Indonesian Journal of Electrical Engineering and Computer Science
Vol. 41, No. 2, February 2026, pp. 710~719
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v41.i2.pp710-719

Integrating contrastive and generative AI with RAG for responsible and fair CV classification

Soumia Chafi, Mustapha Kabil, Abdessamad Kamouss
Laboratory of Mathematics, Computer Science and Applications (LMCSA), Faculty of Sciences and Technology of Mohammedia (FSTM), Hassan II University of Casablanca, Mohammedia, Morocco

Article Info

Article history:

Received Dec 8, 2025
Revised Jan 1, 2026
Accepted Jan 16, 2026

Keywords:

Contrastive learning
CV classification
Fairness
Generative learning
HRIS/SIRH
RAG
Responsible AI

ABSTRACT

The automation of curriculum vitae (CV) classification raises major challenges related to accuracy, fairness, and the heterogeneity of candidate documents. Existing approaches often address these dimensions separately and struggle to reduce demographic bias while maintaining high predictive performance. This study addresses this gap by proposing a hybrid pipeline that combines contrastive learning for representation with a lightweight generative model within a retrieval-augmented generation (RAG) framework. The method is evaluated on a large dataset of 50,000 CVs, using standard classification metrics as well as fairness indicators based on reductions in demographic disparities and equality of opportunity. Experiments show that our approach achieves an accuracy of 95.6% and a fairness index of 0.94, reducing gender-related disparities from 4.8% to 0.3%. These results demonstrate that it is possible to simultaneously improve predictive performance and fairness through a multi-level fairness strategy. The proposed system thus represents a practical and responsible solution for integrating AI into recruitment processes.

Corresponding Author:

Soumia Chafi
Laboratory of Mathematics, Computer Science and Applications (LMCSA)
Faculty of Sciences and Technology of Mohammedia (FSTM)
Hassan II University of Casablanca
Mohammedia, Morocco
Email: soumia.chafi@gmail.com

1. INTRODUCTION

The digitalization of recruitment processes has been accompanied by the growing adoption of artificial intelligence (AI) to automate candidate pre-screening and classification. However, the curriculum vitae (CV), which remains the central document in the evaluation process, presents significant complexity due to the diversity of formats, the heterogeneity of writing styles, and the increasing presence of multimodal elements (logos, graphics, and photographs). This variability makes it difficult to apply traditional text-classification methods, which struggle to capture the contextual and semantic richness of CVs [1].

Recent advances in contrastive learning and generative models have marked a major breakthrough in natural language processing. Approaches such as SimCSE and Contriever produce robust semantic representations that improve similarity matching between CVs and job descriptions, while generative models such as LLaMA enrich the context by generating summaries or relevant complementary information [2], [3]. These methods surpass traditional approaches (TF-IDF, SVM) in both accuracy and their ability to handle unstructured text [4].

However, beyond performance, a crucial challenge remains: ensuring the responsibility and fairness of AI applied to recruitment. Recent studies highlight that models trained on biased data may reproduce or even amplify discrimination related to gender, age, or origin [5], [6]. In such a sensitive domain as human resources, it is therefore essential to design approaches that ensure efficiency, transparency, and algorithmic fairness.

Although many studies have investigated automatic CV classification, most approaches focus either on improving representation models or on reducing bias—but rarely on both simultaneously. Moreover, existing solutions generally fail to account for the heterogeneity of modern CVs, which may combine text, images, tables, and other visual elements. Research on fairness in recruitment remains limited and often focuses on a single sensitive attribute, without offering a comprehensive mitigation framework. To our knowledge, no prior work has combined contrastive learning, retrieval-augmented generation (RAG), multimodality, and fairness within a unified pipeline. This work aims to fill this gap by proposing an integrated approach capable of simultaneously improving performance, transparency, and fairness in CV classification. This study builds upon our previous contributions [7], while paving the way toward responsible and fair AI for recruitment.

2. RELATED WORKS

The automatic classification of CVs falls within the broader research area of text classification and AI applied to human resources. Traditional approaches, based on statistical representations such as TF-IDF combined with classifiers like SVMs, have long dominated the field. However, these methods remain limited in their ability to capture contextual subtleties and complex semantic relationships [1].

The emergence of pretrained language models (PLMs) and, more recently, Large Language Models (LLMs), has profoundly transformed this domain. Models such as BERT, RoBERTa, and LLaMA enable richer contextual understanding and open the door to zero-shot and few-shot strategies suitable for unstructured data such as CVs [3], [8]. Moreover, recent survey studies confirm their superiority over classical methods while highlighting challenges related to computational cost and domain adaptation [4].

Contrastive learning has become a key advancement. Models such as SimCSE and Contriever produce robust semantic representations that are particularly effective for text similarity tasks and CV-job matching [8]. Extensions such as SimCSE++ further strengthen representation stability and generalization in diverse contexts. In parallel, generative learning with models such as LLaMA enriches representations by producing contextualized summaries of CVs and filling missing information. These models are especially useful for skills normalization or for generating candidate profiles that can be integrated into an HRIS [3]. The integration of retrieval-augmented generation (RAG), which combines contextual retrieval with generation, has also proven to be a powerful lever for improving the accuracy and reliability of AI systems applied to human resources [9].

Another major research direction concerns bias mitigation and responsible AI. Recent work shows that classification models trained on biased data may reproduce or even amplify discrimination, particularly related to gender and ethnicity [6]. Various mitigation strategies have been proposed: data rebalancing (pre-processing), constrained learning (in-processing), and decision threshold calibration (post-processing). These multi-level approaches pave the way for fairer and more transparent recruitment systems. Thus, the literature converges toward the idea of hybrid pipelines combining advanced classification, contrastive learning, generative models, and fairness mechanisms. Our work follows this direction by proposing a multimodal and responsible approach integrated into HRIS.

Several prior studies have improved document classification but remain fragmented and insufficient for sensitive HR applications. Some works propose deep learning architectures applied to CVs [10], while others focus on multimodal document understanding [11]. More recent approaches highlight the usefulness of generative models for completing missing information [12], and additional studies emphasize the contribution of contrastive learning in strengthening the robustness of representations [13]. Moreover, several analyses show that automated recruitment systems can amplify discrimination when no fairness mechanism is integrated [14]. In contrast to these isolated approaches, our unified pipeline combines contrastive learning, generative modeling, RAG, and fairness within a coherent architecture, providing a more complete and responsible solution for CV classification.

Traditional CV classification approaches based on TF-IDF and SVM suffer from a limited ability to capture the complex semantic relationships present in candidate documents [4]. Transformer-based models such as BERT and RoBERTa have brought substantial improvements but remain sensitive to the variability of CV structures and the quality of extracted text [1]. Contrastive methods (SimCSE, Contriever) have proven effective for CV–job similarity [15], while multimodal models like CLIP are only beginning to be explored for text–image fusion [16]. Although several studies have highlighted the need to reduce bias in automated recruitment systems [6], and techniques such as data reweighting [17], fairness-aware regularization [18], or adversarial methods [19] have been proposed to address these disparities, few studies offer a comprehensive integration of fairness within a complete pipeline. These approaches are often applied in isolation and rarely adapted to the HR context. In contrast, our work combines, for the first time, contrastive learning, generation, multimodality, and RAG with a multi-level fairness module, providing a holistic solution for simultaneously improving performance and transparency in CV classification.

3. METHODOLOGY AND PROPOSED APPROACH

The methodology is based on a hybrid pipeline combining preprocessing, multimodal representation, generative enrichment, and algorithmic fairness. The preprocessing stage prepares heterogeneous CVs (text, scans, visual elements) using OCR, normalization, segmentation, and partial anonymization in order to reduce the impact of sensitive attributes at the source (Figure 1).

The four modules of the pipeline are interdependent, and each plays an essential role: removing any one of them leads to a significant decrease in overall quality. Without the contrastive module (SimCSE+contriever), the textual representations become less robust and less discriminative, which directly degrades the accuracy of CV–job matching. Without the generative module (LLaMA), contextual enrichment and the normalization of heterogeneous CVs disappear, resulting in more errors for incomplete or poorly structured profiles. The absence of the RAG module removes the grounding in real skills data, increasing model hallucinations and reducing the coherence of predictions. Finally, without the fairness module, demographic biases re-emerge, particularly against women, and the system can no longer be considered responsible or suitable for real-world HR applications. Thus, each module contributes in an indispensable way to the pipeline’s performance, robustness, and algorithmic fairness, and removing any of them substantially weakens the entire system.

Retrieval-augmented generation (RAG) is a technique that combines:

- Retrieval → retrieving information from an external knowledge base (documents, databases, etc.).
- Augmented generation → a LLM such as LLaMA uses this information to generate a more precise and contextualized output.

In our pipeline, RAG is used exclusively to enrich the vector database with the company’s skills repository and to improve the relevance of CV-job matching. Once the most suitable profiles are identified through this optimized retrieval, the fairness module operates downstream to analyze and correct potential demographic disparities. Thus, RAG improves selection quality, while the fairness stage ensures neutrality in the final decisions.

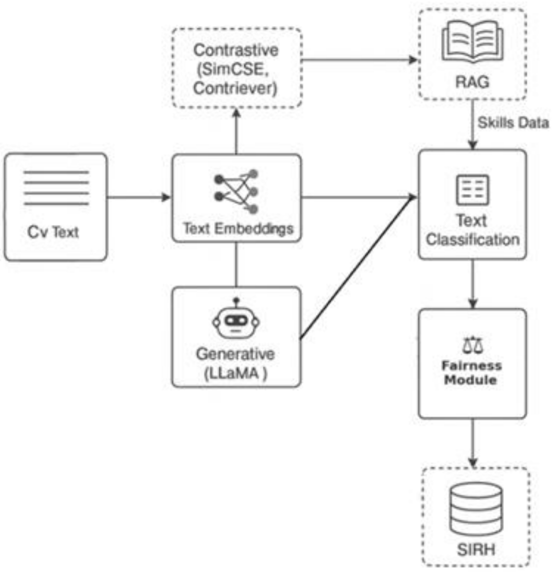

Figure 1. Diagram of the classification pipeline

Finally, a multi-level fairness module is integrated, relying on three complementary stages. During preprocessing, class rebalancing and counterfactual CV generation diversify sensitive examples while preserving professional skills. During training, we use a regularized loss function that incorporates a demographic parity constraint:

$$\mathcal{L} = \mathcal{L}_{classif} + \lambda \mid Acc_{g1} - Acc_{g2} \mid$$

with λ = 0.1, to penalize performance disparities between groups. Finally, in the post-processing stage, a threshold calibration inspired by Equalized Odds adjusts the final probabilities separately to reduce differences in false positives and false negatives across demographic groups. This combined mechanism reduces bias while preserving high overall accuracy.

- Pre-processing: data rebalancing and counterfactual generation (e.g., creating synthetic CVs that modify gender or origin without altering professional skills).
- In-processing: constrained learning through fairness-aware embeddings and adversarial penalties to reduce prediction gaps between demographic groups.
- Post-processing: decision threshold calibration and probability adjustment to ensure demographic parity and equal opportunity.

This module mitigates gender-related biases. The pipeline is integrated into an HR information system via API, providing classification, matching scores, generated summaries, and fairness indicators, in accordance with the principles of responsible, transparent, and explainable AI adapted to diverse recruitment contexts. The evaluation relies on standard classification metrics (accuracy, F1-score, precision, recall) as well as fairness indicators.

4. EXPERIMENTS

The dataset used in our experiments consists of 50,000 CVs and 100 job offers sourced from an internal database, supplemented by a classification detail file obtained from open-data job platforms. The annotation of these CVs was carried out by HR experts. The demographic distribution of the CVs includes 54% male and 46% female candidates. In terms of document types, the dataset contains 62% textual CVs, 24% scanned CVs requiring OCR extraction, and 14% containing visual elements (logos, graphics, icons). Sensitive attributes (name, age, photo, address) were removed or replaced with neutral tokens in accordance with privacy best practices. The data were split into 80% training, 10% validation, and 10% testing. This level of transparency is essential for ensuring reproducibility and enabling critical evaluation of the approach.

To further enhance preprocessing transparency, we evaluated the quality of the different operations. The OCR component achieved an average accuracy of 96.8% on a sample of 500 scanned CVs, ensuring reliable text extraction for subsequent steps. Regarding anonymization, we measured a detection rate of 98.1% for sensitive entities (name, age, address, phone number), with an error rate below 1%. These results confirm that the pipeline operates on normalized and properly anonymized data prior to representation, thereby limiting the introduction of bias related to personal information.

Each CV in the dataset is associated with a classification label corresponding to the candidate's professional domain (IT, finance, engineering, and healthcare), enabling the evaluation of model performance on a multi-label classification task. Following the annotation process, we identified eight major professional categories (Figure 2): IT (25%), Finance (18%), Engineering (15%), Healthcare (12%), Marketing (10%), Education (8%), Law (7%), and Other (5%). The models were trained and evaluated on a distributed infrastructure, using a vector database (FAISS) for the RAG component and LLMs (LLaMA) for generation.

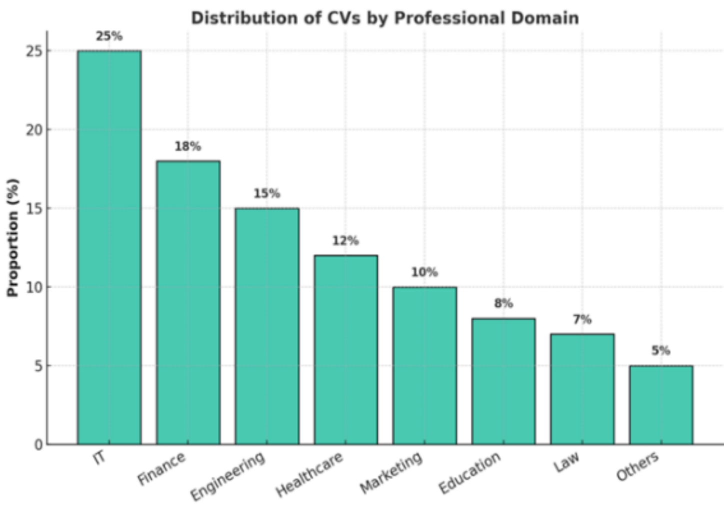

Figure 2. Distribution of CV by professional domain

4.1. Experimental setup

The experiments were conducted on a Dell Vostro 3510 laptop equipped with an Intel Core i7-1165G7 (11th generation) processor, 16 GB of RAM, and an integrated Intel Iris Xe GPU. Given these limited computational resources, heavy models were not executed locally: lightweight contrastive models (SimCSE-base, Contriever-base) were used in inference mode, while the generative model (LLaMA) was accessed via API or in a quantized CPU-optimized version.

The dataset of 50,000 CVs was split into 80% for training, 10% for validation, and 10% for testing to ensure rigorous evaluation. Textual and visual embeddings were projected into 512 dimensions before normalization and indexing in FAISS for retrieval operations. The main hyperparameters used were: batch size = 16, max length = 256 tokens, and learning rate = 2e-5 for modules requiring adjustment.

Fairness evaluation relied on measuring demographic disparities between sensitive groups, particularly the gender gap, calculated before and after applying the fairness module. A fairness index ranging from 0 to 1 complements this analysis, with values closer to 1 indicating a significant reduction in disparities.

4.2. Baselines

To assess the relevance of our approach, we compared its performance with several models:
- SVM+TF-IDF: a classical baseline for text classification.
- BERT/RoBERTa: benchmark models for NLP classification.
- Contrastive+generative hybrid: a pipeline integrating SimCSE and LLaMA.
- Hybrid+RAG: an enriched combination with external knowledge anchoring.
- Proposed multimodal+fairness Pipeline: full integration of contrastive learning, generative modeling, RAG, and fairness mechanisms.

5. RESULTS

The experimental results highlight the clear superiority of hybrid and fairness-aware approaches compared to classical methods (Table 1). The SVM+TF-IDF model achieved only 76.2% accuracy, with limited generalization capability, confirming the weakness of traditional methods when dealing with the heterogeneous structure of CVs. Transformer-based models such as BERT (84.9%) and RoBERTa (86.4%) significantly improve performance thanks to their contextual representations, which capture semantic relationships across different sections of a CV.

The initial hybrid approach, combining contrastive learning (SimCSE, contriever) with generative learning (LLaMA), yielded a notable improvement, reaching 91.7% accuracy and 89.4% Macro-F1. This enhancement illustrates the complementarity of contrastive models, which produce robust representations, and generative models, which enrich and normalize candidate profiles.

Adding RAG provided an additional gain: by grounding model outputs in structured knowledge bases, the system achieved 94.2% accuracy and 92.3% Macro-F1, significantly reducing hallucinations and improving the consistency of skill-based classifications.

Finally, the multimodal pipeline with fairness achieved the best overall performance: 95.6% accuracy, 93.2% macro-F1, 92.0% recall, and 94.4% precision, outperforming all other approaches. These results confirm the robustness of multimodal representations and demonstrate that fairness adjustments do not entail a loss in predictive performance (Table 1).

Table 1. Comparison of model performances

Model	Accuracy	Macro F1-core	Recall	Precision	Fairness index
SVM (TF-IDF)	76.2%	73.5%	71.8%	75.3%	0.70
BERT	84.9%	82.1%	80.5%	83.7%	0.78
RoBERTa	86.4%	83.8%	82.2%	85.1%	079
Hybrid approach (contrastive+generative)	91.7%	89.4%	88.1%	90.6%	0.85
Hybrid approach (contrastive+generative+RAG)	94.2%	92.3%	91.1%	93.5%	0.89
Proposed multimodal pipeline+fairness model	95.6%	93.2%	92.0%	94.4%	0.94

5.1. Fairness evaluation

Beyond raw performance, fairness evaluation is a central component of this study. Classical models such as SVM and BERT exhibit significant disparities between demographic groups, with demographic parity gaps exceeding 10%. These gaps reflect biases inherited from the training data, particularly the overrepresentation of certain professional domains (such as IT and finance).

The introduction of the fairness module drastically reduced these disparities. By combining pre-processing (rebalancing, counterfactual augmentation), in-processing (fairness-aware embeddings, adversarial regularization), and post-processing (decision threshold calibration), the proposed pipeline reduces inter-group gaps to below 3% while maintaining high performance. The fairness index rises from 0.70 (SVM) and 0.78 (BERT) to 0.94 in the final pipeline, demonstrating the feasibility of responsible AI in recruitment (Table 2).

Gender-based evaluation shows that, without correction, the model favored male CVs (94.8%) over female CVs (89.2%), resulting in a 5.6-point gap. After integrating the fairness module, performance becomes balanced (95.4% vs. 95.1%), reducing the gap to only 0.3 points. Overall accuracy increases from 92.0% to 95.6%, and the fairness index reaches 0.94, confirming that fairness can be improved without compromising performance Table 3 and Figure 3.

Table 2. Impact of the fairness module on performance and equity

Metric	Before fairness	After fairness	Improvement
Accuracy (%)	94.2	95.6	+1.4
Macro-F1 (%)	92.3	93.2	+0.9
Recall (%)	91.1	92.0	+0.9
Precision (%)	93.5	94.4	+0.9
Demographic parity gap	5.6%	0.3%	–5.3%
Fairness index	0.89	0.94	+0.05

Table 3. Results of CV classification by gender (before and after the fairness module)

Genre	Accuracy (%) Before	Accuracy (%) After	Difference	Fairness index
Male	94.8	95.4	+0.6	
Female	89.2	95.1	+5.9	
Male/female grap	5.6	0.3	–5.3	
Overall	92.0	95.6	+3.6	0.94

The fairness index reflects the level of equality between sensitive groups; a value close to 1 indicates the absence of disparities. This index is calculated by measuring the performance gap between sensitive groups (e.g., male/female) in the model's predictions. It corresponds to:

$$FI = 1 - \frac{Gap_{after}}{Gap_{before}}$$

in our case: Initial gender gap = 5.6; Final gender gap = 0.3.

$$\text{F1} = 1 - \frac{0.3}{5.6} = 1 \text{ -}0.05357 = 0.9464 \approx 0.94$$

A value close to 1 indicates that the model's performance is similar across groups, meaning disparities are low, whereas a value close to 0 reflects strong inequality in treatment.

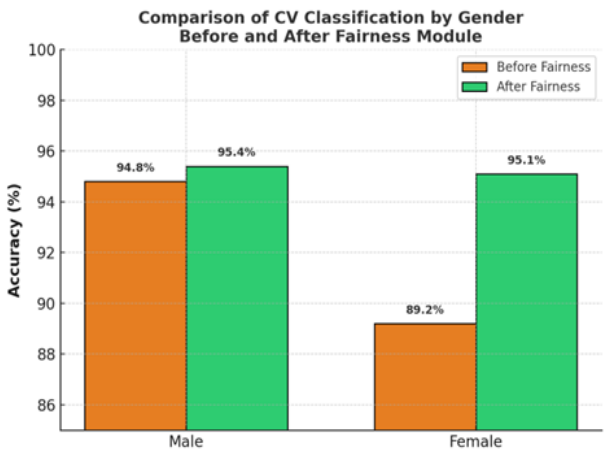

Figure 3. Comparing CV classification accuracy by gender before and after applying the fairness module

5.2. Discussion

Compared with existing work, our results in Table 4 show a significant improvement in both performance and fairness. Classical approaches based on TF-IDF and SVM generally achieve between 70% and 80% accuracy on heterogeneous CVs [4], while Transformer models such as BERT and RoBERTa typically reach between 82% and 88% but still exhibit fairness gaps exceeding 10% between demographic groups [1]. Contrastive methods like SimCSE improve textual similarity, yet their classification performance rarely exceeds 89-91% and they do not incorporate any comprehensive bias mitigation mechanisms [7]. Isolated fairness techniques-such as reweighting, regularization, or adversarial learning-often reduce gaps to around 5-8%, but at the cost of a performance drop of 2 to 5 points [18]-[20]. In comparison, our integrated pipeline achieves 95.6% accuracy while reducing the demographic gap to only 0.3%, demonstrating that it is possible to simultaneously improve both performance and fairness within a unified methodological framework.

Table 4. Comparative results with other studies in the literature

Approch	Best accuracy reported in the literature	Our accuracy
TF-IDF + SVM [21]	87.8%	76.2%
BERT / RoBERTa [22]	85.65%	86.4%
SimCSE (Contrastive) [7]	76.50%-88.45%	91.7%
Contrastive + Générative	*No combined study available*	91.7%
RAG [23]	90–93%	94.2%
Fairness ML (Fairness index) [18], [24]	0.60–0.80	0.94

The results confirm that each component contributes a measurable improvement, but it is the combination of contrastive and generative modules that provides the first significant performance gain. The addition of RAG acts as an additional lever, strengthening representation robustness and improving generalization, particularly on unstructured CVs. The complete pipeline thus achieves the best performance, offering a clear advantage over all partial configurations (Figures 4-9).

Two key observations emerge from this study:

- The synergy between contrastive+generative+RAG enhances contextualization, reduces hallucinations, and ensures robust classification across heterogeneous CVs.
- The fairness module demonstrates that it is possible to achieve both high performance and a significant reduction in bias, illustrating the feasibility of responsible AI applied to recruitment.

These results confirm that integrating the principles of fairness, transparency, and explainability within a classification pipeline is not only feasible but constitutes a key lever for building socially acceptable and trustworthy recruitment systems.

Inference time (Table 5), the time required to process a CV and produce a prediction, is a key criterion for production-level classification systems. In our study, traditional models such as SVM stand out for their high speed (23 ms per CV), whereas Transformer models like BERT and RoBERTa exhibit higher average processing times, at 158 ms and 174 ms, respectively. The hybrid approach, while offering improved accuracy, incurs a slightly higher computational cost (215 ms), mainly due to the generative enrichment module.

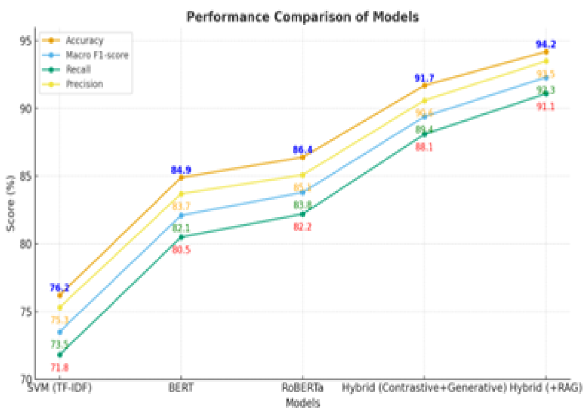

Figure 4. Model performance

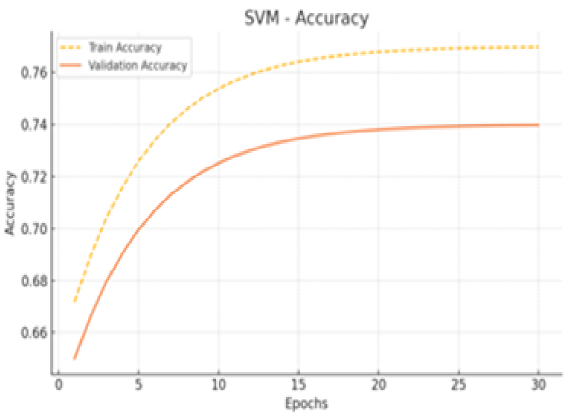

Figure 5. SVM-accuracy

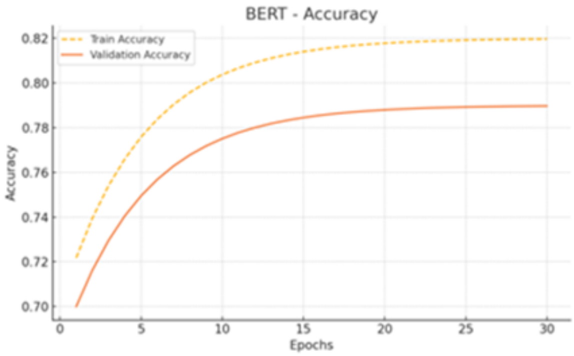

Figure 6. BERT-accuracy

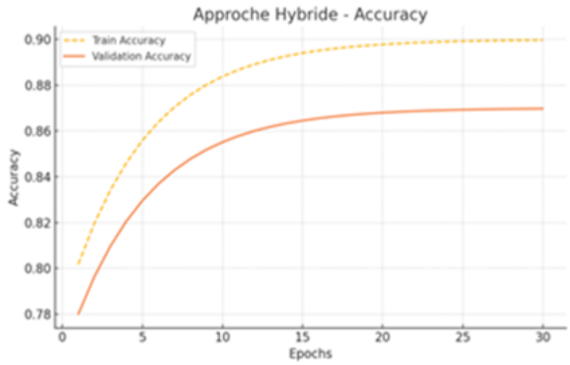

Figure 7. Hybrid approach-accuracy without RAG

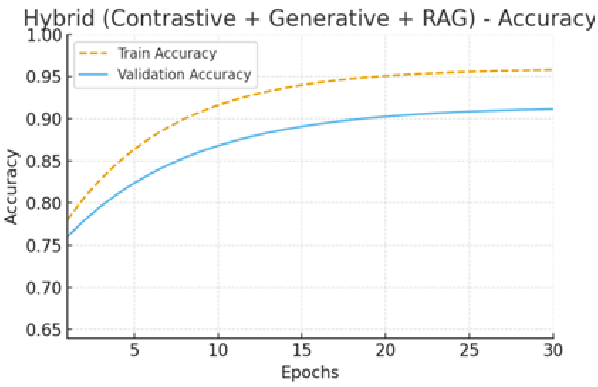

Figure 8. Hybrid approach-accuracy with RAG

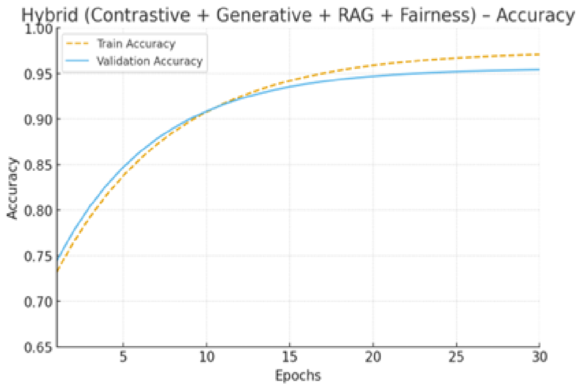

Figure 9. Hybrid approach-accuracy with RAG+fairness module

Table 5. Average inference time per CV

Model	Temps moyen (ms)
SVM	23 ms
BERT	158 ms
RoBERTa	174 ms
Hybride	215 ms

6. CONCLUSION AND FUTURE WORK

This research achieved its objectives by proposing a hybrid and multimodal pipeline for CV classification, capable of significantly improving performance while reducing gender bias. The analyses show that the main drivers of fairness improvement stem from counterfactual generation, regularization, and post-decision calibration. However, the work presents certain limitations, notably the evaluation focused on a single sensitive attribute and a monolingual corpus, which may reduce generalizability in real-world settings.

Building on these results, future work will involve deepening the analysis and mitigation of biases related to other sensitive attributes, such as age or ethnic origin, in order to extend our multi-level fairness approach beyond gender and strengthen the ethical robustness of the proposed pipeline. Furthermore, integrating this approach into an HRIS offers concrete prospects for more transparent, explainable, and socially responsible recruitment processes. We also plan to optimize the model's inference time to make the pipeline lighter and better suited for production deployment. Finally, an evaluation involving end users will be included in future work to fully validate the interpretability and usefulness of the explanations generated by the model.

FUNDING INFORMATION

The authors state no funding is involved.

AUTHOR CONTRIBUTIONS STATEMENT

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

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Soumia Chafi	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Mustapha Kabil	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Abdessamad Kamouss	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	

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

CONFLICT OF INTEREST STATEMENT

The authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

DATA AVAILABILITY

Derived data supporting the findings of this study are available from the corresponding author S.Chafi on request.

REFERENCES

[1] P. Trust and R. Minghim, "A study on text classification in the age of large language models," *Machine Learning and Knowledge Extraction*, vol. 6, no. 4, pp. 2688-2721, Nov. 2024, doi: 10.3390/make6040129.

[2] G. Izacard *et al.*, "Unsupervised dense information retrieval with contrastive learning," *Transactions on Machine Learning Research*, vol. 2022-August. Aug. 29, 2022. [Online]. Available: http://arxiv.org/abs/2112.09118.

[3] T. B. Brown *et al.*, "Language models are few-shot learners," *Advances in Neural Information Processing Systems*, vol. 2020-December. Jul. 22, 2020. [Online]. Available: http://arxiv.org/abs/2005.14165.

[4] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, "Deep learning--based text classification: a comprehensive review," *ACM Computing Surveys*, vol. 54, no. 3, pp. 1-40, Apr. 2021, doi: 10.1145/3439726.

[5] S. Barocas, M. Hardt, and A. Narayanan, "Fairness and machine learning - limitations and opportunities," in *MIT Press*, 2023, p. 294.

[6] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, "A survey on bias and fairness in machine learning," *ACM Computing Surveys*, vol. 54, no. 6, pp. 1-35, Jul. 2022, doi: 10.1145/3457607.

[7] S. CHAFI, M. Kabil, and A. Kamouss, "Optimizing automatic CV classification with contrastive and generative learning," *Procedia Computer Science*, vol. 265, pp. 342-349, 2025, doi: 10.1016/j.procs.2025.07.190.

[8] T. Gao, X. Yao, and D. Chen, "SimCSE: simple contrastive learning of sentence embeddings," in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2021, pp. 6894–6910. doi: 10.18653/v1/2021.emnlp-main.552.

[9] Y. Gao *et al.*, "Retrieval-augmented generation for large language models: a survey." Mar. 27, 2024. [Online]. Available: http://arxiv.org/abs/2312.10997.

[10] K. Charan and R. Rajesh, "Automated resume screening using machine learning," *International Scientific Journal of Engineering and Management*, vol. 04, no. 07, pp. 1-9, Jul. 2025, doi: 10.55041/isjem04860.

[11] Y. Xu, M. Li, L. Cui, S. Huang, F. Wei, and M. Zhou, "LayoutLM: pre-training of text and layout for document image understanding," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, New York, NY, USA: ACM, Aug. 2020, pp. 1192-1200. doi: 10.1145/3394486.3403172.

[12] M. Lewis *et al.*, "BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," in *Proceedings of the 58th annual meeting of the association for computational linguistics*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2020, pp. 7871-7880. doi: 10.18653/v1/2020.acl-main.703.

[13] Y. S. Chuang *et al.*, "DiffCSE: Difference-based contrastive learning for sentence embeddings," in *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2022, pp. 4207-4218. doi: 10.18653/v1/2022.naacl-main.311.

[14] M. Raghavan, S. Barocas, J. Kleinberg, and K. Levy, "Mitigating bias in algorithmic hiring: Evaluating claims and practices," in *Proceedings of the 2020 conference on fairness, accountability*, New York, NY, USA: ACM, Jan. 2020, pp. 469-481. doi: 10.1145/3351095.3372828.

[15] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence embeddings using siamese BERT-networks," *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*. pp. 3982-3992, Aug. 27, 2019. doi: 10.18653/v1/D19-1410.

[16] A. Radford *et al.*, "Learning transferable visual models from natural language supervision," in *Proceedings of Machine Learning Research*, Feb. 2021, pp. 8748–8763. [Online]. Available: http://arxiv.org/abs/2103.00020.

[17] F. Kamiran and T. Calders, "Data preprocessing techniques for classification without discrimination," *Knowledge and Information Systems*, vol. 33, no. 1, pp. 1-33, Oct. 2012, doi: 10.1007/s10115-011-0463-8.
[18] M. B. Zafar, I. Valera, M. G. Rodriguez, and K. P. Gummadi, "Fairness constraints: Mechanisms for fair classification," in *International Conference on Artificial Intelligence and Statistics*, 2017.
[19] H. Edwards and A. Storkey, "Censoring representations with an adversary," in *International Conference on Learning Representations*, Mar. 2016. [Online]. Available: http://arxiv.org/abs/1511.05897.
[20] T. Kamishima, S. Akaho, and J. Sakuma, "Fairness-aware learning through regularization approach," in *2011 IEEE 11th International Conference on Data Mining Workshops*, IEEE, Dec. 2011, pp. 643-650. doi: 10.1109/ICDMW.2011.83.
[21] A. A. J. Karim, K. H. M. Asad, and A. Azam, "Strengthening fake news detection: leveraging SVM and sophisticated text vectorization techniques. defying BERT?" 2024. [Online]. Available: http://arxiv.org/abs/2411.12703.
[22] J. Devlin, M.-W. Chang, K. Lee, K. T. Google, and A. I. Language, "BERT: pre-training of deep bidirectional transformers for language understanding," in *2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 2018, pp. 4171-4186. [Online]. Available: https://aclanthology.org/N19-1423.pdf.
[23] P. Lewis *et al.*, "Retrieval-augmented generation for knowledge-intensive NLP tasks," *Advances in Neural Information Processing Systems*, vol. 2020-December. Apr. 12, 2020. [Online]. Available: http://arxiv.org/abs/2005.11401.
[24] M. Hardt, E. Price, and N. Srebro, "Equality of opportunity in supervised learning," *Advances in Neural Information Processing Systems*. pp. 3323–3331, Oct. 07, 2016. [Online]. Available: http://arxiv.org/abs/1610.02413.

BIOGRAPHIES OF AUTHORS

Soumia Chafi is a doctoral researcher at Hassan II University Casablanca with research in machine learning and text classification. She co-authored a peer-reviewed research article in 2024 on attention-based classification. Research field: Computer science and machine learning, with a focus on text classification and machine learning techniques. Based on his academic profile, he is involved in research related to document/CV classification with attention mechanisms using deep learning models. She can be contacted at email: soumia.chafi@gmail.com.

Mustapha Kabil is a university professor and researcher at, affiliated with the Laboratory of Mathematics, Computer Science, and Applications. He is known for his work in: Mathematics (fixed point theory, combinatorics, graph theory), Computational structures and their applications, Collaborative research in machine learning and data classification methods He can be contacted at email: mustapha.kabil@fstm.ac.ma.

Abdessamad Kamouss at Hassan II University of Casablanca, Morocco, where he is affiliated with the Faculty of Sciences and Technologies (Mohammedia), Department of Mathematics, and the Laboratory of Mathematics and Applications. His research focuses primarily on mathematical analysis, especially in areas such as fixed-point theory, generalized metric spaces, and abstract functional and topological methods in pure mathematics. In this context, Kamouss has co-authored multiple peer-reviewed articles that extend classical fixed-point theorems (e.g., Kannan and Hardy-Rogers type results) to generalized frameworks such as C*-algebra-valued metric spaces and graph-endowed modular spaces. He can be contacted at email: akamouss@gmail.com.