

UNIVERSIDADE DO VALE DO RIO DOS SINOS - UNISINOS
UNIDADE ACADÊMICA DE GRADUAÇÃO
CURSO DE ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

GUILHERME AUGUSTO JAQUES LOPES

**USO DE REDES SOCIAIS PARA A PARTICIPAÇÃO NA GESTÃO PÚBLICA: UMA
PROPOSTA COM BASE NA TÉCNICA DE ANÁLISE DE SENTIMENTOS**

São Leopoldo
2016

GUILHERME AUGUSTO JAQUES LOPES

**USO DE REDES SOCIAIS PARA A PARTICIPAÇÃO NA GESTÃO PÚBLICA: UMA
PROPOSTA COM BASE NA TÉCNICA DE ANÁLISE DE SENTIMENTOS**

Artigo apresentado como requisito parcial
para obtenção do título de Graduado em
Análise e desenvolvimento de sistemas,
pelo Curso de Análise e desenvolvimento
de sistemas da Universidade do Vale do Rio
dos Sinos – UNISINOS.

Orientadora: Prof^a. Ms. Rosemary Francisco

São Leopoldo

2016

USO DE REDES SOCIAIS PARA A PARTICIPAÇÃO NA GESTÃO PÚBLICA: UMA PROPOSTA COM BASE NA TÉCNICA DE ANÁLISE DE SENTIMENTOS

Guilherme Augusto Jaques Lopes^{*}

Rosemary Francisco^{**}

Resumo: A participação popular nos processos decisórios do Governo é essencial em uma democracia, porém, muitas vezes, os cidadãos não conseguem participar nesses processos, ou por falta de conhecimento ou pela burocracia. Um local onde costumam expressar suas opiniões é em redes sociais como o *Twitter*. Este trabalho busca realizar uma pesquisa exploratória para verificar se a técnica de Análise de Sentimentos pode contribuir para uma Gestão Pública mais participativa. O resultado da pesquisa comprovou que é possível utilizar a técnica de Análise de Sentimentos no Orçamento Participativo contribuindo para uma gestão pública mais participativa. Porém, ainda existem lacunas a serem preenchidas como a necessidade de novas pesquisas utilizando dados georeferenciados e também com uma amostra maior de dados.

Palavras-chave: Análise de Sentimentos. Mineração de opinião. Gestão pública participativa. Orçamento Participativo. Redes sociais.

1 INTRODUÇÃO

Os cidadãos estão cada vez mais utilizando as mídias sociais para se comunicarem com seus familiares, amigos, colegas de trabalho, empresas e até mesmo o governo (KAVANAUGH *et al.*, 2012). Como resultado desta utilização, uma grande quantidade de conteúdo é gerado por esses cidadãos e está disponível e acessível, o que representa uma fonte valiosa de dados para identificar a opinião de pessoas sobre diversos assuntos (CERON *et al.*, 2014).

Um problema gerado pela grande quantidade de dados é como será feita a sumarização e organização dos mesmos (LIU, 2010). Neste cenário, é possível utilizar a tecnologia para extrair informações importantes dos textos, como por exemplo o sentimento dos usuários em relação a determinado assunto (CHEN, 2010). Segundo Chen (2010), a Análise de Sentimentos é definida como uma técnica utilizada em mineração de opinião que tem como objetivo identificar sentimentos expressados pelos usuários em seus textos.

^{*} Graduando em Análise e Desenvolvimento de Sistemas pela Universidade do vale do Rio dos Sinos. guiajlopes@gmail.com

^{**} Professora Orientadora.

Uma das áreas onde a Análise de Sentimentos pode ser aplicada é a Gestão Pública, pois, de acordo com Oliveira (2015, p. 44), “um caminho promissor é avaliar quanto a Análise de Sentimentos pode contribuir para promover uma gestão pública mais participativa, estabelecendo uma relação dialógica entre estado e sociedade”.

Atualmente, uma das formas de participação popular é o Orçamento Participativo, onde os cidadãos podem votar em projetos relevantes para seu município, e os que apresentem maior número de votos são executados (PREFEITURA MUNICIPAL DE PORTO ALEGRE, 2011c).

O presente trabalho busca verificar se dados de redes sociais podem contribuir para uma gestão pública mais participativa. Para isso se assume a hipótese de que utilizar dados de redes sociais no Orçamento Participativo contribui para uma gestão pública mais participativa. Para tanto foram desenvolvidos os seguintes objetivos específicos:

- a) Coletar dados no *Twitter* sobre tópicos do Orçamento Participativo;
- b) Selecionar aplicativo para a Análise de Sentimentos e realizar a classificação do conjunto de testes;
- c) Verificar se a Análise de Sentimentos pode ser aplicada no Orçamento Participativo;
- d) Avaliar se a Análise de Sentimentos pode auxiliar na apresentação de informações para tomada de decisão na gestão pública.

Este trabalho trata-se de uma pesquisa exploratória que pode ser classificada como quantitativa no que se refere a identificar a frequência de opiniões positivas, negativas e neutras no conjunto de dados da Análise de Sentimentos e qualitativa no que se diz respeito à coleta e classificação dos dados (RAMBOCAS; GAMA, 2013). A aplicação da técnica de Análise de Sentimentos foi baseada nos estudos de Tsytsarau e Palpanas (2012) e Yoon, Elhadad e Bakken (2013) que consiste nas seguintes etapas: Escolher os temas para Análise de Sentimentos, definir os termos para busca, escolher a fonte dos dados, selecionar aplicativo de mineração de opinião, coletar dados, preparar dados, delimitar conjunto de treino, realizar análise automatizada dos dados e validar os resultados.

O presente trabalho se justifica por uma forte convicção de que as ferramentas de redes sociais podem contribuir para aumentar a transparência e participação popular, além de melhorar os serviços públicos (BONSÓN *et al.*, 2012).

O trabalho, inicialmente, apresenta a revisão de literatura sobre os assuntos necessários para entender ao estudo. Logo após, são apresentados os trabalhos relacionados, método com todos os procedimentos metodológicos aplicados. Por ultimo são apresentados os resultados da pesquisa e uma conclusão sobre o assunto que também indica futuros trabalhos.

2 REFERENCIAL TEÓRICO

Neste capítulo são apresentados os principais conceitos relacionados a esta pesquisa. Serão apresentados assuntos relacionados à gestão pública e à Análise de Sentimentos.

2.1 Gestão Pública participativa

A gestão pública é a área responsável pelo planejamento, implantação e gerenciamento de projetos de políticas públicas (KLOVNISK, 2016). Após a constituição de 1988, criou-se a necessidade de uma maior participação popular nos processos decisórios do governo, visto que a participação é um ponto chave para a renovação da democracia (HELD, 2007).

Para a construção de uma gestão pública democrática é preciso levar em consideração a dimensão econômico-financeira, que trata das finanças públicas e investimentos estatais (PAULA, 2005). Ainda é preciso levar em consideração a dimensão institucional-administrativa que abrange a organização e a articulação dos órgãos que compõem o aparato estatal. Também sendo preciso levar em consideração a dimensão sociopolítica, que trata da relação entre o Estado e a Sociedade, envolvendo a participação dos cidadãos na gestão pública.

O presente trabalho busca auxiliar na relação entre o Estado e a Sociedade por meio da tecnologia. Atualmente uma das maneiras que o cidadão pode interagir com o Estado é através do OP (Orçamento Participativo), onde o cidadão vota onde o dinheiro deve ser investido, selecionando projetos (PREFEITURA MUNICIPAL DE PORTO ALEGRE, 2011b).

2.1.1 Orçamento participativo

O OP é um processo democrático, onde os cidadãos podem definir de forma direta a aplicação de recursos em obras e serviços que serão executados pela administração local. São realizadas reuniões durante as quais a população local define as prioridades para o Município (PREFEITURA MUNICIPAL DE PORTO ALEGRE, 2011b).

O OP é um tema que ganhou grande relevância, principalmente após o OP de Porto Alegre, que despertou interesse nacional e internacional de acadêmicos e gestores públicos (NOVAES; SANTOS, 2014). Atualmente, o OP é uma das principais ferramentas de participação popular nos processos decisórios do governo (PREFEITURA MUNICIPAL DE PORTO ALEGRE, 2011b).

Em Porto Alegre, os cidadãos podem escolher entre 17 prioridades, que são: Saneamento Básico, Política Habitacional, Pavimentação, Transporte e Circulação, Saúde, Assistência Social, Educação, Áreas de Lazer, Esporte e Lazer, Iluminação Pública, Desenvolvimento Econômico, Cultura, Saneamento Ambiental, acessibilidade e mobilidade urbana e juventude. Cada região pode escolher quais são as prioridades. Inicialmente, cada região vota em quatro prioridades, as prioridades globais, atribuindo notas a elas. Após as 17 regiões votarem as prioridades, a soma das notas são as 3 primeiras prioridades do município, que serão as primeiras a serem implementadas no Município (PREFEITURA MUNICIPAL DE PORTO ALEGRE, 2011a).

No ano de 2006, a Prefeitura de Belo Horizonte lançou o OP Digital, com o objetivo de alcançar um número maior de pessoas. O único pressuposto para participar do OP digital é ser eleitor da cidade (SAMPAIO; MAIA; MARQUES, 2010). Com o OP Digital o número de participações aumentou bastante em Belo Horizonte, tanto em 2006 como em 2008, uma vez que é uma forma mais simples de participar nos processos decisórios do governo (SAMPAIO; MAIA; MARQUES, 2010).

2.1.2 Gestão pública e as Redes Sociais

As redes sociais abriram novas possibilidades para o engajamento público na gestão pública (LEE; KWAK, 2012). Além disto, os cidadãos estão confiando mais em utilizar as redes sociais para comunicação com familiares, colegas e com o próprio governo (KAVANAUGH *et al.*, 2012). O Governo tem buscado nesses meios encontrar recursos para melhorar os serviços e as comunicações com os cidadãos,

especialmente em segmentos onde, anteriormente, era difícil a comunicação com a população (BERTOT; JAEGER; HANSEN, 2012).

As redes sociais também começaram a impactar fortemente as decisões políticas no mundo real. Começaram a ser usadas para organizar manifestações, como, por exemplo, a Arab Spring¹ (COTTLE, 2011), para engajar indivíduos em manifestações e também para construir movimentos sociais e partidos políticos, como o Partido Pirata na Suíça e na Alemanha, ou o Movimento Italiano 5 Stelle que usou as redes sociais para decidir a linha política e também selecionar candidatos (SEGERBERG; BENNETT, 2011).

Porém, com o grande número de dados gerados nessas redes sociais - por exemplo, o *Twitter*, em 2015, teve uma média de 500 milhões de *tweets* por dia (STECYK, 2015) - torna-se difícil para um ser humano extrair conteúdos relevantes, sumará-los e organizá-los (LIU, 2010). Nesse contexto, existem algumas tecnologias e técnicas que poderiam ajudar, como a técnica abordada no próximo tópico.

2.2 Análise de sentimento

A Internet mudou a maneira como as pessoas expressam suas visões e opiniões, hoje elas podem publicar revisões de produtos, compartilhar opinião em fóruns e grupos de discussão (LIU, 2010). Estas discussões geram uma grande quantidade de dados, o que representa uma fonte valiosa para identificar opinião de pessoas sobre diversos assuntos. Com isto surge uma nova necessidade, uma vez é muito difícil para um ser humano extrair conteúdos relevantes destas fontes de dados, assim como organizá-las e sumará-las (CERON *et al.*, 2014). Para resolver este problema surge a análise de sentimento, também conhecida como mineração de opinião, é a área que tem como objetivo identificar sentimentos expressos por usuários em seus textos (CHEN, 2010). A análise destes textos resulta na polaridade da opinião que é basicamente sua orientação e normalmente podem ser classificadas em três grupos distintos: positivos, negativos e neutros (LIU, 2010).

Atualmente, existem duas abordagens que se pode utilizar na Análise de Sentimentos. Uma delas é a classificação automática, utilizando a análise léxica, a

¹ Foi uma onda revolucionária contra os governos locais marcada por protestos tanto pacíficos como não pacíficos que começou na Tunísia e se espalhou por países da Liga Árabe.

qual trabalha com dicionários de palavras previamente classificadas como positivas ou negativas (GHIASSI; SKINNER; ZIMBRA, 2013). Já a outra abordagem é a classificação manual de um conjunto de dados, que é chamado de conjunto de treino, esta classificação servirá como base para toda a classificação da base de dados, este método utiliza algoritmo de aprendizagem de máquina para auxiliar a classificação (LIU; ZHANG, 2012).

Algumas diferenças guiam a escolha da abordagem de classificação manual com aprendizagem de máquina ou a totalmente automatizada com a análise léxica. A seguir serão apresentados os conceitos de aprendizagem de máquina e análise léxica e, logo após, um comparativo entre ambos.

2.2.1 Aprendizagem de Máquina

Aprendizagem de máquina é uma subárea de inteligência artificial que busca realizar reconhecimento de padrões. Pode ser dividida em dois tipos, a aprendizagem supervisionada e a aprendizagem não supervisionada (MILAGROS *et al.*, 2016).

Essencialmente a aprendizagem supervisionada busca replicar as decisões de codificação dos codificadores humanos, desta forma, os pontos fortes da codificação são preservados juntamente com a sensibilidade contextual e as sutilezas da linguagem humana (LEWIS; ZAMITH; HERMIDA, 2013).

Existem diferentes tipos de algoritmos supervisionados, como, por exemplo, *Support Vector Machines* (SVN), *Naive Bayes* e *Maximum Entropy* que podem ser utilizados para problemas de classificação (PANG; LEE; VAITHYANATHAN, 2002). Segundo Pang, Lee e Vaithyanathan (2002), dentre os 3 métodos, o que se mostrou mais eficiente foi o SVN, e o menos eficiente foi o *Naive Bayes*, apesar de ter uma diferença mínima. Neste estudo também nota-se que a aprendizagem de máquina apresenta resultados satisfatórios em relação à classificação humana.

Aprendizagem não supervisionada é aquela que não necessita da interação com humanos para treinar a ferramenta. Ela pode ser utilizada com a análise léxica nos seguintes passos: Extrair frases contendo adjetivos ou advérbios, fazer a estimativa semântica de cada frase e classificar a opinião (TURNERY, 2002).

A Aprendizagem não supervisionada também utiliza a Análise Léxica para realizar a classificação dos textos. Além disso, pesquisas comprovam que a

aprendizagem não supervisionada apresenta um resultado inferior à aprendizagem supervisionada, isso se deve principalmente pelo fato dela utilizar dicionários estáticos e não um conjunto aplicado no contexto de uso (TURNERY, 2002).

2.2.2 Análise Léxica utilizando dicionário de palavras

A Análise Léxica tenta suprir o problema do treinamento da máquina utilizando um documento chamado dicionário de palavras, que mostra palavras ou frases consideradas positivas ou negativas (SAIF *et al.*, 2016). Existem alguns dicionários já escritos e disponíveis, cujas palavras podem ser aplicadas em diferentes contextos, como, por exemplo, o *SentiWordNet* (BACCIANELLA; ESULI; SEBASTIANI, 2010).

Uma limitação deste modelo é que ele tem uma completa dependência de palavras ou ferramentas sintáticas que reflitam explicitamente os sentimentos, porém, em muitos casos, os sentimentos das frases estão associados com o seu contexto semântico (CAMBRIA, 2013).

A seguir são listadas as principais diferenças entre este método e o método de Aprendizagem de Máquina supervisionada assim como a justificativa da escolha do método para o presente trabalho.

2.2.3 Análise Léxica utilizando dicionário de palavras x Aprendizagem de Máquina Supervisionada

A aprendizagem de máquina na Análise de Sentimentos pode ser aplicada de duas formas: supervisionada e não supervisionada. A aprendizagem supervisionada apresenta melhores resultados, uma vez que passa por um conjunto de treinamento inicial que pode ajudar a identificar, por exemplo, gírias sobre o assunto específico (MILAGROS *et al.*, 2016).

Comparando a aprendizagem de máquina supervisionada com a análise léxica utilizando dicionário de palavras, nota-se algumas vantagens, a principal delas é justamente o conjunto de treinamento, que apresenta resultados mais precisos, pois é baseado em um contexto específico (TURNERY, 2002). Por outro lado, a análise léxica utilizando dicionário de palavras não necessita de um tempo para realizar o treinamento inicial, o que a torna uma implementação mais rápida e

totalmente automatizada, sendo esta uma das suas principais vantagens (SAIF *et al.*, 2016). O quadro 1 apresenta as principais características das duas modalidades e se ela é ou não implementada.

Quadro 1 - Análise léxica com dicionário de palavras x Aprendizagem de máquina supervisionada

Características	Análise Léxica com Dicionário de Palavras	Aprendizagem de Máquina Supervisionada
Necessidade de treinamento	Não	Sim
Contexto específico de classificação	Não	Sim
Classificação totalmente automatizada	Sim	Não

Fonte: Elaborado pelo autor.

Para o presente trabalho foi selecionada a abordagem de aprendizagem de máquina supervisionada, a qual é capaz de alcançar uma maior precisão para o caso específico.

2.2.4 *Twitter* como fonte de dados

O *Twitter* é uma plataforma de comunicação popular e em rápido crescimento. Os usuários do *Twitter* criam mensagens chamadas de *tweets* para se comunicar com outros usuários por diversas razões e sobre diferentes tópicos. Esses *Tweets*, muitas vezes, contêm informações valiosas sobre as perspectivas e opiniões desses usuários, relacionadas a empresas e a sociedade em geral (GLEASON, 2013).

Tweets são pequenas unidades de texto, de no máximo 140 caracteres, particularizados por serem casual, linguagem compacta com um alto número de gírias, abreviações e *emoticons*², o que pode complicar a técnica de Análise de Sentimento utilizando ferramenta (GHIASSI; SKINNER; ZIMBRA, 2013).

A Análise de Sentimentos, utilizando *Twitter* produziu resultados inconsistentes, o que fez os pesquisadores questionarem a sua efetividade e ficarem cautelosos em relação aos seus resultados (WONG; SEN; CHIAN, 2012). Porém,

² É uma sequência de caracteres ou imagem que traduz ou quer transmitir o estado psicológico ou emotivo de uma pessoa.

uma das vantagens do *Twitter*, que o fez ganhar maior popularidade na Literatura de Análise de sentimentos em relação às outras redes sociais, é que normalmente o acesso aos *Tweets* dos usuários são grátis, sendo possível obtê-los através de uma Interface de Programação de Aplicação (API) (CERON *et al.*, 2014). Outro ponto que fez o *Twitter* se destacar é que sua API tornou fácil a coleta de um grande número de *Tweets*, literalmente centenas de milhões de mensagens (LOTAN *et al.*, 2011).

Um mecanismo interessante para classificar os *Tweets* são as *hashtags*, que são tópicos que começam com o caractere #. Quando essas *tags* são muito comentadas e com um grande número de usuário realizando *posts* com elas, o *Twitter* promove-as aos *Trending Topics*³, conforme ilustrado na Figura 1 (MARTÍNEZ-CÁMARA *et al.*, 2014).

Figura 1 - Trending Topics



Fonte: *Twitter.com* (2016).

Com as *hashtags* é possível encontrar conteúdos sobre diversos assuntos com maior facilidade (COTTLE, 2011).

Existem algumas ferramentas capazes de coletar dados desta plataforma, como, por exemplo, a ferramenta presente no próximo capítulo.

2.2.5 *DiscoverText*

DiscoverText é uma ferramenta na plataforma Web, criada pela *Texifter*, que permite aos usuários realizar análise de textos provenientes de diversas fontes de dados, como, por exemplo, redes sociais, e-mails e arquivos de textos. (TEXIFTER, 2011). Além de analisar textos de mídias sociais e uma grande variedade de outras

³ Trending topics são uma lista em tempo real das palavras mais utilizadas no *Twitter* em todo o mundo.

fontes, o *software* é projetado para melhorar os processos de pesquisa, governamentais e de empresas (SHULMAN, 2011).

Com o *DiscoverText* é possível realizar a classificação dos textos utilizando Aprendizagem de Máquina, assim como realizar a classificação manual do conjunto de treino (SHULMAN, 2014).

O *DiscoverText* possui uma versão gratuita para acadêmicos e pesquisadores o que facilita a utilização em trabalhos de pesquisa científica, além disso, outros trabalhos de análise de sentimentos também o utilizaram e obtiveram resultado satisfatório (OLIVEIRA, 2015).

2.5 Trabalhos relacionados

Oliveira (2015) realizou uma pesquisa dividida em três artigos, onde buscou utilizar a técnica de Análise de Sentimentos na Gestão Pública e Política. Primeiramente, ele realizou uma coleta de dados no *Twitter* e *Facebook*, a partir dos quais comprovou que é possível expressar opiniões sobre diversos assuntos relacionados à administração pública. No segundo artigo ele aplicou os dados de redes sociais na Gestão Social, onde foram coletados dados do *Twitter* sobre os principais projetos do governo como, por exemplo, Bolsa Família. Logo após foi verificado se as estratégias de mídias sociais (*Push*, *Pull*, Rede, Transação) atendem aos princípios de gestão social e se podem ser aplicadas por meio da técnica de Análise de Sentimentos. No último artigo ele aplicou as técnicas de Análise de Sentimentos para verificar a opinião política sobre os principais candidatos das eleições de 2014, obtendo um resultado satisfatório quando comparado com as principais pesquisas de opinião.

No trabalho de Oliveira (2015) não foram levados em consideração modos de participações tradicionais como o orçamento Participativo e é nesta área que o presente trabalho se mostra diferente, uma vez que a Análise de Sentimentos é focada nos tópicos do Orçamento Participativo e busca verificar se é possível aplicar seguindo os principais requisitos pesquisados na literatura.

Tumasjan *et al.* (2010) utilizou o *Twitter* para realizar uma predição eleitoral. Em seu trabalho os autores mostram que o *Twitter* pode expressar opiniões dos cidadãos sobre assuntos políticos. Também descobriram que o *Twitter*, além de mostrar a opinião dos usuários, mostra discussões sobre assuntos políticos.

O estudo de Tumasjan *et al.* (2010) é relacionado com o presente trabalho uma vez que são utilizados dados do *Twitter* para descobrir opiniões políticas, porém eles têm um enfoque um pouco diferente, uma vez que aquele busca realizar uma predição eleitoral, enquanto este busca aplicar os dados principalmente no apoio ao Orçamento Participativo.

Ceron *et al.* (2014) utilizaram redes sociais como *Twitter* e *Facebook* para descobrir as preferências eleitorais dos cidadãos nas eleições da Itália e França. Com isso eles conseguiram descobrir que as preferências *online* tendem a reagir a fatores externos como, por exemplo, Jornais. Também descobriram que quanto mais usuários falam sobre um assunto, maior é a precisão dos resultados em relação a Análise de Sentimentos.

O estudo de Ceron *et al.* (2014) é relacionado com o presente trabalho, uma vez que realiza a Análise de Sentimentos em redes sociais e conclui que a opinião dos usuários de redes sociais podem ser influenciada por fatores externos. A diferença é que o presente trabalho busca aplicar os dados provenientes de redes sociais na gestão pública, enquanto o trabalho de Ceron *et al.* (2014) busca somente verificar se a opinião dos usuários é influenciada por fatores externos. Além disso Ceron *et al.* (2014) aplicam em predição eleitoral, enquanto o presente trabalho busca aplicar os dados na gestão pública.

Portanto, nota-se que existem vários trabalhos que utilizam redes sociais como *Twitter* para coletar as opiniões dos usuários sobre diversos assuntos, como, por exemplo, gestão pública e predição política. O Quadro 2 sumariza o que foi feito em cada trabalho relacionado.

Quadro 2 - Comparativo trabalhos relacionados

Trabalho	Fonte dos dados	Tema principal
Presente trabalho	<i>Twitter</i>	Verificar se é possível ter uma gestão pública mais participativa através da aplicação de Análise de Sentimentos no orçamento participativo.
Ceron et al. (2013)	<i>Twitter e Facebook</i>	Verificar se a opinião dos usuários é influenciada por fatores externos em uma predição eleitoral utilizando Análise de Sentimentos.
Tumasjan et al. (2010)	<i>Twitter</i>	Verificar se é possível utilizar Análise de Sentimentos para realizar uma predição eleitoral
Oliveira (2015)	<i>Twitter e Facebook</i>	São 3 trabalhos, o primeiro o foco é verificar se a técnica de Análise de Sentimentos pode ser utilizada em tópicos da Administração pública. O segundo é verificar se a Análise de Sentimentos pode contribuir para as práticas de gestão social. E o terceiro é uma predição eleitoral

Fonte: Elaborado pelo autor.

3 MATERIAIS E MÉTODOS

O presente trabalho pode ser classificado quanto sua natureza como uma pesquisa aplicada e quanto seus objetivos como uma pesquisa exploratória. Já quanto a abordagem, esta pode ser classificada como qualitativa no que se diz respeito à coleta e classificação dos dados, e quantitativa no que se refere a identificar a frequência de opiniões positivas, negativas e neutras no conjunto de dados da Análise de Sentimentos (RAMBOCAS; GAMA, 2013).

A Análise de Sentimentos exige que sejam seguidas algumas etapas, que são: coleta, codificação e processamento dos dados. Não existe nenhum padrão definido na literatura para realizar a Análise de Sentimentos já que os estudos publicados apresentam métodos e técnicas bastante distintos. Neste trabalho a aplicação da técnica de Análise de Sentimento foi baseada nos estudos de Tsytarau e Palpanas (2012) e Yoon, Elhadad e Bakken (2013) que consistem, basicamente, nas seguintes etapas:

1. Escolher os temas para Análise de Sentimentos;
2. Definir os termos para a busca;
3. Escolher a fonte dos dados;
4. Selecionar aplicativo de mineração de opinião;
5. Coletar dados;

6. Preparar dados;
7. Delimitar conjunto de treino;
8. Realizar análise automatizada dos dados;
9. Validar os resultados.

Os temas para análise de sentimento (etapa 1) foram selecionados com base nos tópicos do orçamento participativo que são: saneamento básico, política habitacional, pavimentação, transporte e circulação, saúde, assistência social, educação, áreas de lazer, esporte e lazer, iluminação pública, desenvolvimento econômico, cultura, saneamento ambiental, acessibilidade e mobilidade urbana e juventude (PREFEITURA MUNICIPAL DE PORTO ALEGRE, 2011a). Dentre estes tópicos foram selecionados 3 que são mais discutidos nas redes sociais, para realizar uma pesquisa exploratória com objetivo de verificar a disponibilidade de utilização da técnica, que são: saúde, educação e transporte público.

Para selecionar os termos para busca (etapa 2), inicialmente, foi utilizada somente a técnica de seleção utilizando *Hashtags*, porém esta técnica não trouxe resultados satisfatórios para alguns tópicos, como, por exemplo, saúde. Muitos *tweets* utilizando a *hashtags* saúde não eram relacionados com a saúde pública. Como pode ser observado no seguinte *Tweet*: “Treino de hoje com o melhor Huhuhu @alepunheiro_personal @eletrofit #drchoque Tá ficando top #saúde” (Usuário A, Twitter, 20/09/2016). Por este motivo, partiu-se para uma abordagem mista, utilizando *hashtags* e também palavras chaves que identificam o tópico. Os termos de busca por tópico estão definidos no Quadro 3.

Quadro 3 - Palavras chaves por tópicos

Tópico	Palavras chaves
Saúde	Saúde, hospitais, #saude, “saude”
Educação	Educação, escolas, #educacao, #educação
Transporte	#transporte, ônibus, trem

Fonte: Elaborado pelo autor.

A fonte dos dados utilizada (etapa 3) foi o *Twitter*, uma vez que é uma ferramenta bastante acessada na internet, tem uma API aberta para consulta aos *Tweets* públicos e possui *Tweets* com, no máximo, 140 caracteres, o que facilita a técnica de Análise de Sentimentos.

A seleção do aplicativo (etapa 4) para realizar a mineração de opinião baseou-se em alguns requisitos básicos. Primeiramente, o aplicativo deveria permitir o tratamento de dados na língua portuguesa e ser capaz de extrair e classificar automaticamente dados do *Twitter* a partir de termos de pesquisa pré-definidos. Além disto, o aplicativo deveria ser gratuito para facilitar futuras pesquisas na área. Nessa etapa foram avaliados o *DiscoverText* e o *RapidMiner*. Ambos possuem licenças gratuitas e integração com o *Twitter*, porém o *RapidMiner* tem uma limitação de número de *Tweets* de apenas 10 mil (RAPIDMINER, 2016), enquanto o *DiscoverText* tem uma limitação de 100 mil. Considerando esses requisitos e limitações, o aplicativo escolhido foi o *DiscoverText* que disponibiliza uma versão gratuita de 30 dias. O aplicativo também se mostrou eficiente e com uma performance satisfatória em pesquisas anteriores (BEYER, 2012; OLIVEIRA, 2015).

A coleta de dados (etapa 5) foi realizada através do aplicativo e o período utilizado para coleta de dados foi entre o dia 20 de setembro de 2016 e 10 de outubro de 2016. Durante esse período houve a coleta total de 82.756 *Tweets*. Os dados brutos por categoria podem ser visualizados na coluna Dados Brutos da Tabela 1.

Após essa etapa foi realizada a preparação dos dados (etapa 6), onde foi utilizada uma ferramenta disponibilizada pelo *DiscoverText* para eliminar os *Tweets* duplicados. Posteriormente, foi realizada uma seleção manual de *Tweets* que realmente expressassem opiniões de cidadãos Brasileiros sobre as áreas públicas, a fim de eliminar ruídos na análise dos dados. Nesta etapa, *Tweets* de meios de comunicação, *Tweets* de outras línguas e *Retweets* foram eliminados. A relação entre os dados brutos e dados processados pode ser observada na Tabela 1 – Relação dados brutos x dados processados por tema.

Tabela 1 - Relação dados brutos x dados processados por tema

Tema	Dados Brutos	Dados processados
Saúde	30.665	458
Educação	18.297	246
Transporte	33.794	521

Fonte: Elaborado pelo autor.

Pode-se observar que nesta etapa houve uma grande redução dos dados brutos para os dados processados, sendo que em outras pesquisas essa relação já foi observada (OLIVEIRA, 2015). Isto pode ser explicado pelo grande número de *Tweets* que não são relacionados aos tópicos, além da grande quantidade de *Tweets* duplicados.

Para esses dados serem utilizados no orçamento participativo, é necessário que os mesmos sejam divididos por cidades. Isto é possível através da *metatag* “*user_location*” disponibilizada pela API do *Twitter*. Porém, ao filtrar os dados processados por esta *metatag*, não se obteve uma quantidade significativa de dados para aplicar a técnica de Análise de Sentimentos. Para comprovar isso, foi selecionada a maior cidade do Brasil em termos de população que é São Paulo como filtro por localização e os resultados podem ser observados na coluna Dados de São Paulo na Tabela 2. Isso ocorre por que somente *Tweets* feitos via dispositivos móveis têm a marcação de localização e também por que essa *metatag* não apresenta resultados únicos, ela pode apresentar dados do Estado, Cidade e até mesmo Bairro. Por este motivo a Análise de Sentimentos foi realizada em cima do conjunto total de dados processados e não dos dados geolocalizados.

Tabela 2 - Relação Dados brutos x Dados processados de São Paulo por tema

Tema	Dados Brutos	Dados de São Paulo
Saúde	30.665	17
Educação	18.297	3
Transporte	33.794	22

Fonte: Elaborado pelo autor.

O conjunto de treino (etapa 7) foi definido em 30% do total de *tweets* por tópico, uma vez que este conjunto de treinamento, segundo a literatura, já é suficiente para alcançar uma assertividade satisfatória na classificação automatizada utilizando-se a técnica de Aprendizagem de Máquina (OLIVEIRA, 2015). O número de itens treinados por tópico pode ser observado na Tabela 3.

Tabela 3 - Conjunto de treino

Tema	Conjunto de treino
Saúde	137
Educação	75
Transporte	157

Fonte: Elaborado pelo autor.

A próxima etapa foi realizar a classificação manual do conjunto de treino, que significa marcar os *tweets* como positivos, negativos ou neutros. Nessa etapa foram classificados manualmente cada conjunto de treino, com o objetivo de estabelecer regras e padrões para a aprendizagem de máquina. Foram utilizados três conjuntos diferentes um para cada tema com o objetivo de alcançar o melhor resultado possível e definir padrões para o tópico em questão. A figura 2 ilustra como funciona essa classificação dentro do aplicativo *DiscoverText*.

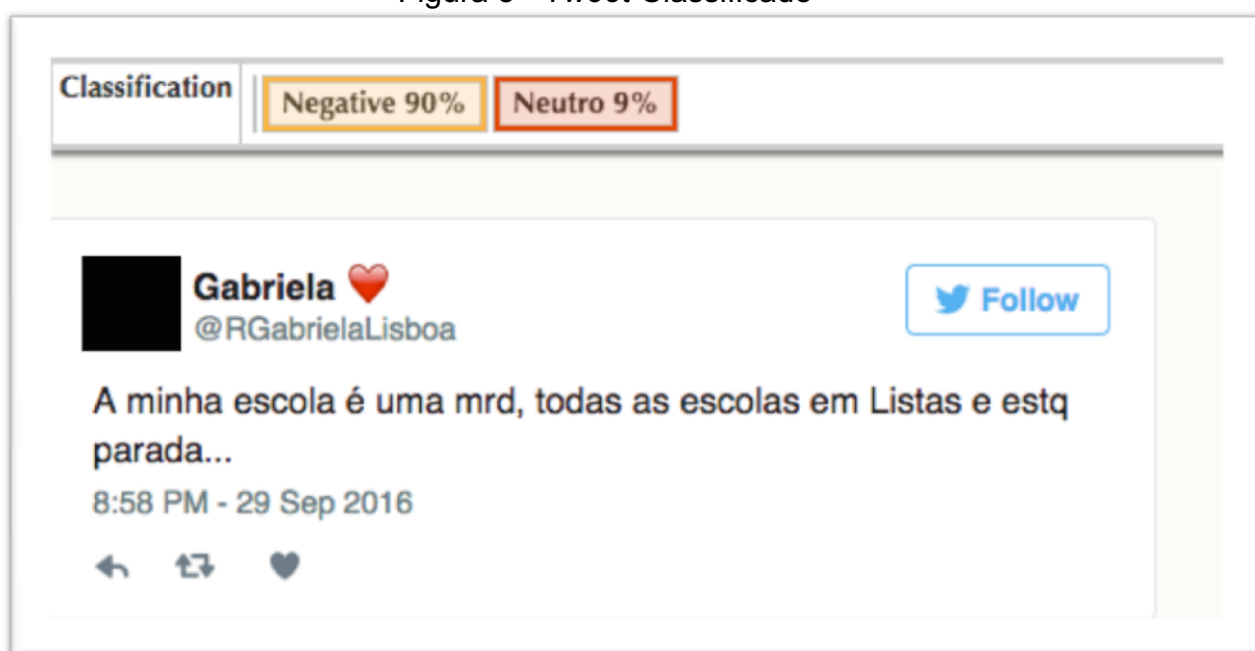
Figura 2 – Polarização



Fonte: DiscoverText.com (2016).

Terminada a classificação manual de cada conjunto de treino, foi realizada a classificação automatizada (etapa 8) de todos os conjuntos de dados, gerando o sentimento positivo, negativo ou neutro sobre cada um dos 3 tópicos selecionados para o orçamento participativo. Esta etapa foi realizada pelo próprio *DiscoverText* que disponibiliza uma ferramenta chamada *Active Learning*, onde é possível realizar esta classificação. A figura 3 ilustra como um dado fica após a classificação automatizada, pode-se observar que é atribuído o percentual de cada polaridade.

Figura 3 - Tweet Classificado



Fonte: DiscoverText.com (2016).

Após a classificação automatizada, foi realizada a validação dos resultados (etapa 9) para saber o nível de precisão do algoritmo. A precisão dos resultados foi calculada com base em uma amostra aleatória simples (COCHRAN, 1997) dos dados classificados de forma automatizada. O tamanho da amostra manteve a mesma proporção do conjunto de treino que foi 30% dos dados brutos. O cálculo para descobrir a precisão foi baseado no modelo de Eirinaki, Pital e Singh (2012) dividindo-se o número total de classificações corretas pelo número total de cada amostra, como pode ser observado na Tabela 4.

Tabela 4 - Precisão

Tópico	Dados classificados automaticamente	Amostra para cálculo de precisão	Dados classificados corretamente	Precisão (%)
Saúde	321	137	112	81,75%
Educação	171	75	59	78,66%
Transporte	364	157	132	84,07%

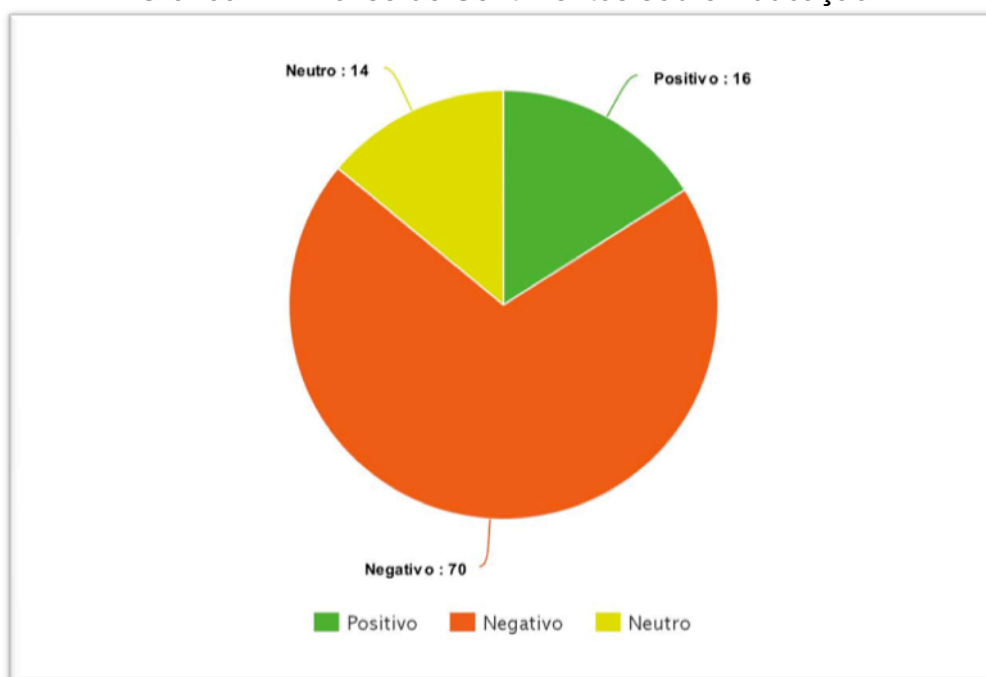
Fonte: Elaborado pelo autor.

A precisão foi calculada realizando uma divisão simples entre a coluna dados classificados corretamente pela coluna Amostra para cálculo de precisão da Tabela 4 multiplicado por 100. (Dados classificados corretamente / Amostra para cálculo de precisão) x 100.

4 ANÁLISE DE DISCUSSÃO DOS RESULTADOS

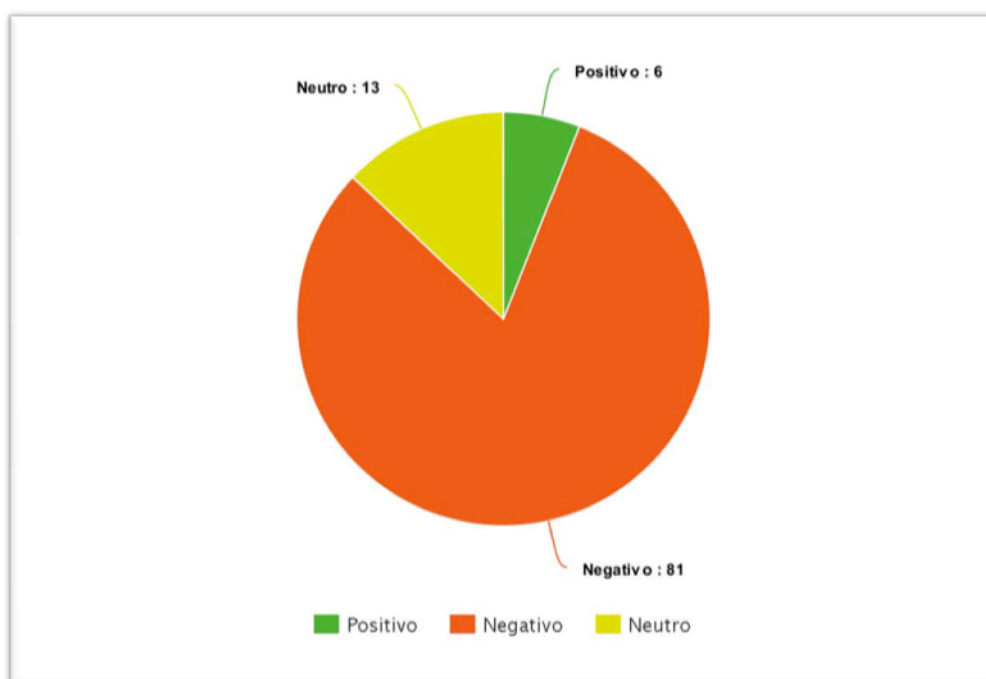
Os resultados da classificação automatizada dos *tweets* mostraram que houve a predominância de opiniões negativas em todos os três tópicos analisados (saúde, educação e transporte público). Educação e transporte público tiveram o mesmo percentual de opiniões positivas, enquanto saúde obteve o menor percentual de aprovação. Os dados de cada um dos tópicos podem ser observados nos gráficos 1, 2, 3. Esses gráficos expressam o percentual de cada uma das três polaridades que, segundo Liu (2010), é a orientação das opiniões e pode ser classificada em três grupos distintos: positivas, negativas e neutras.

Gráfico 1 - Análise de Sentimentos sobre Educação



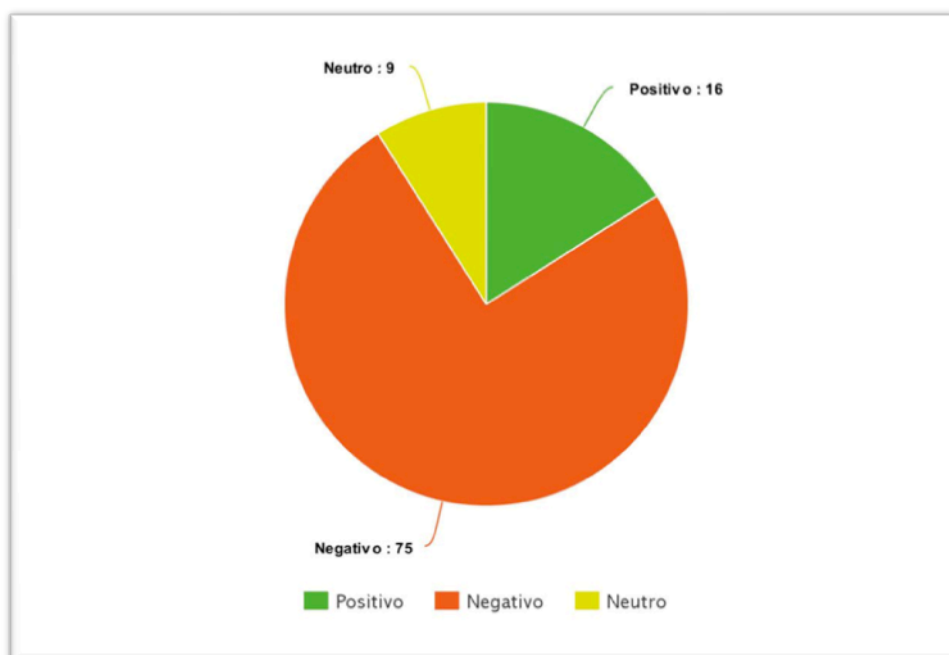
Fonte: Elaborado pelo autor.

Gráfico 2 – Análise de Sentimentos sobre Saúde



Fonte: Elaborado pelo autor.

Gráfico 3 - Análise de Sentimentos Transporte Público



Fonte: Elaborado pelo autor.

Conforme os resultados da Análise de Sentimentos, a área que apresentou maior número de desaprovação foi a Saúde, totalizando 81% do total e com apenas 6% de aprovação do total de *Tweets*. Já o tópico que teve maior nível de aprovação foi a educação, com 16% do total e com 70% de opiniões negativas sobre o assunto. O transporte público manteve-se entre as duas outras, com um total de 75% de opiniões negativas e 16% de opiniões positivas. Exemplos de *tweets* classificados pela técnica de Análise de Sentimentos neste estudo são apresentados no Quadro 4.

Quadro 4 - Exemplos de *Tweets*

Polaridade	<i>Tweets</i>
Positivo	“Ônibus com ar condicionado é outro nível” (Usuário B, <i>Twitter</i> , 29/09/2016)
Negativo	“Só vejo esses políticos falando de escolas integrais. Onde fica a estrutura, refeição, professores e tempo para planejamento?” (Usuário C, <i>Twitter</i> , 29/09/2016)
Neutro	“O estado tem que cuidar de saúde e educação. Só” (Usuário D, <i>Twitter</i> , 29/09/2016)

Fonte: Elaborado pelo autor.

Após os resultados da técnica e Análise de Sentimentos sobre os três tópicos do orçamento participativo em questão, resta saber se essa técnica pode realmente contribuir para uma gestão pública mais participativa e se ela pode ser aplicada no Orçamento Participativo. Essa questão pode ser respondida pelo confronto entre conceitos e aplicações do Orçamento Participativo e Análise de Sentimentos.

Através de dados de redes sociais em conjunto com o orçamento participativo foi possível comprovar alguns pilares do Orçamento Participativo. Conforme o Quadro 4 a aplicação da Análise de Sentimentos possibilita verificar a opinião de cidadãos sobre diversos tópicos do Orçamento participativo, neste estudo foram selecionados três (saúde, educação e transporte público) com o intuito de testar a técnica. Entretanto é necessária cautela ao utilizar estes dados, uma vez que apresentam somente opiniões de usuários do *Twitter* e são limitados ao período da coleta de dados que é entre 20 de setembro de 2016 e 10 de outubro de 2016. Além disto a precisão dos resultados da Análise de Sentimentos não é 100% o que representa outra limitação.

Ao analisar se a ferramenta de Análise de Sentimentos com dados do *Twitter* pode ser utilizada a nível de município, foi descoberto que sim, utilizando a *metatag user_location* do *Twitter* é possível obter somente dados de Municípios desejados. No presente trabalho foram encontradas algumas limitações em relação a obtenção de dados geolocalizados. Primeiramente, nem todos *tweets* apresentam esta *metatag*, uma vez que somente *tweets* enviados via dispositivos móveis apresentam a localização do usuário. Também foi constatado que este dado não é muito preciso, já que apresenta diferentes grafias para a mesma cidade. Em São Paulo observa-se que existem: “São Paulo”, “São Paulo – SP”, “Sao Paulo”, por exemplo. Além disso a quantidade de dados, após a filtragem por cidade, foi muito pequena. Por esse motivo serão necessárias futuras pesquisas para afirmar se existem quantidades significativas de dados por município nas redes sociais para aplicar a técnica de Análise de Sentimentos. Para isto sugere-se utilizar um intervalo de datas maior que a utilizada no presente trabalho.

Além disto, a utilização de redes sociais pode gerar discussão sobre os tópicos apresentados. Isso pode ser observado quando um *Tweet* tem uma resposta chamada de *reply*. Com isso é possível aumentar a quantidade de dados sobre determinado assunto, assim como o número de opiniões. Entretanto, se deve ter

cuidado com a utilização somente desses dados, uma vez que as discussões em redes sociais não podem substituir as reuniões presenciais do Orçamento Participativo, que é uma ferramenta importante. A Análise de Sentimentos deve ser utilizada somente como ferramenta auxiliar visando ter uma quantidade maior de opiniões já que teriam mais cidadãos participando do processo decisório.

O quadro 5 sumariza os principais pilares do orçamento participativo e responde se ele pode ou não ser aplicado através da Análise de Sentimentos no *Twitter*.

Quadro 5 - Confronto entre teoria do Orçamento Participativo e aplicação da Análise de Sentimentos no *Twitter*

Processos do Orçamento Participativo	É possível aplicar através da técnica de Análise de Sentimentos em redes sociais?
Verificar opiniões sobre os tópicos do orçamento participativo	Sim
Saber localização do cidadão	Parcialmente uma vez que existe a metatag de localização, mas não foi possível comprovar a utilização nesse trabalho devido a baixa quantidade de dados.
Gerar discussões sobre ideias	Sim

Fonte: Elaborado pelo autor.

Portanto, nota-se que os principais pilares do orçamento participativo podem ser utilizados através da técnica de Análise de Sentimentos. Neste trabalho foram encontradas algumas dificuldades que não foram apresentadas na literatura referentes ao uso do *DiscoverText* e também da API do *Twitter*.

Primeiramente, um ponto encontrado neste trabalho, o qual não foi explicado no trabalho de Oliveira (2015), é que o aplicativo de Análise de Sentimentos *DiscoverText* não provê inicialmente uma versão gratuita ilimitada para acadêmicos. Ao criar uma nova conta é necessário enviar um formulário de contato pelo site e, somente após a validação da equipe do aplicativo, é liberada uma conta gratuita que tem duração de somente 30 dias. Não existe uma política descrita no site que provê uma licença ilimitada e, nem mesmo explicando que é um trabalho acadêmico, na hora de criação da conta é liberada uma conta ilimitada. Somente após comunicar à

equipe de suporte para retirar dúvidas sobre o aplicativo, a equipe liberou uma conta ilimitada.

Outro ponto neste trabalho que se mostrou diferente do conteúdo encontrado na literatura é que, conforme Lotan et al. (2011), o *Twitter* se destaca por ter uma API de fácil utilização e que torna fácil a coleta de um grande número de *Tweets*, literalmente centenas de milhões de mensagens. Porém, na prática, o resultado foi um pouco diferente, pois existem limitações de número de requisições nessa API e também ela não salva o histórico de *tweets* antigos, o que dificulta a coleta em um período maior, uma vez que é necessário esperar para coletar esses dados. Por exemplo, se é necessário coletar os dados do dia 10/01/2015 até 10/01/2016, é necessário deixar o aplicativo coletando durante esse período. Isso é ruim uma vez que aumenta bastante o tempo necessário para realizar a pesquisa.

Portanto, nota-se que existem algumas restrições no presente trabalho, porém os principais objetivos conseguiram ser cobertos neste trabalho conforme o Quadro 6.

Quadro 6 - Lista de objetivo x Tarefas realizadas

Objetivo Específico	Tarefa Realizada
Coletar dados no <i>Twitter</i> sobre tópicos do Orçamento Participativo	Esta tarefa está apresentada no método e foi possível coletar os dados sobre os três principais tópicos do OP que são: Saúde, educação e transporte público
Selecionar aplicativo para Análise de Sentimentos e realizar a classificação do conjunto de testes	O aplicativo utilizado foi o <i>DiscoverText</i> , além dele foi avaliado o <i>RapidMiner</i> . A escolha pelo <i>DiscoverText</i> foi principalmente pela licença gratuita e por possibilitar uma coleta maior de dados.
Verificar se a Análise de Sentimentos pode ser aplicada no Orçamento Participativo	O quadro 8 mostra que é possível aplicar no Orçamento Participativo, porém com algumas ressalvas. A principal delas é que a quantidade de dados não foi suficiente para aplicar os dados com referências geográficas.
Avaliar se a Análise de Sentimentos pode auxiliar na apresentação de informações para tomada de decisão na gestão pública.	Conforme os Gráficos 1, 2 e 3 é possível sumarizar as informações sobre assuntos relacionados ao Orçamento Participativo, o que pode ajudar na tomada de decisão da gestão pública.

Fonte: Elaborado pelo autor.

5 CONCLUSÃO

Este trabalho propôs que a técnica Análise de Sentimentos, realizada por meio de dados extraídos do *Twitter*, pode ser utilizada como ferramenta de mensuração da opinião pública sobre assuntos relacionados ao Orçamento Participativo visando contribuir para uma gestão pública mais participativa. A revisão da literatura mostrou que a rede social tem um grande potencial para contribuir com o aumento da participação popular em processos decisórios e fortalecer a democracia, mas ainda era necessário transformar a grande quantidade de dados não estruturados em conhecimento, nesse ponto foi selecionada a técnica de Análise de Sentimentos que é a área que tem como objetivo identificar sentimentos expressos por usuários em seus textos (CHEN, 2010).

Os resultados apresentados pela Análise de Sentimentos foram capazes de identificar, com precisão satisfatória, o número de opiniões negativas, positivas e neutras sobre três tópicos do Orçamento Participativo (saúde, educação e transporte público) que mostraram que a maioria da população não está satisfeita nessas áreas.

O objetivo principal era comprovar se a técnica de Análise de Sentimentos pode ajudar a construir uma gestão pública mais participativa através da aplicação no Orçamento Participativo e, conforme os resultados apresentados e os processos definidos na literatura sobre o mesmo, é possível. A técnica de Análise de Sentimentos pode servir como ferramenta auxiliar do orçamento participativo, o que não eliminaria o processo atual, porém adicionaria mais um meio de participação que seria a opinião os cidadãos sobre os temas nas redes sociais.

Entretanto, este estudo possui algumas limitações. A primeira delas é que este estudo se restringe à opinião de cidadãos durante o período de coleta de dados, isso diminui bastante o número de opiniões. Além disso, a coleta de dados foi realizada somente na rede social *Twitter*, o que revela somente opiniões de uma pequena parte da população brasileira, portanto se deve continuar utilizando outros meios de coleta de dados para que se obtenha uma quantidade significativa de opiniões sobre as diversas áreas do Orçamento Participativo. Outra limitação é a precisão da Análise de Sentimentos com dados de redes sociais, uma vez que dados dessa natureza costumam apresentar uma elevada carga de ironia. (MONTOTOYO; MARTÍNEZ-BARCO; BALAHUR, 2012).

Existem lacunas ainda não preenchidas por este estudo que podem ser estudadas em trabalhos futuros. A primeira delas é a aplicação com uma maior quantidade de dados para verificar se é possível utilizar dados *geolocalizados* com uma amostra maior de dados, isto pode ser realizado através da utilização de outras redes sociais e também aumentando o tempo de coleta. Também seria interessante um aplicativo para Análise de Sentimentos que fosse totalmente gratuito, uma vez que os aplicativos atualmente disponíveis necessitam de licença para utilização. Outro objeto de estudo que pode ser abordado são as opiniões sobre os temas do Orçamento Participativo, não somente os 3 abordados neste artigo.

USE OF SOCIAL NETWORKS TO THE PARTICIPATION IN PUBLIC MANAGEMENT: A PROPOSE THROUGH THE SENTIMENT ANALYSIS TECHNIQUE

Abstract: Popular participation in government decision-making is essential in a democracy, but often citizens are unable to participate in these processes, either because of lack of knowledge or through bureaucracy. A place where they usually express their opinions is in social networks like *Twitter*. This work seeks to conduct an exploratory research to verify if the technique of Sentiment Analysis can contribute to a more participative Public Management. The result of the research proved that it is possible to use the Sentiment Analysis technique in the Participatory Budget, contributing to a more participative public management. However, there are still gaps to be filled as the need for new research using georeferenced data and also with a larger sample of data.

Keywords: Analysis of Feelings. Opinion mining. Participatory public management. Participatory budgeting. Social networks.

REFERÊNCIAS

- BACCIANELLA, Stefano; ESULI, Andrea; SEBASTIANI, Fabrizio. SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining sentiwordnet. **LREC**, n. 10, p. 2200-2204, May 2010.
- BERTOT, John Carlo; JAEGER, Paul T.; HANSEN, Derek. The impact of polices on government social media usage: Issues, challenges, and recommendations. **Government Information Quarterly**, Maryland, v. 29, p. 30-40, 2012.
- BEYER, Yngvil. Using Discovertext for large scale Twitter harvesting. **MDR**, Berlim, v. 41, p. 121-125, Dec. 2012.
- BONSÓN, Enrique *et al.* Local e-government 2.0: Social media and corporate transparency in municipalities. **Government Information Quarterly**, Amesterdã, v. 29, n. 2, p. 123-132, Apr. 2012.
- CAMBRIA, Erik. An introduction to concept-level sentiment analysis. **Advances in Soft Computing and Its Applications**, v. 8266, n. 2, p. 478-483, 2013.
- CERON, Andrea *et al.* Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. **New Media & Society**, Milano, v. 16, n. 2, p. 340-358, Mar 2014.
- CHEN, Hsinchun; ZIMBRA, David. AI and opinion mining. **IEEE Intelligent Systems**, v. 25, n. 3, p. 74-80, may/Jun. 2010.
- COCHRAN, William G. **Sampling Techniques**. 3.ed. New York: John Wiley & Sons, 1977.
- COTTLE, Simon. Media and the Arab uprisings of 2011: research notes. **Journalism**, v. 12, n. 5, p. 647-659, 2011.
- EIRINAKI, Magdalini; PISAL, Shamita; SINGH, Japinder. Feature-based Opinion Mining and Ranking. **Journal of Computer and System Sciences**, v. 78, n. 4, p. 1175-1184, Jul. 2012.
- GHIASSI, M.; SKINNER, J.; ZIMBRA, Z. Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. **Expert Systems with Applications**, v. 40, n. 16, p. 6266-6282, Nov. 2013.
- GLEASON, Benjamin. #Occupy Wall Street: exploring informal learning about a social movement on Twitter. **American Behavioral Scientist**, v. 57, n. 7, p. 966-982, 15 Mar. 2013.
- HELD, David. **O que democracia deveria significar hoje?** Belo Horizonte: Paidéia, 2007.

KAVANAUGH, Andrea L. *et al.* Social media use by government: from the routine to the critical. **Government Information Quarterly**, v. 29, n. 4, p. 480-491, Oct. 2012.

KLOVNISK, Yan. **Gestão pública paradoxo entre políticas de governo e de estado**. Brasília: Kindle Edition, 2016.

LEE, Gwanhoo; KWAK, Young Hoon. An Open Government Maturity Model for social media-based public engagement. **Government Information Quarterly**, v. 29, n. 4, p. 492-503, July 2012.

LEWIS, Seth C.; ZAMITH, Rodrigo; HERMIDA, Alfred. Content Analysis in an era of big data: a hybrid approach to computational and manual methods. **Journal of Broadcasting & Electronic Media**, v. 57, n. 1, p. 34-52, 2013.

LIU, Bing. Sentiment Analysis and Subjectivity. **Handbook of Natural Language Processing**. 2. ed. Chicago: Indurkha and F. J. Damerau, 2010. p. 1-38

LIU, Bing; ZHANG, Lei. A survey of opinion mining and sentiment analysis. In: LIU, Bing; ZHANG, Lei. **Mining text data**. Chicago: Springer, 2012. p. 415-463.

LOTAN, Gilad *et al.* The revolutions were tweeted: information flows during the 2011 tunisian and egyptian revolutions. **International Journal of Communication**, v. 5, p. 1375-1405, 2011.

MARTÍNEZ-CÁMARA, Eugênio *et al.* Sentiment analysis in Twitter. **Natural Language Engineering**, v. 20, n. 01, p. 1-28, Jan. 2014.

MILAGROS, Fernández-Gavilanes *et al.* Unsupervised method for Sentiment Analysis in online texts. **Expert Systems with Applications**, v. 58, p. 57-75, 1 Apr. 2016.

MONTOYO, Andrés; MARTÍNEZ-BARCO, Patrício; BALAHUR, Alexandra. Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments. **Decision Support Systems**, v. 53, n. 4, p. 675-679, Nov. 2012.

NOVAES, Flávio Santos; SANTOS, Maria Elisabete Pereira dos. O Orçamento Participativo e a democratização da gestão pública municipal: a experiência de Vitória da Conquista (BA). **Revista Administração Pública**, Rio de Janeiro, v. 8, n. 4, Jul./Ago. 2014.

OLIVEIRA, Daniel José Silva. **Avaliação do método de Análise de Sentimentos em mídias sociais aplicado na gestão social e política**. 2015. 111f. Dissertação (Mestrado) - universidade federal de lavras, 2015. Disponível em: <<http://repositorio.ufla.br/jspui/handle/1/5253>>. Acesso em: 29 Abr. 2016.

PANG, Bo; LEE, Lilian; VAITHYANATHAN, Shivakumar. Thumbs up?: sentiment classification using machine learning techniques. **Proceedings of the ACL-02 conference on Empirical methods in natural language processing - EMNLP '02**, v. 10, p. 79-86, July 2002.

PAULA, Ana Paula Paes de. **Por uma nova gestão pública**: limites e potencialidades da experiência contemporânea. Rio de Janeiro: FGV, 2005. v. 1.

PREFEITURA MUNICIPAL DE PORTO ALEGRE. **Critérios gerais para distribuição de recursos entre regiões e temáticas**. 2011a. Disponível em: <http://www2.portoalegre.rs.gov.br/op/default.php?p_secao=22>. Acesso em: 07 Jun. 2016.

PREFEITURA MUNICIPAL DE PORTO ALEGRE. **Funcionamento Geral**. 2011b. Disponível em: <http://www2.portoalegre.rs.gov.br/op/default.php?p_secao=15>. Acesso em: 07 Jun. 2016.

PREFEITURA MUNICIPAL DE PORTO ALEGRE. **Histórico do Orçamento Participativo**. 2011c. Disponível em: <http://www2.portoalegre.rs.gov.br/op/default.php?p_secao=1>. Acesso em: 25 Jul. 2016.

RAMBOCAS, Meena; GAMA, João. Marketing research: the role of sentiment analysis. **FEP Working Papers**, Porto, n. 489, p. 1-24, Apr. 2013.

RAPIDMINER. **RapidMiner Pricing**. 2016. Disponível em: <<https://rapidminer.com/pricing/>>. Acesso em: 10 ago. 2016.

SAIF, Hassan *et al.* Contextual semantics for sentiment analysis of Twitter. **Information Processing and Management**, v. 52, n. 1, p. 5-19, Jan. 2016.

SAMPAIO, Rafael. Cardoso *et al.* Participação e deliberação na internet: Um estudo de caso do Orçamento Participativo digital de Belo Horizonte. **Opinião Pública**, Campinas, v. 16, n. 2, p. 446-477, Nov. 2010.

SEGERBERG, Alexandra; BENNETT, W. Lance. Social Media and the organization of collective action: using twitter to explore the ecologies of two climate Change Protests. **The Communication Review**, v. 14, n. 3, p. 197-215, 2011.

SHULMAN, Stuart. Discovertext: software training to unlock the power of text. **ACM International Conference Proceeding Series**, Maryland, p. 373, Jun. 2011.

SHULMAN, Stuart. **Historical twitter prize drawing**. 02 Dec. 2014. Disponível em: <<http://discovertext.com/tag/machine-learning/>>. Acesso em: 25 Jun. 2016.

STECYK, Jane. Blog Twitter. **5 things you might not know about Twitter analytics**, 6 Ago. 2015. Disponível em: <<https://blog.twitter.com/2015/5-things-you-might-not-know-about-twitter-analytics>>. Acesso em: 7 Jun. 2016.

TEXIFTER, LCC. **Features**. DiscoverText, 2011. Disponível em: <<http://discovertext.com/features/>>. Acesso em: 25 Jun. 2016.

TSYTSARAU, Mikalai; PALPANAS, Themis. Survey on mining subjective data on the web. **Data Mining and Knowledge Discovery**, v. 24, n. 3, p. 478-514, May 2012.

TUMASJAN, Andranik *et al.* Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment Andranik. **Proceedings of the Fourth International AAI Conference on Weblogs and Social Media**, Conference paper, p. 178-185, Jan. 2010.

TURNEY, Peter D. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. **Proceeding ACL '02 Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL)**, Philadelphia, p. 417-424, July 2002.

WONG, Felix Ming Fai; SEN, Soumya; CHIAN, Mung. Why watching movie tweets won't tell the whole story? **Proceedings of the WOSN Conference**, New York, p. 61-66, 17 Aug. 2012.

YOON, Sunmoo; ELHADAD, Noémie; BAKKEN, Suzanne. A practical approach for content mining of tweets. **American Journal of Preventive Medicine**, New York, v. 45, n. 1, p. 122-129, 2013.