Targeting Political Ads to 3rd Party Voters on Reddit:
An Initial Model

Data Scientists, Inc.

Agenda

- Goal
- Model Approach
- Model Results
- Conclusions
- Further Model Development
- Questions?

Goal

- The USA has two major parties- the Republicans (conservative), and the Democrats (liberal).
- The parties need to sway third-party and undecided voters in order to win the election.
- The biggest third-party in the USA is the Libertarians.
- Their views are based on maximizing individual freedoms and limiting the role of government.
  - Can lean conservative (e.g., decreasing welfare).
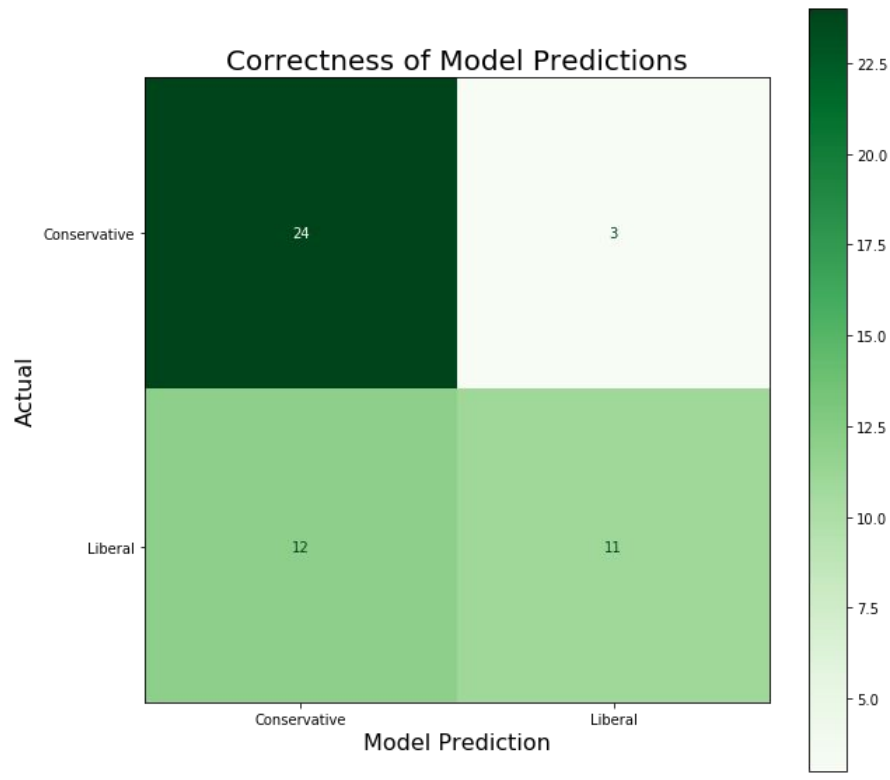 - Can lean liberal (e.g., legalization of cannabis).

Goal

- Determine whether the users of the libertarian subreddit are more liberal or conservative, based on their comments.
- Sell this data to the RNC and DNC so that they can market ads towards the conservative-leaning libertarians and the liberal-leaning libertarians- the RNC and DNC will buy ads to influence these voters in key battleground states for the election.

Model Approach

- Pull comments from the r/politics subreddit and labeling them liberal.
- Pull comments from the r/conservative subreddit and labeling them conservative.
- Parts of the comments that were included:
  - Words (e.g., "healthcare" or "family")
  - Numbers (e.g., "2016" or "1776")
  - Domain names (e.g., NYTimes.com or FoxNews.com)
- Classify a sample of r/libertarian posts as conservative or liberal, and test the model.

Model Results

- The model accurately identifies comments 70% of the time.
- Guessing all the comments were conservative would be correct 54% of the time.

Correctness of Model Predictions

Model Results

- Characteristic words (from up to 1% of comments):
  - Liberal: 'yet', 'fucked', 'days', 'death', 'different', 'donald', 'easy', 'economy', 'elections', 'entire', 'fine', 'fox', 'full', 'white', 'given', 'god', 'gop', 'guns', 'hate', 'healthcare'
  - Conservative: '[new] york', 'gets', 'dying', 'each', 'essential', 'except', 'experts', 'face', 'family', 'fauci', 'fever', 'force', 'global', 'willing', 'god', 'goes', 'guys', 'hell', 'higher', 'hospital'
  - Libertarian: 'italy', 'paid', 'worst', 'gouging', 'data', 'deal', 'demand', 'distancing', 'doctor', 'effective', 'essential', 'false', 'goes', 'gun', 'grocery', 'lower', 'haven'[t]', 'her', 'idiot', 'idiots'

Conclusions

- The words used by liberal and conservative libertarians in the model had significant overlap, but there were some characteristic differences.
- Applying the model by username, 35.5% of libertarian users leaned liberal, and 64.5% leaned conservative.
- Targeting advertisements towards specific libertarian users will be more effective than paying for random advertisements.
- This model could be combined with libertarian users' location data to increase effectiveness of targeted advertisements in battleground states.

Further Model Development

- With more funding, our company will further refine the model for Reddit. Some possible improvements include:
  - Adding far-right and far-left subreddit comment data for training the model.
  - Tuning more components of the model, such as the maximum number of filler words that the model considers.
  - Crowd-sourcing data on political slant for the training dataset, instead of assuming that commenters in r/politics are liberal and commenters in r/conservative are conservative.

Questions?