Lab 10

**Step 1**: Load the given train and test data extracted from the [1994 census bureau database](https://www.census.gov/en.html).

**Step 2:**

* Use the following as input features for training:
  + - * + workclass
        + education\_num
        + age
        + relationship
        + native\_country
        + Occupation
* Use “income\_bracket” as the target variable. We have two class labels:
  + (Income\_bracket > 50k) -> positive label
  + (Income\_bracket <= 50k) -> negative label
* We will evaluate the fairness for the following subgroups: “gender”, “native\_country**”, “**occupation**”.**

**Step 3**: Fit the best KNN classifier for this data, by optimizing for “balanced accuracy”.

**Step 4**: Fit the best SVM classifier for this data, by optimizing for “balanced accuracy”.

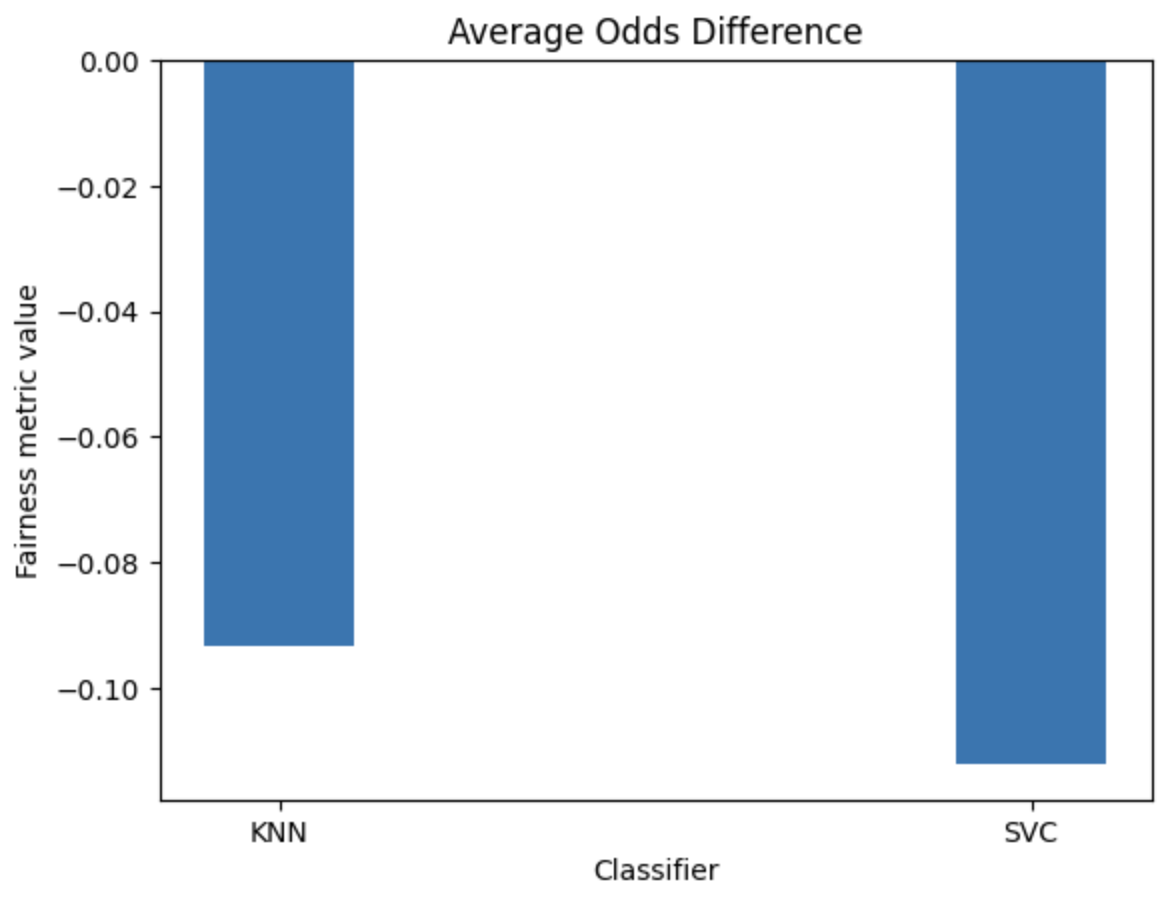
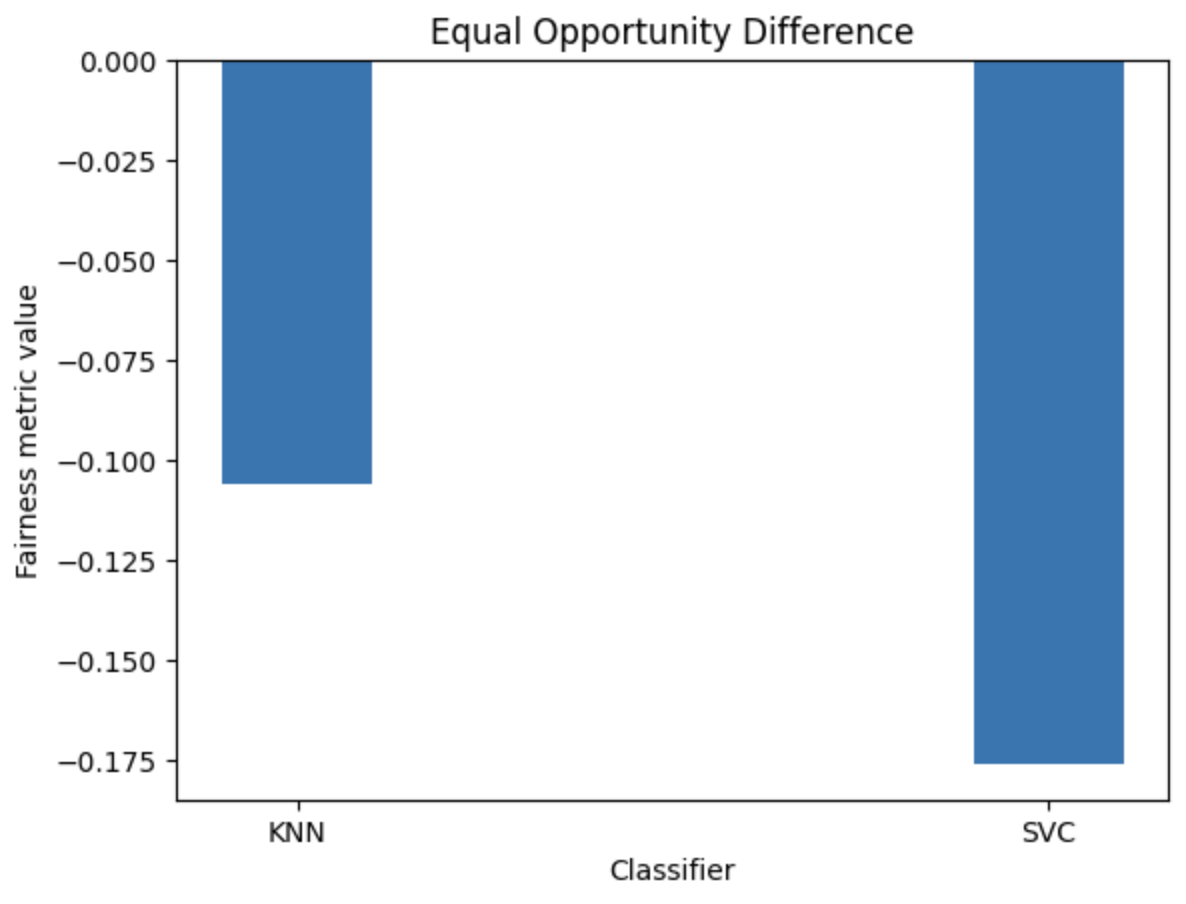
**Step 5**: Compute fairness metrics EOD and AAOD on the test data, for both learnt models.

**Step 6:** Plot bar charts comparing the fairness metrics between the 2 classifiers for “gender”, “native\_country**”, “**occupation**”**

**Step 7**: Answer the following questions for the above data and models:

* Compare the two models in terms of fairness. Which is more fair?
* Which model would you deploy into the real world, and why?
* What real world implications could an unfair model have for this dataset?

Fairness scores for “gender” category



Fairness scores for “occupation”

