Fix Skewness To Achieve Fairness - an innovative approach

Or Aharoni, Bar Katz

March 2023

1 Introduction

Our proposed solution of fairness involves bin splitting a continuous protected variable, followed by checking skewness using the IQR and JB tests, and applying a quantile transformer method. This approach is innovative, particularly compared to other methods. Bin splitting allows us to group the data into multiple bins, addressing extreme values that can significantly affect the data's distribution. The IQR and JB tests enable us to detect heavy-tailed distributions and asymmetry, which are common in real-world data. Checking for skewness using multiple tests makes the approach more robust and able to handle a wider range of skewed distributions. Finally, the quantile transformer method is a powerful tool for transforming the distribution of data, particularly compared to other methods such as log or Box-Cox transformations. It is more robust to outliers and can handle a wide range of distributions, including heavy-tailed distributions. Moreover, the quantile transformer method can transform the data into a normal distribution, which is a key assumption in many statistical and machine learning models. This approach could possibly be a significant contribution to the field of fairness in machine learning and has the potential to improve the accuracy and fairness of machine learning models.

2 results

We will begin be stating that in order to best assess our created tool we used a plotting tool to check for which features have a skew in order to select the mostly skewed features for our assessment of the tool. We added the use of

utility function in order to check for causes of why our project does not work as expected. The utility function that we harness is spearman correlation test, which allows us to measure non-linearity - which could result in our method's failure.
Adult Dataset: Here we used our tool and saw that we have a nice skewed variable called fnlwgt (figure 1), which is the final weight assigned by the goverment of number of people of this specific economic state.

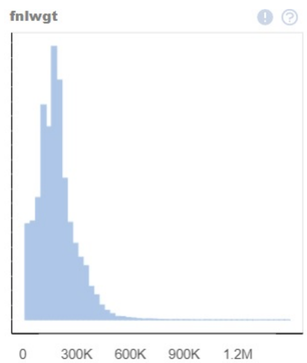
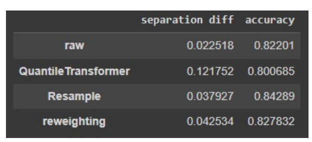


Figure 1: adult data set fnlwgt graph



	separation diff	accuracy
raw	0.022518	0.82201
QuantileTransformer	0.121752	0.800685
Resample	0.037927	0.84289
reweighting	0.042534	0.827832

Figure 2: adult data set fnlwgt results

At figure 2 we can see that our method had poor results compared to other baseline such as resampling and reweighting. We also run spearman correlation and saw that the variables do have a linear relation between them and by doing so we disapproved the idea that our method failed because of non-linearity relation.
Spearman's correlation coefficient: 0.027914179254961
p-value: 6.966135326452799e-10
As the data is quite big and therefore allows the possibility to properly assess the distribution, the only possible outcome that prevented us from succeeding is that the main issue is that there is unmeasured confounding variables that prevents our method from succeeding.

Diabetes Dataset: protected attribute = 'Age' , target variable = 'Outcome'. Here we choose the variable of "age" as we found out that it is skewned (figure 3), as well as that we are quite familiar with the understanding of what "age" means.

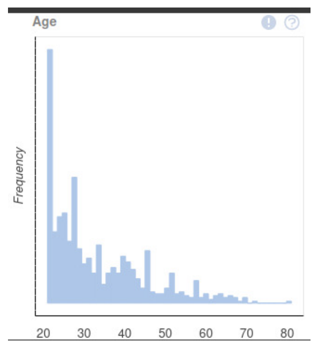


Figure 3: diabetes age graph

	separation diff	accuracy
raw	0.091133	0.779221
QuantileTransformer	0.057065	0.680672
Resample	0.012759	0.722772
reweighting	0.115428	0.713483

Figure 4: diabetes results

At the diabetes dataset we could see that our tool performed better than on the adults dataset (figure 4), the resampling still was ahead of us, yet the reweighting was performing poorly compared to our tool. This gave a rise to an interesting explanation of ours: One advantage of bin splitting and quantile transformation is that they can preserve the original data distribution while achieving fairness, which may be desirable in some cases. Reweighting, on the other hand, can alter the original data distribution, which may affect the performance of models.

Spearman's correlation coefficient: 0.30904026356718634.
p-value: 1.852974424537641e-18.

Heart disease Dataset: This was probably the worst outcome of all for our code. as described below at figure 5:

	separation diff	accuracy
raw	0.160784	0.816667
QuantileTransformer	0.555556	0.711111
Resample	0.1	0.820513
reweighting	0.294118	0.854839

Figure 5: heart disease results

As we also here managed to disregard any beliefs of non-linearity between the protected attribute (oldpeak) and the target variable (condition). Our only explanation from researching the data was that the size of the dataset was a problem, it's possible that there simply wasn't enough data to support the preprocessing methods we used. This can lead to unstable or inconsistent results, as seen by figure 5. A supporting claim is that resample actually performed really well, which supports the claim that there was too less data.

3 Problem description

Our solution aims to improve the data preprocessing stage of the machine learning pipeline, specifically addressing the issue of handling skewed data in the context of fairness. Skewed data can lead to biased models that unfairly disadvantage certain groups of individuals, which is a critical concern in many applications of machine learning.

The problem with traditional approaches to handling skewed data, such as log transformation or removing outliers, is that they can lead to further bias or loss of information in the data. Additionally, these methods may not be effective in handling highly skewed distributions, which are common in real-world datasets.

4 Solution Description

Our solution to address the problem of skewed distributions in the protected attribute involves multiple steps. First, we check the skewness of the protected attribute using the skewness measure. If the absolute value of the skewness exceeds a certain threshold, we apply additional checks using the interquartile range (IQR) and Jarque-Bera (JB) test to ensure the distribution is heavy-tailed.

If these conditions are met, we then perform bin splitting on the continuous variable, dividing it into three quantile-based bins to balance the distribution.

We chose to split the continuous variable into three quantiles because it provides a simple and intuitive way to transform a continuous variable into a categorical variable. Using three bins also strikes a balance between oversimplification and overcomplication. Using too few bins may not capture enough information in the distribution, while using too many bins may result in overfitting or increase the complexity of the resulting model. Additionally, splitting the variable into three quantiles allows for easy interpretation and comparison of the resulting categories, as they can be labeled as "low," "medium," and "high" based on their respective quantiles.

Next, we apply the QuantileTransformer to normalize the data, which can handle the heavy-tailed distribution of the protected attribute. Specifically, we use the 'normal' output distribution parameter to transform the data to have a normal distribution. The normalization process is applied to the minority group (i.e., subgroup of unfairness) and the majority group separately to prevent any bias towards the majority group.

We thought about using Box-Cox but decided to not use it since the data will not always be positive.

Overall, our solution is designed to handle skewed distributions of the protected attribute and produce a more balanced dataset without bias towards any specific group. We chose this approach due to its effectiveness in handling heavy-tailed distributions and its ability to normalize data without introducing bias towards any group.

5 Related Work

Here we will discuss a number of articles - some academic and some are not, as for why we choose to do our project the way we did. the first one is: "feature transformation in machine learning why when and what" - The article discusses the use of feature transformation to address issues such as extreme values, non-linear relationships, numerical stability, and robustness. Quantile transformation is one method that the paper suggested that can be used to transform features and reduce variability due to extreme values. By transforming the feature into a normal distribution, the transformed feature can be better utilized in machine learning algorithms.

"Algorithmic fairness in computational medicine" The paper explores the issue of fairness in predictive models and proposes a method to address it. Specifically, they use a fairness metric based on the equal opportunity principle (Which we then also choose to use) and propose a method to adjust the predictions of the model to improve fairness while maintaining predictive accuracy. Their approach involves modifying the predicted probabilities for the positive class based on the difference in true positive rates between the protected and non-protected groups.

While the paper proposes a novel approach to address fairness in predictive models, we believe that our approach of measuring skewness and using quantileTransformer to match the distribution of the majority and minority could also be effective. By matching the distribution of the protected attribute across subgroups, we can ensure that each subgroup receives equal treatment and avoid potential bias. Additionally, this approach does not require modifying the predicted probabilities or the model itself, making it a simpler and more transparent solution.

There are quite a few online tools we found to measure and mitigate fairness, such as: Fairlearn, Themis-ML and IBM AI Fairness 360. After a bit of searching, we found out that none of the tools mentioned above explicitly use skewness as a measure for handling fairness. These tools typically focus on methods such as reweighting, resampling, and adjusting decision thresholds to achieve fairness in machine learning models. However, it is worth noting that some of these methods may indirectly address skewness in the data through their approach to achieving fairness.

6 Conclusion

There are a few reasons why resampling and reweighting techniques were more successful in achieving fairness compared to our tool of bin splitting and quantileTransformer. First, resampling and reweighting techniques can be more effective in dealing with class imbalance, which is a common issue in datasets where one class is significantly underrepresented. Our tool of bin splitting and quantileTransformer does not directly address this issue, which could lead to poorer performance in achieving fairness when dealing with imbalanced datasets.

Overall, while our tool of bin splitting and quantileTransformer can be a useful tool for achieving fairness in certain situations, further research is needed to determine its effectiveness in different contexts and how it compares to other fairness techniques such as resampling and reweighting.