Model Release Notes

Model: Biden Infrastructure Support
Type: Survey-Response
Geography: United States
Date: July 26, 2021

Model Description

An ensemble method classifier model was created to target and identify people likely to support President Biden's plan to upgrade infrastructure, including transportation, energy, housing, schools, manufacturing, internet, and community health care. The model was constructed using results from an online survey conducted between January 2021 and June 2021. The model scores here are expressed on a 0-100 scale and represent the probability that a person supports President Biden's infrastructure plan. The model was used to score over 261 million voting-age individuals nationwide.

Process Overview

The model was trained on survey respondents who were divided into two groups based on a survey question:

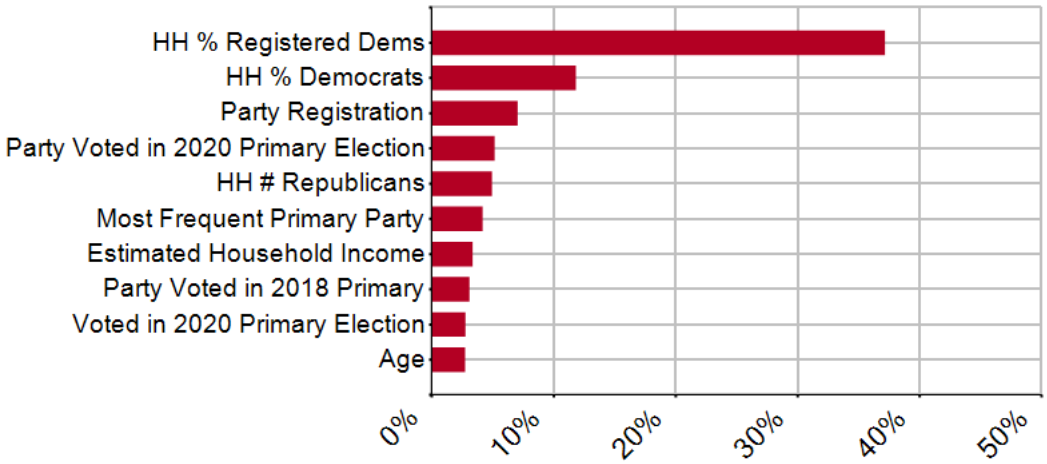
(1) **Targets**: Respondents who support President Biden's infrastructure plan.
(2) **Non-Targets**: Respondents who oppose President Biden's infrastructure plan.

Several sub-models were built using a rules-based classifier on a variety of consumer, political, credit, and demographic variables. In total, over 1,000 variables were considered during model creation.

The model building process determined the most appropriate variables in identifying the differences between target respondents and non-target respondents. The final score was generated using a boosted decision tree classifier. To validate the model, a randomly selected group of survey records were held out from the model building process. These holdout records were then scored with the model and analyzed for accuracy.

Key Variables

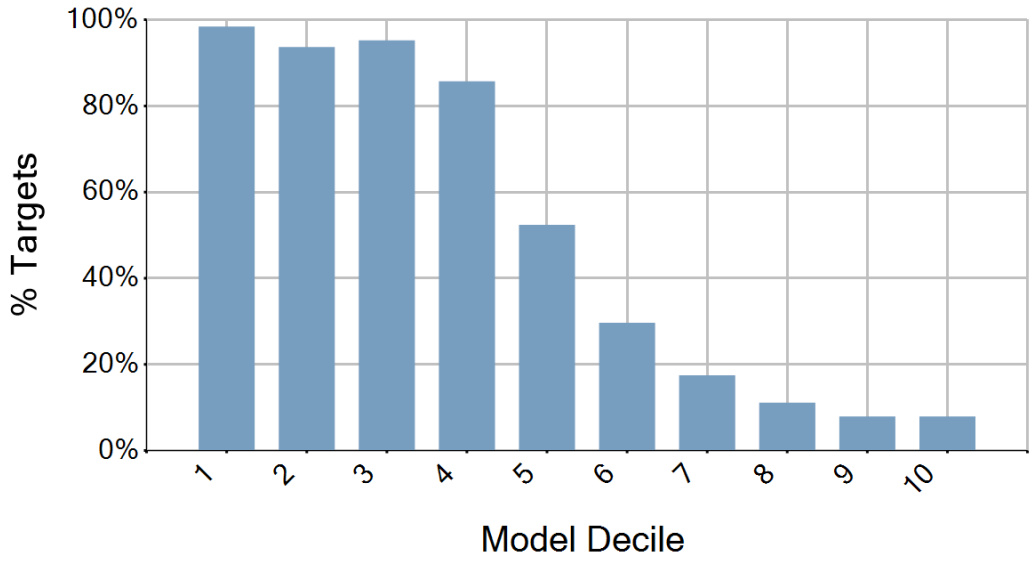
The key variables and relative weights used in the model include:



Validation

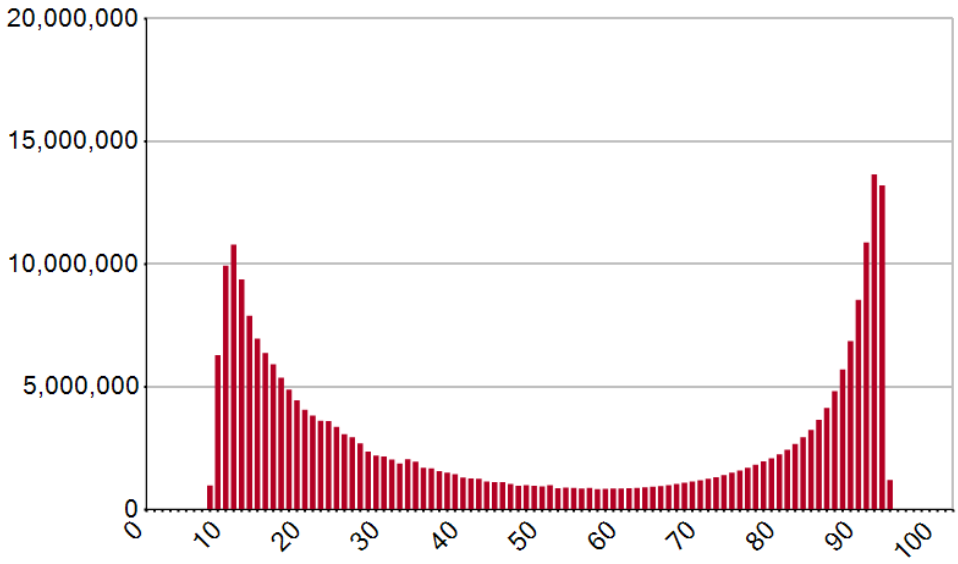
The model was validated by scoring a holdout sample of survey respondents. The holdout records were then ranked by model score, separated into ten deciles, and evaluated. The highest scored records are found in decile 1, while the lowest scored records are found in decile 10.

Successful rank-ordering occurs when a higher percentage of target records are correctly scored higher than non-target records. We expect a valid model to show a stepwise decline from decile 1 through decile 10. Ideal or perfect models show a steep downward slope, with values near 100% in decile 1 and near 0% in decile 10.



Score Distribution

The following chart shows the distribution of model scores for voting-age individuals nationwide. The scores range from 0-100, with higher scores indicating a higher likelihood of supporting President Biden's infrastructure plan.



To Target Top...	Pull Records Above...
10%	85
20%	76
30%	67
40%	59
50%	50
60%	41
70%	33
80%	24
90%	15
100%	7