

# Perfect OLYMPIC PREDICTOR

## OUR IDEA

Using datasets from the Tokyo 2020 Olympic Games, we hypothesize that countries which have won medals in previous Olympics are more likely to win again in the future. Based on this hypothesis, we aim to predict the likelihood of each country winning a medal at the Paris 2024 Olympic Games. Our model considers various influencing factors, including the probability of a country winning any medal in the past and the percentage of winning teams sent from each country.

## OUR TEAM

We are a trio of Masters in Computer Science students, from Northeastern University in Seattle, attending our first hackathon. Click our names to learn more about us:



[Beth Shieh](#)



[Xiaoxue Fang](#)



[Joseph Song](#)

## OUR METHOD

We utilize the Naive Bayes algorithm to train our machine learning model.

$$P(A_1 | B_1, B_2) = P(A_1) \cdot P(B_1 | A_1) \cdot P(B_2 | A_1)$$

$P(A_1 | B_1, B_2)$  = Prediction for a country winning a medal in Paris  
 $P(A_1)$  = Proportion of times a country won a medal in past Olympics  
 $P(B_1 | A_1)$  = Probability of a team to win a medal for their country  
 $P(B_2 | A_1)$  = Proportion of teams being a winning team

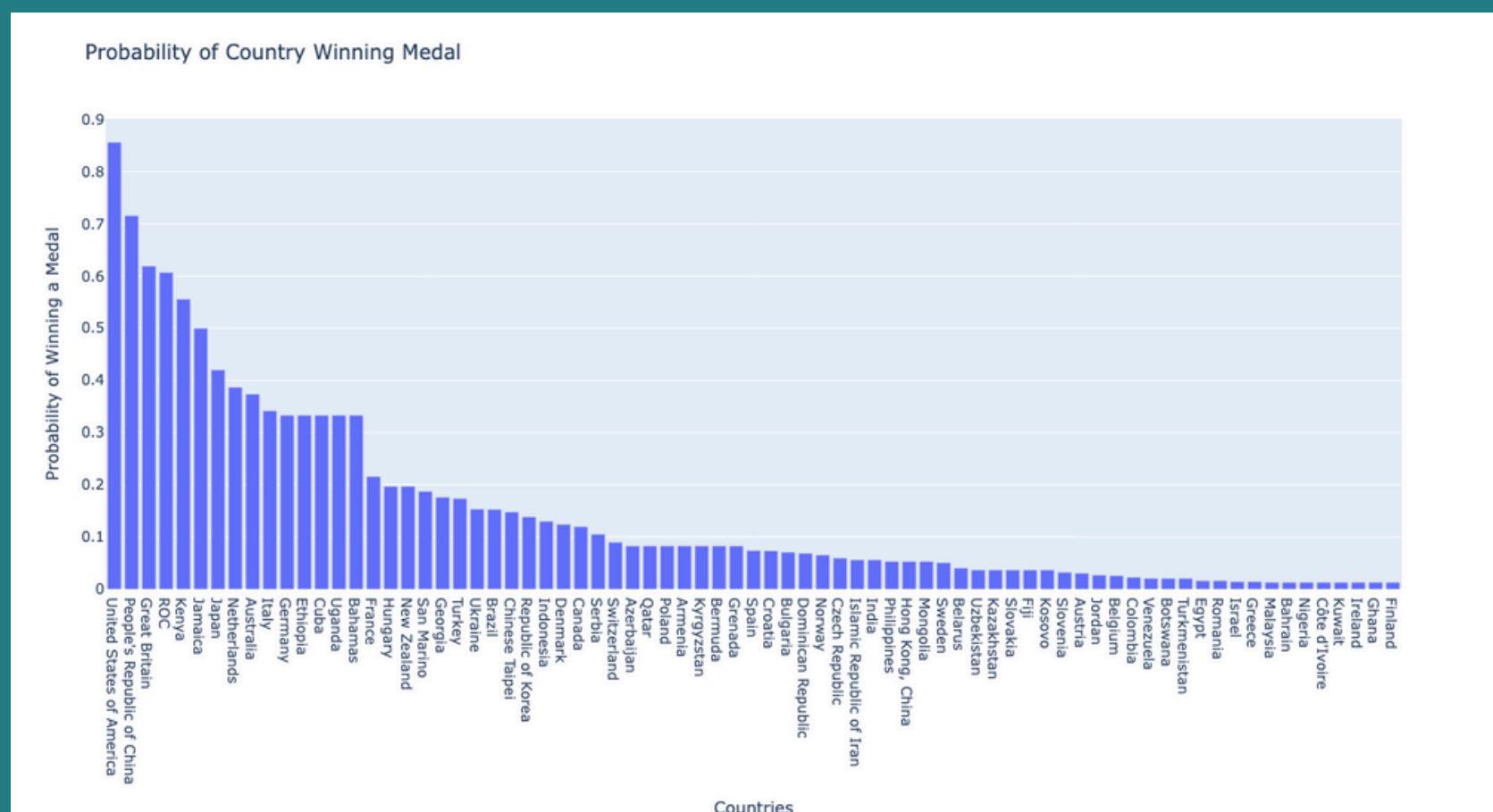
### Steps:

- Data ingestion and cleaning
- Probability calculations
- Generate Naive Bayes dataset
- Create predictions
- Data visualization

By leveraging Naive Bayes and historical Olympic data, this project provides a data-driven approach to predict Olympic outcomes, potentially aiding in strategic planning and preparation for future Olympic Games.

[Click here](#) to check out our full project on GitHub.

## PREDICTION GRAPH



## NAIVE BAYES LEARNING MODEL BUILDING

We built 4 learning models to train on the Tokyo dataset to see if we could build accurately predicting models for future events, given similar data.

## MODEL TRAINING

We used a non-standard 0.4 training split, given that we had a small data set from the Tokyo Olympics to work with.

## CLASS 1 AND 0 ACCURACY, PRECISION, AND RECALL RESULTS

