Election Prediction AI Models

Depending on the data we collect, there are a few different AI models I have researched that could fit our needs. Two algorithms stand out to me, based on their ability to cover simple datasets and complex datasets.

Logistic Regression – simple model used for binary classification. This can be used to predict win or loss. Uses the sigmoid function, which is used to map the predicted values to probabilities. There are different types of LR, however binomial supports two possible types of dependent variables, which in our case should cover the variables “win” or “loss”.

Gradient Boosting – builds upon weak learners sequentially and can handle various types of features. There are 2 boosting algorithms, AdaBoost and Gradient Boosting. If possible, both can be implemented where a registered user can use Gradient Boosting which provides a clearer/accurate result. This is due to Adaboost being more susceptible to noise and outlier in the data. However, with the right sources, there shouldn’t be too many outlying data points.

It’s best to start out with Logistic Regression and determine if the output is precise enough to display the results we want. If LR becomes too simple and cannot handle the information we obtain, then Gradient Boosting should be enough to cover everything relating to election data.