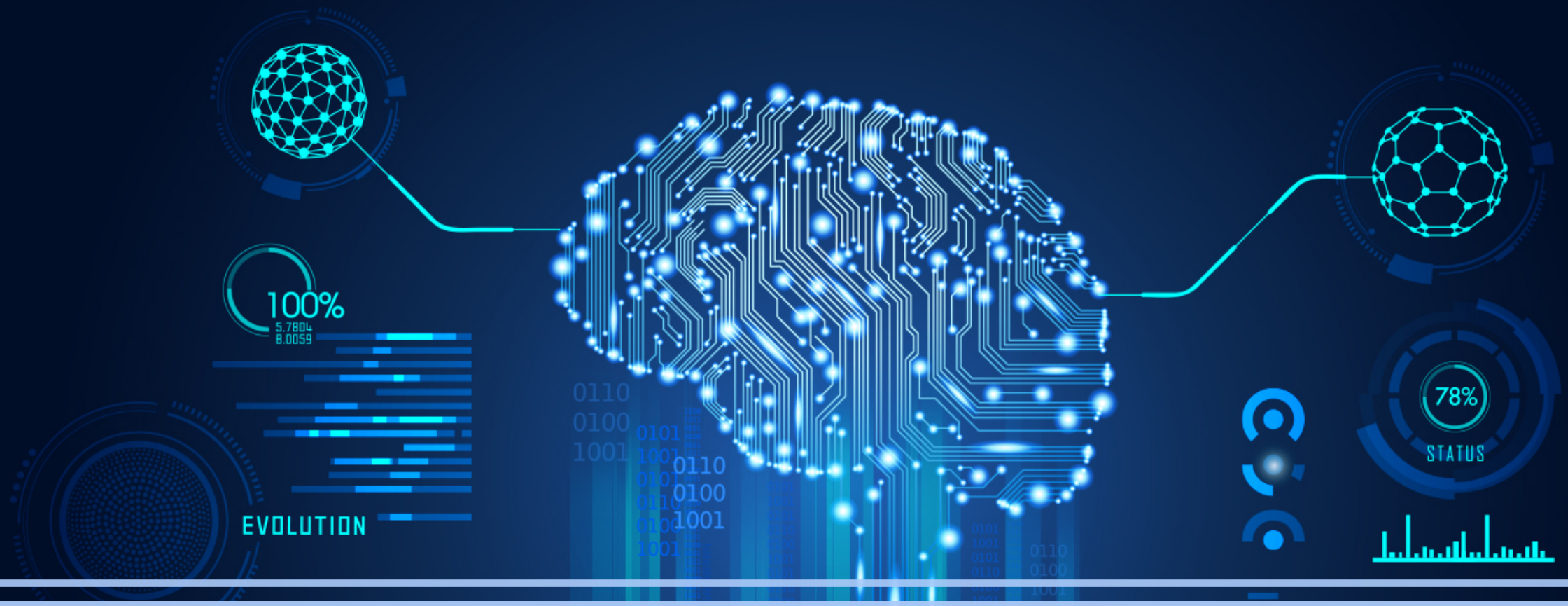


Ethical Bias
Home Mortgage Disclosure Act

Nish Sule, Bryan Tamsir, Victor Yu, Xianyuan Zhou, Xin Zou

AGENDA

DATA PREPARATION

BIAS AND DISCRIMINATION

MODELS

MODELS EVALUATION

SUMMARY

Research Questions

1. Can we identify intrinsic bias within our dataset?

2. Can we mitigate the bias while maintaining or improving a machine learning model's predictive accuracy?

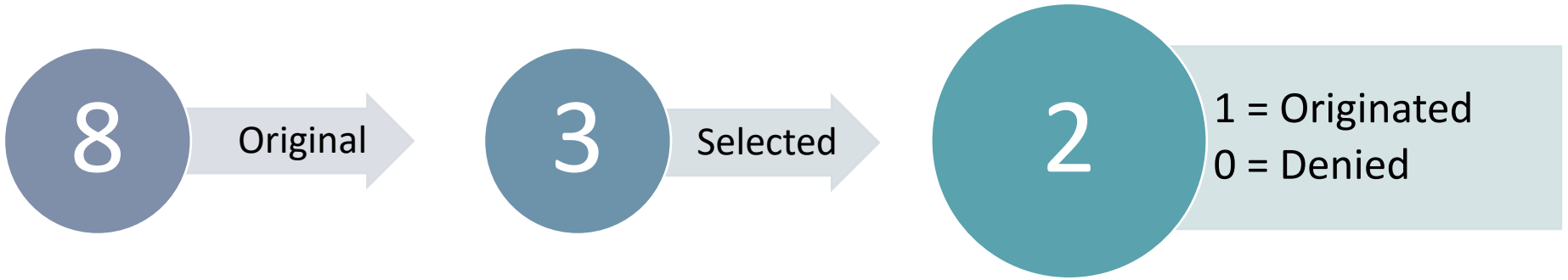
DATA DESCRIPTION

11.7 million loan records reported by **7,062** financial institutions in **2014**.

25 variables

Loan type, Property type, Action type…

TARGET VARIABLE



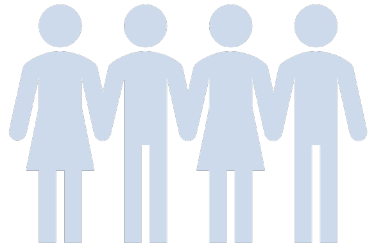
ACTION_TYPE

PROTECTED CLASSES

- Age
- **Race**
- National Origin
- Religious Beliefs
- Gender
- Disability
- Pregnancy
- Veteran Status

PROTECTED CLASSES IN OUR DATA

Applicant’s race, ethnicity, gender.



Co-Applicant’s race, ethnicity, and gender.

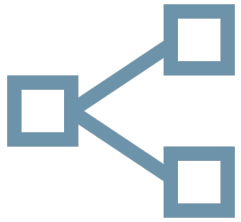
Research Questions #1

Can we identify intrinsic **bias** within our dataset?

WHY RACE?

MEASURE SOCIAL BIAS

Group-level discrimination.



Difference between advantaged class and disadvantaged class.

Values range from -1 to 1.

Mean Difference

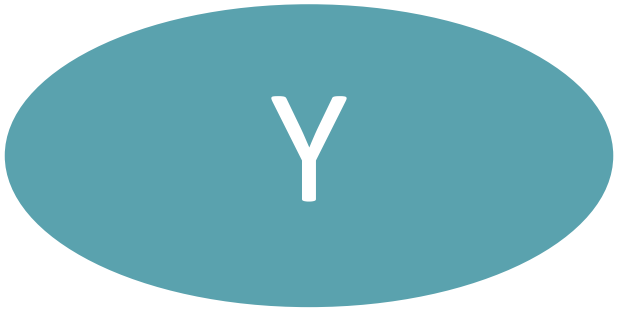
Mean Difference

Disadvantaged Class	Advantaged Class
RACE_1: 12.54% - 95% CI [11.68, 13.40]	RACE_2: -0.08% - 95% CI [-0.39, 0.23]
RACE_3: 17.82% - 95% CI [17.46, 18.17]	RACE_5: -8.52% - 95% CI [-8.75, -8.29]
RACE_4: 8.43% - 95% CI [7.37, 9.48]	

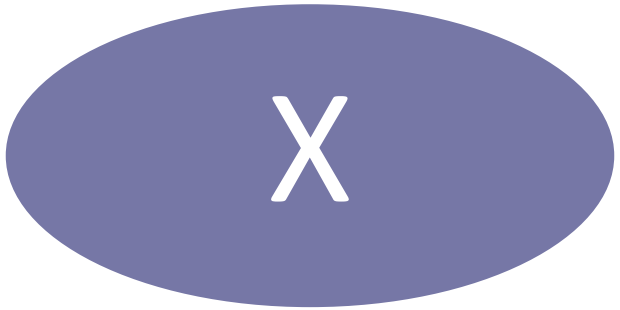
Research Questions #2

Can we **mitigate the bias** while maintaining or improving a machine learning model's predictive **accuracy**?

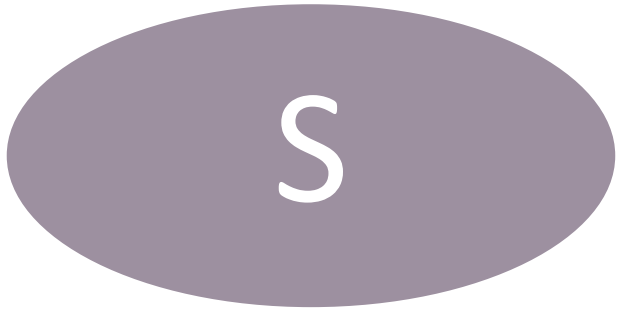
SETUP BEFORE TRAINING



Action(1= Originated, 0=Application denied)



Loan Type, Purpose, Annual Income...



Race (0= advantage group, 1=disadvantage group)

Experimental Procedure

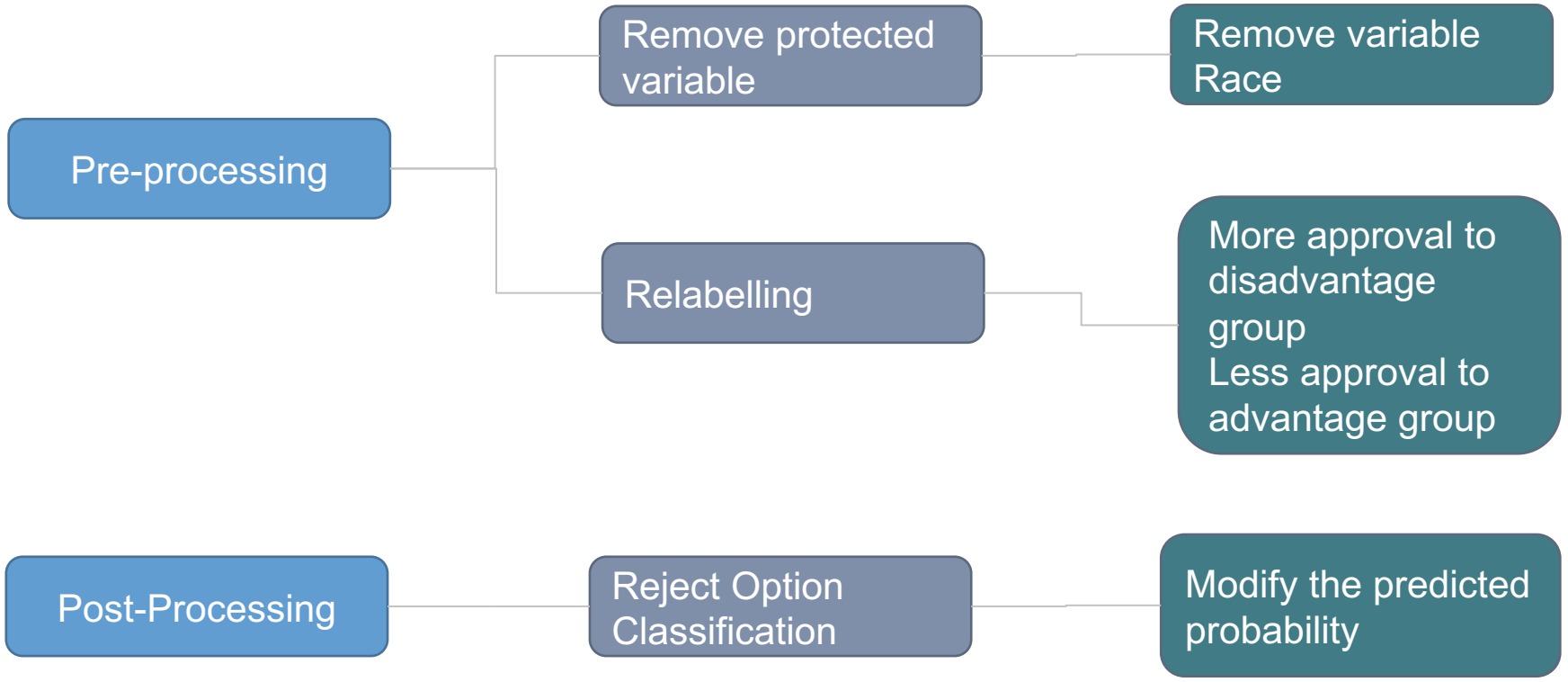
Baseline

Remove Protected Attribute

Relabel Target Variable

Reject-option Classification

Fairness Aware Methods

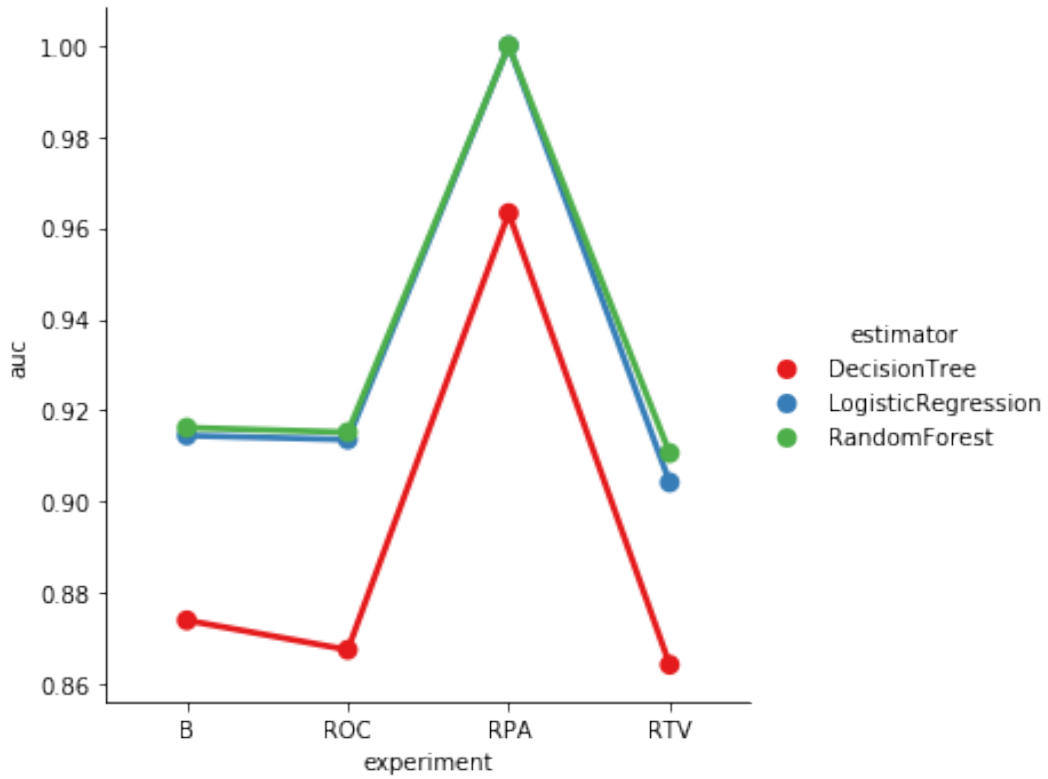


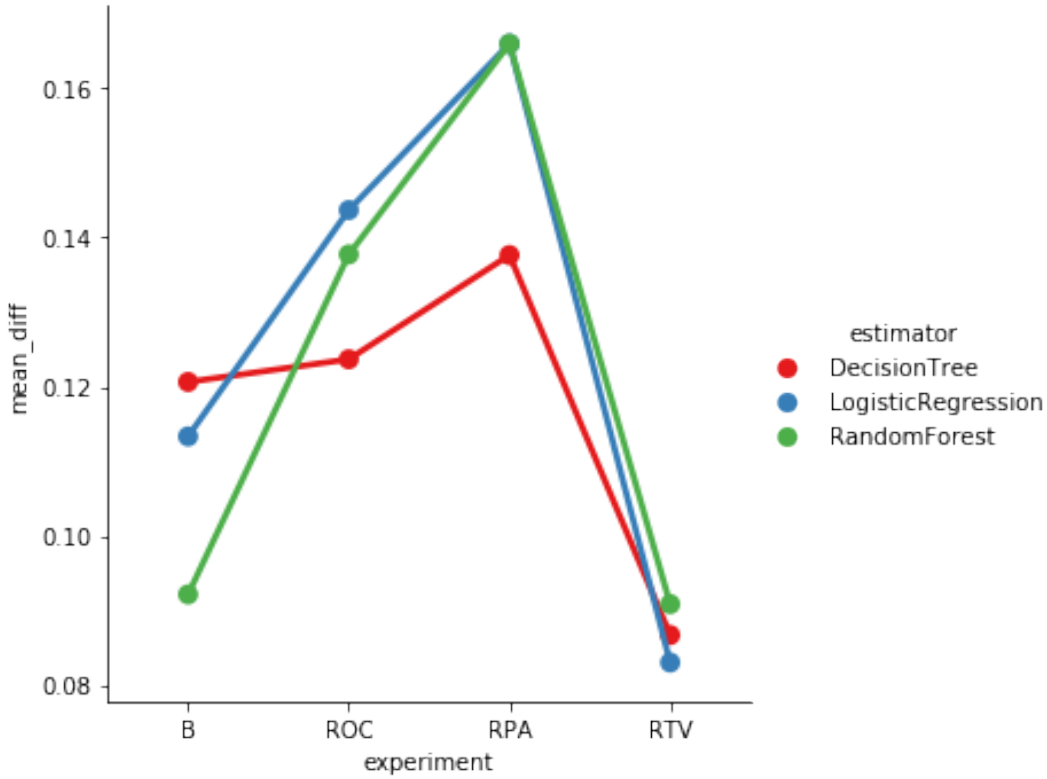
ESTIMATORS

DECISION TREE

LOGISTIC REGRESSION

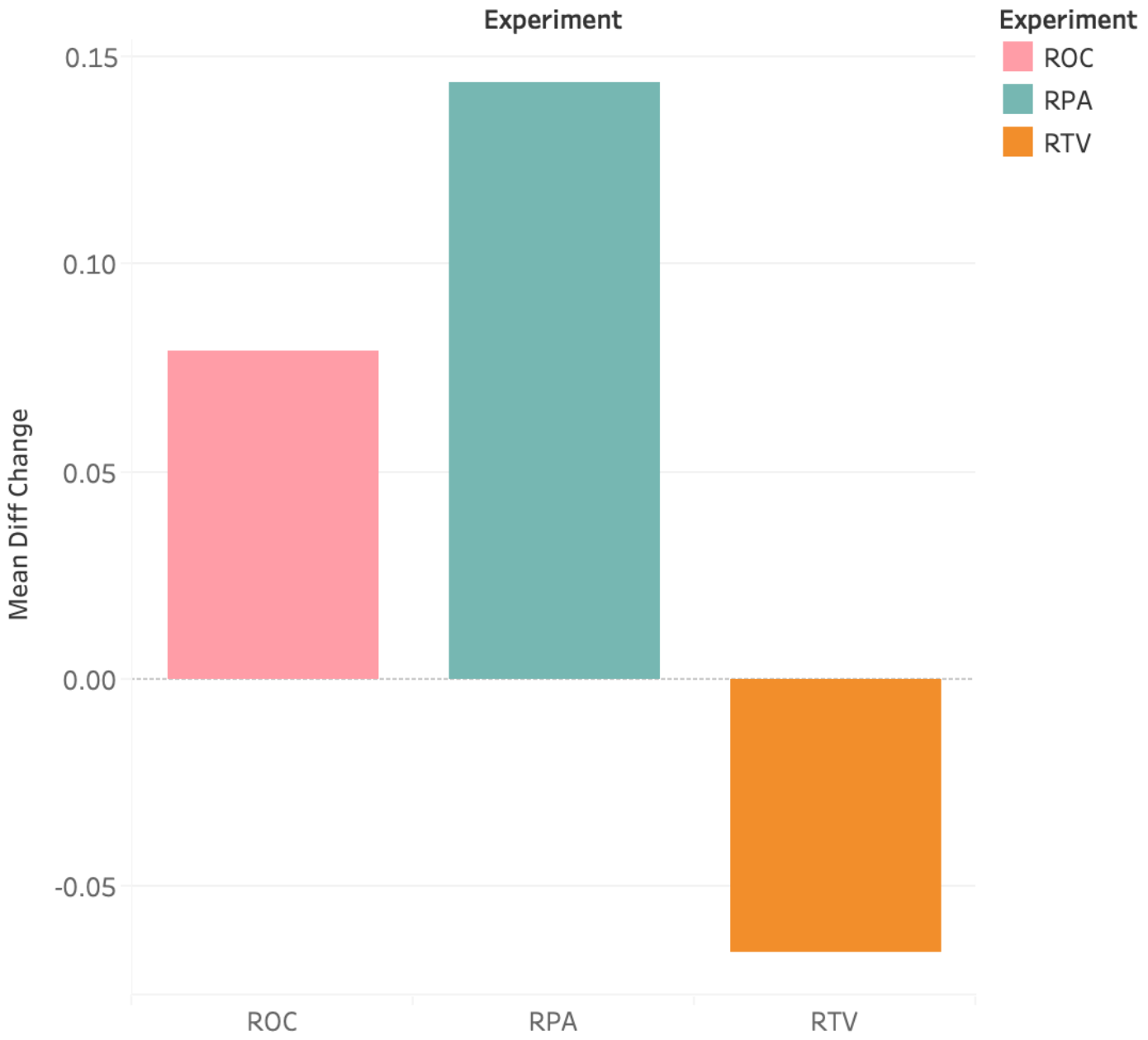
RANDOM FOREST



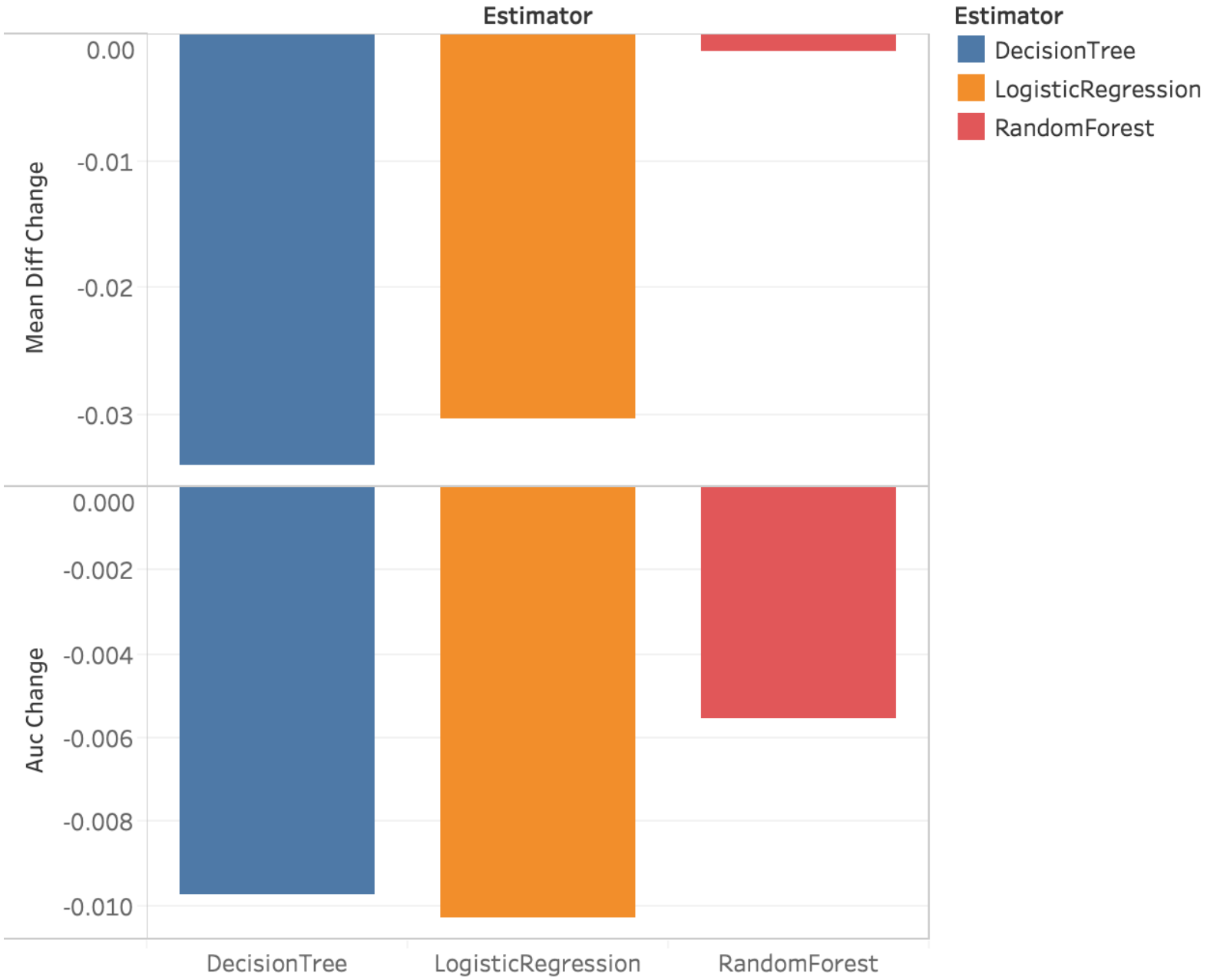


MODEL EVALUATION

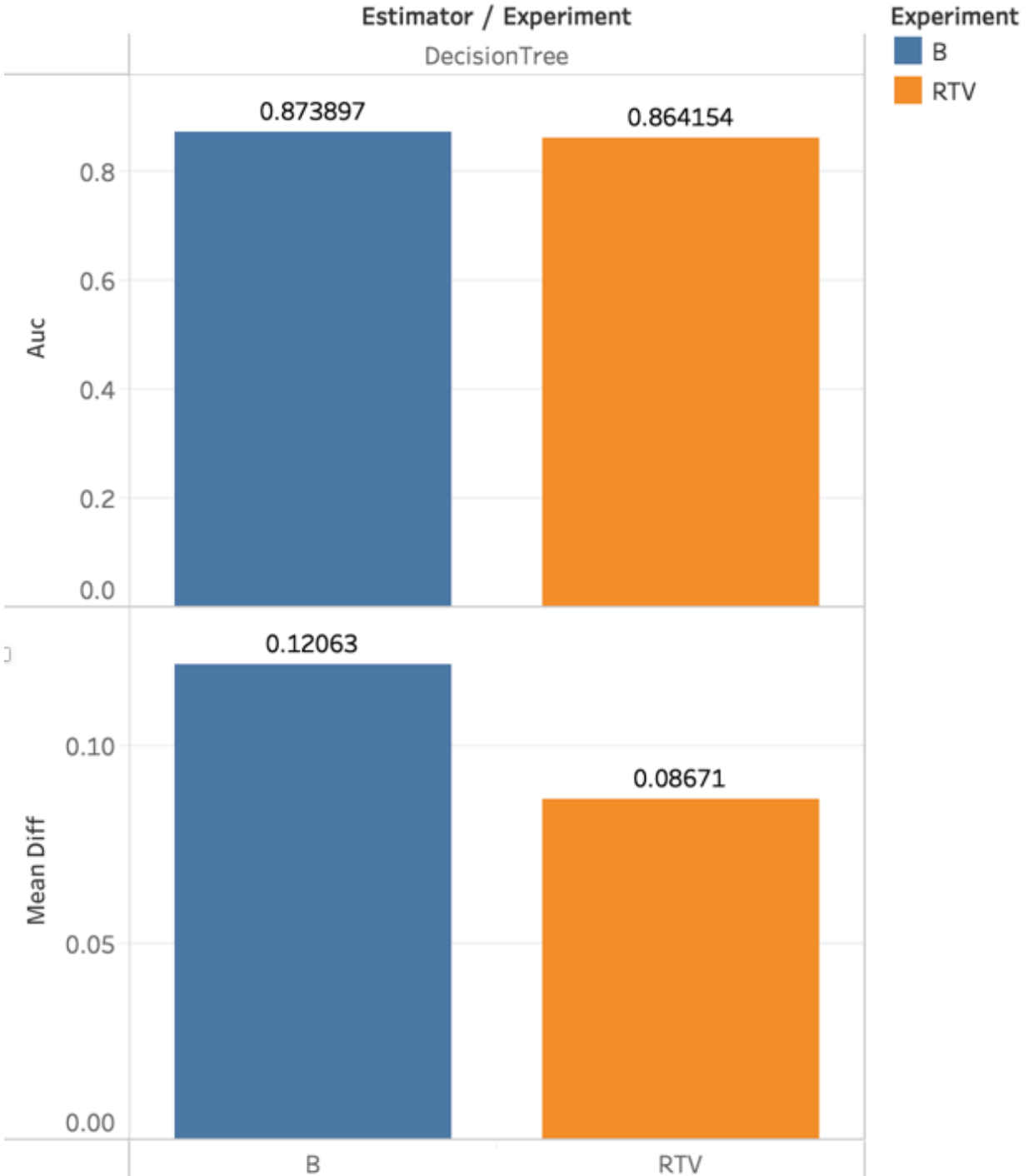
RESULTS COMPARISON



RESULTS COMPARISON



RESULTS COMPARISON



CONCLUSION

Decision Tree

RTV

CHALLENGES

Results changes every time.

Solution:

Increase size of sample data set.

Increase number of splits and repeats.

ROC’s bad performance.

Cause:

The limitation is that model types must be able to produce predicted probability.

TAKEAWAY

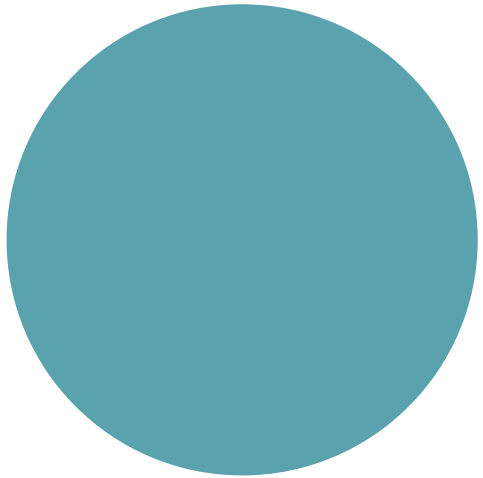
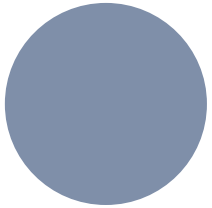
We are at risk of implicating historical biases in socially sensitive data and influence our decision-making.

We use mean difference to measuring the degree of discrimination in the dataset with respect to some discrimination metric and protected class.

We can use Themis-ml package to make a fairness-aware ML Interface. There are several methods including Baseline, Remove Protected Attribute, Relabel Target Variable.

The fairness-utility trade-off can help machine learning practitioners and researchers determine which methods are suitable for their particular data context.

THANK YOU

Q & A