

# Relatório Técnico Análise API Brawl Stars

Vitor Hugo Ferreira<sup>1</sup>

<sup>1</sup>Instituto Federal Goiano - Campus Iporá  
Iporá - Goiás - Brasil

{vitor.hugo2}@estudante.ifgoiano.edu.br

## 1. Introdução

O trabalho tem como objetivo realizar uma análise de desempenho dos personagens (“Brawlers”) do jogo Brawl Stars, com foco na identificação dos mais fortes e dos mais fracos em diferentes modos de jogo. A proposta é utilizar técnicas de análise de dados e aprendizado de máquina para extrair padrões a partir de métricas como win rate (taxa de vitória) e pick rate (frequência de escolha), fornecendo uma visão quantitativa sobre o equilíbrio competitivo entre os personagens.

A motivação para o desenvolvimento deste estudo surge do envolvimento pessoal do autor com o jogo, tanto como jogador há mais de seis anos quanto como participante do cenário competitivo. Essa experiência prática permitiu compreender que a percepção dos jogadores nem sempre reflete o real desempenho estatístico dos personagens, o que reforça a importância de uma abordagem baseada em dados.

Para alcançar esse objetivo, foi aplicada a Regressão Logística, técnica de inteligência artificial, a fim de prever a probabilidade de vitória de cada Brawler com base em suas estatísticas. Além disso, foi desenvolvido um score ponderado de win rate e pick rate, permitindo uma classificação ordenada dos personagens em diferentes contextos de jogo.

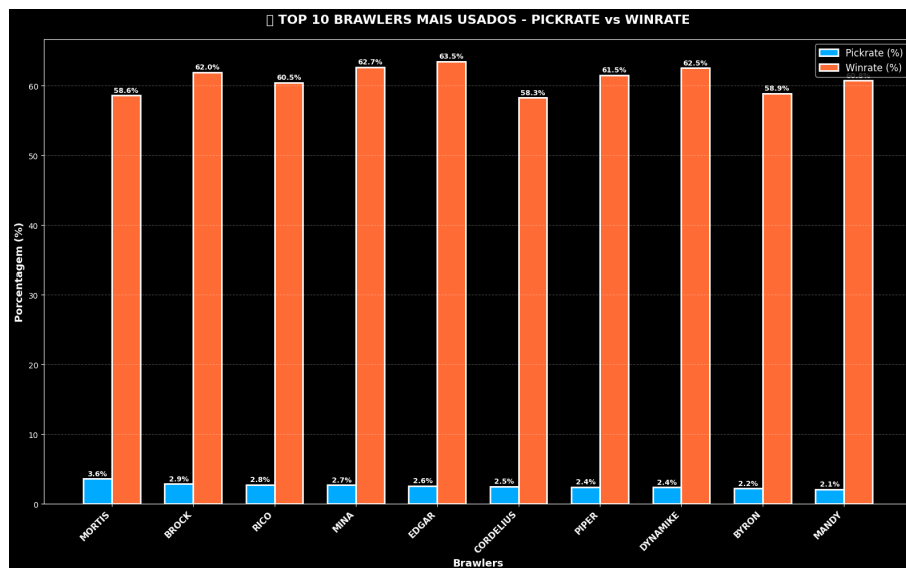
Essa análise busca contribuir para o entendimento do metajogo (metagame) competitivo de Brawl Stars, fornecendo um modelo quantitativo capaz de apoiar tanto jogadores quanto analistas na avaliação de balanceamento e desempenho dos personagens.

## 2. Base de dados e pré-processamento

A base de dados utilizada neste trabalho foi construída a partir da API oficial do Brawl Stars, disponibilizada pela Supercell. A coleta foi realizada em Python utilizando bibliotecas como requests, pandas e json, com autenticação via token para acessar os dados do ranking global. Foram obtidas informações de jogadores de diferentes regiões, incluindo Brasil, Estados Unidos, China, Turquia, México e Japão, com o objetivo de atingir uma amostra ampla e diversificada de partidas. Cada jogador teve seu histórico de batalhas coletado, contendo detalhes como o modo de jogo, os personagens utilizados e o resultado da partida (vitória ou derrota).

Durante o processo de pré-processamento, os dados brutos foram organizados e transformados em tabelas estruturadas, permitindo análises estatísticas por personagem e modo de jogo. Foram calculadas métricas fundamentais como taxa de vitória (winrate) e taxa de uso (pickrate), além de um AI Score que combina ambas as informações de forma ponderada. Todos os dados numéricos foram normalizados para que as variáveis tivessem escalas comparáveis, o que é essencial para o desempenho dos algoritmos de aprendizado

de máquina. O processamento completo e preparação dos dados levou aproximadamente 12 horas, devido ao grande volume de informações analisadas.



**Figure 1. Distribuição das taxas de uso e vitória dos top 10 mais usados (Winrate x Pickrate) por personagem.**

Por fim, foi aplicada uma clusterização exploratória com K-Means, utilizada apenas para obter uma visão geral dos agrupamentos de desempenho entre os personagens, sem constituir o foco principal do estudo. O objetivo central da pesquisa concentrou-se na aplicação de regressão logística para prever probabilidades de vitória de cada personagem em diferentes modos de jogo. Assim, a etapa de pré-processamento consolidou uma base sólida e equilibrada para a etapa de modelagem e análise de resultados.

### 3. Modelo usado (Regressão Logística)

A regressão logística é um dos modelos estatísticos mais tradicionais e amplamente utilizados para classificação binária e previsão de probabilidades. Segundo [Hosmer et al. 2013], ela tem como objetivo estimar a probabilidade de ocorrência de um evento, como vitória ou derrota, a partir de variáveis independentes, ajustando uma função sigmoide que mapeia valores contínuos para o intervalo entre 0 e 1. Essa característica a torna especialmente útil em contextos de previsão de resultados categóricos, sendo uma técnica fundamental em diversas áreas da ciência de dados, desde epidemiologia até aprendizado de máquina supervisionado.

A regressão logística foi aplicada no presente trabalho como núcleo do processo preditivo, sendo responsável por estimar a probabilidade de vitória de cada personagem (*brawler*) em diferentes modos de jogo. O código foi desenvolvido em Python, utilizando a biblioteca `scikit-learn` dentro de um pipeline completo que incluiu etapas de normalização, transformação categórica e calibração de probabilidades. Inicialmente, os dados foram extraídos de um arquivo JSON agregado contendo mais de um milhão de partidas, formando um conjunto proxy balanceado entre vitórias e derrotas. As variáveis categóricas, como o nome do *brawler* e o modo de jogo, foram codificadas por meio de

`OneHotEncoder`, enquanto as variáveis numéricas, como taxas de vitória e número de partidas, foram padronizadas com `StandardScaler`. O conjunto de dados foi então dividido em treino (80%) e teste (20%), garantindo estratificação para preservar a proporção entre vitórias e derrotas.

## Logistic Regression Pipeline

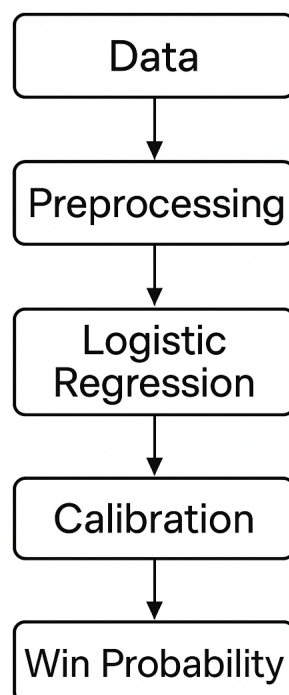


Figure 2. Fluxo do processo de aplicação da Regressão Logística no dataset.

### 3.1. Otimização de hiperparâmetros

A etapa de otimização de hiperparâmetros foi conduzida utilizando a técnica de busca em grade (`GridSearchCV`), amplamente adotada em aprendizado de máquina por permitir a avaliação sistemática de diferentes combinações de parâmetros. No caso da Regressão Logística, o principal hiperparâmetro ajustado foi o coeficiente de regularização ( $C$ ), responsável por controlar o equilíbrio entre o ajuste do modelo aos dados e a generalização para novos exemplos. O objetivo foi reduzir o risco de sobreajuste (*overfitting*), garantindo que o modelo aprendesse padrões reais de desempenho dos *brawlers*, e não apenas ruídos do conjunto de treino. O processo de busca foi realizado em múltiplas iterações, explorando valores crescentes de  $C$  até encontrar aquele que maximizou o desempenho sem comprometer a estabilidade.

### 3.2. Validação cruzada

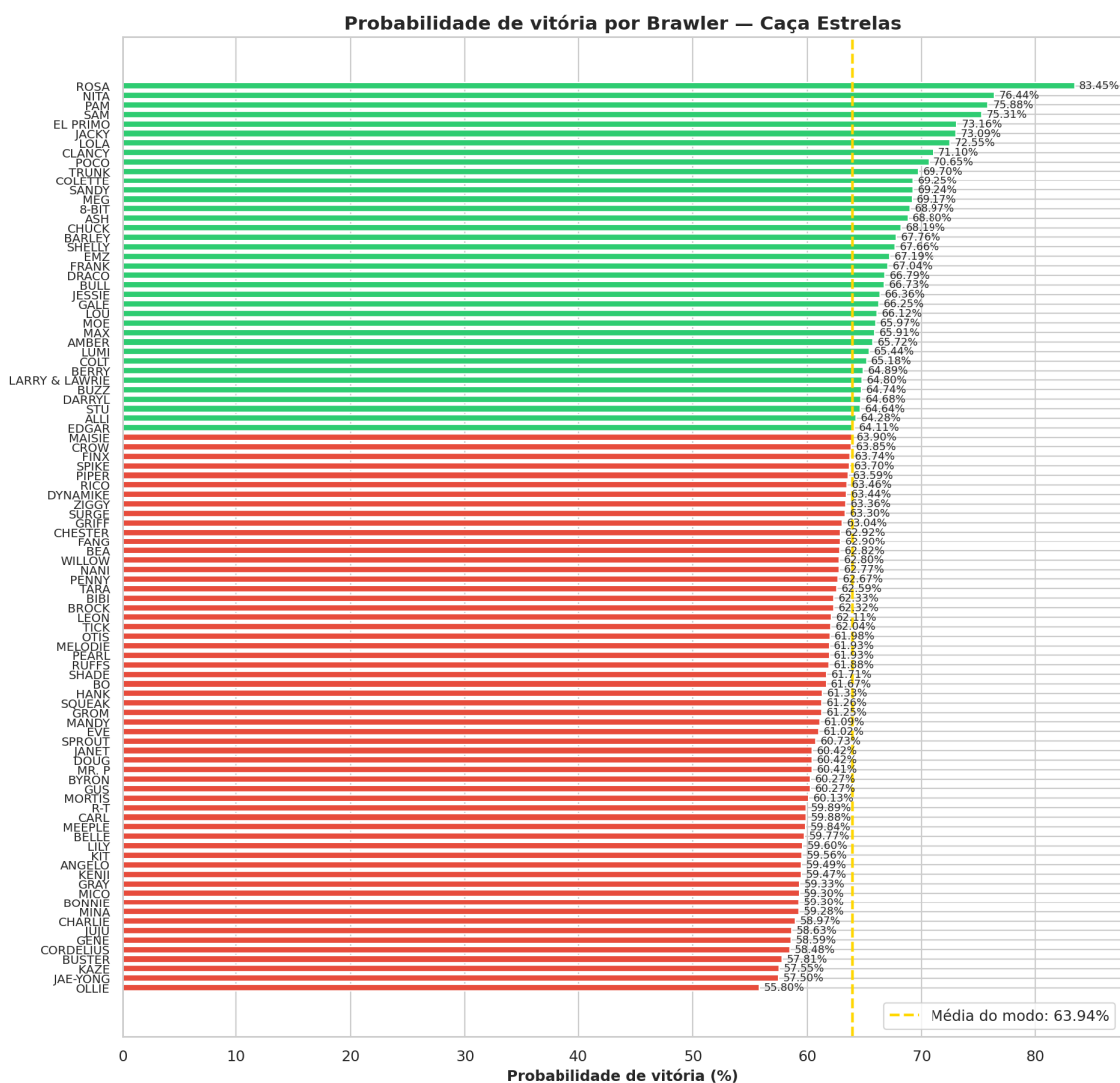
Para assegurar a confiabilidade dos resultados, foi empregada a técnica de validação cruzada (*cross-validation*), dividindo o conjunto de dados em múltiplas partições (*folds*). A cada iteração, uma parte dos dados era reservada para teste, enquanto o restante era utilizado para o treinamento. Essa estratégia reduz a dependência de divisões aleatórias e fornece uma estimativa mais estável do desempenho médio do modelo. Além disso, o processo foi estratificado, preservando a proporção entre vitórias e derrotas em cada subdivisão, o que é essencial para evitar vieses de amostragem em um conjunto de mais de 900 mil partidas. Essa abordagem contribuiu para uma análise mais justa do comportamento do modelo frente a diferentes subconjuntos do jogo.

### 3.3. Métricas de avaliação

O modelo final foi avaliado com base em métricas amplamente utilizadas para tarefas de classificação probabilística: AUC, LogLoss e Accuracy. A métrica AUC (Área Sob a Curva ROC) mede a capacidade do modelo em distinguir entre vitórias e derrotas, enquanto a LogLoss avalia a confiabilidade das probabilidades preditas, penalizando previsões incorretas com alta confiança. Já a Accuracy fornece uma visão mais intuitiva da proporção de acertos totais. No contexto deste trabalho, o modelo alcançou valores de AUC igual a 0.5424, LogLoss de 0.6585 e Accuracy de 0.6254, indicando um desempenho satisfatório, especialmente considerando a natureza competitiva e variável do metajogo do *Brawl Stars*.

## 4. Resultados e interpretação

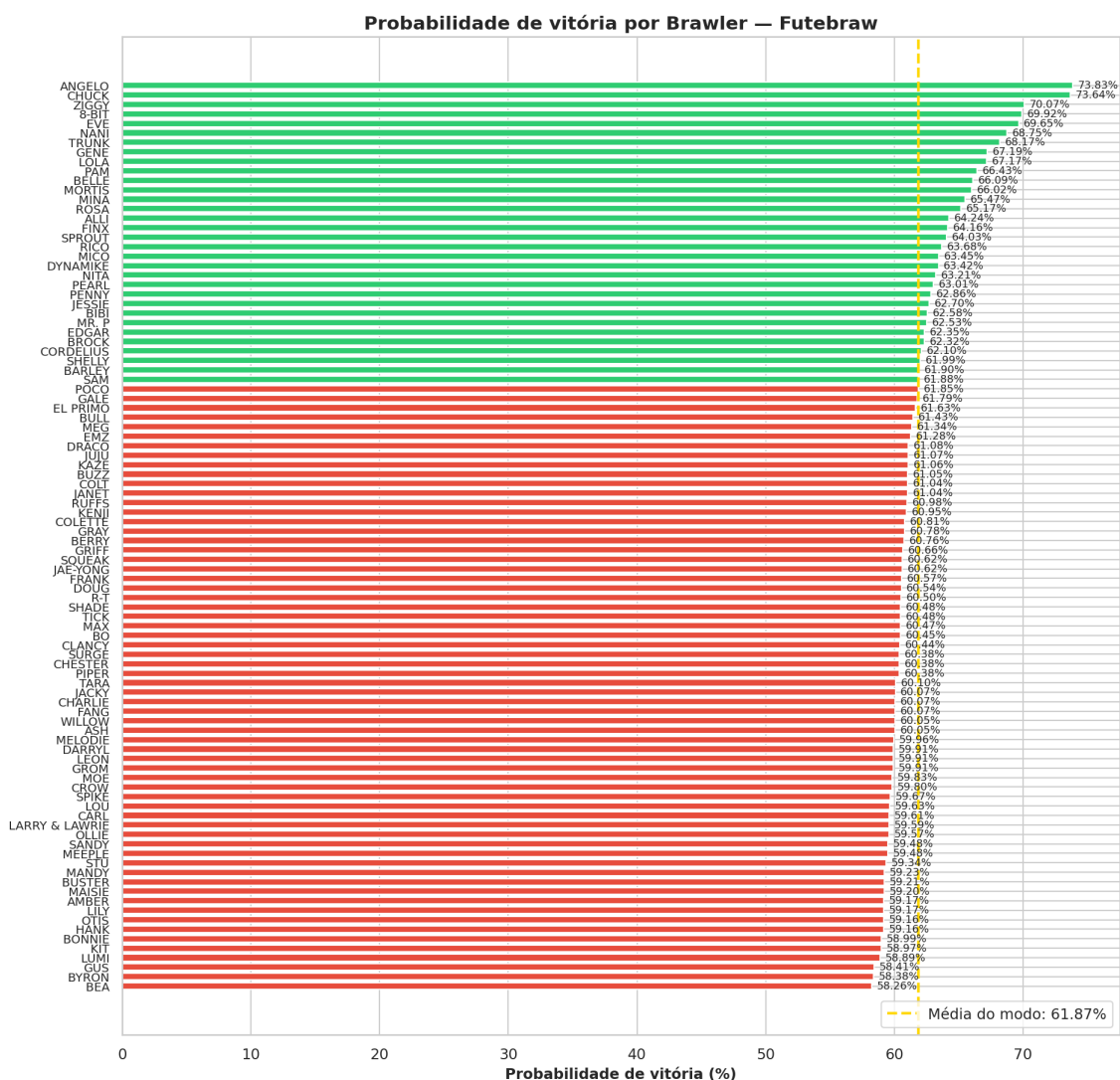
Para compreender o desempenho dos personagens em diferentes contextos de jogo, foram gerados gráficos individuais para cada modo – Caça Estrelas, Futebrawl, Nocaute, Pique-Gema, Roubo e Zona Estratégica – a partir das probabilidades de vitória ( $p_{win}$ ) estimadas pelo modelo de Regressão Logística. Cada gráfico apresenta todos os *brawlers* em ordem crescente de desempenho, com barras verdes representando aqueles com probabilidade acima da média do modo e barras vermelhas indicando desempenho inferior. A linha amarela tracejada marca a média geral de cada modo, funcionando como referência visual para comparação entre personagens. Essa abordagem possibilita identificar, de forma objetiva, quais *brawlers* possuem maior potencial competitivo em determinados modos, permitindo observar tendências de dominância e fragilidade dentro do metajogo atual.



**Figure 3. Probabilidade de vitória por Brawler — Modo Caça Estrelas.**

No modo Caça Estrelas, observa-se o domínio de Nita e Pam, que se destacam por serem brawlers controles, cuja principal vantagem está na capacidade de eliminar adversários e controlar area, um fator essencial neste modo, em que o acúmulo de eliminações define o resultado da partida. Além disso, chama atenção o desempenho da Rosa, que figura no topo do gráfico. Embora sua utilização seja restrita a mapas com grande presença de vegetação, ela apresenta resultados impressionantes, alcançando 83,45% de taxa de vitória.

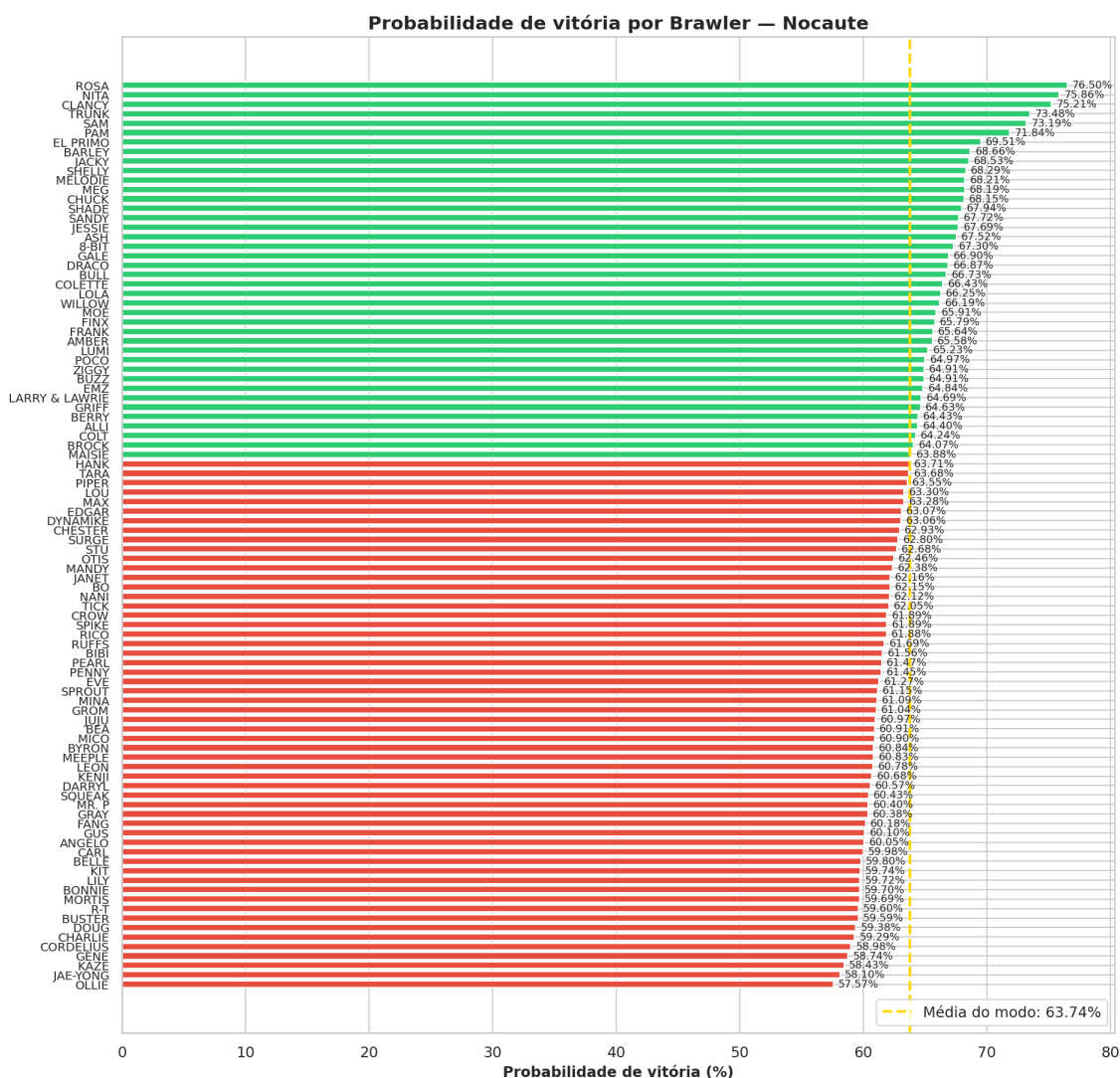
Na faixa inferior do desempenho, encontram-se Jae-Yong e Ollie, que se mostram menos eficazes neste contexto. O Ollie não possui grande destaque, pois depende de se expor excessivamente para conseguir abates, o que o torna vulnerável. Já o Jae-Yong sofre com a desvantagem de ser um suporte e causar pouco dano, tendo dificuldade em enfrentar o controle predominante neste modo, o que limita significativamente sua eficiência em combate.



**Figure 4. Probabilidade de vitória por Brawler — Modo Futebraw.**

No modo Futebraw, observa-se o destaque de Angelo e Chuck. O Angelo se sobressai por seu grande alcance e alto dano individual, enquanto o Chuck demonstra excelente desempenho devido à sua mobilidade, que permite avançar rapidamente e criar oportunidades de gol.

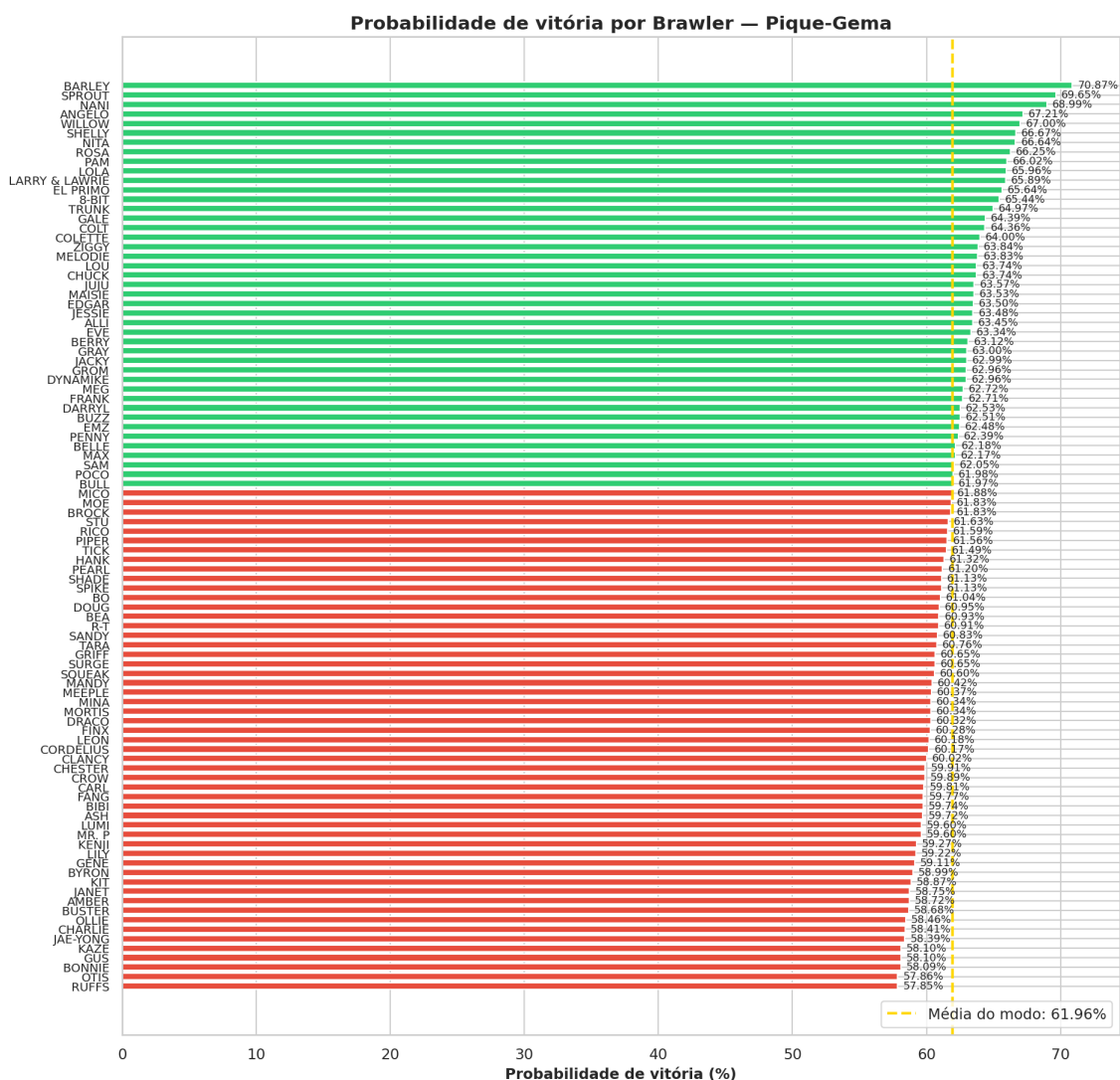
Por outro lado, os três piores desempenhos pertencem a Gus, Byron e Bea. O principal fator para o baixo rendimento desses brawlers é o domínio do Ângelo, considerado o melhor deste modo, que acaba contra-atacando todos eles de forma muito eficiente.



**Figure 5. Probabilidade de vitória por Brawler — Modo Nocaute.**

No modo Nocaute, observam-se as presenças de Rosa, Trunk, Nita e Clancy. Os dois primeiros são tanques que exercem grande pressão, o que é essencial em um modo onde morrer uma única vez pode decidir a partida. Já Clancy e Nita se destacam pelo alto dano e pelo controle de área, características fundamentais para dominar o campo de batalha.

Os piores desempenhos voltam a ser de Ollie e Jae Yong, pelos mesmos motivos observados no modo Caça Estrelas ambos precisam se expor demais para eliminar adversários ou tem pouco dano e acabam sendo punidos facilmente. Esses modos apresentam muitas semelhanças, com a principal diferença sendo que, no Nocaute, basta uma única eliminação para definir a vitória.

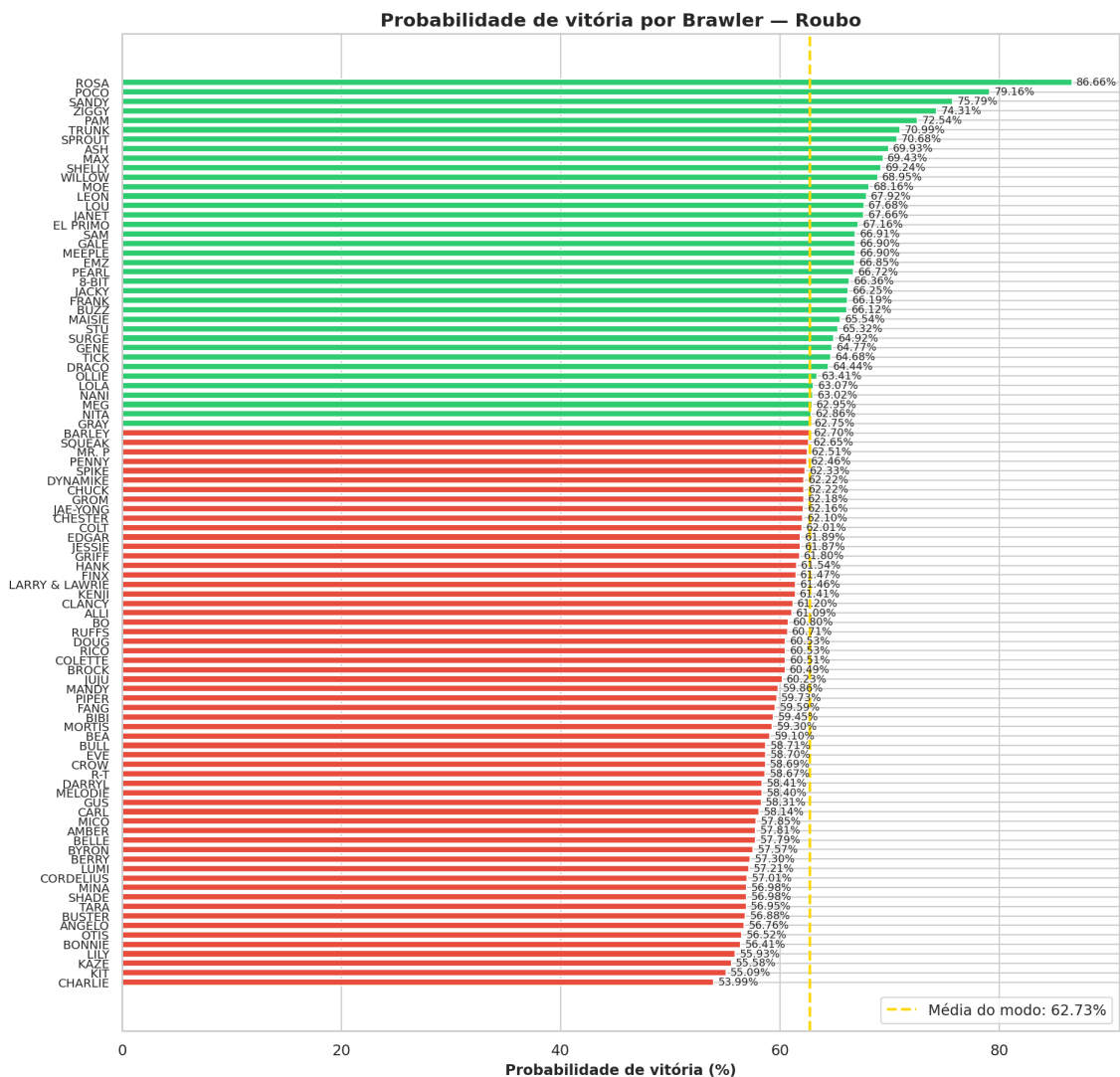


**Figure 6. Probabilidade de vitória por Brawler — Modo Pique-Gema.**

No modo Pique-Gema, observam-se Barley e Sprout no topo do desempenho. Ambos são brawlers lançadores capazes de atacar por trás das paredes, o que lhes confere uma grande vantagem estratégica neste modo, permitindo controlar o centro do mapa e proteger as gemas com segurança.

Por outro lado, Ruffs e Otis aparecem entre os piores. Ambos sofrem contra os lançadores dominantes, já que têm dificuldade em acertar ataques através dos obstáculos e acabam recebendo dano constante sem conseguir reagir efetivamente.

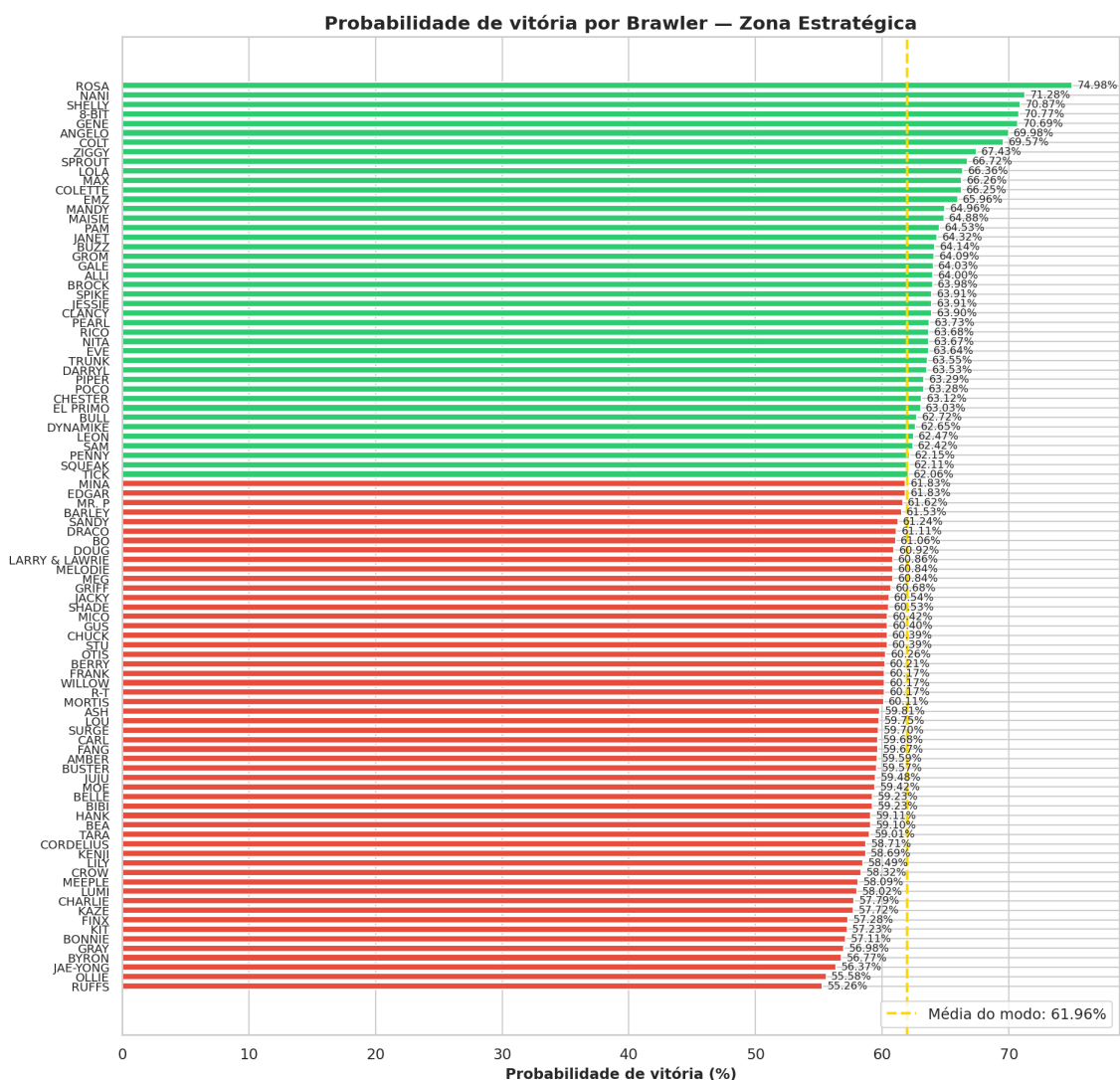




**Figure 7. Probabilidade de vitória por Brawler — Modo Roubo.**

No modo Roubo observa-se um destaque estrondoso da brawler Rosa com 86,66% de vitória, isso se deve a defesa absurda, sua super no game te dá um escudo capaz de tankar 70% de dano por 5 segundos, em segundo vemos Poco ele juntamente com a Rosa formam o combo mais forte do jogo deixando esses dois personagens com taxas de vitória absurdas.

Já na parte de baixo mostra que pior brawler é a Charlie, isso se deve a sua baixa capacidade de causar dano e sua fragilidade, como segundo pior vemos o Kit ele é ruim nesse modo porque possui um alcance muito curto e é counterado pelo topo do meta.



**Figure 8. Probabilidade de vitória por Brawler — Modo Zona Estratégica.**

No modo Zona Estratégica, a brawler Rosa se destaca como a melhor opção, principalmente devido à sua alta capacidade de defesa e sustentabilidade nas zonas de controle. Ao lado dela, Nani também é uma escolha poderosa, notável pela sua grande quantidade de dano a longas distâncias e a utilidade estratégica de sua Habilidade Estelar de teletransporte.

No entanto, na parte inferior da lista, com as piores probabilidades de vitória, encontramos Ruffs, que se mostra um suporte bastante inconsistente neste modo. Além dele, Ollie (presumindo que seja "Gale" ou outro brawler de dano baixo/controlado) está em desvantagem devido ao seu dano ser considerado baixo em comparação com os tanques mais dominantes no meta atual

## 5. Conclusão e próximos passos

Este estudo alcançou seu objetivo de fornecer uma análise de desempenho quantitativa dos Brawlers de Brawl Stars, utilizando a Regressão Logística para estimar a probabilidade de vitória em diferentes modos de jogo. Os resultados confirmam que o metajogo é

altamente contextual, com alguns Brawlers demonstrando domínio esmagador em modos específicos, como a Rosa em Roubo (86,66% de  $p_{win}$ ) e Zona Estratégica, graças à sua defesa e sustentabilidade. Da mesma forma, Nita e Pam se destacam no Caça Estrelas por seu controle de área. Essa análise estatística válida, em grande parte, a percepção competitiva da comunidade, mas aprofunda-a ao fornecer uma classificação objetiva e preditiva. Por outro lado, a identificação consistente de Brawlers como Ollie e Jae-Yong nas faixas de baixo desempenho em múltiplos modos como Caça Estrelas e Nocaute aponta para possíveis desequilíbrios que merecem atenção por parte dos desenvolvedores, sendo personagens que sofrem por depender de exposição ou por terem baixo dano em relação ao meta dominante.

Para elevar a profundidade e a relevância desta análise, uma das principais ideias para o futuro do projeto é expandir o foco da performance individual dos Brawlers para a dinâmica coletiva das partidas. Seria pertinente investigar a possibilidade de refinar o modelo preditivo para analisar a sinergia entre composições de equipe (trios) e identificar os confrontos diretos (matchups ou counters) mais eficazes para cada personagem. Isso poderia ser explorado através da expansão do dataset para incluir a formação completa das equipes em cada partida, permitindo a aplicação de modelos de Machine Learning mais robustos, como árvores de decisão ou redes neurais. Essa potencial expansão permitiria prever a chance de sucesso de uma composição inteira contra formações adversárias específicas, fornecendo insights estratégicos de nível competitivo (metajogo) e de seleção de personagens (draft) ainda mais valiosos.

## References

Hosmer, D. W., Lemeshow, S., and Sturdivant, R. X. (2013). *Applied Logistic Regression*. John Wiley & Sons, Hoboken, NJ, 3rd edition.