LOWERING GENDER BIAS USING SAMPLING TECHNIQUES

Kaan Gögcay, Master of Applied Intelligence, Hogeschool van Amsterdam, Amsterdam, Netherlands

INTRODUCTION

This study explores whether sampling techniques can reduce gender bias in facial age prediction models trained on the MORPH-II dataset. With 86.1% male and 13.9% female images, the dataset's imbalance often leads to unequal model performance across genders.

THE STATISTICS

Random undersampling is a simple yet effective technique for reducing gender bias in age prediction models, maintaining reasonable accuracy while promoting fairness. In contrast, oversampling using external datasets may increase bias if dataset characteristics differ. These findings highlight the importance of dataset consistency and careful selection of mitigation strategies in developing fair facial age prediction models.
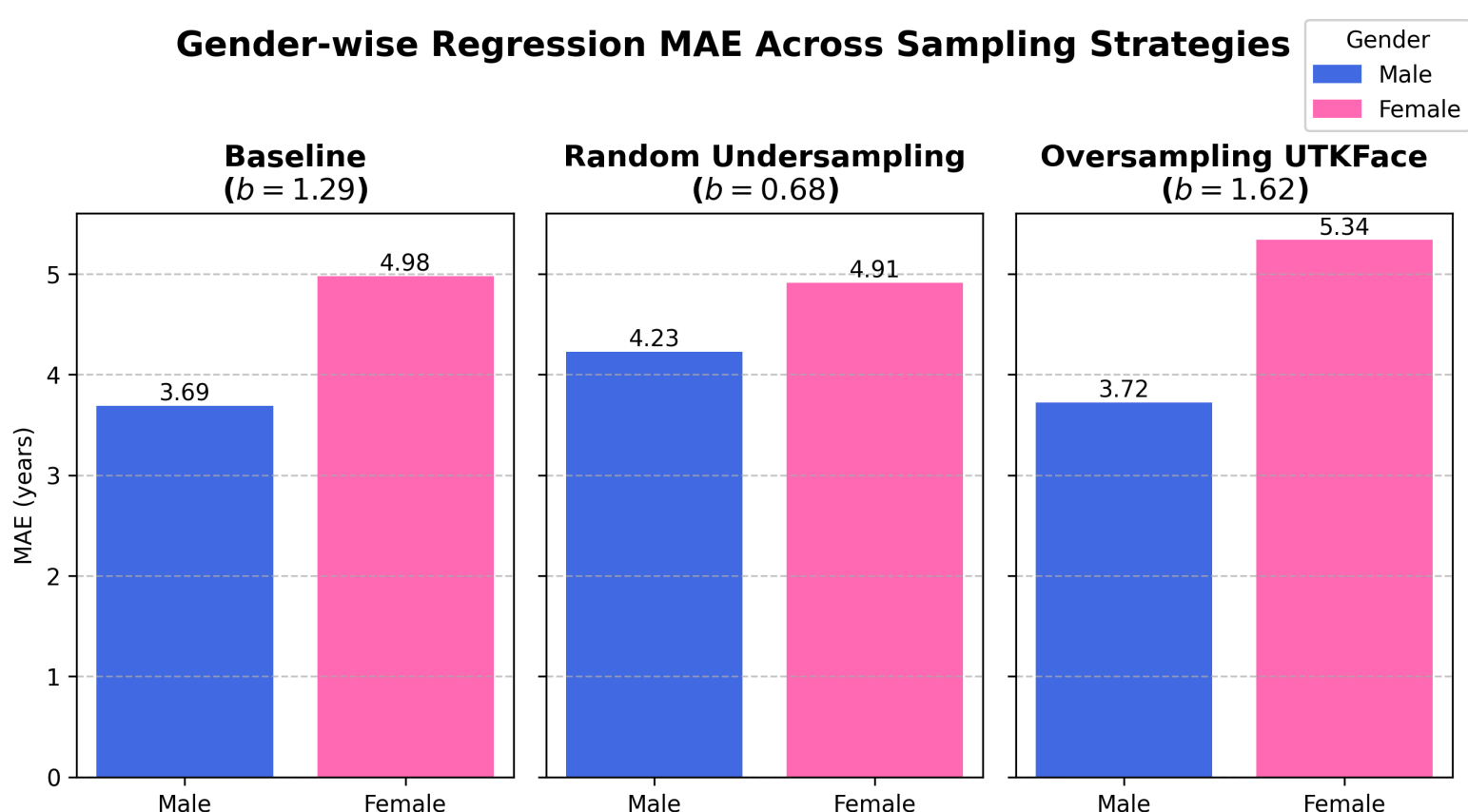
METHODS

Three workflows were designed: baseline (original MORPH-II dataset), random undersampling for balanced gender distribution, and traditional oversampling with female images from UTKFace. Models were trained as classifiers over 62 age classes but evaluated using regression outputs to measure performance and gender bias via Mean Absolute Error (MAE).

RESULTS

The baseline model showed a gender bias of 1.29 MAE. Random undersampling reduced this to 0.68, indicating fairer performance but slightly higher overall error. Oversampling increased bias to 1.62 due to inconsistencies between MORPH-II's standardized mugshots and UTKFace's more variable image conditions.

Gender-wise Regression MAE Across Sampling Strategies

Gender: Male, Female

Baseline ($b = 1.29$) — Male: 3.69, Female: 4.98
Random Undersampling ($b = 0.68$) — Male: 4.23, Female: 4.91
Oversampling UTKFace ($b = 1.62$) — Male: 3.72, Female: 5.34