Work Performed:

- Developed a predictive model to find out about Voter Turnout in the 2014 general elections, using information from 50000 people from the State of Nevada.
- The objective here was to find whether the possible voter would come to vote or stay at Home.

Methodology

- **Data Preprocessing**
    - Handled redundant data by dropping where needed.
    - Converted categorical variables using one-hot encoding.
    - Normalized continuous variables for consistent scale.
- **Feature Selection**
    - Included demographic data (e.g., age, education, net worth).
    - Incorporated historical turnout features (vhXXp, vhXXg) as indicators.
    - Included commercial data signals (e.g., petowner_dog, intrst_nascar_in_hh, etc.).
- **Model Used**

  Used **Logistic Regression** for its interpretability and effectiveness in binary classification tasks like turnout prediction.

  The output included both:

    - A binary prediction (vote)
    - A probability score (vote_prob)

Model Evaluation

Model Performance was evaluated as given below:

- **ROC AUC Score**: ~0.79
- **Accuracy**: ~76%

Output Format

The prediction file includes:

- optimus_id
- All independent variables used
- vote: predicted (1) or not (0)
- vote_prob: predicted probability of voting (0.00 – 1.00)

Conclusion

This model can be further improved with better input data or more geographic and behavioral features.

However, in its current form, it provides strong directional insight for voter targeting strategies.