Model Card

Alaina Hu

April 3, 2024

1 Model Details. Basic information about the model

- Person or organization developing model: Alaina Hu from Hu Private Company
- Model date: March 31, 2024
- Model version: 1
- Model type: Post-stratification prediction model
- Training Algorithms, Parameters, Fairness Constraints: Fairness constraints applied to ensure equitable performance among diverse demographic groups.
- Features: Includes demographic factors (age, gender, race), voting history.
- Paper or Other Resource: Detailed methodology and model insights can be found in our paper, "Analyzing and Forecasting 2020 Voter Behavior"
- Citation Details: HuPrivateCompany Political Analysis Team (2024). "Analyzing and Forecasting 2020 Voter Behavior". HuPrivateCompany Publications.
- License: Distributed under the HuPrivateCompany Data Analysis License, intended for non-commercial, academic use.
- Questions or Comments: For inquiries, contact the Data Analytics Team.

2 Intended Use

- Primary Intended Uses: Analysis and forecast of voter behavior in the 2020 US Presidential Election; informing policy and campaign strategy development
- Primary Intended Users: Political scientists, policy analysts, campaign strategists, and general public.
- Out-of-scope Use Cases: Not intended for individual voter identification, surveillance, or any discriminatory applications.

3 Factors

- Relevant Factors: The model accounts for demographic diversity, gender, age, and responses to political issues.
- Evaluation Factors: Performance evaluated across demographic groups, with additional sensitivity analyses for political affiliation.

4 Metrics

- Model Performance Measures: Accuracy
- Decision Thresholds: Thresholds set to maximize F1 score while maintaining demographic parity.
- Variation Approaches: Performance variation analyzed by demographics

5 Evaluation Data

- Datasets: 2020 US Cooperative Election Study, supplemented by anonymized US voter file for post-stratification.
- Motivation: The CES provides comprehensive insights into voter behavior and demographics, while the voter file enhances model accuracy and applicability.

- Preprocessing: Data anonymization, imputation for missing values, and weighting to correct for sampling biases.

6 Training Data

- Details: Mirrors Evaluation Data in terms of content and preprocessing. The distribution over various factors is balanced to reflect the US voting population as of 2020. # Quantitative Analyses
- Unitary Results: Achieves 82% accuracy overall, with F1 scores above 0.75 across all major demographic groups.

- Intersectional Results: Analysis reveals consistent performance across intersectional demographic groups, with specific attention to historically underrepresented and marginalized populations.

7 Ethical Considerations

The model is designed with privacy and fairness at the forefront. Users are cautioned against using the model in ways that could infringe on individual privacy or reinforce societal biases.

8 Caveats and Recommendations

- Caveats: The model's accuracy is contingent on the representativeness of the 2020 election data; shifts in voter behavior or demographics may affect performance.

- Recommendations: Reevaluations with new data are advised to maintain model relevance and accuracy. Ethical use guidelines should be strictly followed to prevent misuse.