Model Card for 2020 Voter Behavior Analysis Model

Model Details

- **Developers**: National Center for Political Studies Analytics Team
- **Model Version**: 1.0
- **Release Date**: 2024-04-01
- **Model Type**: Supervised Learning - Classification
- **Primary Task**: Predicting voter turnout and party preference based on demographic data and previous voting history.

Intended Use

- **Primary Users**: Academic researchers, political scientists, and policy analysts.
- **Use Cases**: To study factors influencing voter turnout and party preference in the 2020 US elections, identify patterns in political engagement, and contribute to strategies aimed at increasing democratic participation.
- **Out-of-Scope**: Not intended for use in predicting individual voting behavior for targeted advertising, campaign outreach, or any commercial purposes.

Metrics

- **Performance Measures**: Accuracy, Precision, Recall, and F1 Score for classification tasks. Metrics were chosen to balance the importance of correctly predicting voter turnout and preferences against the costs of false predictions.
- **Validation Methods**: 10-fold cross-validation on the training dataset and an independent validation set derived from a separate portion of the CES data not used in training.

Training Data

- **Dataset**: 2020 US CES Voter File Dataset
- **Size**: 200,000 voter profiles
- **Preprocessing**: Missing values imputed where possible, records with missing critical information were excluded. Features engineered from raw data include age groups, geographic regions, and previous election participation.

Ethical Considerations

- **Fairness and Bias**: Models were tested for bias across demographic groups (age, race, gender, geographic location). Adjustments and retraining were performed to mitigate any discovered biases.
- **Privacy**: Training data was anonymized to protect individual privacy, with no personally identifiable information used in the model training process.
- **Transparency**: This model card, along with a detailed datasheet for the dataset, is published to ensure transparency about the model's development and capabilities.

Caveats and Recommendations

- **Caveats**: The model's predictions are based on historical data and may not account for future changes in political trends or voter behavior. Predictive accuracy is higher for general trends than for individual predictions.
- **Recommendations for Users**: Use the model as a tool for understanding broader electoral trends rather than for making decisions about specific individuals. Always consider the context and potential for changes in political dynamics when interpreting model outputs.

Model Maintenance

- **Maintenance Plan**: The model will be reviewed and potentially updated following each major election cycle to incorporate new data and adjust for changing political landscapes.
- **License**: Distributed under the Academic Non-Commercial Use License, permitting non-commercial use, sharing, and adaptation with proper attribution.