Creating Model Card for US Voter Prediction Model

Hailey Jang

April 1, 2024

The US Voter Prediction Model (UVPM) represents a cutting-edge machine learning approach designed to forecast how individuals in the United States might vote, alongside their demographic profiles. This tool is aimed at equipping researchers, policymakers, and data analysts with deeper insights into voting patterns, future election results, and broader socio-political trends. It's important to note that the UVPM is crafted solely for research and analytical endeavors, and its application for personal decision-making or interventions is not advised

Model Details:

- Model Identifier: US Voter Prediction Model (UVPM)
- Type of Model: Supervised Learning, aimed at Classification tasks
- Version Number: 1.0
- Custodian of the Model: Gavin Crooks, associated with the University of Toronto
- Creation Date: 1 April 2024
- Date of Latest Update: 1 April 2024

Intended Use:

The US Voter Prediction Model (UVPM) leverages machine learning to estimate the voting preferences and demographic specifics of American voters. It seeks to arm analysts, policymakers, and scholars with the ability to dissect electoral trends, predict electoral outcomes, and explore the nuances of socio-political dynamics in the U.S. The application of the UVPM is strictly for scholarly and analytical reasons, and it should not be used for decision-making at the individual level.

Metrics:

To measure the effectiveness of the UVPM, several established metrics are employed, including:

- Accuracy: This metric calculates the percentage of predictions the model gets right.

- Precision: This measures the accuracy of positive predictions made by the model.
- Recall: This assesses how many actual positives the model can identify.
- F1 Score: A balanced metric that combines precision and recall.
- Area Under the Receiver Operating Characteristic Curve (AUC-ROC): This metric evaluates the model's ability to distinguish between classes under various thresholds.

Training Data:

The development of the UVPM was based on a rich dataset featuring demographic details, voting records, and socio-economic factors of U.S. voters. The primary data source was the 2020 US Cooperative Election Study, complemented by detailed voter files from a private entity, ensuring a comprehensive understanding of the electorate. The dataset was thoroughly cleaned and preprocessed to meet high-quality standards.

Ethical Considerations:

- Privacy: The model adheres to privacy norms by only dealing with data that has been de-identified and aggregated, ensuring no personal information is involved.
- Fairness: The model strives for unbiased outcomes by incorporating diverse data sources and continuously evaluating for and addressing biases.
- Transparency: Full disclosure of the model's workings, including data sources and performance, is provided to ensure clarity and accountability.

Caveats and Recommendations:

- Interpretation: Users should approach the model's outputs with caution, integrating them with broader contextual knowledge.
- Validation: Ongoing validation and sensitivity checks are recommended to ensure the model remains accurate and fair.
- Application Restrictions: The UVPM is not designed for personal decision-making but rather for generating broader electoral insights. Its probabilistic nature should be carefully considered to avoid misinterpretation.