VerbaNex at TalentCLEF2025: Semantic Matching of Multilingual Job Titles through a Framework Integrating ESCO Taxonomy

Notebook for the TalentCLEF Lab at CLEF 2025

Melissa Moreno Novoa[1], Juan Carlos Martinez-Santos[1], Jairo Serrano[1] and Edwin Puertas[1]

[1]*Universidad Tecnológica de Bolívar, Cartagena, Colombia*

Abstract

The accurate alignment of occupational titles across multiple languages poses a key challenge in modern talent management systems. Linguistic differences, semantic ambiguity, and the absence of structured references hinder the automatic identification of equivalent roles in diverse cultural contexts. Although language models have improved significantly, their performance remains limited. To address this, we propose a multilingual system for the semantic matching of job titles in English, Spanish, and German. The approach combines a pre-trained model with a fine-tuning process, using positive and negative examples generated from occupational family relationships defined by the ESCO taxonomy. The goal is to represent job titles within a shared semantic space that enables meaningful cross-lingual comparisons. The results demonstrate high performance in binary classification (F1 score: 0.9573) and information retrieval tasks. English shows the strongest performance (MAP: 0.5057), followed by Spanish (MAP: 0.4071) and German (MAP: 0.3160), confirming the system's ability to operate effectively in multilingual settings. These findings support the use of semantic representations to enhance linguistic interoperability in human resource applications.

Keywords

Multilingual Job Title Matching, ESCO, Language Models, Human Capital Management

1. Introduction

The contemporary labor market is undergoing a profound transformation driven by technological advances and new labor dynamics. Technologies such as automation and artificial intelligence (AI) are reshaping traditional roles. At the same time, the expansion of remote work has radically changed hiring and performance conditions [1]. In both physical and virtual environments, these transformations have increased the need for adaptability among workers and organizations. Moreover, the rise of remote work modalities has increased access to opportunities, encouraging participation of historically underrepresented groups such as women and minorities [2]. This scenario suggests a labor market that is becoming increasingly inclusive and mediated by emerging technologies. This scenario suggests a labor market that is becoming increasingly inclusive and mediated by emerging technologies. In this context, competitiveness can be defined by the capacity to adapt to hybrid and digitalized environments.

Within this scenario, Large Language Models (LLMs) have gained strategic relevance in talent management for both remote and on-site jobs. Generative AI tools, such as ChatGPT, are revolutionizing key human resources (HR) functions by optimizing processes like recruitment, continuous training, and career development planning. Recent studies emphasize that LLMs enhance operational efficiency in Human Capital Management (HCM) by automating repetitive tasks and accelerating decision-making processes [3]. These technologies enable the rapid processing of large volumes of information, such

CLEF 2025 Working Notes, 9 – 12 September 2025, Madrid, Spain

*Corresponding author.

†These authors contributed equally.

memoeno@utb.edu.co (M. M. Novoa); jcmartinezs@utb.edu.co (J. C. Martinez-Santos); jserrano@utb.edu.co (J. Serrano); epuerta@utb.edu.co (E. Puertas)

0000-0002-1748-2208 (M. M. Novoa); 0000-0003-2755-0718 (J. C. Martinez-Santos); 0000-0001-8165-7343 (J. Serrano); 0000-0002-0758-1851 (E. Puertas)

as résumés and cover letters, thus facilitating the accurate identification of candidates aligned with specific profiles [4]. In turn, they help standardize evaluation processes and reduce biases. In general, the integration of advanced language models into HCM not only streamlines administrative processes and enhances the precision of decision making, but also frees human personnel to focus on strategic tasks of greater value.

However, significant challenges arise with the widespread use of NLP and LLMs in human resource management. One major challenge is multilingualism: while English dominates much of AI development, global organizations require support in multiple languages. Current LLMs often fail in languages other than English, especially in low-resource languages [5], and usually lack sensitivity to local cultural contexts. It makes it difficult to ensure consistent quality in semantic understanding and analysis, particularly in languages such as Spanish, French, or Chinese, an essential requirement for multicultural companies. Another concern involves fairness and algorithmic bias [6]. Ensuring transparency and fairness in automated decision-making is a crucial challenge to avoid perpetuating workplace inequalities. Finally, sector-specific adaptability is also a concern: we must fine-tune general NLP models to diverse industrial domains with distinct vocabularies and competencies. Processing heterogeneous data from fields such as IT, healthcare, or manufacturing requires adapting models to each context, as a single NLP algorithm may not perform well across all sectors without additional training [7]. In summary, overcoming language, fairness, and sector specialization barriers is crucial for implementing AI systems in HR responsibly and inclusively.

In response to these challenges, international initiatives offer both conceptual and practical support. ESCO (European Skills, Competences, Qualifications, and Occupations) stands out as a standardized and multilingual taxonomy at the European level. ESCO functions as a shared dictionary of skills, qualifications and occupations, available in the 27 official EU languages [8]. This classification provides uniform descriptions of job profiles and skills, highlighting relationships between professions and competencies. By serving as a shared reference, ESCO facilitates communication between the worlds of education and employment, helping to harmonize talent supply and demand across Europe. That is, employers, jobseekers, and employment services from different countries can communicate effectively when describing positions or skills based on a shared ontology. It is particularly valuable in multilingual settings, where, for example, a job posting described in Italian or Polish can be understood and matched with candidate profiles in English or Spanish via ESCO cross-referencing.

Finally, advances in semantic representation are empowering truly multilingual and intelligent HCM systems. In particular, the use of pre-trained language models enables the generation of vector representations of texts, job offers, profiles, and skills that capture their meaning uniformly across languages. It has enabled automated recommendation and matching functions in multiple languages. Recent studies have developed approaches in which a multilingual encoder is pre-trained using unsupervised learning on large datasets of job postings and skills, followed by contrastive fine-tuning using pairs of similar and dissimilar job offers based on the ESCO taxonomy. This two-phase training strategy significantly improves the quality of the resulting representations, more closely aligning the semantic vectors of job posts with related meanings. As a result, the same model can map job offers and candidates from different languages into a shared semantic space, enabling direct comparisons[9].

2. Related Work

Research in HCM and HR using NLP techniques has grown significantly over the past five years. Among the most prominent approaches is the normalization and recommendation of job titles in multilingual and noisy environments, addressed through contrastive learning strategies that leverage occupational relationships in taxonomies such as ESCO Figure 1, thus improving semantic alignment [10, 11, 12]. Additionally, hierarchical classification directly explores modeling structural levels of taxonomies [13] [14].

In addition, other lines of work apply these techniques to tasks such as job vacancy fraud detection [18], demonstrating their versatility. Semantic representations generated by pre-trained models such as

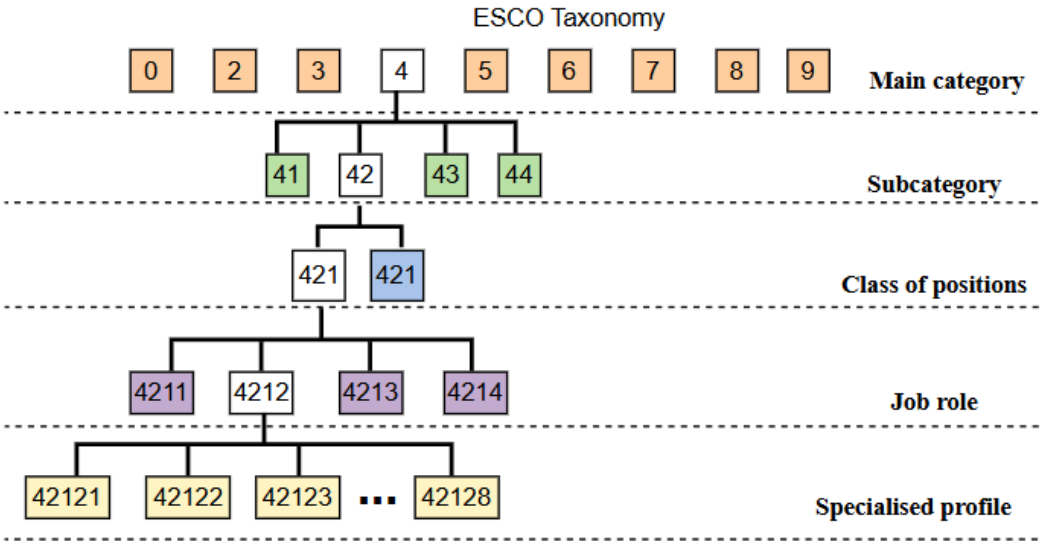

Figure 1: Taxonomy ESCO.

BERT or SBERT, fine-tuned to the occupational domain through skill co-occurrence or textual matching, have also been highlighted [15, 16, 19, 17] . Table 1 summarizes the contributions and techniques employed in these studies. Thus, the present article is situated within this line of research, adopting the approach of semantic representations generated by pre-trained models.

3. Data

The corpus used in Task A Multilingual Job Title Matching consists of job titles in three main languages: English, Spanish, and German. These titles originate from various occupational domains and professional sectors and have been collected and processed to facilitate the identification and comparison of equivalent occupations across languages. The organization responsible for the task provided the corpus used for the task.

The training set uses public terminologies, ensuring that the included job titles are representative of a wide range of occupational areas and aligned with standard market nomenclature. In contrast, The validation and test sets were manually annotated by domain experts appointed by the organizers of the shared task. As described in the TalentCLEF overview paper [9], the annotation process was carried out by the task organizers using specialized tools, and included several quality control stages to ensure the fidelity of semantic relationships across languages.

3.1. Training Set

The training data is presented in tabular format, with three separate files, one for each language. Each entry consists of an occupational family identifier, an occupation identifier, and a pair of related job titles, both associated with the same occupational family. In total, there are three training files: Spanish with 28,880 entries, German with 23,023, and English with 20,724. Altogether, 72,627 job title pairs were grouped by family, where each pair reflects a potential semantic relationship or equivalence between professional roles. This structure allows for the analysis of both textual overlaps and semantic proximity between associated terms.

To illustrate the lexical characteristics of the training data, Figure 5 combines three complementary perspectives: Jaccard similarity (word overlap) Figure 2, Levenshtein similarity (character-level variation) Figure 3, and cosine similarity using TF-IDF (semantic proximity) Figure 4. Most job pairs exhibit low Jaccard similarity, indicating little direct vocabulary overlap. However, Levenshtein similarity reveals shared orthographic patterns, such as gender variations or syntactic changes. In contrast, cosine

Table 1
Summary of related work on multilingual job title normalization and recommendation

Main Author	Contribution and Taxonomy Used	Techniques and Models Applied
Deniz et al. (2024) [10]	Multilingual encoder for job titles using co-occurrence of skills and job pairings from ESCO.	Doc2Vec, NT-Xent contrastive learning, XLM-RoBERTa, MPNet, mUSE-CNN.
Zhang et al. (2023) [11]	Multilingual domain-adapted pretraining with ESCO occupation and skill relations.	ESCOXLM-R: XLM-Rlarge with MLM and ESCO Relation Prediction.
Decorte et al. (2021) [15]	JobBERT: job title embeddings derived from associated skills, no explicit taxonomy.	BERT with gating attention, MLP, negative skip-gram.
Bocharova et al. (2023) [16]	VacancySBERT: semantic embeddings for job titles using skill-based training.	Siamese SBERT, contrastive ranking loss.
Lavi et al. (2021) [17]	conSultantBERT: multilingual matching of CVs and job vacancies in real-world settings.	Fine-tuned Sentence-BERT, binary classification, regression via cosine similarity.
Beręsewicz et al. (2024) [13]	Multilingual hierarchical classification aligned to ISCO for vacancy statistics.	Transformer models (XLM-R, HerBERT), top-down and bottom-up strategies.
Safikhani et al. (2023) [14]	Automatic hierarchical classification of occupations based on the German KldB taxonomy.	BERT, GPT-3, level-wise hierarchical prediction and feature integration.
Taneja et al. (2025) [18]	Fraud detection in job postings using contextualized BERT; no taxonomy used.	Fine-tuned BERT on EMSCAD dataset, binary classification without rebalancing.
Rosenberger et al. (2025) [12]	CareerBERT: job category recommendation from CVs using ESCO taxonomy.	Sentence-BERT, semantic mapping, ranking via MRR, and expert human validation.
Gutiérrez et al. (2025) [19]	JAMES: job title normalization using multi-aspect graph embeddings and co-attention.	Hyperbolic Graph Embeddings, BERT, syntactic similarity, collaborative reasoning.

similarity highlights thematic connections between occupations, even when explicit vocabulary is not shared.

It should be noted that this analysis was conducted on a combined multilingual corpus, which may introduce noise into the lexical metrics. Equivalent titles in different languages tend to obtain low scores in Jaccard or Levenshtein because of the lack of textual matching.

An exploratory analysis of the corpus was also conducted, focusing on job title length and lexical distribution by language. On average, Spanish titles tend to be longer in both word and character count, which may indicate greater specificity or redundancy in descriptive expressions. German titles are more concise in terms of token count but often feature complex compound words. English titles fall between the two in both length and lexical complexity.

The corresponding word clouds illustrate the most frequent terms in job titles for each language. In Spanish, technical and operational terms predominate, with explicit gender marking Figure 6. In English, there is a noticeable tendency toward technical and managerial roles, often expressed using gender-neutral language Figure 7. In German, despite some encoding issues affecting specific characters, similar patterns can be observed, including terms such as director and engineer appearing in both their masculine and feminine forms 8.

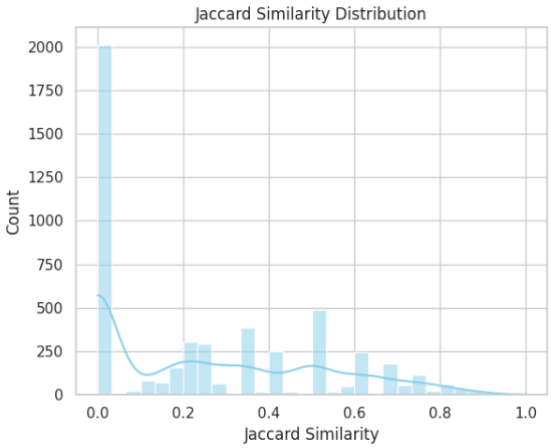

Figure 2: Jaccard similarity: lexical overlap

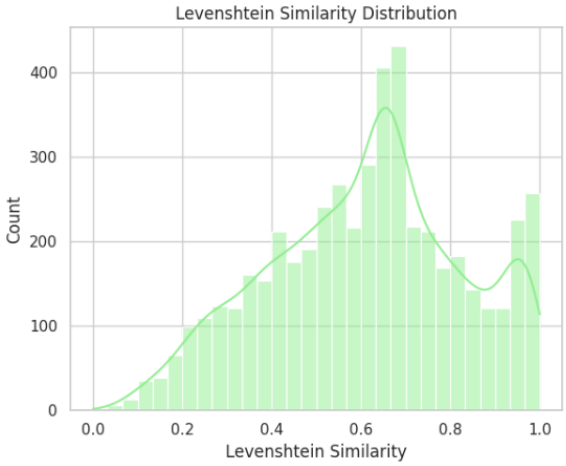

Figure 3: Levenshtein similarity: character-level variation

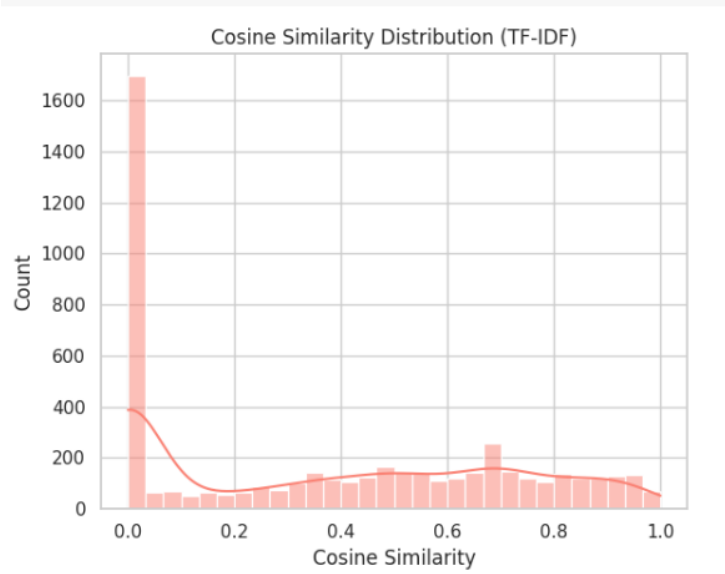

Figure 4: TF-IDF cosine similarity: semantic closeness

Figure 5: Lexical similarity in training data: overlap, variation, and semantic proximity between job title pairs

3.2. Validation Set

The validation set is structured into three files per language: queries, corpus elements, and qrels. This setup simulates an information retrieval scenario in which the system's ability to identify relevant occupations based on a given title is evaluated.

The queries file contains unique identifiers linked to job titles used as queries. The corpus elements file includes identifiers for occupations that form the pool of potentially relevant candidates. Finally, the qrels file defines the binary relevance relationship between each query and the corpus elements, assigning a value of 1 when a relevant relationship exists and 0 otherwise.

This structure enables precise evaluation of retrieval models, as it allows for a clear distinction between relevant and non-relevant predictions based on expert-labeled data.

Figure 6: Spanish word cloud.

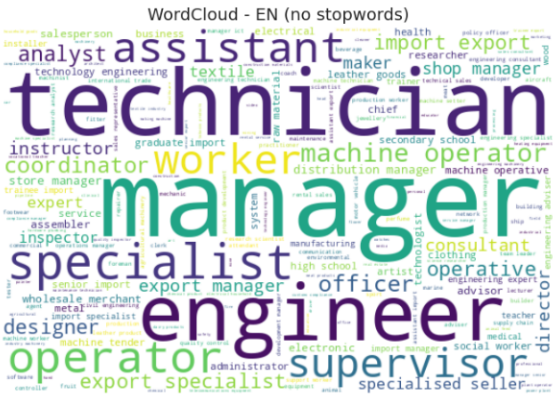

Figure 7: English word cloud.

Figure 8: German word cloud.

3.3. Test Set

Finally, the test set follows a structure similar to the validation set, with separate files for queries and corpus elements, organized by language. Unlike the validation set, no relevance file is provided in this phase, as participants are required to generate their own predictions in TREC format and submit them for external evaluation.

This phased structure, supervised training, labeled validation, and blind testing supports the progressive development of robust models capable of generalizing effectively in multilingual and semantically ambiguous contexts, such as job title alignment.

4. Proposed System Architecture

Figure 9 illustrates the complete pipeline developed for the training, evaluation, and deployment of a multilingual job title matching model. This process is structured into multiple stages that cover the entire workflow from initial data ingestion to final export of results in TREC format. The key components of this architecture are detailed below.

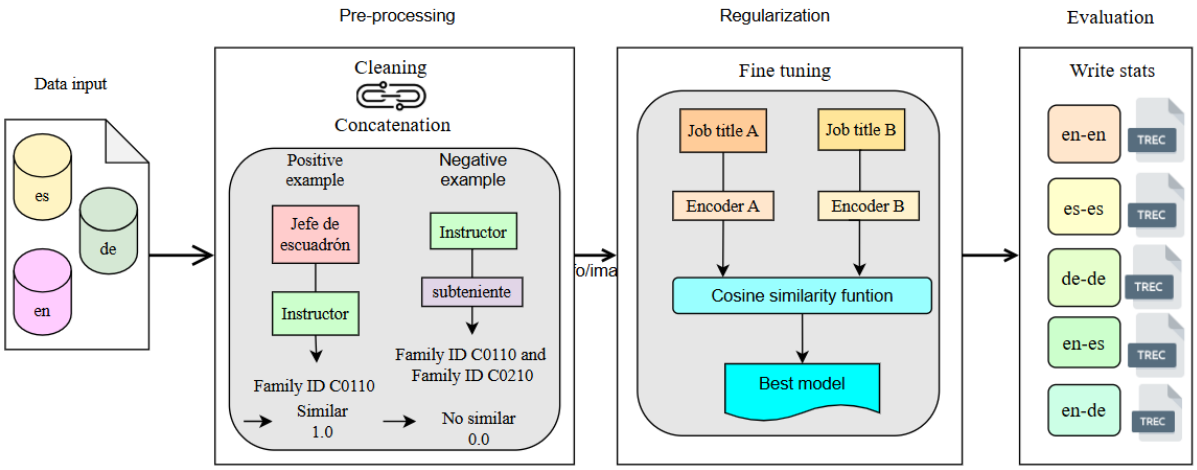

Figure 9: Pipeline

4.1. Data Loading by Language

The process begins with the loading of annotated files in three languages: Spanish (es), English (en), and German (de). Each file contains pairs of job titles along with their corresponding family identifiers, which represent semantically related occupational groupings. Language labels are added and an initial text normalization is performed, such as removing extra whitespace and converting all text to lowercase, to ensure consistency across the data set.

4.2. Concatenation of Multilingual Data

Once individually processed, the three datasets are concatenated into a single *DataFrame*. This step enables multilingual unified training, which is essential for building a model capable of functioning effectively regardless of the language of the job title.

4.3. Generation of Positive and Negative Examples

Two types of job title pairs are generated. Positive examples are composed of titles belonging to the same family identifier, indicating high semantic similarity, and are labeled with 1.0. Negative examples are created through a "hard-negative sampling" strategy, where a job title is paired with another randomly selected from a different family, and assigned a label of 0.0. These pairs are transformed into Input Example objects, encapsulating the pair of texts and their respective labels.

4.4. Model Training (Fine-tuning)

The pre-trained model from Sentence-BERT Transformers [20] was fine-tuned using a contrastive learning approach with CosineSimilarityLoss. This loss function optimizes the cosine similarity between sentence embeddings by increasing the similarity of positive job-title pairs while decreasing it for negative ones. The training process included periodic evaluations by language (es, en, de) using a custom callback to monitor multilingual performance.

4.5. Multilingual Evaluation and Model Saving

Model evaluation is conducted on two complementary levels. First, binary classification assesses the model's ability to accurately distinguish between pairs of similar and dissimilar job titles using standard evaluation metrics such as precision, precision, recall, F1 score, and AUC. Second, an Information Retrieval (IR) evaluation measures the model's practical utility in search and recommendation tasks. This evaluation employs metrics such as Mean Average Precision (MAP), Mean Reciprocal Rank (MRR),

and Precision at K (P@K), computed per language using the queries, corpus, and qrels datasets that simulate real-world search scenarios. The fine-tuned model is saved permanently, allowing it to be reused without retraining, which is crucial for future inference tasks or production deployment.

4.6. Cross-lingual Inference and TREC Export

In the final step, ranking files are generated in TREC format to evaluate the performance of the model in monolingual and cross-lingual retrieval scenarios. All possible combinations between query and corpus languages (e.g., en-en, es-es, de-de, en-es, and en-de) are processed, and cosine similarity scores are used to produce ranked outputs. These results are saved in `.trec` files for standardized evaluation.

5. Experimental Results

During the training process, the multilingual model was fine-tuned using a cosine similarity loss, resulting in progressive improvements across all evaluation metrics. As shown in Table 2, the training loss steadily decreased from 0.0697 to 0.0339 in the final step, indicating stable convergence of the model. Simultaneously, both precision and F1-score increased, ultimately surpassing 95%, with a final F1-score of 0.9573 and a precision of 0.9574, confirming the model's ability to effectively distinguish between similar job title pairs.

In terms of information retrieval (IR), the results reveal a notable difference in performance across languages. According to Table 3, the model achieved higher performance in English (EN), with a MAP value of 0.5057 and MRR of 0.7856, whereas Spanish (ESP) and German (GE) yielded MAP values of 0.4071 and 0.3160, respectively. The P@k metrics reinforce this trend: English achieved a P@1 of 0.6286, compared to 0.1297 in Spanish and 0.2956 in German. This suggests that, while the model is robust, its performance is influenced by linguistic characteristics and potential differences in data quality across languages.

Furthermore, the results of the system evaluation presented in Table 4 show that the average MAP in the three languages (en, es, de) was 0.36, which is consistent with the values obtained during training, particularly when adjusting for variability between pairs of monolingual and cross-lingual. Notably, better performance was observed in English-English matching (0.408) compared to Spanish-Spanish (0.348) and German-German (0.324), while cross-lingual matches also demonstrated competitive performance, such as in the English-German case (0.344).

Overall, these results validate the effectiveness of the proposed model in both binary classification and information retrieval tasks. The metrics obtained demonstrate that the system is capable of identifying semantic similarities between professional titles in different languages, with outstanding performance across the evaluated languages. These findings support the usefulness of the adopted approach for multilingual occupational alignment tasks, while also highlighting opportunities for improvement in the representation of lower-performing languages.

Table 2
Training Results

Step	Training Loss	Accuracy	F1-score	Precision	Recall	AP	MCC
1000	0.0697	0.9394	0.9045	0.9338	0.8770	0.9663	0.8610
5000	0.0482	0.9561	0.9320	0.9447	0.9197	0.9819	0.8998
10000	0.0385	0.9665	0.9489	0.9451	0.9528	0.9899	0.9239
12256	0.0339	0.9721	0.9573	0.9574	0.9571	0.9924	0.9364

Table 3
Training Results — Information Retrieval (IR) Metrics

Metric	es(final)	en (final)	de (final)
MAP	0.4071	0.5057	0.3160
MRR	0.5521	0.7856	0.4905
P@1	0.1297	0.6286	0.2956
P@5	0.6259	0.6743	0.5143
P@10	0.5962	0.5819	0.5044

Table 4
Evaluation Results for *VerbaNex*

Metric	Score
Avg. MAP (en, es, de)	0.36
MAP (en-en)	0.408
MAP (es-es)	0.348
MAP (de-de)	0.324
MAP (en-es)	0.335
MAP (en-de)	0.344

6. Future Work

Considering the results obtained, there remain several opportunities to further enhance the proposed system. One potential direction involves optimizing the model's performance in lower-performing languages, such as German. This could be addressed through the use of supportive machine translation techniques, synthetic data augmentation, or language-specific fine-tuning, which would allow for better adaptation to the linguistic and semantic particularities of each language.

Additionally, it is worth exploring more advanced methods for generating negative examples, such as those based on dynamic embeddings or supervised contrastive learning strategies, to increase semantic discrimination between non-equivalent titles. This may contribute to improvements in both binary classification and retrieval metrics.

Equally important is the integration of complementary structured information, such as task descriptions, required skills, or hierarchical relationships from the ESCO system. Incorporating these signals could enrich the semantic representation of job titles and enhance the model's generalization capabilities in real-world scenarios.

7. Conclusion

By virtue of the results obtained, it is inferred that the proposed system evidences a robust ability to identify semantic similarities between multilingual job titles, with outstanding performance in English. However, areas of opportunity have been identified in languages with lower performance, such as German, suggesting the need to implement language-specific strategies, including data augmentation or targeted optimization techniques.

At the research level, it has been concluded that the incorporation of more refined negative examples, as well as the integration of additional structured data, such as task descriptions or occupational hierarchies, can enrich the semantic representation and optimize the generalization of the model in real contexts. As can be seen from the above, future directions reinforce the value of the adopted approach and point to its evolution towards more accurate, equitable, and adaptive systems for multilingual occupational alignment.

Declaration on Generative AI

During the preparation of this investigation, ChatGPT (OpenAI) was used for the revision of translations into English, as well as for grammatical and spelling correction. After using this tool, the content was reviewed and edited as necessary, and full responsibility for the content of the publication is assumed.

Acknowledgments

The authors express their gratitude to the Call 933 "Training in National Doctorates with a Territorial, Ethnic and Gender Focus in the Framework of the Mission Policy — 2023" of the Ministry of Science, Technology and Innovation (Minciencia). In addition, we thank the team of the Artificial Intelligence Laboratory VerbaNex https://github.com/VerbaNexAI, affiliated with the UTB, for their contributions to this project.

References

[1] Diversity and technology-challenges for the next decade in personnel selection (2023). URL: https://onlinelibrary.wiley.com/doi/10.1111/ijsa.12439. doi:10.1111/ijsa.12439.

[2] D. H. Hsu, P. Tambe, Remote work and job applicant diversity: Evidence from technology startups, SSRN Electronic Journal (2021). URL: https://papers.ssrn.com/abstract=3894404. doi:10.2139/SSRN.3894404.

[3] J. Sun, Research on the application of large language models in human resource management practices, International Journal of Emerging Technologies and Advanced Applications 1 (2024) 1–8. URL: https://www.ijetaa.com/article/view/125. doi:10.62677/IJETAA.2408125.

[4] S. Abdelhay, M. S. R. AlTalay, N. Selim, A. A. Altamimi, D. Hassan, M. Elbannany, A. Marie, The impact of generative ai (chatgpt) on recruitment efficiency and candidate quality: The mediating role of process automation level and the moderating role of organizational size, Frontiers in Human Dynamics 6 (2024) 1487671. doi:10.3389/FHUMD.2024.1487671/BIBTEX.

[5] C. Meinhardt, H. B. U. Zaman, T. Friedman, S. T. Truong, D. Zhang, E. Cryst, V. Marivate, S. K. J. N. Pava, Mind the (language) gap: Mapping the challenges of llm development in low-resource language contexts, Frontiers in Human Dynamics (2025).

[6] Z. Chen, Ethics and discrimination in artificial intelligence-enabled recruitment practices, Humanities and Social Sciences Communications 10 (2023) 1–12. URL: https://www.nature.com/articles/s41599-023-02079-x. doi:10.1057/S41599-023-02079-X;SUBJMETA=4000,4001,4014,4045;KWRD=BUSINESS+AND+MANAGEMENT,SCIENCE.

[7] N. Otani, N. Bhutani, E. Hruschka, Natural language processing for human resources: A survey, in: W. Chen, Y. Yang, M. Kachuee, X.-Y. Fu (Eds.), Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 3: Industry Track), Association for Computational Linguistics, Albuquerque, New Mexico, 2025, pp. 583–597. URL: https://aclanthology.org/2025.naacl-industry.47/.

[8] European Commission, ¿qué es esco? clasificación europea capacidades/competencias, cualificaciones y ocupaciones (esco), 2025. URL: https://esco.ec.europa.eu/es/about-esco/what-esco, accessed: 2025-06-04.

[9] Gasco, Luis and Fabregat, Hermenegildo and García-Sardiña, Laura and Estrella, Paula and Deniz, Daniel and Rodrigo, Álvaro and Zbib, Rabih, Overview of the TalentCLEF 2025 Shared Task: Skill and Job Title Intelligence for Human Capital Management, in: International Conference of the Cross-Language Evaluation Forum for European Languages, Springer, 2025.

[10] D. Deniz, F. Retyk, L. García-Sardiña, H. Fabregat, L. Gasco, R. Zbib, Combined unsupervised and contrastive learning for multilingual job recommendation | european skills, competences, qualifications and occupations (esco), ???? URL: https://esco.ec.europa.eu/en/about-esco/publications/publication/combined-unsupervised-and-contrastive-learning-multilingual-job.

[11] M. Zhang, R. van der Goot, B. Plank, Escoxlm-r: Multilingual taxonomy-driven pre-training for the job market domain, Proceedings of the Annual Meeting of the Association for Computational Linguistics 1 (2023) 11871–11890. URL: https://arxiv.org/pdf/2305.12092. doi:10.18653/v1/2023.acl-long.662.

[12] J. Rosenberger, L. Wolfrum, S. Weinzierl, M. Kraus, P. Zschech, Careerbert: Matching resumes to esco jobs in a shared embedding space for generic job recommendations, Expert Systems with Applications 275 (2025) 127043. URL: https://www.sciencedirect.com/science/article/pii/S0957417425006657?utm_source=chatgpt.com. doi:10.1016/J.ESWA.2025.127043.

[13] M. Beręsewicz, M. Wydmuch, H. Cherniaiev, R. Pater, Multilingual hierarchical classification of job advertisements for job vacancy statistics, 2024. URL: https://arxiv.org/abs/2411.03779. arXiv:2411.03779.

[14] P. Safikhani, H. Avetisyan, D. Föste-Eggers, D. Broneske, Discover artificial intelligence automated occupation coding with hierarchical features: a data-centric approach to classification with pre-trained language models, Discover Artificial Intelligence 3 (123) 6. URL: https://doi.org/10.1007/s44163-023-00050-y. doi:10.1007/s44163-023-00050-y.

[15] J.-J. Decorte, J. V. Hautte, T. Demeester, C. Develder, Jobbert: Understanding job titles through skills, 2021. URL: https://arxiv.org/abs/2109.09605. arXiv:2109.09605.

[16] M. Y. Bocharova, E. V. Malakhov, V. I. Mezhuyev, Vacancysbert: the approach for representation of titles and skills for semantic similarity search in the recruitment domain vitaliy i. mezhuyev 2), Online) Applied Aspects of Information Technology 6 (2023) 52–59. URL: https://doi.org/10.15276/aait.06.2023.4. doi:10.15276/aait.06.2023.4.

[17] D. Lavi, V. Medentsiy, D. Graus, consultantbert: Fine-tuned siamese sentence-bert for matching jobs and job seekers, CEUR Workshop Proceedings 2967 (2021). URL: https://arxiv.org/pdf/2109.06501.

[18] K. Taneja, J. Vashishtha, S. Ratnoo, Fraud-bert: transformer based context aware online recruitment fraud detection, Discover Computing 28 (2025) 1–16. URL: https://link.springer.com/article/10.1007/s10791-025-09502-8. doi:10.1007/S10791-025-09502-8/TABLES/6.

[19] M. Yamashita, J. T. Shen, T. Tran, H. Ekhtiari, D. Lee, James: Normalizing job titles with multi-aspect graph embeddings and reasoning, 2023. URL: https://arxiv.org/abs/2202.10739. arXiv:2202.10739.

[20] N. Reimers, I. Gurevych, Sentence-bert: Sentence embeddings using siamese bert-networks, in: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 2019. URL: http://arxiv.org/abs/1908.10084.