

I-URIP 2025
**INTERNATIONAL UNDERGRADUATE RESEARCH
INNOVATION, INVENTION AND DESIGN**

**UNDERGRADUATES
PIONEERING
TOMORROW'S
BREAKTHROUGH**

E-BOOK OF EXTENDED ABSTRACTS

END-TO-END TALENT RECRUITMENT FRAMEWORK FOR TALENT CORPORATION MALAYSIA BERHAD

Anis Aqilah Mohd Asri¹, Siti Zanariah Satari^{1*}, and Ainil Afiqah Mohd Hamdan²

¹ Centre for Mathematical Sciences, Universiti Malaysia Pahang Al-Sultan Abdullah, Lebuhr Persiaran Tun Khalil Yaakob, 26300 Kuantan, Pahang, Malaysia

² Talent Corporation Malaysia Berhad, Level 6, Surian Tower, No.1 Jalan PJU 7/3, Mutiara Damansara, 47810 Petaling Jaya, Selangor, Malaysia

*zanariah@umpsa.edu.my

ABSTRACT

This project explores the application of advanced data analytics, with a focus on deep learning model, to enhance the talent recruitment process. It proposes an end-to-end talent recruitment framework powered by Bidirectional Encoder Representations from Transformers (BERT), a deep learning model under Artificial Intelligence (AI) for natural language processing (NLP) to analyse and match candidate resumes with job descriptions. A system named Candidates Evaluation List for Employment Selection and Talent Acquisition Ranking (CELESTAR) was developed as the main product of this framework. Using datasets of job descriptions and candidate resumes, the BERT model was integrated into the framework to capture contextual insights and semantic meaning, enabling more accurate candidate-job matching and improving the identification of high-potential candidates. Data visualization dashboards were also used to present the results clearly and support Human Resource (HR) decision-making. The implementation of CELESTAR system has demonstrated its ability to streamline recruitment by reducing manual screening workload, shortening decision-making time, improving matching accuracy, and contributing to cost and resource efficiency. Most importantly, the integration of AI ensures a fairer and data-driven selection process, aligning with the New Industrial Master Plan (NIMP) 2030, which promotes the shift towards digitalisation in the workforce. This project benefits not only Talent Corporation Malaysia Berhad but also other organisations seeking to modernize and optimize their recruitment process.

Keywords: *Talent Recruitment Framework, BERT Algorithm, Candidate-Job Matching, Resume Screening, Human Resource*

INTRODUCTION

Recruitment today is no longer just about filling vacancies. Recruiting the right talent is not a sprint but a marathon, and in today's technology-driven world, the competition for top talent has become increasingly intense. Matching the right person to the right job is more critical than ever. Talent recruitment plays a vital role in helping organisations select the best candidates to achieve organisational goals while supporting individual career growth. In Malaysia, Talent Corporation Malaysia Berhad (TalentCorp) is responsible for managing and enhancing national talent development. To support this mission effectively, TalentCorp itself must first implement a strong and efficient recruitment system to ensure that its own workforce consist of qualified and well-matched employees. By recruiting the right talent, TalentCorp can successfully nurture and support the wider national talent ecosystem.

Traditional recruitment approaches consumes significant time when narrowing down resumes from large applicant pools (Tallapragada et al., 2023). Therefore, developing an accurate candidate-job matching model is critical as it influences the identification of high-potential candidates. However, many existing models face limitations in understanding the semantic relationships between job description and candidate resumes, leaving notable gaps in accurately identifying top talent (Deshmukh & Bajaj, 2024). Current platforms such as JobStreet are effective in attracting applicants and displaying matched skills, but they still require manual review and shortlisting, offering no ranking features to speed up decisions-making. As a result, HR department is left to manually filter large volumes of resumes, which is both time consuming

and prone to human error (Tallapragada et al., 2023). Traditional keyword matching and Boolean search techniques lack the ability to capture deeper contextual meaning (Mochol et al., 2007), making it difficult to identify high-potential candidates and often leading to mismatched hiring (Malherbe et al., 2014; Chen et al., 2018). Moreover, many existing studies focus only on job-matching models without addressing how they fit into a complete end-to-end recruitment framework.

Therefore, this project aims to proposed an integrated talent recruitment framework that combine two different workflow into a comprehensive end-to-end talent recruitment framework. The framework emphasises the candidate-job matching model, which utilises a deep learning model, Bidirectional Encoder Representations from Transformer (BERT) to calculate matching degree scores between candidate resumes and job descriptions. From this model, a system called CELESTAR was developed to automate and enhance the user experience. This system ensures that companies are able to hire the most suitable candidates for their specific job openings. Additionally, the framework integrates visualisation dashboard to provide actionable insights and a comprehensive solution for recruiters. By achieving these objectives, this project supports TalentCorp's mission to optimize the talent recruitment process, while also contributing to Malaysia's talent development, digitalisation agenda, and workforce sustainability.

RESULTS AND DISCUSSION

This section explains the results obtained and discussion, focusing on how each of the four objectives was successfully achieved. The first objective was to design and propose an end-to-end talent recruitment framework tailored for TalentCorp. The proposed framework is a combination of TalentCorp's existing hiring process and the integration of a BERT-based job matching model (Tallapragada et al., 2023) to enhance candidate screening and selection. The proposed framework encompasses the entire hiring process, divided into 16 distinct stages and categorised into four phases which are Preparation Phase, Execution Phase, Evaluation Phase, and Finalisation Phase. Figure 1 shows the proposed end-to-end talent recruitment framework designed for TalentCorp.

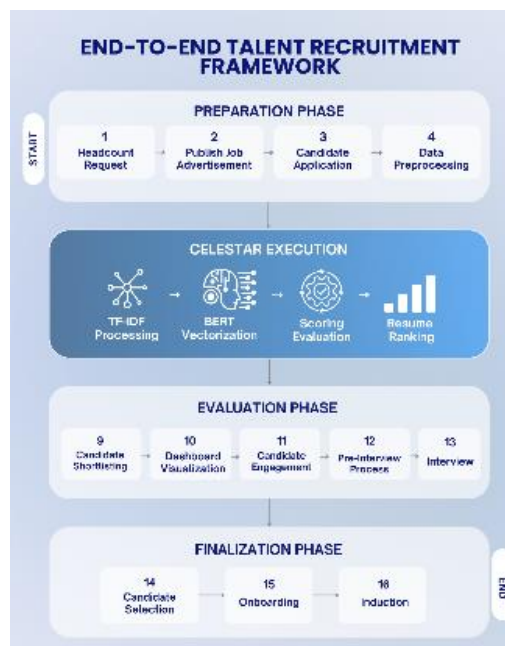


Figure 1: The Proposed End-To-End Talent Recruitment Framework for TalentCorp

To achieve the second objective, real recruitment data were collected from JobStreet between January and May 2025. During this period, a total of six job postings were published under TalentCorp's company profile, with four still available for application and the remaining two classified as expired. In total, 351 candidate resumes were collected from the JobStreet platform where 175 resumes from the expired postings and 176 resumes from the active ones. The job posting data was scraped directly from the JobStreet platform, while the candidate resumes were downloaded from the JobStreet employer portal. After cleaning

and preprocessing, these datasets were trained and used to develop a hybrid candidate-job matching model using the Term Frequency-Inverse Document Frequency (TF-IDF) and BERT. The TF-IDF highlights important keywords in resumes and job description. While BERT, understands the meaning and context of sentences. By combining keyword importance with deeper language understanding, the system produced more accurate matching results compared to traditional keyword searches. Figure 2 shows the CELESTAR system interface, which is available in both desktop and mobile view. Either they are used in the desktop or mobile, the interface for both remains clean, user-friendly and easy to use.

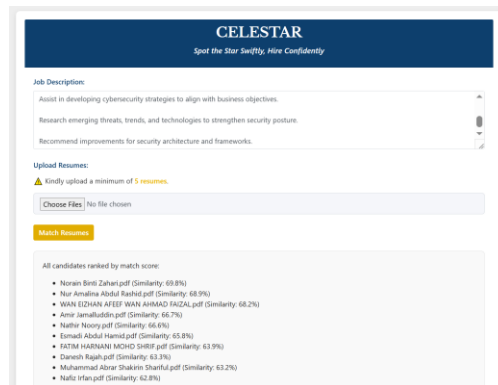


Figure 2: CELESTAR system interface displaying candidate ranking results

The third objective focused on identifying high-potential candidates. The trained model was integrated into a CELESTAR system. When the system tested on a closed job posting, the candidate ultimately hired by HR also appeared within CELESTAR's top 10 ranking, showing strong alignment with human judgment. While in an ongoing cybersecurity hiring, CELESTAR's top 20 shortlisted candidates significantly overlapped with HR's manual screening results. A similarity threshold of 60 percent was used, based on references from past studies that applied cosine similarity scores and identified 0.6 as an effective level for text classification and semantic matching tasks (Alsaket and El-Khatib, 2019). Candidates with scores above this level were considered strong matches. This threshold ensured that the shortlisting process was neither too strict nor too lenient. The list of high potential candidates and their scores was then downloaded into a CSV file and loaded into Tableau for dashboard visualisation. Figure 3 shows the dashboard visualisation of high potential candidates and job application insights.

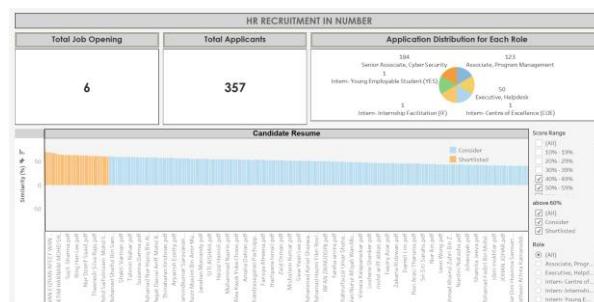


Figure 3: Dashboard visualization of high potential candidates and job application insights

The fourth objective was achieved by presenting CELESTAR's results through an interactive dashboard. The dashboard provided HR with a clear overview of candidate rankings, job application distributions, and filtering options that allowed deeper analysis. This transformed technical similarity scores into simple and actionable insights that recruiters could easily interpret, supporting both technical and non-technical decision-making. From Figure 3, candidates highlighted in orange represent shortlisted candidates and those in blue represent candidates who may be considered. Finally, feedback was gathered from three internal stakeholders, including the Head of Department and two Assistant Vice Presidents. The feedback questions focused on the proposed framework, the CELESTAR system, the identification of high-potential candidates, and the dashboard visualisation. All respondents answered positively to each question, confirming that the framework and system are effective and practical for real-world HR use. Figure 4 shows

the internal stakeholder feedback on the proposed framework, the effectiveness of the BERT model, the candidate identification in the CELESTAR system, and the visualisation dashboard.

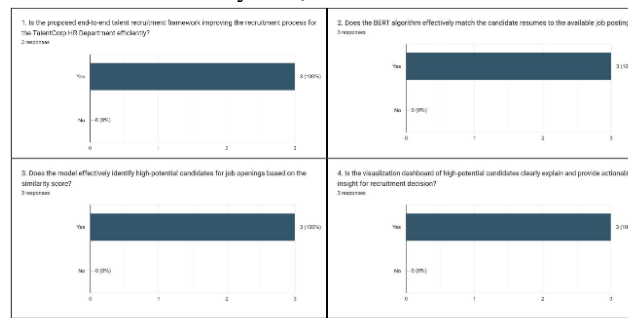


Figure 4: Internal Stakeholder Feedback Summary

Overall, the proposed end-to-end talent recruitment framework and CELESTAR system have shown that recruitment can be made more systematic, efficient and effective. By combining deep learning model with user friendly system interface and insightful dashboard, the solution reduces HR manual screening, speeds up decision-making, and ensures fairer shortlisting of candidates. This project uniqueness lies in combining the TF-IDF model and the BERT model within an end-to-end framework, supported by a user-friendly interface and insightful dashboards. Importantly, the outcomes align with Malaysia's digitalisation agenda under NIMP 2030 and contribute to the Sustainable Development Goals, namely SDG 8 on Decent Work by promoting fairer and more efficient hiring, SDG 9 on Industry Innovation by embedding AI into recruitment processes.

CONCLUSION

This project successfully achieved its objectives of proposing an end-to-end talent recruitment framework, developing the CELESTAR system, identifying high-potential candidates, and presenting the high-potential candidates through dashboard visualisation. The proposed framework provided a structured workflow from headcount request to induction session, while CELESTAR system, powered by BERT model, demonstrated reliable matching accuracy by aligning candidate resumes with job descriptions and supporting HR in making faster, fairer, and more data-driven hiring decisions. The dashboards further enhanced decision-making by transforming similarity scores into clear, actionable insights for both technical and non-technical stakeholders. Collectively, these outcomes reduce manual workload, improve recruitment efficiency, and lower costs, proving the effectiveness of integrating AI into recruitment. The novelty of this work lies in combining advanced deep learning with interactive dashboards in a real-world HR context, positioning CELESTAR as a practical solution beyond academia. Importantly, this project aligns with the New Industrial Master Plan (NIMP) 2030 and contributes to the Sustainable Development Goals (SDGs) by promoting digitalisation, fair employment practices, and inclusive workforce participation. The findings and system developed hold potential to benefit not only Talent Corporation Malaysia Berhad but also wider industries, SMEs, and educational institutions seeking to modernise recruitment and talent placement.

ACKNOWLEDGEMENT

The authors would like to express their gratitude and thanks to Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA) for the academic support provided throughout this project. Appreciation is also extended to Talent Corporation Malaysia Berhad for the opportunity to collaborate and contribute to the development of this project.

REFERENCES

- Alsaket, O., & El-Khatib, H. (2019). A Semantic Similarity Measure for Improving the Performance of Text Clustering [Technical report]. Zarqa University.
https://zu.edu.jo/UploadFile/PaperFiles/PaperFile_22_48.pdf

- Chen, J., Zhang, C., & Niu, Z. (2018). A Two-Step Resume Information Extraction Algorithm. *Mathematical Problems in Engineering*, 2018. <https://doi.org/10.1155/2018/5761287>
- Deshmukh, S., & Bajaj, A. (2024). CareerBoost: A Hybrid RAG-NLP Job Recommendation Framework. 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), I-SMAC 2024 - Proceedings, 853–858. <https://doi.org/10.1109/I-SMAC61858.2024.10714727>
- Malherbe, E., Diaby, M., Cataldi, M., Viennet, E., & Aufaure, M. (2014). Field selection for job categorization and recommendation to social network users. 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 588–595. <https://doi.org/10.1109/asonam.2014.6921646>
- Mochol, M., Wache, H., & Nixon, L. (2007). Improving the Accuracy of Job Search with Semantic Techniques. In *Business information systems* (pp. 301–313). https://doi.org/10.1007/978-3-540-72035-5_23
- Tallapragada, V. V. S., Raj, V. S., Deepak, U., Sai, P. D., & Mallikarjuna, T. (2023). Improved Resume Parsing based on Contextual Meaning Extraction using BERT. *Proceedings of the 7th International Conference on Intelligent Computing and Control Systems, ICICCS 2023*, 1702–1708. <https://doi.org/10.1109/ICICCS56967.2023.10142800>