Summary of our use of the Aequitas Tool:

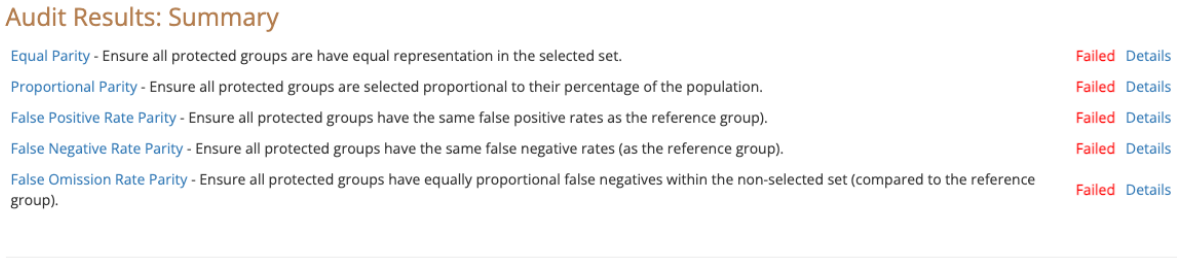
Disparity Intolerance decides the fairness threshold, we used 80% because that was the default value given by Aequitas. We also tested it lower than 80% and more tests passed for each model when we lowered it to 10%, significantly less tests passed when we used 50%, and all tests failed when we increased it from 80%. With the Disparity Intolerance set to 80% if the disparity for the tested group is within 80%-125% the test will pass.

We used the majority group in each attribute as the reference class
Reference Groups:
Gender: Female
Age: 33.50-54.00
Residence type: Urban

Even though our interventions are assistive we will display both parts of the decision tree so we have both false positive and false negative rate parity displayed. We decided to perform all parity tests except for false discovery rate because our models are not tied to a hypothesis so there is no null hypothesis to be tested in the false discovery rate.

What follows is the overall test results for the 3 models recreated from the research article, the results for the 2 models we created after only the initial bias mitigation was employed, and finally the results for the 2 models we created after all of the bias mitigation.

Neural Network (research article):

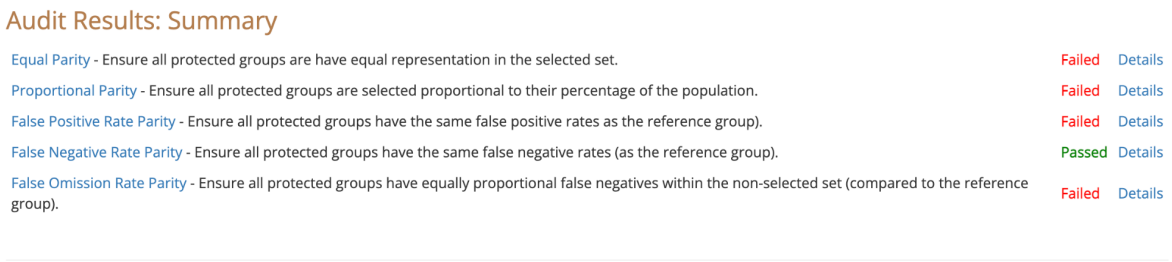


For each attribute:

- Equal Parity:
    - Gender: Failed
        - Male with 0.79x
    - Age: Failed
        - 54-68 with 1.53x
        - 68-82 with 2.47x
    - Residence Type: Failed
        - Rural with 0.56x
- Proportional Parity:
    - Gender: Passed
    - Age: Failed
        - 54-68 with 1.67x
        - 68-82 with 2.56x
    - Residence Type: Failed
        - Rural with 0.57x
- False Positive Rate Parity:
    - Gender: Failed
        - Male with 0.64x
    - Age: Failed
        - 54-68 with 1.52x
        - 68-82 with 1.40x
    - Residence Type: Failed
        - Rural with 0.56x
- False Negative Rate Parity:
    - Gender: Failed
        - Male with 1.29x
    - Age: Failed
        - 54-68 with 0.73x
        - 68-82 with 0.41x

        - Residence Type: Failed
            - Rural with 1.50x
- False Omission Rate Parity:
    - Gender: Failed
        - Male with 2.00x
    - Age: Failed
        - 54-68 with 5.49x
        - 68-82 with 11.43x
    - Residence Type: Failed
        - Rural with 0.78x

Random Forest (research article):

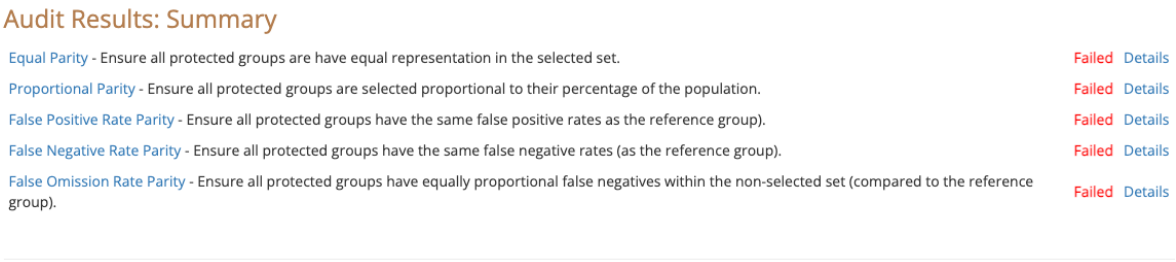


For each attribute:

- Equal Parity:
    - Gender: Passed
    - Age: Failed
        - 54-68 with 10.00x
        - 68-82 with 10.00x
    - Residence Type: Passed
- Proportional Parity:
    - Gender: Passed
    - Age: Failed
        - 54-68 with 10.00x
        - 68-82 with 10.00x
    - Residence Type: Passed
- False Positive Rate Parity:
    - Gender: Failed
        - Male with 1.35x
    - Age: Failed

        - 54-68 with 10.00x
        - 68-82 with 10.00x
    - Residence Type: Passed
- False Negative Rate Parity:
    - Gender: Passed
    - Age: Passed
    - Residence Type: Passed
- False Omission Rate Parity:
    - Gender: Passed
    - Age: Failed
        - 54-68 with 10.00x
        - 68-82 with 10.00x
    - Residence Type: Passed

Decision Tree (research article):



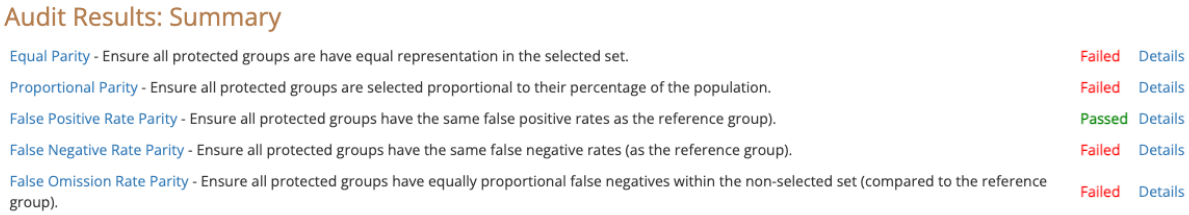
For each attribute:

- Equal Parity:
    - Gender: Failed
        - Male with 0.70x
    - Age: Failed
        - 54-68 with 4.14x
        - 68-82 with 5.43x
    - Residence Type: Passed
- Proportional Parity:
    - Gender: Passed
    - Age: Failed
        - 54-68 with 4.00x
        - 68-82 with 5.34x
    - Residence Type: Passed

- False Positive Rate Parity:
    - Gender: Failed
        - Male with 0.69x
    - Age: Failed
        - 54-68 with 4.19x
        - 68-82 with 4.07x
    - Residence Type: Passed
- False Negative Rate Parity:
    - Gender: Passed
    - Age: Failed
        - 54-68 with 0.56x
        - 68-82 with 0.32x
    - Residence Type: Failed
        - Rural with 1.31x
- False Omission Rate Parity:
    - Gender: Failed
        - Male with 1.31x
    - Age: Failed
        - 54-68 with 23.03x
        - 68-82 with 36.43x
    - Residence Type: Failed
        - Rural with 1.74x

Comparing the bias between the 3 research article models:
Of the three, the random forest had the least bias present when run through the Aequitas tool. It passed a lot of the tests for the gender and residence type attributes but even the least biased model consistently failed with the age attribute. The neural network and decision tree models consistently failed with all 3 tested attributes. In general across the three models and all the tests ran, the residence type attribute was the least biased as it least frequently failed tests, while the gender attribute was moderately more biased, and the age attribute was by far the most biased of the three as it failed nearly every test.

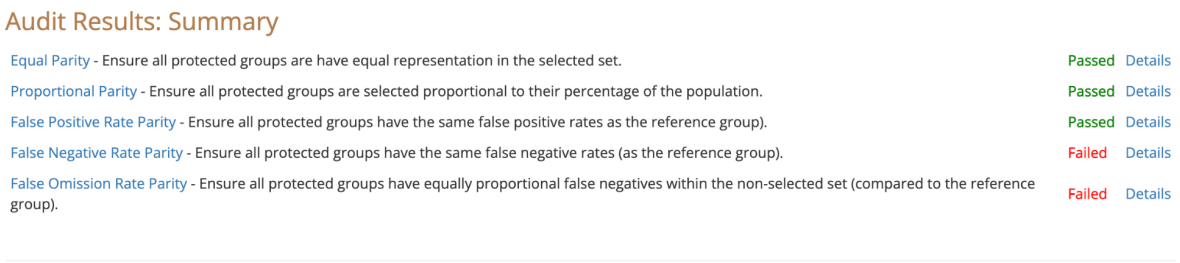
KNN Model (after only the initial bias mitigation):



For each attribute:

- Equal Parity:
    - Gender: Passed
    - Age: Failed
        - 0-33 with 1.20x
    - Residence Type: Passed
- Proportional Parity:
    - Gender: Passed
    - Age: Failed
        - 0-33 with 1.26x
    - Residence Type: Passed
- False Positive Rate Parity:
    - Gender: Passed
    - Age: Passed
    - Residence Type: Passed
- False Negative Rate Parity:
    - Gender: Passed
    - Age: Failed
        - 0-33 with 1.58x
    - Residence Type: Passed
- False Omission Rate Parity:
    - Gender: Passed
    - Age: Failed
        - 54-68 with 2.75x
        - 0-33 with 2.16x
    - Residence Type: Passed

Random Forest Model (after only the initial bias mitigation):



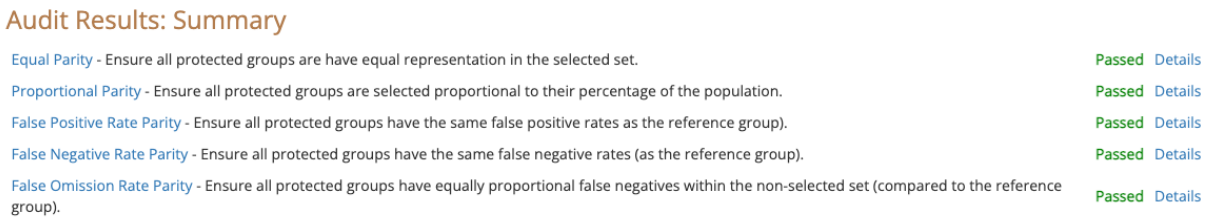
For each attribute:

- Equal Parity:
    - Gender: Passed
    - Age: Passed
    - Residence Type: Passed
- Proportional Parity:
    - Gender: Passed
    - Age: Passed
    - Residence Type: Passed
- False Positive Rate Parity:
    - Gender: Passed
    - Age: Passed
    - Residence Type: Passed
- False Negative Rate Parity:
    - Gender: Passed
    - Age: Failed
        - 0-33 with 1.31x
    - Residence Type: Passed
- False Omission Rate Parity:
    - Gender: Passed
    - Age: Failed
        - 0-33 with 1.74x
    - Residence Type: Passed

Takeaways after the initial bias mitigation:
The only attribute that was failing any tests was the age attribute, mainly because of the 0-33 age range. We knew that if we could fix this part of the data then we could pass all the tests. This is when we decided to “clean” the data again and make new age range bins that essentially excluded those younger than 25. This limits the scope of the application of our predictor models

but we find that it's a necessary step to take in order to mitigate bias. In the grand scheme of the project, only being applicable to people over 25 when the topic is stroke is not too much of a blow to the overall efficacy in the goal.

KNN Model (with complete bias mitigation):



Random Forest Model (with complete bias mitigation):