A high-contrast silhouette image showing a hand from the left, wearing a watch, placing a ballot into a dark ballot box. The background is white, and the ballot box is a solid black shape at the bottom.

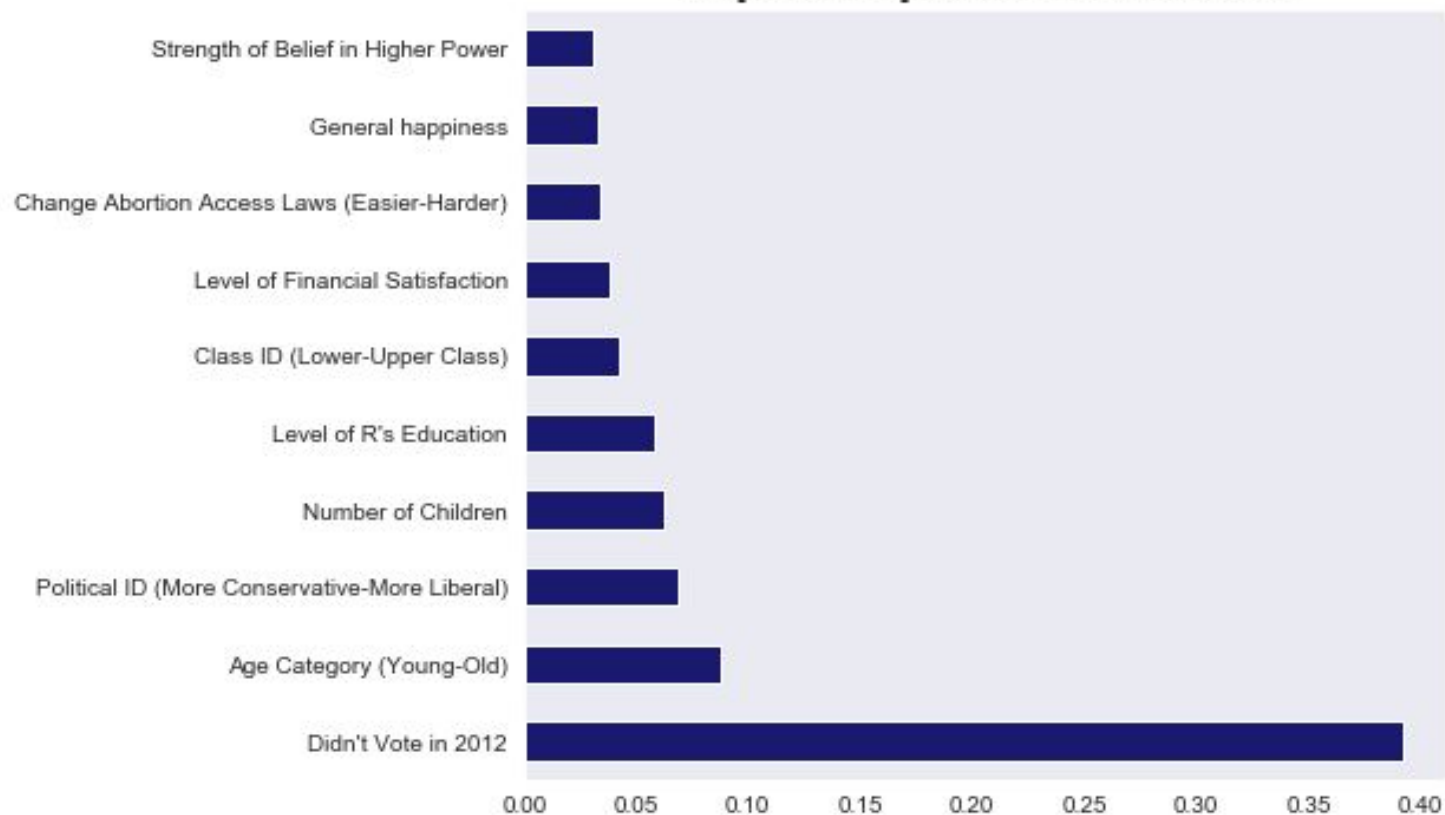
Boosting Voter Turnout with Results from the General Social Survey

Avonlea Fisher
October 2020

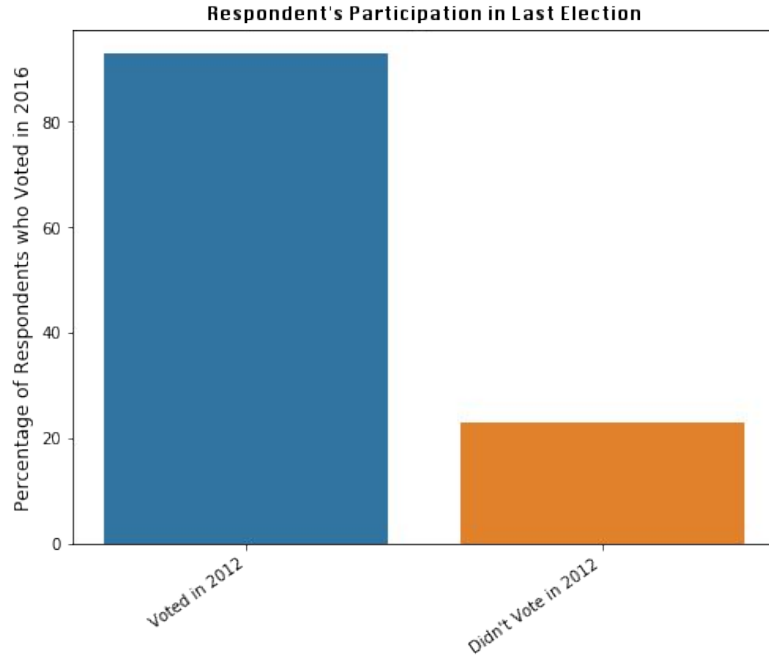
The Data

- 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

Top 10 important features



Percentage of 2016 Voters by Predictor

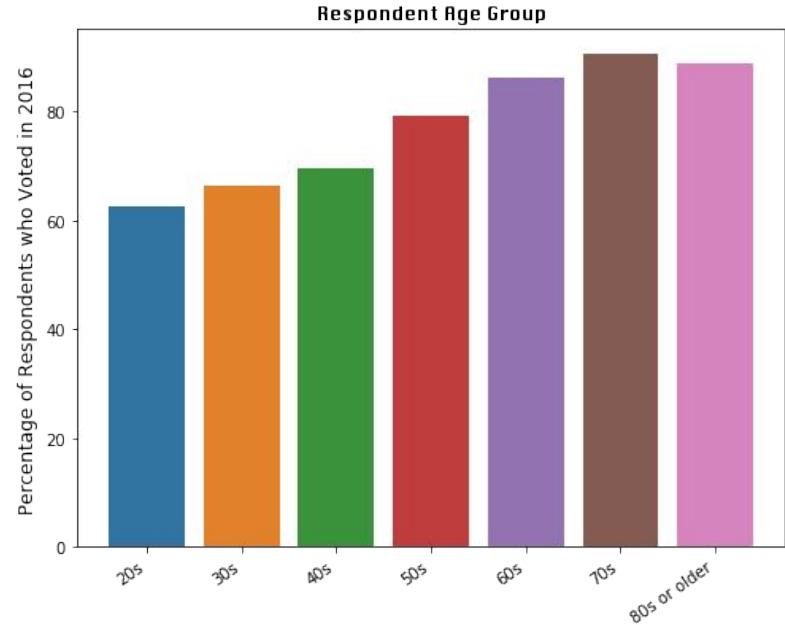


- Just over 20% of those who didn't vote in 2012 voted in 2016.
- Recommendation: engage voters who didn't vote in 2016 and/or other recent elections.

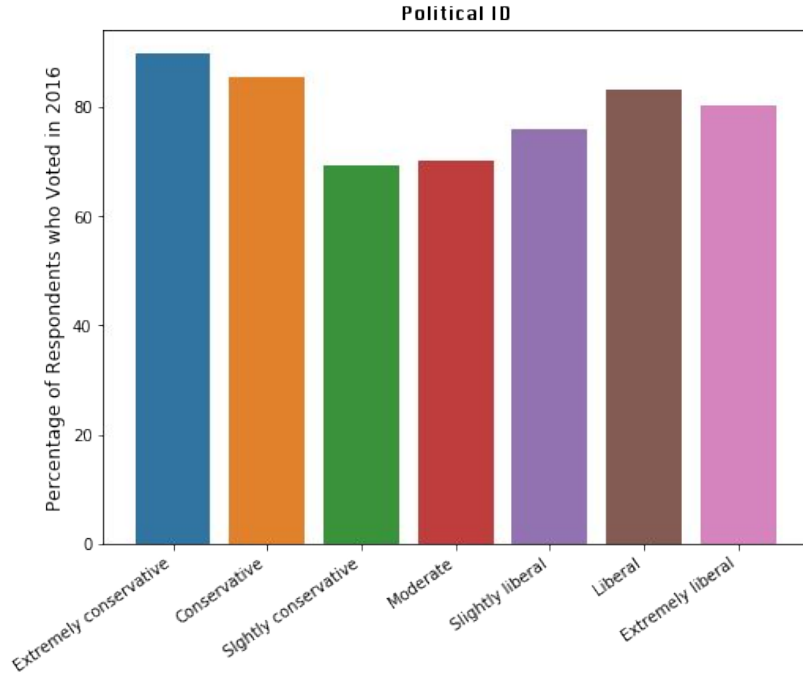


Percentage of 2016 Voters by Predictor

- The percentages of voters in their 20s and 30s were the lowest.
- Recommendation: focus on reaching younger voters.



Percentage of 2016 Voters by Predictor

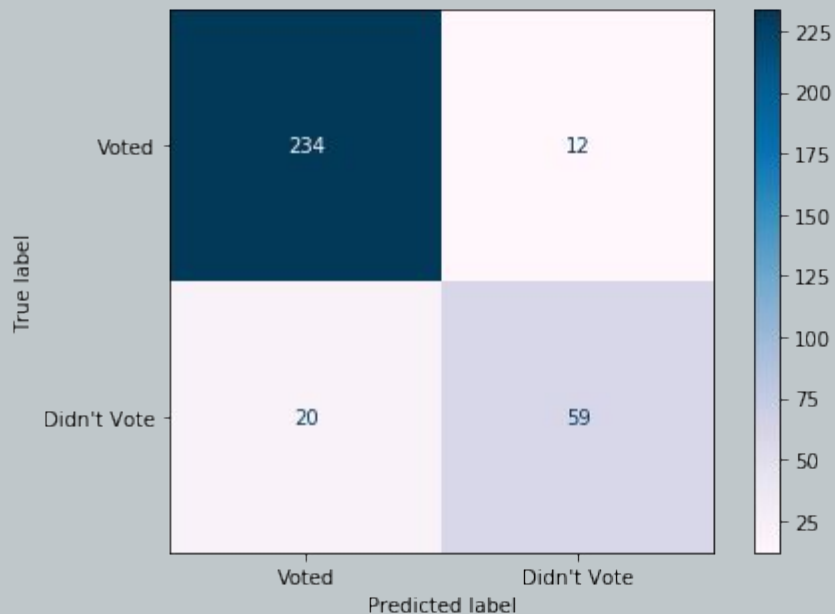


- Moderates and slightly liberal or conservative voters had the lowest turnout.
- Recommendation: engage voters who don't place themselves on the more extreme ends of this spectrum.



Best Classification Model: Random Forest

Random Forest Test Confusion Matrix

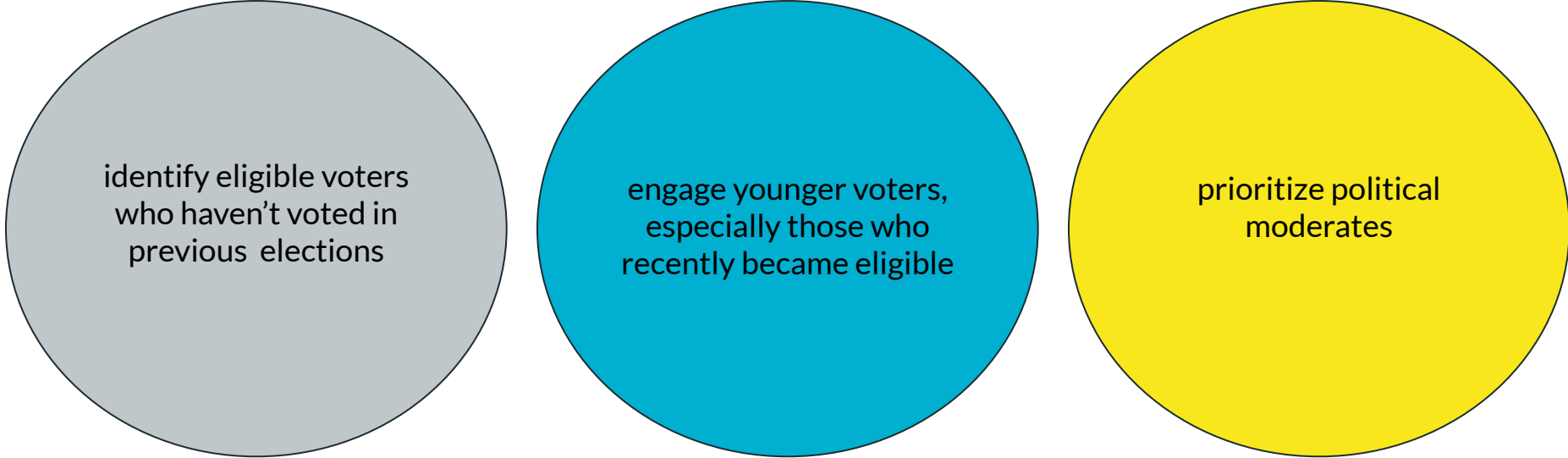


Test Data Accuracy: 90%

**Percentage of correctly
classified non-voters: 74%**

Recommendations Recap

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



identify eligible voters
who haven't voted in
previous elections

engage younger voters,
especially those who
recently became eligible

prioritize political
moderates

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?

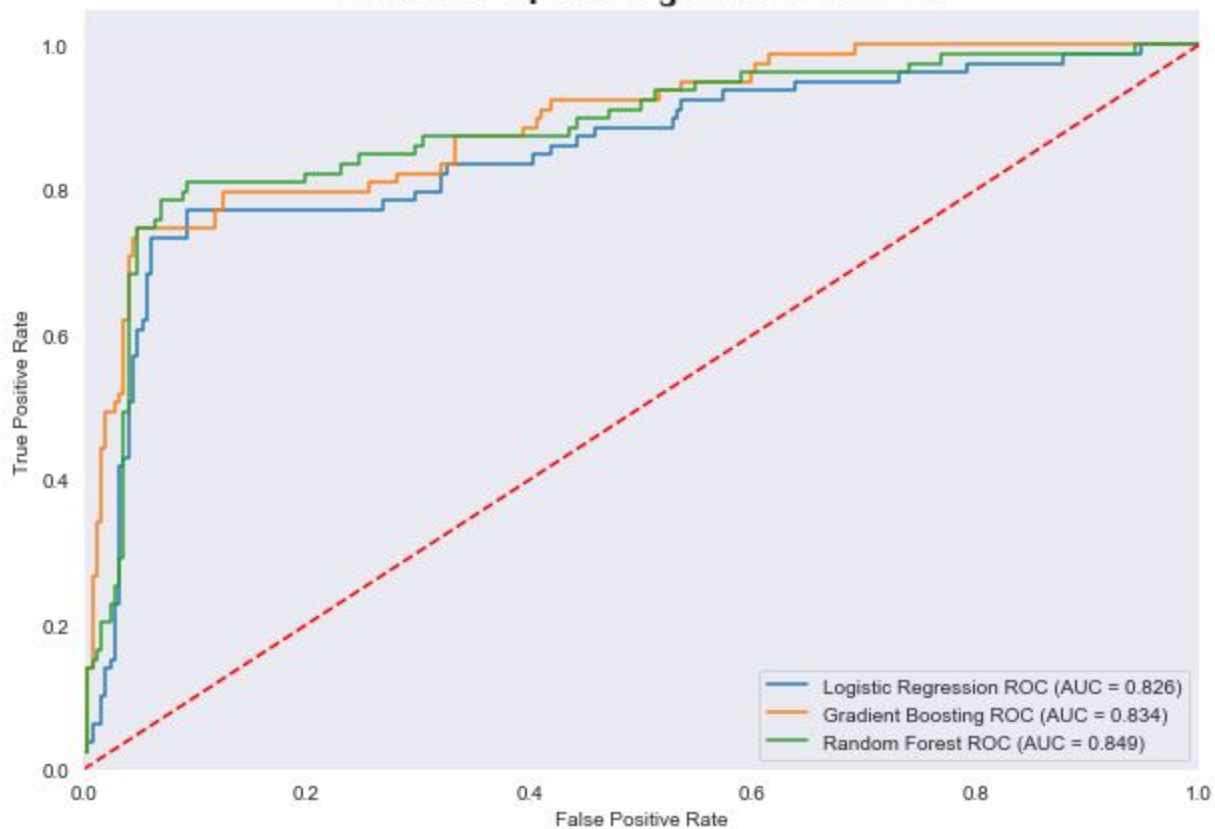


Thank you!

Appendix



Receiver Operating Characteristic



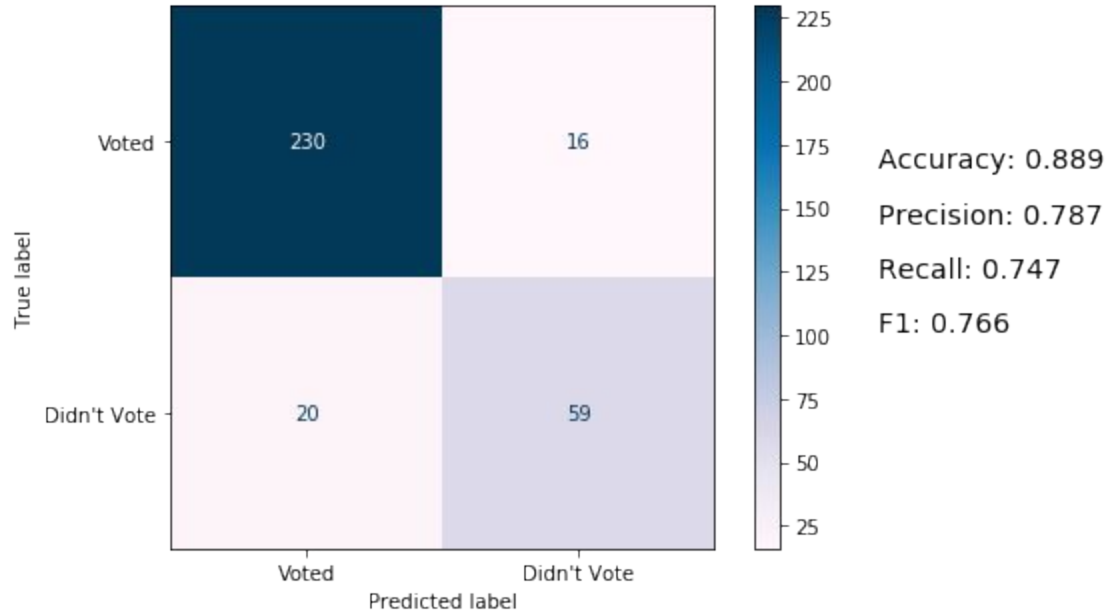
Recall Scores

Logreg: 0.734

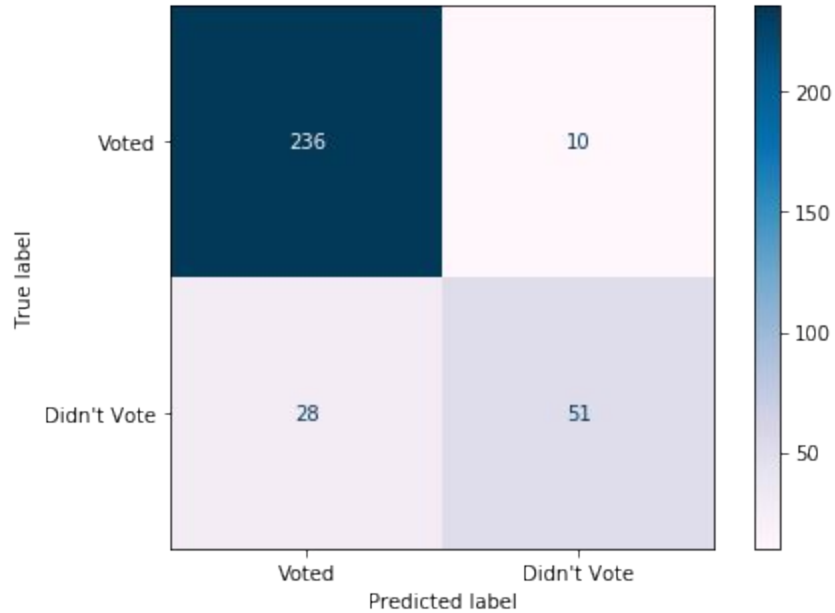
Gradient Boosting: 0.709

Random Forest: 0.747

Logistic Regression Test Confusion Matrix



Gradient Boosting Test Confusion Matrix



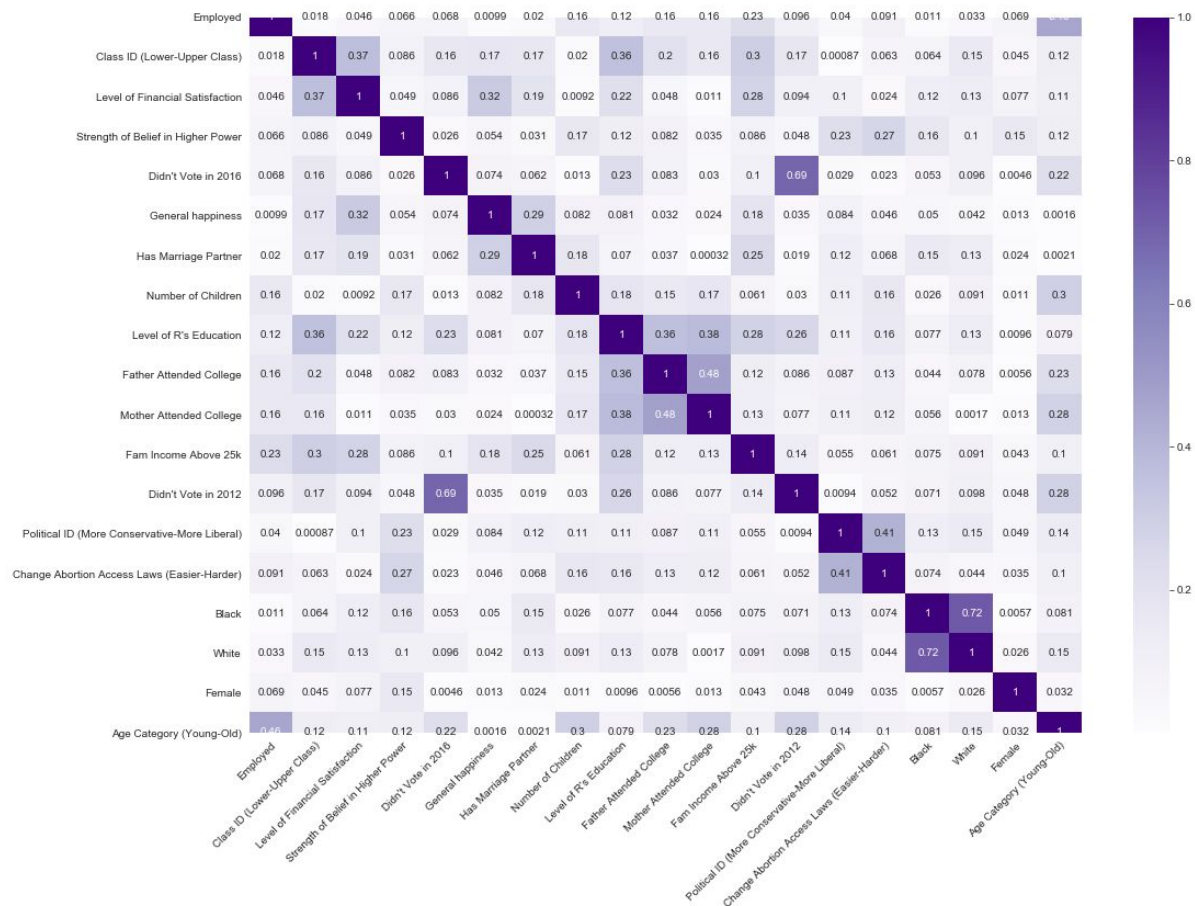
Accuracy: 0.883

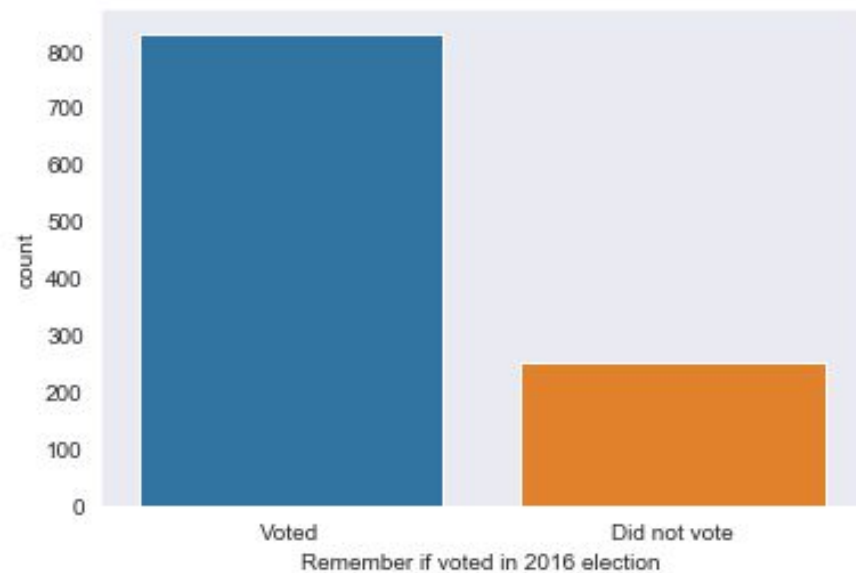
Precision: 0.836

Recall: 0.646

F1: 0.729

Variable Correlations





Credits

- Presentation theme from [PoweredTemplate.com](https://www.PoweredTemplate.com)
- Mail-in ballot image by Tiffany Tertipes on [Unsplash](https://www.Unsplash.com)

