Ranking job-resume similarity using cross-lingual bi-encoding

TalentCLEF Subtask A at CLEF 2026

Daniel-Antoniu Dumitru[1]

[1] "Alexandru Ioan Cuza" University, Carol I 11 Boulevard, 700506 Iași, Romania

Abstract

The paper has the purpose to describe a teacher-student model approach for ranking similarities between job descriptions and resumes. The model supports both English and Spanish data, the teacher being the model that processes the similarity and the student being the model that is used for the cross-lingual processing. A bi-encoder is used to separate job description and resume embeddings. Triplets store for a job description a positive resume match and a negative resume match such that the model can learn aspects that do not represent a fit. Hard negative mining is used to distinguish resumes with similar skills but with different experience levels.

Keywords

JobBERT, bi-encoding, hard negative mining, knowledge distillation, triplets

1. Introduction

The matching between job descriptions and resumes has become an important research topic in Human Capital Management. Recruitment platforms must process large numbers of resumes and job offers and determine which candidates are the most suitable for a position. Traditional approaches relied mainly on keyword matching and manually designed features. Recent advances in Natural Language Processing and Transformer models allow the comparison of resumes and job descriptions at a semantic level, making possible the identification of relevant skills, qualifications and professional experience even when different words are used.

TalentCLEF was introduced as a benchmark for skill intelligence and job title intelligence tasks in Human Capital Management [1, 2]. The competition provides multilingual datasets and evaluation protocols for different recruitment-related problems. Task A focuses on contextualized job-person matching, where the objective is to rank resumes according to their relevance for a job description [3]. The task includes English and Spanish documents and evaluates both monolingual and cross-lingual retrieval scenarios.

One of the first directions explored in the literature is the learning of job title representations. JobBERT learns job title embeddings using job-skill associations extracted from hundreds of millions of job advertisements instead of manually annotated pairs [4]. Similar ideas are proposed in one research article of Zbib et al.[5], where noisy skill labels are used to create auxiliary job representations and train multilingual job title encoders. Both studies show that skill information can be used as weak supervision and can outperform traditional lexical retrieval methods.

Several works extend these ideas to multilingual settings. Deniz et al. combine weak supervision and contrastive learning to create multilingual job title representations aligned across eleven languages [6]. JobBERT-V3 further extends JobBERT using multilingual MPNet and synthetic translations for English, German, Spanish and Chinese [7]. Similar conclusions are obtained by ESCOXLM-R, which uses the ESCO taxonomy during pretraining and improves multiple job-related tasks in several European languages [8]. The importance of multilingual representations is also demonstrated by the MELO benchmark, where embedding models outperform lexical matching methods in cross-lingual occupation

CLEF 2026 Working Notes, 21 – 24 September 2026, Jena, Germany

daniel.dumitru@info.uaic.ro (D. Dumitru)

0000-0002-0071-7891 (D. Dumitru)

normalization tasks [9]. Additional work shows that contrastive multilingual training can reduce language bias and improve retrieval quality between different languages [10].

Another important research direction is skill extraction and skill prediction. Bhola et al. formulate skill prediction as an extreme multi-label classification problem and use BERT representations to predict more than two thousand skills from job descriptions [11]. Anand et al. focus on ranking skills according to their importance for a job title and show that weak supervision can be used to estimate skill relevance without manual annotation [12]. Several studies use the ESCO taxonomy as a source of structured information. Examples include multilingual skill extraction pipelines [13], joint extraction and classification of competences [14], and domain-adapted multilingual models such as ESCOXLM-R [8]. Other studies investigate the extraction of occupations and skills using Transformer models, domain adaptation and low-resource learning approaches [15, 16, 17]. These works show that domain-specific training generally improves extraction quality, although lightweight Transformer models often obtain competitive results.

The problem of matching candidates to jobs has also received significant attention. Zhu et al. use convolutional neural networks to jointly learn representations of resumes and job requirements from historical applications [18]. Jiang et al. extend this idea by combining structured information, textual information and behavioral information extracted from recruitment histories [19]. Chen et al. introduce graph neural networks and professional network information to improve person-job fit prediction [20]. Liu et al. represent jobs using titles, skills and locations in a common embedding space and show that geographic information can improve recommendation quality [21].

Recent studies increasingly use Transformer architectures for resume-job matching. Li et al. propose section-aware BERT models capable of processing long resumes and modeling interactions between resume sections and job descriptions through attention mechanisms [22]. CareerBERT learns a shared embedding space between resumes and occupations using a Siamese architecture and demonstrates that domain-adaptive pretraining improves retrieval quality [23]. Similar results are obtained by Smart-Hiring, which combines resume parsing, skill extraction and semantic matching using sentence embeddings [24]. JobFormer further incorporates skill-aware supervision and Transformer architectures to improve recommendation performance [25]. Zero-shot approaches have also been investigated, showing that compact sentence-transformer models can provide meaningful resume-job rankings even without task-specific training [26].

Several studies investigate complementary problems that are useful for job-person matching. JAMES combines graph embeddings, contextual embeddings and syntactic similarity for job title normalization [27]. Job Description Aggregation Networks learn job title representations directly from job descriptions and show that complete descriptions can provide richer information than isolated skills [28]. Other works focus on industrial sector classification [29], soft skill extraction [30], multilingual occupation classification [31], and general NLP applications for Human Resources [32].

The TalentCLEF 2025 overview shows that the best-performing systems rely mainly on encoder-based architectures, contrastive learning objectives, cross-encoder reranking and data augmentation techniques [1]. The survey of NLP methods for Human Resources reaches similar conclusions and identifies skill extraction, occupation normalization, multilingual retrieval and resume-job matching as the main research directions in the field [32].

Overall, the literature shows several common trends. First, skill information is frequently used as supervision for learning job and resume representations [4, 5, 12]. Second, contrastive learning is one of the most commonly used training strategies for retrieval and ranking tasks [6, 7, 10]. Third, multilingual and cross-lingual representations have become increasingly important due to the international nature of recruitment platforms [9, 6, 7]. Finally, recent systems tend to use bi-encoder architectures that embed jobs and candidates in a common vector space and compare them using similarity measures [23, 24, 26].

Despite the progress obtained so far, ranking complete resumes against complete job descriptions remains a challenging task. Resumes contain information about education, certifications, skills and experience, while job descriptions may contain different levels of detail and requirements. TalentCLEF 2026 Task A addresses this problem directly by evaluating the ranking of candidates for a given position

Table 1
Example of a job description- resume pair from the dataset

Input type	Text
Job description	A Fitness Coach is responsible for helping clients achieve their fitness goals by designing and leading group or individual fitness programs. The role requires instruction on exercises, proper form, and injury prevention techniques, while motivating clients and monitoring their progress to adjust fitness plans as needed.
Resume	Proficient in Injury Prevention, Motivation, Nutrition, Health Coaching, Strength Training, with mid-level experience in the field. Holds a Bachelor's degree. Holds certifications such as Certified Personal Trainer (CPT) by NASM. Skilled in delivering results and adapting to dynamic environments.

in multilingual scenarios [3]. Therefore, methods based on domain-specific encoders, contrastive learning and multilingual representation learning represent promising directions for this task.

2. Training dataset construction

The training dataset was extracted from the Kaggle platform [33], containing 9 different columns: name of the applicant, age of the applicant, gender, race, ethnicity, resume, job role, job description, and best match. The best match label represents a binary classifier, where 0 means that the resume does not fit the job description and 1 means that the resume does fit the job description. The dataset contains 10.000 samples of pairs between a person and their resume and a job place. From those 10.000 pairs, there are 547 unique values for the names of the applicants.

Regarding the age of the applicants, the dataset is approximately evenly distributed, with the ages in the range 25-55. Each range of 3 consecutive years has approximately the same number of samples, between 888 and 1009. The only exception is at the range 52-55, where an increase in the number of the samples can be seen, with 1.369 applicants. The gender of the candidates is evenly distributed, with 51% of the people being male and 49% of the people being female. The race is also evenly distributed and divided in 3 different categories: 34% are mongoloid/asian, 33% are white/caucasian and the other 33% are negroid/black. The ethnicity of the candidates is diverse, having 21 different categories, with the most common one being vietnamese. The applicants have applied to 51 different job roles, each one of them covering approximately 2% of the database. The matching between a job description and a resume is approximately equally distributed, with 51% of the pairs not being a match and 49% being a match.

According to the documentation of the Kaggle dataset, the data was obtained either from publicly available job applications or through synthesis. The resumes and job descriptions are anonymized, synthesized or derived from public websites.

Table 1 presents an example of a job description-resume pair used in the matching task.

Based on the data provided by TalentCLEF 2026 for the development, a prompt was used in Claude to synthesize and format the resume and job description of the training data. In the prompt, the .csv file with the training data, 3 job description examples and 3 resume examples were attached as files. The model used was Claude Sonnet 4.6. To format the data, the following prompt was given:

```
I want to rewrite the content a little bit, such that both the job description
and the resume to be the same in structure and composition as the job
descriptions and resumes in the development set. To have an idea, I will
send to you some examples of resumes from the development set (49, 67, 78,
237) and some job descriptions from the development set (29243, 32447, 35129,
36044). I want in the job_applicant_dataset.csv to modify the "Resume" column
content to fit with the examples I gave it to you and the same with the job
description. In the "Job Description" you will add in the beginning for each
```

`sample the corresponding "Job Roles". Rewrite the content and eventually add additional synthetic information for each sample such that it fits naturally with the examples that I gave to you and respect the structure. Do not modify anything else. These elements that will be modified will be replaced in each sample in the dataset. The rest of the dataset will remain exactly the same. Generate the new .csv file such that I can download it.`

After the formatting step, a resume and a job description looks like in the Table 2.

Table 2: Example of a resume–job description pair from the dataset after formatting

Input type	Text
Job description	**Fitness Coach** A Fitness Coach is responsible for helping clients achieve their fitness goals by designing and leading group or individual fitness programs. You will provide instruction on exercises, proper form, and injury prevention techniques, encouraging clients to push their limits while maintaining a focus on their well-being. **About the Role** The role requires a passion for health and fitness, a strong understanding of exercise physiology, and the ability to motivate and inspire others. You will also monitor clients progress and make adjustments to their fitness plans as needed to ensure continuous improvement. **Required Skills** – Injury Prevention – Motivation – Nutrition – Health Coaching – Strength Training **Responsibilities** – A Fitness Coach is responsible for helping clients achieve their fitness goals by designing and leading group or individual fitness programs. – You will provide instruction on exercises, proper form, and injury prevention techniques, encouraging clients to push their limits while maintaining a focus on their well-being. – The role requires a passion for health and fitness, a strong understanding of exercise physiology, and the ability to motivate and inspire others. – You will also monitor clients progress and make adjustments to their fitness plans as needed to ensure continuous improvement. **What We Offer** – An intellectually stimulating and mission-driven environment – Structured onboarding and ongoing mentorship programmes – Opportunities to contribute to high-impact projects – Competitive remuneration and work-life balance support **How to Apply** Please submit your application through the platform with a current CV and a brief description of your experience related to this role.

Resume	**Daisuke Mori**
	243 Hill Street
	Amsterdam, North Holland
	+31 20 018-1590
	`daisuke_mori@protonmail.com`
	`https://www.linkedin.com/in/daisukemori`
	Professional Summary
	Motivated Fitness Coach with 2–4 years of hands-on experience in Injury Prevention, Motivation, Nutrition. Adept at applying technical knowledge to solve complex problems and contribute to team objectives. Eager to continue growing professionally in a challenging and collaborative environment.
	Professional Experience
	Fitness Coach
	Elevate Wellness Centre – Sydney, NSW
	June 2022 – Present
	– Designed and delivered personalised Injury Prevention programmes for a client base of 12+ individuals, achieving an average 41% improvement in fitness benchmarks.
	– Led group fitness classes of up to 11 participants, maintaining a 18% retention rate through high-quality instruction.
	– Mentored junior trainers and standardised coaching methodologies across the facility's Motivation programme offerings.
	– Developed nutritional guidance and holistic wellness plans in collaboration with healthcare professionals, increasing Nutrition client satisfaction scores.

Input type	Text
	Fitness Intern Momentum Sports Club – Brussels, Brussels June 2020 – June 2022 – Monitored client progress and adapted programmes to ensure continuous improvement and goal attainment. – Designed customised Motivation training plans incorporating injury prevention and progressive overload principles. – Delivered one-on-one Injury Prevention sessions and group classes, receiving consistently positive client reviews. **Assistant Trainer** Peak Performance Centre – Stockholm, Stockholm County June 2019 – June 2020 – Conducted fitness assessments and recorded Motivation baseline data for new clients. – Maintained equipment cleanliness and contributed to a safe and welcoming facility environment. – Supported senior trainers in delivering Injury Prevention sessions and preparing client programme materials. **Education** **Bachelor of Science in Kinesiology** Peking University – Amsterdam, North Holland Graduated: May 2022 **Skills** – Injury Prevention – Motivation – Nutrition – Health Coaching – Strength Training **Certifications** – Certified Personal Trainer (CPT) by NASM

This step was done to ensure that the structure of the training data is the same as the structure of the development and test datasets, making easier the prediction and ranking of the pairs. The formatting was also used to standardize the structure of a resume, giving additional information about the candidate, such as studies made, achieved certifications, skills possessed, and learned abilities.

The Spanish training dataset was made by translating the resume, job description, and job title fields of the English dataset to Spanish. For that, MarianMTModel is used[34], a sequence-to-sequence Transformer model specialized in the translation of text data. Chunks are used for the translation, such that the model does not receive a quantity of information that is too big. The maximum chunk dimension is of 450 tokens. A preprocessing function is used afterwards to encode the characters in the "latin-1" format and decode it in the "utf-8" format. This step was made because of the presence of unrecognized characters after the translation, such as "Ã³", "Ã©", or "Ã±". The encoding and decoding helps in translating the unrecognized characters in valid Spanish characters. In the end, the resumes and job descriptions were formatted such that they have the same structure as their English version.

3. Methods

3.1. Technologies used

The scripts for training the model were written in Python, version 3.12.13. The main libraries used were transformers, torch and sklearn. The scripts were executed in both Kaggle and Google Colab. For Kaggle, GPU T4 x 2 was used as an accelerator, whereas in Google Colab, the Pro version was used, and the A100 GPU was used as accelerator. The execution of the final version of the script took approximately 35 minutes.

3.2. Architecture

For the results presented in the tables 3 and 4, the English development dataset given by the competition was used for evaluation.

Initially, 3 different models were used to test which one has the best results: JobBERT [4], JobBERTv3 [7], and RoBERTa [35]. In the experiment, only the development set given by the organizers was used, with 472 pairs classified as positive matches. To have also negative examples, all of the possible

Table 3
Performance of transformer models

Model	MAP	MRR	P@1	P@5	P@10
JobBERT	0.9044	0.9500	0.9000	0.8400	0.6400
JobBERTv3	0.9010	1.0000	1.0000	0.8400	0.6400
RoBERTa	0.7258	0.9333	0.9000	0.9000	0.8200

Table 4
Performance based on the number of epochs

Nunber of epochs	MAP	MRR	P@1	P@5	P@10
4	0.7917	0.9333	0.9000	0.9200	0.8900
6	0.7588	0.9333	0.9000	0.9200	0.9100
8	0.6984	0.8833	0.8000	0.9200	0.8600
10	0.6897	0.8833	0.8000	0.9000	0.8400

pairs between a job description and a resume were created and the ones that were not already in the development dataset classified to 1 were labeled to 0, consisting of a total number of 3728 pairs. The data was divided into subsets as follows: 70% of the data was used for training, 10% of the data was used for validation and 20% of the data was used for evaluation. The results are presented in Table 1.

Since JobBERT obtained the best result for the MAP metric, this model was used afterwards in the architecture of the system.

The fine-tuning of the model is based on a ranking similarity to a number in the range [0,1]. For that, the model creates triplets of job descriptions and resumes in the following format: a job description, a resume that is a match for the job and a resume that is not a match for the job. A bi-encoder is used to compare the similarity between the components: one encoder for job descriptions and one encoder for resumes. Each document is represented as a sequence of token vectors. These vectors are then compared such that the model can find elements that recommend or does not recommend a candidate for a job.

To experiment the results obtained for different hyperparameters of the same system, multiple number of epochs were tested. Table 2 presents the result of the experiment.

Since the best results for the MAP metric were obtained on 4 epochs, this number was used in further development.

To have cross-lingual support, the model adopts a knowledge distillation approach with student and teacher models [36], where the teacher is represented by the JobBERT model and the student is represented by the paraphrase-multilingual-mpnet-base-v2 Sentence-Transformer model [37]. The student is used to receive the similarities and differences between jobs and resumes and learn these aspects for documents that are written in the Spanish language. The model uses a contrastive learning approach, where the loss function is a sum of 3 different loss values: cross entropy loss, margin loss and distillation loss.

To calculate the cross entropy loss, the similarity matrix between the job description embeddings and the transpose of the positive resumes embeddings is calculated. The matrix is used as input, where for each row the softmax function is applied to calculate the probability for each element. The batch size is used as target for the cross-entropy function and a label smoothing of 0.05 is applied in the loss calculation.

The margin loss is a triplet margin loss function, having as parameters the embeddings of the job description, positive resume, and negative resume. It is calculated based on the following formula:

$$L(a, p, n) = \max\{d(a_i, p_i) - d(a_i, n_i) + \text{margin}, 0\} \tag{1}$$

where the distance is the Minkowski distance:

$$d(x_i, y_i) = \|x_i - y_i\|_p \tag{2}$$

The margin was set to 0.2 and the p parameter is set to 2, the default value.

The distillation loss is calculated by encoding the raw job description and the raw positive resume with the teacher model. These embeddings are then used to calculate the similarity matrix and the softmax for each row of the matrix. For the student model, the embeddings used in the previous loss functions are used to calculate the similarity matrix and the softmax followed by a logarithm for each row of the matrix. The probabilities obtained for both the teacher and the student model represent the parameters for the Kullback-Leibler divergence loss, result used as distillation loss.

The final loss formula is a sum of the previously presented loss functions:

$$L = L_{\text{CE}} + 0.1\, L_{\text{margin}} + 0.3\, L_{\text{distill}} \tag{3}$$

After the training of the model, an iterative hard negative mining function is implemented and a second training is executed with 2 rounds, each one of them on 4 epochs. This step is implemented to differentiate same job positions, but that require different experience grades[38].

Figure 1 presents the architecture of the final model.

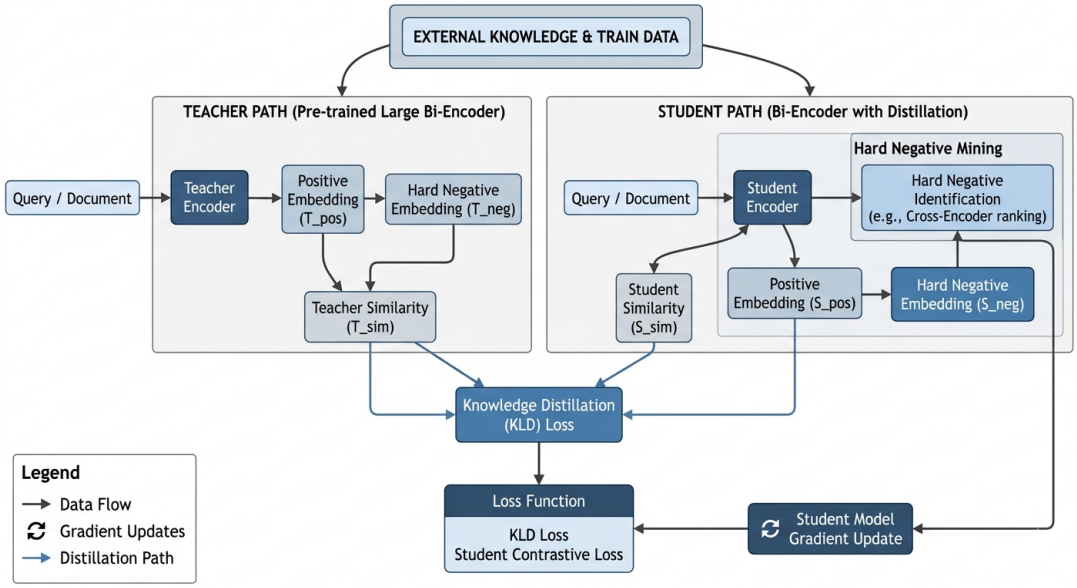

Figure 1: Architecture of the final model

3.3. Hyperparameter setting

For the execution of the system, it was used a batch size of 8, a learning rate of 1e-5, a weight decay of 0.01, and a number of 4 epochs, as a result of the previous experiment. A tau initial value of 0.07 is set as the temperature value, and a value of 2.0 is associated as the temperature for the knowledge distillation. For the hard negative mining, 2 rounds are executed, each one with 4 epochs, and a batch size of 16.

4. Discussion of the results

For the following results, the Kaggle dataset was used for training, and the development and test datasets given by the competition organizers were used for validation and testing.

The model achieved an average MAP value of 0.56, with the following MAP value for each specific case: 0.553 for jobs and resumes in English, 0.566 for jobs and resumes in Spanish, and 0.551 for jobs

Table 5
Comparison of the results

Model	Average MAP	MAP(en-en)	MAP(es-es)	MAP(en-es)
Teacher-student bi-encoder with HNM	0.56	0.553	0.566	0.551
Baseline	0.38	0.399	0.365	0.355

in English and resumes in Spanish. These results exceed the baseline given by the organizers of the TalentCLEF, as presented in figure 3.

The usage of bi-encoding improves the results substantially, allowing the comparison between the embeddings of the jobs and the ones of the resumes. The triplet grouping contributes in the learning of the similarities and the differences, such that the model can distinguish which resumes represent a fit and which resumes are not a match. The teacher-student approach helps in processing resumes in languages other than English (e.g. Spanish) and determining if they are a fit for a job. The hard negative mining function differentiates persons that have similar skills for the same job, but have differences when it comes to previous experience.

For future improvement, the model can be fine-tuned such that it gives a more accurate percentage of the fit between a job and a resume when the job description has many details. If the description has a lot of bullet points, previous experience and certificates requests, the model tends to assign a higher probability score for a resume (around 70%-75%), even if the resume is for a domain that is not correlated with the job description. In contrast, if the job description consists of only a couple of bullet points, the model assigns a lower percentage (around 40%-50%) for a resume with a domain that is not correlated with a job description. The model can also be fine-tuned such that it has not any bias for factors such as age, gender, and race.

5. Conclusions

The matching between job descriptions and resumes remains an important research topic in Human Capital Management. While there are a lot of research papers and articles that use several implementations and solutions, there are areas that can be improved, such as eliminating gender, race, or age bias, and adapting the models to less spoken languages. There are several methods, such as encoders, contrastive learning and multilingual representation, that have proven to be a promising direction for this topic.

Declaration on Generative AI

During the preparation of this work, the author(s) used GPT-5.3-mini in order to: **Summarize the bibliography articles** and **Paraphrase and reword**. Further, the author(s) used Gemini-3.5-Flash for figure 1 in order to: **Generate images**. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] L. Gasco, H. Fabregat, L. García-Sardiña, P. Estrella, D. Deniz, A. Rodrigo, R. Zbib, Overview of the TalentCLEF 2025: Skill and Job Title Intelligence for Human Capital Management, in: International Conference of the Cross-Language Evaluation Forum for European Languages, 2025, pp. 464–485.

[2] L. Gasco, H. Fabregat, L. García-Sardiña, P. Estrella, W. Veys, C. P. Carrino, M. De Lange, D. Deniz Cerpa, A. Rodrigo, J.-J. Decorte, R. Zbib, Overview of the TalentCLEF 2026: Skill and Job Title Intelligence for Human Capital Management, in: International Conference of the Cross-Language Evaluation Forum for European Languages, 2026.

[3] H. Fabregat, L. García-Sardiña, P. Estrella, L. Gasco, C. P. Carrino, D. Deniz Cerpa, A. Rodrigo, R. Zbib, Overview of TalentCLEF 2026: Task A — Contextualized Job-Person Matching, 2026.

[4] J.-J. Decorte, J. Van Hautte, T. Demeester, C. Develder, JobBERT: Understanding Job Titles through Skills, in: International workshop on Fair, Effective And Sustainable Talent management using data science (FEAST @ ECML-PKDD 2021), ECML-PKDD, 2021. URL: https://arxiv.org/abs/2109.09605.

[5] R. Zbib, L. Alvarez Lacasa, F. Retyk, R. Poves, J. Aizpuru, H. Fabregat, V. Simkus, E. García-Casademont, Learning Job Titles Similarity from Noisy Skill Labels, in: International Workshop on Fair, Effective And Sustainable Talent management using data science (FEAST @ ECML-PKDD 2022), ECML-PKDD, 2022. URL: https://arxiv.org/abs/2207.00494.

[6] D. Deniz, F. Retyk, L. García-Sardiña, H. Fabregat, L. Gasco, R. Zbib, Combined Unsupervised and Contrastive Learning for Multilingual Job Recommendation, in: Proceedings of the 4th Workshop on Recommender Systems for Human Resources (RecSys-in-HR 2024), volume 3788 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2024. URL: https://ceur-ws.org/Vol-3788/RecSysHR2024-paper_3.pdf.

[7] J.-J. Decorte, M. De Lange, J. Van Hautte, Multilingual JobBERT for Cross-Lingual Job Title Matching, in: Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2025), CEUR Workshop Proceedings, CEUR-WS.org, 2025. URL: https://ceur-ws.org/Vol-4038/paper_367.pdf.

[8] M. Zhang, R. van der Goot, B. Plank, ESCOXLM-R: Multilingual Taxonomy-driven Pre-training for the Job Market Domain, in: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Toronto, Canada, 2023, pp. 11871–11890. URL: https://aclanthology.org/2023.acl-long.662. doi:10.18653/v1/2023.acl-long.662.

[9] F. Retyk, L. Gascó, C. P. Carrino, D. Deniz, R. Zbib, MELO: An Evaluation Benchmark for Multilingual Entity Linking of Occupations, in: Proceedings of the 4th Workshop on Recommender Systems for Human Resources (RecSys in HR 2024), volume 3788 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2024. URL: https://ceur-ws.org/Vol-3788/RecSysHR2024-paper_2.pdf.

[10] N. Laosaengpha, T. Tativannarat, A. Rutherford, E. Chuangsuwanich, Mitigating Language Bias in Cross-Lingual Job Retrieval: A Recruitment Platform Perspective (2025). URL: https://arxiv.org/abs/2502.03220.

[11] A. Bhola, K. Halder, A. Prasad, M.-Y. Kan, Retrieving Skills from Job Descriptions: A Language Model Based Extreme Multi-Label Classification Framework, in: Proceedings of the 28th International Conference on Computational Linguistics, International Committee on Computational Linguistics, Barcelona, Spain (Online), 2020, pp. 5832–5842. URL: https://aclanthology.org/2020.coling-main.513. doi:10.18653/v1/2020.coling-main.513.

[12] S. Anand, J.-J. Decorte, N. Lowie, Is it Required? Ranking the Skills Required for a Job-Title, arXiv preprint arXiv:2212.08553 abs/2212.08553 (2022). URL: https://arxiv.org/abs/2212.08553.

[13] H. Kavas, M. Serra-Vidal, L. Wanner, Multilingual Skill Extraction for Job Vacancy–Job Seeker Matching in Knowledge Graphs, in: Proceedings of the Workshop on Generative AI and Knowledge Graphs (GenAIK), Association for Computational Linguistics, 2025, pp. 146–155. URL: https://aclanthology.org/2025.genaik-1.15/.

[14] Q. Li, C. Lioma, Joint Extraction and Classification of Danish Competences for Job Matching, in: Advances in Information Retrieval: 45th European Conference on IR Research (ECIR 2023), Proceedings, Part III, volume 13981 of *Lecture Notes in Computer Science*, Springer, 2023, pp. 32–45. URL: https://arxiv.org/abs/2410.22103. doi:10.1007/978-3-031-28238-6_38.

[15] L. Lange, H. Adel, J. Strötgen, Boosting Transformers for Job Expression Extraction and Classification in a Low-Resource Setting, 2021. URL: https://arxiv.org/abs/2109.08597.

[16] L. Vásquez-Rodríguez, B. Audrin, S. Michel, S. Galli, J. Rogenhofer, J. N. Cusa, L. van der Plas, Hardware-Effective Approaches for Skill Extraction in Job Offers and Resumes (2024). URL: https://ceur-ws.org/Vol-3788/RecSysHR2024-paper_9.pdf.

[17] T. Reiser, J. Dörpinghaus, P. Steiner, Detecting Occupations in German Texts: Challenges and Data, in: Proceedings of the 2nd International Workshop on AI in Society, Education and Educational

Research (AISEER 2025), volume 4138 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2025. URL: https://ceur-ws.org/Vol-4138/paper7.pdf.

[18] C. Zhu, H. Zhu, H. Xiong, C. Ma, F. Xie, P. Ding, P. Li, Person-Job Fit: Adapting the Right Talent for the Right Job with Joint Representation Learning, 2018. URL: https://arxiv.org/abs/1810.04040.

[19] J. Jiang, S. Ye, W. Wang, J. Xu, X. Luo, Learning Effective Representations for Person-Job Fit by Feature Fusion, Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM '20) (2020) 2549–2556. URL: https://doi.org/10.1145/3340531.3412717. doi:10.1145/3340531.3412717.

[20] H. Chen, L. Du, Y. Lu, Q. Fu, X. Chen, S. Han, Y. Kang, G. Lu, Z. Li, Professional Network Matters: Connections Empower Person-Job Fit (2023). URL: https://arxiv.org/abs/2401.00010.

[21] M. Liu, J. Wang, K. Abdelfatah, M. Korayem, Tripartite Vector Representations for Better Job Recommendation, 2019. URL: https://arxiv.org/abs/1907.12379.

[22] C. Li, E. Fisher, R. Thomas, S. Pittard, V. Hertzberg, J. D. Choi, Competence-Level Prediction and Resume & Job Description Matching Using Context-Aware Transformer Models, 2020. URL: https://arxiv.org/abs/2011.02998.

[23] 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://doi.org/10.1016/j.eswa.2025.127043. doi:10.1016/j.eswa.2025.127043.

[24] K. Khelkhal, D. Lanasri, Smart-Hiring: An Explainable End-to-End Pipeline for CV Information Extraction and Job Matching (2025). URL: https://arxiv.org/abs/2511.02537.

[25] Z. Guan, J.-Q. Yang, Y. Yang, H. Zhu, W. Li, H. Xiong, JobFormer: Skill-Aware Job Recommendation with Semantic-Enhanced Transformer (2024). URL: https://arxiv.org/abs/2404.04313.

[26] J. Kurek, T. Latkowski, M. Bukowski, B. Świderski, M. Łępicki, G. Baranik, B. Nowak, R. Zakowicz, Ł. Dobrakowski, Zero-Shot Recommendation AI Models for Efficient Job–Candidate Matching in Recruitment Process (2024). URL: https://www.researchgate.net/publication/379122952_Zero-Shot_Recommendation_AI_Models_for_Efficient_Job-Candidate_Matching_in_Recruitment_Process. doi:10.3390/app14062601.

[27] 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.

[28] N. Laosaengpha, T. Tativannarat, C. Piansaddhayanon, A. Rutherford, E. Chuangsuwanich, Learning Job Title Representation from Job Description Aggregation Network, in: Findings of the Association for Computational Linguistics: ACL 2024, Association for Computational Linguistics, 2024, pp. 1319–1329. URL: https://aclanthology.org/2024.findings-acl.77. doi:10.18653/v1/2024.findings-acl.77.

[29] R. Fechner, J. Dörpinghaus, A. Firll, Classifying Industrial Sectors from German Textual Data with a Domain Adapted Transformer, 2023. URL: https://www.researchgate.net/publication/374717212_Classifying_Industrial_Sectors_from_German_Textual_Data_with_a_Domain_Adapted_Transformer. doi:10.15439/2023F6694.

[30] L. Sayfullina, E. Malmi, J. Kannala, Learning Representations for Soft Skill Matching, 2018. URL: https://arxiv.org/abs/1807.07741.

[31] A.-S. Gnehm, S. Clematide, Improving Occupational ISCO Classification of Multilingual Swiss Job Postings with LLM-Refined Training Data, in: Findings of the Association for Computational Linguistics: ACL 2025, Association for Computational Linguistics, 2025, p. 21834–21847. URL: https://aclanthology.org/2025.findings-acl.1124/. doi:10.18653/v1/2025.findings-acl.1124.

[32] N. Otani, N. Bhutani, E. Hruschka, Natural Language Processing for Human Resources: A Survey (2025). URL: https://arxiv.org/abs/2410.16498.

[33] S. K. Nellore, recruitment dataset, https://www.kaggle.com/datasets/surendra365/recruitement-dataset, 2025.

[34] M. Junczys-Dowmunt, R. Grundkiewicz, T. Dwojak, H. Hoang, K. Heafield, T. Neckermann, F. Seide, U. Germann, A. F. Aji, N. Bogoychev, A. F. T. Martins, A. Birch, Marian: An Efficient Neural Machine Translation Framework in C++, in: Proceedings of ACL 2018, System Demonstrations,

Association for Computational Linguistics, Melbourne, Australia, 2018, pp. 116–121. URL: https://aclanthology.org/P18-4020. doi:10.18653/v1/P18-4020.
[35] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, V. Stoyanov, RoBERTa: A Robustly Optimized BERT Pretraining Approach, arXiv preprint arXiv:1907.11692 (2019). URL: https://arxiv.org/abs/1907.11692.
[36] G. Hinton, O. Vinyals, J. Dean, Distilling the Knowledge in a Neural Network, arXiv preprint arXiv:1503.02531v1 (2015). URL: https://arxiv.org/abs/1503.02531.
[37] 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 and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Association for Computational Linguistics, 2019, pp. 3982–3992. URL: https://aclanthology.org/D19-1410/. doi:10.18653/v1/D19-1410.
[38] F. Schroff, D. Kalenichenko, J. Philbin, FaceNet: A Unified Embedding for Face Recognition and Clustering, arXiv preprint arXiv:1503.03832v3 (2015). URL: https://arxiv.org/abs/1503.03832.

A. Appendix

The datasets and the code necessary to reproduce the system are available at: https://github.com/Dumitru-Daniel-Antoniu/TalentCLEF.