Balancing Accuracy and Fairness

Predictive Models for Credit Approval

Mahidol University
Information and Communication Technology
ITCS227 Introduction to Data Science
Summer Semester 2025

Agenda

Introduction

- Key Facts about Mana Outlet
- Illustration: Fairness Decision Tree?
- Research Question(s)

Methodology

- Dataset
- Data Cleaning and Transformation
- Machine Learning Algorithms (Interpretable) and Fairness Metrics

Results and Findings

- Data Analysis and Results
- Feature Selection and Importance
- Fairness Evaluation and Research Question(s) Discussion

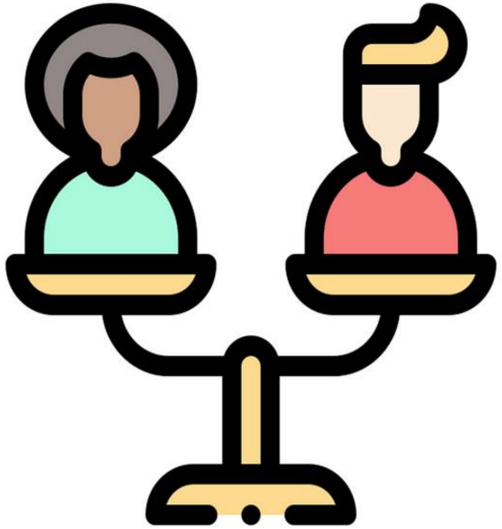
Conclusion and Outlook

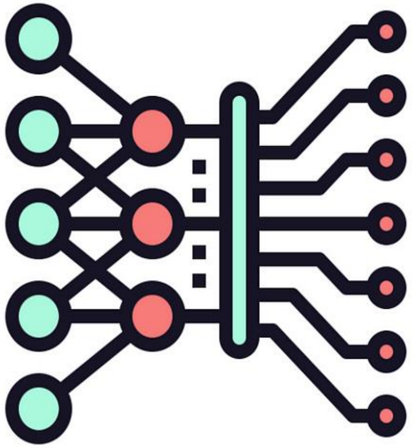
- Expected Outcomes and deliverables

Introduction

Motivation: Data Science and Fairness

- Data Science **aims** to uncover **actionable insights hidden** in an **organization's data**
- Methods of **Mathematics**, **Statistics** and **Programming** as well as **Machine Learning** and **AI** are used for this purpose
- Many applications, but especially in **fields** where **important decisions** are made **about people**





- **Fairness**: The quality of **treating people equally** or in a **way that is right** or **reasonable** (Cambridge Dictionary)
- **Can Machine Learning Models that learn from data become discriminatory and unfair?**

Data Science and Fairness (2)

- **Peela, H. V., Gupta, T., Rathod, N., Bose, T. & Sharma, N. (2021).** Prediction of Credit Card Approval. International Journal Of Soft Computing And Engineering, 11(2), 1–6. https://doi.org/10.35940/ijsce.b3535.0111222.

→ **Best Machine Learning Model achieved an Accuracy of 86%**

- **Alagic, A., Zivic, N., Kadusic, E., Hamzic, D., Hadzajlic, N., Dizdarevic, M. & Selmanovic, E. (2024).** Machine Learning for an Enhanced Credit Risk Analysis: A Comparative Study of Loan Approval Prediction Models Integrating Mental Health Data. Machine Learning And Knowledge Extraction, 6(1), 53–77. https://doi.org/10.3390/make6010004

→ **Best Machine Learning Model achieved an Accuracy of 85%**

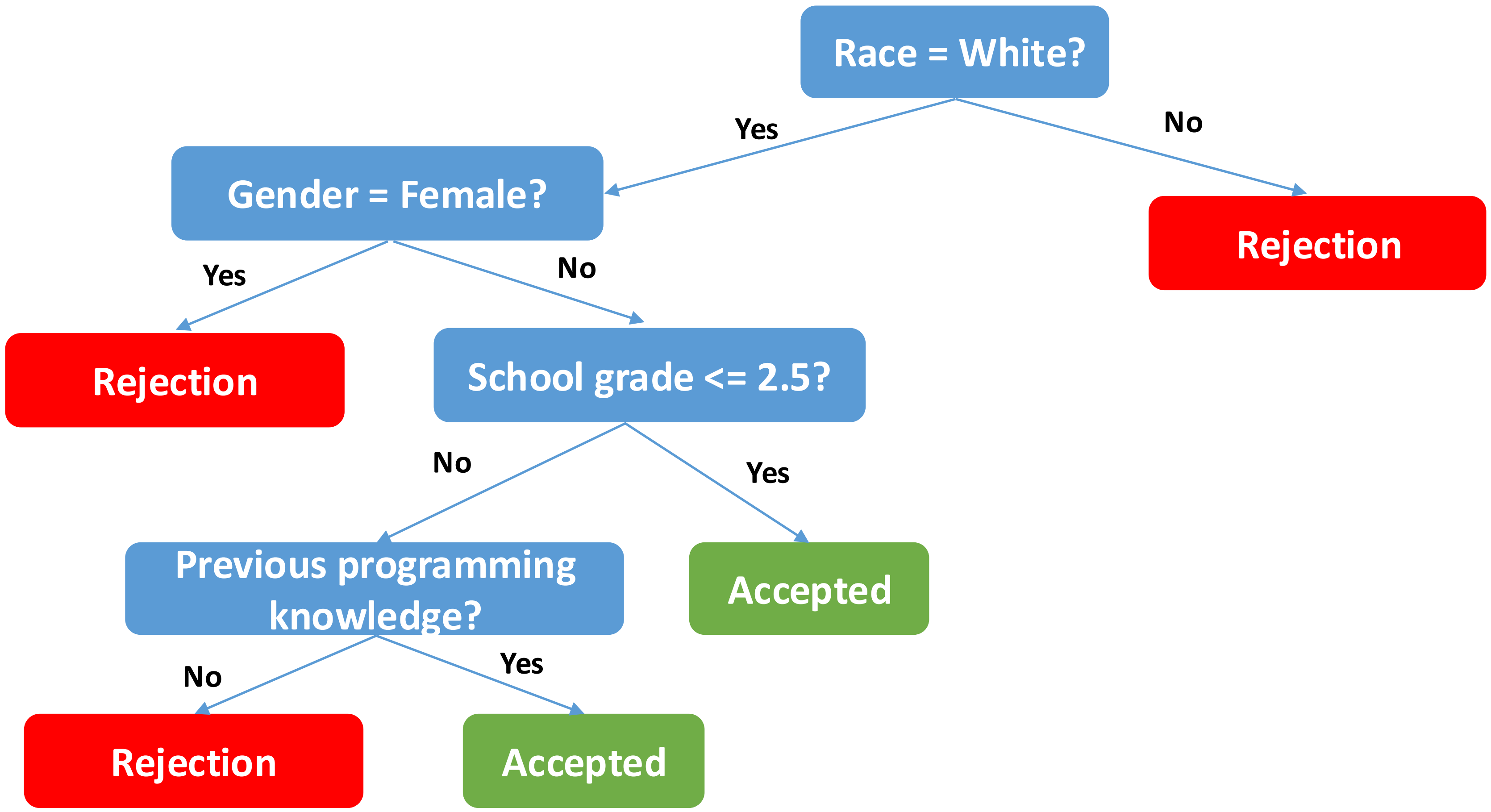
- **Sarkar, T., Rakhra, M., Sharma, V. & Singh, A. (2024).** An Empirical Comparison of Machine Learning Techniques for Bank Loan Approval Prediction. International Conference On Communication, Computer Sciences And Engineering, 137–143. https://doi.org/10.1109/ic3se62002.2024.10593355

→ **Best Machine Learning Model achieved an Accuracy of 90.01%**

→ **Fairness not evaluated**

Problem Statement: Fairness Decision Tree?

Example: Admission to Computer Science degree programme (Accuracy: 95%)



Research Question(s)

Reserach Question 1

Feature Selection and Importance

- Are sensitive attributes strongly considered?
- Can we simply remove them (without consequences for accuracy)

Research Question 2

Metrics to measure the Fairness of Models

- What metrics are there to measure fairness?
- Are these metrics meaningful and useful?

Research Question 3

Trade-Off between Accuracy and Fairness?

Methodolgy

Dataset: Adult Census Income

- Based on 1994 U.S. Census Bureau CPS data
- Contains socio-demographic characteristics and income information
- Sample of approx. 32,561 people
- Data comes from surveys of households in the USA

→ **Ideal to analyze Fairness, as it contains sensitive attributes (age, race and sex)**

No.	Attribute	Original Type	Range	Type Used
1	age	continuous	17–90	categorical
2	workclassge	categorical	1–8	categorical
3	final weight (fnlwgt)	continuous	12,285–1,484,705	numeric
4	education	categorical	1–16	categorical
5	education-num	continuous	1–16	categorical
6	marital-status	categorical	1–7	categorical
7	occupation	categorical	1–14	categorical
8	relationship	categorical	1–6	categorical
9	race	categorical	1–5	categorical
10	sex	categorical	1–2	categorical
11	capital-gain	continuous	0–99,999	numeric
12	capital-loss	continuous	0–4356	numeric
13	hours-per-week	continuous	1–99	categorical
14	native-country	continuous	1–41	categorical
15	class	categorical	1–2	categorical

→ **Idea: Using the Dataset for Credit Approval**

Data Cleaning & Transformation

Analyze data set and transform features

- Binary categorial features transform to 0 and 1
- Non-binary categorial features with One-Hot Encoder
- Standardize numerical features
- Transform target variable to 0 and 1

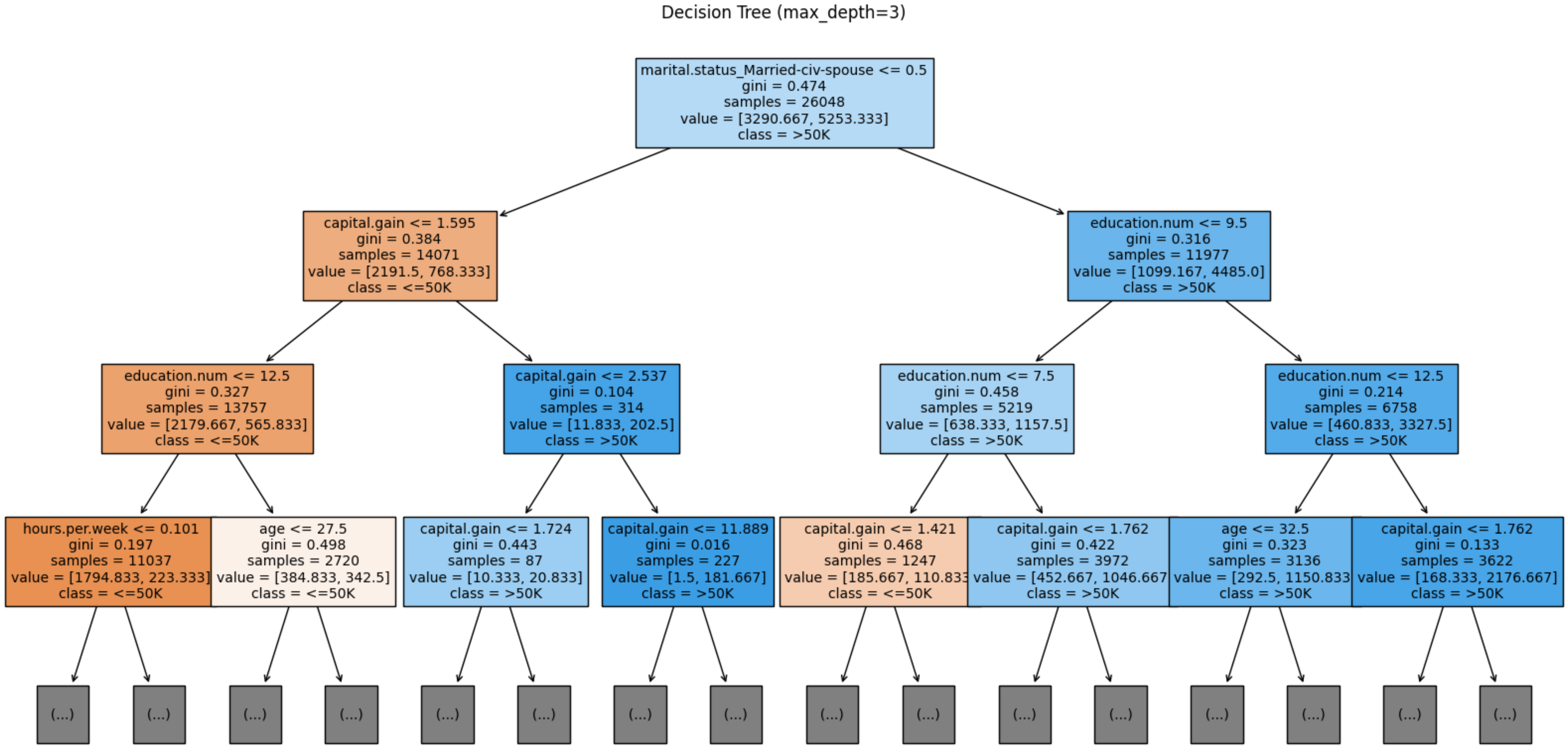
Replace missing values according to assumptions

- **Workclass** is mostly Private → Replace “?” with Private
- **Occupation** not known → New category of “Unknown” for “?”

Remove unnecessary features and duplicates/outliers from the data

- Delete final-weight and education (education-num instead used)

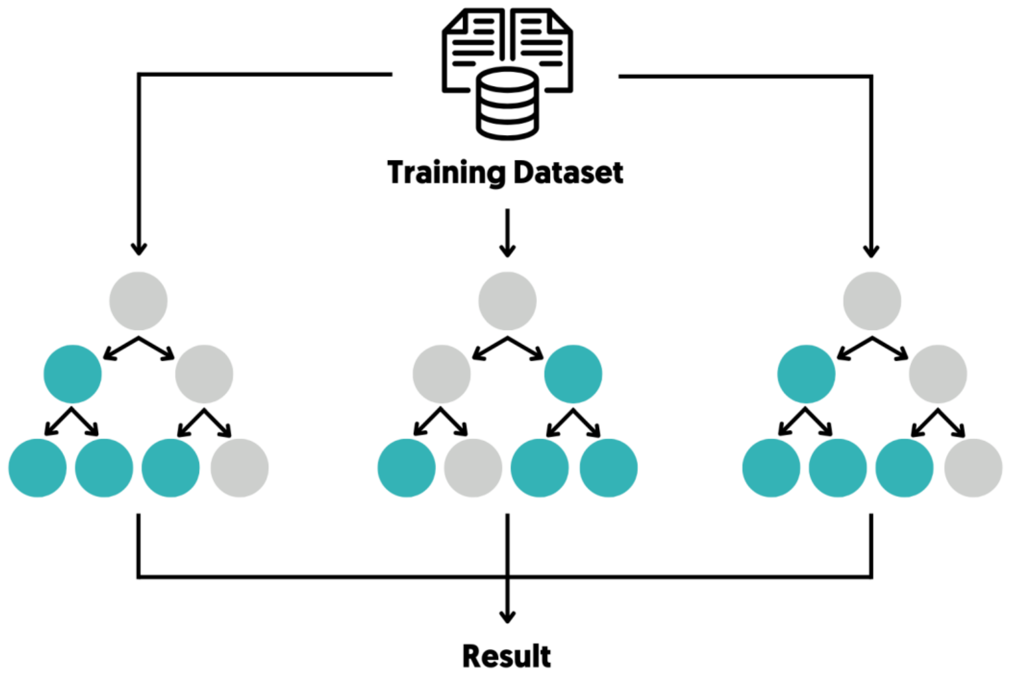
Decision Tree



Random Forest

We used Random Forest to **evaluate decision process and predict** whether an individual earns **more than $50K per year**.

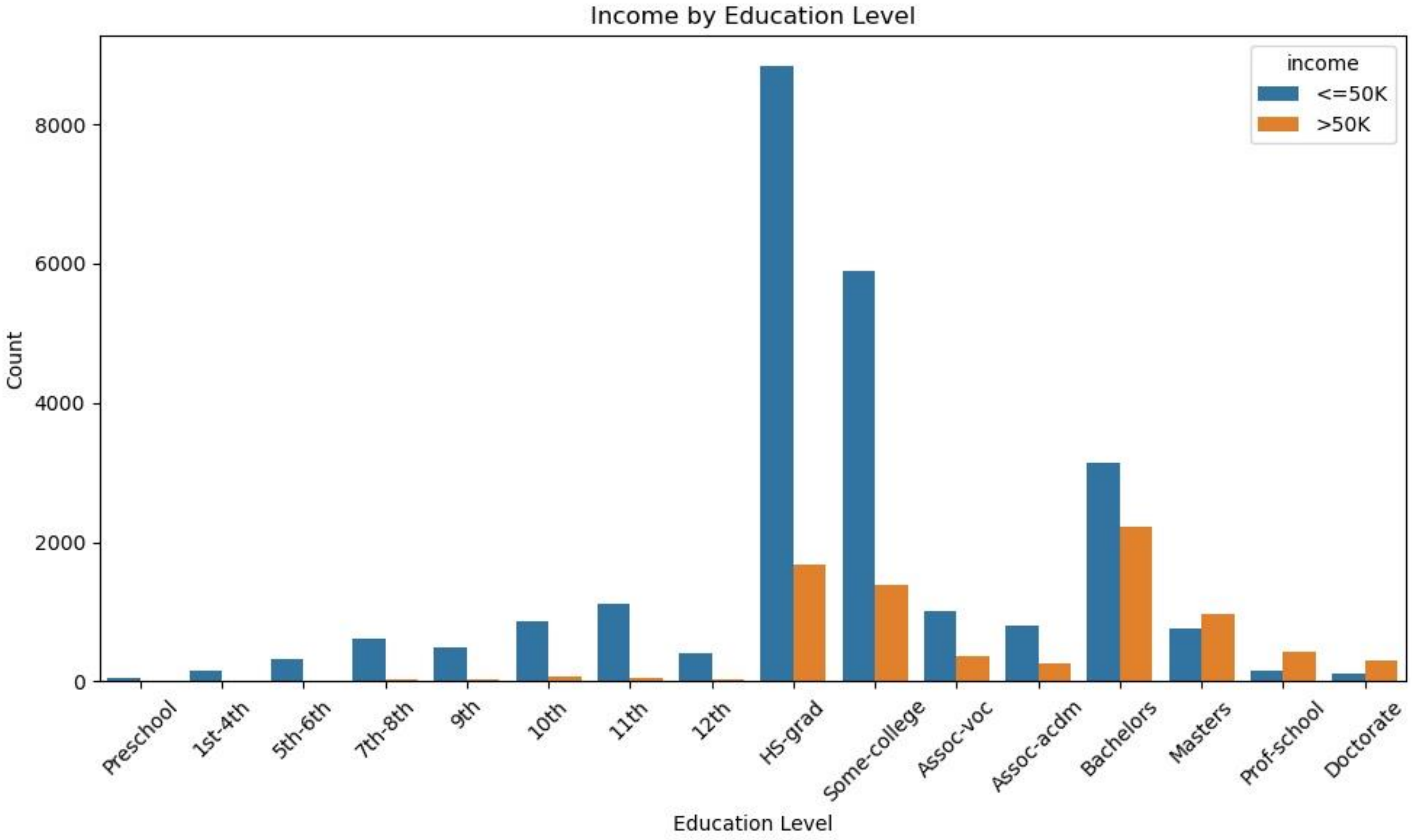
- We used Random Forest to find patterns in the data and improve accuracy.
- This model performs better than Logistic Regression in terms of accuracy but is harder to understand.



Results and Findings

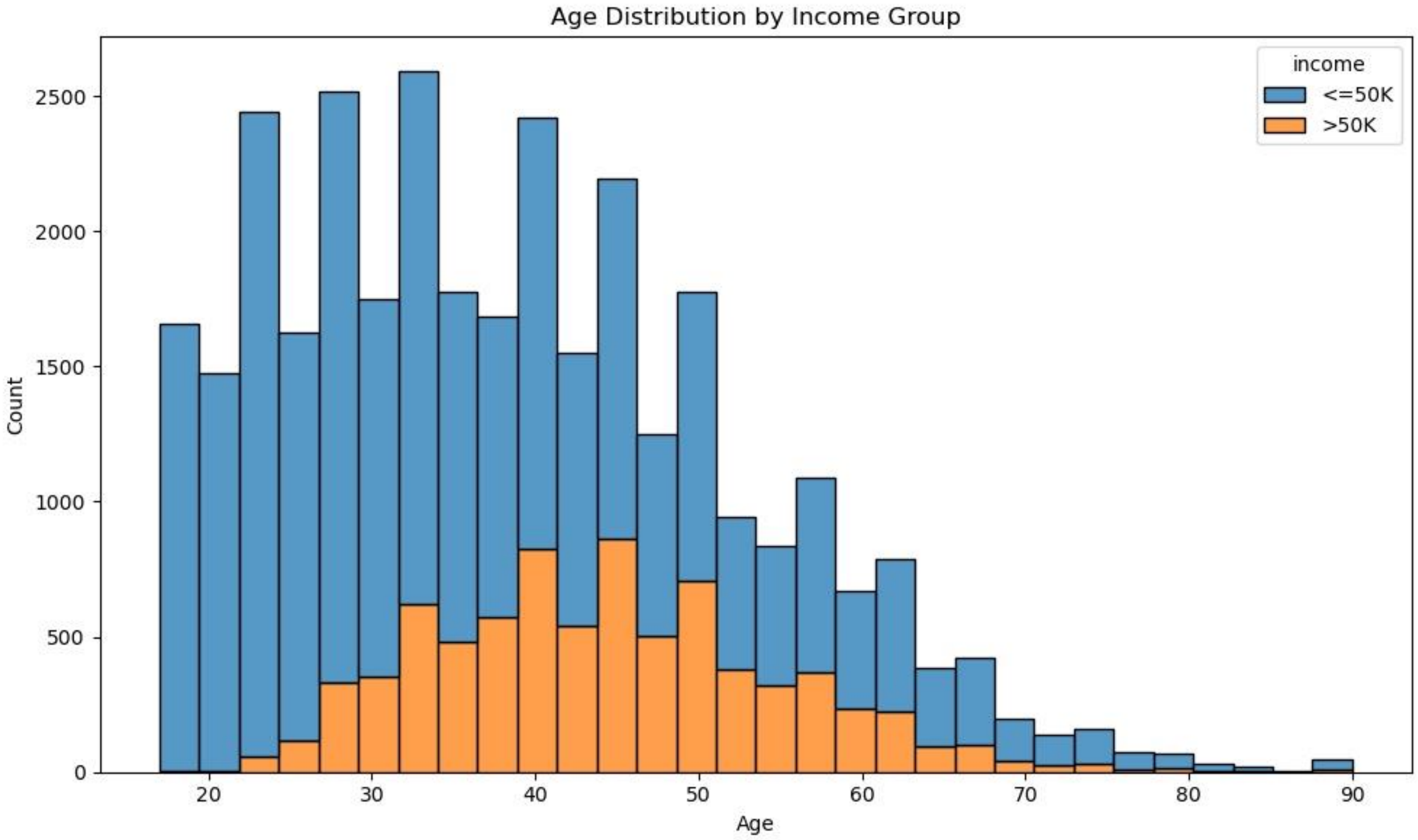
Data Analysis

The effect of education level on income



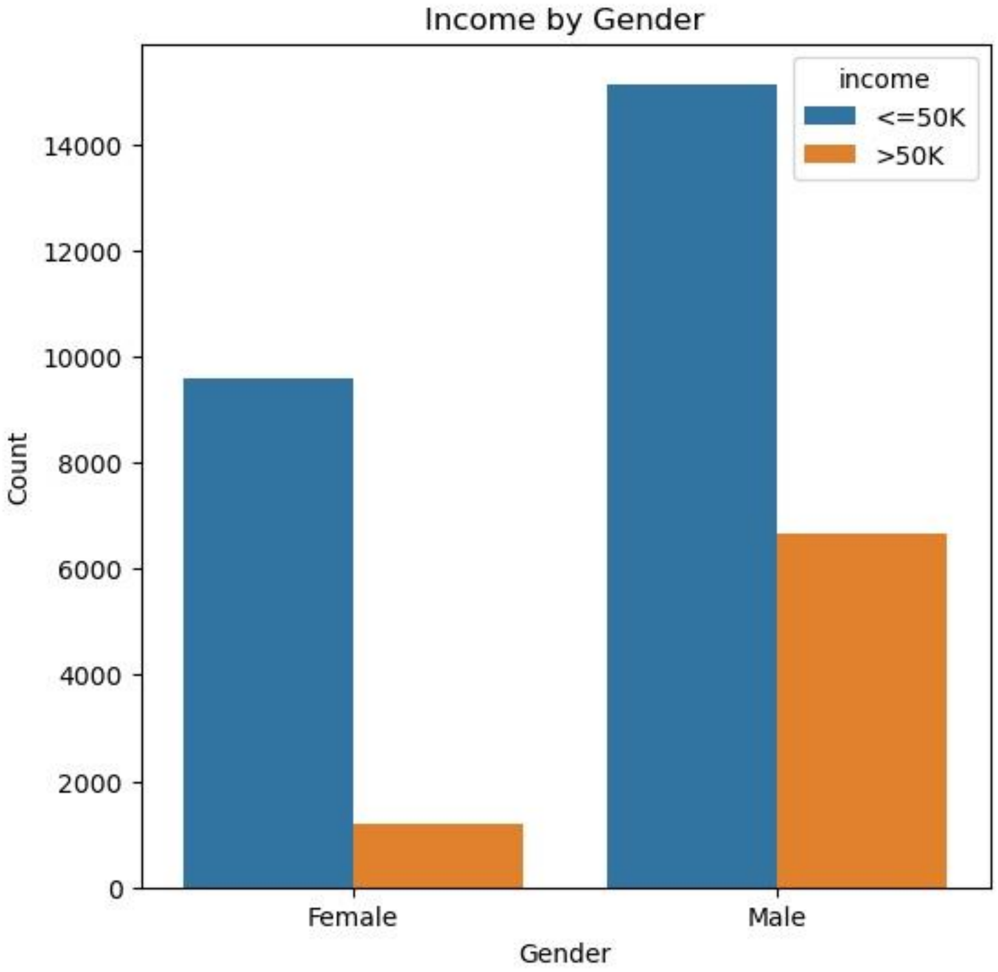
Data Analysis

The connection between age and income



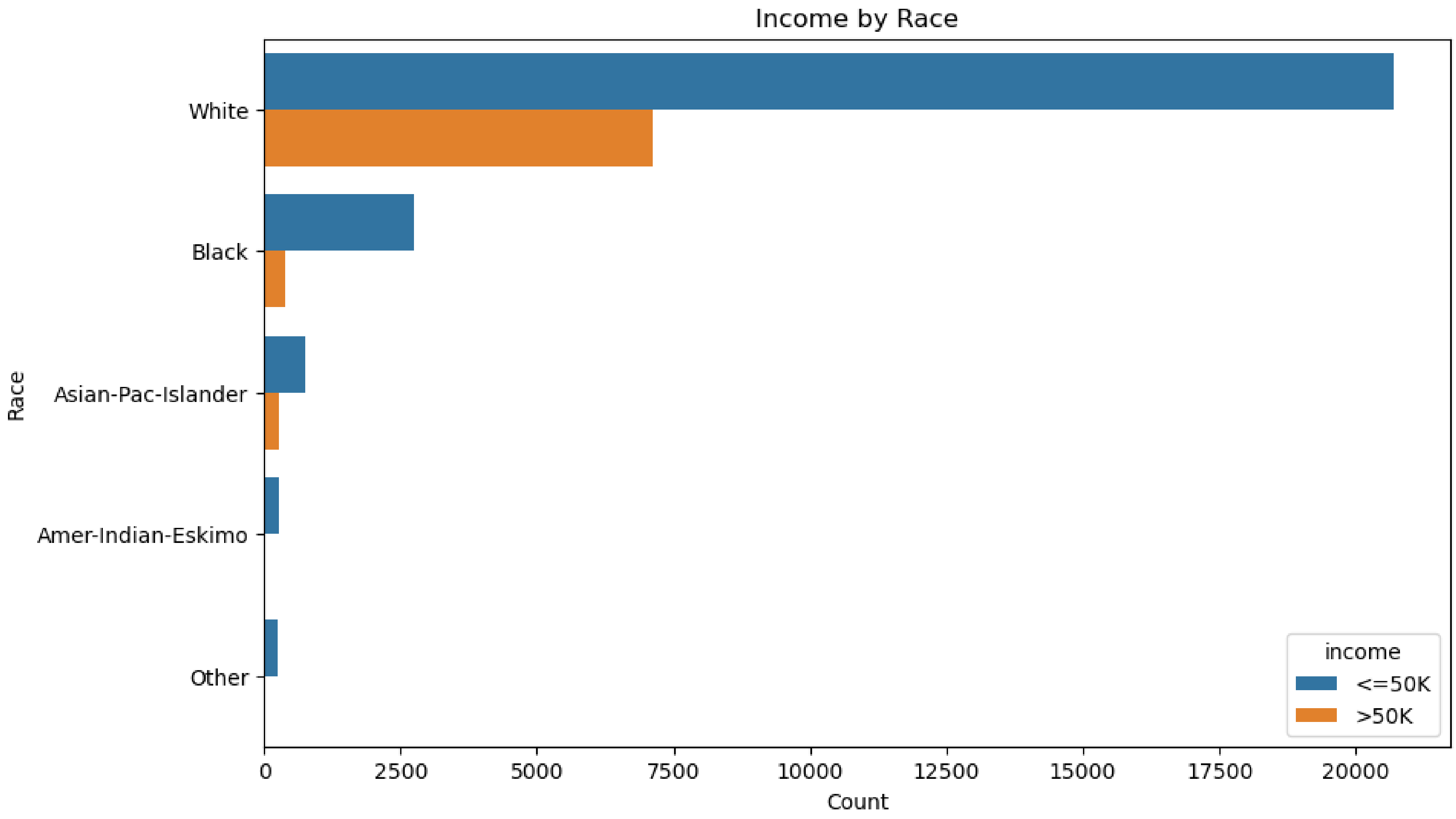
Data Analysis

Income differences between genders

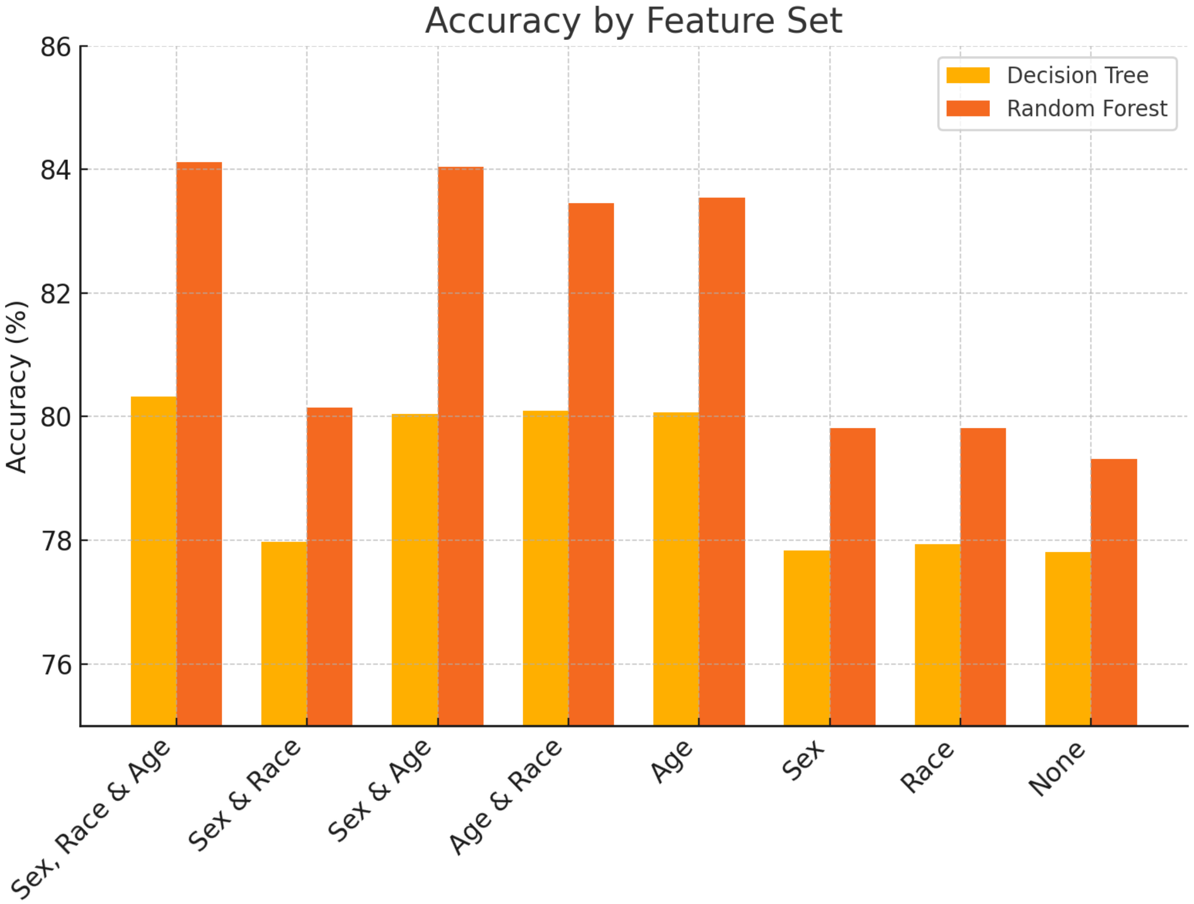


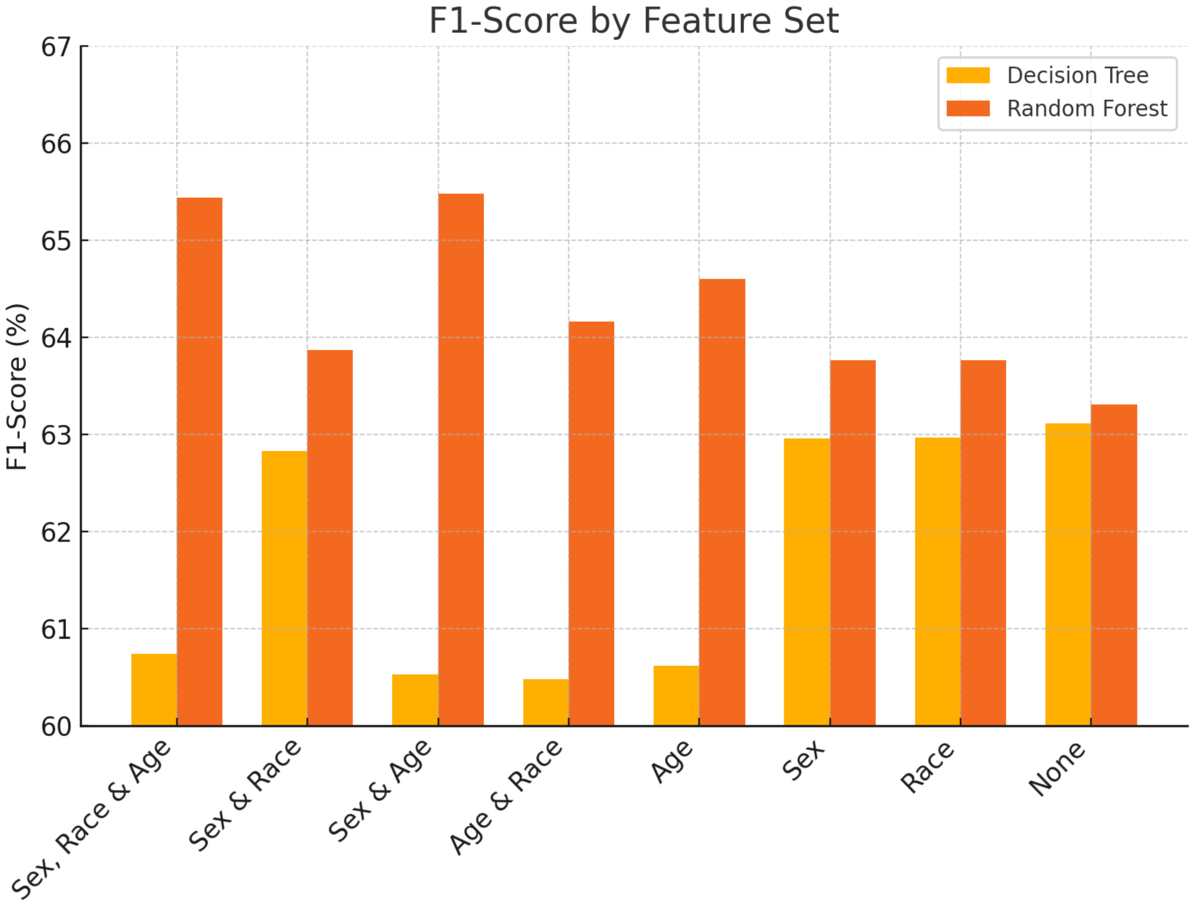
Data Analysis

Income differences among races and ethnic groups



Results – Accuracy & F1-Score

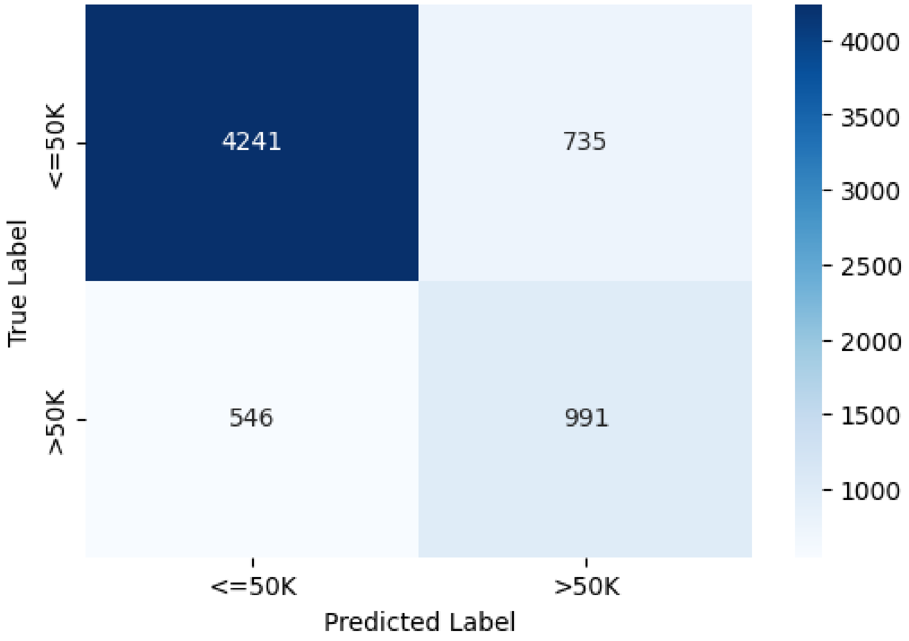




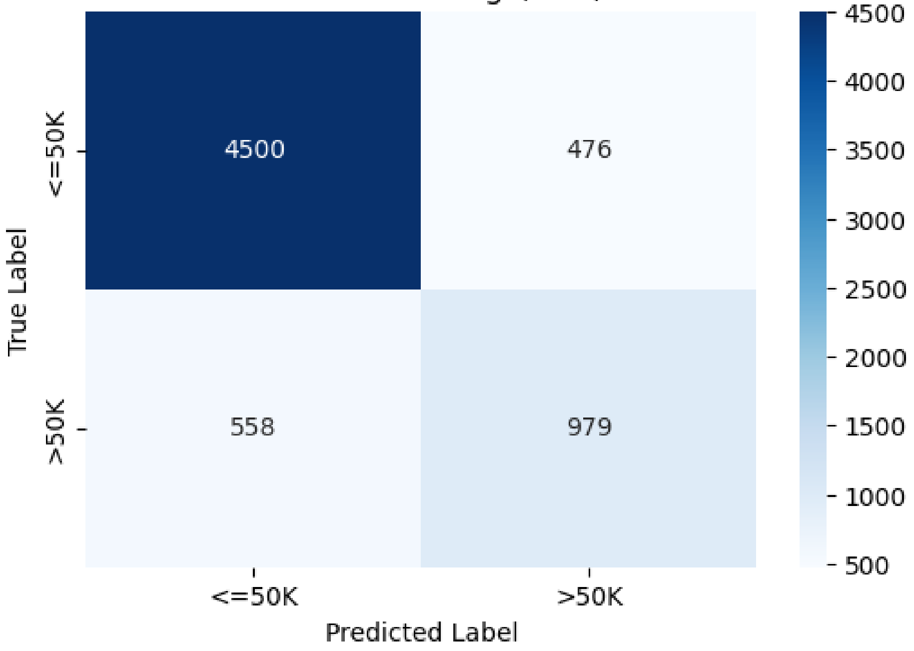
→ **F1-Score worse than Accuracy – Class Imbalance is the reason**
→ **Age much more important for Accuracy & F1-Score than Sex & Race**

Results – Confusion Matrices

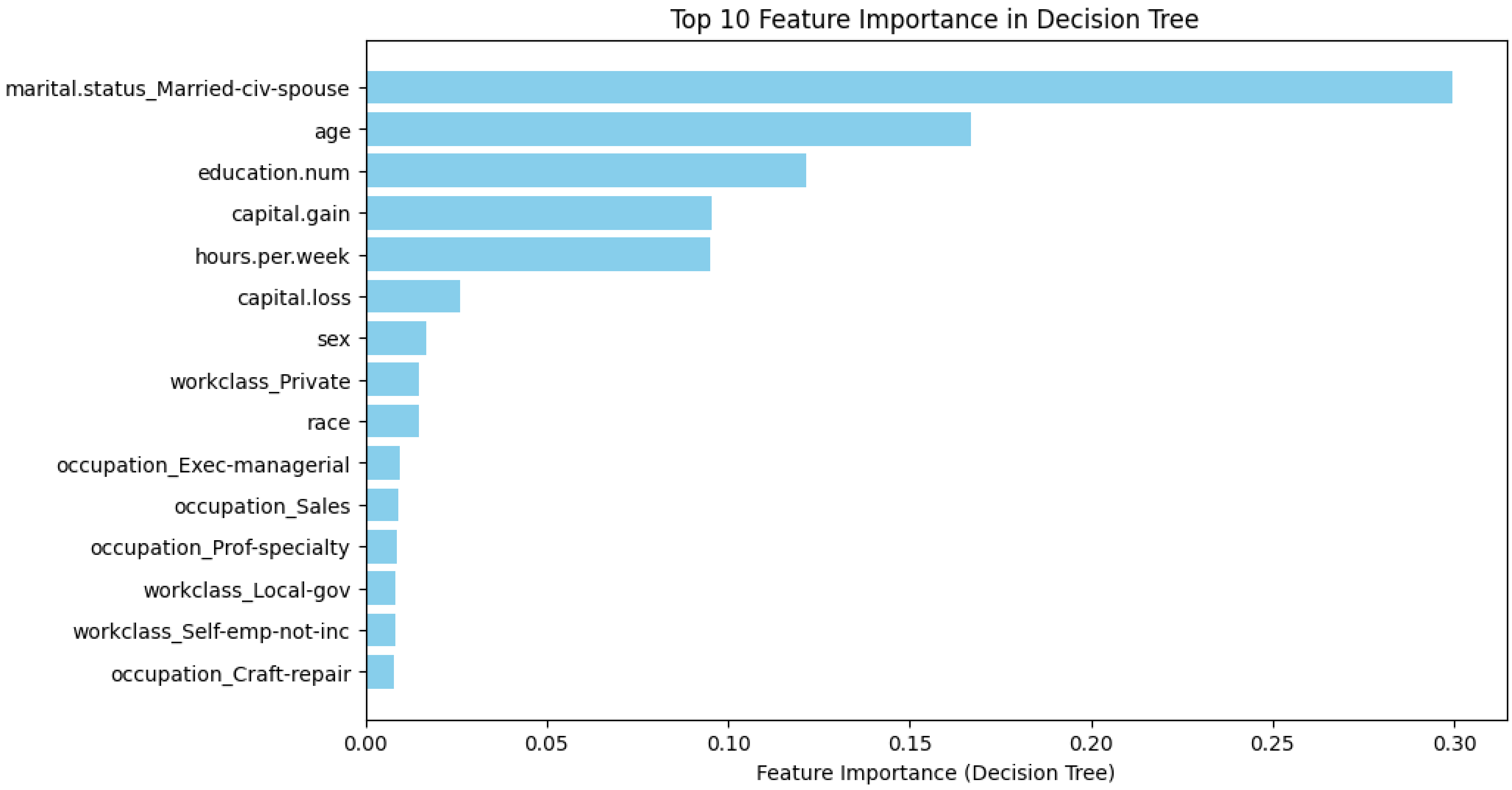
Confusion Matrix for Decision Tree with age, sex, race as sensitive features



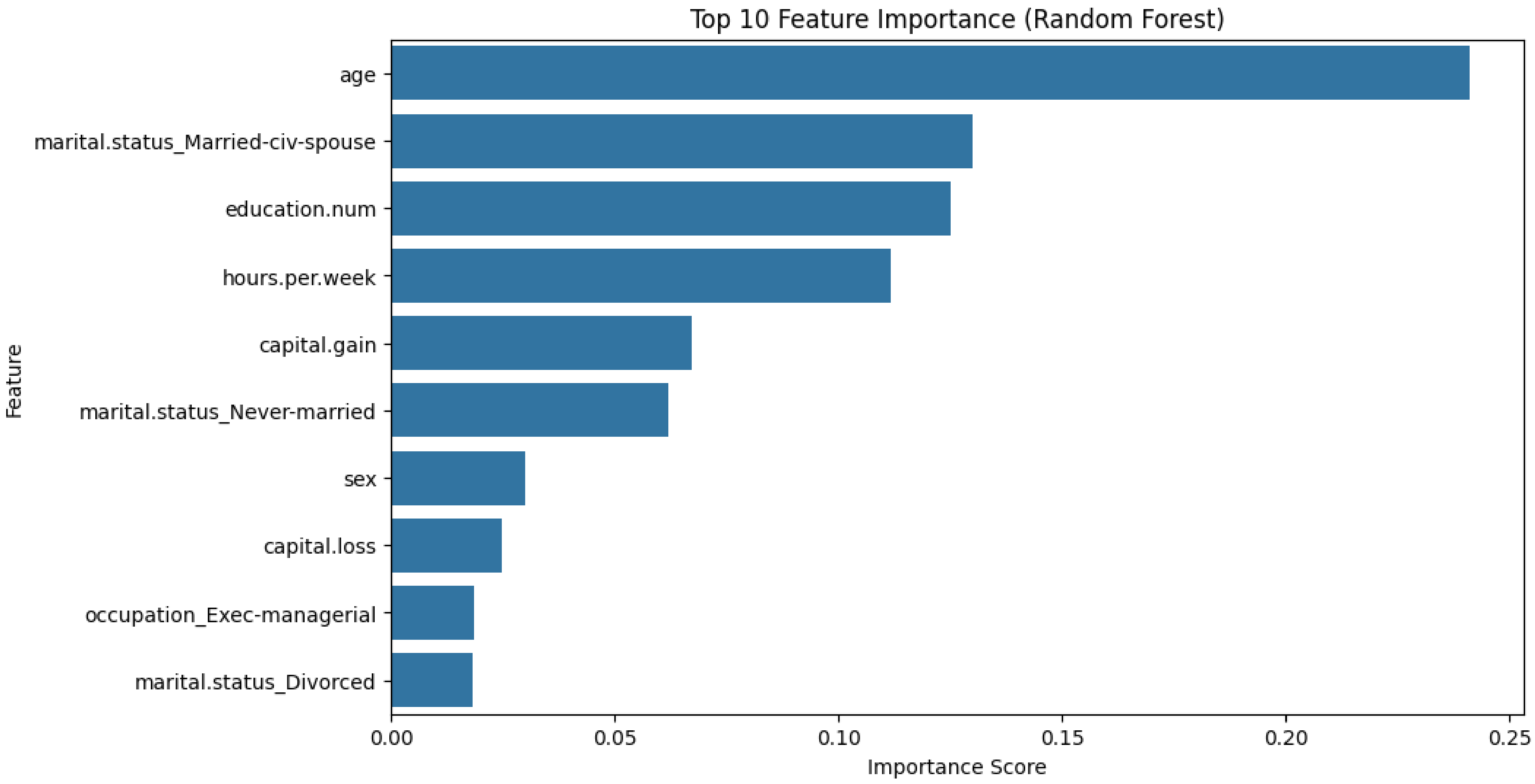
Confusion Matrix for Random Forest with age, sex, race as sensitive features



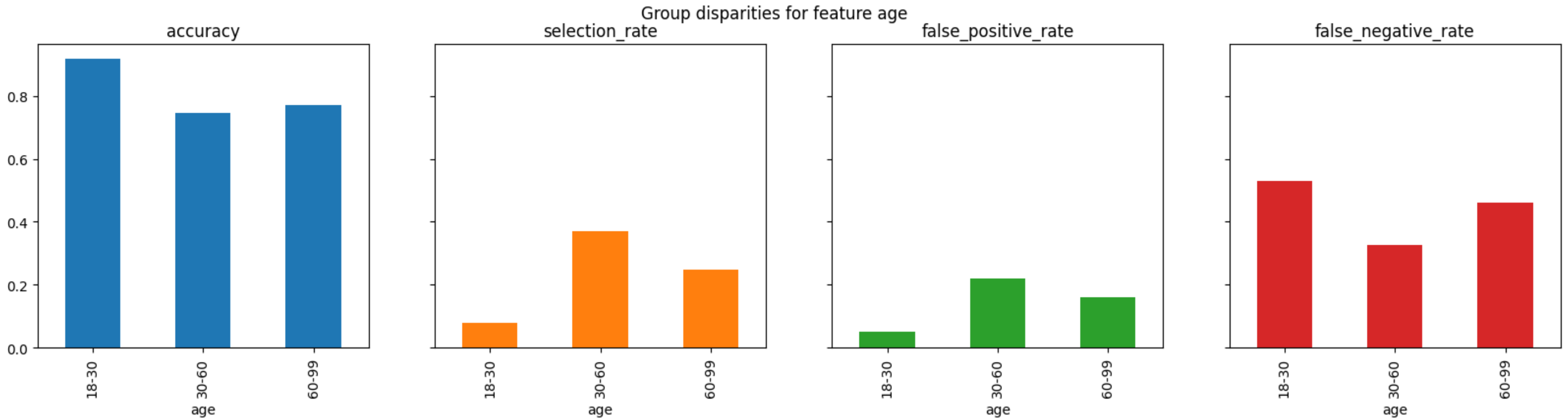
Feature Importance (Decision Tree)



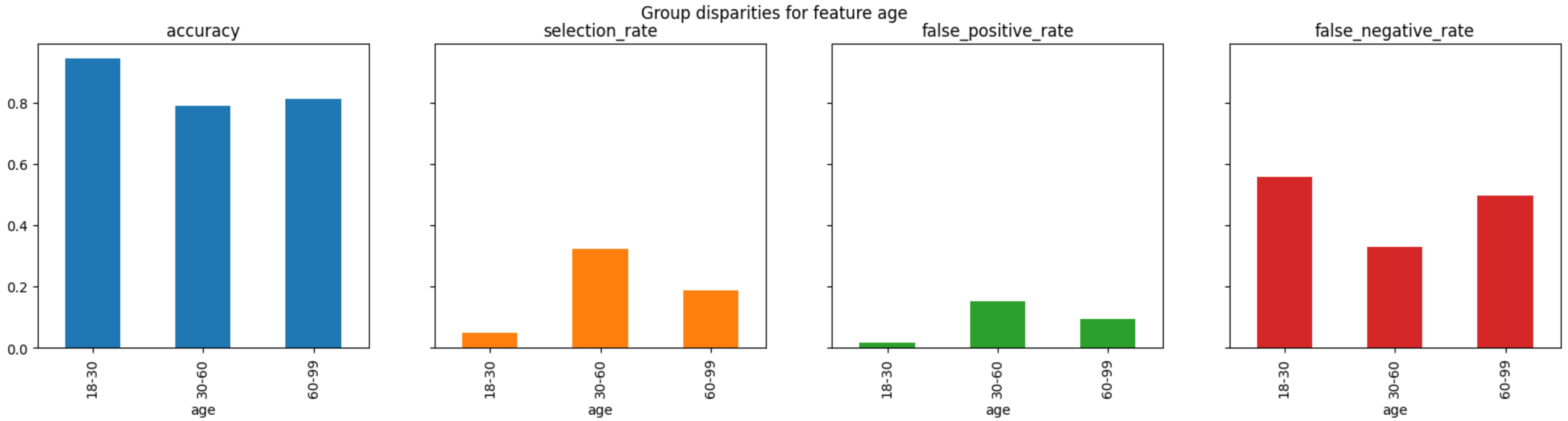
Feature Importance (Random Forest)



Fairness Evaluation (Age)

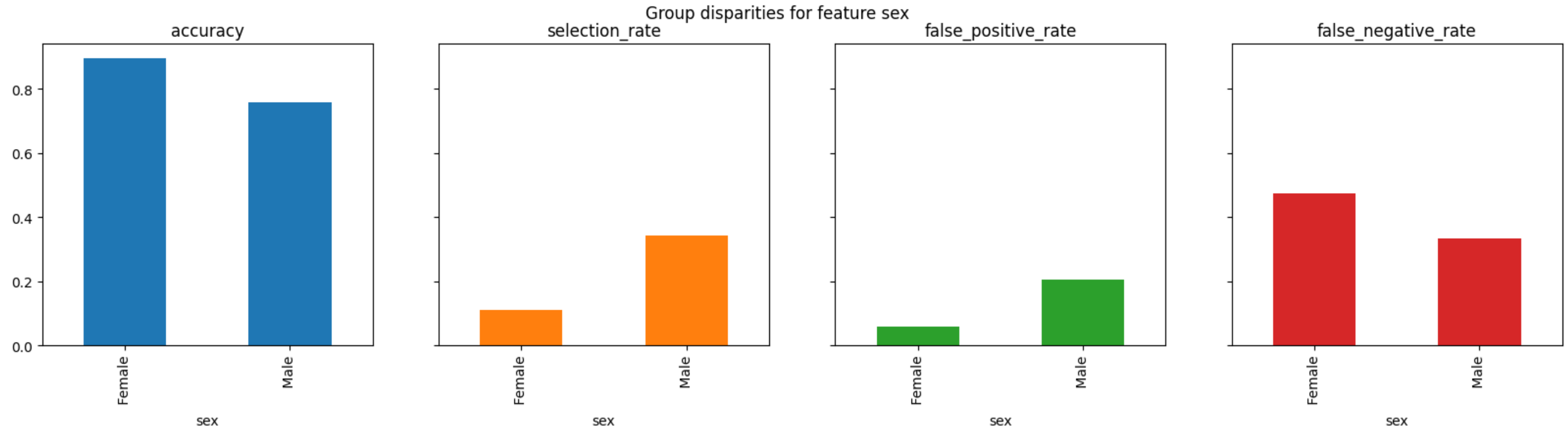


Decision Tree

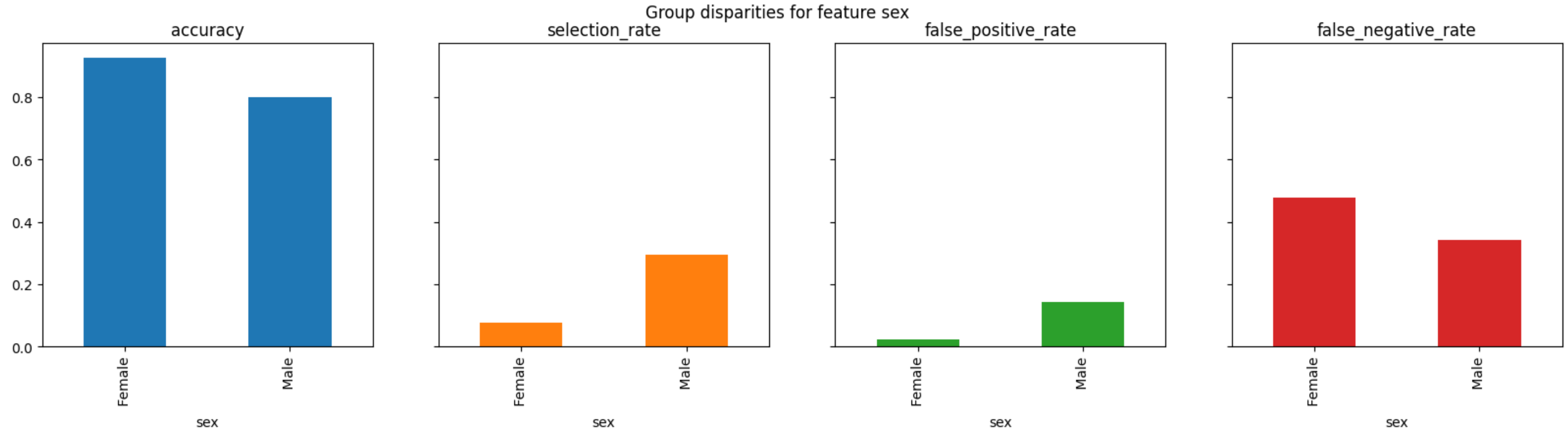


Random Forest

Fairness Evaluation (Sex)

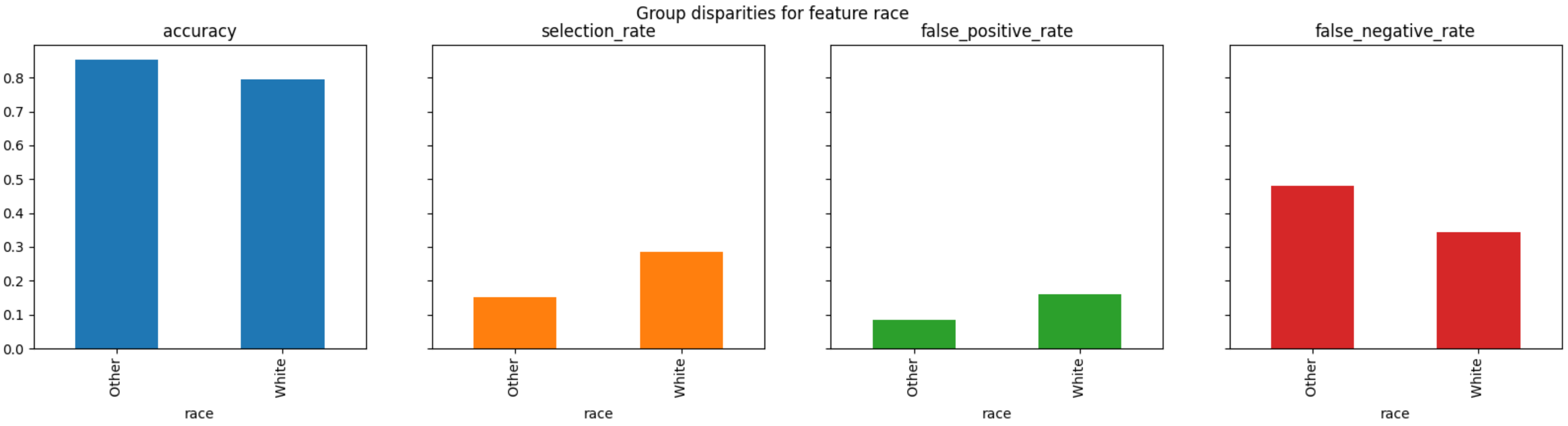


Decision Tree

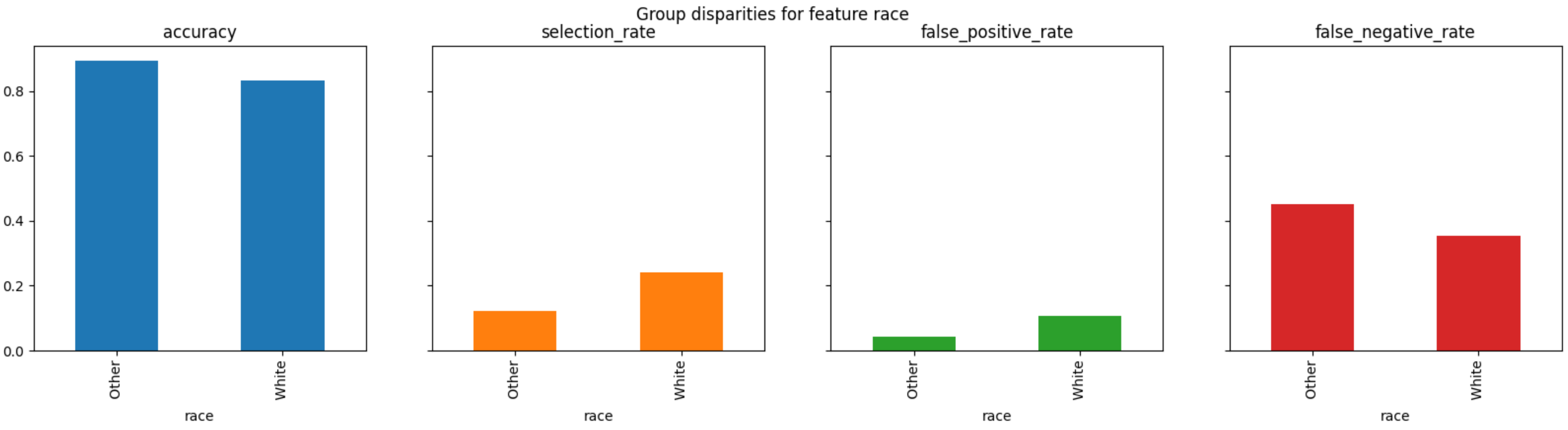


Random Forest

Fairness Evaluation (Race)



Decision Tree



Random Forest

Discussion: Research Question(s)

Reserach Question 1

Feature Selection and Importance

- Are sensitive attributes strongly considered?
  - **Yes, but not strongly as awaited**
- Can we simply remove them (without consequences for accuracy)
  - **No!**

Research Question 2

Metrics to measure the Fairness of Models

- What metrics are there to measure fairness?
  - **Selection Rate, False Negative & False Positive**
- Are these metrics meaningful and useful?
  - **Yes!**

Research Question 3

Trade-Off between Accuracy and Fairness?

- Models can **have different accuracy** but still have **same fairness**
- Fairness depends highly on **dataset** and **class balance**
- **Visualization** of **Feature Importance** helpful for first insights
- **Dataset:** Bias in dataset is easily learned by models
- **Class Imbalance:** Certain imbalance can decrease fairness of model

Conclusion & Outlook

Conclusion

- In conclusion, our group learned how predictive models for credit approval can prove to be both accurate and fair, using the Adult Census Income dataset. In the real world, sensitive attributes such as age, race, and sex are a significant roles, making this dataset ideal for analyzing the balance between accuracy and fairness in machine learning.

- We performed data analysis, implemented Decision Tree and Random Forest models, and evaluated fairness metrics to identify the best approach. Through this process, we gained valuable insights into the challenges and importance of balancing accuracy and fairness in predictive modeling.

Literatures

Literatures:

- **Sarkar, T., Rakhra, M., Sharma, V. & Singh, A. (2024).** An Empirical Comparison of Machine Learning Techniques for Bank Loan Approval Prediction. International Conference On Communication, Computer Sciences And Engineering, 137–143. https://doi.org/10.1109/ic3se62002.2024.10593355
- **Alagic, A., Zivic, N., Kadusic, E., Hamzic, D., Hadzajlic, N., Dizdarevic, M. & Selmanovic, E. (2024).** Machine Learning for an Enhanced Credit Risk Analysis: A Comparative Study of Loan Approval Prediction Models Integrating Mental Health Data. Machine Learning And Knowledge Extraction, 6(1), 53–77. https://doi.org/10.3390/make6010004
- **Peela, H. V., Gupta, T., Rathod, N., Bose, T. & Sharma, N. (2021).** Prediction of Credit Card Approval. International Journal Of Soft Computing And Engineering, 11(2), 1–6. https://doi.org/10.35940/ijsce.b3535.0111222.

Questions & Answer

Nichakorn Wisalkamol
Teerada Chatrattanawuth
Thanaporn Artidtayamontol
Rawisara Chantravutikorn
Emmanuel Gyau-Amoakoh