Exercise

Bias Mitigation

Exercises

1. Name 3 models of fairness, explain them briefly. When should which models be used?

2. To understand the data plot

The first 10 datasets and all column names

3. Read data `creditworthiness.csv`

Define ‚IS_DEFAULT' as target

Split the data into training and test data

4. How many women and how many men repay the loan late?

Plot the confusion matrix

5. Define as proteced group ‚GENDER'

6. Define privileged and unprivileged groups

7. Quantify the bias by using

8. Quantify the model bias

Use the `CatBoostClassifier`

Give larger weights to positive datasets: 0.9

Exercises

9. plot the confusion matrix

10. Compare the results of using standard weights with using your weights

11. In which classification phases we can reduce the bias?

12. What are advantages and disadvantages of the bias mitigation methods in

- Preprocessing
- In-processing
- Post-processing?

13. What is the greatest disadvantage of all bias mitigation methods?

14. Use the method Reweighing to reduce the bias

15. Compare the results with and without reweighting

16. Use the In-Processing method Prejudice Remover for bias reduction

17. Compare the results to the Pre-Processing mitigation method

18. Use the Post-Processing Bias reduction method Equalized odds postprocessing

19. Compare the results to the results of your In-Processing method

Bias Mitigation