87**
Mechanical Engineering	0.82	**0.95**
Operations Manager	**0.90**	0.86
Software Testing	**0.96**	0.90

Note: Bold values indicate the best nDCG score for each domain.

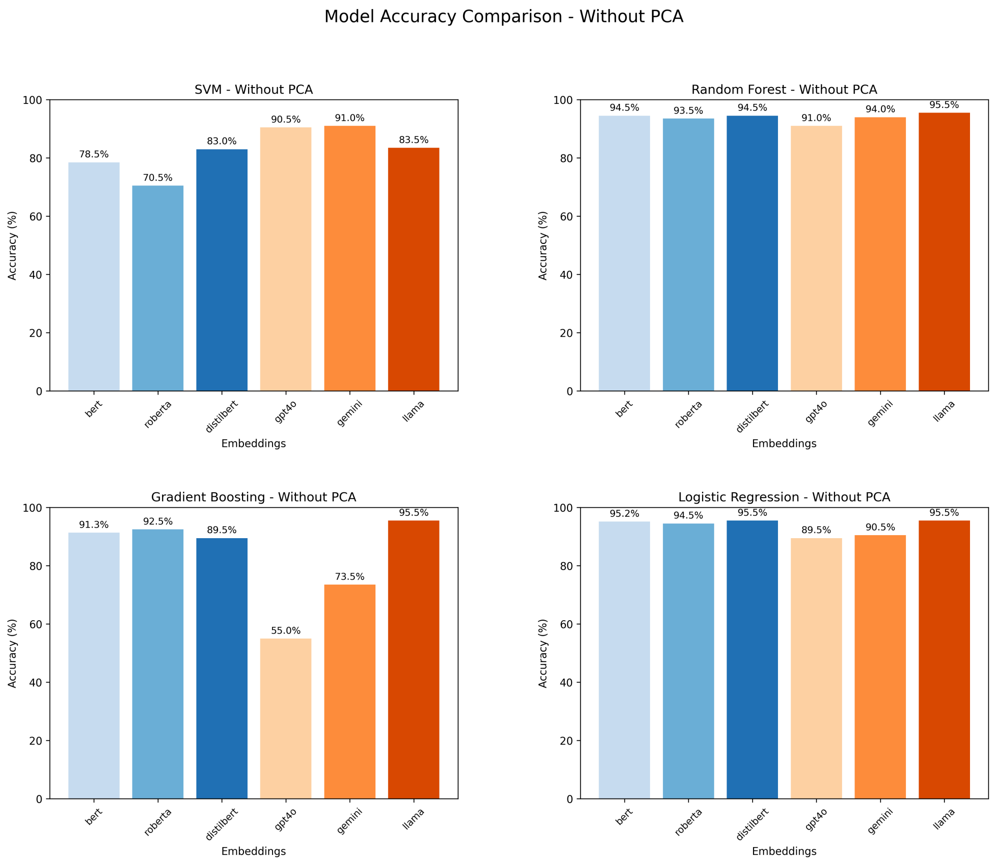

Figure 4. Model accuracy comparison without PCA.

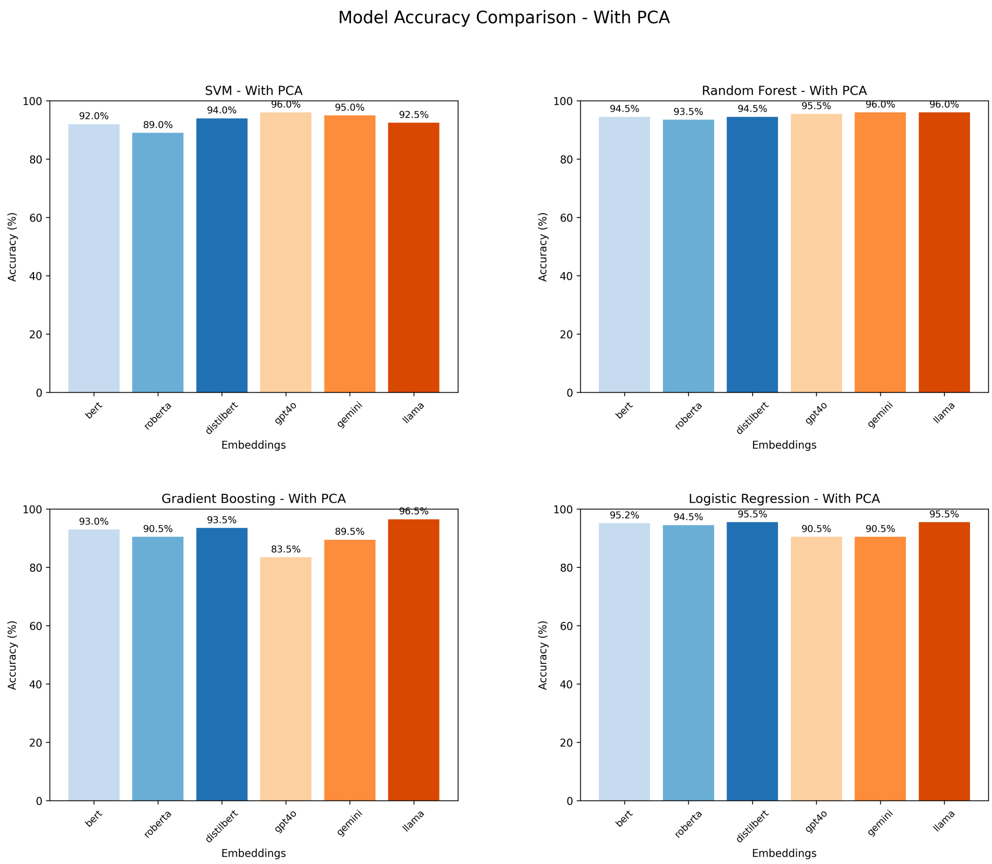

Figure 5. Model accuracy comparison with PCA.

Table 3. Comparison of RBO scores across domains.

Domain	RBO (ATS)	RBO (Proposed)
Data Science	0.84	**0.92**
Health and Fitness	0.69	**0.80**
Mechanical Engineering	0.88	**1.00**
Operations Manager	0.86	**0.97**
Software Testing	**1.00**	0.96

Note: Bold values indicate the best RBO score for each domain.

4. Results and Evaluation

The evaluation of Resume2Vec and the ATS was conducted using two ranking metrics, namely Normalized Discounted Cumulative Gain (nDCG) and Rank-Biased Overlap (RBO), across five distinct job categories: Data Science, Health and Fitness, Mechanical Engineering, Operations Management, and Software Testing. The results provide an insight into the comparative performance of the two approaches and highlight the strengths of Resume2Vec in most categories, as shown in Figure 6.

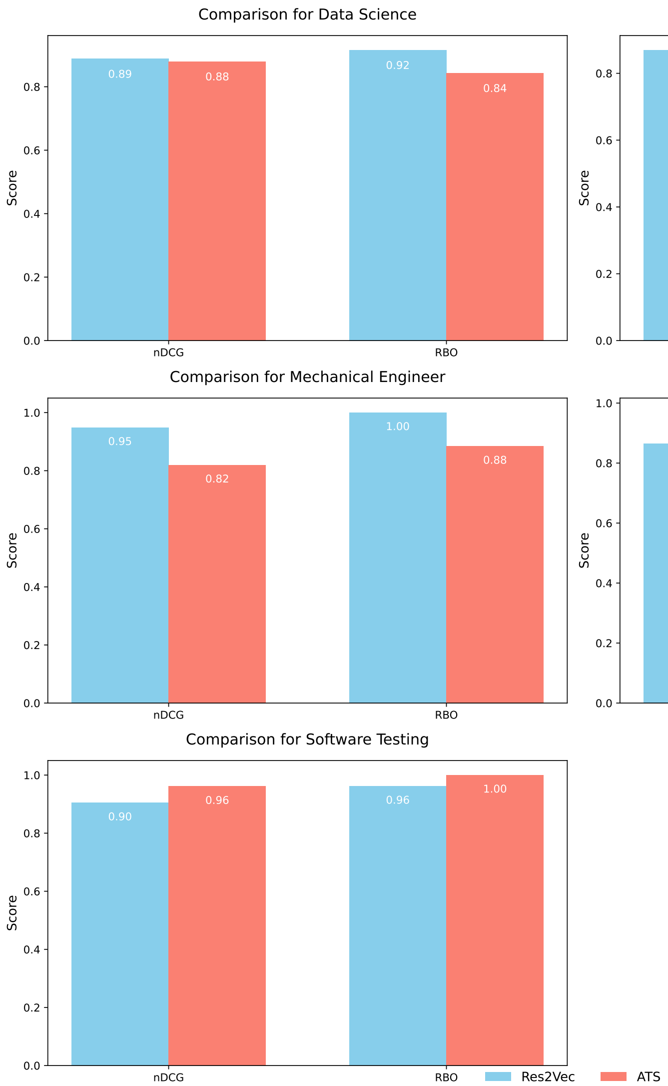

Figure 6. Comparison of Resume2Vec and ATS performance across various metrics (nDCG and RBO) for different job categories.

In the Data Science category, Resume2Vec demonstrated a slight advantage in both metrics. It achieved an nDCG score of 0.89 compared to the ATS's score of 0.88, reflecting a marginal improvement of 1.14%. For the RBO metric, Resume2Vec outperformed the ATS with a score of 0.92, representing a significant 9.52% improvement over the ATS's score of 0.84. These results indicate that Resume2Vec aligned better with human rankings in this domain.

For Health and Fitness, Resume2Vec exhibited a stronger advantage. Its nDCG score was 0.87, outperforming the ATS's score of 0.83 by 4.82%. The improvement was even more pronounced for the RBO metric, where Resume2Vec achieved a score of 0.80, outperforming the ATS's score of 0.69 by 15.94%. These findings highlight Resume2Vec's ability to consistently produce rankings that were more closely aligned with human judgments in this category.

The largest performance gains for Resume2Vec were observed in the Mechanical Engineering category. It achieved an nDCG score of 0.95, which was 15.85% higher than the ATS's score of 0.82. In terms of RBO, Resume2Vec achieved a perfect score of 1.00, reflecting a 13.64% improvement over the ATS's score of 0.88. These results underscore the effectiveness of Resume2Vec in capturing human preferences in highly technical fields like Mechanical Engineering.

In the Operations Management category, the ATS showed a marginal advantage in the nDCG metric, achieving a score of 0.90 compared to Resume2Vec's score of 0.86. This difference represents a slight 4.44% improvement for the ATS. However, Resume2Vec still outperformed the ATS in the RBO metric, achieving a score of 0.97 compared to the ATS's score of 0.86, marking a 12.79% improvement. This marginal edge for the ATS in nDCG could be attributed to the domain's inherent nature, where operational workflows and performance evaluations may favor the ATS's ability to prioritize direct relevance. In contrast, Resume2Vec excelled in preserving the relative ranking order, as evidenced by its consistently higher RBO scores. This suggests that Resume2Vec is better suited for tasks requiring fine-grained rank alignment, while ATSs may provide value in environments where specific job–requirement matching is paramount.

In the Software Testing category, the ATS again demonstrated a slight advantage in nDCG, achieving a score of 0.96 compared to Resume2Vec's score of 0.90, resulting in a 6.25% improvement. Despite this, Resume2Vec achieved a perfect score of 1.00 in the RBO metric, reflecting a 4.17% improvement over the ATS's score of 0.96. The stronger nDCG performance of the ATS in this category might have stemmed from how resumes and job descriptions in software testing often emphasize keyword matches or technical skills explicitly listed in job requirements. The ATS's reliance on keyword relevance likely contributed to this outcome. However, Resume2Vec's superior RBO score indicates that it provided better alignment with overall human preferences, which are not solely dictated by keyword matches but rather by broader contextual factors.

The aforementioned results substantiate our research hypothesis that the transformer-based embeddings in Resume2Vec more successfully capture candidate–job relevance compared to conventional ATS keyword-based filtering. Resume2Vec consistently exhibited excellent performance in rank-based concordance with human judgments, as evidenced by its elevated RBO ratings across all job categories. Although the ATS demonstrated marginal benefits in nDCG for Operations Management and Software Testing, these were context-dependent and did not surpass Resume2Vec's overarching advantages in preserving human-aligned rankings. The notable performance improvements noted in Mechanical Engineering, Health and Fitness, and Data Science further substantiate that embedding-based techniques offer a more comprehensive depiction of candidate–job compatibility than inflexible keyword-based filtering. The findings substantiate the assertion that Resume2Vec's transformer-based embeddings produce rankings that more closely correspond with human assessments, hence confirming the fundamental research hypothesis.

Overall, the results show that Resume2Vec consistently outperformed the ATS in the RBO metric across all categories. This suggests that Resume2Vec is better at maintaining agreement with human preferences in ranking tasks. While Resume2Vec also demonstrated superior nDCG scores in most categories, the ATS showed slight improvements in nDCG in two domains: Operations Management and Software Testing. These instances of the ATS's better performance were likely influenced by domain-specific characteristics, such as the explicitness of technical requirements or the prevalence of structured, keyword-driven evaluation criteria in resumes and job descriptions. However, these advantages are narrow and do not diminish Resume2Vec's broader strengths. Resume2Vec consistently demonstrated its ability to align with more nuanced human judgments, as reflected in its consistently higher RBO scores.

Furthermore, the findings suggest that Resume2Vec is a highly effective approach for ranking tasks, particularly in scenarios where maintaining the relative order of rankings is critical. Even in domains where the ATS shows slight advantages, these were limited to specific metrics and contexts, whereas Resume2Vec offered a more comprehensive and reliable performance across multiple job categories. Its performance across multiple domains underscores its potential for broader adoption in resume-ranking applications, while also pointing to opportunities for further optimization in specific contexts.

5. Discussion

The findings of this study underscore the transformative potential of Resume2Vec in redefining modern recruitment processes by addressing critical limitations of conventional Applicant Tracking Systems (ATS). By employing transformer-based architectures such as BERT, RoBERTa, DistilBERT, GPT-4, Gemini, and Llama, Resume2Vec demonstrates a clear advantage in aligning candidate profiles with job descriptions through semantic embedding techniques. Unlike traditional ATS platforms that rely heavily on keyword matching, Resume2Vec captures nuanced contextual relationships, enabling a more holistic assessment of candidates' qualifications.

Resume2Vec's performance gains were particularly pronounced in technical domains like Mechanical Engineering and Health and Fitness, where it outperformed traditional the ATS by significant margins in metrics including Normalized Discounted Cumulative Gain (nDCG) and Rank-Biased Overlap (RBO). These findings highlight Resume2Vec's ability to process complex and unstructured resumes effectively. Additionally, Resume2Vec's embeddings consistently aligned with human preferences, as indicated by higher RBO scores across all evaluated domains. This alignment with human judgment underscores the framework's capability to mitigate the systemic biases often associated with keyword-based ATS systems. By focusing on neural embeddings, Resume2Vec promotes diversity and inclusion, ensuring that candidates are evaluated based on the full spectrum of their qualifications, rather than narrowly defined terms.

However, the study also identified specific contexts where Resume2Vec demonstrated room for improvement. In domains like operations management and software testing, the traditional ATS systems marginally outperformed Resume2Vec in nDCG scores. This outcome can be attributed to domain-specific characteristics, where job descriptions often rely on explicitly stated keywords or technical skills, making the ATS's keyword matching slightly more effective. This highlights the need for further optimization in embedding techniques to enhance their adaptability in domains where structured, keyword-driven evaluations dominate.

While the use of students as evaluators provided valuable insights, due to their direct exposure to job search processes and familiarity with industry expectations, it presents certain limitations. Unlike HR professionals, students may lack comprehensive hiring experience, particularly in assessing soft skills, long-term career trajectories, and company-specific hiring nuances. Additionally, their evaluations may reflect a job-seeker perspective rather than that of a recruiter, potentially influencing ranking decisions.

To address this, future studies should incorporate HR professionals as a validation group to compare their assessments with student rankings. This would help determine the degree of alignment between the two evaluator groups and refine the model's effectiveness in replicating professional hiring judgments. By integrating HR insights, Resume2Vec could further enhance its alignment with real-world recruitment decision-making, while maintaining its efficiency in candidate screening.

Despite these limitations, the current approach aligns with the broader goal of demonstrating Resume2Vec's ability to mirror human decision-making in resume evaluations. By leveraging domain-specific student evaluators, this study effectively showcases how the model captures nuanced criteria that go beyond keyword matching, supporting its application as an intelligent resume screening tool. The findings indicate that while student evaluations serve as an effective proxy for human assessment, further validation with experienced hiring professionals will strengthen the generalizability and credibility of the model's performance.

From a practical standpoint, Resume2Vec offers significant advantages for real-world recruitment workflows. The system can be deployed as an API service, allowing seamless integration into existing ATS platforms, or as a standalone tool for HR teams to manually upload resumes and job descriptions. For companies that utilize platforms like LinkedIn, Greenhouse, or Workday, Resume2Vec can function as a plug-in for retrieving candidate profiles and ranking them based on semantic similarity. Unlike traditional ATSs that require extensive manual configuration, Resume2Vec operates with pretrained transformer models, enabling rapid deployment with minimal set up effort. By leveraging transformer embeddings, the system reduces recruiter workload, improves candidate–job alignment, and enhances scalability in high-volume hiring scenarios. Additionally, Resume2Vec contributes to fairer hiring practices by mitigating the biases present in keyword-based filtering, ensuring a more equitable selection process.

These practical implications align with the broader evolution of AI-enabled recruitment. Recent studies have emphasized that AI-driven tools have transitioned from a "nice-to-have" feature to a strategic necessity in talent acquisition, due to the increasing importance of human capital in organizational success [29]. AI recruitment tools are now central to digital hiring strategies, as they enable companies to screen large applicant pools efficiently, while reducing the cognitive biases that impact human decision-making. By integrating Resume2Vec into recruitment workflows, organizations can achieve similar improvements in hiring speed, accuracy, and fairness.

Beyond its technical advantages, Resume2Vec has significant managerial implications. Traditional ATSs often filter candidates based on rigid criteria, potentially eliminating strong applicants who use different terminology in their resumes. The ability of Resume2Vec to capture semantic meaning helps HR managers make more data-driven decisions, ensuring that recruitment efforts are aligned with long-term organizational goals rather than arbitrary keyword filters. Moreover, AI-driven recruitment frameworks have been shown to improve candidate experiences, as fairer and more transparent selection mechanisms enhance job seekers' perceptions of employers. This fosters stronger employer branding, making companies more attractive to high-quality candidates.

Additionally, Resume2Vec can play a crucial role in assessing cross-cultural compatibility, a critical factor in modern hiring. Organizations are increasingly recognizing that successful hiring goes beyond matching skills to job descriptions—it involves ensuring that candidates fit within the team's cultural dynamics and communication styles. AI-based tools are now being explored for their potential to evaluate alignment with organizational culture, which can help mitigate turnover and improve team cohesion [30]. Resume2Vec can be adapted to incorporate culture-fit metrics, offering a more holistic approach to hiring that extends beyond hard skills to consider interpersonal and team dynamics.

The scalability of Resume2Vec is another critical advantage. By leveraging transformer-based architectures capable of processing large-scale datasets, Resume2Vec supports efficient recruitment workflows without sacrificing accuracy or alignment with human preferences. Moreover, its ability to handle unstructured data and adapt to diverse resume formats makes it highly versatile across industries. This versatility positions Resume2Vec as a robust tool for organizations navigating competitive talent-acquisition landscapes.

Despite its successes, several areas warrant future exploration. Enhancing the interpretability of embeddings, particularly in sensitive domains, could improve recruiter trust and adoption. Additionally, integrating multimodal data—such as combining resumes with interview transcripts or performance reviews—could further refine the accuracy of candidate evaluations. Addressing computational efficiency in handling large datasets and expanding the framework's capabilities for multilingual and cross-cultural contexts also represent valuable avenues for improvement.

In conclusion, this discussion reaffirms Resume2Vec's potential to revolutionize recruitment by providing an inclusive, scalable, and context-aware solution for matching candidates to job descriptions. While challenges remain in specific contexts, the framework's adaptability and performance gains highlight its promise as a next-generation recruitment tool, paving the way for more equitable and efficient hiring practices.

6. Conclusions

Resume2Vec represents a significant advancement in recruitment technology, addressing the limitations of conventional ATS platforms through the integration of transformer-based models like BERT, RoBERTa, DistilBERT, GPT-4, Gemini, and Llama. By employing neural embeddings to capture semantic relationships between resumes and job descriptions, Resume2Vec demonstrates superior performance in aligning candidate profiles with organizational requirements. Its ability to process unstructured data and adapt to varying resume formats underscores its versatility across diverse professional domains.

The framework's consistent outperformance of traditional ATSs in metrics such as nDCG and RBO highlights its capacity to deliver nuanced, human-aligned evaluations. While domains like Mechanical Engineering and Health and Fitness showcased the strongest gains, areas such as operations management and software testing revealed opportunities for further optimization. These findings reinforce Resume2Vec's role in fostering diversity, reducing bias, and enhancing recruitment efficiency on a large scale.

The empirical results robustly supported our research hypothesis, indicating that the transformer-based embeddings in Resume2Vec surpass conventional ATS keyword-based filtering in assessing candidate–job relevance. In various categories, Resume2Vec demonstrated greater concordance with human assessments, as indicated by its elevated RBO ratings. Although the ATS maintained some advantages in nDCG in certain domains, Resume2Vec's consistently superior rank-based agreement demonstrates that embedding-based methodologies effectively capture more profound semantic connections between resumes and job descriptions. These findings confirm that utilizing transformer-based embeddings improves recruiting procedures by offering a more thorough and human-centric evaluation framework, reinforcing the fundamental premise of this study.

Future research could explore deeper integrations of multimodal data, such as incorporating structured candidate profiles and recruiter feedback alongside resumes, to refine the shortlisting process before candidates reach the interview stage. Since ATSs play a crucial role in pre-screening, integrating additional sources of information could improve decision-making accuracy, while ensuring fairer evaluations. Additionally, Resume2Vec could be extended to assess cross-cultural compatibility, helping organizations determine whether a candidate aligns with the values, communication styles, and team dynamics of the company. This would move beyond purely skill-based matching, offering a more holistic approach to recruitment. Addressing these aspects will further enhance Resume2Vec's applicability in modern hiring environments. By addressing these areas, Resume2Vec can continue to refine its capabilities, ensuring its relevance in an ever-evolving job market. As organizations prioritize inclusive and scalable hiring practices, Resume2Vec offers a transformative solution to meet modern recruitment challenges with precision and adaptability.

Author Contributions: Conceptualization, R.V.K.B.; methodology, R.V.K.B., N.R.M. and S.R.; writing original draft, R.V.K.B., N.R.M. and S.P.K.; writing—review and editing, R.V.K.B., N.R.M., S.P.K., B.L. and T.X; visualization, R.V.K.B. and N.R.M.; data curation, H.P., S.C.D. and S.P.K.; resources, S.P.K.; supervision, B.L. and T.X.; project administration, B.L. and T.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (IRB) of the University of North Texas (protocol code IRB-24-372, approved on 5 August 2024).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data used in this study cannot be shared or made publicly available due to ethical and privacy considerations. As the data collection process was conducted under Institutional Review Board (IRB) approval, access to the dataset is restricted to ensure compliance with participant confidentiality and ethical research guidelines.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Albaroudi, E.; Mansouri, T.; Alameer, A. A Comprehensive Review of AI Techniques for Addressing Algorithmic Bias in Job Hiring. *AI* **2024**, *5*, 383–404. [CrossRef]
2. Strohmeier, S.; Piazza, F. Artificial intelligence techniques in human resource management—A conceptual exploration. In *Intelligent Techniques in Engineering Management*; Intelligent Systems Reference Library; Springer: Cham, Switzerland; 2015; Volume 87, pp. 149–172. [CrossRef]
3. Le Q.; Mikolov, T. Distributed representations of sentences and documents. In Proceedings of the International Conference on Machine Learning, PMLR, Beijing, China, 22–24 June 2014; pp. 1188–1196.
4. AL-Qassem, A.H.; Agha, K.; Vij, M.; Elrehail, H.; Agarwal, R. Leading Talent Management: Empirical investigation on Applicant Tracking System (ATS) on e-Recruitment Performance. In Proceedings of the 2023 International Conference on Business Analytics for Technology and Security (ICBATS), Dubai, United Arab Emirates, 7–8 March 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–5.
5. Li, C.; Fisher, E.; Thomas, R.; Pittard, S.; Hertzberg, V.; Choi, J.D.Competence-level prediction and resume & job description matching using context-aware transformer models. *arXiv* **2020**, arXiv:2011.02998.
6. Parasurama, P.; Sedoc, J. Degendering Resumes for Fair Algorithmic Resume Screening. *arXiv* **2021**, arXiv:2112.08910.
7. Deshmukh, A.; Raut, A. Enhanced Resume Screening for Smart Hiring Using Sentence-Bidirectional Encoder Representations from Transformers (S-BERT). *Int. J. Adv. Comput. Sci. Appl.* **2024**, *15*, 8. [CrossRef]
8. James, V.; Kulkarni, A.; Agarwal, R. Resume Shortlisting and Ranking with Transformers. In Proceedings of the International Conference on Intelligent Systems and Machine Learning, Hyderabad, India, 16–17 December 2022; Springer Nature Switzerland: Cham, Switzerland, 2022; pp. 99–108.

9. Heakl, A.; Mohamed, Y.; Mohamed, N.; Elsharkawy, A.; Zaky, A. ResuméAtlas: Revisiting Resume Classification with Large-Scale Datasets and Large Language Models. *Procedia Comput. Sci.* **2024**, *244*, 158–165. [CrossRef]
10. Vishaline, A.R.; Kumar, R.K.; Sai Pramod, V.V.N.S.; Vignesh, K.V.K.; Sudheesh, P. An ML-based Resume Screening and Ranking System. In Proceedings of the 2024 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT), Karaikal, India, 4–5 July 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 1–6.
11. Minvielle, L. How AI Is Shaping Applicant Tracking Systems. Geekflare. 2025. Available online: https://geekflare.com/guide/ats-and-ai/ (accessed on 15 January 2025).
12. Lavi, D.; Medentsiy, V.; Graus, D. consultantbert: Fine-tuned siamese sentence-bert for matching jobs and job seekers. *arXiv* **2021**, arXiv:2109.06501.
13. Skondras, P.; Zervas, P.; Tzimas, G. Generating Synthetic Resume Data with Large Language Models for Enhanced Job Description Classification. *Future Internet* **2023**, *15*, 363. [CrossRef]
14. Pendhari, H.; Rodricks, S.; Patel, M.; Emmatty, S.; Pereira, A. Resume Screening using Machine Learning. In Proceedings of the 2023 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI), Chennai, India, 21–23 December 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–5.
15. Bharadwaj, S.; Varun, R.; Aditya, P.S.; Nikhil, M.; Babu, G.C. Resume Screening using NLP and LSTM. In Proceedings of the 2022 International Conference on Inventive Computation Technologies (ICICT), Lalitpur, Nepal, 20–22 July 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 238–241.
16. Kavas, H.; Serra-Vidal, M.; Wanner, L. Using Large Language Models and Recruiter Expertise for Optimized Multilingual Job Offer–Applicant CV Matching. In Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, Jeju, Republic of Korea, 3–9 August 2024; pp. 8696–8699.
17. Ransing, R.; Mohan, A.; Emberi, N.B.; Mahavarkar, K. Screening and Ranking Resumes using Stacked Model. In Proceedings of the 2021 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT), Mysuru, India, 10–11 December 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 643–648.
18. Zaroor, A.; Maree, M.; Sabha, M. JRC: A job post and resume classification system for online recruitment. In Proceedings of the 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI), Boston, MA, USA, 6–8 November 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 780–787.
19. Nasser, S.; Sreejith, C.; Irshad, M. Convolutional Neural Network with Word Embedding Based Approach for Resume Classification. In Proceedings of the 2018 International Conference on Emerging Trends and Innovations In Engineering And Technological Research (ICETIETR), Ernakulam, India, 11–13 July 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6.
20. Koh, N.H.; Plata, J.; Chai, J. BAD: BiAs Detection for Large Language Models in the context of candidate screening. *arXiv* **2023**, arXiv:2305.10407.
21. Gan, C.; Zhang, Q.; Mori, T. Application of LLM agents in recruitment: A novel framework for resume screening. *arXiv* **2024**, arXiv:2401.08315. [CrossRef]
22. Pombo, L.M. Landing on the Right Job: A Machine Learning Approach to Match Candidates with Jobs Applying Semantic Embeddings. Master's Thesis, Universidade NOVA de Lisboa, Lisboa, Portugal, 2019.
23. Bird, S.; Klein, E.; Loper, E. *Natural Language Processing with Python*; O'Reilly Media, Inc.: Newton, MA, USA, 2009.
24. Willis, K.L.; Holmes, B.; Burwell, N. Improving Graduate Student Recruitment, Retention, and Professional Development During COVID-19. *Am. J. Educ. Res.* **2022**, *10*, 81–84. [CrossRef]
25. Hung, S.-P.; Huang, H.-Y. Forced-Choice Ranking Models for Raters' Ranking Data. *J. Educ. Behav. Stat.* **2022**, *47*, 603–634. [CrossRef]
26. Shahriar, S.; Lund, B.D.; Mannuru, N.R.; Arshad, M.A.; Hayawi, K.; Bevara, R.V.K.; Mannuru, A.; Batool, L. Putting GPT-4.0 to the Sword: A Comprehensive Evaluation of Language, Vision, Speech, and Multimodal Proficiency. *Appl. Sci.* **2024**, *14*, 7782. [CrossRef]
27. Bevara, R.V.K.; Mannuru, N.R.; Karedla, S.P.; Xiao, T. Scaling Implicit Bias Analysis Across Transformer-Based Language Models Through Embedding Association Test and Prompt Engineering. *Appl. Sci.* **2024**, *14*, 3483. [CrossRef]
28. Bevara, R.V.K.; Wagenvoord, I.; Hosseini, F.; Sharma, H.; Nunna, V.; Xiao, T. *Census2Vec: Enhancing Socioeconomic Predictive Models with Geo-Embedded Data*; Intelligent Systems Conference; Springer Nature Switzerland: Cham, Switzerland, 2024; pp. 626–640.
29. Black, J.S.; van Esch, P. AI-enabled recruiting: What is it and how should a manager use it? *Bus. Horizons* **2020**, *63*, 215–226. [CrossRef]
30. Dospinescu, O.; Dospinescu, N. Workaholism in IT: An analysis of the influence factors. *Adm. Sci.* **2020**, *10*, 96. [CrossRef]