



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE  
CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA  
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

**JAILSON BARROS DA SILVA JUNIOR**

**LEAGUE OF LEGENDS:  
PREDIÇÃO DE RESULTADOS EM TEMPO REAL**

**CAMPINA GRANDE - PB**

**2023**

**JAILSON BARROS DA SILVA JUNIOR**

**LEAGUE OF LEGENDS:  
PREDIÇÃO DE RESULTADOS EM TEMPO REAL**

**Trabalho de Conclusão Curso apresentado ao Curso Bacharelado em Ciência da Computação do Centro de Engenharia Elétrica e Informática da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.**

**Orientador : Cláudio Elízio Calazans Campelo**

**CAMPINA GRANDE - PB**

**2023**

**JAILSON BARROS DA SILVA JUNIOR**

**LEAGUE OF LEGENDS:  
PREDIÇÃO DE RESULTADOS EM TEMPO REAL**

**Trabalho de Conclusão Curso  
apresentado ao Curso Bacharelado em  
Ciência da Computação do Centro de  
Engenharia Elétrica e Informática da  
Universidade Federal de Campina  
Grande, como requisito parcial para  
obtenção do título de Bacharel em  
Ciência da Computação.**

**BANCA EXAMINADORA:**

**Cláudio Elízio Calazans Campelo  
Orientador – UASC/CEEI/UFCG**

**Cláudio de Souza Baptista  
Examinador – UASC/CEEI/UFCG**

**Francisco Vilar Brasileiro  
Professor da Disciplina TCC – UASC/CEEI/UFCG**

**Trabalho aprovado em: 28 de Junho de 2023.**

**CAMPINA GRANDE - PB**

## RESUMO

Este artigo apresenta um estudo sobre a predição de resultados em partidas do jogo eletrônico *League of Legends (LoL)* utilizando técnicas de aprendizado de máquina. Com o objetivo de explorar a capacidade de prever resultados em tempo real, considerando diferentes variáveis e estágios da partida, destacamos o uso de dados inéditos como parte fundamental desse processo. Com o aumento da popularidade do *LoL* e a realização de torneios, surgiram também as apostas relacionadas ao jogo, tornando ainda mais relevante a investigação nessa área. Diversos modelos foram avaliados e os resultados foram encorajadores. O modelo baseado em *Random Forest* obteve o melhor desempenho, alcançando uma acurácia média de 81,57% em estágios intermediários da partida, quando a porcentagem de tempo decorrido estava entre 60% e 80%. Por outro lado, os modelos de Regressão Logística e *Gradient Boosting* mostraram-se mais eficazes em estágios iniciais do jogo, com resultados promissores. Esse estudo contribui para o campo de aprendizado de máquina aplicado a jogos eletrônicos, fornecendo *insights* valiosos sobre a predição em tempo real no *League of Legends*. Os resultados obtidos podem ser relevantes tanto para os jogadores que desejam aprimorar suas estratégias quanto para a indústria de apostas relacionada ao jogo.

# **LEAGUE OF LEGENDS: REAL-TIME RESULT PREDICTION**

## **ABSTRACT**

This paper presents a study on the prediction of outcomes in matches of the electronic game League of Legends (LoL) using machine learning techniques. With the aim of exploring the ability to predict real-time results, considering different variables and stages of the match, we highlight the use of unpublished data as a fundamental part of this process. With the increasing popularity of LoL and the emergence of tournaments, betting related to the game has also emerged, making the investigation in this area even more relevant. A variety of models were evaluated and the results were encouraging. A model based on Random Forest showed the best performance, achieving an average accuracy of 81.57% in intermediate stages of the match when the percentage of elapsed time was between 60% and 80%. On the other hand, the Logistic Regression and Gradient Boosting models proved to be more effective in early stages of the game, with promising results. This study contributes to the field of machine learning applied to electronic games, providing valuable insights into real-time prediction in League of Legends. The results obtained may be relevant for both players seeking to improve their strategies and the betting industry related to the game.

# League of Legends: Predição de Resultados em Tempo Real

Jailson Barros da Silva Junior  
Departamento de Sistemas e Computação  
Universidade Federal de Campina Grande  
Campina Grande, Paraíba - Brasil  
jailson.junior@ccc.ufcg.edu.br

Cláudio Campelo  
Departamento de Sistemas e Computação  
Universidade Federal de Campina Grande  
Campina Grande, Paraíba - Brasil  
campelo@computacao.ufcg.edu.br

## RESUMO

Este artigo apresenta um estudo sobre a predição de resultados em partidas do jogo eletrônico *League of Legends* (LoL) utilizando técnicas de aprendizado de máquina. Com o objetivo de explorar a capacidade de prever resultados em tempo real, considerando diferentes variáveis e estágios da partida, destacamos o uso de dados inéditos como parte fundamental desse processo. Com o aumento da popularidade do LoL e a realização de torneios, surgiram também as apostas relacionadas ao jogo, tornando ainda mais relevante a investigação nessa área. Diversos modelos foram avaliados e os resultados foram encorajadores. O modelo baseado em *Random Forest* obteve o melhor desempenho, alcançando uma acurácia média de 81,57% em estágios intermediários da partida, quando a porcentagem de tempo decorrido estava entre 60% e 80%. Por outro lado, os modelos de Regressão Logística e *Gradient Boosting* mostraram-se mais eficazes em estágios iniciais do jogo, com resultados promissores. Esse estudo contribui para o campo de aprendizado de máquina aplicado a jogos eletrônicos, fornecendo *insights* valiosos sobre a predição em tempo real no *League of Legends*. Os resultados obtidos podem ser relevantes tanto para os jogadores que desejam aprimorar suas estratégias quanto para a indústria de apostas relacionada ao jogo.

## PALAVRAS-CHAVE

Predição de Resultados, League of Legends, Aprendizado de Máquina, Modelos de Predição, Estratégias de Jogo, Apostas

## 1 INTRODUÇÃO

*League of Legends* (LoL) é um dos jogos eletrônicos mais populares do mundo atualmente. Existem mais de 150 milhões de jogadores registrados e em média mais de 100 milhões de jogadores ativos mensalmente e mais de 10 milhões diariamente [1]. *League of Legends* é um jogo do gênero *Multiplayer Online Battle Arena* (MOBA) desenvolvido pela *Riot Games* em 2009. No jogo, duas equipes se enfrentam com o objetivo de destruir as estruturas da equipe adversária, sendo o principal alvo uma estrutura chamada *Nexus*. As equipes são compostas por cinco jogadores, cada um controlando um campeão/personagem, e eles batalham em um mapa com três rotas e uma selva, com cada jogador desempenhando um papel específico na partida. Existem muitos objetivos e variáveis importantes no jogo que impactam diretamente a vitória, desde a destruição de

torres até a eliminação de dragões pelo mapa, gerando vantagens para a equipe na busca pela vitória.

Com a fama do jogo e o crescimento dos eSports [13, 5], surgiu o Campeonato Mundial de *League of Legends* (*League of Legends World Championship*), também conhecido como *Worlds*. O evento acontece anualmente e reúne centenas de jogadores de todo o mundo, com o torneio sendo realizado em diferentes cidades a cada ano. O primeiro torneio foi organizado pela *Riot Games* em 2011 e teve uma premiação de US\$100.000 [18]. Atualmente, o torneio está em sua 12ª edição, que ocorreu em setembro de 2022, com uma premiação total de US\$2.225.000 [15].

Com a popularidade dos torneios e as premiações em volta dele, surgiram as apostas e os bolões dentro da comunidade e até mesmo pela própria *Riot Games*. Quando o *Worlds* começa, a *Riot Games* cria um bolão para os jogadores de *League of Legends*, onde aqueles que acertam as previsões dos resultados das partidas recebem recompensas dentro do jogo. Além disso, existem sites de apostas que incluem categorias de eSports e permitem que os usuários apostem em partidas de *League of Legends*. Sites como *Betway* e *Rivalry* são exemplos de plataformas onde é possível fazer apostas em torneios de LoL.

Prever os resultados de uma partida de LoL pode ser bastante desafiador. Muitos fatores podem afetar o resultado final, como conclusão de objetivos, desde objetivos da selva como dragões e o Arauto até objetivos principais como a destruição de torres e inibidores. A vantagem que uma equipe tem sobre a outra em termos de ouro (a moeda do jogo), itens ou nível dos campeões também é levada em consideração. Todas essas variáveis podem mudar ao longo da partida, tornando ainda maior o desafio de prever o resultado final.

Este artigo apresenta uma abordagem baseada em aprendizado de máquina para a previsão em tempo real dos resultados de partidas de LoL com a maior precisão possível. Ênfase é dada ao uso de dados recentes e exclusivos como parte fundamental desse processo, fornecendo uma análise abrangente e atualizada do jogo. Além disso, tem como objetivo identificar a melhor técnica ou combinação delas, considerando o uso desses dados distintos. A eficácia da abordagem também é avaliada em diferentes momentos durante a partida, garantindo uma abordagem dinâmica e atualizada para a previsão de resultados no contexto de *League of Legends*.

O restante deste artigo está estruturado da seguinte forma. A Seção 2 apresenta informações de contexto. Em seguida, a Seção 3 discute os trabalhos relacionados. Na sequência, são descritos os dados e métodos utilizados na Seção 4. Posteriormente, a Seção 5 discute os resultados. Por fim, a Seção 6 conclui o artigo e aponta direções futuras.

Os autores retêm os direitos, ao abrigo de uma licença Creative Commons Atribuição CC BY, sobre todo o conteúdo deste artigo (incluindo todos os elementos que possam conter, tais como figuras, desenhos, tabelas), bem como sobre todos os materiais produzidos pelos autores que estejam relacionados ao trabalho relatado e que estejam referenciados no artigo (tais como códigos fonte e bases de dados). Essa licença permite que outros distribuam, adaptem e evoluam seu trabalho, mesmo comercialmente, desde que os autores sejam creditados pela criação original.

## 2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção, fornecemos um contexto que serve como base para entender as análises e modelos desenvolvidos neste estudo. Primeiramente, apresentamos uma explicação sobre o gênero *MOBA* (*Multiplayer Online Battle Arena*) e, em seguida, aprofundamos na compreensão do jogo *League of Legends*, que é o foco central deste trabalho.

### 2.1 Gênero *MOBA*: *Multiplayer Online Battle Arena*

O gênero *MOBA*, ou *Multiplayer Online Battle Arena*, é um dos gêneros mais populares no cenário competitivo de jogos eletrônicos. Os jogos *MOBA* se destacam pela combinação de elementos de estratégia em tempo real e ação intensa, proporcionando uma experiência dinâmica e desafiadora para os jogadores. Nesse gênero, as partidas são disputadas entre duas equipes, cada uma composta por um grupo de jogadores que controlam um único personagem, conhecido como herói ou campeão. Cada herói possui habilidades únicas e desempenha um papel específico na equipe, como ataque, defesa ou suporte.

O objetivo central dos jogos *MOBA* é destruir a base inimiga enquanto defende a própria. As bases são protegidas por torres defensivas e unidades controladas pelo computador que tentam impedir o avanço da equipe adversária. Durante a partida, os jogadores devem trabalhar em conjunto, planejar estratégias, coordenar ataques e tomar decisões táticas para obter vantagens e superar os oponentes. A comunicação efetiva e a sincronização entre os membros da equipe são cruciais para o sucesso no jogo.

Além disso, os jogos *MOBA* apresentam elementos como progressão dos personagens ao longo da partida, coleta de recursos, a conquista de objetivos no mapa e tomada de decisões estratégicas. Esses elementos contribuem para a profundidade estratégica e a complexidade das partidas, tornando o gênero *MOBA* uma modalidade de jogo altamente competitiva e envolvente.

### 2.2 *League of Legends*

Entre os inúmeros títulos do gênero *MOBA*, o *League of Legends* (*LoL*) se destaca como um dos jogos eletrônicos mais populares e influentes da atualidade. Desenvolvido pela *Riot Games*, o *LoL* foi lançado em 2009 e conquistou uma enorme base de jogadores em todo o mundo.

No *League of Legends*, duas equipes compostas por cinco jogadores cada enfrentam-se em um mapa estrategicamente dividido em três rotas principais (superior, central e inferior) e uma área chamada selva. Cada jogador assume o controle de um campeão com habilidades únicas, responsável por desempenhar um papel específico dentro da equipe. Os campeões são selecionados no início da partida e possuem diversos estilos de jogo, como lutadores, magos, atiradores, tanques e suportes.

O principal objetivo no *LoL* é destruir o *Nexus* inimigo, uma estrutura localizada na base da equipe adversária. Para alcançar esse objetivo, os jogadores devem avançar pelas rotas, eliminar lacaios controlados pelo computador, destruir torres defensivas e objetivos neutros espalhados pelo mapa, como dragões e o Barão Nashor. Essas conquistas fornecem recursos e vantagens estratégicas, como

ouro, experiência e *buffs*, que podem ser utilizados para fortalecer a equipe e aumentar as chances de vitória.

O *League of Legends* se destaca pela variedade de campeões, estratégias de equipe, combinações de habilidades e complexidade estratégica. Cada partida é única, exigindo rápida adaptação, tomada de decisões táticas e trabalho em equipe eficiente. A profundidade do jogo e a constante evolução do *meta game* (estratégias dominantes) contribuem para sua longevidade e contínua popularidade no cenário dos eSports.

## 3 TRABALHOS RELACIONADOS

Esta seção analisa pesquisas relacionadas à predição de resultados em jogos eletrônicos, com foco em *MOBA* e *League of Legends*. Esses estudos exploram diferentes maneiras de prever quem vencerá as partidas e entender o que afeta o sucesso das equipes. Essa revisão nos dá uma ideia do que os pesquisadores têm estudado nesse campo.

### 3.1 Predição de Resultados em Jogos *MOBA*

Na área de predição de resultados em jogos *MOBA*, especificamente no contexto de *DOTA 2*, foram realizados estudos que contribuem para o avanço desse campo. Ke et al. [10] propõem uma abordagem inovadora para identificar e definir as lutas em equipe como eventos cruciais durante as partidas de *DOTA 2*. O principal objetivo desse estudo é explorar o potencial das informações das lutas em equipe na predição em tempo real dos resultados das partidas. Utilizando diferentes modelos de redes neurais recorrentes, o estudo obteve uma precisão de mais de 70% ao considerar todas as informações das lutas em equipe até os primeiros 32 minutos de cada partida.

Outro estudo realizado no contexto de *DOTA 2* investigou o uso de modelos de aprendizado de máquina, desde regressão linear supervisionada até modelos de aprendizado profundo, como Redes Neurais e *Long Short-Term Memory (LSTM)* [3]. Os resultados obtidos demonstraram que os modelos de aprendizado profundo alcançaram altas taxas de precisão, com médias de 82% para regressão linear e até 93% para modelos *LSTM*. Além disso, foram realizadas análises considerando diferentes estágios de predição futura e correlação temporal, revelando a importância desses aspectos na obtenção de resultados mais precisos.

Em um contexto diferente, no jogo *Heroes of the Storm (HOTS)*, Swidler [17] explorou a possibilidade de prever os resultados das partidas. Os jogadores têm a opção de enviar arquivos de replay para um site chamado *HOTSLogs* [9], que analisa os jogos e fornece estatísticas relevantes. Embora a *Blizzard Entertainment*, desenvolvedora do jogo, não forneça um histórico de partidas para download, o *HOTSLogs* oferece uma maneira de obter um resumo dos últimos 30 dias de jogos. Esses dados foram usados para desenvolver um modelo que prevê com precisão mais de 62% das partidas e estima a probabilidade de vitória de uma equipe em um determinado jogo.

### 3.2 Predição de Resultados de *League of Legends*

Os artigos analisados nesta seção abordam a predição dos resultados das partidas de *League of Legends* usando diferentes abordagens de aprendizado de máquina. Em comum, todos os estudos têm como objetivo identificar padrões e variáveis que possam influenciar os

resultados das partidas, a fim de desenvolver modelos de predição precisos.

Alguns trabalhos abordam a predição dos resultados das partidas de *League of Legends* utilizando redes neurais [12, 11]. Eles exploram diferentes conjuntos de características, como Dragões, Nível, Arautos do Vale, Torres, ouro, abates, assistências, torres destruídas e dragões. Esses estudos demonstram a capacidade das redes neurais em alcançar altas taxas de precisão na predição, variando de 75,1% a 93,75%.

Por outro lado, outros estudos utilizam algoritmos de árvore de decisão para prever os resultados [6,4]. Esses estudos consideram características como duração, abates, torres e ouro. Ambos obtiveram resultados promissores, com precisões médias de 80% e taxas de precisão variando de 68,33% a 85,17%.

Por sua vez, Silva et. al. [16] compara diferentes tipos de Redes Neurais Recorrentes (RNNs) na predição de resultados. Uma RNN simples alcançou maior precisão e foi selecionada para experimentos adicionais com diferentes intervalos de tempo. A precisão variou de 63,91% a 83,54%, dependendo do intervalo de tempo considerado. O estudo sugere o uso de RNNs para analisar *power spikes* e tomar decisões estratégicas no jogo.

Por sua vez, Silva et al. [16] comparam diferentes tipos de Redes Neurais Recorrentes (RNNs) na predição de resultados. Uma RNN simples obteve maior precisão e foi selecionada para experimentos adicionais com diferentes intervalos de tempo. A precisão variou de 63,91% a 83,54% dependendo do intervalo de tempo considerado. O estudo sugere o uso de RNNs para analisar *power spikes* e tomar decisões estratégicas no jogo.

Do et al. [7] destacam a importância da seleção de campeões e da experiência do jogador na predição dos resultados. Um modelo de aprendizado de máquina usando uma Rede Neural Profunda (DNN) obteve uma precisão de 75,1% ao considerar a experiência do jogador com o campeão como uma das características. O estudo enfatiza a necessidade de sistemas de matchmaking justos que levem em conta a seleção de campeões dos jogadores.

Em resumo, as abordagens existentes compartilham o objetivo de prever os resultados das partidas de *League of Legends*, mas diferem em suas abordagens e características consideradas. Redes neurais e algoritmos de árvore de decisão têm mostrado resultados promissores, enquanto a análise de intervalos de tempo e a experiência do jogador com o campeão também foram abordadas. Esses estudos fornecem *insights* valiosos para a comunidade de jogadores e desenvolvedores, estabelecendo uma base sólida para trabalhos futuros no campo de predição de resultados em jogos competitivos como *League of Legends*.

A Tabela 1 apresenta uma comparação dos trabalhos relacionados, destacando as técnicas utilizadas e os dados considerados em cada estudo.

## 4 METODOLOGIA

A seção de Metodologia descreve a abordagem adotada nesta pesquisa para a predição de jogos de *League of Legends*. É apresentado o conjunto de dados utilizado, as técnicas de aprendizado de máquina aplicadas e as métricas de avaliação utilizadas. O objetivo é explorar diferentes abordagens e fornecer previsões precisas para melhorar a compreensão e o desempenho nesse contexto de jogos competitivos.

Ref	Técnicas					Dados		
	DNN	Árvores de Decisão	RNN	Regressão Logística	RF	Gradient Boosting	Pre-Game	In-Game
[12]	X							X
[6]		X			X			X
[11]				X		X	X	X
[16]			X				X	X
[7]	X						X	X
[4]		X						X
[8]	X							X

**Tabela 1: Tabela Comparativa de Trabalhos Relacionados (Técnicas e Dados)**

Esta seção de Metodologia descreve a abordagem adotada nesta pesquisa para a predição de jogos de *League of Legends*. É apresentado o conjunto de dados utilizado, as técnicas de aprendizado de máquina aplicadas e as métricas de avaliação empregadas. O objetivo é explorar diferentes abordagens e fornecer previsões precisas para aprimorar a compreensão e o desempenho nesse contexto de jogos competitivos.

### 4.1 Coleta dos Dados

Os dados foram coletados utilizando a API oficial da *Riot Games* [2], desenvolvedora do jogo *League of Legends*. A API fornece acesso a uma variedade de informações relacionadas ao jogo, como estatísticas das partidas, detalhes dos jogadores, resultados e outros dados relevantes para o estudo. Através da API, foi possível obter dados atualizados e precisos para análise e modelagem.

O processo de coleta de dados envolveu o uso da biblioteca *riotwatcher* [14], uma biblioteca em Python desenvolvida especificamente para interagir com a API oficial da *Riot Games*. Essa biblioteca facilitou o envio de requisições e a manipulação dos dados retornados pela API, agilizando o processo de coleta.

É importante destacar que a coleta de dados por meio da API da *Riot Games*, com o auxílio da biblioteca *riotwatcher*, garantiu a confiabilidade e integridade dos dados, uma vez que as informações foram obtidas diretamente da fonte oficial do jogo. Além disso, todas as políticas de privacidade e termos de uso da API foram estritamente seguidos durante o processo de coleta.

### 4.2 Conjunto de dados

O conjunto de dados utilizado neste estudo representa uma amostra representativa de partidas ranqueadas no *League of Legends*, abrangendo jogadores de diferentes níveis de habilidade, que vão desde o



elo Ferro até o Diamante. As 64.556 partidas coletadas fornecem uma visão abrangente e diversificada do cenário competitivo durante o período de 12/01/2023 a 18/05/2023.

Cada partida registrada no conjunto de dados contém uma riqueza de informações valiosas para análise e modelagem. Entre as variáveis disponíveis, temos informações sobre abates de campeões, incluindo o número de abates realizados pelas equipes azul e vermelha, o primeiro abate que ocorreu em cada partida, bem como o desempenho dos jogadores em assegurar abates e assistências.

Além disso, temos dados sobre a conquista de objetivos importantes durante as partidas, como eliminação de dragões, torres e inibidores. Esses elementos estratégicos são cruciais para determinar a vantagem de uma equipe e sua capacidade de controlar o mapa do jogo.

Outras variáveis relevantes incluem o total de ouro adquirido por cada equipe, o número de minions abatidos tanto nas rotas quanto na selva, o nível médio dos jogadores e muitas outras métricas que nos permitem compreender melhor a dinâmica das partidas e os fatores que podem influenciar o resultado final.

Ao explorar este conjunto de dados diversificado, busca-se identificar padrões, tendências e correlações entre as variáveis, a fim de desenvolver modelos de previsão mais precisos e robustos para prever os resultados das partidas do *League of Legends*. Essa análise detalhada do conjunto de dados é essencial para compreender o jogo e desenvolver estratégias mais eficazes.

### 4.3 Pré-processamento dos dados

Durante o processo de pré-processamento dos dados, foram adotadas várias abordagens para garantir a qualidade e consistência do conjunto de dados. Um dos primeiros passos envolveu a conversão dos valores de tempo, expressos em milissegundos, para minutos. Essa transformação permitiu uma melhor compreensão da duração da partida e facilitou a análise dos eventos em diferentes momentos do jogo.

Posteriormente, o conjunto de dados foi dividido em quatro partes, representando diferentes percentuais da duração total das partidas: 20%, 40%, 60% e 80%. Essa divisão estratégica permitiu uma análise mais detalhada dos acontecimentos em diferentes estágios das partidas, possibilitando a identificação de padrões e comportamentos específicos em cada intervalo de tempo.

O conjunto de dados foi então dividido em quatro intervalos de tempo, representando diferentes percentagens da duração total da partida: 20%, 40%, 60% e 80%. Essa divisão estratégica foi realizada para examinar o desempenho do modelo ao lidar com diferentes quantidades de informações disponíveis sobre o jogo. Embora a intuição nos leve a esperar que o modelo faça previsões melhores à medida que mais informações são fornecidas, é importante destacar que o desempenho do modelo nem sempre apresenta um crescimento linear à medida que a quantidade de informações aumenta. Essa abordagem analítica nos permite explorar como o modelo se comporta em diferentes contextos temporais.

Para garantir a confiabilidade e a qualidade dos dados, foram tomadas medidas para lidar com possíveis problemas de qualidade dos dados. Como parte desse processo, as partidas com duração inferior a 5 minutos foram excluídas do conjunto de dados. Essa

decisão foi tomada para eliminar instâncias que provavelmente são reinícios de partida (remakes) em vez de partidas reais.

Outro passo importante foi a transformação de variáveis booleanas em valores numéricos binários, atribuindo o valor 0 para “falso” e o valor 1 para “verdadeiro”. Essa abordagem simplifica a representação dos dados e permite a aplicação direta de técnicas de modelagem e análise numérica.

Além disso, variáveis redundantes ou opostas foram removidas. Por exemplo, as variáveis *redWin* e *redFirstBlood* foram removidas, pois fornecem informações semelhantes às demais variáveis e não acrescentam valor adicional à análise. Essa remoção de variáveis redundantes resultou em um conjunto de dados mais conciso, focado nas informações mais relevantes para o estudo.

### 4.4 Análise Exploratória dos Dados

Nesta subseção, realizaremos uma análise exploratória dos dados para compreender as características gerais, identificar padrões e tendências e extrair *insights* relevantes para nossa análise. Utilizaremos técnicas de visualização e estatísticas descritivas para explorar os dados em detalhes e obter uma visão abrangente do conjunto de dados.

Agora apresentamos uma visão geral concisa das principais descobertas e observações derivadas de nossa Análise Exploratória de Dados.

Durante a análise exploratória dos dados, observou-se que a taxa de vitória é de 50,5% para a equipe azul e 49,5% para a equipe vermelha. Isso indica que não há um desequilíbrio significativo entre as equipes, com um resultado relativamente equilibrado.

Em relação à duração das partidas, a mais longa registrada no conjunto de dados é de 67,35 minutos, enquanto a mediana da duração das partidas é de 30,01 minutos. Isso sugere que a maioria das partidas tem uma duração em torno dessa marca.

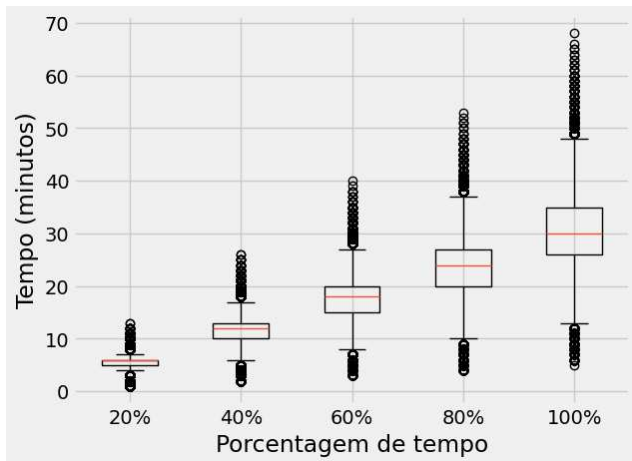
A Figura 1 apresenta a distribuição das durações das partidas de *LoL*, em minutos, tanto para a partida completa (100%), quanto para cada uma das porções consideradas nas análises preditivas (20%, 40%, 60%, 80%). Esses valores nos fornecem *insights* sobre a progressão do tempo de jogo ao longo das partidas.

Ao analisar as correlações entre as variáveis e a vitória da equipe azul em diferentes percentagens de tempo, encontramos os seguintes resultados nas Tabelas 2, 3, 4 e 5.

Variáveis Positivas	Correlação	Variáveis Negativas	Correlação
blueFirstBlood	0,14	redFirstBlood	-0,15
blueChampionKill	0,10	redChampionKi	-0,12
blueTotalGold	0,04	redTotalGold	-0,07

**Tabela 2: Correlações entre Variáveis e Vitória do Time Azul - 20% das Partidas**

Ao analisar as Tabelas 2, 3, 4 e 5, podemos observar que as correlações entre as variáveis e a vitória da equipe azul variam de acordo com o tempo decorrido na partida. Nos estágios iniciais, como mostrado na Tabela 2, o desempenho dos campeões da equipe azul em abates e a conquista do primeiro abate são relevantes. Conforme a partida avança, as Tabelas 3, 4 e 5 demonstram que



**Figura 1: Variação dos Tempos de Partidas de LoL nos Diferentes Intervalos de Tempo**

Variáveis Positivas	Correlação	Variáveis Negativas	Correlação
blueChampionKill	0,16	redChampionKill	-0,18
blueFirstBlood	0,16	redDragonKill	-0,17
blueDragonKill	0,16	redFirstBlood	-0,16

**Tabela 3: Correlações entre Variáveis e Vitória do Time Azul - 40% das Partidas**

Variáveis Positivas	Correlação	Variáveis Negativas	Correlação
blueDragonKill	0,29	redDragonKill	-0,31
blueChampionKill	0,22	redChampionKill	-0,24
blueFirstBlood	0,16	redTowerKill	-0,19

**Tabela 4: Correlações entre Variáveis e Vitória do Time Azul - 60% das Partidas**

Variáveis Positivas	Correlação	Variáveis Negativas	Correlação
blueDragonKill	0,43	redDragonKill	-0,45
blueChampionKill	0,29	redTowerKill	-0,38
blueTowerKill	0,27	redChampionKill	-0,31

**Tabela 5: Correlações entre Variáveis e Vitória do Time Azul - 80% das Partidas**

a conquista de Dragões e a destruição de torres pela equipe azul também se tornam significativas. Por outro lado, o desempenho da equipe vermelha em abates, a conquista de Dragões e a destruição de torres pela equipe vermelha têm um impacto negativo na vitória da equipe azul. Esses *insights* nos ajudam a compreender quais aspectos são mais determinantes para o sucesso da equipe azul em diferentes momentos do jogo.

## 4.5 Seleção de Características

Ao analisar a correlação entre as variáveis e a vitória da equipe azul, observamos que diferentes variáveis apresentaram diferentes níveis de correlação em diferentes intervalos de tempo das partidas. De maneira geral, as variáveis *blueChampionKill*, *blueDragonKill* e *blueFirstBlood* mostraram uma correlação positiva mais forte com a vitória da equipe azul em quase todos os intervalos de tempo analisados. Isso sugere que ter um maior número de abates de campeões, objetivos relacionados aos dragões e garantir o primeiro abate estão associados a uma maior probabilidade de vitória para a equipe azul.

No entanto, é importante destacar que correlação não implica causalidade direta. Embora essas variáveis apresentem uma correlação significativa, outros fatores podem influenciar o resultado da partida. Portanto, é necessário considerar essas variáveis em conjunto com outras informações relevantes para construir modelos de previsão mais precisos.

Adicionalmente, durante a análise dos diferentes intervalos de tempo, observamos algumas diferenças nas correlações das variáveis. Por exemplo, no cenário de 40% das partidas, o abate inicial (*firstBlood*), o número de abates e a conquista de dragões apresentaram correlações iguais com a vitória da equipe azul, indicando a importância desses eventos no desempenho da equipe.

No intervalo de 60% das partidas, as variáveis relacionadas aos abates de campeões, dragões e ao primeiro abate mostraram correlações mais significativas com a vitória da equipe azul. Isso pode sugerir que esses objetivos são cruciais para o sucesso da equipe nessa etapa do jogo.

No intervalo de 80% das partidas, observamos um fortalecimento das correlações entre abates de campeões, dragões e torres conquistadas pela equipe azul. Esses resultados podem indicar que em partidas mais avançadas, esses aspectos se tornam ainda mais determinantes para a vitória.

Portanto, considerando essas análises de correlação em diferentes intervalos de tempo, podemos concluir que certas variáveis possuem uma relação mais significativa com a vitória da equipe azul em diferentes estágios das partidas. Essas informações são valiosas para a construção de modelos de previsão que levem em conta essas características específicas para prever o resultado das partidas de *League of Legends* com maior precisão.

## 4.6 Construção e Avaliação dos Modelos de Predição

Nesta subseção, descrevemos a metodologia adotada para a construção e avaliação dos modelos de previsão para as partidas de *League of Legends*, levando em consideração a porcentagem de tempo decorrido nas partidas. Além dos algoritmos de aprendizado de máquina mencionados anteriormente, consideramos essa nova variável como um atributo importante para melhorar a precisão dos modelos.

Inicialmente, dividimos o conjunto de dados em conjuntos de treinamento e teste, seguindo a mesma abordagem mencionada anteriormente. No entanto, desta vez estamos incluindo a porcentagem de tempo decorrido como um atributo adicional para treinar e avaliar os modelos.

Utilizamos os seguintes algoritmos de aprendizado de máquina: Regressão Logística, *Random Forest*, *Naive Bayes*, *Gradient Boosting*, Rede Neural e *Bagging*. Esses modelos foram selecionados devido à sua capacidade de lidar com problemas de classificação e por serem amplamente utilizados em estudos de predição.

Para avaliar os modelos, utilizamos a técnica de validação cruzada (*cross validation*) para garantir uma avaliação robusta, considerando a porcentagem de tempo decorrido como parte dos dados de entrada. Isso nos permite capturar padrões e comportamentos ao longo do tempo e melhorar a capacidade preditiva dos modelos.

Durante a avaliação dos modelos, estamos aplicando métricas como acurácia, precisão, *recall*, *F1-score* e *AUC-ROC*. Além disso, estamos considerando o desempenho e a capacidade preditiva dos modelos levando em conta a porcentagem de tempo decorrido.

Ao comparar os resultados obtidos pelos diferentes modelos, estamos considerando não apenas a acurácia e as métricas tradicionais, mas também a capacidade dos modelos de capturar e interpretar as mudanças ao longo do tempo. Isso nos permite identificar o modelo mais adequado para prever os resultados das partidas de *League of Legends*, considerando a influência da porcentagem de tempo decorrido.

Ao incluir a porcentagem de tempo decorrido como um atributo adicional, obtemos *insights* mais precisos e relevantes para prever os resultados das partidas. Estamos combinando algoritmos de aprendizado de máquina, técnica de validação cruzada, métricas de avaliação e a consideração da porcentagem de tempo decorrido para construir modelos de previsão mais eficazes e confiáveis.

## 5 RESULTADOS E DISCUSSÃO

Durante a análise das correlações entre as variáveis e a vitória da equipe azul, observamos que essas correlações variam ao longo da duração da partida. Nas fases iniciais, os abates realizados pelos campeões da equipe azul e a conquista do primeiro abate estavam fortemente correlacionados com a vitória. Conforme a partida avançava, a obtenção de Dragões e a destruição de torres pela equipe azul também mostravam uma correlação positiva com o resultado.

Em relação à importância das variáveis no modelo de *Random Forest*, observamos que ela varia ao longo do tempo na partida. Especificamente, em 20% do tempo decorrido, a variável mais importante para prever a vitória da equipe azul foi “redTotalGold”. Isso significa que o ouro total obtido pela equipe vermelha desempenhou um papel crucial na determinação do resultado nessa fase inicial da partida. Essa observação é evidenciada na Figura 2, onde podemos ver o gráfico das 12 principais variáveis com as maiores importâncias nesse cenário.

Nos intervalos de tempo de 40%, 60% e 80%, a importância das características sofreu mudanças. Em 40% do tempo, a variável “blueTotalGold” se tornou a mais relevante, como ilustrado na Figura 3. Isso indica que o ouro total adquirido pela equipe azul foi um fator determinante para sua vitória nessa fase intermediária da partida. Em 60% do tempo, pôde-se observar que a variável “blueTotalGold” permaneceu a mais importante, como mostrado na Figura 4, reforçando sua relevância nessa etapa do jogo. Para 80% do tempo, a variável “blueChampionKill” apresentou a maior importância,

como mostrado na Figura 5, enfatizando a importância dos abates executados pelos campeões da equipe azul nessa etapa da partida.

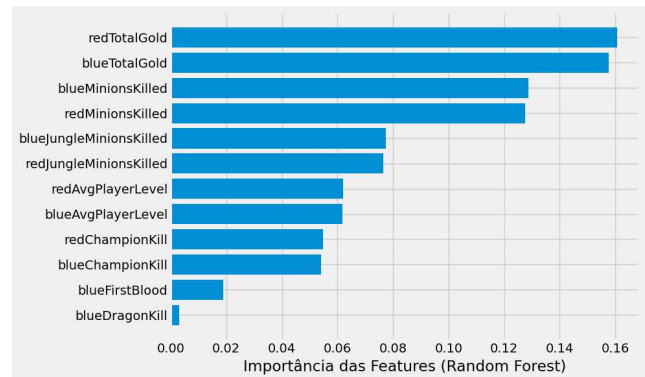


Figura 2: Gráfico de Importância das Características em 20% do Tempo

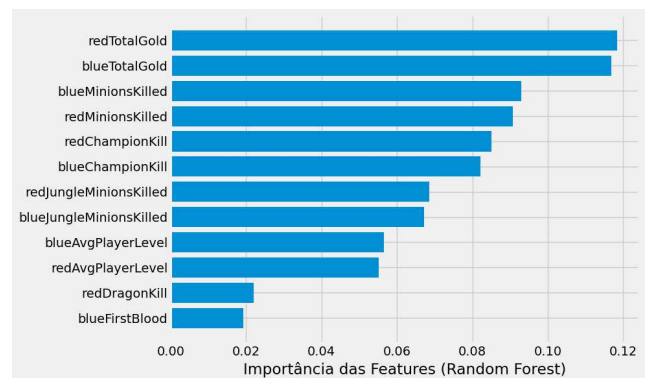


Figura 3: Gráfico de Importância das Características em 40% do Tempo

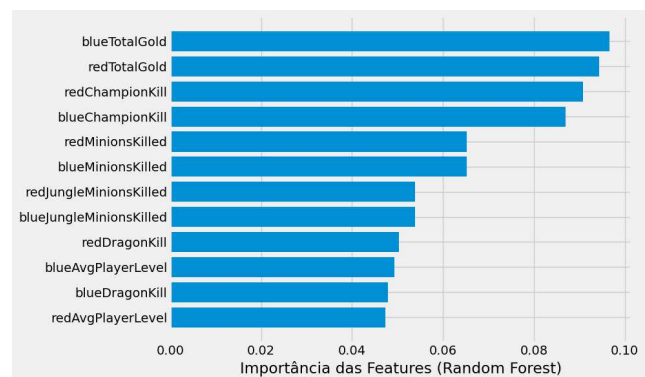
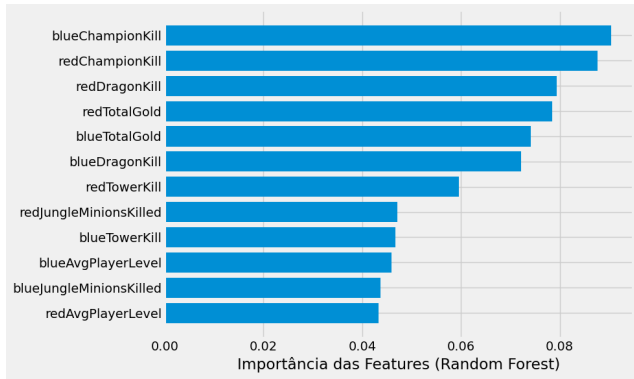


Figura 4: Gráfico de Importância das Características em 60% do Tempo



**Figura 5: Gráfico de Importância das Características em 80% do Tempo**

Esses resultados destacam como as características relevantes para a vitória da equipe azul podem variar ao longo do tempo. Portanto, é essencial considerar as diferentes fases da partida ao construir modelos de previsão e desenvolver estratégias para ter sucesso em momentos distintos do jogo. Além disso, ao comparar os resultados das correlações com a importância das variáveis obtida por meio do *Random Forest*, podem ser observadas diferenças significativas. Enquanto as correlações fornecem informações sobre a relação direta entre as variáveis e a vitória, a importância das características indica o grau de contribuição de cada variável para a precisão do modelo de previsão.

É interessante observar que, em alguns casos, uma variável pode ter uma correlação forte com a vitória da equipe azul, mas uma importância relativamente baixa de acordo com o *Random Forest*. Isso pode ser explicado pelo fato de que o *Random Forest* avalia a importância de uma variável considerando não apenas sua relação direta com a variável-alvo, mas também sua interação com outras variáveis e sua capacidade de reduzir a incerteza do modelo.

Portanto, é importante considerar tanto as correlações quanto a importância das características ao interpretar os resultados. Enquanto as correlações ajudam a identificar as relações diretas entre as variáveis e a vitória, a importância das características fornece *insights* sobre o impacto relativo de cada variável no desempenho geral do modelo de previsão. Essa combinação de informações nos permite obter uma compreensão mais abrangente e precisa dos fatores que influenciam o resultado das partidas de *League of Legends*.

Com base nos resultados obtidos pela aplicação dos modelos de previsão construídos, avaliamos o desempenho de cada um deles considerando a porcentagem de tempo decorrido nas partidas de *League of Legends*. As métricas de avaliação utilizadas foram acurácia, precisão, *recall*, *F1-score* e *AUC-ROC*.

A Tabela 6 apresenta as métricas de avaliação para cada modelo em diferentes porcentagens de tempo decorrido.

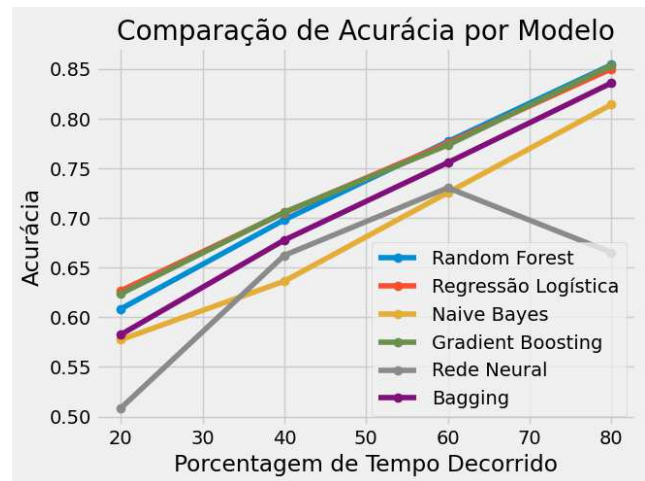
Ao comparar os resultados entre os modelos, observamos que o *Random Forest* apresentou a melhor performance em todas as métricas de avaliação quando consideramos uma porcentagem de tempo decorrido entre 60% e 80%. No entanto, é importante ressaltar que para uma porcentagem de tempo decorrido de 20%, outros modelos como a Regressão Logística e o *Gradient Boosting* tiveram resultados

Modelo	Porcentagem de Tempo	Acurácia	Precisão	Recall	F1-score	AUC-ROC
Random Forest	20%	0,6085	0,6150	0,5979	0,6063	0,6086
	40%	0,6982	0,7027	0,6957	0,6992	0,6982
	60%	0,7770	0,7836	0,7706	0,7770	0,7771
	80%	0,8543	0,8615	0,8472	0,8543	0,8543
Regressão Logística	20%	0,6268	0,6266	0,6431	0,6348	0,6267
	40%	0,7053	0,7027	0,7204	0,7114	0,7052
	60%	0,7755	0,7786	0,7752	0,7769	0,7755
	80%	0,8502	0,8572	0,8435	0,8503	0,8503
Naive Bayes	20%	0,5774	0,5591	0,7662	0,6464	0,5758
	40%	0,6364	0,6299	0,6762	0,6522	0,6361
	60%	0,7256	0,7228	0,7394	0,7310	0,7255
	80%	0,8143	0,8142	0,8186	0,8164	0,8143
Gradient Boosting	20%	0,5937	0,6112	0,5265	0,5660	0,5936
	40%	0,6772	0,6832	0,6633	0,6731	0,6771
	60%	0,7563	0,7648	0,7461	0,7553	0,7562
	80%	0,8304	0,8371	0,8222	0,8296	0,8304
Rede Neural	20%	0,5086	0,5064	0,9973	0,6717	0,5044
	40%	0,6624	0,8101	0,4315	0,5631	0,6643
	60%	0,7304	0,6653	0,9361	0,7778	0,7286
	80%	0,6652	0,9942	0,3380	0,5045	0,6680
Bagging	20%	0,5826	0,6021	0,5077	0,5508	0,5832
	40%	0,6778	0,7031	0,6249	0,6617	0,6783
	60%	0,7558	0,7837	0,7122	0,7463	0,7562
	80%	0,8360	0,8616	0,8038	0,8317	0,8363

**Tabela 6: Métricas de Avaliação dos Modelos - Porcentagem de Tempo Decorrido nas Partidas**

superiores. Em geral, os modelos tiveram um desempenho melhor ao considerar a porcentagem de tempo decorrido como atributo adicional em seus processos de aprendizado.

A Figura 6 exibe a acurácia dos modelos em diferentes porcentagens de tempo decorrido.



**Figura 6: Gráfico de Acurácia dos Modelos**

Podemos observar uma tendência de aumento da acurácia dos modelos à medida que a porcentagem de tempo decorrido aumenta. Isso indica que as informações sobre o tempo decorrido nas partidas são relevantes para prever os resultados.

## 6 CONCLUSÕES E TRABALHOS FUTUROS

Com base nos resultados obtidos e nas análises realizadas, podemos concluir que incluir a porcentagem de tempo decorrido nas partidas de *League of Legends* como um atributo adicional tem um impacto significativo no desempenho dos modelos de previsão. Os modelos apresentaram melhoria geral em suas métricas de avaliação ao considerar essa informação temporal.

O modelo *Random Forest* mostrou-se o mais eficaz em termos de desempenho, especialmente quando a porcentagem de tempo decorrido estava entre 60% e 80%. No entanto, nas fases iniciais do jogo (porcentagem de tempo decorrido de 20%), outros modelos como Regressão Logística e *Gradient Boosting* tiveram resultados superiores. Isso indica que diferentes modelos podem ser mais adequados dependendo da fase da partida.

Considerando uma acurácia de 60%, podemos afirmar que há potencial para fazer previsões em jogos que tenham passado apenas 20% do tempo. No entanto, é importante observar que prever resultados em partidas de *League of Legends* ainda apresenta desafios significativos devido à natureza dinâmica do jogo. Portanto, é necessário considerar outros fatores além da acurácia do modelo, como a análise de outras métricas de desempenho e a compreensão das limitações do estudo.

Os *insights* obtidos revelaram a importância de diferentes variáveis ao longo do tempo. Variáveis relacionadas a eliminações de campeões, dragões e torres conquistadas tiveram um impacto maior nas previsões nas fases iniciais, enquanto variáveis como ouro total e first blood se tornaram mais relevantes à medida que o tempo decorrido aumentou. Essa evolução do jogo destaca a importância de considerar diferentes estágios das partidas ao construir modelos de previsão.

Embora os resultados sejam promissores, ainda há espaço para melhorias. A natureza dinâmica e complexa do *League of Legends* apresenta desafios significativos para a modelagem e previsão de resultados. Portanto, é crucial considerar outras variáveis e informações adicionais, além de explorar novas abordagens e conjuntos de dados, para aprimorar ainda mais os modelos de previsão.

Uma perspectiva interessante para trabalhos futuros é abordar o problema de prever os resultados das partidas de *League of Legends* como uma série temporal. Essa abordagem pode permitir capturar padrões emergentes e tendências no jogo, levando em consideração a natureza dinâmica e em constante evolução das partidas.

## REFERÊNCIAS

- [1] [n. d.]. ActivePlayer - League of Legends. <https://activeplayer.io/league-of-legends/>. Acesso em: 16 de maio de 2023.
- [2] [n. d.]. API oficial da Riot Games. <https://developer.riotgames.com/>. Acesso em: 20 de abril de 2023.
- [3] Kodirjon Akhmedov and A. Phan. 2021. Machine learning models for DOTA 2 outcomes prediction. (2021). <https://arxiv.org/abs/2106.01782>
- [4] Farnod Bahrololloomi, Fabio Klonowski, Sebastian Sauer, R. Horst, and R. Dörner. 2023. E-Sports Player Performance Metrics for Predicting the Outcome of League of Legends Matches Considering Player Roles. *SN Computer Science* (2023). <https://doi.org/10.1007/s42979-022-01660-6>
- [5] Sebastian Block and Florian Haack. 2021. eSports: a new industry. In *SHS Web of Conferences*, Vol. 92. EDP Sciences, 04002.
- [6] Alexandre C. S. Cruz, T. G. Do Rêgo, Telmo De M. Filho, and Yuri Malheiros. 2021. League of Legends: An Application of Classification Algorithms to Verify the Prediction Importance of Main In-Game Variables. *Anais Estendidos do XX Simpósio Brasileiro de Games e Entretenimento Digital (SBGames Estendido 2021)* (2021). [https://doi.org/10.5753/sbgames\\_estendido.2021.19662](https://doi.org/10.5753/sbgames_estendido.2021.19662)
- [7] Tiffany D. Do, Seong Ioi Wang, Dylan S. Yu, Matthew G. McMillian, Ryan P. McMahan, and Ryan P. McMahan. 2021. Using Machine Learning to Predict Game Outcomes Based on Player-Champion Experience in League of Legends. *arXiv: Learning* (2021). <https://doi.org/10.1145/3472538.3472579>
- [8] Raluca D. Gaina. 2018. League of Legends: A Study of Early Game Impact. (2018).
- [9] HOTSLogs. [n. d.]. HOTSLogs. Disponível em: <https://www.hotlogs.com/default>. Acesso em: 08 de junho de 2023.
- [10] Cheng Hao Ke, Haozhang Deng, Congda Xu, Jiong Li, Xingyun Gu, Borchuluun Yadamsuren, Diego Klabjan, Rafet Sifa, Anders Drachen, and Simon Demediuk. 2022. DOTA 2 match prediction through deep learning team fight models. *2022 IEEE Conference on Games (CoG)* (2022). <https://doi.org/10.1109/cog51982.2022.9893647>
- [11] Lucas Lin. 2016. League of Legends Match Outcome Prediction. (2016).
- [12] Si-Jae No, Yoo-Jin Moon, and Young-Ho Hwang. 2021. Quantitative Analysis for Win/Loss Prediction of 'League of Legends' Utilizing the Deep Neural Network System through Big Data. *Journal of the Korea Society of Computer and Information* (2021). <https://doi.org/10.9708/jksci.2021.26.04.213>
- [13] Anders Hval Olsen. 2015. The evolution of eSports: An analysis of its origin and a look at its prospective future growth as enhanced by Information Technology Management tools. *arXiv preprint arXiv:1509.08795* (2015).
- [14] pseudonym117. [n. d.]. Riot-Watcher. <https://github.com/pseudonym117/Riot-Watcher>. Acesso em: 20 de abril de 2023.
- [15] Showmetech. [n. d.]. Worlds 2022: Mundial de LoL promete ser único. <https://www.showmetech.com.br/worlds-2022-mundial-de-lol-prometer-ser-unico/>. Acesso em: 16 de maio de 2023.
- [16] Antonio Luis Cardoso Silva, G. Pappa, and L. Chaimowicz. 2018. Continuous Outcome Prediction of League of Legends Competitive Matches Using Recurrent Neural Networks. (2018).
- [17] Jesse Swidler. 2017. Predicting the outcome of online Heroes of the Storm Matches. (2017). <https://github.com/jswidler/HOTSNet/blob/master/paper/HOTSNetPaper.pdf>
- [18] Wikipedia. [n. d.]. Campeonato Mundial de League of Legends. [https://pt.wikipedia.org/wiki/Campeonato\\_Mundial\\_de\\_League\\_of\\_Legends](https://pt.wikipedia.org/wiki/Campeonato_Mundial_de_League_of_Legends). Acesso em: 16 de maio de 2023.