

Previsão de Vitórias em Partidas de CS:GO

Felipe Basilio da Silva

Setembro 2024

1 Introdução

Este projeto tem como objetivo desenvolver um modelo de aprendizado de máquina para prever o resultado de partidas de *Counter-Strike: Global Offensive* (CS:GO). Com a ascensão dos *eSports*, o interesse em análises preditivas de jogos competitivos tem crescido substancialmente. O problema a ser resolvido é um típico problema de classificação binária, onde o objetivo é prever se o primeiro time vencerá uma partida com base em um conjunto de variáveis explicativas que representam o desempenho dos jogadores, as estratégias das equipes e a dinâmica do jogo.

2 Descrição dos Dados

Os dados utilizados neste projeto incluem informações detalhadas sobre partidas de CS:GO entre os anos de 2017 e 2021, extraídos da HLTV. As variáveis disponíveis são:

- **Dados sobre os times:** Ranking mundial (*world rank*), total de pontos, e estatísticas agregadas de desempenho por mapa.
- **Desempenho por mapa:** Pontuação em cada mapa (M1, M2, M3), bem como os lados jogados (*CT/T*) e o resultado por metade.
- **Pick/Ban:** Sequência de mapas escolhidos e banidos por cada time.
- **Estatísticas detalhadas dos jogadores:** *Kill/Death ratio* (KD), *Average Damage per Round* (ADR), *Kill Assist Survival Trade* (KAST), e *Rating 2.0* tanto para os lados *CT* quanto *T*.

Essas variáveis formam a base para a criação do modelo preditivo.

3 Caracterização do problema e modelos a serem testados

O problema consiste em prever se o primeiro time (*first team*) vencerá a partida. O modelo a ser desenvolvido é de **classificação binária**, onde a variável alvo é o resultado da partida (*first_team_won* = 1 ou 0).

Os modelos que serão testados incluem:

- **Random Forest:** Um modelo de ensemble que combina várias árvores de decisão para melhorar a precisão.
- **Regressão Logística:** Um modelo simples e eficiente para problemas de classificação binária.
- **KNN (K-Nearest Neighbors):** Um algoritmo de classificação que utiliza a proximidade de vizinhos para determinar a classe de uma nova instância.
- **SVM (Support Vector Machines):** Um modelo que encontra a melhor fronteira de decisão entre as classes, maximizando a margem entre elas.