

Previsão do time vencedor de partidas de Dota 2

Camilo S. Melgaço¹, Jean P. V. Batinga¹
Thiago R. da Silva¹, Victor A. J. da Cruz¹

¹ Departamento de Computação
Centro Federal de Educação Tecnológica de Minas Gerais
Belo Horizonte, MG – Brasil

Abstract. *Dota 2 is a complex electronic game, where each match has several variables and virtually infinite possible ways to reach the final result. This paper shows how to use the artificial intelligence methods Random Forest and K-Nearest Neighbors to train a model and predict the winner of matches using as features the heroes chosen and the advantages in net worth and experience of each team. The results show that only the heroes are not enough to get a good prediction, but using the advantages the results are good, and 75% can be achieved combining both advantages and heroes as features.*

Resumo. *Dota 2 é um jogo eletrônico complexo, no qual cada partida tem inúmeras variáveis e virtualmente infinitos meios de se chegar a um dos dois possíveis resultados finais. Este trabalho mostra como usar os métodos de inteligência artificial Random Forest e K-Nearest Neighbors para treinar um modelo e prever o vencedor de partidas utilizando como features os heróis selecionados e as vantagens em patrimônio e experiência de cada time. Os resultados mostram que apenas os heróis não são suficientes para uma boa predição do vencedor, mas utilizando as vantagens os resultados são satisfatórios, e combinando ambos, pode-se prever com 75% de precisão o vencedor de uma partida.*

1. Introdução

Dota 2 é um jogo eletrônico onde 5 jogadores enfrentam outros 5 em uma guerra de território, cujo objetivo é destruir a base do time adversário. Essa guerra se estende ao longo de brigas por objetivos parciais até que o objetivo final possa ser atingido. Durante esse tempo, batalhas entre os times são travadas, rendendo ouro e experiência para os jogadores que vencem, e custando ouro e tempo fora da partida para os que perdem. Dota é um jogo extremamente complexo e com muitas variáveis, sendo tão inesperado que algumas vezes o time que esteve ganhando durante toda a partida, acaba perdendo, mesmo tendo ampla vantagem em patrimônio (ouro) e experiência. Este trabalho tenta prever o vencedor de uma partida de Dota 2 usando métodos de inteligência artificial.

2. Sobre Dota 2

Cada time tem 5 jogadores, que durante a fase de escolha de herói, podem selecionar livremente seu herói dentre um vasto conjunto. Todos os jogadores escolhem do mesmo conjunto, e um herói pode ser escolhido por apenas um jogador. Heróis tem características diferentes entre si, sendo únicos na sua forma de contribuir para o time, se especializando em funções diferentes. Mesmo assim, o mesmo herói pode ser utilizado de diferentes

formas, e cada jogador escolhe como deseja empregar as características daquele que escolheu.

Após a escolha, cada jogador começa com o mesmo patrimônio e nenhuma experiência. Estes devem adquirir ouro e experiência matando pequenas criaturas que se distribuem por todo o mapa. Cada jogador tem seu próprio nível de experiência e patrimônio, sendo independente dos outros. Partidas tem muitas variáveis, mas duas delas são amplamente usadas para definir a situação atual dos times: patrimônio e experiência, que são respectivamente os patrimônios e experiência de cada jogador do time somados.

3. Metodologia

Todo o código e desenvolvimento encontra-se em repositório público, contendo instruções para execução e reprodução dos resultados [Melgaço et al. 2021]. Para reproduzir os resultados, deve-se baixar o código fonte e a base de dados da YASP [Cui et al. 2015], extraí-la e renomear o arquivo para "bigdata.json", colocá-lo na pasta "data" do código fonte e executar o *notebook Jupyter* do código fonte.

A seguir serão explicados os detalhes sobre a base de dados utilizada, o tratamento desses dados e a análise dos mesmos.

3.1. Base de dados

Neste trabalho, usaremos uma base de dados disponibilizada pelo YASP [Cui et al. 2015], hoje chamado OpenDota [OpenDota 2015], contendo as últimas 500 mil partidas processadas pelo serviço no ano de 2015. Detalhes sobre o significado de cada campo podem ser encontrados na documentação da API do OpenDota [OpenDota 2021].

3.2. Tratamento dos dados

O jogo é atualizado em intervalos não regulares, e heróis tem suas características alteradas, e isso altera a dinâmica de como eles se encaixam nas partidas. Existem também vários tipos de partidas, com diferentes modos de se jogar. Para garantir a consistência das análises, elas serão feitas sempre extraíndo apenas partidas do mesmo tipo, e com a mesma versão de atualização do jogo. As partidas analisadas nesse trabalho seguem os seguintes critérios:

- Todos os jogadores são humanos (não bots);
- Tipo de partida é competitiva e ranqueada;
- Todas as partidas são da mesma versão;
- Todos os jogadores ficaram na partida até o final.

A base de dados é muito grande, então é processada iterativamente até que o número de partidas desejado seja atingido. Cada partida que atende aos critérios é salva em uma nova base de dados, contendo apenas as *features* a serem utilizadas na análise. Essa nova base é reduzida e posteriormente pode ser carregada para análise.

3.3. Análise dos dados

Após as partidas alvo serem extraídas, a análise é feita sobre essas partidas. Serão feitas três análises variando as *features*:

1. Apenas os heróis de cada equipe;

2. Apenas as vantagens de cada equipe;
3. Heróis e vantagens de cada equipe;

Como as vantagens ao final da partida são muito indicativas de seu resultado, serão utilizadas as vantagens que cada time possuía no meio da partida. Para as partidas analisadas, o organizador de partidas do jogo automaticamente escolhe jogadores de um mesmo nível de habilidade, sendo as partidas niveladas. Logo, características dos jogadores não serão consideradas.

Cada análise apresenta os resultados de dois métodos:

- Random Forest: são geradas várias árvores de decisão com parâmetros variando de forma aleatória, e é escolhida a que melhor se ajusta ao modelo.
- K-Nearest Neighbors: usa uma métrica de proximidade com outras entradas para classificar.

4. Resultados

Foram extraídas 10 mil partidas da versão de jogo 6.85. Os resultados podem ser reproduzidos usando estes parâmetros ao executar o algoritmo. Cada resultado será apresentado em uma subseção de acordo com as *features* utilizadas.

4.1. Apenas heróis

Esses são os resultados utilizando apenas heróis como *features*.

4.1.1. Random Forest

	precision	recall	f1-score	support
0	0.50	0.43	0.47	11985
1	0.54	0.60	0.57	13015
accuracy			0.52	25000
macro avg	0.52	0.52	0.52	25000
weighted avg	0.52	0.52	0.52	25000

Figure 1. Métricas

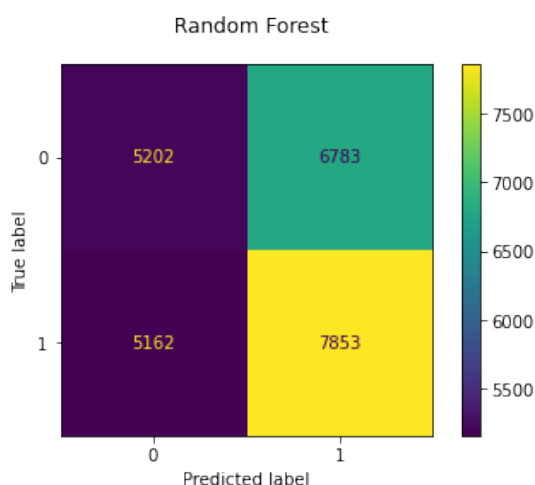


Figure 2. Matriz de confusão

4.1.2. K-Nearest Neighbors

	precision	recall	f1-score	support
0	0.48	0.46	0.47	11985
1	0.52	0.55	0.53	13015
accuracy			0.50	25000
macro avg	0.50	0.50	0.50	25000
weighted avg	0.50	0.50	0.50	25000

Figure 3. Métricas

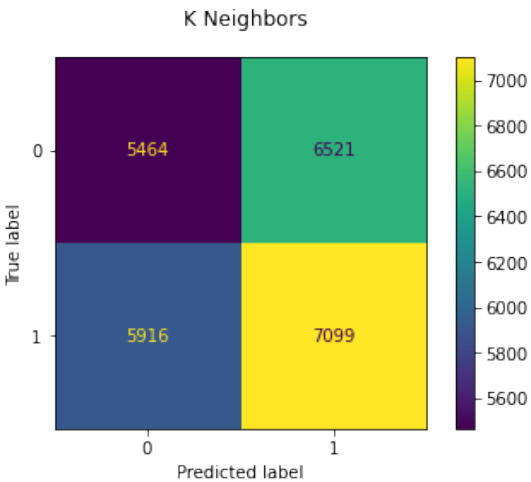


Figure 4. Matriz de confusão

4.2. Apenas vantagens

Esses são os resultados utilizando apenas vantagens (patrimônio e experiência) como *features*.

4.2.1. Random Forest

	precision	recall	f1-score	support
0	0.48	0.46	0.47	11985
1	0.52	0.55	0.53	13015
accuracy			0.50	25000
macro avg	0.50	0.50	0.50	25000
weighted avg	0.50	0.50	0.50	25000

Figure 5. Métricas

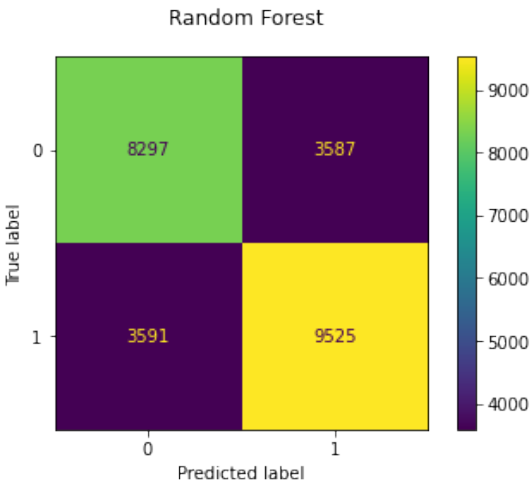


Figure 6. Matriz de confusão

4.2.2. K-Nearest Neighbors

	precision	recall	f1-score	support
0	0.69	0.69	0.69	11884
1	0.72	0.72	0.72	13116
accuracy			0.71	25000
macro avg	0.71	0.71	0.71	25000
weighted avg	0.71	0.71	0.71	25000

Figure 7. Métricas

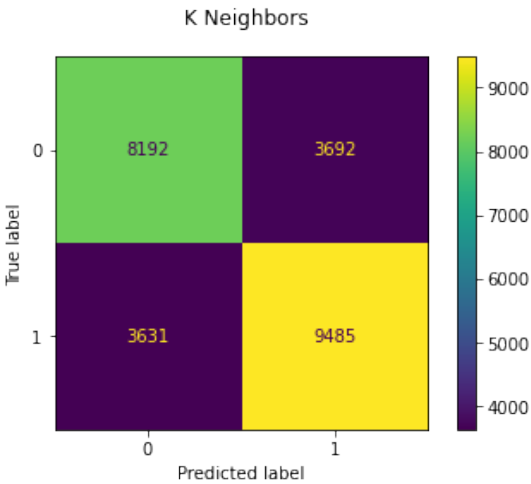


Figure 8. Matriz de confusão

4.3. Heróis e vantagens

Esses são os resultados utilizando heróis e vantagens (patrimônio e experiência) como *features*.

4.3.1. Random Forest

	precision	recall	f1-score	support
0	0.74	0.73	0.73	11979
1	0.75	0.76	0.76	13021
accuracy			0.75	25000
macro avg	0.75	0.75	0.75	25000
weighted avg	0.75	0.75	0.75	25000

Figure 9. Métricas

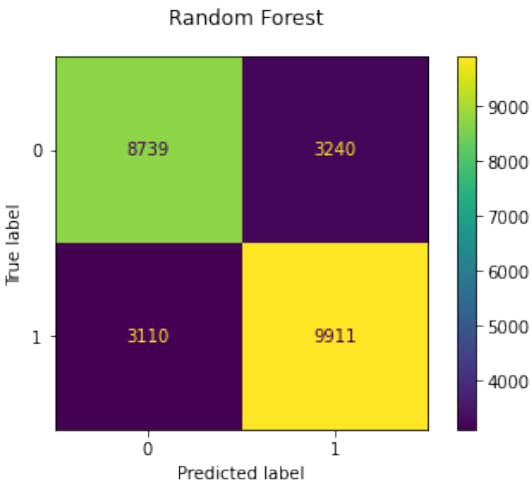


Figure 10. Matriz de confusão

4.3.2. K-Nearest Neighbors

	precision	recall	f1-score	support
0	0.70	0.69	0.70	11979
1	0.72	0.73	0.72	13021
accuracy			0.71	25000
macro avg	0.71	0.71	0.71	25000
weighted avg	0.71	0.71	0.71	25000

Figure 11. Métricas

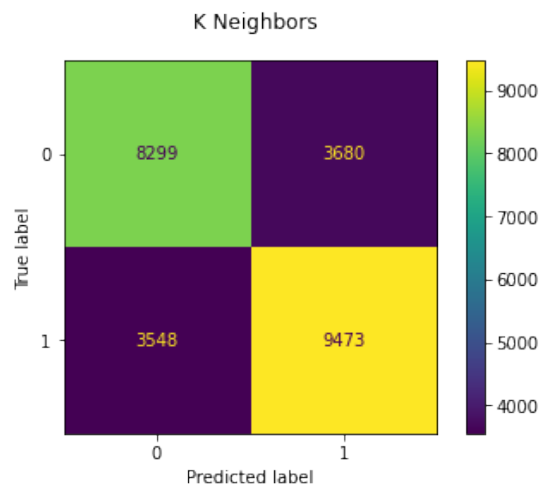


Figure 12. Matriz de confusão

5. Discussão

Como pode ser verificado nos resultados, os métodos tem resultados parecidos, tendo diferença significativa apenas na análise considerando ambos heróis e vantagens. Sendo assim, utilizaremos os resultados do método Random Forest para discussão, que se apresentou como melhor método neste trabalho.

Os resultados estão de acordo com o esperado. No Dota 2, os heróis são extremamente versáteis e podem desenvolver papéis distintos dependendo da necessidade da partida. Com isso, não se pode afirmar se um time é completo ou não (considerando competências necessárias para atingir objetivos) apenas analisando os heróis escolhidos. Com isso, analisando apenas os heróis, o resultado é o mesmo de uma escolha aleatória quando se tem duas possibilidades: 50%. Ao se considerar vantagens, temos uma considerável melhora, e podemos dizer que o time que está ganhando a partida, provavelmente vai ganhar, com 71% de precisão. Isso se dá pois a vantagem em patrimônio e experiência é um bom indicador de qual time tem o controle da partida. Assim, analisando as partidas considerando ambos os conjuntos de *features*, temos uma acurácia de 75%, um resultado expressivo para um número tão grande de partidas analisadas, onde cada uma tem tantas variáveis. Estes números representam bem a dinâmica do jogo, pois mostram que mesmo que um determinado time tenha o controle em dado momento, com heróis balanceados, a partida não está ganha, ainda pode ser virada.

References

- Cui, A., Chung, H., and Hanson-Holtry, N. (2015). Yasp december 2015 500k data dump. *Academic Torrents*. <https://academictorrents.com/details/384a08fd7918cd59b23fb0c3cf3cflaea3ea4d42>.
- Melgaço, C. S., Batinga, J. P. V., da Silva, T. R., and da Cruz, V. A. J. (2021). dota2predictor. *GitHub*. <https://github.com/camilosm/dota2predictor>.

OpenDota (2015). December 2015 data dump. *The OpenDota Blog*. <https://blog.opendota.com/2015/12/20/datadump/>.

OpenDota (2021). Documentation. *OpenDota API*. <https://docs.opendota.com/#tag/matches>.