Fairness in Credit Scoring

Khoza N. Ndlovu N., Mustapha N. Ngcobo S.

Council for Scientific and Industrial Research, South Africa

Introduction

Credit scoring plays a crucial role in determining an individual's creditworthiness and has a significant impact on their financial opportunities. Traditional credit scoring models have long been the standard in evaluating credit risk, relying on historical financial data and statistical algorithms. However, concerns have emerged regarding the fairness and potential biases embedded within these models, leading to a growing need for assessing and implementing fairness in credit scoring.

Fairness in credit scoring refers to the equitable treatment of individuals from diverse backgrounds, regardless of protected characteristics such as race, gender, or age. Biases in credit scoring models can inadvertently lead to discrimination and disparities in access to credit, perpetuating existing social and economic inequalities. Recognizing the importance of fair credit assessment, organizations are increasingly seeking ways to address these biases and promote equal opportunities for all individuals.

The research aims to explore the assessment and implementation of fairness in credit scoring. We will delve into the various methods and techniques used to assess fairness, and the strategies for implementing fair credit scoring models. By understanding the underlying issues and adopting fair credit assessment practices,

Data

Age is considered as the protected attribute contributing to bias during training and unfairness in prediction

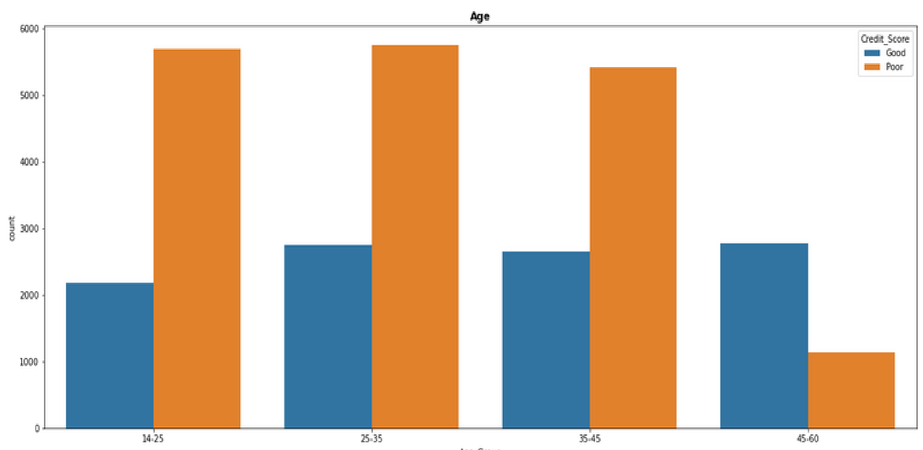

Figure 1. Age Distribution

Based on figure 1, the threshold at the age of 45 years was considered to differentiate between the unprivileged group (45 and under) and the privileged group (over 45) in relation to the target variable (Credit)

Methodology

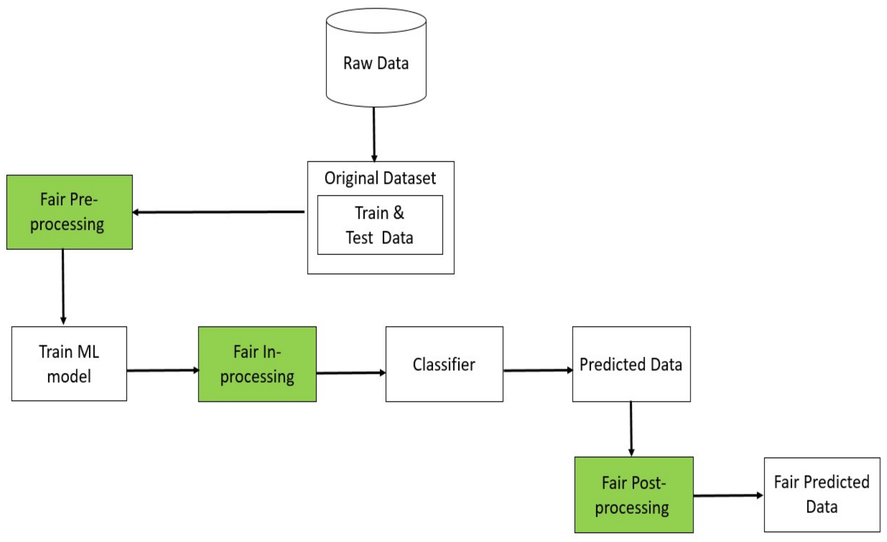

Figure 2. Pipeline of Fairness in ML

Figure 2 shows fairness integration in ML which involves different stages where fairness measures and bias mitigation methods are incorporated in ML. The methods are divided into three stages, pre-processing, in-processing and post-processing methods.

- **Pre-processing techniques** are applied before the training of the classifier to modify the dataset and features in a way that reduces the influence of protected attributes on the learning process.
- **In-processing techniques** integrate bias mitigation directly into the training process of the classifier to create a fairer classifier directly.
- **Post-processing techniques** are applied after the classifier has made predictions, they adjusts the predictions made by the classifier to achieve fairness objectives after the prediction phase.

Model

Table 1. Model Performance Results

Model	Accuracy %
Random Forest	**94.46**
XGBoost	94.37
Gradient Boosting	87.91
Decision Treee	86.80
Neural Network	86.52

Experiment and Results

Bias assessment metrics play a crucial role in credit scoring to identify and quantify potential biases in the predictive models or decision-making processes. These metrics help evaluate whether the models or decisions exhibit unfair or discriminatory behavior towards certain groups based on protected attributes such as age. Here are some common bias assessment metrics used in credit scoring:

Table 2. Bias assessment metrics

Bias Metrics	Protected Attribute	Fairness Interval
Average Odds Difference	Age	(-0.1, 0.1)
Equal Opportunity Difference	Age	(-0.1, 0.1)
Statistical Parity Difference	Age	(-0.1, 0.1)
Disparate Impact	Age	(0.8, 1.2)
Theil Index	Age	≈ 0
False Positive Rate Difference	Age	(-0.1, 0.1)
False Negative Rate Difference	Age	(-0.1, 0.1)

To evaluate the performance of different bias mitigation methods, we used fairness assessment metrics to measure the model's fairness and potential bias showed in Table 3

Table 3. Credit Dataset Results

Mitigation	Technique	AOD	EOD	SPD	DI	TI	FPRD	FNRD	Acc
No mitigation	None	-0.129	-0.0525	-0.448	0.406	0.033	-0.201	0.057	94.48
Reweighting	Pre	-0.127	-0.055	-0.444	0.411	0.034	-0.199	0.055	94.11
Learning Fair Representations	Pre	0.007	0.007	0.005	1.005	0.048	0.008	-0.007	100.0
Disparate Impact Remover	Pre	-0.159	-0.051	-0.432	0.455	0.051	-0.268	0.051	88.82

- For no bias mitigation technique, the results suggest some disparity in the treatment of the privileged group compared to the unprivileged group.
- The results for reweighting show that there are still some disparities in the treatment and outcomes for the privileged group compared to the unprivileged group. The privileged group generally has higher probabilities of favourable outcomes, but there are differences in various fairness metrics.
- For Learning Fair Representations, the results indicate that applying the Learning Fair Representations has helped reduce disparities and promote fairness in the credit scoring classification for the privileged group. The metrics show improvements in terms of reducing disparities in different fairness metrics.
- For Disparate Impact Remover, the results indicate that applying the Disparate Impact Remover has helped reduce disparities and promote fairness in the credit scoring classification for the privileged group. The metrics show improvements in terms of reducing disparities in different fairness metrics.

Conclusion

To address these biases and promote fairness, we explored pre-processing bias mitigation methods, including Reweighing, Learning Fair Representations (LFR), and Disparate Impact Remover. The results showed that each method contributed to reducing disparities and promoting fairness to some extent. Reweighing and Learning Fair Representations helped reduce the level of disparities across multiple fairness metrics, while the Disparate Impact Remover worked towards equalizing the impact of protected attributes on the prediction outcomes.

However, it's important to note that despite the improvements from pre-processing methods, there is still a need for in-processing and post-processing bias mitigation techniques to further enhance fairness in credit scoring which is a continuation of this paper.

nokuthulakhoza489@gmail.com