

Prevendo o resultado de partidas de Dota 2

Victor B. R. Jorge

Universidade Federal de Minas Gerais

Belo Horizonte, Minas Gerais, Brazil

victorbrjorge@dcc.ufmg.br

ABSTRACT

Neste trabalho, apresentamos um método para se prever resultados de partidas de popular jogo de computador Dota 2. Detalhamos a complexidade de se obter um previsor para o problema devido ao tamanho do espaço de soluções e apresentamos alternativas para contornar esse problema.

Além disso, discutimos o desempenho de algoritmos de aprendizado de máquina ao resolver o problema. Ademais, apresentamos alternativas de como codificar e selecionar features, além de comparar algoritmos e métricas de avaliação para o problema.

1 INTRODUÇÃO

Dota 2 é um jogo de computador do estilo *Multiplayer Online Battle-Arena* (MOBA) e possui, atualmente, uma base de aproximadamente 15 milhões de jogadores únicos, a segunda maior do mundo. Além de abrigar partidas entre jogadores casuais, Dota 2 é um dos principais jogos do cenário dos esportes eletrônicos - os *e-sports*. Atrás das premiações milionárias, vários jogadores se tornaram profissionais do jogo e seguem rotinas de verdadeiros atletas. Para se ter uma ideia da rentabilidade do Dota 2 como esporte, o último campeonato mundial - o *The International 2016* - teve uma premiação total na casa dos 20 milhões de dólares.

Cada partida de Dota 2 consiste em dois times de 5 jogadores cada. Cada lado, Radiant e Dire, tem como objetivo destruir a fortaleza inimiga - o *Ancient*. Cada jogador controla um personagem, conhecido como herói. Cada herói possui uma série de habilidades únicas - geralmente quatro - que são acessadas e melhoradas ganhando-se níveis. Para subir de nível, os heróis devem adquirir experiência (XP). Existem três maneiras de se obter XP: matando heróis inimigos, matando criaturas inimigas, e matando criaturas neutras. Além da experiência, os heróis adquirem também ouro. O ouro pode ser adquirido de maneira análoga à XP, porém, para um herói receber o ouro relativo ao abate de uma criatura ou herói inimigo, ele deve ser o responsável pelo abate (último ataque). Com ouro, o herói pode adquirir itens que o darão vantagem na batalha para destruir o *Ancient* inimigo.

Dota 2 engloba vários modos de jogo. Esses modos se diferenciam, basicamente, na maneira em que os heróis podem ser escolhidos pelos jogadores. Neste trabalho, focaremos nos dois principais modos de jogo: *All Pick* e *Captains Mode*. O *All Pick* é o modo de jogo principal das partidas casuais. Nele, cada jogador tem direito a escolher seu herói a partir de todos 112 heróis disponíveis no jogo. Quando um herói é escolhido por um jogador, ele se torna indisponível aos outros jogadores. Portanto, mesmo que em times diferentes, não é possível haver heróis duplicados em uma partida. Já o *Captains Mode* é o modo utilizado nas partidas profissionais e de campeonatos. Nesse modo, cada time possui a figura do capitão, que é o jogador responsável por escolher todos os heróis do time.

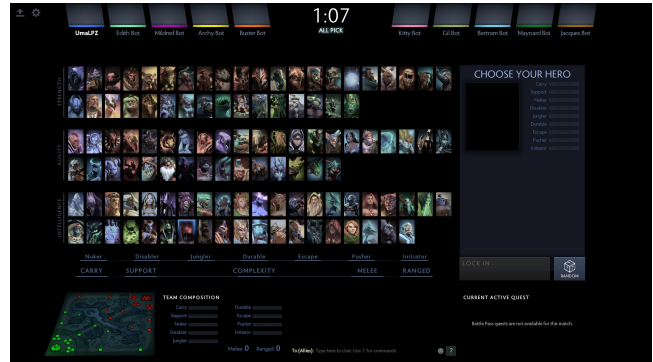


Figure 1: Tela de seleção de heróis do modo All Pick.

Além disso, no *Captains Mode* os capitães têm direito a retirar um certo número de heróis da partida.

Os jogadores levam em consideração diversos fatores antes de realizar a escolha de seu herói e não somente as habilidades individuais daquele herói. A sinergia do herói escolhido com o resto do time e o potencial de contra-ataque aos heróis do outro time também é levado em consideração. Na verdade, uma combinação sinérgica de heróis é um dos fatores mais importantes em uma partida de Dota 2. Além, claro, da habilidade dos jogadores para trabalhar em equipe para conquistar os objetivos. Sendo assim, uma boa escolha de heróis pode até compensar uma diferença de habilidade entre os jogadores dos dois times.

Podemos visualizar esse problema como um problema clássico de classificação binária. O que nos chama a atenção é o tamanho do espaço de soluções: o número de possíveis combinações única de heróis chega à ordem de grandeza de 10^6 . Portanto, quase todos os jogos são únicos e merecem ser analisados. A tarefa de prever o resultado de um jogo somente através da escolha de heróis é desafiadora, porque ela tenta modelar estatisticamente um conhecimento obtido por jogadores experientes ao longo de vários anos.

2 TRABALHOS RELACIONADOS

Dota Picker [1] é uma aplicação que usa métricas especificadas para recomendação de heróis. As duas métricas disponíveis são vantagens de heróis (confronto entre dois heróis) e taxas de vitórias. Infelizmente, os criadores não divulgaram detalhes específicos de sua abordagem nem sua precisão.

DotaBuff [2] é um site que fornece várias estatísticas do Dota 2. Atualmente, não fornece nenhuma forma de previsão de vitória ou recomendação de herói, mas suas várias visualizações de dados são úteis.

O artigo How Does He saw me? [3] descreve um sistema de recomendação de heróis Dota 2 que depende de um preditor de

vitórias. Especifica dois modelos possíveis para a previsão de vitórias. O primeiro usa regressão logística com um vetor de herança binária. O segundo usa a classificação do vizinho mais próximo usando um peso personalizado para especificar a distância entre as equipes.

O artigo To Win or Not To Win [4] descreve um preditor de vitórias Dota 2. Especifica dois modelos possíveis para a previsão de vitórias. O primeiro usa regressão logística com um vetor de herança binária. O segundo combina esse preditor com uma métrica de aptidão genética, ponderando cada preditor igualmente na previsão final.

3 CONHECIMENTO PRÉVIO

Para este trabalho, é fundamental um entendimento mais profundo das mecânicas e dos papéis do herói em uma partida de Dota 2. Só assim será possível estabelecer um conceito de sinergia e contra-ataque entre os times. Cada herói possui um papel dentro do jogo. Este papel, porém, não está atrelado somente ao herói em si, mas sim à composição do seu time. Apesar de alguns heróis possuírem habilidades que os encaixam melhor em alguns papeis, um herói pode alterar seu papel dependendo da estratégia do seu time no jogo. Os papeis básicos dos heróis são:

- **Core** - É um herói que possui a prioridade no acúmulo de ouro e experiência no time. Deve ser capaz, através de suas habilidades e itens, de infringir danos ao time adversário. Um time bem formado possui, em média, 3 cores.
- **Support** - Responsável por proporcionar um jogo tranquilo aos cores, garantindo que eles consigam acumular o ouro e experiência necessários para o time vencer a partida. Além disso, são os responsáveis por comprar os itens de uso comum, como itens que proporcionam visão extra no mapa do jogo. Um time bem formado, possui, em média, 2 supports.

Como dito anteriormente, um herói pode cumprir tanto a função de Core quanto de Support em diferentes partidas. Entretanto, há heróis que possuem habilidades que os tornam mais adequados para uma função ou outra. Sendo assim, fica fácil perceber a importância de um time bem formado. Um time com 5 Cores, fracassará na maioria das vezes pela falta de heróis que ofereçam suporte para eles. Assim como um time com 5 supports fracassará por falta de heróis que consigam implicar grande quantidade de dano ao time inimigo.

4 CONJUNTO DE DADOS

O conjunto de dados utilizado neste trabalho consiste de 50mil partidas disponível no Kaggle ¹. As partidas contemplam o período de x a y. Alguns filtros foram aplicados no dados de modo a retirar partidas que são pouco significativas para nosso problema. Foram aplicados os seguintes filtros:

- Somente partidas dos modos All Pick e Captains Mode foram analisadas. Os outros modos não representam a escolha de heróis de maneira deliberada pelos jogadores. Nesses modos, em especial, um herói tem a mesma probabilidade de aparecer em qualquer um dos times.
- Todos os jogos analisados terminaram com os 10 jogadores jogando a partida. Jogos em que há abandono de jogadores

Table 1: Estatísticas pós filtros

	Radiant	Dire	Total
Qtd. partidas	18960	19761	38721
Herói mais escolhido	Pudge	Oracle	Oracle

de algum dos times não representam o jogo em seu potencial pleno e introduzem um viés à predição.

- Todos os jogos possuem uma duração mínima de 10 minutos. No Dota 2, é praticamente impossível se ganhar um jogo em que todos os jogadores estão se esforçando para serem vitoriosos em menos de 10 minutos. Esse tipo de partida não representa como os heróis realmente funcionam e interagem entre si.

Após a filtragem das partidas, nosso conjunto de dados possui as seguintes características mostradas na Tabela 1.

5 METODOLOGIA

5.1 Vetor de features

Para codificar a escolha de heróis como uma feature do nosso conjunto de dados, foi utilizado um vetor $x \in \mathbb{R}^{224}$ que mapeia a escolha de heróis da seguinte maneira: As 112 primeiras posições do vetor, correspondem à escolha de heróis do Radiant. A posição do vetor corresponde ao herói de mesmo *id*. Ela será 1 caso o herói seja escolhido e 0 caso contrário. O mesmo raciocínio vale para as 112 posições finais do vetor, porém essas posições tratam do time Dire. Essa técnica de codificação de atributos categóricos é conhecida como one-hot encoding. Além disso, para cada partidas, temos um label y que é igual a 1 caso o Radiant seja o vencedor da partida e 0 caso contrário.

Um fato interessante a se notar é o viés que o time pode dar aos dados. Teoricamente, no Dota 2 o mapa é simétrico e nenhum dos times possui alguma vantagem por simplesmente pertencer ao Radiant ou Dire. Portanto, para evitar que haja configurações de heróis que apareçam mais em um time que outro, todas as partidas foram replicadas invertendo-se a ordem dos vetores que representam as escolhas de heróis dos times e do label y .

5.2 Previsão com dados pós partida

Primeiramente, foram utilizados alguns dados pós partida para se treinar um modelo capaz de prever o resultado de uma partida. Além da escolha de heróis, foram utilizadas a diferença de ouro e experiência finais entre os dois times. Esse experimento tem por objetivo servir como ponto de comparação, já que sabemos que é uma tarefa relativamente simples prever o resultado de uma partida conhecendo as diferenças entre o acúmulo de recurso final dos dois times. Para esse experimento, utilizamos o algoritmo Naive Bayes, por ser bastantes simples e requerer poucos hiperparâmetros. Utilizando-se de validação cruzada com 5 conjuntos, obtivemos uma acurácia média de 96,72% com esses dados.

Como era de se esperar, realizar previsões com os dados pós partida é uma tarefa relativamente simples.

¹<https://www.kaggle.com/devinanzelmo/dota-2-matches>

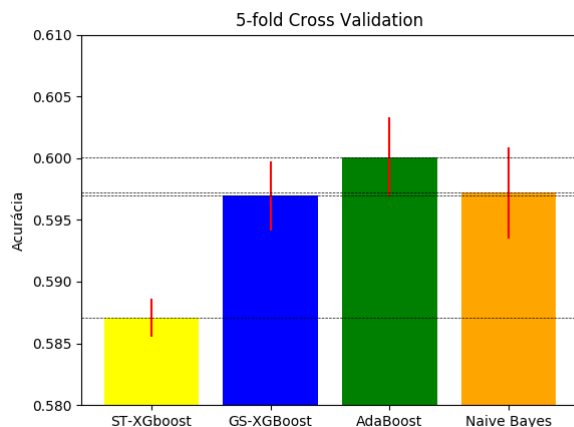


Figure 2: Acurácia dos modelos de previsão.

5.3 Previsão pela escolha de heróis

Neste experimento, temos como objetivo prever o time vitorioso somente com os dados pré jogo, ou seja, a escolha dos heróis. Como foi dito anteriormente, o espaço de soluções desse problema é gigantesco. Portanto, mesmo sendo um problema de classificação binária, é de extrema complexidade. Foram utilizados três algoritmos nessa etapa: XGBoost, AdaBoost e Naive Bayes.

XGBoost é um algoritmo de Boosting que engloba técnicas de Bagging e descida de gradiente. É conhecido pelo alto desempenho aliado à precisão de classificação. É considerado estado da arte para dados estruturados. Porém, requer um extensivo refinamento de hiper-parâmetros.

O AdaBoost foi o primeiro algoritmo de Boosting a ser usado comercialmente. É bastante conhecido pela sua precisão de classificação e por possuir poucos hiper-parâmetros a serem refinados.

Naive Bayes é um algoritmo de classificação estatístico que utiliza distribuições de probabilidades para realizar suas previsões. É bastante conhecido por ser simples e possuir um ótimo custo benefício já que não requer hiper-parâmetros.

Levando em consideração as características dos algoritmos citados acima, foi realizado um GridSearch para se extrair hiper-parâmetros ideais para o XGBoost. Um GridSearch consiste na busca por melhores hiper-parâmetros de maneira empírica. Ou seja, executando o algoritmo com diferentes hiper-parâmetros e os avaliando através de alguma métrica. O desempenho dos algoritmos pode ser avaliado na Figura 2.

6 CONCLUSÕES

Começamos este trabalho com uma simples dúvida, será possível prever com precisão o resultado de uma partida de DotA? E nós respondemos com sucesso a pergunta criando um preditor com 60% de precisão somente com informações sobre as escolhas de heróis.

Para alcançar esse objetivo, exploramos o conjunto de dados da perspectiva dos heróis que estão no jogo, bem como as estatísticas de recursos pós-jogo e encontramos relacionamentos interessantes entre escolhas de heróis. Além disso, obtivemos uma clara

indicação de que Dota, no seu núcleo, pode ser visto como um jogo de recursos em que ter mais recursos leva a uma vitória quase sempre.

Nosso preditor está longe de ser perfeito e isso ocorre porque uma partida de Dota quase nunca é decidida na fase de escolha, mas sim durante o jogo e. Portanto, mais informações sobre o que acontece durante o jogo aumentariam o poder do nosso preditor e melhorariam bastante o desempenho.

7 TRABALHOS FUTUROS

Apesar dos resultados promissores apresentados neste trabalho, podemos trabalhar para encontrar métricas pré partida que ajudem a melhorar a acurácia de classificação. Métricas que consigam quantificar conceitos como sinergia e contra-ataque entre heróis com certeza ajudariam o classificador a entender a semântica do problema.

Além disso, uma feature que consiga mensurar a habilidade do jogador com o herói que ele escolheu evitaria que jogadores com pouca habilidade com um certo herói afetem a performance da classificação.

Ademais, o uso de redes neurais para realizar a previsão também pode ser uma boa estratégia para o problema. Redes neurais podem ser capazes de inferir a semântica do jogo sem necessitar de features adicionais.

Por fim, acreditamos que ainda há bastante espaço para a aplicação de aprendizado de máquina nessa área.

REFERENCES

- [1] Dota 2014. Dota Picker. (2014). <http://dotapicker.com>.
- [2] DOTABUFF 2014. Dotabuff - Dota 2 Statistics. (2014). <http://dotabuff.com>.
- [3] How Does He Saw Me? 2013. A Recommendation Engine for Picking Heroes in Dota 2. (2013). <http://cs229.stanford.edu/proj2013/PerryConley-HowDoesHeSawMeARecommendationEngineForPickingHeroesInDota2.pdf>.
- [4] To win or not to win? 2013. A prediction model to determine the outcome of a DotA2 match. (2013). https://cseweb.ucsd.edu/~jm-cauley/cse255/reports/wi15/Kaushik_Kalyanaraman.pdf.