



Removing Bias in AI

Team Anti-Bias

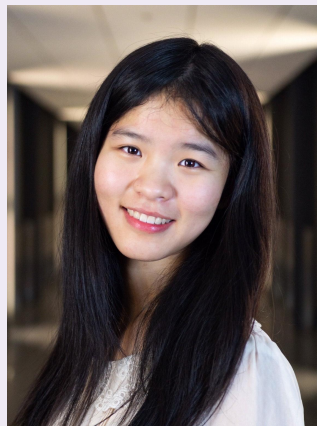
Amy Russ, Chang Liu



Meet the Team



Amy Russ
Vice President
Intermediate



Chang Liu
AI Masters Student
Intermediate

Bias in AI

The problem

- Bias in humans translates to bias in AI
- Often happens during data collection
- Current world is increasingly fueled by AI, and AI is fueled by data
- Results in systematic discrimination based on age, race, gender and more



Settings and Conditions that Demonstrated Bias

- For the x-axis we checked all filters, and then specifically drilled down to the family size.

choose x-axis, field you want to compare across

x-axis

sex



inspecting distribution along dimension: sex

Choose filter (optional), choose sub-group to focus on

filter by category

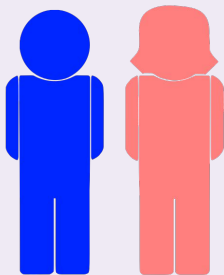
family_size



Key Findings

Selection Bias

- More female data points than male
- More new immigrants than locals
- Ethnic group 0 has twice the data compared to ethnic group 2
- Skewed data: more people applied in the morning v.s.the afternoon
- Underrepresentation of single female data points



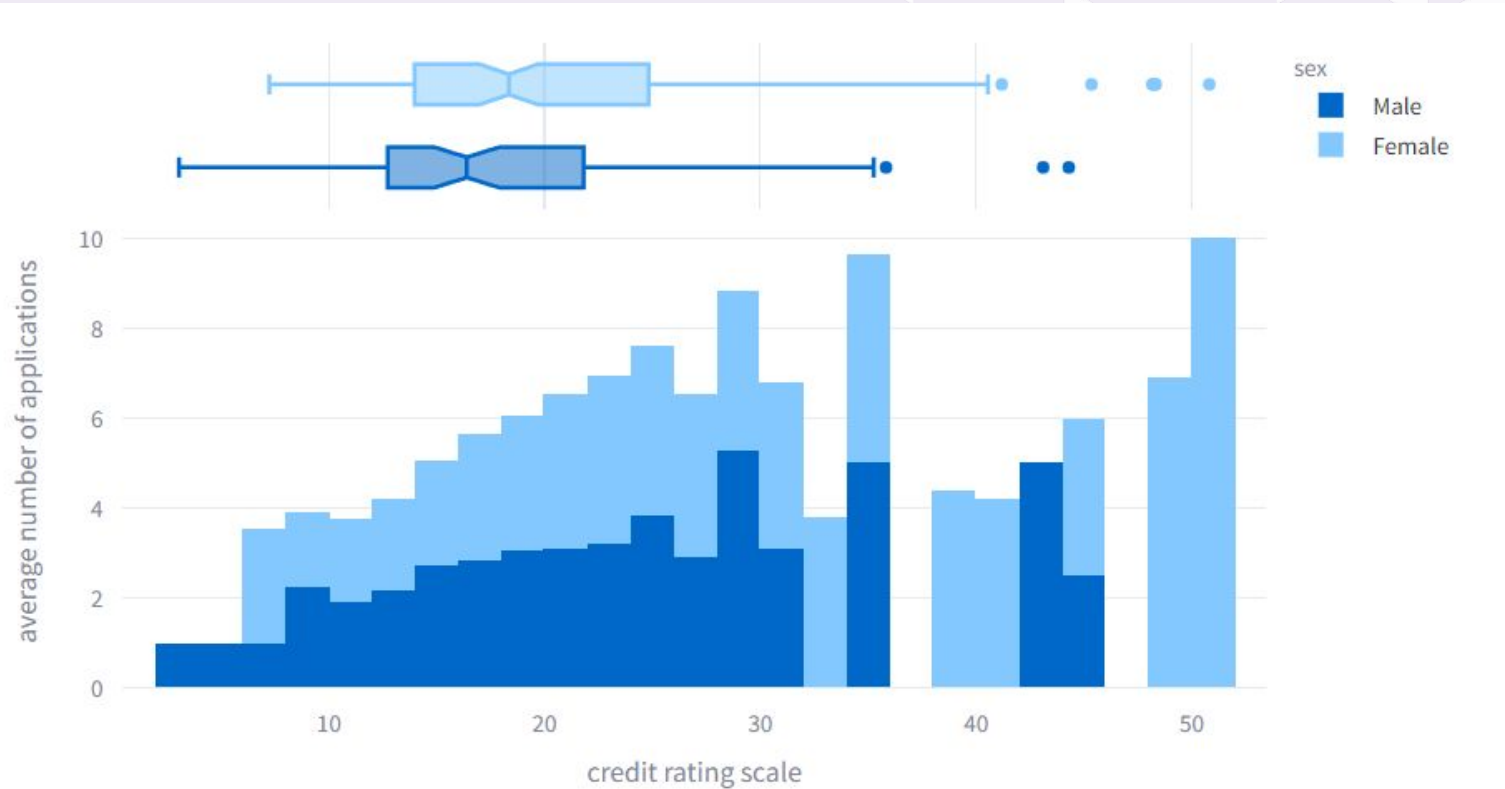
Exclusion Bias

- Only two genders are found in the data: male and female
- Non-binary individuals are not represented
- Only 4 ethnic groups are represented (there's a lot more ethnic groups in the world)

Join Round II of Data Collection!

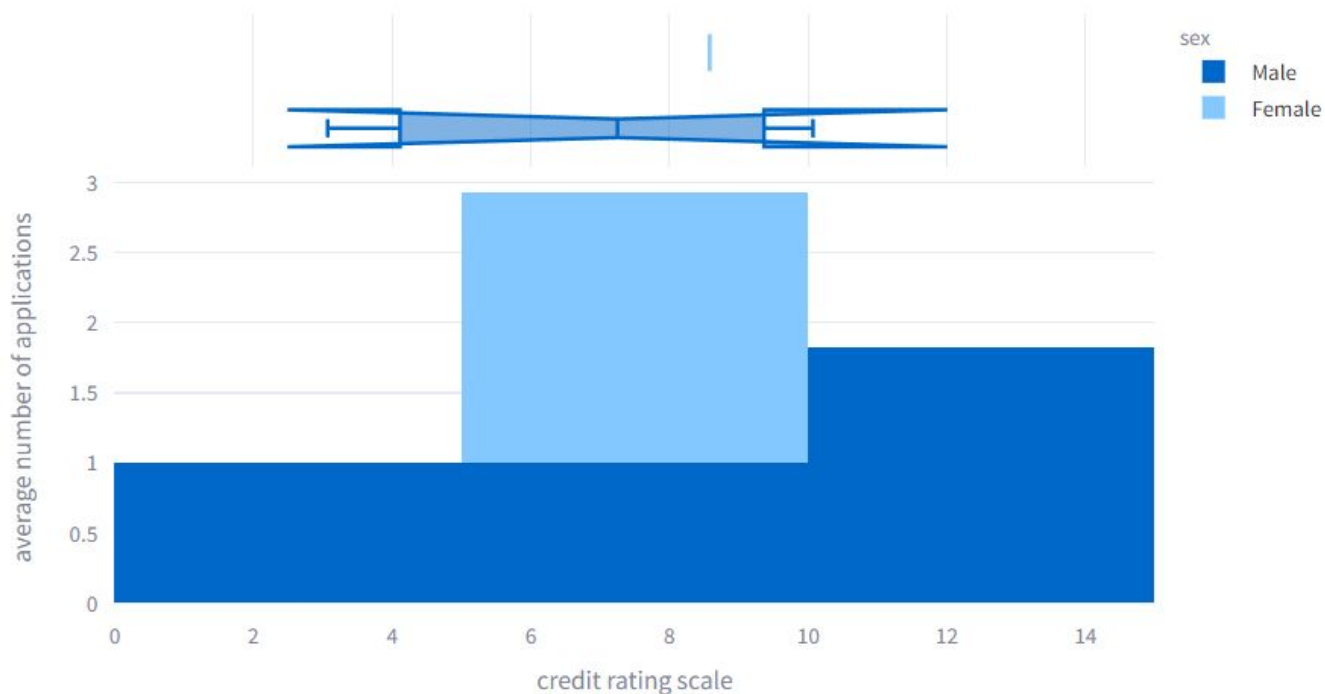


Selection Bias: More Female Data Points

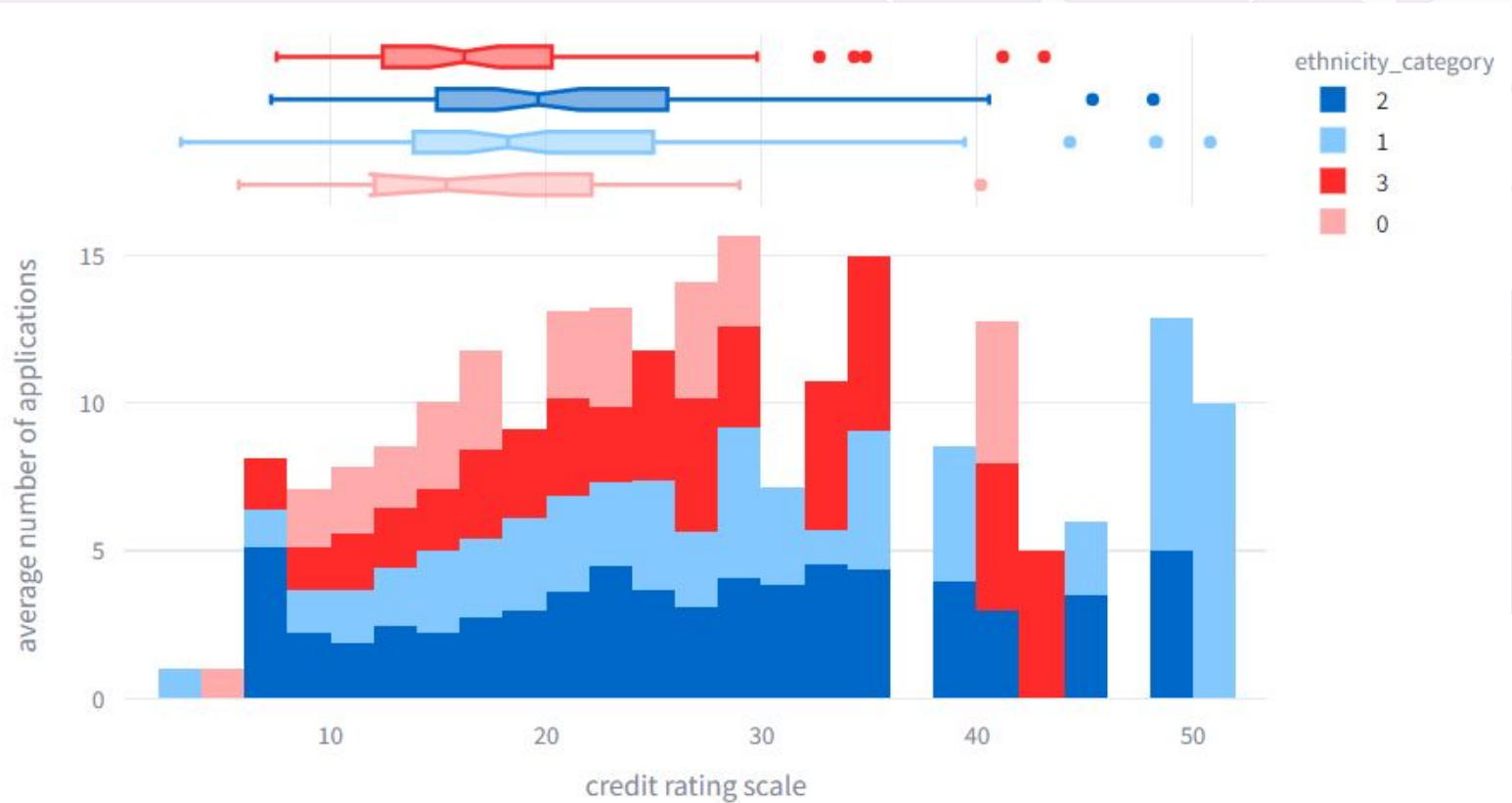


Single Male vs. Single Female (Only 1 Data Point)

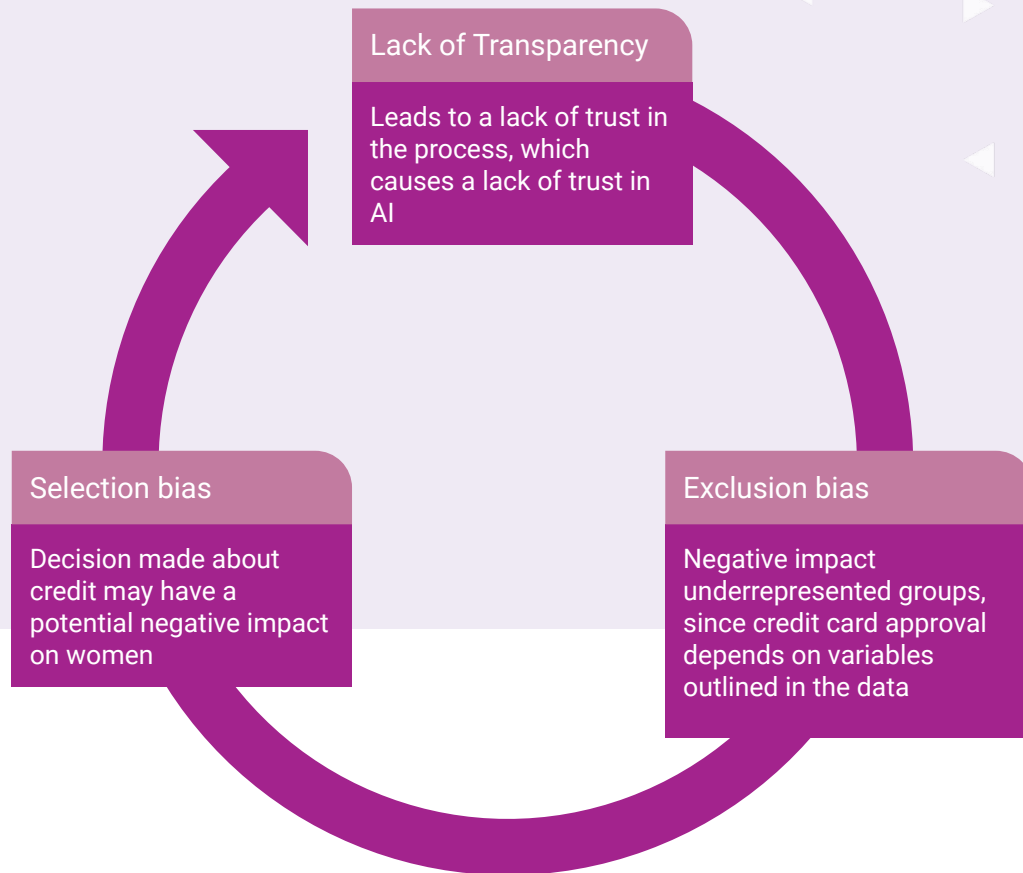
filter applied - family_size : 1



Credit Score Disparity Between Ethnicities



Bias: Societal Impacts



How do we Combat bias in AI?

01

Create Evaluation Metrics

- Metrics should be relevant to both the industry and the model type
- Should be agreed upon by all key stakeholders within the organization
- Should be managed by a data steward—someone within the organization who is a Subject Matter Expert (SME)
- The model should be reviewed by multiple groups within the organization
- The data collection methodology should also be reviewed

02

Use human reviewers

03

Re-evaluate the model and data based on societal shifts and on a set timeline

- The model should be reevaluated based on rules and regulations that govern financial models and information use (i.e., the Consumer Financial Protection Act, etc.).

Proposed Solution

1

Review data source for bias

Appears to be a lack of understanding the population and business case

2

Document data collection and cleaning methods




Maintain transparency through all aspects of the data handling process

3

Implement model review process

Continuous modeling based on changing regulatory requirements, demographics, etc. to ensure model performance and accuracy

Solutions: Implementation and Monitoring

- Review data source for bias 
 - Implementation of data governance console monitoring (Tool: Collibra)
 - Ensure vendor SOP including wording around data quality and how it will be monitored by the vendor
- Document data collection and cleaning methods 
 - Implement organizational change to data as a product mindset, which tends to increase documentation
 - Require SMEs to follow a set documentation process
- Implement model review process 
 - Create a model review process based on the importance of the model, the level of security around the data, and the type of decisions to be made from the data

The above solutions must constantly be monitored and evaluation on a basis that minimizes the risk of possible organizational loss (e.g. loss from government fines, etc.

Pros and Cons of Potential Solutions

Pros 🌟

- Solutions will allow for more transparency and mitigate the two identified biases
- Ensure compliance with any federal regulations
- Increase trust in the model and the potential business decisions made from the model

Cons 💔

- Implementation takes time and consensus—slows down go-to-market for the solution
- Solutions focus primarily on the data, and not the model itself
- Requires skill set that may not be available within the organization

Next Steps

- Follow-up to understand how the data was collected
- Follow-up with engineers to understand if there's also bias in the model architecture
- Work with organization to implement data governance procedures for models
- Advocate for AI Fairness and removing Bias in AI!



Thank You for Organizing the Hackathon!

We all had lots of fun!