IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 14, No. 4, August 2025, pp. 3404~3411
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i4.pp3404-3411

Comparing bidirectional encoder representations from transformers and sentence-BERT for automated resume screening

Asmita Deshmukh, Anjali Raut Dahake
Department of Computer Science and Engineering, HVPM's College of Engineering and Technology, Sant Gadge Baba Amravati University, Amravati, India

Article Info

Article history:

Received Apr 16, 2024
Revised Jun 11, 2025
Accepted Jul 10, 2025

Keywords:

Bidirectional encoder representations from transformers
Cosine similarity
Resume screening
Sentence bidirectional encoder representations from transformers
Talent acquisition

ABSTRACT

In today's digital age, organizations face the daunting challenge of efficiently screening an overwhelming number of resumes for job openings. This study investigates the potential of two state-of-the-art natural language processing models, bidirectional encoder representations from transformers (BERT) and sentence-BERT (S-BERT), to automate and optimize the resume screening process. The research addresses the need for accurate, efficient, and unbiased candidate evaluation by leveraging the power of these transformer-based language models. A comprehensive comparison between BERT and S-BERT is performed, evaluating their performance across multiple metrics, including accuracy, screening time, correlation with job descriptions, and ranking quality. The findings reveal that S-BERT outperforms BERT, achieving higher accuracy (90% vs. 86%), faster screening time (0.061 seconds vs. 1 second per resume), and stronger correlation with job descriptions (0.383855 vs. 0.1249). S-BERT though has a smaller vector size of 384 enables capturing richer semantic information compared to BERT's vector size of 768, contributing to its superior performance. The study provides insights into the strengths and limitations of each model, offering valuable guidance for organizations seeking to streamline their talent acquisition processes and enhance candidate selection through automated systems.

Corresponding Author:

Asmita Deshmukh
Department of Computer Science and Engineering, HVPM's College of Engineering and Technology
Sant Gadge Baba Amravati University
Amravati, Maharashtra, India
Email: asmitadeshmukh7@gmail.com

1. INTRODUCTION

The rapid proliferation of digital technologies has transformed the landscape of talent acquisition [1], presenting both opportunities and challenges for organizations. As job postings receive an influx of applications [2], the traditional manual approach to resume screening becomes increasingly inefficient and susceptible to biases [3]. This predicament has ignited a growing demand for automated solutions that can streamline the resume screening process [4], ensuring fair candidate evaluation [5].

The natural language processing (NLP) [6], a branch of artificial intelligence, has emerged as a powerful tool in this domain [7], offering language models capable of understanding and analyzing textual data with remarkable accuracy [8]. Among the most promising NLP models are bidirectional encoder representations from transformers (BERT) [9] and sentence-BERT (S-BERT) [10], which have demonstrated impressive performance [11] in various text analysis tasks [12]. This study aims to critically evaluate the potential of BERT [13] and S-BERT for automated resume screening [14], addressing the what, why, and how of this research endeavor [15]. Specifically, we investigate the ability of these models to extract relevant information from resumes [16], assess their suitability for job descriptions [17], and rank candidates based on their qualifications [18]. By conducting a comprehensive comparison between BERT and S-BERT [19], we seek to identify the model [20] that excels in accuracy [21], efficiency [22], and ranking quality, ultimately providing organizations with a robust solution for optimizing their talent acquisition processes [23].

The motivation behind this research stems from the pressing need to alleviate the time-consuming and error-prone nature of manual resume screening, while mitigating potential biases that may arise from human subjectivity [24]. By leveraging the power of advanced NLP models [25], organizations can streamline their hiring processes, reduce time-to-hire, and enhance candidate selection, thereby gaining a competitive advantage in today's dynamic job market. To achieve our research objectives, we employed a rigorous experimental approach, utilizing a large dataset of resumes and job descriptions. The methodology involved preprocessing the data, generating embeddings using BERT and S-BERT, and computing cosine similarities between resume and job description embeddings. Extensive quantitative and qualitative analyses were conducted to evaluate the model's performance, considering metrics such as accuracy, screening time, correlation with job descriptions, and ranking quality.

The subsequent sections of this paper provide a detailed overview of the research methodology, including data collection, preprocessing techniques, and experimental procedures. The results and discussion section presents a comprehensive analysis of the findings, highlighting the strengths and limitations of each model, and drawing insights from a broader scientific context. Additionally, potential limitations of the study are addressed, and implications for future research are explored. Finally, the conclusion summarizes the key contributions and underscores the significance of this research in advancing the field of automated resume screening and talent acquisition.

2. METHOD

A comprehensive comparison between BERT and S-BERT for automated resume screening was conducted using a robust and replicable methodology. This section outlines the experimental procedure, including data collection, preprocessing steps, and the implementation details of both models.

2.1. Data collection and preprocessing

The study utilized a diverse dataset comprising of 223 resumes and 7 job descriptions, obtained from various online platforms such as LinkedIn, Google Forms, and Freshers World. The resumes were initially converted to PDF format and subsequently to Excel format to ensure compatibility with the Python programming language used for data processing. Preprocessing steps were performed to prepare the data for analysis. For the S-BERT model, sentences containing sets of 10 words were extracted from the resumes to generate embeddings. Conversely, the BERT model underwent lemmatization and stemming as preprocessing steps, followed by the removal of stop words and repetitions. The keywords extracted from the resumes were then utilized for embedding generation with BERT. This process was replicated for the job descriptions, ensuring consistency in the data representation across both models.

2.2. Experimental procedure

The experimental procedure involved several distinct stages, each designed to evaluate the performance of BERT and S-BERT in the context of automated resume screening.

– Stage 1 (keyword extraction and embedding generation): in the first stage, the system extracted sentences and top keywords from the resumes. For S-BERT, sentences containing sets of 10 words were extracted, while for BERT, top keywords were identified after removing common punctuations and stop words. The extracted data was organized into a panda DataFrame and exported to an Excel file.
– Stage 2 (BERT analysis): in the second stage, the BERT model and a pre-trained tokenizer were utilized to compute the cosine similarity between a job description (search query) and the top keywords extracted from the resumes. This stage involved integrating various libraries, including transformers for BERT, openpyxl for Excel handling, NumPy for numerical operations, and torch for interaction with PyTorch. The cosine similarity between the query embeddings and each resume was calculated using the sklearn function.
– Stage 3 (S-BERT analysis): the third stage employed the MiniLM model from the Sentence Transformers library to compute the cosine similarity between the search query and sentences stored in an Excel file. This involved loading the MiniLM model, accessing the Excel file with openpyxl, and calculating the cosine similarity between the query embeddings and each resume.

– Stage 4 (ranking and evaluation): after computing cosine similarities for both BERT and S-BERT, the resumes were ranked based on their similarity scores. The model's0 performance was evaluated using metrics such as accuracy, screening time per resume, and correlation with job descriptions. Additionally, three independent HR managers cross-verified the results to ensure the accuracy of the rankings.

The flow diagram presented in Figure 1 delineates the step-by-step process of conducting a comparative analysis between BERT and S-BERT for automated resume screening. Initially, a pool of over 223 resumes was collected from various platforms like LinkedIn, Google Forms, and Freshers World, often in differing formats. These resumes undergo conversion to PDF and subsequently to Excel format for compatibility with Python. S-BERT is employed to extract sentences containing sets of 10 words from the resumes for embedding generation. Conversely, BERT undergoes lemmatization and stemming as preprocessing steps, followed by the removal of stop words and repetitions. The keywords extracted from the resumes are then utilized for embedding generation with BERT. This process is replicated for the job description, and embedding vectors from both BERT and S-BERT are employed to compute cosine similarity with the job description, thereby facilitating resume ranking. To ensure the accuracy of the rankings, three independent HR managers cross-verify the results. This comprehensive approach underscores the system's reliability and efficacy in automating the resume screening process.

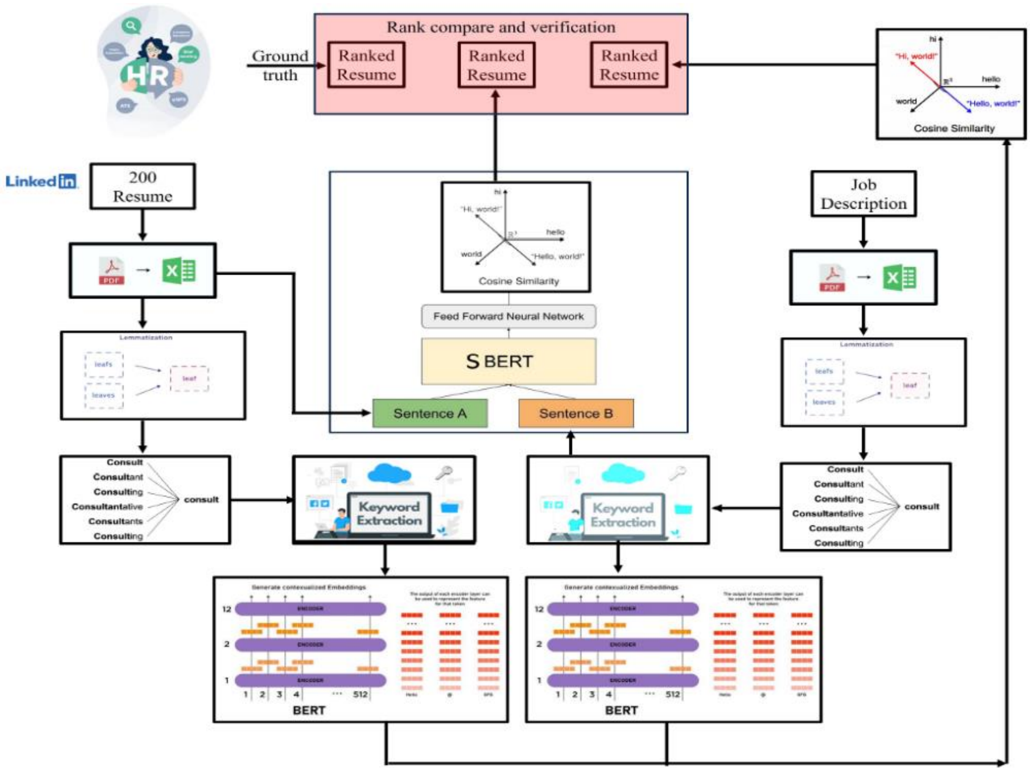


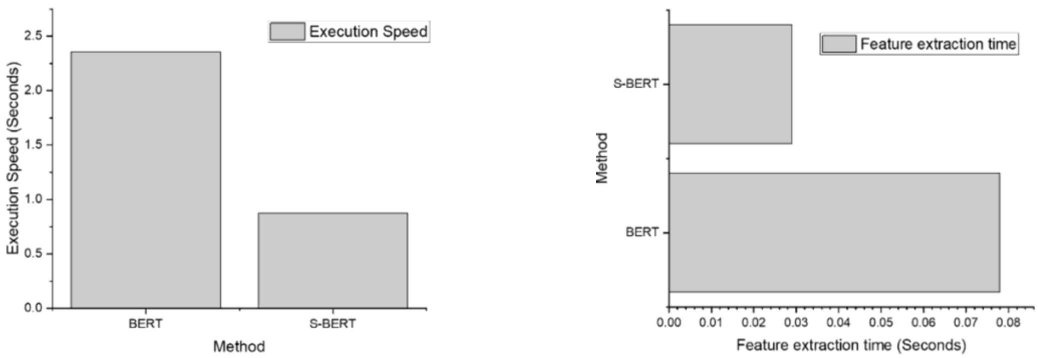
Figure 1. Flow diagram to conduct comparative study of BERT and S-BERT for automated resume screening

3. RESULTS AND DISCUSSION

The comparison between BERT and S-BERT for automated resume screening revealed several key findings across multiple evaluation metrics. Initially, both models analyzed an identical dataset comprising 223 resumes and 7 job descriptions to ensure a fair comparison. A notable difference lies in the feature vector size employed by each model. While BERT produced embeddings of size 768, S-BERT generated more compact vectors of size 384. This divergence suggests that S-BERT may capture more detailed and contextual information from the resumes, potentially facilitating better matches between candidates and job requirements. Furthermore, our study found that S-BERT exhibited superior efficiency, with a screening time of 0.061 seconds per resume, compared to BERT's longer screening time of 1 second per resume. This discrepancy in processing speed underscores S-BERT's potential to enhance the overall productivity of talent acquisition processes.

Crucially, S-BERT achieved a higher accuracy rate of 90% in shortlisting resumes, outperforming BERT's accuracy rate of 86%. This finding implies that S-BERT may possess a better ability to discern relevant information and accurately match candidates to job descriptions, thereby improving the quality of candidate selection. Moreover, our analysis of similarity scores revealed a substantial gap between the two models. While BERT yielded a similarity score of 0.3, S-BERT exhibited a significantly higher score of 0.599. This result suggests that S-BERT excels in capturing semantic relationships and contextual nuances, leading to more precise matches between resumes and job descriptions. However, it is important to acknowledge the limitations of our study. While our dataset comprised a diverse range of resumes and job descriptions, further research with larger and more varied datasets is necessary to validate and generalize our findings across different industries and job roles.

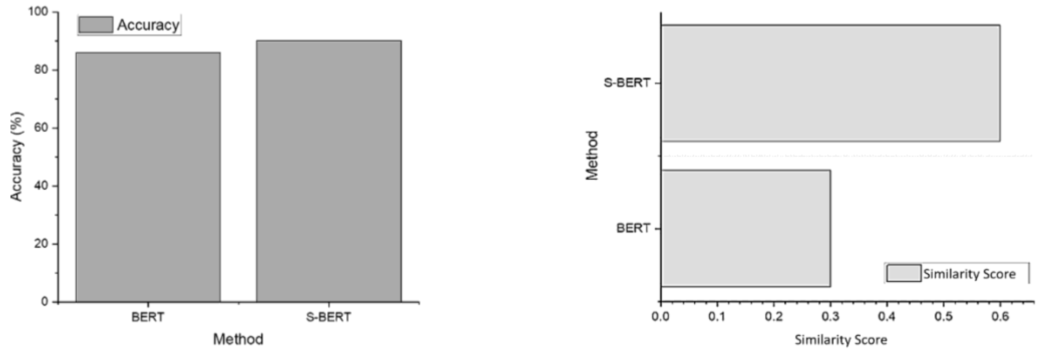
Figure 2 offers a detailed comparison between BERT and S-BERT in terms of execution speed in seconds (Figure 2(a)) and feature extraction speed (Figure 2(b)). BERT exhibits an execution speed of 2.362 seconds, significantly slower than S-BERT's 0.87 seconds, demonstrating the latter's superior efficiency. Moreover, in feature extraction speed per resume, BERT requires about 0.078 seconds, while S-BERT showcases an impressive 0.029 seconds. This substantial difference underscores S-BERT's enhanced computational capabilities, making it a more suitable choice for tasks requiring swift and accurate execution and feature extraction.



(a) (b)

Figure 2. Comparison of BERT and S-BERT with respect to (a) execution speed in seconds and (b) feature extraction speed

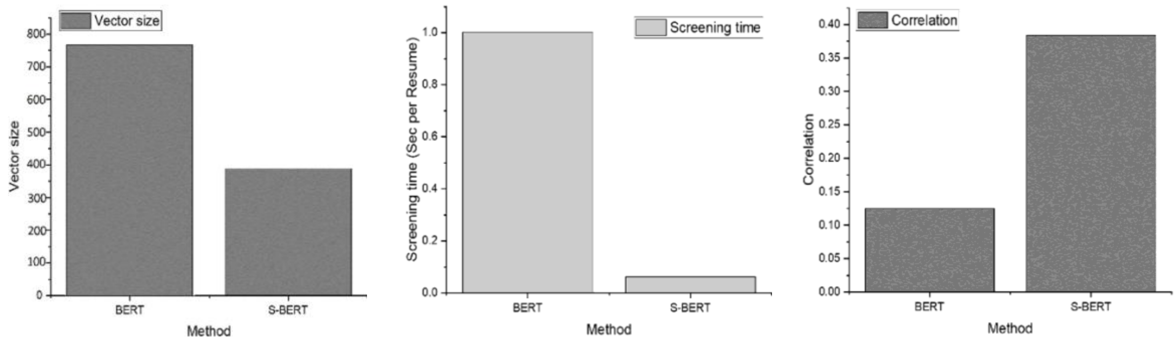
Figure 3 depicts bar charts comparing the performance of BERT and S-BERT across two key metrics of accuracy (Figure 3(a)) and similarity score (Figure 3(b)). The accuracy metric, displayed on the vertical axis labeled "Accuracy (%)", evaluates the effectiveness of the algorithms in shortlisting resumes compared to HR manager results. A higher value signifies superior performance. S-BERT achieves a higher accuracy (90%) compared to BERT (86%), as illustrated in the chart. The similarity score, measured on the vertical axis labeled "Similarity Score", assesses the cosine similarity between embeddings of job descriptions and resume embeddings. Again, a higher value indicates better performance. The chart demonstrates that S-BERT achieves a higher similarity score (0.599) compared to BERT (0.3). Overall, the findings suggest that S-BERT outperforms BERT in terms of both accuracy and similarity score.



(a) (b)

Figure 3. Comparison of BERT and S-BERT with respect to (a) accuracy and (b) similarity score

Figure 4 compares BERT and S-BERT across three key metrics of vector size (Figure 4(a)), screening time per resume (Figure 4(b)) and correlation with job description (Figure 4(c)). The vector size, shown on the leftmost vertical axis labeled “Vector size”, indicates the compactness of the representation, with lower values suggesting greater compactness. The comparison reveals that while BERT requires a vector size of 768, S-BERT requires 384 bytes of vector to generate embeddings at the final output. Screening time, measured on the center vertical axis labeled “Screening time (sec per resume)”, represents processing speed, where lower values signify faster processing.



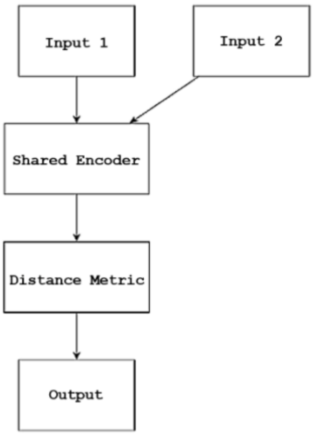
(a) (b) (c)

Figure 4. Comparison of BERT and S-BERT with respect to (a) vector size, (b) screening time in seconds per resume, and (c) correlation

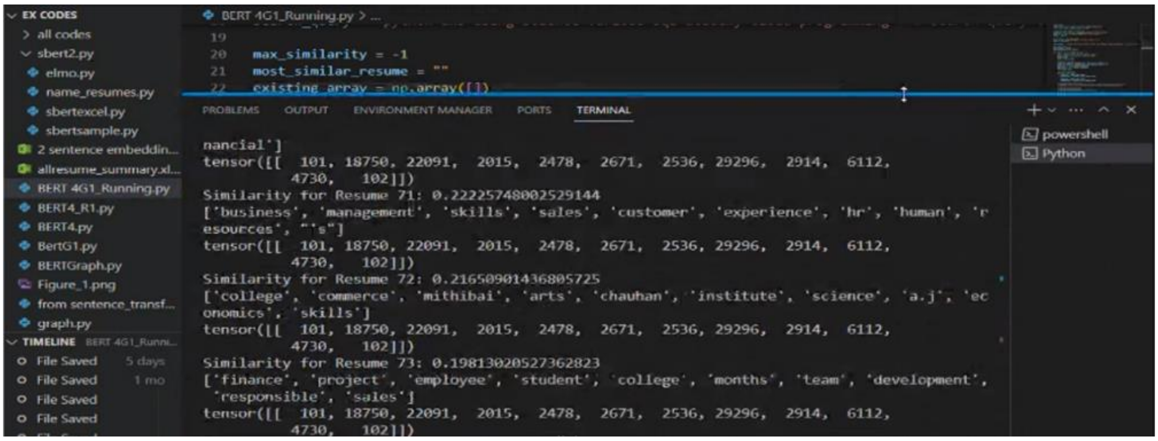
The chart demonstrates that S-BERT (0.061 seconds) significantly outperforms BERT (1 second) in terms of screening resumes. Correlation with job description, depicted on the rightmost vertical axis labeled “Correlation”, reflects the strength of correlation, with higher values indicating a stronger correlation. The analysis illustrates that S-BERT (0.383855) exhibits a stronger correlation with job descriptions compared to BERT (0.1249). In summary, the figure suggests that S-BERT is more efficient than BERT for applications like resume screening, where processing speed is crucial. However, BERT may be preferred if a more compact vector size is necessary, such as for execution on a mobile device locally. Despite BERT’s larger vector size of 768 compared to S-BERT’s 384, S-BERT offers faster screening times and demonstrates a stronger correlation with job descriptions.

Figure 5 illustrates the resume matching process using both BERT and S-BERT algorithms. In Figure 5(a), a schematic overview of the resume matching process is presented, where both the resume and job description are utilized for embedding calculation. These embeddings are then compared using cosine similarity, and the resumes are ranked accordingly based on their similarity scores. In Figure 5(b), a snapshot of the output screen during the execution of BERT shows the extraction and processing of significant keywords. However, the processing speed with the BERT algorithm appears to be slower. In contrast, Figure 5(c) showcases a snapshot of the output screen after the completion of S-BERT execution.

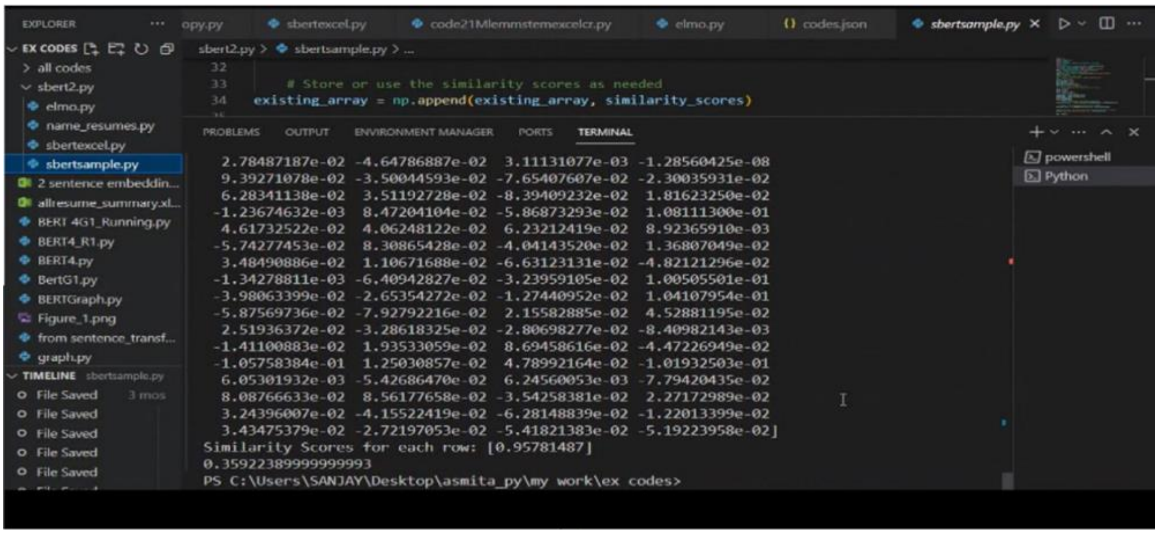
Additionally, our evaluation focused primarily on quantitative metrics such as accuracy, screening time, and similarity scores. Future studies could explore qualitative aspects, such as the interpretability and explainability of the models’ decision-making processes, to ensure fairness and transparency in the recruitment process. Nonetheless, our findings hold significant implications for recruiters and HR professionals. By leveraging S-BERT’s strengths, organizations can streamline their hiring processes, leading to enhanced candidate selection, reduced time-to-hire, and improved employee satisfaction and retention. Furthermore, the potential for more accurate matches between candidates and job descriptions can contribute to a better alignment of skills and organizational needs, ultimately driving productivity and organizational success. While our study provides valuable insights into the comparative performance of BERT and S-BERT for resume screening, it has limitations in terms of the dataset size and scope. Future research could evaluate these models on larger, more diverse datasets across various fields and domains. Additionally, exploring ensemble techniques combining BERT and S-BERT or integrating other state-of-the-art language models could further enhance accuracy. The implications are significant, as implementing automated resume screening based on S-BERT can streamline hiring, reduce bias, and promote workforce diversity. However, addressing interpretability, domain-specific terminology, and end-to-end system integration remains crucial for practical deployment.



(a)



(b)



(c)

Figure 5. Flow chart with output of (a) generalized approach for resume matching, (b) output screen capture while BERT is running, and (c) output screen capture after S-BERT execution completed

4. CONCLUSION

In today's talent acquisition landscape, where organizations face the daunting task of efficiently screening numerous resumes, our study demonstrates the superiority of the S-BERT model over BERT for automated resume screening. S-BERT excels in accuracy, efficiency, and contextual understanding, as shown by its higher similarity scores. S-BERT's inspite of having smaller feature vector size allows for more detailed resume analysis, and its faster screening time enhances recruitment productivity. Its higher accuracy and superior similarity scores enable better matching of candidates with job descriptions, leading to more precise candidate selection. While our findings are promising, further research is needed to validate and expand them across diverse datasets and industries. Future studies could also explore qualitative aspects such as model interpretability to ensure fairness in the recruitment process. Our research significantly advances talent acquisition by demonstrating the potential of advanced models like S-BERT to improve hiring efficiency. Utilizing S-BERT can enhance candidate selection, employee satisfaction, and retention, ultimately driving organizational success.

ACKNOWLEDGMENTS
Authors would like to thank colleagues at HVPM's College of Engineering and Technology.

FUNDING INFORMATION
Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT
This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Asmita Deshmukh	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			
Anjali Raut Dahake		✓				✓				✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT
Authors state no conflict of interest.

DATA AVAILABILITY
The authors confirm that the data that support the findings of this study are available from the corresponding author, [AD], upon reasonable request.

REFERENCES
[1] S. Zu and X. Wang, "Resume information extraction with a novel text block segmentation algorithm," *International Journal on Natural Language Computing*, vol. 8, no. 5, pp. 29–48, Oct. 2019, doi: 10.5121/ijnlc.2019.8503.
[2] W. Lakshan, K. Prabuddhi, E. Bandara, and R. R. P. De Zoysa, "Revolutionizing the hiring process with automated evaluation and behavioral analysis-intellihire," *International Research Journal of Innovations in Engineering and Technology*, vol. 7, no. 11, pp. 307–314, 2023, doi: 10.47001/IRJIET/2023.711042.
[3] B. Wang and C.-C. J. Kuo, "SBERT-WK: a sentence embedding method by dissecting BERT-based word models," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 2146–2157, 2020, doi: 10.1109/TASLP.2020.3008390.
[4] N. Thakur, N. Reimers, J. Daxenberger, and I. Gurevych, "Augmented SBERT: data augmentation method for improving bi-encoders for pairwise sentence scoring tasks," in *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2021, pp. 296–310, doi: 10.18653/v1/2021.naacl-main.28.
[5] V. Yadav and S. Bethard, "A survey on recent advances in named entity recognition from deep learning models," *Computation and Language*, vol. 7, no. 1, pp. 41525–41550, 2019.
[6] J. Seo, S. Lee, L. Liu, and W. Choi, "TA-SBERT: token attention sentence-BERT for improving sentence representation," *IEEE Access*, vol. 10, pp. 39119–39128, 2022, doi: 10.1109/ACCESS.2022.3164769.
[7] Y. Santander-Cruz, S. Salazar-Colores, W. J. Paredes-García, H. Guendulain-Arenas, and S. Tovar-Arriaga, "Semantic feature extraction using SBERT for dementia detection," *Brain Sciences*, vol. 12, no. 2, Feb. 2022, doi: 10.3390/brainsci12020270.
[8] J. Mullis, C. Chen, B. Morkos, and S. Ferguson, "Deep neural networks in natural language processing for classifying requirements by origin and functionality: an application of BERT in system requirements," *Journal of Mechanical Design*, vol. 146, no. 4, Apr. 2024, doi: 10.1115/1.4063764.

[9] Y. Kim *et al.*, "A pre-trained BERT for Korean medical natural language processing," *Scientific Reports*, vol. 12, no. 1, 2022, doi: 10.1038/s41598-022-17806-8.
[10] Y. Gu *et al.*, "Domain-specific language model pretraining for biomedical natural language processing," *ACM Transactions on Computing for Healthcare*, vol. 3, no. 1, pp. 1–23, Jan. 2022, doi: 10.1145/3458754.
[11] R. Gupta, "Bidirectional encoders to state-of-the-art: a review of BERT and its transformative impact on natural language processing," *Informatics Economics Management*, vol. 3, no. 1, pp. 311–320, Mar. 2024, doi: 10.47813/2782-5280-2024-3-1-0311-0320.
[12] A. Özçift, K. Akarsu, F. Yumuk, and C. Söylemez, "Advancing natural language processing (NLP) applications of morphologically rich languages with bidirectional encoder representations from transformers (BERT): an empirical case study for Turkish," *Automatika*, vol. 62, no. 2, pp. 226–238, Apr. 2021, doi: 10.1080/00051144.2021.1922150.
[13] B. Min *et al.*, "Recent advances in natural language processing via large pre-trained language models: a survey," *ACM Computing Surveys*, vol. 56, no. 2, pp. 1–40, Feb. 2024, doi: 10.1145/3605943.
[14] A. Gupta, A. Chadha, and V. Tewari, "A natural language processing model on BERT and YAKE technique for keyword extraction on sustainability reports," *IEEE Access*, vol. 12, pp. 7942–7951, 2024, doi: 10.1109/ACCESS.2024.3352742.
[15] R. Vedapradha, R. Hariharan, and R. Shivakami, "Artificial intelligence: a technological prototype in recruitment," *Journal of Service Science and Management*, vol. 12, no. 3, pp. 382–390, 2019, doi: 10.4236/jssm.2019.123026.
[16] B. Ele, A. Ele, and F. Agaba, "Intelligent-based job applicants' assessment and recruitment system," *British Journal of Computer, Networking and Information Technology*, vol. 6, no. 1, pp. 25–46, Jan. 2024, doi: 10.52589/BJCNIT-ISOBAMPD.
[17] A. Ujlayan, S. Bhattacharya, and Sonakshi, "A machine learning-based AI framework to optimize the recruitment screening process," *International Journal of Global Business and Competitiveness*, vol. 18, no. S1, pp. 38–53, Jan. 2024, doi: 10.1007/s42943-023-00086-y.
[18] B. Kinge, S. Mandhare, P. Chavan, and S. Chaware, "Resume screening using machine learning and NLP: a proposed system," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, pp. 253–258, Apr. 2022, doi: 10.32628/CSEIT228240.
[19] V. Rajath, R. T. Fareed, and S. Kaganurmath, "Resume classification and ranking using KNN and cosine similarity," *International Journal of Engineering Research & Technology*, vol. 10, no. 8, pp. 192–195, 2021, doi: 10.17577/IJERTV10IS080057.
[20] D. Pandita, "Talent acquisition: analysis of digital hiring in organizations," *SAMVAD: SIBM Pune Research Journal,* vol. 18, pp. 66–72, Sep. 2019, doi: 10.53739/samvad/2019/v18/146242.
[21] M. Mokoatle, V. Marivate, D. Mapiye, R. Bornman, and V. M. Hayes, "A review and comparative study of cancer detection using machine learning: SBERT and SimCSE application," *BMC Bioinformatics*, vol. 24, no. 1, Mar. 2023, doi: 10.1186/s12859-023-05235-x.
[22] L. George and P. Sumathy, "An integrated clustering and BERT framework for improved topic modeling," *International Journal of Information Technology*, vol. 15, no. 4, pp. 2187–2195, Apr. 2023, doi: 10.1007/s41870-023-01268-w.
[23] J. Wang *et al.*, "Utilizing BERT for information retrieval: survey, applications, resources, and challenges," *ACM Computing Surveys*, vol. 56, no. 7, pp. 1–33, Jul. 2024, doi: 10.1145/3648471.
[24] B. Juarto and A. S. Girsang, "Neural collaborative with sentence BERT for news recommender system," *JOIV : International Journal on Informatics Visualization*, vol. 5, no. 4, Dec. 2021, doi: 10.30630/joiv.5.4.678.
[25] N. Reimers and I. Gurevych, "Sentence-BERT: sentence embeddings using siamese BERT-networks," *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, pp. 3982–3992, 2019, doi: 10.18653/v1/d19-1410.

BIOGRAPHIES OF AUTHORS

Asmita Deshmukh is Assistant Professor at K. J. Somaiya Institute of Technology, Mumbai, India. She is currently pursuing Ph.D. in Computer Science Engineering from Sant Gadge Baba Amravati University. She obtained her Master of Engineering in Computer Engineering from the University of Mumbai and Bachelor of Engineering in Computer Science Engineering from Amravati University in 1996 and 2012. Her main area of research interests is natural language processing, deep learning, data science, and big data. She can be contacted at email: asmitadeshmukh7@gmail.com.

Anjali Raut Dahake is Professor and In-charge Principal at HVPM's College of Engineering and Technology, Amravati, India. She completed her Ph.D. in Computer Science Engineering from Sant Gadge Baba Amravati University and obtained her Master of Engineering in Computer Science Engineering and Bachelor of Engineering in Computer Science Engineering from Amravati University in 1994 and 2013. Her main area of research interests is data mining. She can be contacted at email: anjali_dahake@rediffmail.com.