Predictive Modeling on Real Hiring Data: Exploring Salary Stereotypes and Bias Mitigation

Lyudmil Stamenov lyudmil.stamenov@studio.unibo.it

April 3, 2025

Abstract

This project focuses on developing a predictive model using real hiring data from AKKODIS to explore and address salary stereotypes and potential biases in recruitment processes. The objective is to analyze historical hiring data to identify patterns that may indicate bias related to salary offers (RAL) and develop models that promote fairness and transparency. By leveraging machine learning techniques, we aim to detect underlying trends that may not be immediately visible through traditional analysis. The findings from this study can help organizations implement data-driven policies to mitigate bias and ensure equitable salary distribution.

1 Objectives

The primary objectives of this study are:

- Analyze the existing hiring dataset to detect potential biases in salary offers using statistical analysis, exploratory data analysis (EDA), and visualization techniques.
- Explore previous research conducted on the same dataset, focusing on bias detection and mitigation methods such as fairness-aware machine learning and ethical AI frameworks.
- Develop predictive models to identify and quantify salary disparities using different machine learning models or techniques.
- Propose and implement bias mitigation strategies, including reweighting, adversarial debiasing, and post-processing adjustments to ensure fairness.
- Evaluate model fairness and accuracy using relevant metrics such as disparate impact, equalized odds, demographic parity, and fairness-aware performance measures.