ISSN: 2641-2039
DOI: 10.33552/GJES.2025.12.000788
Global Journal of Engineering Sciences
Iris Publishers

Research Article

Development of Large Language Models Based Employee-Task Matching in the Consulting Sector

Onur Aygün[1], Ceren Ulus[2] and Mehmet Fatih Akay[3]*

[1]*Innovance Information Technologies, Department of Software Development, Istanbul, Turkey*

[2]*EFA Innovation Consulting and Technology, Department of R&D, Adana, Turkey*

[3]*Cukurova University, Faculty of Engineering, Department of Computer Engineering, Adana, Turkey*

***Corresponding authors:** Mehmet Fatih Akay, Cukurova University, Faculty of Engineering, Department of Computer Engineering, Adana, Turkey

Received Date: November 11, 2025

Published Date: November 20, 2025

Abstract

Nowadays, with the advancement of technology, many new sectors are emerging, and some are gaining increasing importance. In the consulting sector, which provides businesses with the guidance they need, providing the right service to customers quickly is one of the most important strategies. The quality of the services provided directly impacts customer loyalty and business success. Therefore, it is essential to assign the right service to the right personnel. The study aimed to ensure effective employee-task matching, and an approach combining Sentence-Bidirectional Encoder Representations from Transformers (SBERT) and Large Language Models (LLMs) has been proposed for this purpose. SBERT measured the semantic similarity between each candidate's resume and the requirements of the requested position with high accuracy. Integrating the Meta-Llama-3-8B-Instruct, Mistral-7B-Instruct-v0.2, and Phi-3 Mini language models enabled contextual analysis of consulting requests and comprehensive assessment of candidate experiences. LLM-based assessment enabled context-sensitive and in-depth analysis of positions and candidate profiles, not just at the keyword level. Thus, by effectively using unstructured text alongside structured data, much more accurate and explainable matching results have been achieved. The results showed that although the model training time has been longer than other models, Meta-Llama-3-8B-Instruct achieved the highest coverage and accuracy. It provided more justification and technical detail in the suitability scores. As a result of the study, LLMs generated comments for each resume, including detailed justifications and strengths/weaknesses. The LLM approach, applied in conjunction with SBERT, provided both extensive preliminary screening and in-depth qualitative analysis.

Keywords: Employee-Task Matching; Large Lanaguage Models; Consulting Sector

Introduction

Businesses need external expert support in response to the challenges they face in their fields of activity, changing market dynamics, and strategic growth goals. Consulting services, which emerge to address this need, are gaining significant importance today as guidance processes provided by individuals or organizations specialized in a specific field, leveraging their knowledge, experience, and analytical perspectives. The service provided to customers must be delivered in the most efficient, rapid, and accurate

manner. This results in positive customer feedback and contributes significantly to brand image. Furthermore, matching the right personnel to the job is a key factor that directly impacts employee motivation. Employees are able to produce informed and swift work on a topic within their expertise. This maximizes team efficiency. In this context, matching the right personnel with the right competencies is crucial. Improperly matching personnel with service personnel leads to services that fail to meet customer needs and expectations, prolonging the work process. This reduces customer satisfaction and, consequently, damages the company's corporate reputation. Assigning unsuitable personnel to the service area disrupts project management processes and lowers quality standards. This leads to additional resource and time requirements, increasing costs and negatively impacting operating budgets.

During the employee-task matching process, the consultant's academic and professional expertise, previous experience in the sectors they have worked in, their technical knowledge and social skills, as well as their ability to communicate effectively and their past successes are evaluated. Furthermore, the client requesting the service is thoroughly analyzed for their sectoral position, organizational structure, the problem they face, and their priorities and expectations in seeking solutions. This ensures a high level of rapport between the client and the consultant, enabling the consultant to quickly adapt to the process and effectively implement strategic business planning. Therefore, numerous parameters are considered for accurate matching. Accurately analyzing these parameters is time-consuming and cannot yield accurate results using manual methods. Therefore, automatic matching of personnel and services is necessary [1].

Employee-task matching is addressed as a classification problem throughout the academic literature, and studies have been conducted using various machine learning algorithms. However, sufficient labeled historical data may not be available for this classification problem. This leads to overfitting in machine learning methods, weakening the generalization capabilities of the model. Furthermore, inefficiency occurs in scenarios where the model is not dynamically updated. In this context, LLMs, which have become popular today, are gaining importance due to their many advantages for solving problems with contextual, multidimensional, and diverse data, such as employee-task matching. LLMs have high semantic association, natural language understanding, and generalization capabilities. Furthermore, LLMs can directly process natural language texts such as employee resumes, job descriptions, or performance ratings, and can more accurately capture the contextual fit between an employee's past experiences and the requirements in the job description [2].

This study aimed to perform employee-task matching using SBERT and three different language models for the consulting sector. For this purpose, Meta-Llama-3-8B-Instruct, Mistral-7B-Instruct-v0.2, and Phi-3 Mini have been used. A two-stage evaluation method has been implemented. In the first stage, the query sentence, based on the position name and role technical description, has been vectorized using the SBERT model and compared with the embeddings of each resumeto calculate the SBERT similarity score. Additionally, domain-specific bonus points have been assigned based on the technologies and terms used within the resume. This is not only focused on linguistic similarity but also on content-oriented technical suitability.

This study is organized as follows: Section 2 includes relevant literature. Methodology is presented in Section 3. Employee-task matching models are presented in Section 4. Results and discussion are given in Section 5. Section 6 concludes the paper.

Literature Review

[3] presented an LLM-based, end-to-end zero-shot system for skill extraction from job descriptions. They synthetically generated training data for European Skills, Competences, Qualifications and Occupations (ESCO) skills. They then trained a classifier model to extract skill phrases from job postings. A similarity collector has been used to generate skill candidates for reranking with the help of LLM. Synthetic data provided a 10-point performance improvement in the Relevant Precision at 10 (RP10) metric compared to previous remotely supervised approaches. The reranking feature of the GPT-4 language model has been also incorporated into the study. This integration increased the RP10 value by more than 22 points compared to previous methods. Experimental results show that weaker LLM models perform better than natural language commands thanks to framing the task in pseudo-programming format when initializing the language model.

[4] proposed an LLM approach that fine-tuned the baseline resume-based model. Table data obtained from the survey has been converted into text files resembling resumes. These text files have been used to predict the next token, and the LLMs have been retrained with fine-tuning. The resulting fine-tuned LLM has been fed into an occupation model. The results showed that its predictive performance has been superior to all previous models.

[5] compared GPT-4 and human ratings on 736 resumes submitted to job postings in different fields using real-world evaluation criteria. The effect of prompt engineering techniques on GPT-4 ratings has been examined. Differences between GPT-4 and human ratings have been also analyzed across race and gender groups. It has been observed that the scores generated by the LLM showed a low correlation with human ratings, and the two approaches cannot be used interchangeably. On the other hand, the integration of rapid engineering techniques such as Chain of Thought (CoT) has been observed to improve the performance of LLM ratings. Experimental results indicate that LLM scores do not significantly differ from human evaluations.

[6] aimed to evaluate recommendation quality and generate large-scale negative signals while maintaining cost-effectiveness. To this end, a fine-tuned LLM-based approach has been presented. A classifier has been trained using the labels generated by the LLM. This classifier improved the performance of recommendation systems.

[7] presented an LLM-based recruitment recommendation sys-

tem that enables contextual understanding and matching of job offers with resumes. Unlike traditional classification approaches, the proposed system captures the complex relationships between candidate qualifications and job requirements. The system is personalized by analyzing resumes and job descriptions. Successful recommendations are provided. The system's success has been tested on an open dataset of resumes and job offers. The results show that the proposed system outperforms traditional deep learning models such as BERT.

[8] This approach aims to overcome the shortcomings of traditional methods in employee-task matching and to improve the interpretation of deep learning approaches. A new employee-task matching approach based on the BM25 algorithm and a pre-trained language model has been proposed. In the first stage, the BM25 algorithm pre-screened resumes based on job descriptions. The match degree of keywords has been then calculated. This degree has been used to eliminate candidates that did not match the job requirements. Using BERT, resumes that passed the pre-screening have been ranked and the most suitable candidates have been identified. Experimental results demonstrated the effectiveness of the proposed approach.

[9] A two-stage training framework has been proposed to match job descriptions with suitable candidates. First, a contrastive learning approach has been presented to train the model on a dataset derived from real-world matching rules, such as geographic proximity and research area overlap. While the method has been found to be effective, it has been observed that the model primarily learned the patterns defined by these matching rules. Second, a novel preference-based fine-tuning method called Rank Preference Optimization (RankPO), inspired by Direct Preference Optimization (DPO), has been presented. Experimental results showed that the first-stage model demonstrated high performance on rule-based data, but its robust textual understanding remained weak. After fine-tuning using the RankPO method, the model's performance on the original tasks has been observed to improve.

[10] LLM and Graph Neural Network (GNN) based Signal Integration for Talent and Recruiters method has been proposed to overcome the problems of cold start, filter bubbles, and biases affecting candidate-job matching while developing recommendation algorithms for the LinkedIn job matching system. LLMs have been used to extract semantic representations of textual data such as member profiles and job postings. GNNs have been used to mitigate cold start problems through network effects by modeling complex relationships in the network structure. STAR also included features such as adaptive sampling and versioning. The proposed approach provides a methodology for creating embedded representations in industrial applications.

[11] proposed a DeepSeek-based employee-task matching approach. The proposed system extracted skills from job descriptions and resumes using Natural Language Processing (NLP) and semantic embedding techniques. The outputs have been converted into skill vectors, and employee-task matching scores have been calculated. Examining the results and the resulting analyses, it has been observed that the proposed system has high potential.

[12] presented an approach that incorporates the Automatic Semantic Taxonomy Enrichment Methodology (ASTEM) and the Role Competency Embedding-Based (RCE) framework to reduce manual effort, increase the precision of role-competency matches, and support data-driven decision-making processes. A qualitative case study has been conducted on a large company operating in the aerospace, defense, and security sectors.

[13] aimed to improve students' career prospects in three steps, and an LLM-based system has been proposed for this purpose. Resume analysis identified suitable job opportunities for students, and skill gaps have been identified using resumes, class assignments, and similar external resources. Students have been provided with customized learning paths. The proposed system accurately matched students' profiles with jobs in real-world environments. True skill gaps have been identified while reducing false positives.

[14] A computational system capable of extracting ESCO skills from textual data using NLP techniques, sentence embeddings, HDBSCAN-based clustering, and LLMs has been presented. The similarity between input queries and ESCO skills has been assessed using three different approaches. First, only NLP has been presented; only LLM suggestions have been presented through the official DeepSeek Application Programming Interface (API); and a hybrid approach combining NLP and LLM outputs has been presented. HDBSCAN-based clustering has been then used to group similar skills into coherent clusters. Deepseek-V3 has been integrated through the official API to improve and validate skill mappings. Results indicate that the system has the potential to automate skill extraction with high accuracy for universities, students, and employers. Furthermore, the ESCO API performed only comparable to the LLM approach for English input. Performance has been observed to degrade for Portuguese input.

[15] aimed to develop an LLM-based resume-job intelligent matching system specific to the hotel industry. An analysis of the system's matching performance across different position types is presented. A chain hotel group has been used as a case study. 1,847 records have been obtained between 2014 and 2024 across six main categories: front office, housekeeping, food and beverage, sales, logistics, and management. Through controlled experiments and ablation studies, the performance differences of LLMs in resume screening across various position types have been evaluated. The BERT model has been fine-tuned with Low-Rank Adaptation (LORA) for domain adaptation. Additionally, a multidimensional matching algorithm incorporating skill, experience, and soft skill dimensions has been developed. The results showed that the system achieved higher matching accuracy for more standardized positions than for positions requiring strong customization. Furthermore, a statistically significant performance improvement in the F1 score has been observed compared to traditional Term Frequency – Inverse Document Frequency (TF-IDF) methods.

Methodology

SBERT

SBERT replaces a pre-trained BERT with Siamese and ternary network structures to generate comparable semantically meaningful sentence embeddings using only cosine similarity. The Siamese architecture is a computationally efficient BERT. Using a single copy of the pre-trained BERT requires running all possible combinations of sentence pairs from a dataset to generate a representation for the sentence pairs. SBERT first averages a pair of BERT embeddings into fixed-size sentence embeddings. It uses the two sentence embeddings and the element-wise difference between them. It can also run a structured softmax layer for classification and regression tasks [16].

Meta-Llama-3-8B-Instruct

The model is 8B and 70B in size and consists of pre-trained and instruction-tuned generative text. Meta-Llama-3-8B-Instruct instruction-tuned models are optimized for conversational use cases and outperform most open-source conversational models [17].

Mistral 7B

Mistral 7B utilizes Grouped Query Attention (GQA) and Sliding Window Attention (SWA). GQA significantly improves inference speed. By reducing memory requirements during decoding, it enables higher batch sizes and thus higher throughput. This makes the model important for real-time applications. SWA is designed to process longer sequences more efficiently with lower computational cost. This feature circumvents a common limitation of LLMs. These attention mechanisms collectively contribute to the improved performance and efficiency of Mistral 7B. Mistral 7B is released under the Apache 2.0 license. Furthermore, the model allows fine-tuning across a wide range of tasks [18].

Phi-3 Mini

The Phi-3-mini model stands out today as a transformer decoder architecture with a default context length of 4K. To maximize its benefits to the open-source community, Phi-3-mini is built on a similar block structure to Llama-2. It uses the same tokenizer with a vocabulary size of 320641. All packages developed for the Llama-2 model family can be directly adapted to Phi-3-mini. The model uses a latent size of 3072, 32 headers, and 32 layers. The Phi-3-small model uses the TikToken tokenizer for better multilingual tokenization, with a vocabulary size of 1003522 and a default context length of 8192. The model also follows the standard decoder architecture of a 7D model class with 32 heads, 32 layers, and a latent size of 4096 [19].

EMPLOYEE-TASK MATCHING MODELS

The top 25 resumes that best matched the position description have been identified using SBERT. For this purpose, SBERT models have been integrated with the Sentence Transformers library via HuggingFace. Paraphrase-multilingual-MiniLM-L12-v2 has been selected as the model, and the Sentence Transformers 2.6.1 framework has been used. The embedding size has been set to 384, and Mean Pooling has been applied. Embedding vectors have been normalized using the L2 unit length. A batch size of 16 has been chosen for the training process. Cosine Similarity has been used as the similarity metric.

Meta-Llama-3-8B-Instruct, Mistral-7B-Instruct-v0.2, and Phi-3 Mini have been used as language models. LLM models have been run locally with LM Studio in. gguf format, eliminating the need for an online API. All models have been tested with quantized (Q4) versions to operate in CPU and low-RAM environments. A two-stage evaluation method has been implemented within the study. First, query sentences generated based on the position name and technical description of the role have been vectorized using the SBERT model. Then, SBERT similarity scores have been calculated by comparing them with the embeddings of each resume. Bonus points have been assigned based on the technologies and terms included in the resume, specific to the relevant domain. This approach prioritized not only linguistic similarity but also content-focused technical suitability. Secondly, candidates have been pre-selected based on SBERT and bonus scores. The highest-scoring candidates from the top 50 resumes have been analyzed in detail using the LLM. For each resume, the LLM has been provided with a specially prepared prompt, including the position information, technical requirements, and resume content. A suitability score, descriptive justification, strengths (pros), and weaknesses (cons) have been generated, ranging from 0 to 100. Each candidate has been evaluated three-dimensionally using the SBERT similarity score, content-based bonus score, and LLM score. These three components have been weighted between 30% and 70%. Thus, the final score has been calculated, and the highest-scoring candidates have been prioritized. All LLMs have been run in GGUF format using local inference based on llama.cpp. Quantized (Q4) versions have been preferred, resulting in high accuracy with low memory usage. LLM layer details are given in Table 1.

Table 1: LLM's layer details.

Model	GGUF File	Number of parameters	Quantization	N_ctx	N_GPU_layers	N_threads	Seed
Meta-Llama-3-8B-Instuct	Meta-Llama-3-8B-Instuct-Q4_K_M.gguf	~8B	Q4_K_M	4096	32	8	42
Mistral-7B-Instruct-v0.2	Mistral-7b-insctuct-v0.2.Q4_k_S.gguf	~87B	Q4_K_S	4096	32	8	42
Phi-3 Mini (Microsoft)	Phi-3-mini-4k-instruct-q4.gguf	~3.8B	Q4_0	4096	0	8	42

The resume content, position, and role information have been input to the LLM using a specially designed prompt. A sample query is shown in Figure 1.

A sample query output is provided in Figure 2.

Pozisyon: Java, PL SQL Developer
Programming Languages: Java, Javascript, Angular
Frameworks: Spring, SpringBoot, Hibernate, JPA, Struts, Oracle Forms, Oracle ADF
Databases and Query Languages: Oracle, PLSQL
CI/CD and DevOps Tools: Jenkins, Bamboo, Docker, Kubernetes
Other Tools and Technologies: Maven, Git, Webservices (SOAP and REST), Elastic, Keycloak, RabbitMQ, CAS
TDD (Junit, Mockito, PowerMock), Swagger, IBM Content Manager, XML, JSX, JSON, YAML
Application Servers: Weblogic

Figure 1: A sample query.

Final Skor	55 / 100
Justification	Adayın geçmişte Java geliştirici olarak çalıştığı ve PL/SQL bilgisine sahip olduğu belirtilmiş. Ancak son yıllarda geliştirici olarak aktif rol almadığı, Java/PLSQL ile güncel deneyim eksikliği ve ihtiyaç duyulan teknolojilerin (Spring, Angular vb.) CV'de yer almaması nedeniyle düşük puan verilmiş.
Açıklama	Aday, geçmişte Java geliştiricisi olarak çalışmış ve PL/SQL teknolojisini belirtmiş. Ancak son iş deneyimlerinin analist ve proje yönetimi odaklı olması, teknik yeterliliğini sınırlandırıyor. İlgili araçlar ve framework'ler açıkça CV'de belirtilmemiş.
Avantajlar (Pros)	- Java geliştirici geçmişi \n- PL/SQL bilgisi mevcut
Dezavantajlar (Cons)	- Son deneyimler teknik değil (analist/Scrum Master) \n- Java/PLSQL ile güncel çalışma yok \n- Gerekli framework ve araçlar CV'de eksik (Spring, Hibernate, Docker, Angular vb.)

Figure 2: A sample query output.

Results and Discussion

– Meta-Llama-3-8B-Instruct provided the highest coverage and accuracy. It provided more justification and technical detail in its suitability scores. The model training time was longer than the other models.

– Mistral 7B was not as deep as Meta-Llama-3-8B-Instruct, but it provided balanced results. The model training time was shorter than Meta-Llama-3-8B-Instruct.

– Phi-3 Mini runs fast and consumes very little memory, but the answers have been very summary and superficial. LLM generated detailed justifications and strengths/weaknesses for each resume, along with commentary.

– The SBERT + LLM approach provided both extensive preliminary screening (SBERT) and in-depth qualitative analysis.

The relevant position to be matched with the employee has been identified as Java, PL SQL Developer. The results of the six candidates with the highest scores for this position are presented in Table 2. Table 2 shows the final scores obtained from the Meta-Llama-3-8B-Instruct model, which demonstrated the highest performance within the scope of the study.

If the resume included domain and technology keywords related to the position (e.g., Java, Spring Boot, PL/SQL, Oracle, REST API), a bonus of +0.05 points has been added for each unique match. In scenarios where the same keyword appeared multiple times, only a single match has been considered. In this context, the total bonus score increased linearly with the number of keywords identified. The final SBERT+Bonus Score for each resume is calculated using the following formula:

$$SBERT+Bonus\ Score=SBERT\ Cosine+(0.05\times\ number\ of\ matches)$$

The response output generated by the model has been divided into five main sections: Final Score (0–100), Rationale, Explanation, Advantages, and Disadvantages. The Final Score has been calculated using the following formula. This value has been added together with the score generated by the large language model (LLM) to form the final score. The final score has been determined using the following formula:

$$Final\ Score = 0.7 \times LLM\ Score + 0.3 \times (SBERT+Bonus\ Score \times 100)$$

Table 2: The results obtained with Meta-Llama-3-8B-Instruct.

Position	Resume File	SBERT + Bonus Score	LLM Score	Final Score
Java, PL SQL Developer	462.txt	1,5397	78	100,76
Java, PL SQL Developer	276.txt	1,245	80	93,35
Java, PL SQL Developer	282.txt	1,0701	78	86,7
Java, PL SQL Developer	426.txt	1,0644	78	86,53
Java, PL SQL Developer	18.txt	1,0225	78	85,25
Java, PL SQL Developer	439.txt	0,9825	78	84,07

Based on the results:

– Candidate number 462.txt had the highest SBERT + Bonus Score of 1.5397, with a Final Score of 100.76. This result indicates that the candidate accurately aligned with the position requirements in terms of textual similarity.

– Candidates in files 282.txt, 426.txt, and 18.txt received scores in the 86–85 range due to low SBERT scores despite similar LLM scores. This suggests poor textual cohesion or task-related semantic similarity.

– Candidate number 439.txt had the lowest SBERT + Bonus Score (0.9825) and the lowest Final Score of 84.07.

– Considering all these results, it can be said that the SBERT + Bonus Score has a distinctive effect in measuring position-specific textual cohesion.

Conclusion

In today's digital and economic age, the consulting industry is critical its knowledge-based decision-making processes. Sustainable customer satisfaction is a strategic imperative for businesses. Matching the right personnel with the right role is essential to improve service quality and efficiency. This study proposes an approach combining SBERT and LLMs for employee-task matching. As a result, the LLMs generated comments for each resume, including detailed justifications and strengths/weaknesses. The LLM approach, combined with SBERT, provided both comprehensive preliminary screening and in-depth qualitative analysis. Employee-task matching studies in the literature typically use traditional machine learning algorithms or semantic similarity measures. In contrast, this study leverages the Transformer-based SBERT model and the contextual interpretation and inference-generating capabilities of LLMs. The combined use of these two models contributes to quantitative similarity measurement and qualitative descriptive assessment. Additionally, by combining SBERT and LLM outputs with a weighted formula in the final score calculation, more balanced results have been achieved. The language models used included Meta-Llama-3-8B-Instruct, Mistral-7B-Instruct-v0.2, and Phi-3 Mini, which significantly contribute to the novelty of the study. This allows multiple comparisons of LLM contextual performance. All these elements distinguish the study from the existing literature.

Acknowledgement

None.

Conflicts of Interest

No conflicts of interest.

References

1. Ali M (2025) Enhancing job-skill matching with LLM-driven data augmentation and fine-tuned Bi-encoders. CLEF (Working Notes).
2. Kavas H, Serra-Vidal M, Wanner L (2024) 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 pp. 3-9.
3. Clavié B, Soulié G (2023) Large language models as batteries-included zero-shot esco skills matchers. arXiv preprint arXiv:2307.03539.
4. Athey S, Brunborg H, Du T, Kanodia A, Vafa K (2024) Labor-llm: Language-based occupational representations with large language models. arXiv preprint arXiv:2406.17972.
5. Vaishampayan S, Leary H, Alebachew Y B, Hickman L, Stevenor B, et al. (2025) Human and LLM-Based Resume Matching: An Observational Study. In Findings of the Association for Computational Linguistics: NAACL 2025 pp. 4808-4823.
6. Pei Y, Pang Y W, Cai W, Sengupta N, Toshniwal D (2024) Leveraging LLM generated labels to reduce bad matches in job recommendations. In Proceedings of the 18th ACM Conference on Recommender Systems pp. 796-799.
7. Abidi R, Inoubli W, Ayadi MG (2024) Leveraging Large Language Models (LLMs) to Match Job Offers with Candidate CVs. In International Conference on Management of Digital (pp. 387-400). Cham: Springer Nature Switzerland.
8. Tang J, Chen H, Chen Z, Pan J, He Y, Zhao J, et al. (2023, December) A Person-job Matching Method Based on BM25 and Pre-trained Language Model. In Proceedings of the 2023 6th International Conference on Machine Learning and Natural Language Processing, pp. 78-83.
9. Zhang Y, Wang M, Wang Y, Wang X (2025) RankPO: Preference Optimization for Job-Talent Matching. arXiv preprint arXiv:2503.10723.
10. Liu P, Arora R, Shi X, Le BH, Shen Q, et al. (2025) A Scalable and Efficient Signal Integration System for Job Matching. In Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2, pp. 4659-4669.
11. Wang B (2025) Research on Intelligent Competency Modeling and Job-Person Matching Mechanism Based on the DeepSeek Model. Educational Innovation Research 3(5): pp. 6-12.
12. Barba G, Corallo A, Lazoi M, Lezzi M (2025) Large language models for competence-based HRM: A case study in the aerospace industry. Journal of Innovation & Knowledge 10(5): 100780.
13. Patel M, Tyagi V (2024) Enhancing Career Development Utilizing LLM for Targeted Learning Pathway. Journal of the Korea Information Processing Society 13(9): pp. 460-467.

14. Ferreira A, Ribeiro F, Neves A (2025, July) Mapping ESCO Skills Taxonomy to Educational Offers Using Natural Language Processing and Large Language Models. In 2025 6th International Conference of the Portuguese Society for Engineering Education (CISPEE) IEEE pp. 1-7.

15. Sun J (2025) Large Language Model-based Resume-Job Intelligent Matching Algorithm and Its Adaptability Case Study Across Different Positions in the Hotel Industry. International Journal of Emerging Technologies and Advanced Applications 2(7): pp. 1-7.

16. Choi H, Kim J, Joe S, Gwon Y (2021, January) Evaluation of bert and albert sentence embedding performance on downstream nlp tasks. In 2020 25th International conference on pattern recognition (ICPR) IEEE: pp. 5482-5487.

17. Zhang P, Shao N, Liu Z, Xiao S, Qian H, et al. (2024) Extending llama-3's context ten-fold overnight. arXiv preprint arXiv:2404.19553.

18. Chaplot DS (2023) Albert q. jiang, alexandre sablayrolles, arthur mensch, chris bamford, devendra singh chaplot, diego de las casas, florian bressand, gianna lengyel, guillaume lample, lucile saulnier, lélio renard lavaud, marie-anne lachaux, pierre stock, teven le scao, thibaut lavril, thomas wang, timothée lacroix, william el sayed. arXiv preprint arXiv:2310.06825, 3.

19. Abdin M, Aneja J, Behl H, Bubeck S, Eldan R, et al. (2024) Phi-4 technical report. arXiv preprint arXiv:2412.08905.