Classifying opinions on Reproductive Rights

Business Problem - Advertising by Planned Parenthood

- Identify groups that are more likely to support legal abortion

- Develop classification model to predict a whether a given person would support legal abortion

- Formulate a strategy for optimal advertising
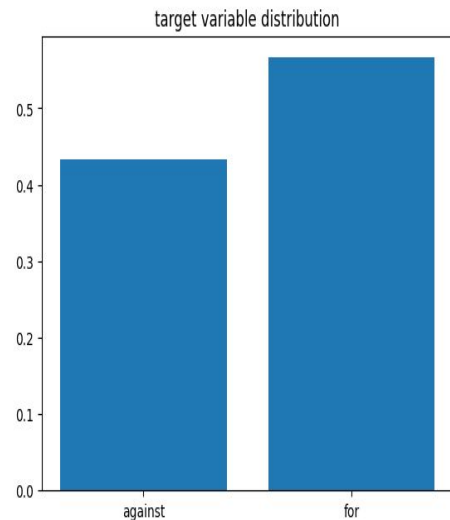
The Data

Source:

GSS by NORC at Uchicago

"The most rigorous, widely used data on the attitudes, opinions, and behaviors of the American public"

- The NORC website

About the dataset

- Results from **survey** conducted in **2018**
- More than 2000 respondents
- Random sampling
- Responses to **23 select questions** were examined

target variable distribution

Methods

- Experimented with **4 different classifiers**

- Modified input information to **highlight outliers**

- Evaluated performance by testing model on **unseen data**

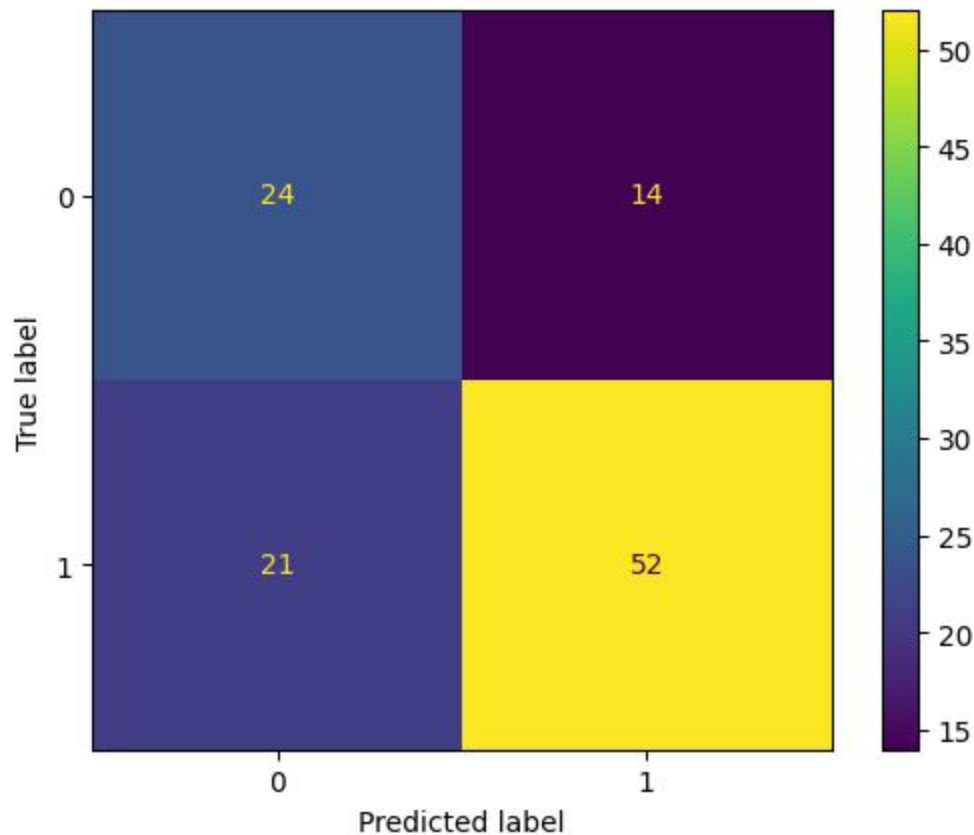Final model: **Random Forests Classifier**

About the Random Forest Classifier

- **Combines** predictions from multiple **independent models**
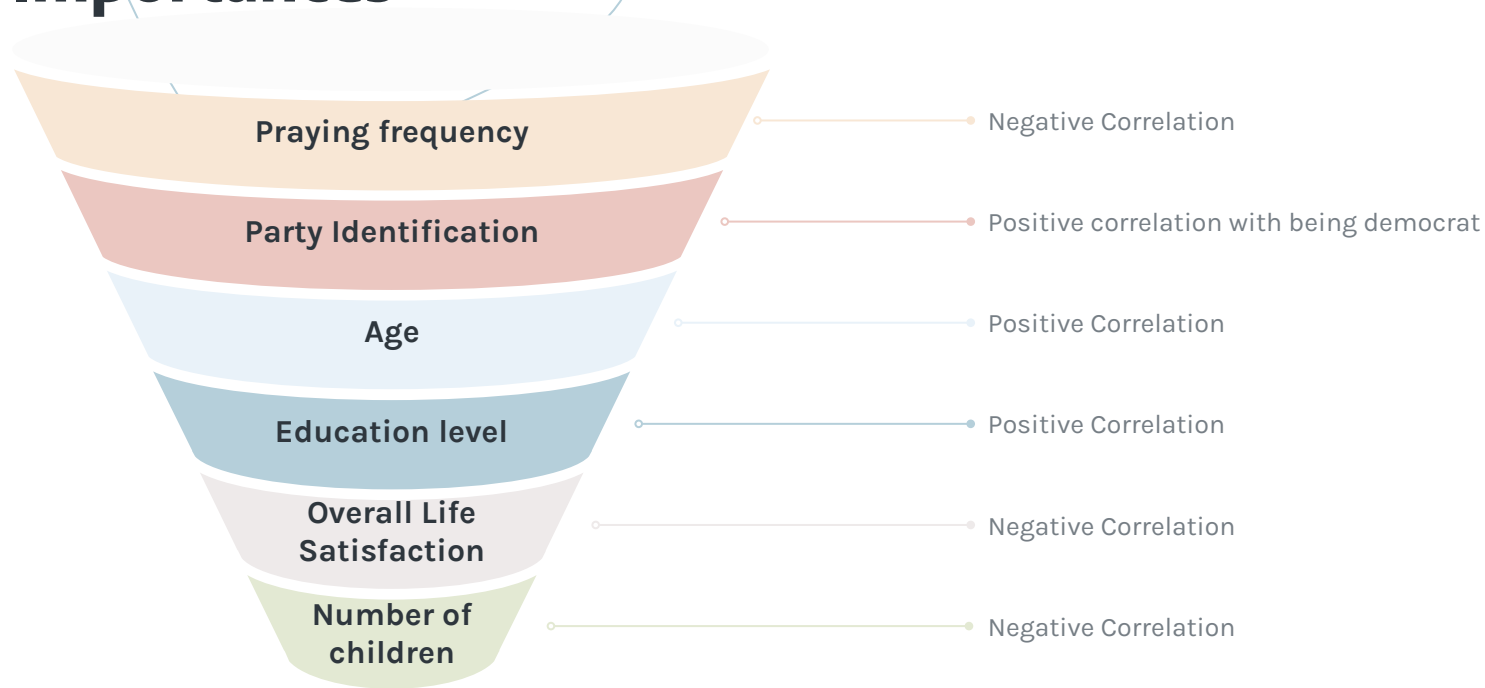- Each model determines **optimal ways to categorize data**

Random Forest Simplified
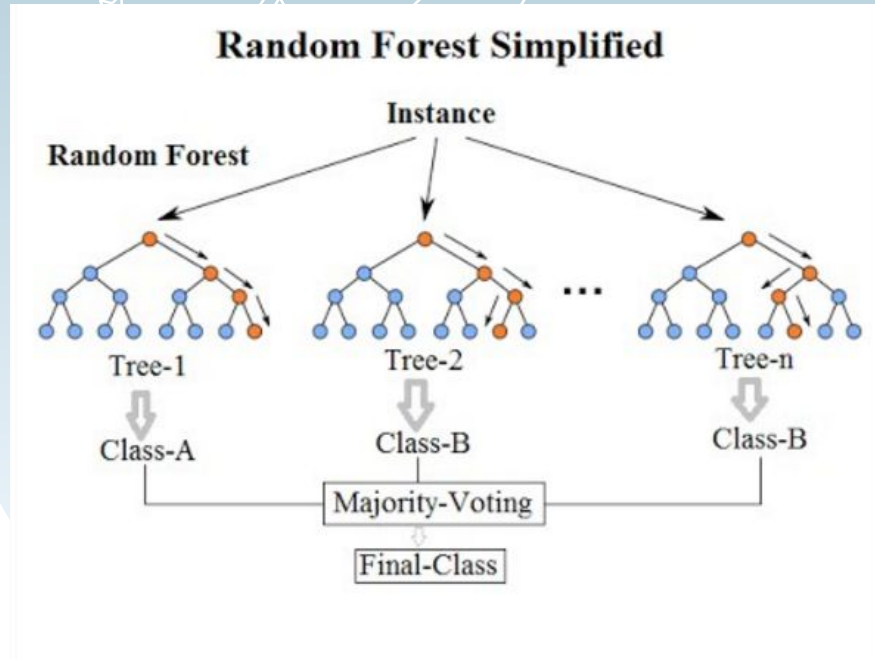
Results- 71% of supportive persons captured

Classification results of final model

Feature Importances

Praying frequency — Negative Correlation

Party Identification — Positive correlation with being democrat

Age — Positive Correlation

Education level — Positive Correlation

Overall Life Satisfaction — Negative Correlation

Number of children — Negative Correlation

Recommendations

Avoid targeting very religious groups

Eg. Southwest regions of the US

Look for people who are more mature

Eg. Facebook over Instagram

Pray **Age**

Educ **Party ID**

Look for people who are highly educated

Eg. College towns

Look for people identifying as democrats

Eg. urban cities with a diverse population

➔ Can also use model to weed out prospective customers directly

Thanks!

Any questions?

You can find me at @beatrix.wmh@hotmail.com