Exploring Fusion Techniques in Multimodal AI-Based Recruitment: Insights from FairCVdb

Paper ID: 68

Swati Swati1, Arjun Roy2 and Eirini Ntoutsi

1Research Institute CODE, University of the Bundeswehr Munich, Germany, Institute of Computer Science, Free University Berlin, Germany

INTRODUCTION & MOTIVATION

Research Objective: Investigate the fairness and bias implications of Fusion Approaches in multimodal AI systems.

Real-World Application: Multimodal AI-based recruitment systems:

- With the increasing use of multimodal decision-making algorithms, there are rising concerns about transparency and discrimination, especially affecting specific social groups.
- Majority of existing fairness-aware learning approaches focus on single modalities (tabular, images, or text), but there is a need to understand bias and discrimination in a multimodal context.

EXPERIMENTAL SETUP

Dataset: FairCVdb1 for fairness study:
- Synthetic research dataset: 24,000 profiles which contain rich multimodal information tailored to assess fairness and bias aspects in AI-driven recruitment algorithms.

Methodology: Recruitment model to predict scores based on candidate resumes, following the methodology from Peña et al. (2023)2.

Evaluation Metrics:
- Mean Absolute Error (MAE) to measure prediction error.
- Kullback-Leibler (KL) divergence to measure differences between demographic distributions.

RESULTS & CONCLUSIONS

Multimodal Fusion Strategies: Early and Late:

Early Fusion (Feature-Level Fusion):
- Occurs at beginning, typically before the data is fed into a neural network.
- Advantageous when model-scale relationships are simple.

Late Fusion (Classifier-Level Fusion):
- Occurs at the final decision-making stage, after each modality has been processed separately and the decision scores have been calculated.
- Advantageous when modalities exhibit highly distinct data characteristics.

Key Conclusions:
- Fusion techniques play a crucial role in addressing fairness and bias in multimodal AI. Nonetheless, they have the potential to amplify biases from individual modalities, and blindly fusing them may not lead to optimal results.
- Early fusion closely mimics ground truth for both demographics and achieves lowest MAEs by incorporating unique characteristics of each modality effectively. It yields fairer solutions even in the presence of demographic biases.
- Late fusion leads to higher over-generalized mean scores, resulting in higher MAEs.

Future Directions:
- Bias-aware fusion strategies: Mid-fusion may enhance fairness and accuracy by strategically selecting and combining modalities.
- Test the applicability of these findings across diverse datasets and domains beyond hiring for broader impact and relevance.

Ethics statement: Understanding the risks of using simulated or synthetic data is crucial for fairness, transparency, and effectiveness in automated hiring processes.

For code and additional insights, visit: https://github.com/Swati1729/Multimodal-AI-Based-Recruitment-FairCVdb or write to: swati.swati@uni-bw.de

This research work is funded by the European Union under the Horizon Europe MMAMCIH project. Grant Agreement ID: 101073285. Also in part funded by Horizon Europe Project MAMCIH is supported by UKRI grant number 10041914 (Talisman Research LTD). The research is also supported by the EU Horizon Europe project STELARK, Grant Agreement ID: 101070122.