

# Jogadores Tóxicos em Partidas de Jogos *Online*

*Joaquim Neto e Kazuki Yokoyama*

*07 de outubro de 2016*

## Introdução

Juntamente com a popularização das redes sociais na internet, veio a popularização dos jogos online. Nestes jogos, o comportamento abusivo de uma parcela dos jogadores é um problema constante. Dentro da comunidade destes jogos, esse tipo de comportamento é denominado tóxico, e pode destruir completamente o esforço de equipe entre os jogadores, arruinando a partida dos jogadores afetados, o que pode levar a uma gama de problemas, desde jogadores desistiram de consumir estes jogos, até a problemas psicológicos resultantes do bullying, nos casos mais sérios.

Consideramos um jogador tóxico, um jogador que incorre em comportamento abusivo durante o jogo. Neste trabalho, nós consideramos que jogadores tóxicos contaminam o ambiente ao seu redor, tornando ambiente de jogo pior para todo. Dizemos que ocorre uma contaminação tóxica, quando as ações de um jogador tóxico alteram negativamente o desempenho e/ou o comportamento de outros jogadores com quem ele interage.

## Objetivos

Pretendemos aqui primeiramente avaliar, através da análise dos dados das partidas e dos jogadores, qual é a extensão da contaminação tóxica aos outros jogadores. Para isso, analisaremos os dados com a finalidade de gerar um modelo para uma métrica de toxicidade. Depois tentaremos descobrir, entre os diferentes tipos de ofensas catalogadas pelo jogo, quais delas geram maior contaminação tóxica. Assim, podemos resumir os objetivos primários do trabalho como sendo

- Avaliar a extensão da contaminação tóxica a outros jogadores a partir dos dados da partida;
- Descobrir que tipos de ofensas geram mais contaminação.

## Objetivos da mineração

- Definir três categorias de jogadores distintas: tóxicos, contaminados e limpos;
- Definir uma métrica do desempenho dos jogadores;
- Mostrar que o jogador tóxico tem uma influência negativa no desempenho dos outros jogadores;
- Criar métrica para quantificar a contaminação dos jogadores considerados contaminados.

A fim de conseguirmos tais objetivos, precisamos primeiramente definir o que caracteriza um jogador tóxico, um jogador contaminado, e um jogador não contaminado (limpo). Após isso, teremos que definir uma métrica que quantifique o desempenho dos jogadores para podermos mostrar que o jogador tóxico exerce uma influência negativa no desempenho de outros jogadores. Finalmente, com essas informações, poderemos criar uma métrica que quantifique o grau da contaminação que o jogador tóxico exerce sobre os contaminados.

## Dados

Nosso trabalho tem como base os dados de partidas do jogo *online League of Legends* - LoL. O *dataset* é formado por denúncias de comportamento tóxico (abusivo) por jogadores do jogo. O *dataset* completo contém

1.46 milhões de casos de denúncias, que podem conter uma ou mais partidas com o jogador denunciado, e no total possui mais de 10 milhões de partidas, tornando-o grande demais para o escopo do nosso trabalho.

Devido a grande extensão dos dados, obtemos uma amostra, com aproximadamente 37 mil partidas. Todas as informações estavam originalmente no formato JSON. Para podermos trabalhar mais facilmente com os dados, escolhemos os atributos que consideramos relevantes ao nosso problema, e os dividimos em dois arquivos CSV: um contendo os dados inerentes à própria partida (`matches.csv`), outro contendo os dados dos jogadores de cada partida (`players.csv`), cujos atributos estão descritos abaixo:

### **matches.csv**

- *case* - caso dentro do dataset a partida pertence;
- *match* - numeração da partida dentro do caso;
- *premade* - se foi uma partida combinada anteriormente com os outros jogadores do time (1), ou se foi uma partida aonde o jogo combinou os jogadores (0);
- *most.common.offense* - qual foi o tipo de ofensa mais comum do jogador denunciado que foi reportada pelos outros jogadores da partida;
- *reports.allies* - quantidade de denúncias feitas pelo time aliado do jogador denunciado;
- *reports.enemies* - quantidade de denúncias feita pelo time inimigo ao jogador denunciado;
- *reports.case* - quantidade total de denuncias em todas as partidas do caso;
- *time.played* - tempo de jogo decorrido nessa partida.

### **players.csv**

- *case* - caso dentro do dataset a partida pertence;
- *match* - numeração da partida dentro do caso;
- *relation.offender* - se este jogador está no mesmo time do denunciado (*ally*), em time diferente (*enemy*) ou se é o próprio denunciado (*offender*);
- *champion* - com qual personagem do jogo este jogador está jogando. Note que dois jogadores no mesmo time não podem ter o mesmo personagem;
- *kills* - quantidade de vezes que este jogador matou outros jogadores durante a partida;
- *deaths* - quantidade de vezes que este jogador morreu na partida;
- *assists* - quantidade de vezes que este jogador proveu algum tipo de assistência para o abate de um jogador inimigo, mas não realizou o abate ele mesmo;
- *gold* - quantidade de ouro acumulada durante a partida. Normalmente é usado para medir o desempenho de um jogador;
- *outcome* - Se este jogador ganhou a partida (*Win*), se ele a perdeu (*Loss*) ou se ele saiu do jogo antes dele acabar (*Leave*).

## **Transformação dos dados**

Para a nossa análise exploratória, utilizamos a linguagem R. Neste processo de análise exploratória, analisamos e associamos os dados de ambos os datasets construídos (players e matches) de maneiras diferentes, das quais apenas os resultados mais relevantes, e os que contribuem para uma compreensão melhor do dataset foram reportados aqui. Nossos objetivos nessa etapa são: descobrir possíveis padrões que sejam de futuro interesse, e procurar dados que suportem nossas hipóteses.

Abaixo disponibilizamos o resumo de informações de ambos os arquivos obtidos:

### **matches.csv**

##	case	match	premade
----	------	-------	---------

```

## Min.      : 10000    Min.      :0.000    Min.      :0.00000
## 1st Qu.: 39925    1st Qu.:0.000    1st Qu.:0.00000
## Median :400824    Median :1.000    Median :0.00000
## Mean    :291928    Mean    :1.553    Mean    :0.04564
## 3rd Qu.:479911    3rd Qu.:3.000    3rd Qu.:0.00000
## Max.     :530999    Max.     :4.000    Max.     :1.00000
##
##           most.common.offense reports.allies reports.enemies
## Offensive Language :11837    Min.      :0.000    Min.      :0.0000
## Intentionally Feeding: 9444    1st Qu.:1.000    1st Qu.:0.0000
## Verbal Abuse       : 6961    Median :1.000    Median :0.0000
## Negative Attitude  : 4478    Mean    :1.376    Mean    :0.6486
## Assisting Enemy Team : 3298    3rd Qu.:2.000    3rd Qu.:1.0000
## Spamming           :  875    Max.     :4.000    Max.     :5.0000
## (Other)            :  225
## reports.case      time.played
## Min.      : 1.000    Min.      : 378
## 1st Qu.: 5.000    1st Qu.:1481
## Median : 6.000    Median :1928
## Mean    : 7.741    Mean    :1967
## 3rd Qu.:10.000    3rd Qu.:2363
## Max.     :30.000    Max.     :5362
##

```

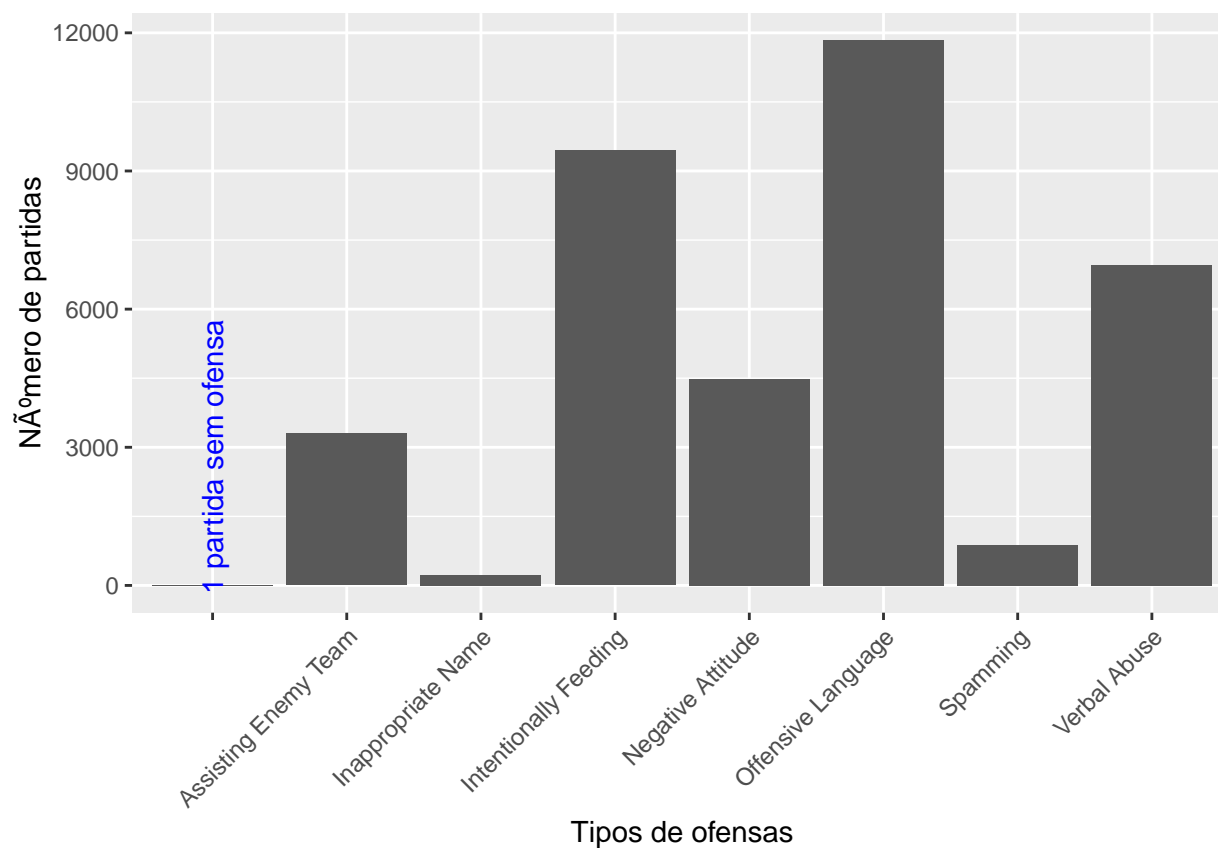
## players.csv

```

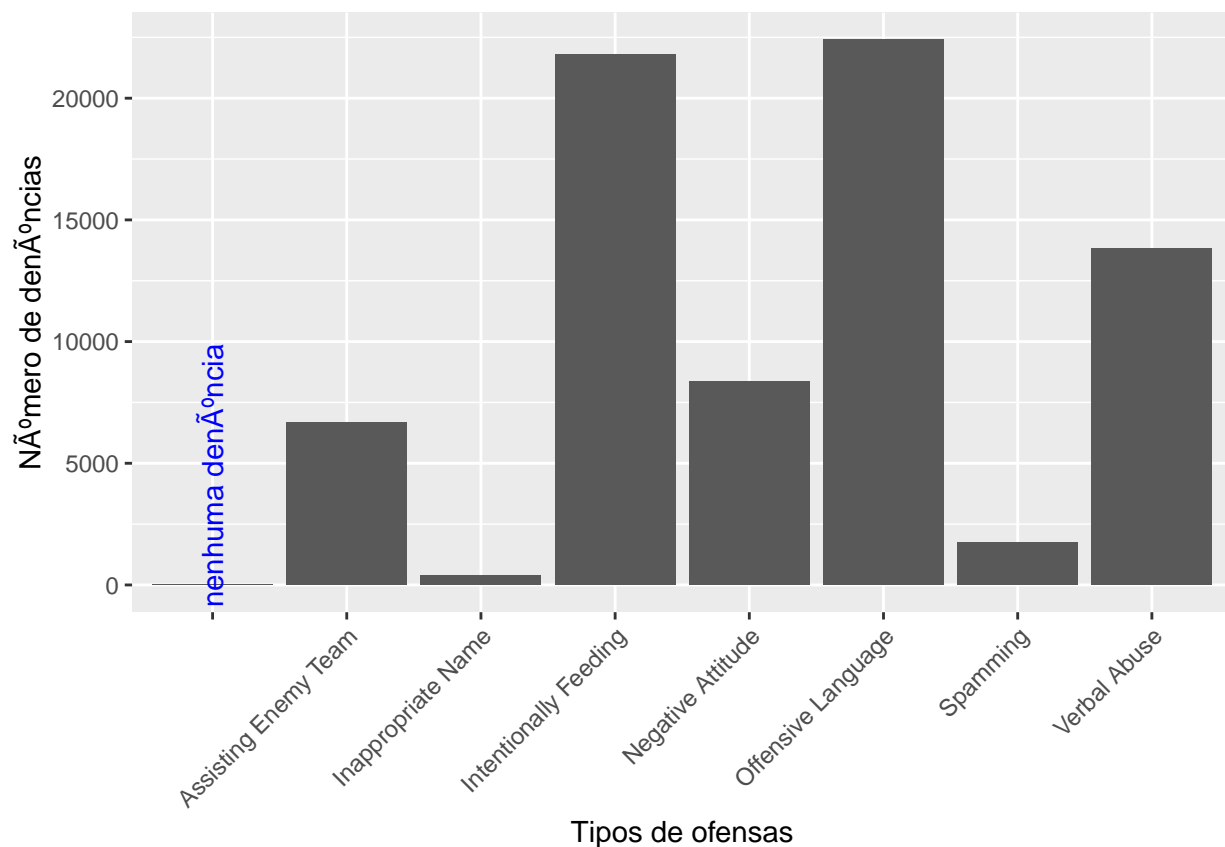
##           case           match      relation.offender      champion
## Min.      : 10000    Min.      :0.000           : 103    Ashe      : 9783
## 1st Qu.: 39923    1st Qu.:0.000      ally   :143817    Ezreal    : 9290
## Median :400832    Median :1.000      enemy   :179541    Lee Sin   : 9005
## Mean    :292188    Mean    :1.551      offender: 37117    Teemo     : 8858
## 3rd Qu.:479923    3rd Qu.:3.000           Blitzcrank: 8367
## Max.     :530999    Max.     :4.000           Jax       : 8039
##                                     (Other)   :307236
##           kills           deaths           assists           gold
## Min.      : 0.000    Min.      : 0.000    Min.      : 0.000    Min.      : 1027
## 1st Qu.: 2.000    1st Qu.: 3.000    1st Qu.: 4.000    1st Qu.: 6647
## Median : 5.000    Median : 5.000    Median : 7.000    Median : 9178
## Mean    : 5.802    Mean    : 5.818    Mean    : 8.402    Mean    : 9551
## 3rd Qu.: 8.000    3rd Qu.: 8.000    3rd Qu.:12.000    3rd Qu.:12039
## Max.     :49.000    Max.     :37.000    Max.     :80.000    Max.     :38097
##
##           outcome
## Leave: 4115
## Loss :176394
## Win  :180069
##
##
##
##

```

## Gráficos

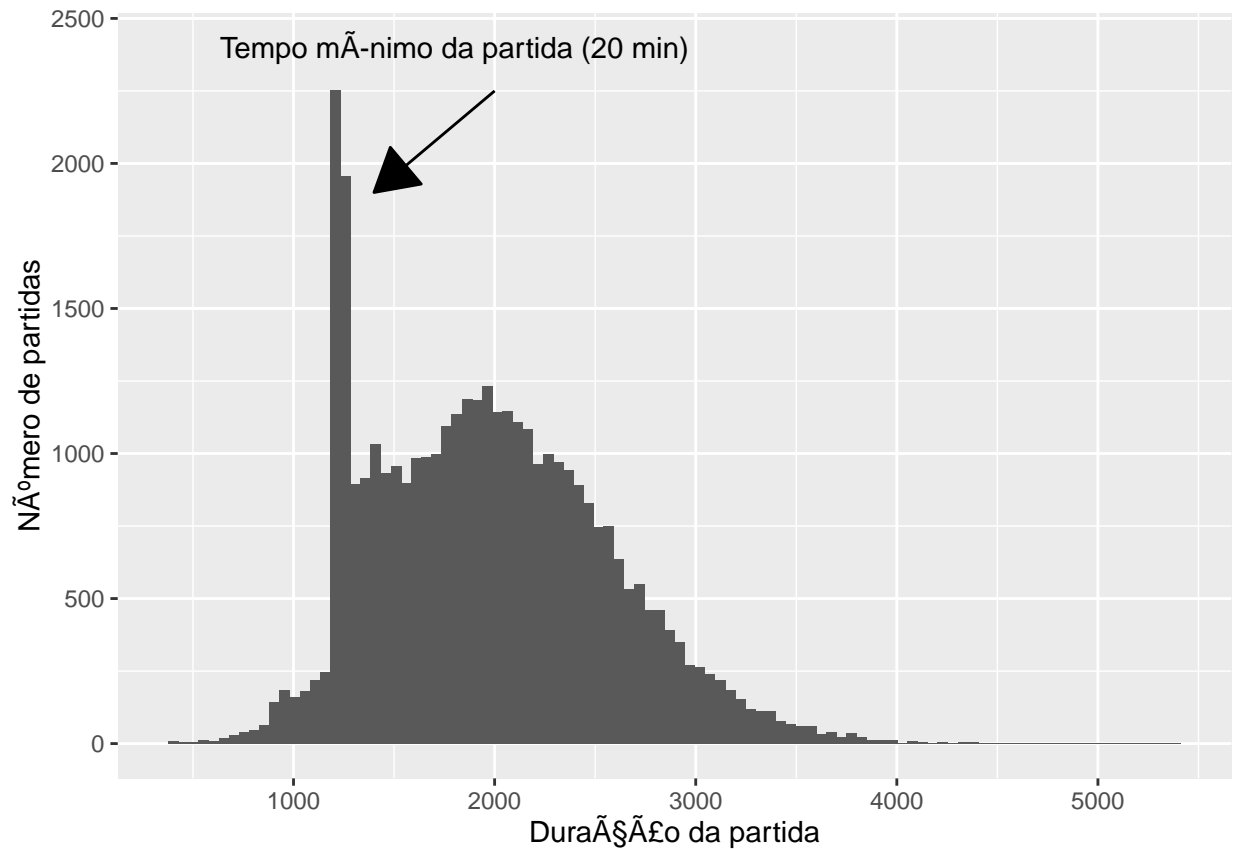


Neste gráfico, nós exploramos quais os tipos de comportamento tóxico mais incidentes nas partidas. No eixo X são listados todos os tipos diferentes de denúncias catalogados pelo jogo, e no eixo Y nós mostramos em quantas partidas no dataset o jogador ofensor foi denunciado por cada um dos motivos.

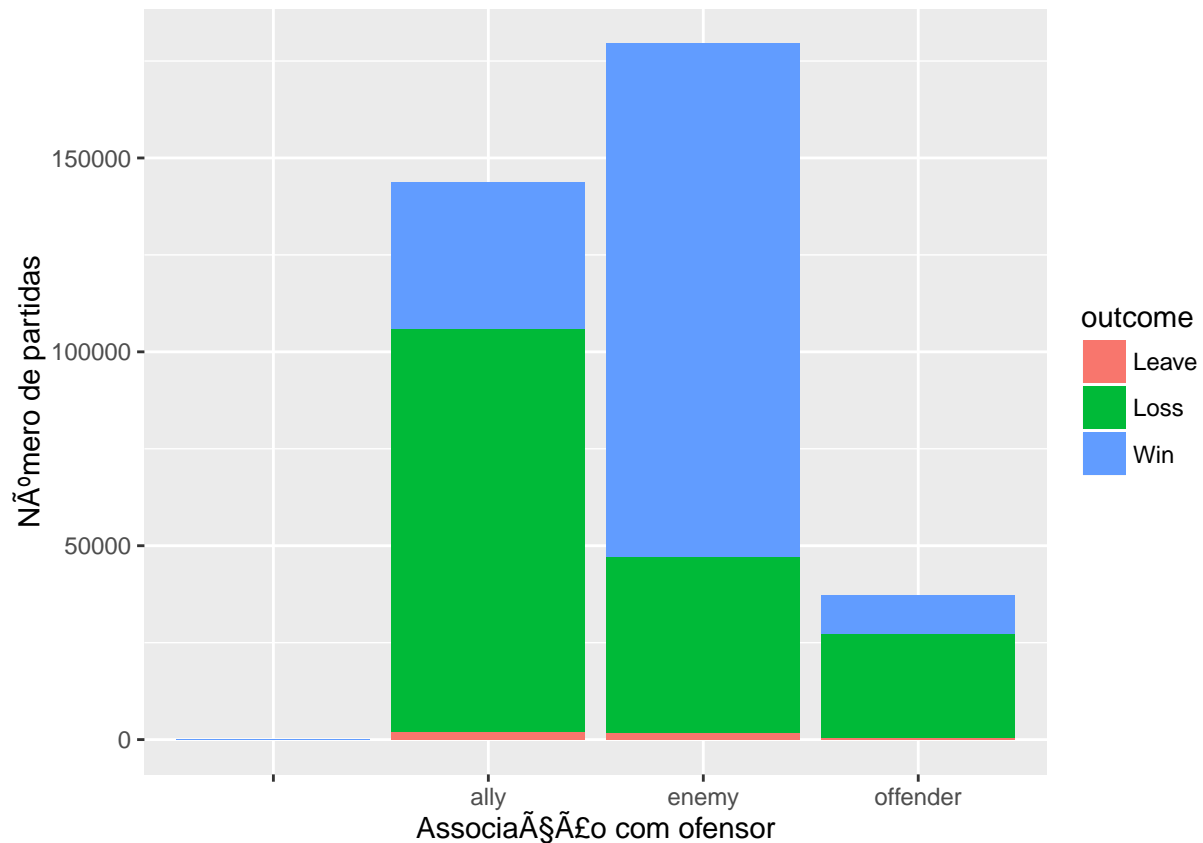


Este gráfico é bem similar ao anterior, com a diferença de que ao invés de ser o número de partidas em que o jogador ofensor foi denunciado por cada tipo de ofensa, nós contamos o número total de denúncias feitas pelos jogadores.

Essa comparação tem como objetivo explicitar que, pela análise exploratória, “Intentionally Feeding” é uma ofensa mais grave que as demais, pois ela tem uma média maior de denúncias por partida. Isso pode nos ajudar a fazer uma avaliação futura do modelo de métrica de toxicidade. Também fica evidente que os demais atributos mantêm a mesma proporção.



Fazendo o histograma da variável “time.played”, é fácil ver que há uma anomalia por volta de 1200 e 1300 segundos (~20min). Isso representa que muitas partidas terminam neste período. Supomos que isso ocorra devido a mecânica de rendimento do jogo, aonde a partir de 20 minutos de partida um time pode escolher se render e acabar com a partida. Pode se fazer várias análises futuras de proveito sobre essa anomalia, como por exemplo, descobrir se essas partidas tem uma incidência maior de toxicidade nelas.



Nesta comparação entre os campos 'relation.offender' e 'outcome', nós buscamos saber se existe uma relação entre a presença do ofensor em um time, e a taxa de vitórias do time. Aqui encontramos que um time com um jogador tóxico(ally), perde consideravelmente mais partidas do que um time sem um jogador tóxico. Isso é um indício claro do que hipotetizamos acima, que o jogador tóxico tem uma influência negativa sobre o desempenho de seu time.

Tempo de jogo dos aliados do ofensor

```
## time.played
## Min. : 404
## 1st Qu.:1562
## Median :2053
## Mean :2076
## 3rd Qu.:2510
## Max. :5362
```

Tempo de jogo dos inimigos do ofensor

```
## time.played
## Min. : 378
## 1st Qu.:1468
## Median :1897
## Mean :1938
## 3rd Qu.:2318
## Max. :4768
```

Existe uma diferença significativa entre o tempo de jogo de aliados do jogador tóxico e inimigos deste quando estes ganham uma partida. Os aliados(ally) atingem um tempo médio de 2013 segundos, contra 1964 segundos

dos inimigos(enemy), de onde nós podemos deduzir que os aliados de um jogador tóxico, mesmo quando ganham uma partida, tem uma dificuldade maior de ganhar, corroborando ainda mais nossa hipótese de que o jogador tóxico influencia negativamente o desempenho do seu time.

## Conclusão

Da nossa análise acima, podemos concluir claramente que o jogador tóxico influencia negativamente no desempenho da equipe. Também podemos concluir que os aliados de um jogador tóxico são suas principais ‘vítimas’, assim fazendo sentido considerá-los como jogadores ‘contaminados’. Atualmente não há nenhuma evidência da influência do jogador tóxico no desempenho do time adversário, então decidimos considerá-los como jogadores ‘limpos’.

Também a partir da análise dos dados, ficou mais claro o que poderia ser obtido através deles, e que objetivos deveríamos almejar. Também conseguimos confirmar importantes hipóteses iniciais, e adquirimos mais confiança de que nossos objetivos finais são sim, possíveis de se conseguir a partir destes dados

Finalmente, gostaríamos de agradecer a (Blackburn et. al. 2014) por ter gentilmente cedido o dataset utilizado neste trabalho.