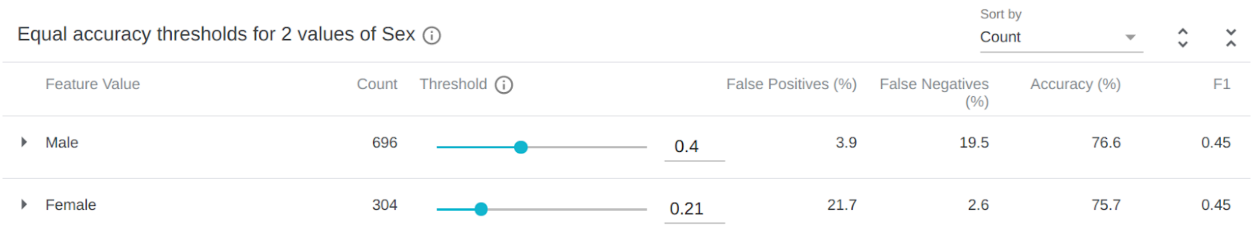
# Forum: Fairness

When optimizing for Equal Accuracy in Google's What-If tool for fairness metrics, we get the results shown below. Take a closer look at the distribution of *inaccurate* predictions for males vs females - What do you notice?

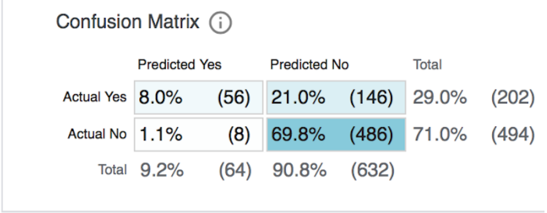


Equal accuracy thresholds for 2 values of Sex

Sort by: Count

| Feature Value | Count | Threshold | False Positives (%) | False Negatives (%) | Accuracy (%) | F1 |
|---|---|---|---|---|---|---|
| Male | 696 | 0.4 | 3.9 | 19.5 | 76.6 | 0.45 |
| Female | 304 | 0.21 | 21.7 | 2.6 | 75.7 | 0.45 |

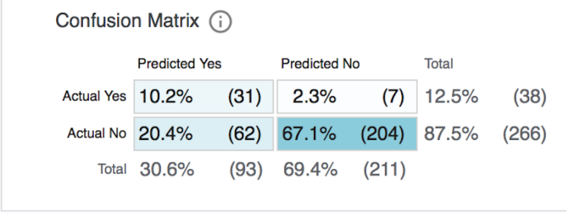
## Confusion Matrix for Males



Confusion Matrix

| | Predicted Yes | Predicted No | Total |
|---|---|---|---|
| Actual Yes | 8.0% (56) | 21.0% (146) | 29.0% (202) |
| Actual No | 1.1% (8) | 69.8% (486) | 71.0% (494) |
| Total | 9.2% (64) | 90.8% (632) | |

## Confusion Matrix for Females



Confusion Matrix

| | Predicted Yes | Predicted No | Total |
|---|---|---|---|
| Actual Yes | 10.2% (31) | 2.3% (7) | 12.5% (38) |
| Actual No | 20.4% (62) | 67.1% (204) | 87.5% (266) |
| Total | 30.6% (93) | 69.4% (211) | |

Now let's suppose we are approving loans on the basis of positive predictions. How might optimizing the model for Equal Accuracy be problematic? Do you think it would be better to avoid false positives more than false negatives or vice versa? Why? In which situations do you think one would be more costly than the other?