Title of Your Report
Your subtitle

Junming Zhang, Hairong Sun, Xiaoxi Bai, Yangyang Liu

Sunday, November 01, 2020

Abstract

First sentence. Second sentence. Third Sentence. Fourth Sentence.

key words: USA 2020 election

Please click **"here"** to access the GitHub repository for all work.

Title of your Report

Name(s) of Author(s)

Date

Model

This is my first line. This is my second line. This is my third line.

Model Specifics

```
##      sex              age_group              race              hispan
##  Length:3467       Length:3467        Length:3467        Length:3467
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##   education            state            vote_trump
##  Length:3467       Length:3467        Min.   :0.0000
##  Class :character   Class :character   1st Qu.:0.0000
##  Mode  :character   Mode  :character   Median :1.0000
##                                        Mean   :0.5097
##                                        3rd Qu.:1.0000
##                                        Max.   :1.0000

##      sex          age_group                                   race
##  female:1591    >= 70  :381    american indian or alaska native:  24
##  male  :1876    18 ~ 20: 23    asian or pacific                : 136
```

```
##              20 ~ 30:328     black                    : 366
##              30 ~ 40:697     other                    : 174
##              40 ~ 50:701     white                    :2767
##              50 ~ 60:590
##              60 ~ 70:747
##           hispan                    education          state
##  cuban       :  19    at most high school   : 584   CA    : 376
##  mexican     : 238    bachelor              :1620   NY    : 314
##  not hispanic:3059    graduate              : 757   FL    : 292
##  other       : 148    tertiary (not bachelor): 506  TX    : 226
##  puerto rican:   3                                  IL    : 157
##                                                     OH    : 155
##                                                     (Other):1947
##     vote_trump
##  Min.   :0.0000
##  1st Qu.:0.0000
##  Median :1.0000
##  Mean   :0.5097
##  3rd Qu.:1.0000
##  Max.   :1.0000
##

## # A tibble: 1 x 1
##   prop_vote_trump
##             <dbl>
## 1           0.510

## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: vote_trump ~ sex + age_group + race + hispan + education + (1 |
##     state)
##    Data: survey_set
##
##      AIC      BIC   logLik deviance df.resid
##   4335.9   4458.9  -2147.9   4295.9     3447
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.2688 -1.0089  0.5580  0.8129  4.6751
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  state  (Intercept) 0.0565   0.2377
## Number of obs: 3467, groups:  state, 51
##
## Fixed effects:
##                                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)                      2.10152    0.73666   2.853 0.004334 **
## sexmale                          0.41511    0.07578   5.478 4.30e-08 ***
## age_group18 ~ 20                -1.62472    0.66708  -2.436 0.014868 *
## age_group20 ~ 30                -0.23053    0.17010  -1.355 0.175340
## age_group30 ~ 40                -0.01900    0.13696  -0.139 0.889689
## age_group40 ~ 50                 0.20558    0.13742   1.496 0.134634
## age_group50 ~ 60                 0.15819    0.14086   1.123 0.261423
```

```
## age_group60 ~ 70                   0.02292    0.13392    0.171 0.864082
## raceasian or pacific             -1.50903    0.49078   -3.075 0.002107 **
## raceblack                        -3.21613    0.49016   -6.561 5.33e-11 ***
## raceother                        -1.00184    0.48213   -2.078 0.037715 *
## racewhite                        -0.48299    0.45368   -1.065 0.287054
## hispanmexican                    -1.74771    0.58520   -2.987 0.002822 **
## hispannot hispanic               -1.11824    0.56529   -1.978 0.047908 *
## hispanother                      -1.30600    0.59093   -2.210 0.027100 *
## hispanpuerto rican               -1.54304    1.45366   -1.061 0.288470
## educationbachelor                -0.49200    0.10938   -4.498 6.86e-06 ***
## educationgraduate                -0.46929    0.12589   -3.728 0.000193 ***
## educationtertiary (not bachelor) -0.45434    0.13576   -3.347 0.000818 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation matrix not shown by default, as p = 19 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it

## convergence code: 0
## Model failed to converge with max|grad| = 0.00781382 (tol = 0.002, component 1)
```

ROC curve
AUC = 0.6451397

PR curve
AUC = 0.6063118
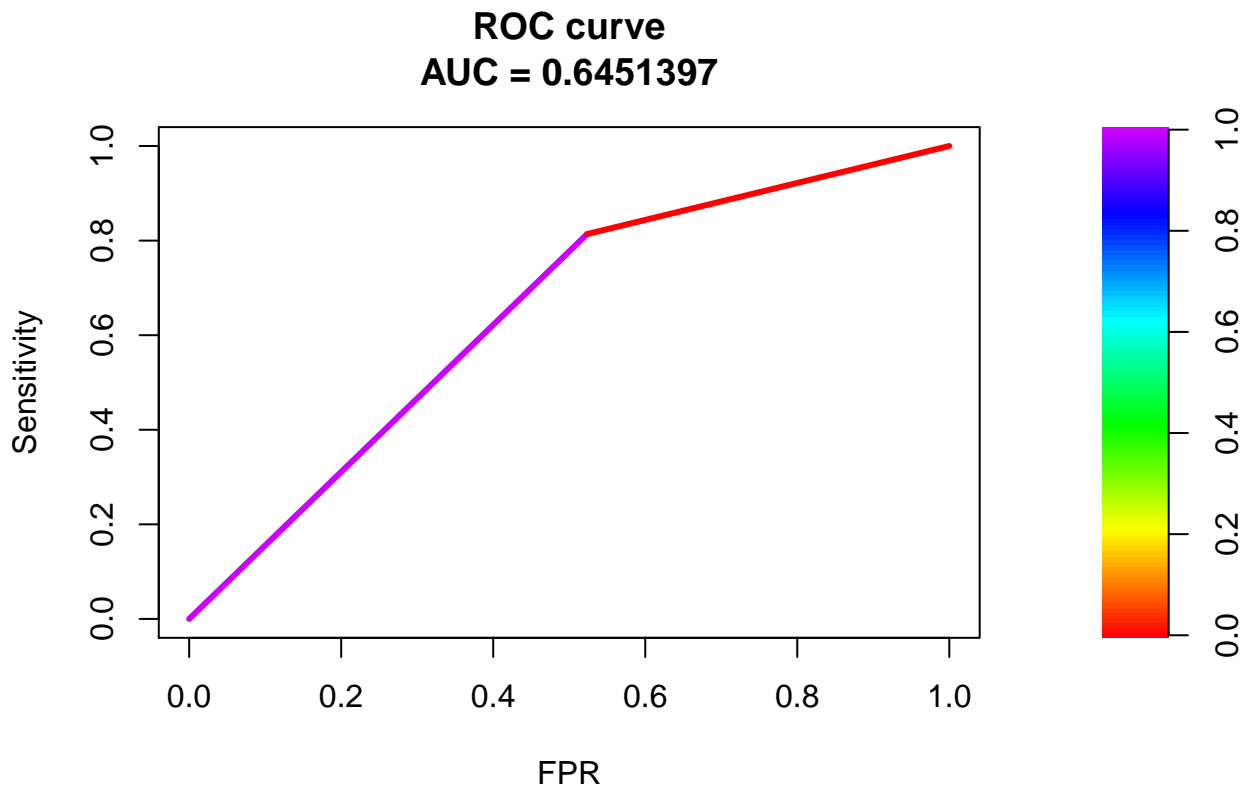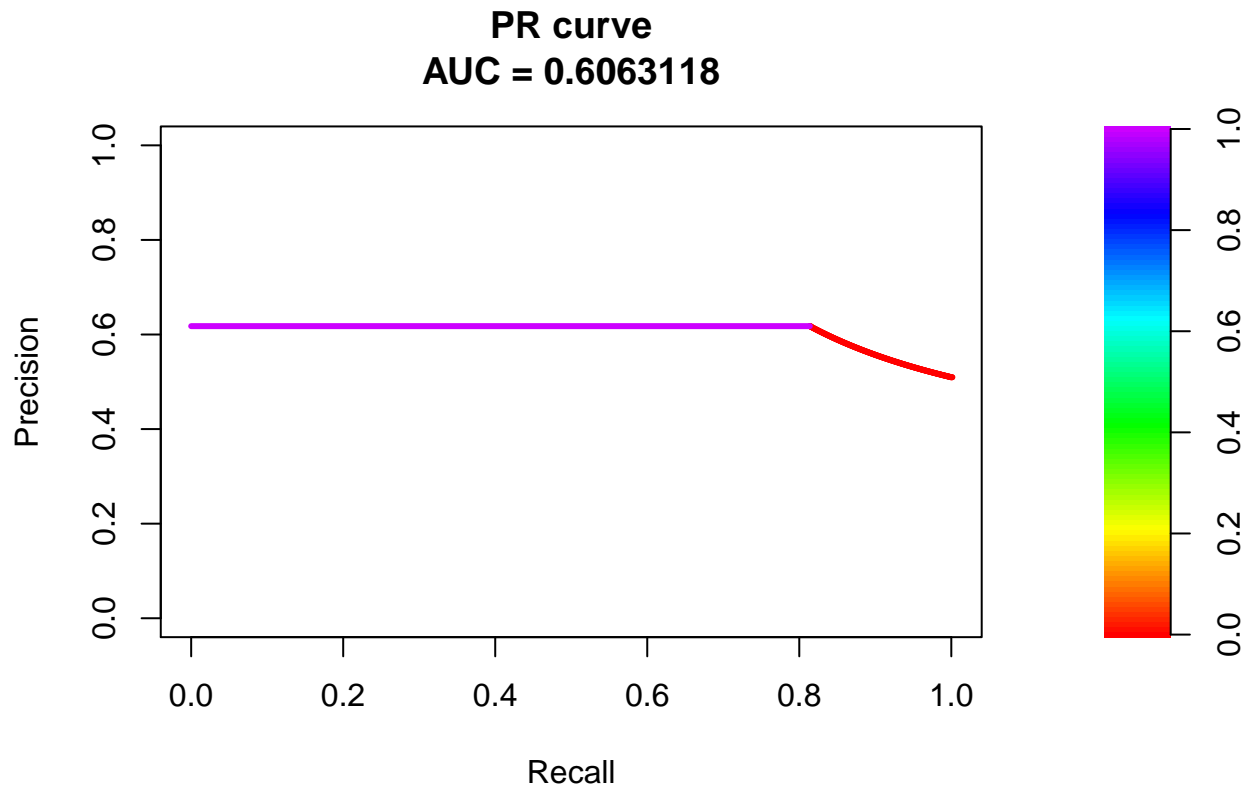

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0     1
##          0  810   329
##          1  890  1438
##
##                Accuracy : 0.6484
##                  95% CI : (0.6322, 0.6643)
##     No Information Rate : 0.5097
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.2921
##
##  Mcnemar's Test P-Value : < 2.2e-16
##
##             Sensitivity : 0.4765
##             Specificity : 0.8138
##          Pos Pred Value : 0.7112
##          Neg Pred Value : 0.6177
##              Prevalence : 0.4903
##          Detection Rate : 0.2336
##    Detection Prevalence : 0.3285
##       Balanced Accuracy : 0.6451
##
##        'Positive' Class : 0
##
```

Post-Stratification

```
##      sex          age_group                                            race
##  female:23559   >= 70  :8211   american indian or alaska native:  117
##  male  :19888   18 ~ 20: 596   asian or pacific                :13784
##                 20 ~ 30:3203   black                           : 3706
##                 30 ~ 40:5419   other                           : 5017
##                 40 ~ 50:8260   white                           :20823
##                 50 ~ 60:9541
##                 60 ~ 70:8217
##           hispan                          education           state
##  cuban        : 1474   at most high school    :21528   CA     :11347
##  mexican      : 6264   bachelor               :11340   NY     : 4836
##  not hispanic:30272    graduate               : 5746   FL     : 4710
##  other        : 5277   tertiary (not bachelor): 4833   TX     : 3600
##  puerto rican:  160                                     NJ     : 2153
##                                                         IL     : 1636
##                                                         (Other):15165
##      perwt
##  Min.   :   5.23
##  1st Qu.: 308.57
##  Median : 444.55
##  Mean   : 568.17
##  3rd Qu.: 679.90
##  Max.   :9607.51
##
## # A tibble: 51 x 2
##     state trump_predict
##     <fct>         <dbl>
##  1 AK            0.468
##  2 AL            0.500
##  3 AR            0.560
##  4 AZ            0.484
##  5 CA            0.402
##  6 CO            0.487
##  7 CT            0.381
##  8 DC            0.370
##  9 DE            0.379
## 10 FL            0.545
## # ... with 41 more rows
```

Results

```
## # A tibble: 2 x 2
##   elected       total_electoral_votes
##   <chr>                         <dbl>
## 1 Donald Trump                    116
## 2 Joe Biden                       422
```

Discussion

Weaknesses

Next Steps

References

1. Tausanovitch, Chris and Lynn Vavreck. 2020. Democracy Fund + UCLA Nationscape, October 10-17, 2019 (version 20200131). Retrieved from [URL].
2. Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Matthew Sobek. IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D010.V10.0
3. PUMS USA, University of Minnesota, www.ipums.org.
4. Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686
5. Stefan Milton Bache and Hadley Wickham (2014). magrittr: A Forward-Pipe Operator for R. R package version 1.5. https://CRAN.R-project.org/package=magrittr
6. Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.
7. Jens Keilwagen, Ivo Grosse and Jan Grau (2014). Area under Precision-Recall Curves for Weighted and Unweighted Data. PLOS ONE (9) 3.
8. Jan Grau, Ivo Grosse, and Jens Keilwagen (2015). PRROC: computing and visualizing precision-recall and receiver operating characteristic curves in R. Bioinformatics (31) 15, pp. 2595-2597.R package version 1.3.1.
9. Max Kuhn (2020). caret: Classification and Regression Training. R package version 6.0-86. https://CRAN.R-project.org/package=caret
10. United States Electoral College Votes by State. (n.d.). Retrieved November 01, 2020, from https://www.britannica.com/topic/United-States-Electoral-College-Votes-by-State-1787124