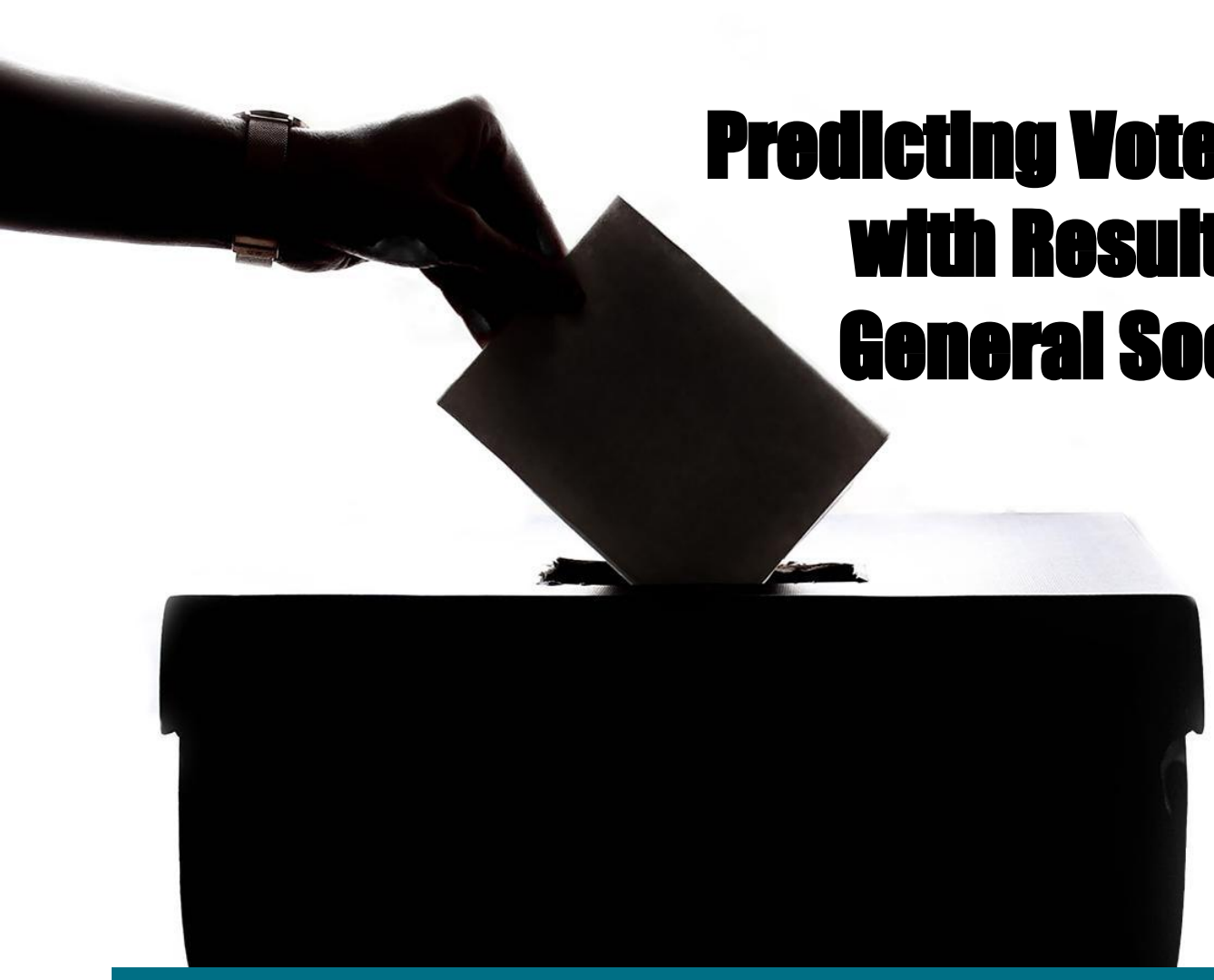


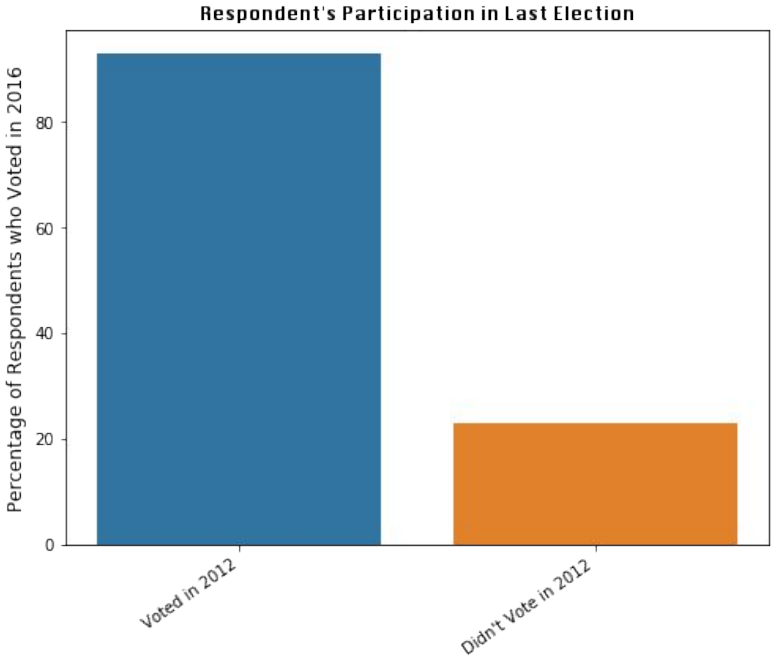
Predicting Voter Participation with Results from the General Social Survey

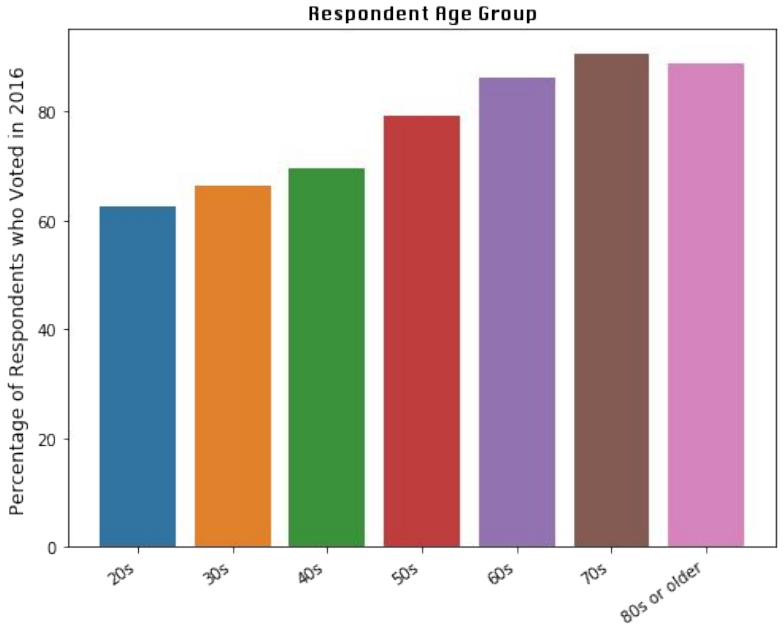
Avonlea Fisher
October 2020

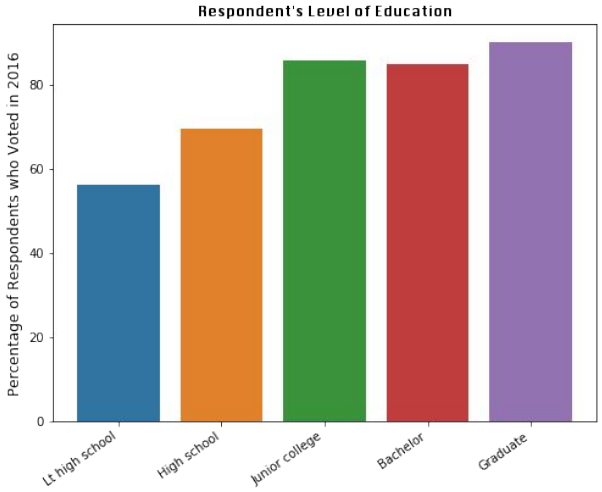
The Data

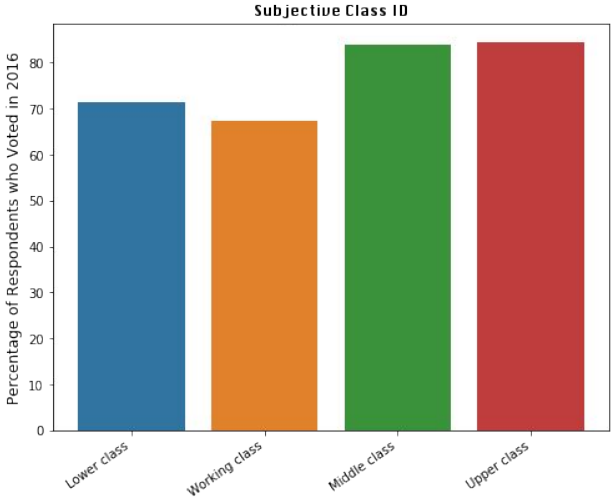
- 2018 General Social Survey (GSS)
- Variables: respondents' participation in last two elections, socio-economic info, and political attitudes
- After data cleaning:
  - 1,083 rows (survey respondents)
  - 18 columns (survey questions)
  - only eligible voters who remembered voting history

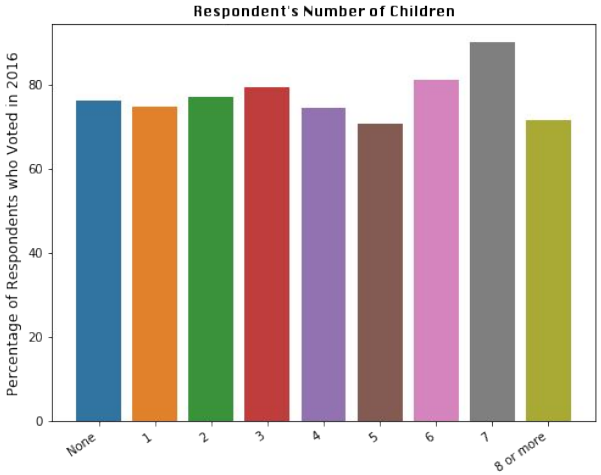
Percentage of 2016 Voters by Predictor

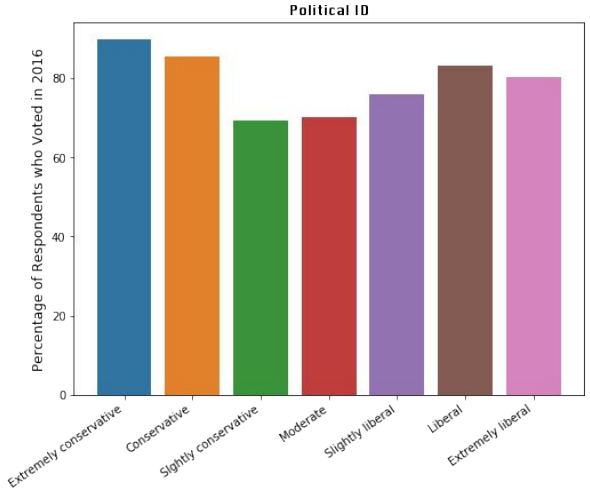




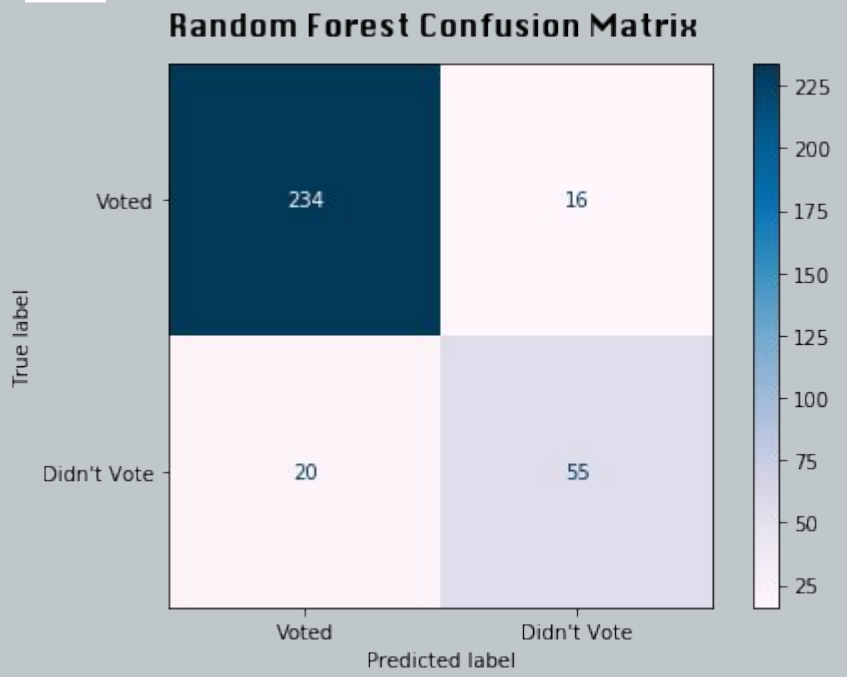

Respondent's Level of Education
Percentage of Respondents who Voted in 2016
0
20
40
60
80
Lt high school
High school
Junior college
Bachelor
Graduate


Subjective Class ID
Percentage of Respondents who Voted in 2016
0
10
20
30
40
50
60
70
80
Lower class
Working class
Middle class
Upper class


Respondent's Number of Children
Percentage of Respondents who Voted in 2016
0
20
40
60
80
None
1
2
3
4
5
6
7
8 or more


Political ID
Percentage of Respondents who Voted in 2016
0
20
40
60
80
Extremely conservative
Conservative
Slightly conservative
Moderate
Slightly liberal
Liberal
Extremely liberal

Best Classification Model: Random Forest

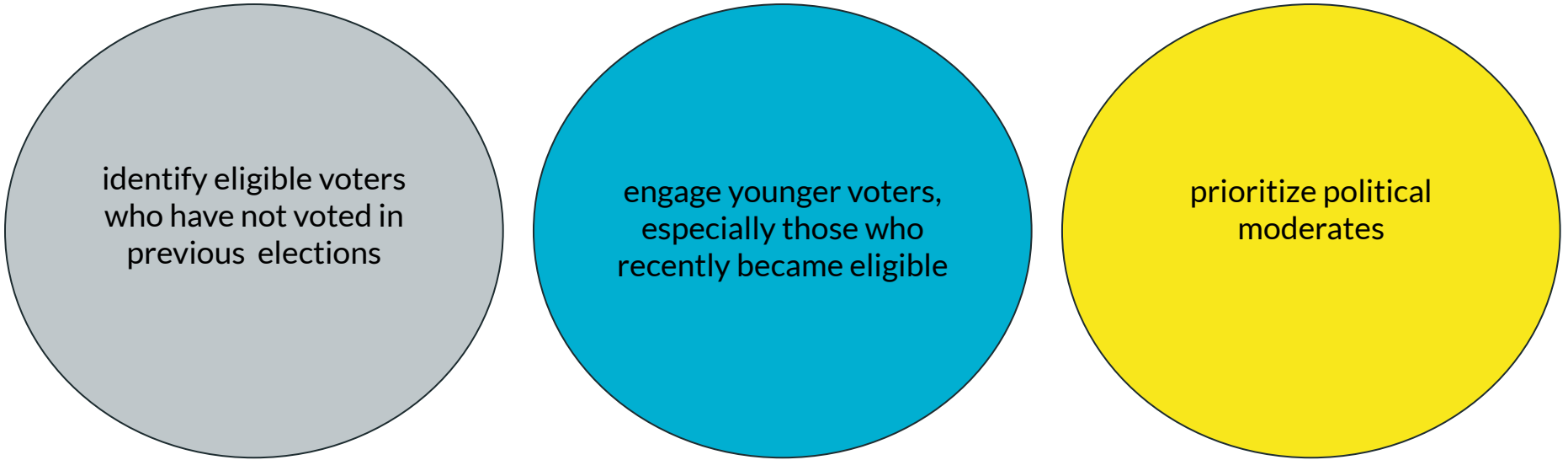


Accuracy: 89%

True Positive Rate: 73%

Recommendations

To maximize the number of potential non-voters that you reach:



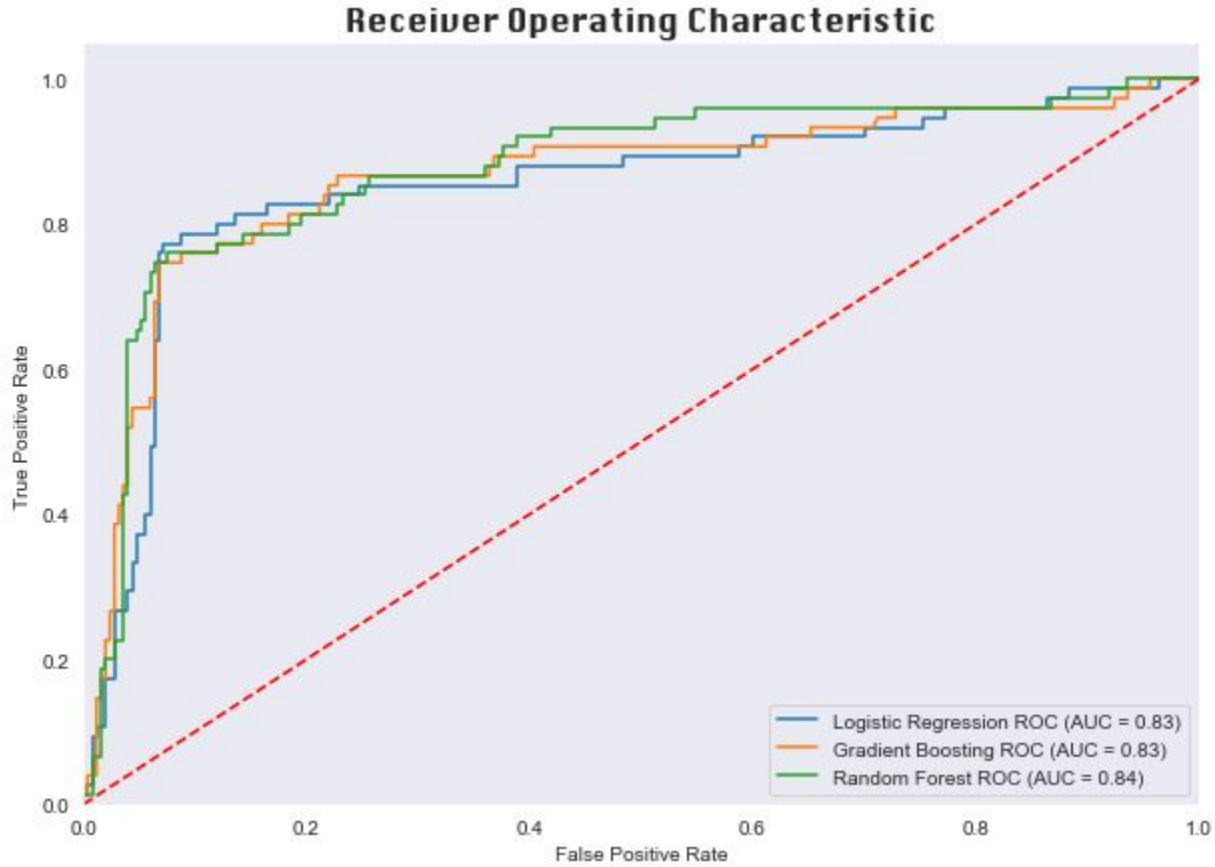
Questions for Future Work

- Which voter engagement strategies are most effective at increasing voter turnout?
- What are the most significant barriers to voting among eligible voters who don’t vote, but report an interest in doing so?


Return Envelope
OFFICIAL ELECTION MAIL
Authorized by the U.S. Postal Service
BUSINESS REPLY MAIL
FIRST-CLASS MAIL
PERMIT NO. 74
POSTAGE WILL BE PAID BY ADDRESSEE
NO POSTAGE NECESSARY IF MAILED IN THE UNITED STATES

Thank you!

Appendix



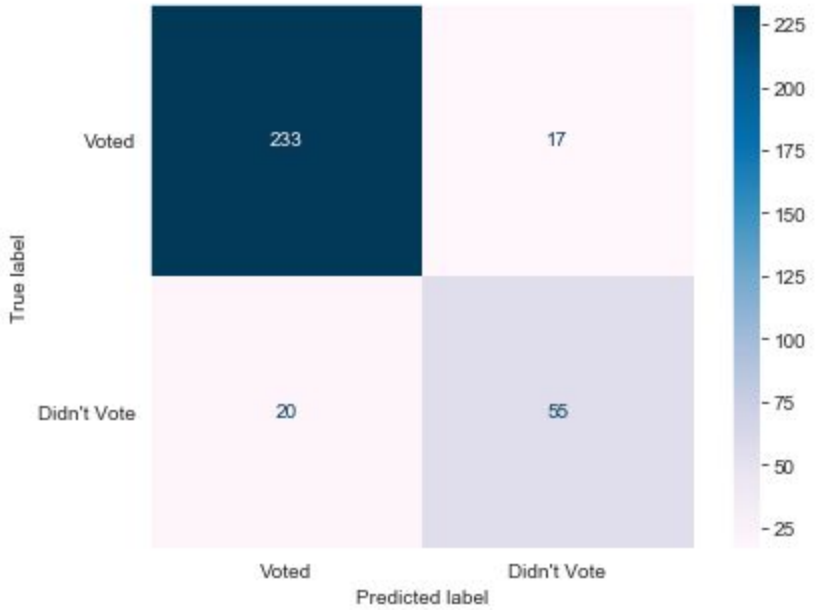
Recall Scores

Logreg: 0.733

Gradient Boosting: 0.72

Random Forest: 0.733

Logistic Regression Confusion Matrix



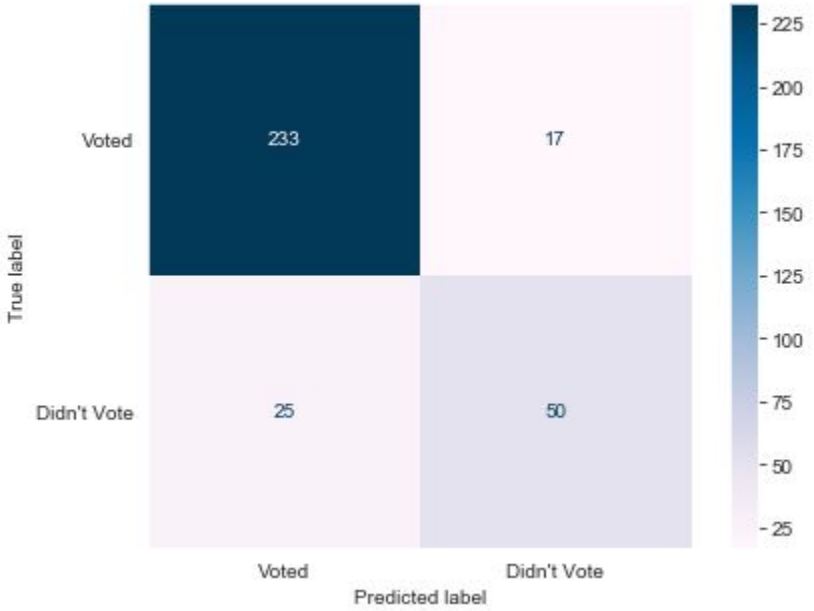
Accuracy: 0.8862

Precision: 0.7639

Recall: 0.7333

F1: 0.7483

Gradient Boosting Confusion Matrix

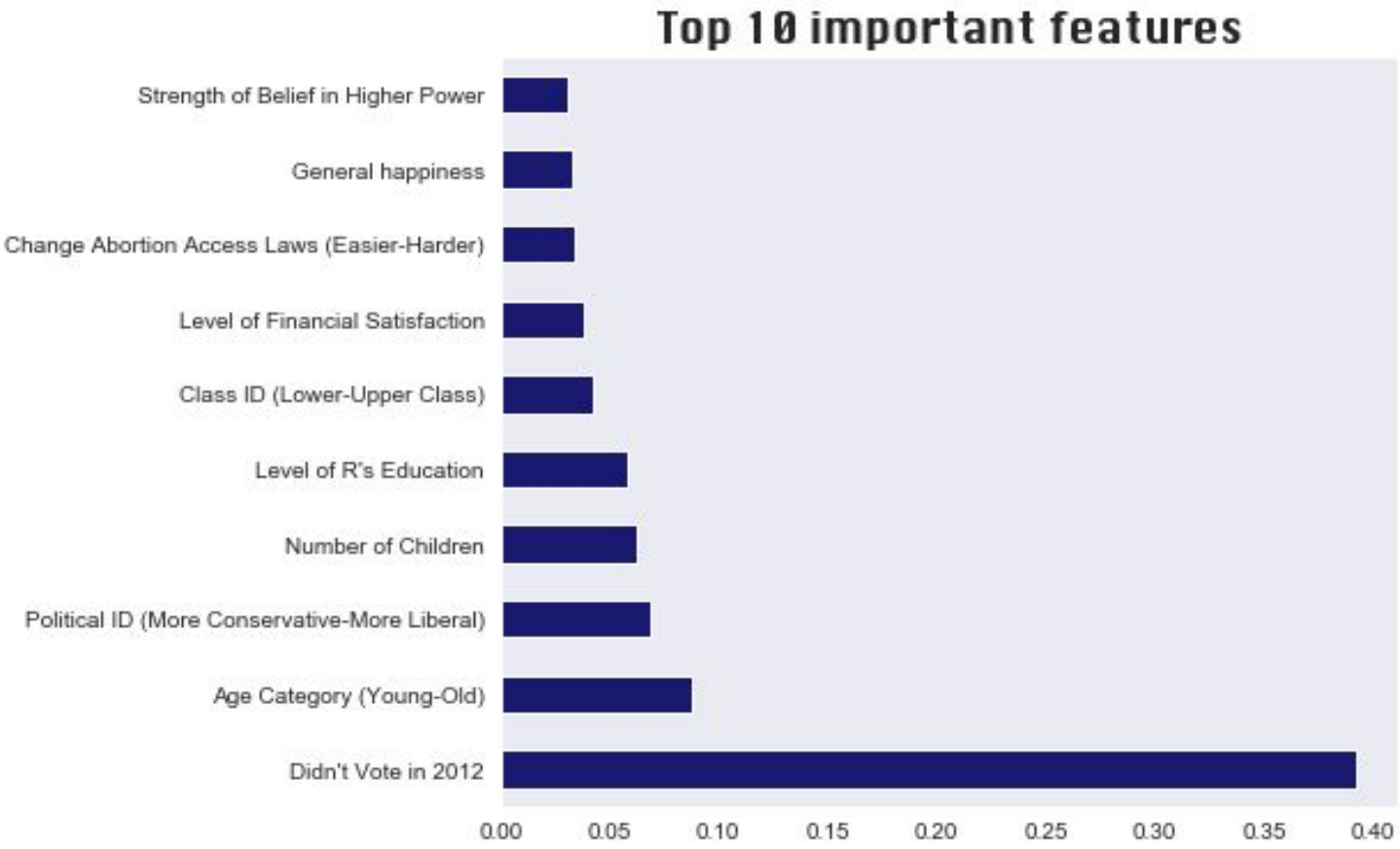


Accuracy: 0.871

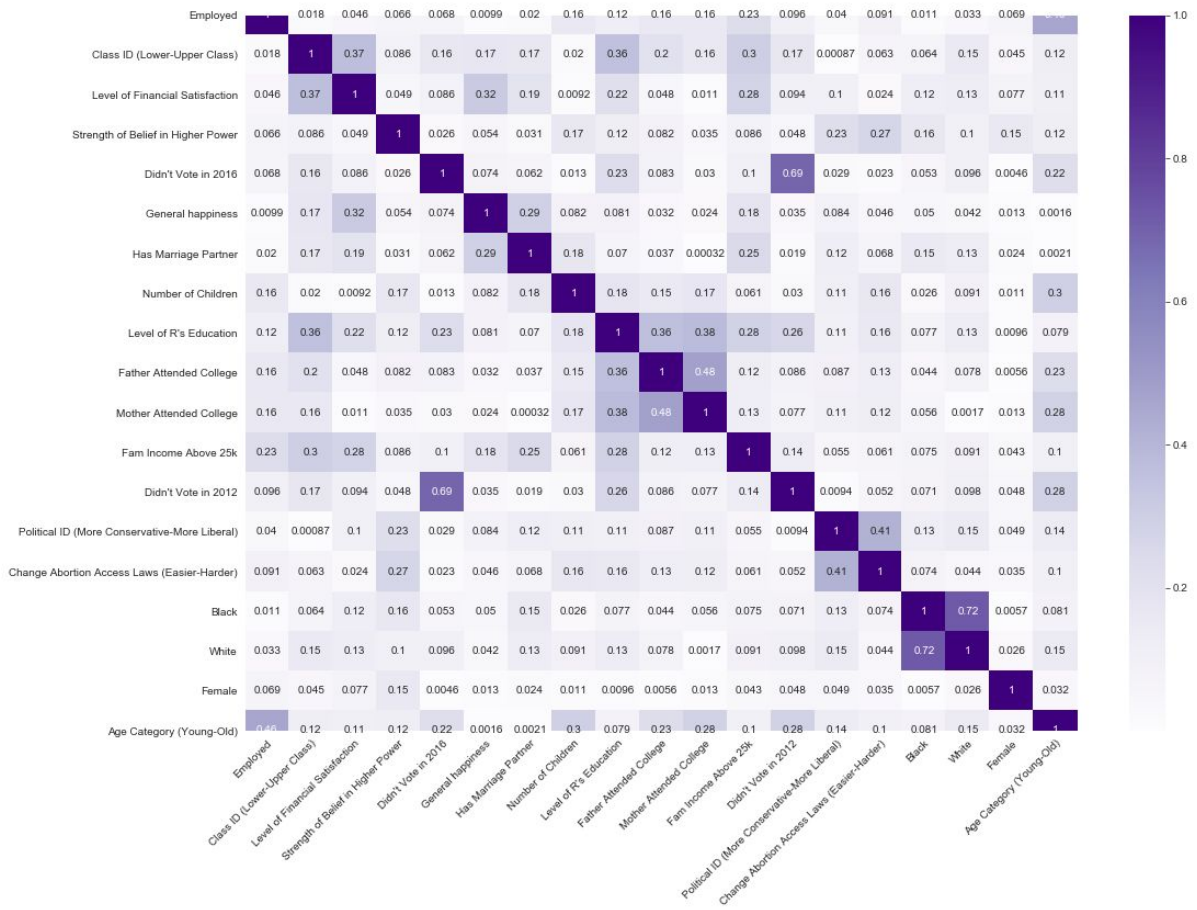
Precision: 0.746

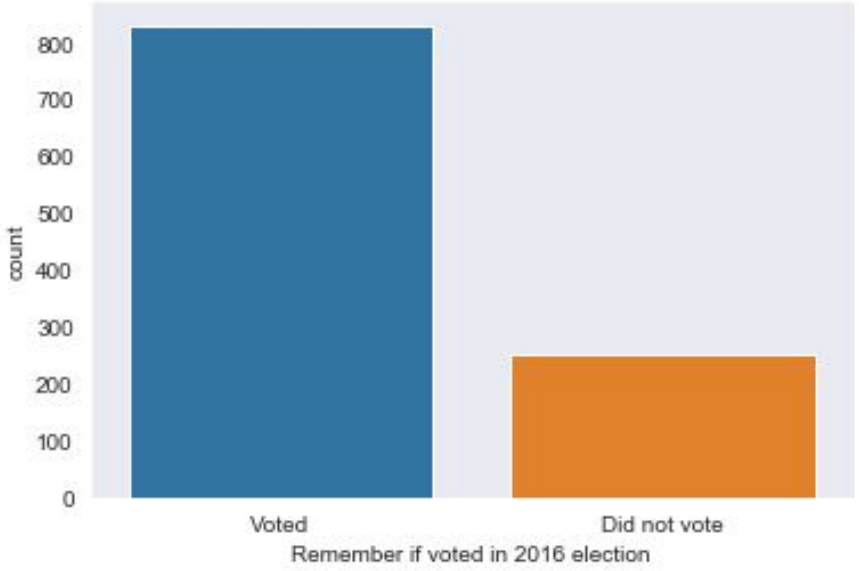
Recall: 0.667

F1: 0.704


Top 10 important features
Strength of Belief in Higher Power
General happiness
Change Abortion Access Laws (Easier-Harder)
Level of Financial Satisfaction
Class ID (Lower-Upper Class)
Level of R's Education
Number of Children
Political ID (More Conservative-More Liberal)
Age Category (Young-Old)
Didn't Vote in 2012
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40

Variable Correlations




800
700
600
500
400
300
200
100
0
count
Voted
Did not vote
Remember if voted in 2016 election

Credits

- Presentation theme from PoweredTemplate.com
- Mail-in ballot image by Tiffany Tertipes on Unsplash