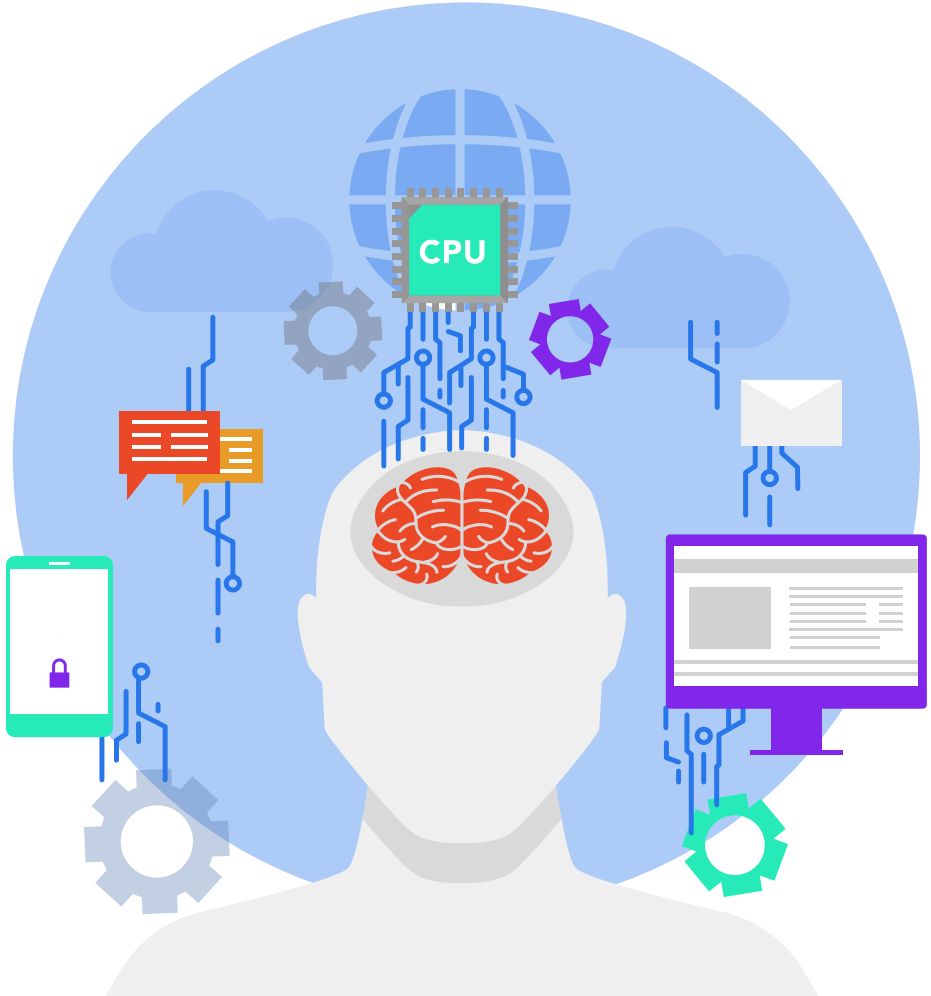



# Ethical Bias in Machine Learning

Team: Supernovas
Group 3

Priyanka Murugan
Akshay Prasath
Aishwarya

# Agenda

01 **Motivation**

Why is there a need for fairness in machine learning?

02 **Problem Statement**

What is the problem we are trying to solve?

03 **Data and its features**

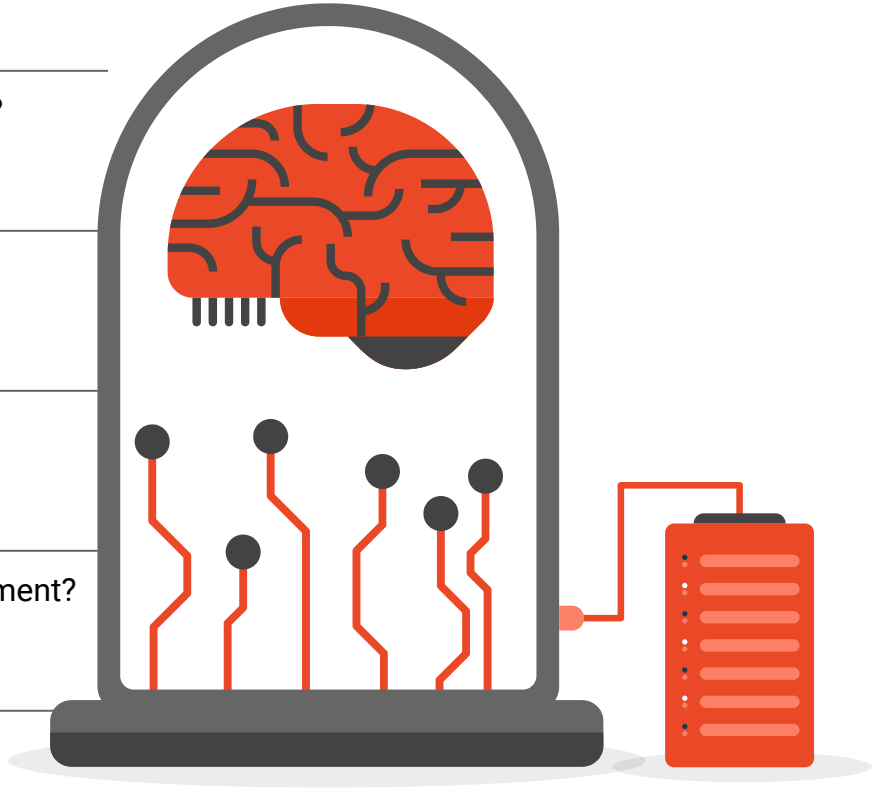
What are the features of data used?

04 **Matching and Prediction Methodology**

What are the methodologies used to prove the statement?

05 **Results and Next steps**

What are the findings and where do we head next?



# MOTIVATION

## AI bias in COMPAS

- Correctional Offender Management Profiling for Alternative Sanctions
- US court systems
- Likelihood of defendant being rearrested

### What went wrong?

- Predicted **twice** as many **false positives** for black offenders than others

## Gender bias in lending

- AI systems for loan approvals and amounts
- European banks

### What went wrong?

- Black and Hispanics **80% more likely to be rejected**
- Women approved amounts **14,000 euros lower on average**

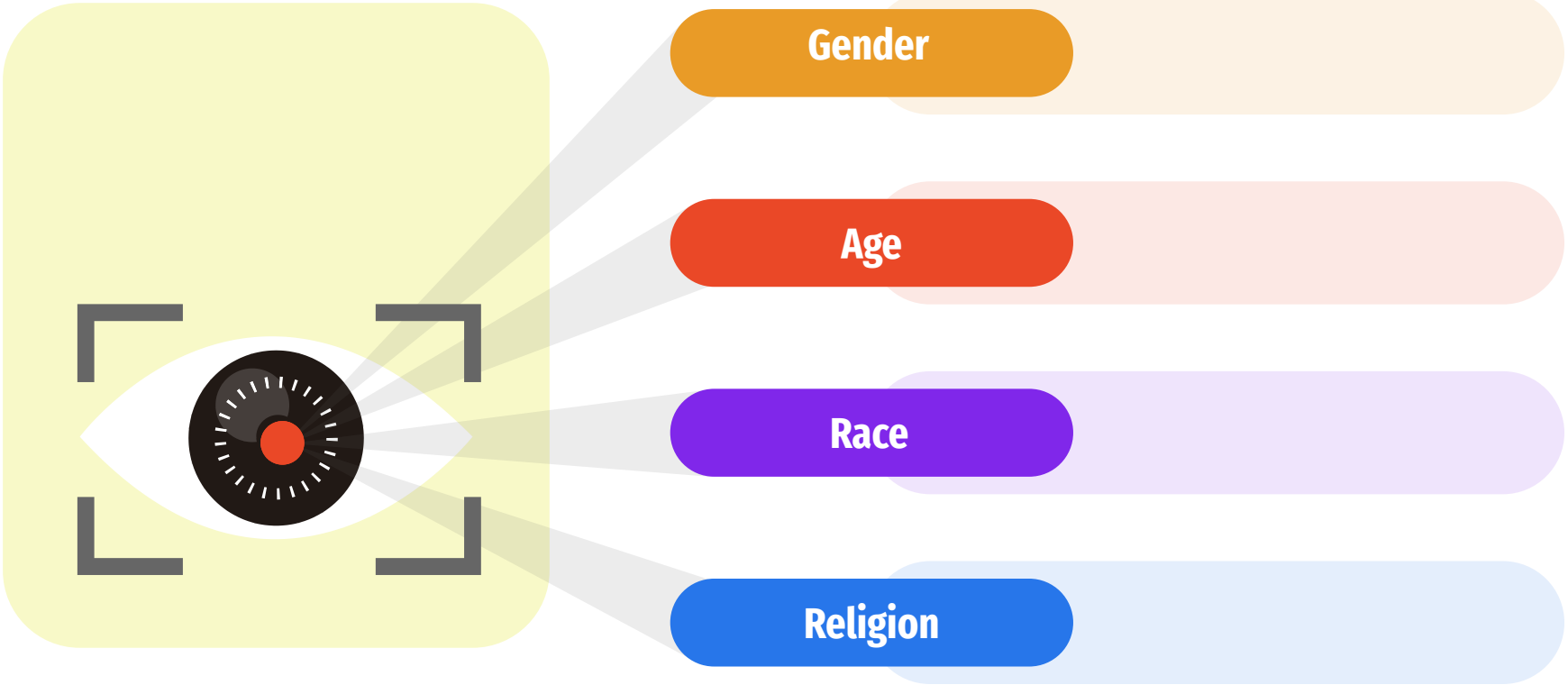
## Bias in hiring

- Resume ranking
- Amazon's recruiting engine
- Trained on historical data

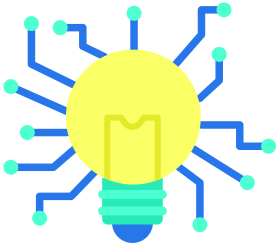
### What went wrong?

- Penalized the word "**women's**"
- Little significance to skills **not** found on men's resumes

# Sources of Bias



# Problem Statement

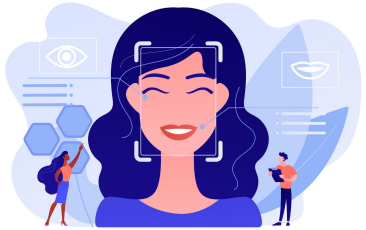


- **Matching:** Can the users be matched on data other than demographic?
  Spending behavior?
  Social Network?

- **Prediction:** Can we predict user behavior without using demographic data?
  How well?

# Input data

**Data Sources**



**Face ++**

**Demographic**
(Race, Gender, Age)

**Venmo**


**Behavioural**
(Type of activity, no.of friends)

**Social Network**

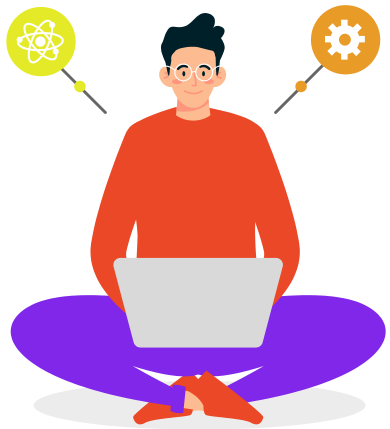
**Transactional**

# Behavioural Features

## Social Network

Identifies the user social network behavior

- Friends
- Friends of Friends
- Clustering coefficient



## Transactional

Identifies the user spending behaviour

- Recency
- Frequency
- Activity

# Matching Methodology

**Deviance Comparison**
Demographic vs
Behavioural

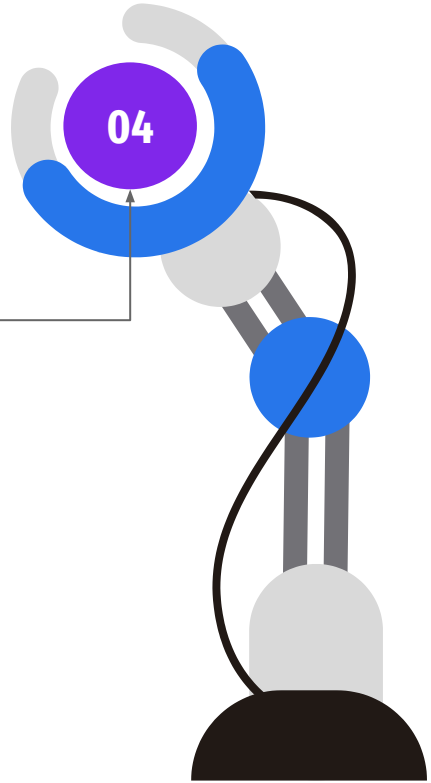
**Identify Matched Pairs**

K-Nearest Neighbour

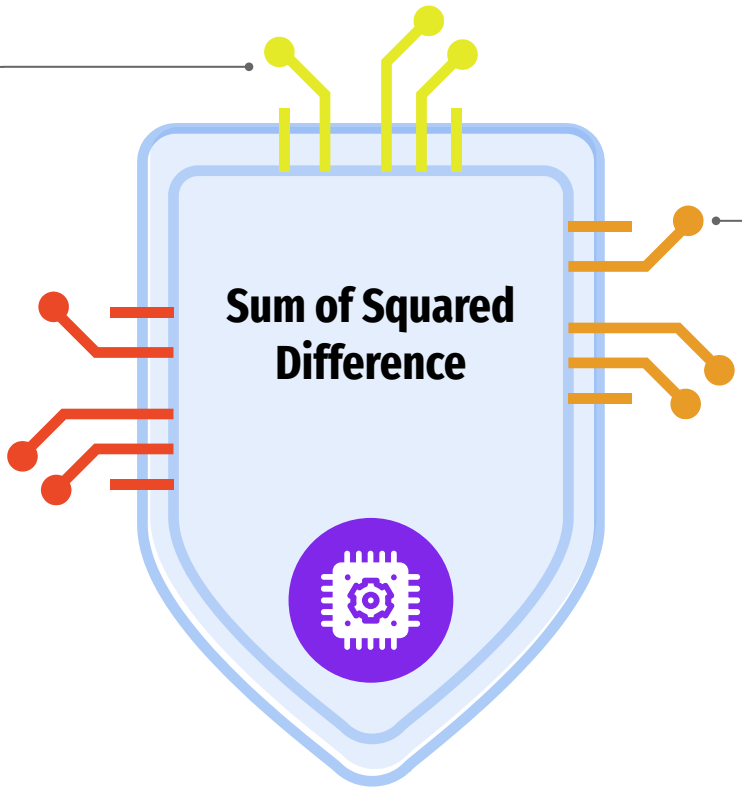
**Propensity score**

Logistic Regression

**Data Processing**
Social network and
transactional features



# Results Comparison




01

**Demographic**

SSD: 0.006e-4

02

**Activity & Network**

SSD: 1.6e-4

03

**Baseline (Demographic and Behavioural)**

SSD: 3.2e-4

# Prediction Methodology

**Comparing MSE**
Demographic vs
Transactional

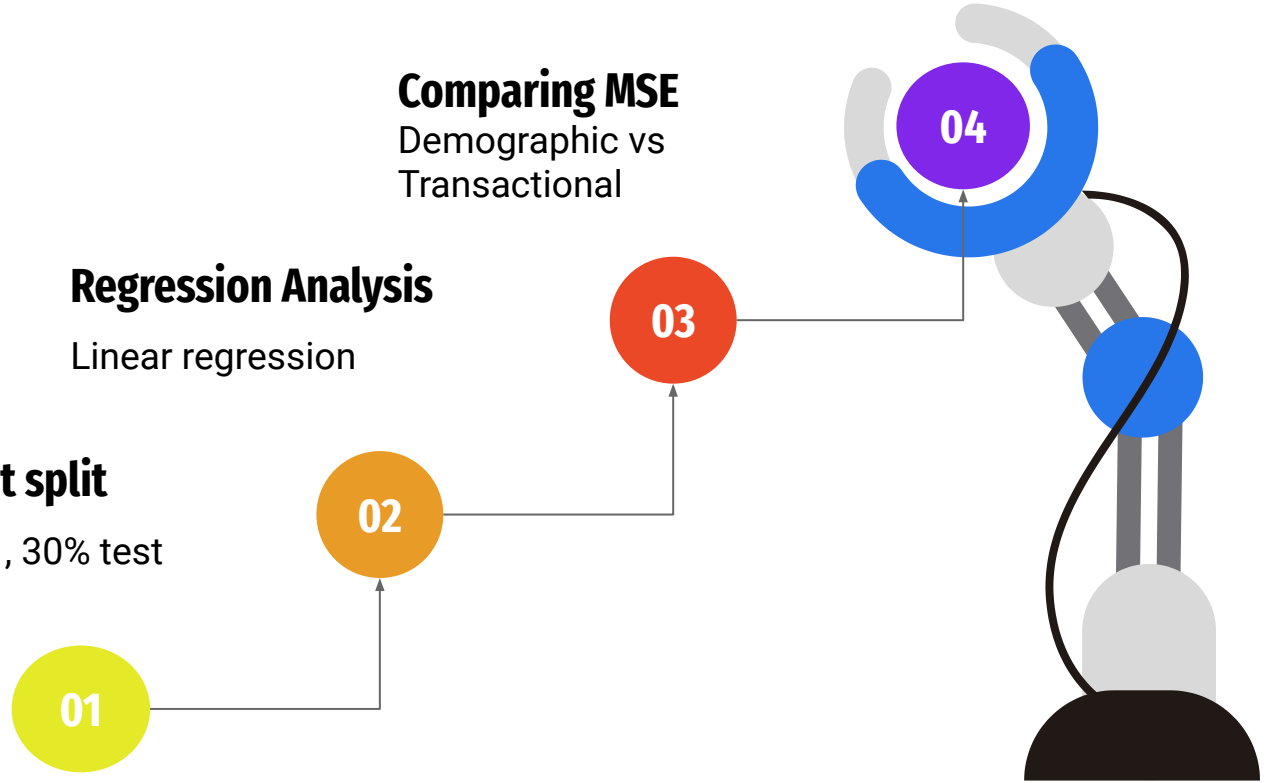
**Regression Analysis**

Linear regression

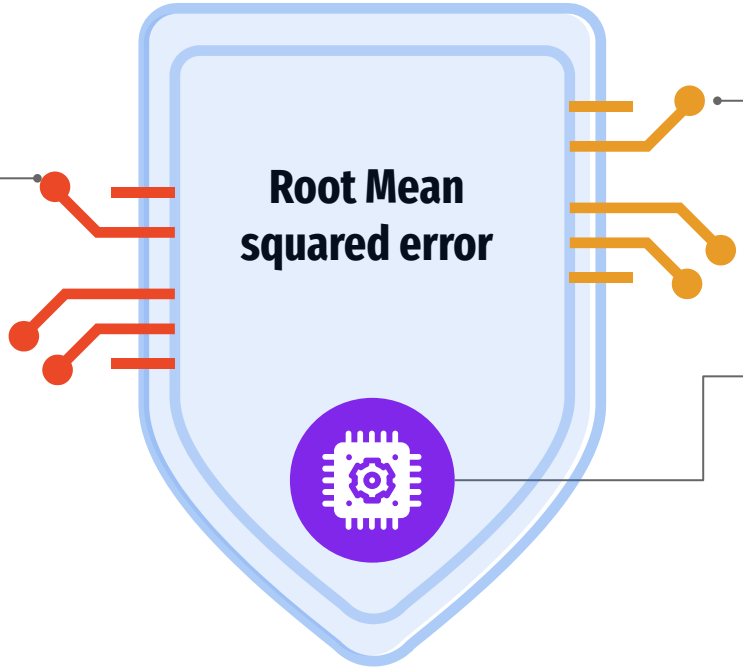
**Train test split**

70% train , 30% test

**Data Processing**
y-Number of transactions
x-Behavioral & Demographic



# Results Comparison




01

**Demographic, Transactional & Network**

RMSE: 0.42
OOS R^2: 94.22%

02

**Transactional**

RMSE: 0.48
OOS R^2: 92.7%

03

**Transactional & Network**

RMSE: 0.38
OOS R^2: 94.56%

# Next Steps

## What about new users?

### Use average of existing users

- Average of transactional and social network behavior of existing users

### How long to collect data?

- Compare different user lifetimes

## Relevant Feature

### Social Network

- Social network activity
- Lifetime of network
- Frequency of transaction of network

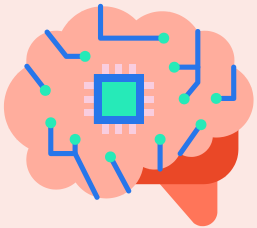
### Transactions

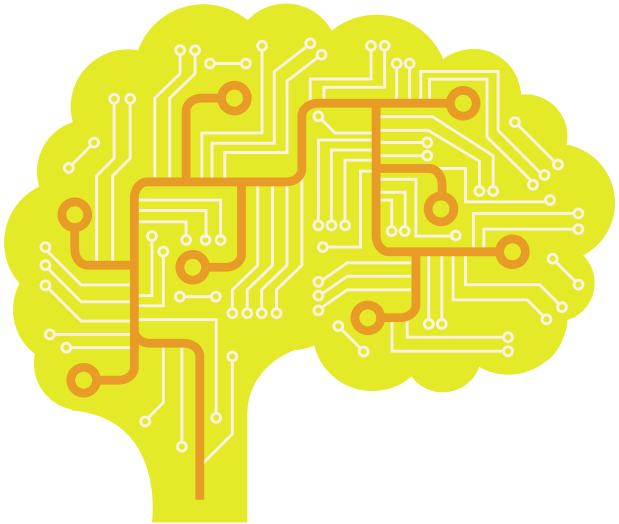
- User persona based on trend in transaction
- Remove variables correlated with demographics

## Custom Metrics

### User similarity

- Cosine similarity
- Euclidean distance
- Jaccard similarity





# Thank You!

# Questions?