

# CLAIRVOYANCE

## Predizendo resultados em League of Legends

João F. B. S. Martins

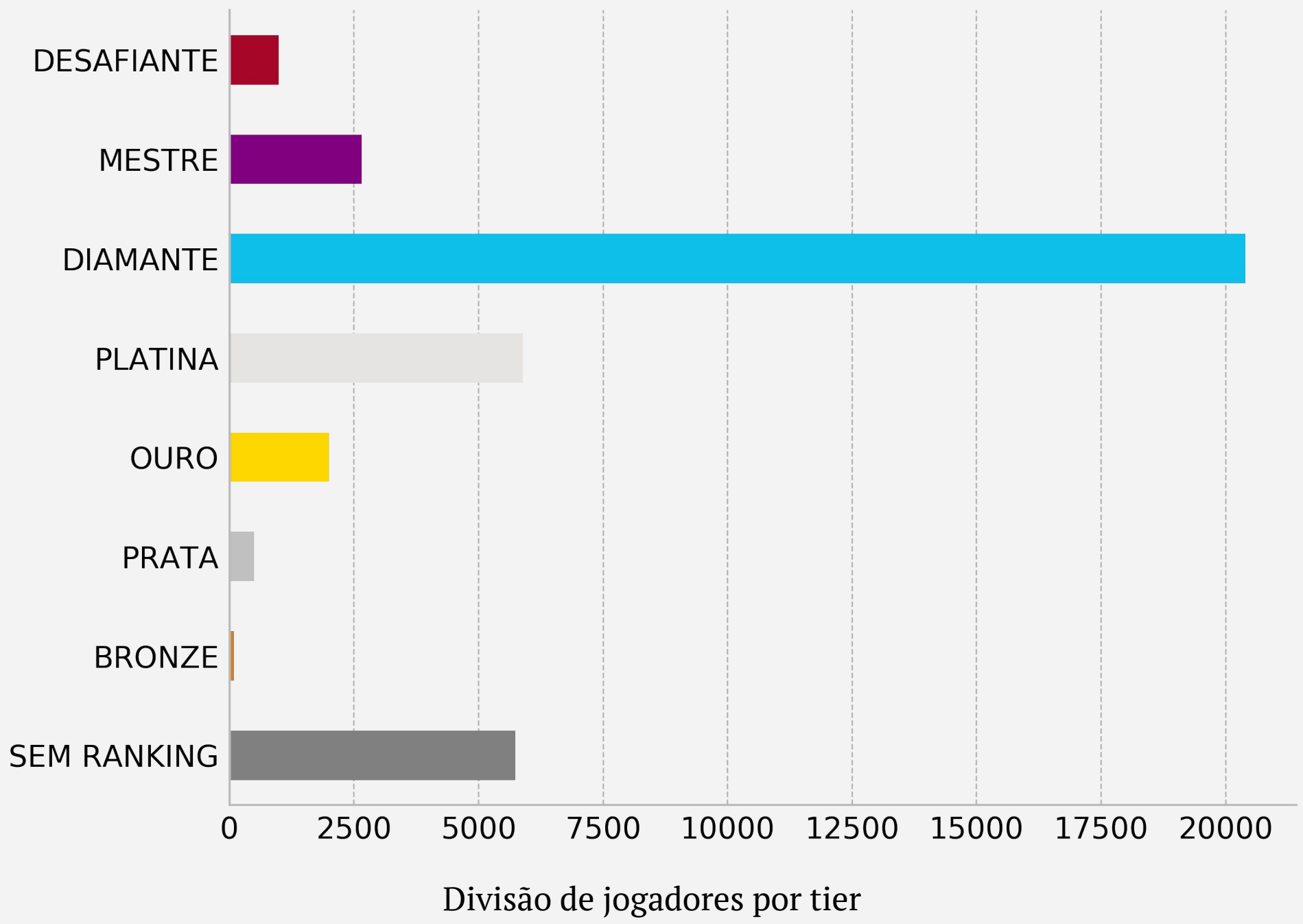
### Motivação

Atualmente League of Legends é o jogo mais jogado do planeta com mais de 100 milhões de jogadores ativos mensalmente. Além de uma comunidade casual muito grande, o jogo também conta com um cenário competitivo muito desenvolvido, tendo distribuído mais de 5 milhões de dólares em prêmios no último campeonato mundial. Este projeto foi desenvolvido com o objetivo de fornecer a essa comunidade uma ferramenta de predição de resultados em tempo real, utilizando dados pregame.

### Dados

Foi utilizado um wrapper da API da Riot Games, desenvolvido em Python e nomeado Cassiopeia, para coleta dos dados. Foram coletadas um total de 40228 partidas da season 2017, distribuídas por 4 regiões do mundo: América do Norte, Brasil, Europa Ocidental e Coreia do Sul.

A coleta dos dados foi feita em “teia” utilizando o tier Mestre como seed inicial para priorizar jogadores mais habilidosos.



### Metodologia

Com os dados coletados foram modelados vários datasets utilizando tanto features pregame como in-game. Ao agregar por time e fazer a diferença deles foram criados alguns atributos adicionais para aumentar a acurácia das predições. Alguns dos atributos pregame, como a escolha dos campeões, tiveram que ser representados em one-hot encoding.

As features de dados in-game representam o acumulado dos mesmos em intervalos de 10 minutos. São eles(por minuto e por jogador): creeps mortos, dano recebido, ganho de ouro, ganho de experiência.

Devido à natureza tabular dos dados, os modelos de aprendizado utilizados foram todos baseados em árvores de decisão. Foram escolhidos dois modelos de boosting (AdaBoost e XGBoost) e um de bagging (Random Forests).

### Análise Experimental

O uso das features de histórico requer muita coleta e pré-processamento, portanto foram criados dois datasets unificados dos dados pregame, um que as leva em conta e outro que não. Para ajuste dos hiperparâmetros dos modelos foi então executado um gridsearch sobre os datasets criados.

Atributos por base de dados	AdaBoost	XGBoost	Random Forests
Campeões por time	52.41% (0.46%)	52.89% (0.30%)	51.11% (0.63%)
Feitiços de invocador por time	51.24% (0.37%)	51.14% (0.23%)	50.61% (0.41%)
Talentos de invocador por time	52.37% (0.48%)	52.53% (0.15%)	50.61% (0.43%)
Tipos/capacidade de dano por campeão	51.01% (0.38%)	51.39% (0.25%)	50.31% (0.55%)
Soma das maestrias de campeão por invocador	51.38% (0.30%)	51.37% (0.33%)	50.29% (0.68%)
Maestria de invocador com campeão escolhido	58.78% (0.72%)	58.63% (0.62%)	54.68% (0.58%)
Vitórias por invocador	66.46% (2.52%)	66.91% (2.59%)	62.41% (2.53%)
Vitórias para campeão escolhido por invocador	89.08% (1.93%)	89.29% (1.99%)	87.74% (2.12%)
Acurácia média	59.07%	59.23%	57.18%

Acurácia dos modelos para dados pregame isolados

Dataset	AdaBoost	XGBoost	Random Forests
Exceto histórico(pré-ajuste)	59.44%	<b>59.77%</b>	54.58%
Exceto histórico(ajuste)	60.62%	61.42%	<b>64.80%</b>
Todos os atributos(pré-ajuste)	89.08%	<b>89.29%</b>	87.68%
Todos os atributos(ajuste)	89.41%	<b>89.5%</b>	89.1%

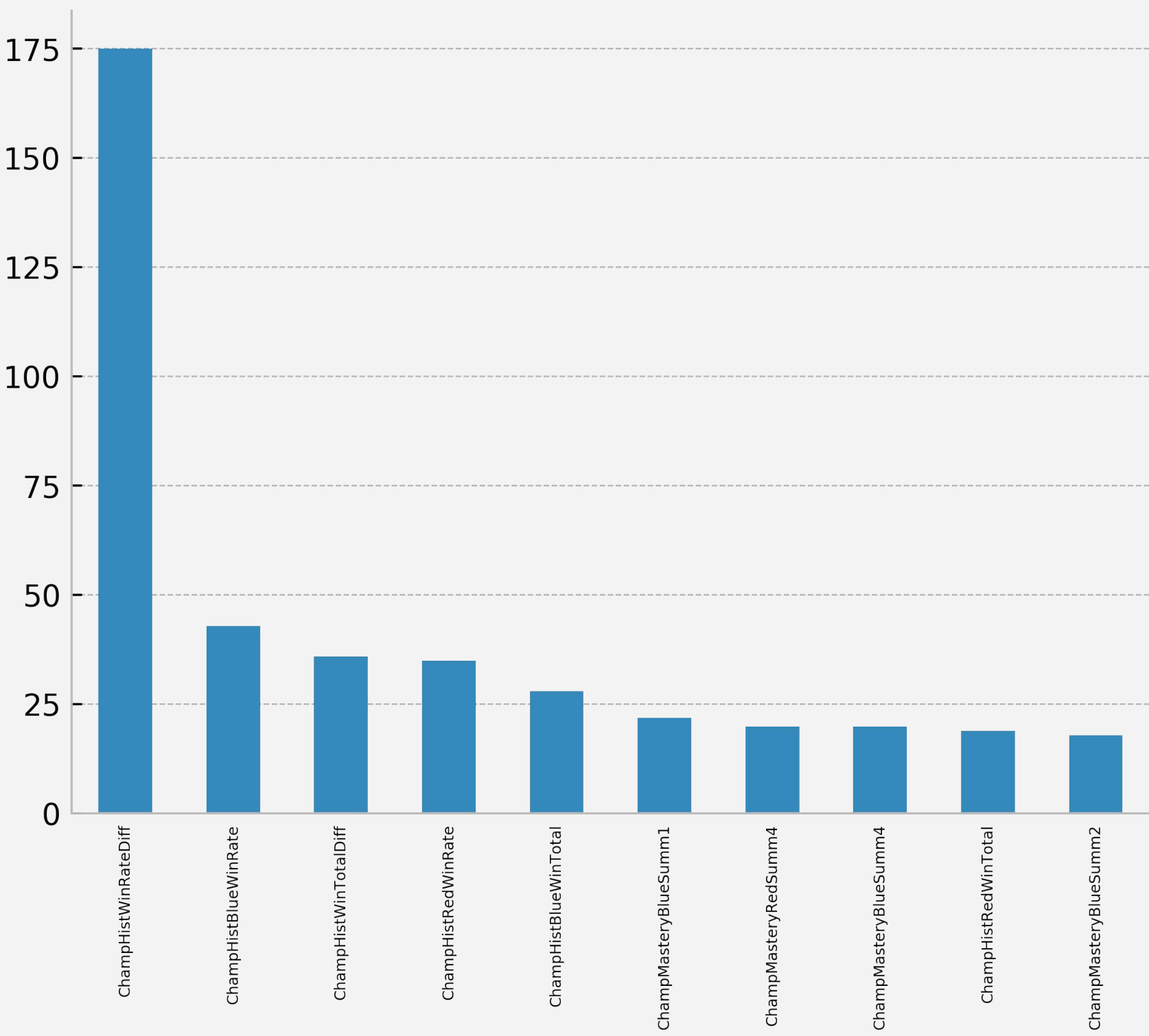
Agregados de atributos e ajuste de hiperparâmetros

Usando como base de comparação, os resultados pregame foram melhores que os resultados in-game até 30 minutos de partida.

Intervalos de tempo	AdaBoost	XGBoost	Random Forests
0 a 10 minutos	71.80% (0.82%)	71.93% (0.98%)	71.89% (1.01%)
0 a 20 minutos	81.95% (0.75%)	82.03% (0.79%)	81.85% (0.69%)
0 a 30 minutos	87.98% (0.50%)	88.09% (0.47%)	87.88% (0.47%)
0 até o fim do jogo	90.37% (0.51%)	90.57% (0.46%)	90.35% (0.51%)

Acurácia dos modelos para dados ingame

Relação dos 10 atributos mais importantes para os dados pregame



### Conclusão e Trabalhos Futuros

O uso das features de histórico retornou uma acurácia maior do que o de dados in-game até 30 minutos, porém se mostrou inviável de obter em tempo real. No entanto foi alcançada 64.80% de acurácia com dados que poderiam ser coletados em tempo hábil.

Para que esse processo realmente se torne viável, uma interface será criada para automatizar a coleta dos dados para os jogadores e campeões presentes em uma partida, da qual o usuário participará.

Também serão coletados dados de jogos profissionais para uma análise das tendências e do mercado de apostas em jogos competitivos.



DCC

DEPARTAMENTO DE  
CIÊNCIA DA COMPUTAÇÃO