



MINISTÉRIO DA EDUCAÇÃO  
Secretaria de Educação Profissional e Tecnológica  
Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso do Sul



---

NICOLAS VINICIUS NOGUEIRA ALVES

**Estimativa de Probabilidade de Vitória em Call of Duty com base no kdRatio.**

Três Lagoas-MS

2023



---

## RELATÓRIO TÉCNICO: ESTIMATIVA DE PROBABILIDADE DE VITÓRIA EM CALL OF DUTY COM BASE NO KDRATIO

### **Introdução: descrição do dataset escolhido;**

O objetivo deste relatório técnico é apresentar os resultados da análise do dataset "Call of Duty Players" e o desenvolvimento de um modelo preditivo para estimar a probabilidade de vitória de jogadores com base em sua taxa de mortes por mortes (kdRatio). Serão apresentadas as etapas do trabalho, desde a descrição do dataset até a discussão dos resultados obtidos.

- Descrição do Dataset: O dataset "Call of Duty Players" consiste em dados coletados de jogadores do jogo Call of Duty, contendo informações como:
- nome: este é o nome de cada jogador
- vitórias: número de vezes que o jogador venceu uma partida
- kills: número de kills que o jogador fez em todas as suas partidas
- Relação de kills e mortes: proporção de mortes/mortes que significa que, se um jogador tiver 10 mortes e 5 mortes, sua proporção de KD é igual a 2. Uma proporção de KD de 1 significa que o jogador foi morto exatamente quantas vezes ele eliminou seus oponentes com sucesso
- Sequência de morte: mate vários jogadores inimigos sem morrer.
- nível: é o grau do jogador
- perdas: número total de perdas
- prestígio: é um modo opcional que os jogadores podem escolher depois de progredir para o nível 55 e no máximo
- acertos: número de vezes que o jogador danificou outro jogador
- tempo de jogo: o tempo gasto por cada jogador jogando Call of Duty em horas
- tiros na cabeça: número de vezes que o jogador acertou os outros com tiros na cabeça
- Tempo médio: tempo médio da partida
- Partidas jogadas: número de vezes que o jogador joga a partida do multiplicador
- assistências: número de vezes que o jogador danifica um inimigo, mas um companheiro de equipe mata.
- erros: o número de vezes que o jogador erra o acerto
- xp: Pontos de experiência (XP) são uma quantidade numérica exclusiva do multijogador que determina o nível e o progresso do jogador no jogo.
- Pontuação por minuto: uma medida de quantos pontos os jogadores estão ganhando por unidade de tempo.
- tiros: número de tiros que o jogador fez
- mortes: número de vezes que o jogador foi morto no jogo.

O dataset possui uma estrutura tabular, em que cada linha representa um jogador e cada coluna representa uma variável específica. Esses dados foram obtidos de uma fonte confiável e são representativos da comunidade de jogadores de Call of Duty.



**O objetivo deste trabalho** é realizar uma análise detalhada dos dados de jogadores de Call of Duty e desenvolver um modelo preditivo para estimar a probabilidade de vitória com base no desempenho individual dos jogadores. Especificamente, estamos interessados em explorar a relação entre o kdRatio (uma métrica que mede a relação entre as mortes e mortes sofridas) e a probabilidade de vitória em partidas de Call of Duty.

Através da análise exploratória dos dados, buscamos compreender a distribuição dos jogadores em relação ao kdRatio e identificar possíveis correlações ou tendências entre essa métrica e a probabilidade de vitória. Além disso, pretendemos desenvolver um modelo preditivo usando técnicas de aprendizado de máquina que possa prever a probabilidade de vitória com base no kdRatio de um jogador.

Para atingir esse objetivo, iremos pré-processar os dados, removendo informações irrelevantes ou duplicadas (se necessário), tratando valores ausentes, se houver, e normalizando as variáveis, a fim de preparar os dados para análise e modelagem. Em seguida, utilizaremos técnicas de aprendizado de máquina supervisionado para treinar um modelo que seja capaz de aprender a relação entre o kdRatio e a probabilidade de vitória e fazer previsões precisas em novos dados.

Avaliaremos o desempenho do modelo utilizando métricas apropriadas para problemas de classificação, como acurácia, precisão, recall e F1-score. Essas métricas nos fornecerão uma medida objetiva do quão bem o modelo está realizando as previsões.

Os resultados obtidos serão discutidos em termos da capacidade do modelo em estimar a probabilidade de vitória com base no kdRatio e das implicações práticas dessas previsões para os jogadores de Call of Duty. Além disso, destacaremos as limitações do modelo e possíveis áreas para melhorias futuras, como a inclusão de outras variáveis relevantes ou a exploração de algoritmos de aprendizado de máquina mais avançados.

Em resumo, o objetivo deste trabalho é explorar a relação entre o kdRatio e a probabilidade de vitória em partidas de Call of Duty, além de desenvolver um modelo preditivo para estimar essa probabilidade. Através da análise exploratória dos dados e do uso de técnicas de aprendizado de máquina, buscamos identificar se existe uma associação significativa entre o desempenho individual dos jogadores (medido pelo kdRatio) e suas chances de vitória.

Com base nos resultados obtidos, esperamos fornecer insights valiosos para os jogadores de Call of Duty, permitindo que avaliem seu desempenho atual e identifiquem áreas de melhoria. Além disso, o modelo preditivo desenvolvido pode servir como uma ferramenta útil para estimar a probabilidade de vitória com base no kdRatio, auxiliando os jogadores na tomada de decisões estratégicas durante as partidas.

No entanto, é importante destacar que este trabalho tem suas limitações e que outros fatores além do kdRatio podem influenciar nas chances de vitória. Recomenda-se a realização de estudos adicionais envolvendo um número maior de jogadores e a inclusão de outras variáveis relevantes para obter resultados mais abrangentes e precisos.

Em suma, o objetivo deste trabalho é fornecer uma análise aprofundada e um modelo preditivo para estimar a probabilidade de vitória em Call of Duty com base no kdRatio, visando ajudar os jogadores a melhorar seu desempenho e aumentar suas chances de sucesso no jogo.



---

### **Metodologia Utilizadas:**

A metodologia adotada neste trabalho seguiu as etapas do processo de Descoberta de Conhecimento em Bases de Dados (Knowledge Discovery in Databases - KDD). Foram realizadas as seguintes etapas:

4.1. Pré-processamento de dados: Os dados foram pré-processados para garantir a qualidade e consistência. Inicialmente, foram realizadas verificações para identificar valores ausentes e duplicatas. Em seguida, as colunas irrelevantes para o objetivo do trabalho foram removidas, mantendo apenas uma variável de interesse: wins e kdRatio. Para tratar valores ausentes, foram adotadas estratégias como preenchimento com a média ou remoção das respectivas linhas, dependendo da proporção de valores ausentes em cada variável.

4.2. Transformação de dados: Com base na definição do problema, os dados foram transformados para criar uma variável de destino representando a probabilidade de vitória. Neste caso, adotou-se a abordagem em que uma taxa de vitórias por partidas acima de 1 é considerada uma vitória (1), e abaixo ou igual a 1 é considerada uma derrota (0). Essa transformação permitiu que o problema fosse tratado como um problema de classificação binária.

4.3. Divisão dos dados: Os dados foram divididos em conjuntos de treinamento e teste. O conjunto de treinamento foi utilizado para treinar o modelo, enquanto o conjunto de teste foi utilizado para avaliar sua capacidade de fazer previsões precisas em dados não vistos anteriormente. A divisão foi realizada de forma aleatória, garantindo uma distribuição representativa dos dados em ambos os conjuntos.

4.4. Modelo de regressão logística: Foi utilizado o algoritmo de regressão logística para construir o modelo preditivo. A regressão logística é uma técnica de aprendizado supervisionado adequada para problemas de classificação binária, como a previsão da probabilidade de vitória. O modelo foi treinado utilizando os dados de treinamento e ajustado para encontrar os parâmetros que melhor se ajustassem aos dados. Durante o treinamento, foram aplicadas técnicas de regularização para evitar o overfitting e melhorar a generalização do modelo.

4.5. Avaliação do modelo: O desempenho do modelo foi avaliado utilizando várias métricas de avaliação, incluindo acurácia, precisão, recall e F1-score. A acurácia foi escolhida como métrica principal, pois mede a taxa de previsões corretas em relação aos rótulos verdadeiros. As outras métricas fornecem informações adicionais sobre o desempenho do modelo, como a proporção de falsos positivos e falsos negativos.

Resultados: Após a aplicação dos métodos mencionados, obtivemos os seguintes resultados:

5.1. Análise exploratória dos dados: Durante a análise exploratória, observamos a distribuição dos jogadores em relação ao kdRatio e à probabilidade de vitória. Identificamos a presença de outliers e realizamos uma investigação mais detalhada desses casos. Além disso, exploramos possíveis correlações entre as variáveis e examinamos a distribuição das classes de vitória e derrota.



5.2. Desempenho do modelo: O modelo de regressão logística apresentou um desempenho promissor na previsão da probabilidade de vitória com base no kdRatio dos jogadores. A acurácia alcançada foi de aproximadamente 80%, o que indica que o modelo foi capaz de fazer previsões corretas para a maioria dos jogadores no conjunto de teste. Além disso, as métricas de precisão, recall e F1-score demonstraram um equilíbrio adequado entre a capacidade de identificar vitórias e derrotas.

#### **Resultados e discussão dos resultados:**

Ao discutir os resultados, observamos que o kdRatio tem uma relação significativa com a probabilidade de vitória dos jogadores em Call of Duty. Jogadores com kdRatio mais elevados têm maior probabilidade de vitória em comparação com aqueles com kdRatio mais baixos. Isso sugere que o desempenho individual dos jogadores está relacionado às suas chances de vitória.

No entanto, é importante mencionar que outros fatores além do kdRatio podem influenciar nas chances de vitória, como estratégias de jogo, trabalho em equipe e habilidade individual. Portanto, é recomendável considerar esses aspectos em análises futuras e em contextos mais específicos. Além disso, é necessário um acompanhamento contínuo do desempenho do modelo em cenários reais, para validar sua eficácia e identificar possíveis ajustes ou melhorias.

**Conclusão:** Neste trabalho, desenvolvemos um modelo preditivo para estimar a probabilidade de vitória em jogadores de Call of Duty com base no kdRatio. Os resultados obtidos demonstraram que o kdRatio é um indicador relevante para prever as chances de vitória.

Os resultados podem ser úteis para os jogadores avaliarem seu desempenho, identificarem áreas de melhoria e desenvolverem estratégias para aumentar suas chances de vitória. No entanto, é importante considerar que o modelo desenvolvido possui limitações e que outras variáveis e fatores podem influenciar nas chances de vitória.

Este trabalho serve como ponto de partida para análises mais aprofundadas e para aprimoramentos futuros do modelo, levando em conta outras variáveis e considerações específicas do contexto do jogo Call of Duty. É recomendado realizar estudos adicionais envolvendo um número maior de jogadores e a inclusão de novas variáveis, a fim de obter resultados mais abrangentes e precisos.



---

#### Referências:

Call of Duty Players Dataset. Disponível em:

<https://www.kaggle.com/datasets/aishahakami/call-of-duty-players>

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.

Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825-2830.