

# Sentimentos em Sequências de Ações para Mensuração da Capacidade de Mobilização dentro de Perfis de Instituições Públicas de Ensino no Twitter

Guilherme Augusto da Silva Ferreira<sup>1</sup>, Alan Keller Gomes<sup>1</sup>

<sup>1</sup>Instituto Federal de Educação, Ciência e Tecnologia de Goiás (IFG) - Câmpus Inhumas  
Av. Universitária, s/nº, Vale das Goiabeiras, Inhumas-GO – Brasil – CEP: 75400-000

guilhermeadsf@gmail.com, alan.gomes@ifg.edu.br

**Abstract.** *In this paper, supported by Pierre Bourdieu's theory, we combine sentiment analysis within actions sequence learning to measure the mobilization capacity contained in profiles of public educational institutions on Twitter. The results show that the mobilization capacity (or incapacity) is independent of the size of network of connections (number of followers), corroborating Bourdieu's theories. In addition, we present an indicator of the mobilization capacity which considers the number of social interactions (tweets) of each profile.*

**Resumo.** *Neste artigo, sob a luz das teorias de Pierre Bourdieu, a análise de sentimentos é combinada com o aprendizado de sequências de ações para medir a capacidade de mobilização contida nos perfis de instituições públicas de ensino no Twitter. Os resultados mostram que a capacidade (ou incapacidade) de mobilização é independente do tamanho da rede de conexões (número de seguidores), corroborando com as teorias de Bourdieu. Além disso, é apresentado um indicador da capacidade de mobilização que leva em conta o número de interações sociais (tweets) de cada perfil.*

## 1. Introdução

A Sociologia Digital é um ramo de pesquisa que trata tanto dos aspectos e impactos sociais das tecnologias de comunicação digital quanto da aplicação de tecnologias digitais a metodologias de pesquisa em ciências sociais. Sistemas teóricos capazes de apoiar os estudos das redes sociais digitais, tais como o arcabouço teórico de Pierre Bourdieu [Bourdieu 1981][Bourdieu 1986] [Bourdieu et al. 1991], dão suporte a esse novo ramo de estudo, trazendo para as pesquisas em ciências sociais uma abordagem relacional do estudo da vida social [Ignatow and Robinson 2017].

As Redes Sociais Online (RSO) são redes sociais digitais que facilitam o estabelecimento e o fortalecimento de laços sociais no mundo virtual. Essas redes incorporam em suas funcionalidades o caráter lúdico e simétrico que define a sociabilidade nas redes sociais do mundo real, ou seja, o fortalecimento dos laços sociais que se dão tanto sob o aspecto do divertimento, como também, sob o aspecto da igualdade de poder para execução das ações [Santos and Cypriano 2011].

O desenvolvimento de aplicações e pesquisas que utilizam dados de Redes Sociais Online (RSO) enfrentam o desafio da conciliação entre quais dados são permitidos

capturar e a utilidade (significado) desses dados. Como consequência, a elaboração de estratégias direcionadas tanto para a captura de dados relevantes e quanto para a interpretação do resultado do processamento desses dados, é importante e altamente recomendável. Um caminho possível e interessante para a superação desses desafios é a apropriação do arcabouço teórico de Pierre Bourdieu no estudo das RSO [Ignatow and Robinson 2017].

Sob a perspectiva da Sociologia Digital, funcionalidades como curtir, compartilhar, comentar, publicar, apontam para práticas de sociabilidade em rede [Ignatow and Robinson 2017]. Essas práticas são constituídas de formas de comunicação on-line, que por sua vez, formam e dão manutenção na dinâmica das relações entre os agentes (i.e., perfis institucionais) e sua rede de contatos (i.e., seguidores) no campo (i.e., RSO) [Sterne 2003] [Santos and Cypriano 2014].

As práticas de sociabilidade em rede se cristalizam a partir de trocas realizadas no campo [Bourdieu 1981] [Bourdieu et al. 1991]. A realização das trocas advém do reconhecimento mútuo de integrantes da rede de contatos e dependem da capacidade de mobilização do agente [Bourdieu 1986]. Dentro das RSO, as trocas são realizadas a partir da execução de ações como publicar comentar, curtir e compartilhar. Portanto, agentes (perfis institucionais) com capacidade de mobilização da sua rede são capazes de estimular seus seguidores para realizarem trocas a partir da execução de ações (funcionalidades) que são próprias da RSO (campo).

Na pesquisa aqui apresentada, são analisadas práticas de sociabilidade que constam em perfis de instituições públicas de ensino na Rede Twitter. Uma característica peculiar dos perfis institucionais é que o agente que promove as trocas entre sua rede de contatos é o único provedor de publicações de conteúdo dentro da rede. No Twitter, os conteúdos são publicados principalmente em texto com limitação de 280 caracteres.

Em RSO como o Twitter, a Análise de Sentimentos (AS) é uma abordagem comumente empregada para extrair sentimento e/ou emoção expressos nas publicações [Neri et al. 2012]. Esse tipo de análise é empregada principalmente para entender o senso comum sobre um determinado acontecimento, por exemplo, um evento com grande repercussão junto à opinião pública [Kansaon et al. 2019]. Geralmente na AS, investiga-se a polaridade da opinião [Benevenuto et al. 2015], que é referente à atribuição do sentimento, percepção ou atitude que estão expressos em um texto, no caso do Twitter, em um tweet. O sentimento (polaridade de opinião) é classificado como positivo, negativo ou neutro [Ribeiro and Silva 2018].

O presente trabalho tem como objetivo mensurar a capacidade de mobilização contida em perfis de instituições públicas de ensino no Twitter. Dados relativos às práticas de sociabilidade em rede (conteúdo do tweet e ações subjacentes), que ocorrem dentro desses perfis, são coletados. No processamento desses dados, a análise de sentimentos é aplicada a cada tweet coletado e a polaridade do sentimento é inserida como atributo no conjunto de dados.

No passo seguinte, esses dados processados são utilizados na extração automática e na mensuração de interações sociais no formato de sequência de ações. O resultado da mensuração das sequências é utilizado como um indicador da capacidade de mobilização da rede de conexões de cada perfil institucional, que leva em conta o número de interações sociais (tweets).

A principal contribuição desta pesquisa é a abordagem empregada para mensuração da capacidade de mobilização. É importante observar que os dados coletados se referem a informações públicas e não pertencem ao perfil individual de qualquer seguidor dos perfis institucionais. Os dados coletados são mantidos sem identificação do perfil do seguidor que os produziu. O perfil institucional é o único agente que publica tweets, portanto, é o único agente provedor de estímulos às práticas de sociabilidade em rede. Sendo assim, as interações sociais extraídas automaticamente desses dados retratam padrões de comportamento social impessoalizados e coletivos que ocorrem dentro do Twitter.

Este trabalho está organizado da seguinte forma: Seção 2 detalha a metodologia empregada; Seção 3 apresenta e discute os resultados obtidos; Seção 4 elenca trabalhos relacionados; e Seção 5 apresenta as considerações finais.

## **2. Metodologia**

### **2.1. Recorte da pesquisa:**

Dentro de uma RSO, a ação de publicar é ação que inicia uma interação social, também chamada de *estímulo social* [Skinner 1953]. A partir da ação de publicar, as ações como comentar, curtir e compartilhar ficam disponíveis e podem ser executadas pelos demais usuários conectados ao provedor do estímulo social.

Sob a perspectiva da Sociologia Digital [Ignatow and Robinson 2017], práticas de sociabilidade em rede dos agentes (i.e., perfis institucionais) e sua rede de conexões (i.e., seguidores) no campo (i.e., RSO) são cristalizadas a partir de funcionalidades de relacionamento digital (publicar, comentar, curtir, compartilhar).

Para realizar a análise das práticas de sociabilidade em rede, foram selecionadas 4 instituições públicas federais de ensino da região Centro-oeste do Brasil. Cada instituição tem autonomia para gerir seu perfil no Twitter, seguindo diretrizes estabelecidas pelo Governo Federal<sup>1</sup>. O único provedor de estímulos sociais (publicações de tweets) é a instituição de ensino. O público conectado nessas páginas são, em sua imensa maioria, estudantes, professores e servidores da instituição.

### **2.2. Identificação dos dados para captura e Análise de sentimentos:**

O arcabouço teórico de Pierre Bourdieu, no âmbito da Sociologia Digital [Ignatow and Robinson 2017], foi utilizado no direcionamento dos esforços na etapa de identificação dos dados a serem capturados. As funcionalidades de interação e comunicação que apontam para as práticas de sociabilidade em rede, foram identificadas, para esta pesquisa, como sendo as ações de publicar (tweet), curtir, comentar e retuitar. A ação de retuitar equivale a uma ação de compartilhamento dentro da RSO (compartilhamento interno) de um tweet.

Com o objetivo de identificar a polaridade da opinião expressa em cada tweet publicado pelo perfil institucional, foi feita a AS fazendo uso da API IBM Watson [Nasukawa and Yi 2003]. Essa API realiza a classificação do sentimento contido no tweet como positivo, negativo ou neutro. No estudo aqui apresentado não é realizada a análise de sentimentos em comentários ou em retweets.

<sup>1</sup>[www.governodigital.gov.br/transformacao/compras/orientacoes/orientacoes-redes-sociais-1](http://www.governodigital.gov.br/transformacao/compras/orientacoes/orientacoes-redes-sociais-1)

### 2.3. Coleta, transformação e carregamento dos dados:

Na etapa de coleta, transformação e carregamento dos dados identificados anteriormente foi empregado o método ETL (Extract/Transform/Load) [Kimball and Ross 2002]. O acesso aos dados é feito por intermédio do uso de API<sup>2</sup>. Na versão gratuita, a API do Twitter permite a coleta apenas do número de curtidas ( $L$ ) e do número de retweets ( $Rt$ ), ficam de fora da coleta os comentários e retweets com comentários (disponíveis para versões pagas da API).

Scripts em Python<sup>3</sup> foram implementados para o processamento do método ETL e os dados foram armazenados em um banco de dados MySQL<sup>4</sup>. Foram coletados os números máximos de tweets mais recentes (Tweets), permitidos pela versão gratuita da API do Twitter, em 15 de maio de 2019. Foi realizada a AS de cada tweet e a polaridade do sentimento foi armazenada no banco de dados como um atributo do tweet. Uma base de dados foi construída para cada perfil institucional.

Foi necessário transformar os dados capturados via API, e assim, prepará-los para etapa de extração de sequências. Uma restrição imposta pelo algoritmo de extração de sequências é que os dados utilizados sejam não numéricos. Para atender essa restrição, os dados capturados foram transformados e carregados em uma outra base de dados com valores não-numéricos.

### 2.4. Extração e mensuração de sequências das ações:

Uma implementação do algoritmo GSP [Agrawal and Srikant 1995] foi elaborada para o aprendizado automático das sequências das ações. Originalmente, o GSP não manipula valores numéricos e o número de sequências extraídas depende de um limiar (frequência mínima) arbitrada pelo usuário. Na implementação aqui realizada, suporte mínimo considerado para extração de sequências candidatas foi  $\alpha \geq 0$ , portanto, todas sequências que possam constar na base de dados foram extraídas. São geradas sequências de tamanho 1 até as de tamanho igual ao número de colunas da base, sem valores de atributos repetidos. Por fim, o número de ocorrências (frequência) de cada sequência é calculado.

### 2.5. Validação:

A validação dos resultados foi feita a partir de consultas à base de dados elaboradas em Linguagem de Consulta Estruturada - SQL<sup>5</sup>. Nessas consultas, são realizadas contagens da classificação do sentimento de cada tweet juntamente com número de curtidas e retweets iguais a 0 ou diferentes de 0. Havendo coincidência entre a frequência de ocorrência de cada sequência e a contagem retornada pelas consultas SQL, esses resultados são considerados válidos.

## 3. Resultados e Discussão

### 3.1. Resultados da identificação dos dados e do ETL

Como resultado das fases de identificação dos dados a serem capturados no Twitter e aplicação do ETL, foram construídas 4 bases de dados, uma para cada perfil de cada instituição. Nessas bases foram armazenadas informações como nome do perfil, identificador

<sup>2</sup>API Twitter - <https://developer.twitter.com/content/developer-twitter/en.html>

<sup>3</sup><https://www.python.org>

<sup>4</sup><https://www.mysql.com>

<sup>5</sup><https://www.oracle.com/br/database/technologies/appdev/sql.html>

do perfil, identificador da postagem, conteúdo da postagem, tipo de mídia publicada na mensagem, sentimento contido no conteúdo da postagem, data e hora da publicação, número de curtidas, número de compartilhamentos internos (retweets). Essas informações foram coletadas com a versão gratuita da API do Twitter, portanto, ficam de fora da coleta os comentários e retweets com comentários (disponíveis para versões pagas da API).

A primeira base contém dados capturados de 3.218 tweets publicados pelo perfil @IFG\_Goias (Perfil 1). A segunda base contém dados capturados de 1.977 tweets publicados pelo perfil @IFGoiano (Perfil 2). A terceira base contém dados capturados de 3.211 tweets publicados pelo perfil @ufg\_oficial (Perfil 3). A quarta base contém dados capturados de 3.229 tweets publicados pelo perfil @unb\_oficial (Perfil 4). Cada linha dessas bases de dados corresponde ao registro cristalizado de um tweet e as ações subjacentes executadas por seguidores dos respectivos perfis.

Sentimento, Curtida, Retweet positivo, L, !Rt neutro, L, Rt negativo, !L, !Rt !class, !L, !Rt
---

**Listagem 1. Exemplo de Base de Dados não numérica**

Foi necessário realizar mais uma etapa de transformação e carregamento dos dados para que bases com valores não numéricos fossem construídas. Valores maiores que 0 de curtidas foram representados por L e os valores iguais a 0 representados por !L. Os valores maiores que 0 de retweets foram representados por Rt e os valores iguais a 0 representados por !Rt. Além dos valores não numéricos de curtidas e retweets, cada base não numérica foi armazenado o resultado da análise do sentimento contido em cada tweet (positivo, negativo, neutro ou !class = não classificado). Um exemplo de como os dados estão organizados em uma base não numérica é apresentado na Listagem 1. A partir de cada base de dados não numéricos foi realizado o aprendizado de sequência de ações.

### 3.2. Resultados da extração e mensuração das sequências

De cada base de dados não numérica (Perfil 1, Perfil 2, Perfil 3 e Perfil 4) foram extraídas sequências de tamanho 1, 2 e 3. Nas Tabelas 1, 2 e 3, a primeira coluna apresenta as sequências extraídas. Na segunda coluna a frequência de cada sequência para Perfil 1, na terceira coluna as frequências para Perfil 2, na quarta coluna, as frequências para Perfil 3, e na quinta coluna, as frequências para Perfil 4. Os valores percentuais são calculados em relação ao total de tweets capturados em cada perfil.

**Tabela 1. Sequências Tamanho 1: Sentimentos e suas frequências**

Sequência	Perfil 1	Perfil 2	Perfil 3	Perfil 4
< <i>positivo</i> >	1.319 (41%)	860 (43,4%)	1.834 (57,1%)	1.274 (39,45%)
< <i>neutro</i> >	1.778 (55%)	960 (48,5%)	1.172 (36,5%)	1.804 (55,87%)
< <i>negativo</i> >	117	125	202	149
< <i>!class</i> >	4	32	3	2
<b>Tweets</b>	3.218	1.977	3.211	3.229
<b>Seguidores</b>	6.392	4.611	200.967	94.886

Os valores da frequência de sequências de tamanho 1 são apresentados na Tabela 1. Essas sequências permitem verificar a quantidade de tweets classificados de acordo com o sentimento analisado no conteúdo da postagem. A sequência <!class> indica a quantidade de tweets não classificados.

**Tabela 2. Sequências Tamanho 2 e suas frequências**

Sequência	Perfil 1	Perfil 2	Perfil 3	Perfil 4
< L, Rt >	572	102	1.222	2.080
<!L, Rt >	880	289	87	44
< L, !Rt >	425	198	1.448	819
<!L, !Rt >	1.341 (46,67%)	1388 (70,2%)	454 (14,14%)	286 (8,86%)
Tweets	3.218	1.977	3.211	3.229
Seguidores	6.392	4.611	200.967	94.886

Os valores da frequência de sequências de tamanho 2 são apresentados na Tabela 2. Essas sequências permitem verificar a capacidade e a incapacidade de mobilização, sem levar em consideração o sentimento contido em cada tweet.

**Tabela 3. Sentimentos nas Sequências Tamanho 3 e suas frequências**

Sequência	Perfil 1	Perfil 2	Perfil 3	Perfil 4
< positivo, L, Rt >	272	51	798	846
< positivo, !L, Rt >	366	120	51	18
< positivo, L, !Rt >	186	99	783	317
< positivo, !L, !Rt >	495 (15,38%)	590 (29,84%)	202 (6,29%)	93 (2,88%)
Tweets < positivo >	1.319	860	1.834	1.274
< neutro, L, Rt >	280	34	360	1.141
< neutro, !L, Rt >	483	151	30	23
< neutro, L, !Rt >	219	87	579	467
< neutro, !L, !Rt >	796 (24,7%)	688 (34,8%)	203 (6,3%)	173 (5,36%)
Tweets < neutro >	1.778	960	1.172	1.804

Os valores da frequência de sequências de tamanho 3 são apresentados na Tabela 3. Essas sequências permitem verificar a capacidade e a incapacidade de mobilização levando-se em conta o sentimento contido no conteúdo de cada tweet. Nesta tabela, os valores de Tweets <positivo> e Tweets <neutro> correspondem aos valores de <positivo> e <neutro> apresentados na Tabela 1.

### 3.3. Validação dos resultados

A validação dos resultados apresentados nas Tabelas 1, 2 e 3 foi realizada a partir de consultas em SQL feitas em cada uma das 4 bases com dados capturados via API do Twitter. Cada sequência extraída foi mapeada em uma consulta SQL.

```

< positivo, L, Rt >
SELECT COUNT(*) FROM 'perfil' _ WHERE
_____sentimento_=_ 'positivo' AND
AND num_likes <> 0 AND num_retweets <> 0

```

#### Listagem 2. consultas em SQL

Na Listagem 2 é apresentado o mapeamento da sequência de tamanho 3  $\langle \text{positivo}, L, Rt \rangle$ . Cada consulta SQL foi executada em cada base de dados. Os resultados das consultas são coincidentes com os valores das frequências apresentados nas Tabelas 1, 2 e 3. Sendo assim, os resultados foram considerados válidos.

### 3.4. Discussão

Observando-se os valores da Tabela 1, é possível notar que mais de 90% dos tweets publicados em cada perfil tem sentimento (polaridade de opinião) positivo ou neutro. Esse alinhamento na polaridade de opinião é reflexo do cumprimento das diretrizes para tratamento dos conteúdos de mídia social estabelecidas pelo Governo Federal. Além disso, esse alinhamento permite que os valores obtidos pelo indicador da capacidade de mobilização possam ser comparados dentro de um campo (RSO) em que os agentes (perfis institucionais) estão em igualdade de poder [Santos and Cypriano 2011].

De acordo com Bourdieu [Bourdieu 1986], a mobilização de uma rede de contatos se dá a partir de trocas. As trocas instituem o reconhecimento mútuo entre indivíduos em uma rede de contatos. Dentro de uma RSO, trocas são realizadas a partir de funcionalidades que viabilizam a execução de ações como publicar, curtir, compartilhar [Santos and Cypriano 2014]. Portanto, agentes com capacidade de mobilização da sua rede de conexões são capazes de fazer os usuários conectados (seguidores) realizar trocas, executando ações subsequentes a uma publicação. Por outro lado, agentes com incapacidade de mobilização da sua rede de conexões não são capazes de fazer seus seguidores realizarem trocas.

Observando-se a Tabela 2, a incapacidade de mobilização é indicada pela sequência  $\langle !L, !Rt \rangle$ , ou seja, foi realizada uma publicação que não provocou realização de trocas entre a rede de conexões, portanto, não foram executadas ações subjacentes a uma publicação. Com os maiores valores do indicador de incapacidade de mobilização estão, respectivamente, Perfil 2 = 70,2% e Perfil 1 = 46,67% (ou seja, menores valores do indicador da capacidade de mobilização). Com menores valores do indicador da incapacidade de mobilização estão, respectivamente, Perfil 4 = 8,86% e Perfil 3 = 14,14% (ou seja, maiores valores do indicador da capacidade de mobilização). Portanto, os perfis com maior capacidade são Perfil 4 (91,14%) e Perfil 3 (85,86%).

O indicador de capacidade (ou incapacidade) de mobilização, levando-se em conta o sentimento contido no conteúdo de cada tweet, pode ser observado a partir dos valores apresentados na Tabela 3. Tanto observando-se Tweets  $\langle \text{positivo} \rangle$  quanto observando-se Tweets  $\langle \text{neutro} \rangle$ , os perfis com maior capacidade de mobilização são Perfil 4 e Perfil 3, e com menor capacidade de mobilização são Perfil 2 e Perfil 1, confirmando o resultado obtido a partir da análise dos valores da Tabela 2.

Deve ser ressaltado que o indicador de capacidade (ou incapacidade) de mobilização é obtido a partir do número de ocorrências (frequência) das sequências analisadas, portanto, esses indicadores levam em conta o número de interações sociais de cada perfil. Mesmo se o número de interações sociais capturadas de um determinado perfil for discrepante dos demais (por exemplo, Perfil 2 com 1.977 tweets), o indicador respeita a proporcionalidade do número de interações sociais capturadas e permite a comparação entre valores percentuais.

É importante observar o número de interações sociais na leitura do indicador. Por

exemplo, ainda observando o Perfil 2 com 1.977 tweets na Tabela 2, que está alinhado aos demais perfis no que se refere à polaridade de opinião, além de ter um número de interações sociais menor que os demais, 70,2% dessas interações sociais não estimulam a realização de trocas entre os seguidores conectados no perfil. Isso indica que o perfil é pouco utilizado pela instituição na publicação de conteúdo, e ainda, que os seguidores interagem pouco entre si, possivelmente, pelo motivo de que o conteúdo publicado não está sendo interessante para os seguidores.

Ainda segundo [Bourdieu 1986], a capacidade de mobilização de um agente depende do tamanho da sua rede de conexões. Os resultados apresentados na Tabela 2 corroboram esse aspecto da teoria de Bourdieu, pois o Perfil 4 que tem a maior capacidade de mobilização tem um número de seguidores (94.886) menor que o número de seguidores (200.967) do Perfil 3.

Por fim, Bourdieu associa essa capacidade de mobilização ao Capital Social acumulado pelo agente no campo. Em um trabalho futuro, os resultados aqui apresentados serão explorados para se apurar esse tipo de capital.

#### **4. Trabalhos Relacionados**

A pesquisa aqui apresentada está alinhada com trabalhos que analisam interações entre usuários na rede Twitter tendo como suporte o arcabouço teórico de Pierre Bourdieu.

As interações entre usuários provocadas pelo uso de hashtags no Twitter tem sido analisadas com apoio de [Bourdieu et al. 1977]. Por exemplo, [Page 2012] estuda o papel das hashtags no Twitter e [Lindgren and Lundström 2011] analisam os protocolos culturais e sociais da hashtag *#WikiLeaks* no Twitter. Nesses trabalhos não são abordadas em profundidade a capacidade de mobilização dos agentes provedores das hashtags. Além de avançar nesse aspecto, o estudo aqui apresentado se destaca por incorporar a polarização da opinião e aprendizado de sequências de ações na análise da capacidade de mobilização.

Acontecimentos de repercussão junto à opinião pública também são estudados com o suporte das teorias de Bourdieu. Apoiado em [Bourdieu et al. 1977], [Murthy 2015] descobriu em seu estudo que candidatos que usaram o Twitter apenas para causar burburinho, não traduziram seu capital político em sucesso nas urnas. [Recuero et al. 2014] contam com o suporte de [Bourdieu 1977] para compreender o uso do Twitter na narração dos protestos e mobilização de pessoas ocorridos no Brasil em junho de 2013. Em outro estudo, [Recuero 2015] explora o discurso sobre o Dia da Consciência Negra no Brasil, através do Twitter, com apoio em outro trabalho de [Bourdieu et al. 1991]. Novamente, a presente pesquisa se destaca desses estudos por tratar em profundidade a capacidade de mobilização de agentes no Twitter, além de incorporar a polarização da opinião e aprendizado de sequências de ações.

A utilização de aprendizado de máquina para resolver problemas de classificação de dados capturados do Twitter tem sido extensamente estudado nos últimos anos. Por exemplo, no trabalho de [Pennacchiotti and Popescu 2011] são utilizados dados sobre orientação política ou étnica, o comportamento do usuário, a estrutura da rede e o conteúdo linguístico para classificar perfis de usuários do Twitter.

No que se refere a análise de sentimentos utilizando algoritmos de aprendizado de máquina, [Neethu and Rajasree 2013] identificaram o efeito de informações de domínio na classificação de tweets como positivos e negativos, e ainda, na extração da opinião

dos usuários sobre determinados produtos. Um análise dos algoritmos de classificação de sentimentos em tweets em português do Brasil é apresentada em [Kansaon et al. 2019].

Os trabalhos relacionados que utilizam algoritmos de aprendizado de máquina na classificação de dados e análise de sentimentos no Twitter estão voltados para o tratamento de tarefas preditivas de aprendizado. A pesquisa aqui apresentada se destaca por apresentar uma abordagem que incorpora a análise de sentimentos (classificação) no aprendizado de sequências de ações, tratando a análise da capacidade (ou incapacidade) de mobilização como tarefa descritiva de aprendizado de máquina.

## 5. Considerações Finais

Foi apresentado neste trabalho que o arcabouço teórico de Pierre Bourdieu pode ser utilizado no direcionamento da coleta de dados relativos às práticas de sociabilidade em rede, que constam de perfis de instituições públicas federais de ensino na rede Twitter. As informações coletadas são processadas e utilizadas na extração automática e na mensuração de interações sociais no formato de sequência de ações. A mensuração das sequências é utilizada como um indicador da capacidade (ou incapacidade) de mobilização da rede de conexões de cada perfil institucional.

Foi apresentado que os perfis institucionais analisados possuem alinhamento na polaridade de opinião no conteúdo dos tweets publicados. O indicador da capacidade (ou incapacidade) de mobilização leva em conta o número de interações sociais (tweets) de cada perfil. Sendo assim, o indicador respeita a proporcionalidade do número de interações sociais capturadas e permite a comparação de valores percentuais entre perfis que estão em igualdade de poder dentro da rede Twitter.

Apesar de serem preliminares, os resultados obtidos corroboram com a teoria de Bourdieu, mostrando que a capacidade de mobilização de um agente independe do tamanho da sua rede de conexões (número de seguidores).

O escopo da pesquisa aqui apresentada está limitado a análise das ações executadas por seguidores dentro de perfis institucionais no Twitter. Apesar disso, deseja-se que a metodologia empregada possa ser generalizada para ser aplicada na análise da capacidade de mobilização de qualquer agente, dentro de uma RSO, que tenha uma rede de conexões.

## Referências

- Agrawal, R. and Srikant, R. (1995). Mining sequential patterns. In *Proc. of the Inter. Conf. on Data Engineering (ICDE '95)*, pages 3–14. IEEE Computer Society.
- Benevenuto, F., Araújo, M., and Ribeiro, F. (2015). Sentiment analysis methods for social media. In *Proc. of Braz. Symposium on Multimedia and the Web, WebMedia 2015*, pages 11–11, New York, NY, USA. ACM.
- Bourdieu, P. (1977). *A economia das trocas simbólicas*. São Paulo: Perspectiva.
- Bourdieu, P. (1981). Men and machines. In *Advances in Social Theory and Methodology*, pages 304–317. Rout. & Keg. Paul, Boston.
- Bourdieu, P. (1986). The forms of capital. In *Handbook of Theory and Research for the Sociology of Education*, pages 241–258. Greenwood Press, New York.
- Bourdieu, P., Passeron, J.-C., and Nice, R. (1977). Education, society and culture. *Trans. Richard Nice. London: SAGE Pub.*

- Bourdieu, P., Thompson, J., Raymond, G., and Adamson, M. (1991). *Language and Symbolic Power*. Social theory. Harvard University Press.
- Ignatow, G. and Robinson, L. (2017). Pierre bourdieu: theorizing the digital. *Infor., Commun. & Society*, 20(7):950–966.
- Kansaon, D., Brandão, M., and Pinto, S. (2019). Análise de algoritmos de classificação para detecção de emoções em tweets em português brasileiro. *iSys - Revista Brasileira de Sistemas de Informação*, 12(3).
- Kimball, R. and Ross, M. (2002). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*. John Wiley & Sons, Inc., New York, NY, USA.
- Lindgren, S. and Lundström, R. (2011). Pirate culture and hacktivist mobilization: The cultural and social protocols of #wikileaks on twitter. *New Media & Society*, 13(6):999–1018.
- Murthy, D. (2015). Twitter and elections: are tweets, predictive, reactive, or a form of buzz? *Information, Communication & Society*, 18(7):816–831.
- Nasukawa, T. and Yi, J. (2003). Sentiment analysis: Capturing favorability using natural language processing. In *Proceedings of the 2Nd International Conference on Knowledge Capture, K-CAP '03*, pages 70–77, New York, NY, USA. ACM.
- Neethu, M. and Rajasree, R. (2013). Sentiment analysis in twitter using machine learning techniques. In *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, pages 1–5. IEEE.
- Neri, F., Aliprandi, C., Capeci, F., Cuadros, M., and By, T. (2012). Sentiment analysis on social media. In *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 919–926.
- Page, R. (2012). The linguistics of self-branding and micro-celebrity in twitter: The role of hashtags. *Discourse & Communication*, 6(2):181–201.
- Pennacchiotti, M. and Popescu, A.-M. (2011). A machine learning approach to twitter user classification. In *Fifth Inter. AAAI Conference on Weblogs and Social Media*.
- Recuero, R. (2015). Discutindo análise de conteúdo como método: o #diadaconsciencia-negra no twitter. *Cadernos de Estudos Lingüísticos*, 56(2):289–309.
- Recuero, R., Zago, G., and Bastos, M. T. (2014). O discurso dos #protestosbr: análise de conteúdo do twitter. *Galáxia (São Paulo)*, 14(28):199–216.
- Ribeiro, A. P. and Silva, N. F. F. (2018). Um estudo comparativo sobre métodos de análise de sentimentos em tweets. *Revista de Sistemas de Informação da FSMA*, 1(22):35–48.
- Santos, F. C. and Cypriano, C. P. (2014). Redes sociais, redes de sociabilidade. *Revista Brasileira de Ciências Sociais*, 29:63–78.
- Santos, F. C. d. and Cypriano, C. P. (2011). Blogs e wikis: duas formas de colaboração em redes sociais. *Revista Ciência em Movimento*, 13:07 – 19.
- Skinner, B. F. (1953). *Science and Human Behavior*. New York: Macmillan.
- Sterne, J. (2003). Bourdieu, technique and technology. *Cultural Studies*, 17:367–389.