The Oval Office Prediction

2024 U.S. Election Prediction

COGNIFY

AGENDA

01
Introduction
Overview of the project and its goals

02
Data Collection
Details about sources, cleaning, and preprocessing.

03
Exploratory Data Analysis
Key insights and trends observed in the data.

04
Modelling and Evaluation
Algorithms used and performance metrics

05
Deployment and Challenges
Different challenges faced and deployment.

06
Conclusion
Conclusion and Potential improvements.

MEET OUR TEAM

MODEL DEVELOPMENT

Abdullah Ifteqar
Sub-Team Lead

Subash Pavan
Data Scientist & PPT Expert

Amruth Raj
DevOps Engineer

Sangeeth
Full Stack Developer

Saurav Risal
Database Manager

Alen C Saji
Data Scientist & PPT Expert

DATA COLLECTION / EDA

Clifford Addison
Sub-Team Lead

Aparna J R
Documentation Expert

Yugahang Limbu
Data Analyst

Sumith Padma PV
Data Engineer

Nischal Pradhan
UI/UX Designer

Riya Shah
Data Analyst

Serene Sebastian
Data Analyst

MODEL DEVELOPMENT

Abdullah Ifteqar
Sub-Team Lead

ubash Pavan
ta Scientist & PPT Expert

Amruth Raj
DevOps Engineer

Sangeeth
Full Stack Developer

aurav Risal
Database Manager

Alen C Saji
Data Scientist & PPT Expert

01 Introduction

2024 U.S. Election

Global Impact:

- International relations, National security concerns, trade agreements, economic policies, economic policies, social dynamics, etc.

Kamala D. Harris

OBJECTIVE

- A machine learning model that can accurately predict the winning party based on census - demographics, socioeconomic indicators, and polling data.
- The insights generated can help stakeholders, including campaign strategists, analysts, and policymakers, make informed decisions.

Donald J. Trump

Approach

Data Collection & EDA

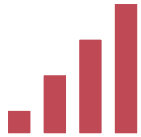
- Data Pre-processing
- Exploratory Data Analysis (EDA)
- Feature engineering & selection

Model Development & Evaluation

- Random Forest
- XGBoost
- Model Ensembling
- Performance Evaluation

Deployment

- Database setup
- Google Cloud Platform (GCP)



Dashboard

- Interactive
- Accessibility features

02 Data Collection

https://census.gov

Census data, voter registrations, age, employment status, etc were sourced from official government databases. United States Census Bureau.

https://fivethirtyeight.com

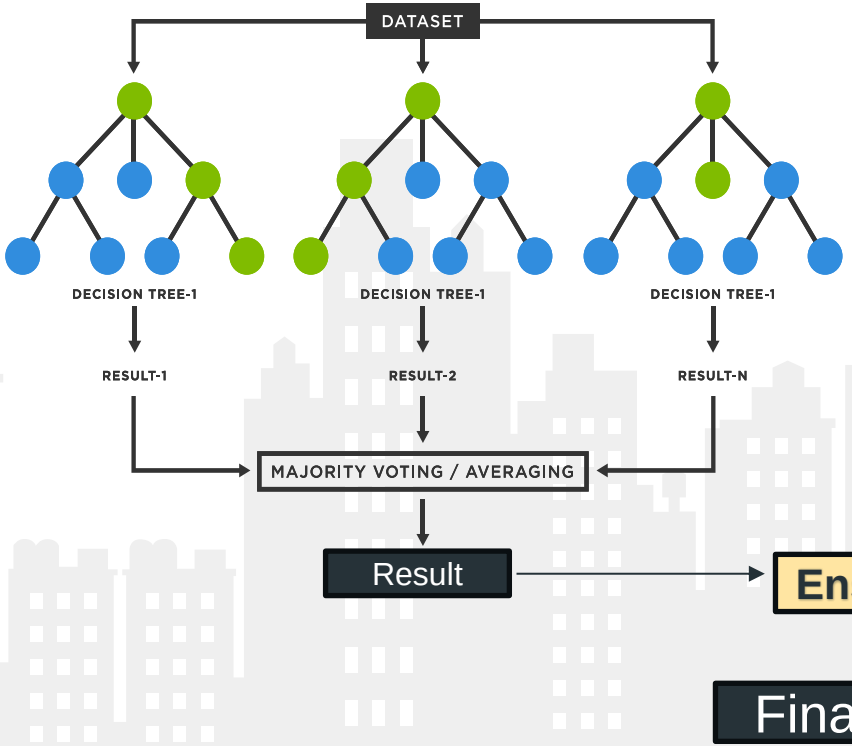
Election forecasts were sourced from 538, a well-known political forecasting site.

03 Exploratory Data Analysis

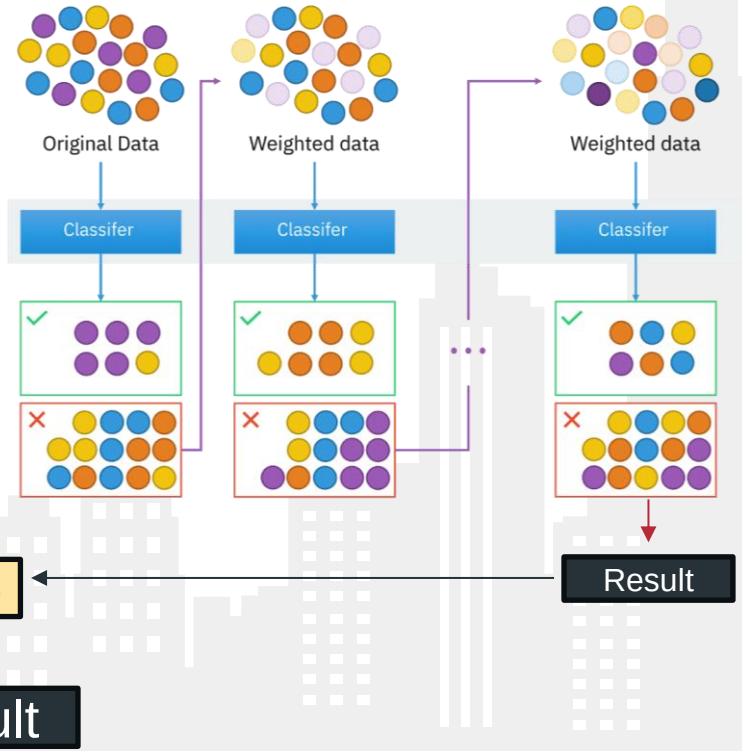
Demographic Impact	Higher education and diversity trends suggest shifts towards Democratic preferences in key states.
Economic Factors	Middle-income groups lean towards Democrats, Democrats, while high-income groups prefer prefer Republicans.
State Variations	California dominates with Democratic support, while Florida and Texas are key key battlegrounds.

Model Building

Random Forest Model



XGBoost Model

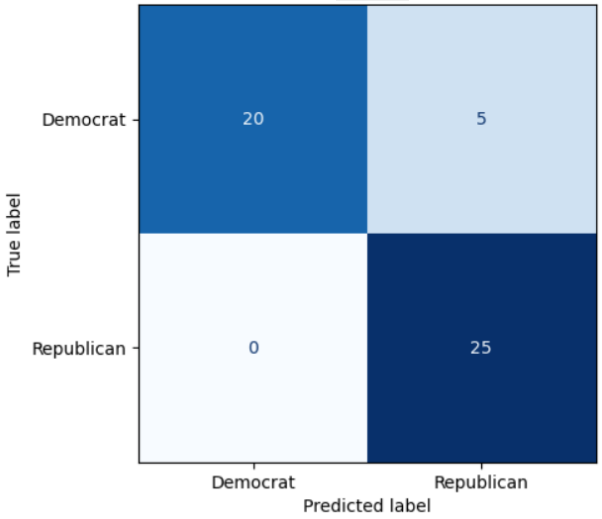


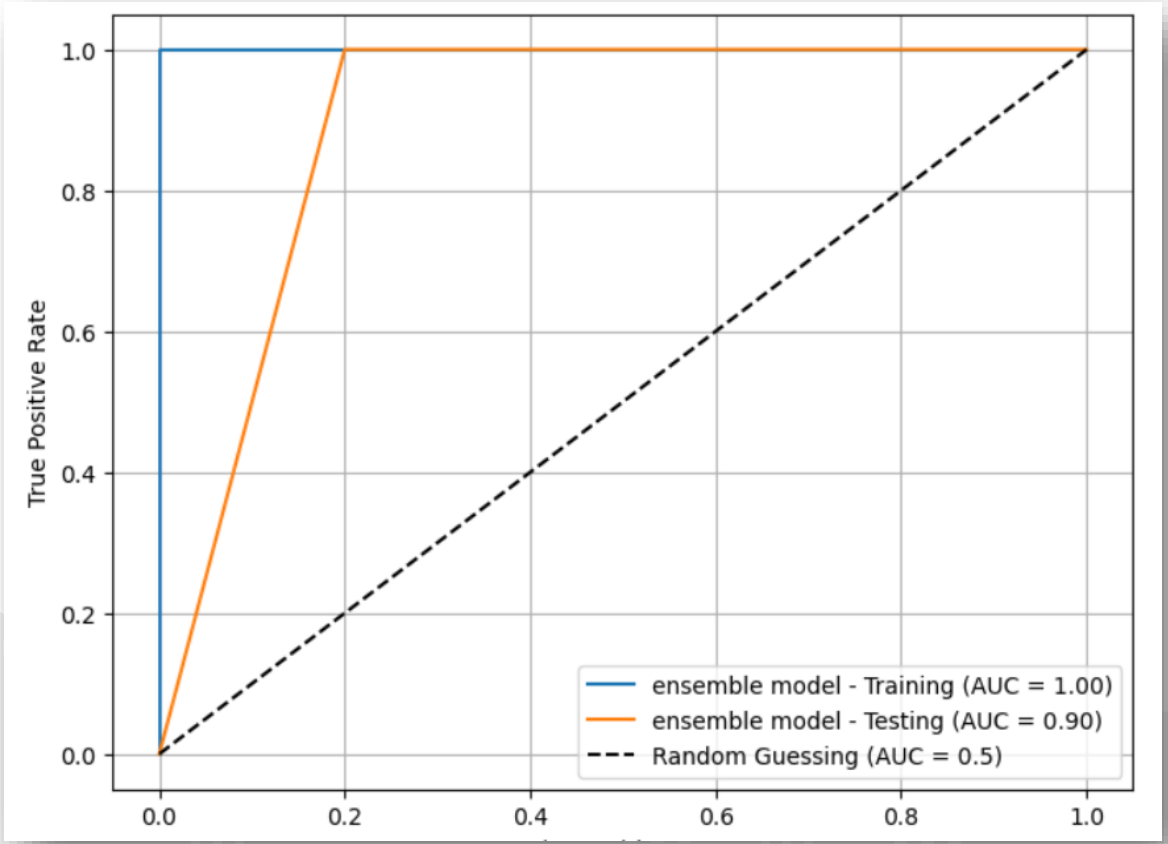
Performance Evaluation

Ensemble Model

Accuracy : **0.90**

F1- Score : **0.89**



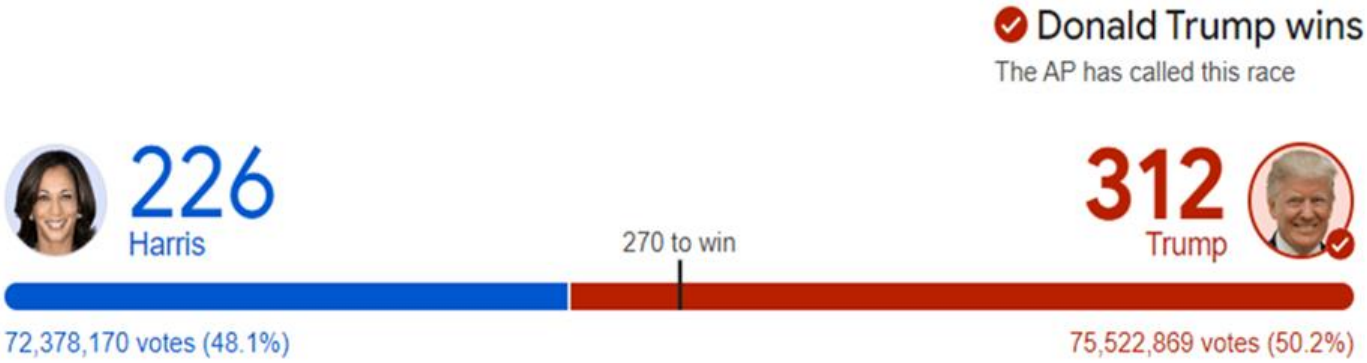


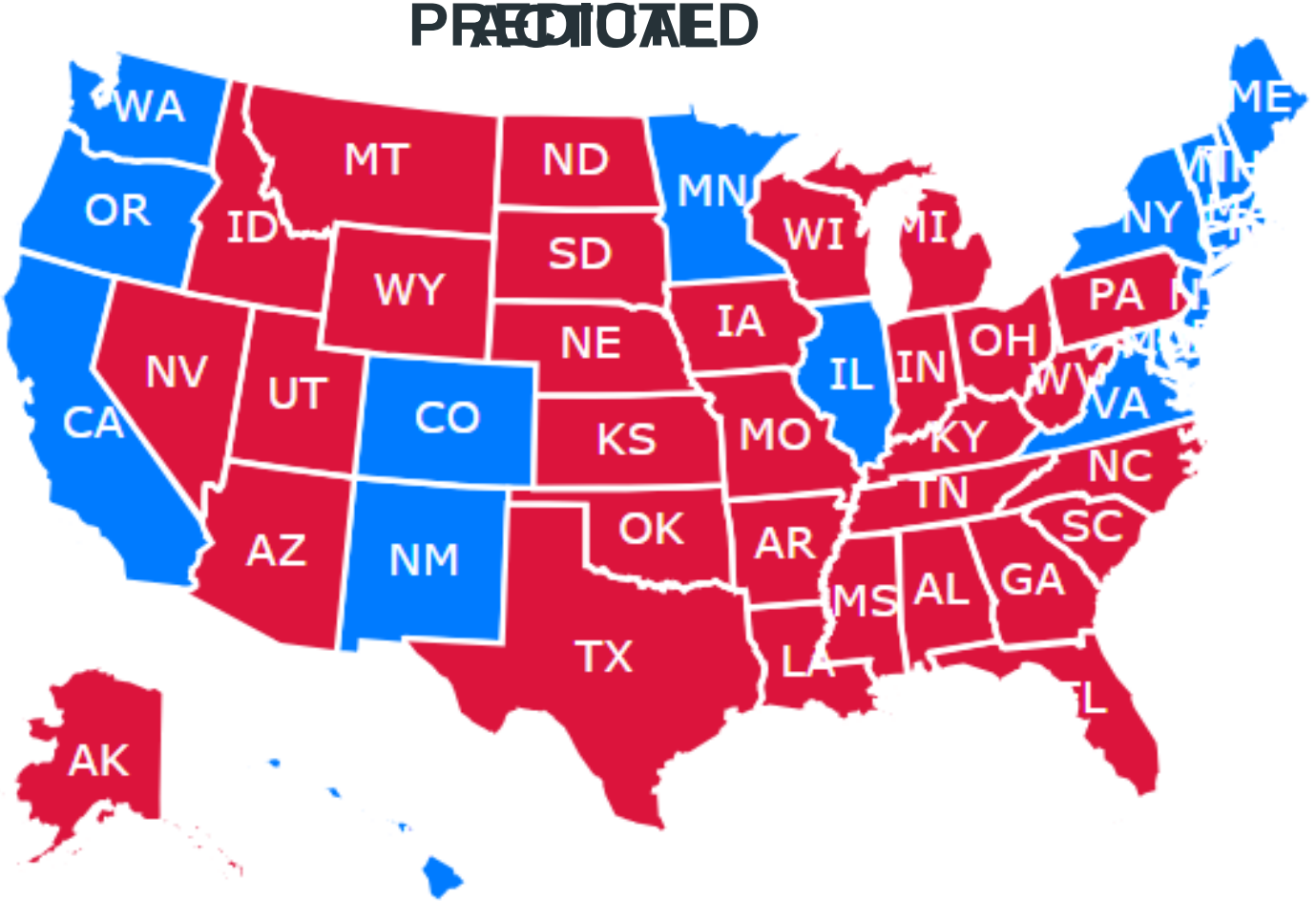
Predicted Result (November 3 - 96% accuracy)

	party	electoral_college_seats
0	Democrat	236
1	Republican	302

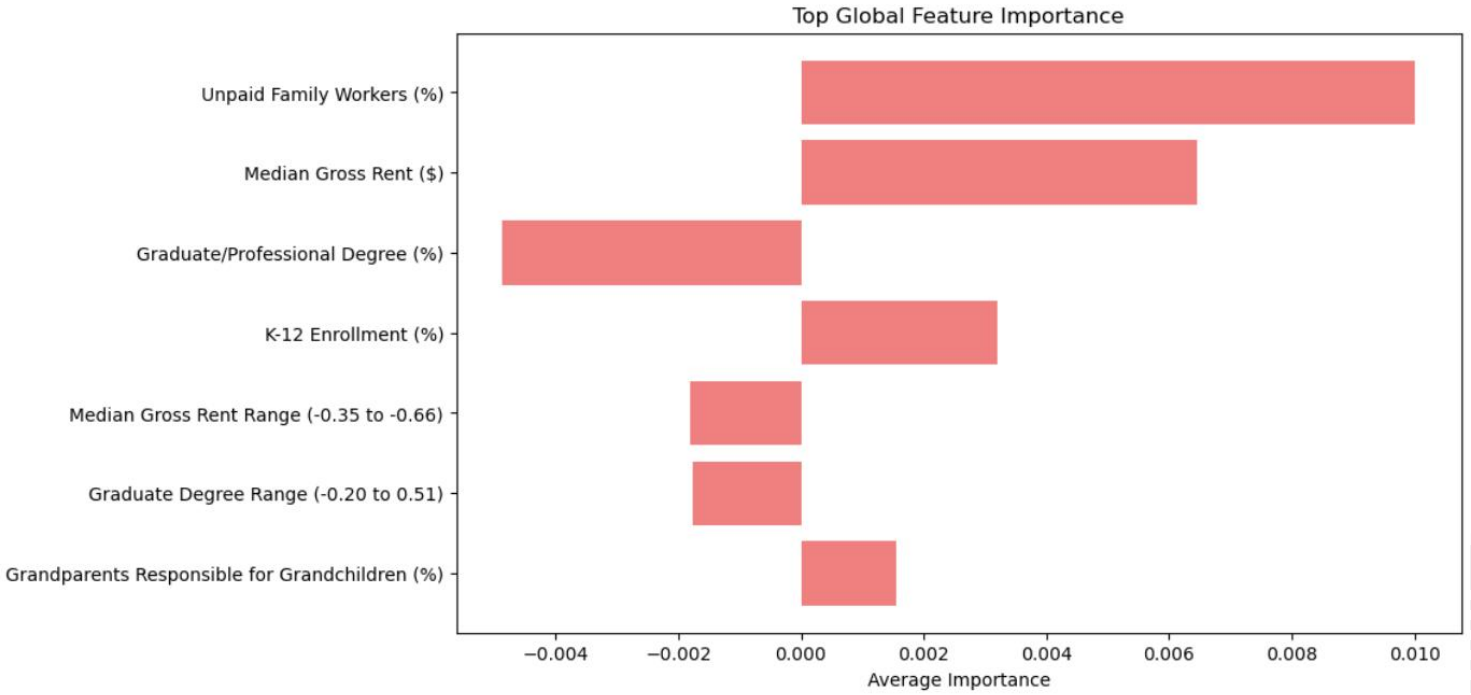
Party	Difference
Democrat	+10
Republican	-10

Actual Result (November 5)

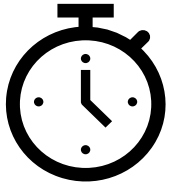



PREDICTED
WA
OR
CA
ID
NV
MT
WY
UT
AZ
CO
NM
ND
SD
NE
KS
OK
TX
MN
IA
MO
AR
LA
WI
IL
MS
MI
IN
KY
TN
AL
OH
WV
GA
FL
SC
NC
VA
PA
NY
ME
AK

Important features

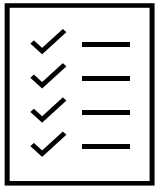


Model Robustness and Responsiveness



Unit Testing

Completed in 0.360 second, which means functionalities are stable and efficient.



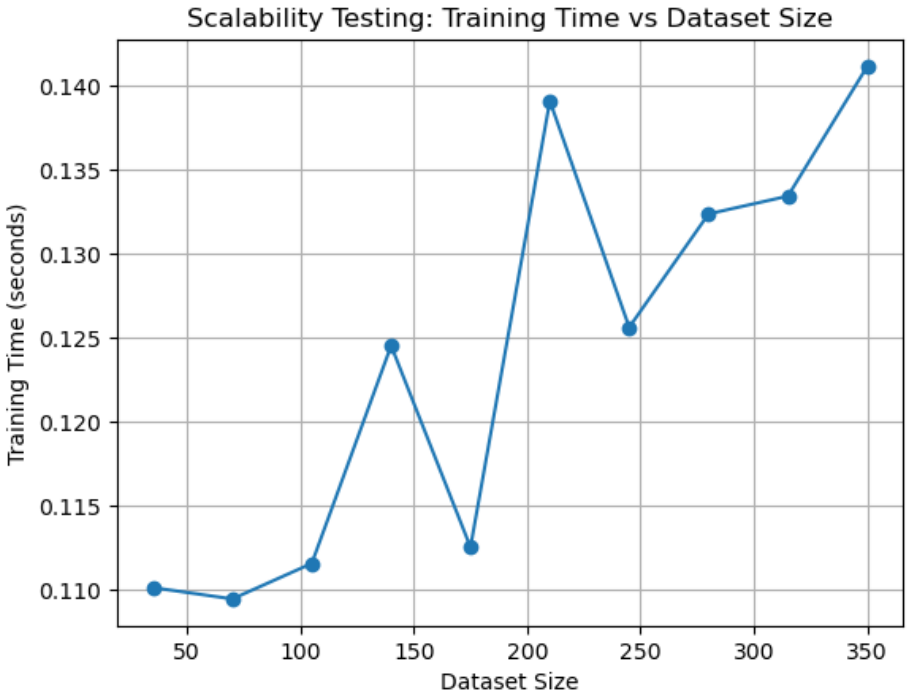
Edge Case Testing

Model successfully handled potential edge cases like errors and empty datasets.

Speed/Latency

Training latency at 0.1269 seconds and single prediction latency at 0 seconds

Scalability and Extensibility



Training Time

- Model was trained in 0.1611 seconds.

Load Testing

- Model handles 1000 calls in 5.0759 seconds.

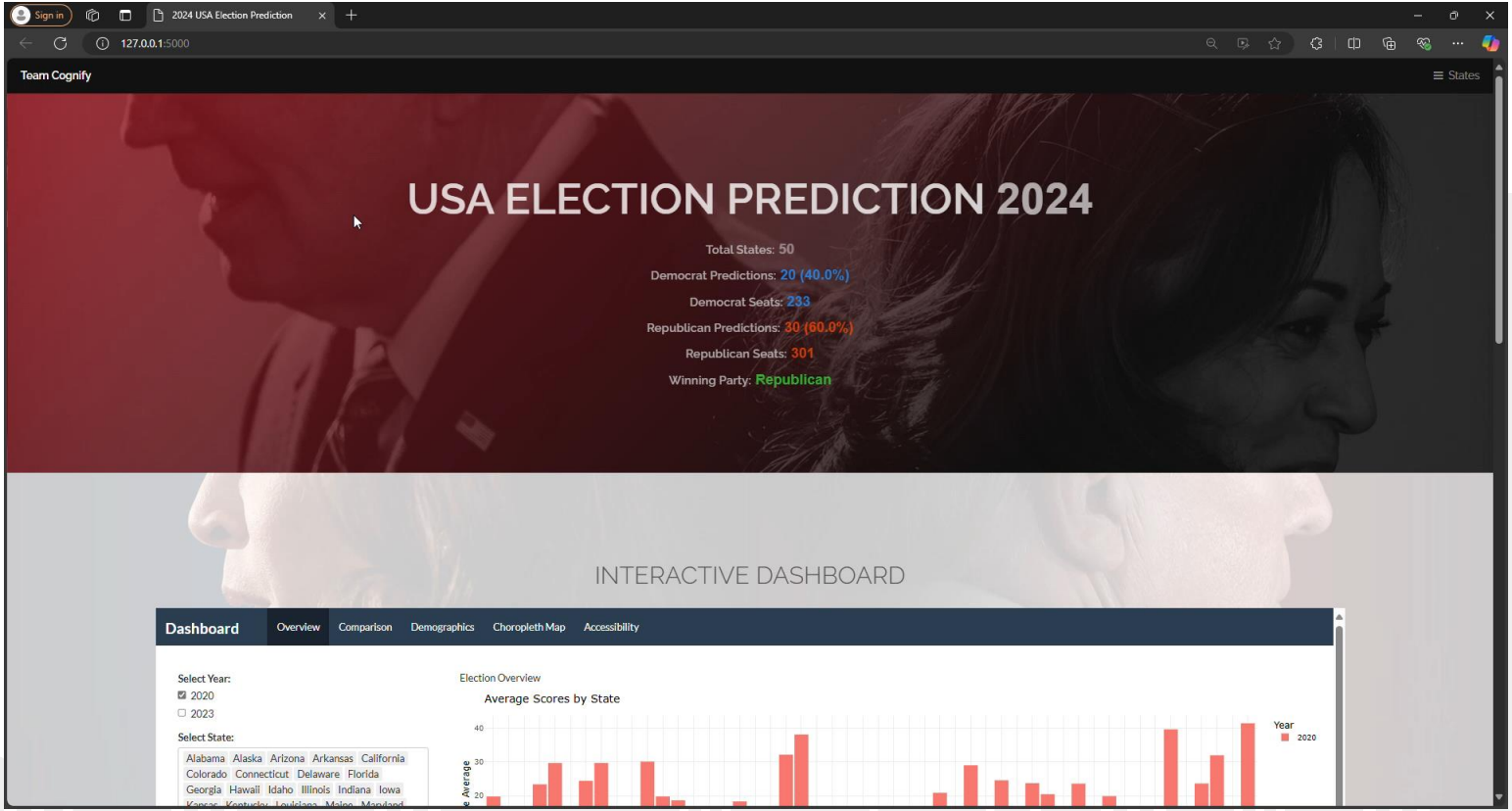
05 Model Deployment

Google Cloud Platform (GCP)

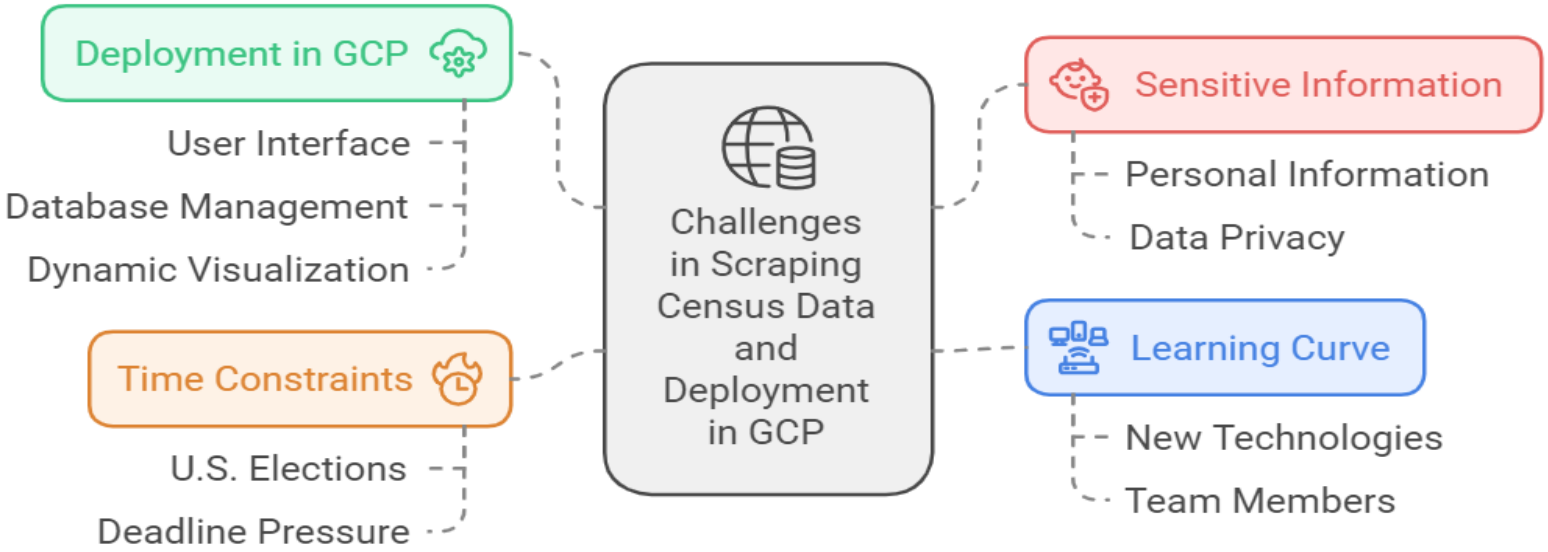
https://ovalofficeproject.nn.r.appspot.com/

Deployed the election prediction model to provide real-time predictions, ensuring scalability, and ease of access.

The deployment utilizes Google App Engine for API hosting, and BigQuery for analyzing large datasets.


Sign in
2024 USA Election Prediction
127.0.0.1:5000
Team Cognify
States
USA ELECTION PREDICTION 2024
Total States: 50
Democrat Predictions: 20 (40.0%)
Democrat Seats: 233
Republican Predictions: 30 (60.0%)
Republican Seats: 301
Winning Party: Republican
INTERACTIVE DASHBOARD
Dashboard
Overview
Comparison
Demographics
Choropleth Map
Accessibility
Select Year:
2020
2023
Select State:
Alabama Alaska Arizona Arkansas California
Colorado Connecticut Delaware Florida
Georgia Hawaii Idaho Illinois Indiana Iowa
Election Overview
Average Scores by State
Year
2020
40
30
20

Challenges


Deployment in GCP
User Interface
Database Management
Dynamic Visualization
Time Constraints
U.S. Elections
Deadline Pressure
Challenges in Scraping Census Data and Deployment in GCP
Sensitive Information
Personal Information
Data Privacy
Learning Curve
New Technologies
Team Members

06 Conclusion

Model Accuracy:
Achieved almost 97% accuracy in predicting the election results.

Future Potential:
With continuous improvements, the model can provide even more precise predictions.

Real-World Applications:
Offers valuable insights for future election forecasting and decision-making.

Resources


jupyter


Jira


Google Cloud


NumPy


pandas


Streamlit


SQL


GitHub


R

Thank You
Any Q/A