



RITA GEORGINA GUIMARÃES

**APRIMORAMENTO DA ANÁLISE DE SENTIMENTOS EM
REDES SOCIAIS UTILIZANDO ANÁLISE LÉXICA E PERFIL
DE USUÁRIO**

LAVRAS – MG

2017

RITA GEORGINA GUIMARÃES

**APRIMORAMENTO DA ANÁLISE DE SENTIMENTOS EM REDES SOCIAIS
UTILIZANDO ANÁLISE LÉXICA E PERFIL DE USUÁRIO**

Dissertação de Mestrado apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia de Sistemas e Automação, área de concentração em Automação e Instrumentação de Sistemas, para a obtenção do título de Mestre.

Prof. DSc. Demóstenes Zegarra Rodríguez
Orientador

Prof. DSc. Renata Lopes Rosa
Coorientadora

LAVRAS – MG

2017

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Guimarães, Rita Georgina

Aprimoramento da Análise de Sentimentos em Redes Sociais Utilizando Análise Léxica e Perfil de Usuário / Rita Georgina Guimarães. – Lavras : UFLA, 2017.

64 p. : il.

Dissertação (mestrado acadêmico).–Universidade Federal de Lavras, 2017.

Orientador: Prof. DSc. Demóstenes Zegarra Rodríguez.
Bibliografia.

1. Mineração de Opiniões. 2. Twitter. 3. Aprendizagem de Máquina. I. Rodríguez, Demóstenes Zegarra. II. Rosa, Renata Lopes. III. Título

RITA GEORGINA GUIMARÃES

**APRIMORAMENTO DA ANÁLISE DE SENTIMENTOS EM REDES SOCIAIS
UTILIZANDO ANÁLISE LÉXICA E PERFIL DE USUÁRIO**

Dissertação de Mestrado apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia de Sistemas e Automação, área de concentração em Automação e Instrumentação de Sistemas, para a obtenção do título de Mestre.

APROVADA em 21 de Agosto de 2017.

Prof. DSc. Daniel Furtado Leite DEG / UFLA

Prof. DSc. Ahmed Ali Abdalla Esmin DCC / UFLA

Prof. DSc. Dárlinton Barbosa Feres Carvalho DCC / UFSJ

Prof. DSc. Demóstenes Zegarra Rodríguez
Orientador

Prof. DSc. Renata Lopes Rosa
Coorientadora

**LAVRAS – MG
2017**

Chegar até aqui não foi nada fácil e se hoje comemoro uma conquista, esta se deve àqueles que estiveram ao meu lado em todos os momentos; que fizeram dos meus sonhos os seus e dos meus objetivos sua própria luta. Quero compartilhá-la com vocês, pessoas tão especiais, que não pouparam esforços para que o sorriso que hoje tenho no rosto fosse possível.

AGRADECIMENTOS

Dois anos se passaram e muitos desafios foram superados. Hoje comemoro mais esta vitória, mas sozinha, não teria conseguido. Agradeço a todos que, de forma direta ou indireta, me ajudaram a realizar este sonho.

Gostaria de agradecer especialmente à Deus que me presenteou com um sonho que hoje se torna realidade. Ao meu orientador e minha coorientadora pelos ensinamentos e dedicação, foi uma etapa desafiadora e não teria conseguido concluí-la sem o apoio de vocês.

Ao meu querido esposo, pelo apoio e compreensão; e aos meus pais, que me acompanharam nesta trajetória, me apoiando e incentivando. Várias vezes usei vocês como escudo, em que despejava minhas frustrações, mas o amor sempre foi maior e no momento seguinte vocês estavam lá para me reerguer através do seu apoio incondicional. Nos momentos importantes, suportaram minha ausência e nos dias de fracasso, respeitaram meus sentimentos e enxugaram minhas lágrimas.

Não posso deixar de mencionar minha amada família, irmãos, tios, primos, avós e minha querida filha, pelo carinho e por nunca me deixarem desistir.

Agradeço aos novos amigos que conquistei e que agora fazem parte da minha vida. Sou o resultado da confiança e da força de cada um de vocês.

Meu agradecimento se estende também à CAPES pelo apoio financeiro.

*“Perder-se também é caminho”
(Clarice Lispector)*

RESUMO

As redes sociais online possuem grande quantidade de dados disponíveis para serem explorados de forma cada vez mais abrangente. Através da análise de sentimentos, é possível utilizar estes dados para desenvolver aplicações úteis no cotidiano das pessoas. Porém, ainda existem falhas neste tipo de análise, seja pelo número restrito de palavras contidas nos dicionários ou por não considerarem os mais diversos parâmetros que podem influenciar no sentimento final de uma frase como: perfil dos usuários, pontuação das frases, localização geográfica ou a frequência que se utiliza as redes sociais, por exemplo. Resultados mais confiáveis podem ser obtidos aproveitando o maior número destes parâmetros ou agrupando aqueles mais adequados. Primeiramente, este trabalho sugere que o resultado da análise de sentimentos seja obtido após a pontuação dos advérbios, que tem como função inverter ou intensificar o sentimento final de uma frase. Em seguida, trabalha a proposta de considerar características do perfil do usuário, como faixa etária e gênero para determinar o valor de sentimento de cada frase postada em uma rede social. É apresentado um Sistema de Recomendação (SR), baseado na análise de sentimentos das frases extraídas das redes sociais, a partir de um algoritmo que considera a pontuação dos advérbios. Também foi realizada uma análise detalhada com 7000 frases para determinar quais características seriam mais relevantes, como o uso de pontuação, número de caracteres, compartilhamento de mídias, assuntos, entre outros; e quais poderiam ser desconsiderados. Diferentes algoritmos de aprendizagem de máquina foram testados em busca do melhor resultado para classificação de usuários por faixa etária. Através da pontuação das frases considerando os advérbios foi possível obter o erro máximo absoluto correspondente a 0.21, inferior aos resultados apresentados por outras ferramentas de análise de sentimentos. Para classificação dos usuários por faixa etária a Rede Neural Convolutiva Profunda (DCNN) teve o melhor desempenho, atingindo uma precisão de 0.95 nos testes de validação. Além disso, para validar a utilidade do modelo proposto para classificar grupos etários, ele é implementado no *Enhanced Sentimeter Metric* (eSM), e os resultados da métrica eSM quando as informações do grupo etário não estão disponíveis, são superados com o método proposto de classificação de grupos etários por meio de aprendizagem de máquina. Os avanços e testes realizados mostram que as ferramentas de análise de sentimentos em textos extraídos das redes sociais, são capazes de apresentar resultados cada vez mais confiáveis e realistas.

Palavras-chave: Mineração de Opiniões. Redes Sociais. Twitter. Sistemas de Recomendação. Aprendizagem de Máquina.

ABSTRACT

Social networks have a large amount of data available to be exploited in an increasingly comprehensive way. Sentiment analysis uses this data to develop useful applications in people's daily lives. However, there are still failures in this type of analysis, either by the restricted number of words contained in dictionaries or by not considering all parameters that can influence the final sentiment of a sentence such as: users profile, punctuation, geographical location or social networks using frequency, for example. More reliable results can be obtained by taking advantage of the greater number of these parameters or by grouping the most suitable parameters. First, this work suggests that the result of sentiment analysis is influenced by the adverbs punctuation, whose proposal is to reverse or intensify the final sentiment of a sentence depending of the adverbs. We work on the proposal to consider user profile characteristics such as age and gender to determine the sentiment value of each sentence posted on a social network. A Recommendation System (SR) was presented, based on the sentiment analysis of sentences extracted from social networks, from an algorithm that considers the adverbs punctuation. We also performed a detailed analysis with 7000 sentences to determine which characteristics would be more relevant, such as punctuation, number of characters, media sharing, subjects, among others; and which characteristic could be disregarded. Different machine learning algorithms were tested in search of the best result for classifying the users by age group. Through the punctuation of the sentences considering the adverbs, it was possible to obtain the absolute maximum error corresponding to 0.21, an inferior value compared to the results presented by other tools of sentiment analysis. In order to classify users by age group, the Deep Convolutional Neural Network (DCNN) had the best performance, reaching an accuracy of 0.95 in the validation tests. In addition, to validate the utility of the proposed model to classify age groups, the model was implemented in the Enhanced Sentimeter Metric (eSM), and the eSM metric results when the age group information was not available, were improved with the proposal of age group classification of this work. The advances and tests carried out show that the tools for sentiment analysis in texts extracted from social networks presented increasingly reliable and realistic results.

Keywords: Mining of Opinions. Social networks. Twitter. Recommendation System. Machine Learning.

LISTA DE FIGURAS

Figura 2.1 – Topologia de Redes Neurais Convolucionais	24
Figura 2.2 – Escala de Intensidade Sentimental Positiva ou Negativa de uma Palavra. . .	25
Figura 2.3 – Emoções Contendo Sentimentos Positivos e Negativos.	26
Figura 2.4 – Gráfico das Sete Emoções de Lui (2013).	27
Figura 3.1 – Algoritmo Utilizado para Obtenção do Valor Final de Sentimento da Frase.	40
Figura 3.2 – Implementação do SR em um Dispositivo Eletrônico.	41
Figura 3.3 – Representação da Metodologia Utilizada.	43
Figura 3.4 – Representação dos Dados Normalizados	45
Figura 3.5 – Testes de Aplicação em um Dispositivo Móvel.	47
Figura 4.1 – Representação da Árvore de Decisão	51

LISTA DE TABELAS

Tabela 2.1 – Palavras positivas e negativas	17
Tabela 2.2 – Exemplos de palavras e seus respectivos valores de sentimento	18
Tabela 2.3 – Comparação entre as ferramentas	36
Tabela 3.1 – Exemplos de alguns advérbios e seus respectivos valores	39
Tabela 4.1 – Pontuação de sentimento atribuída pelas ferramentas	49
Tabela 4.2 – Avaliação de desempenho das ferramentas de análise de sentimento	49
Tabela 4.3 – Resultados da aprendizagem de máquina para classificação da faixa etária, considerando todos os parâmetros – fase de treinamento	51
Tabela 4.4 – Resultados da aprendizagem de máquina para classificação da faixa etária, descartando parâmetros irrelevantes - fase de treinamento	52
Tabela 4.5 – Resultados da aprendizagem de máquina para classificação por faixa etária - fase de validação	53
Tabela 4.6 – Avaliação do desempenho da métrica de sentimento, que considera o mo- delo proposto para prever a faixa etária	53

SUMÁRIO

1	INTRODUÇÃO	11
1.1	Motivações e Justificativa	11
1.2	Objetivos	14
1.2.1	Objetivos Específicos	14
1.3	Contribuições	15
1.4	Estrutura da Dissertação	15
2	REFERENCIAL TEÓRICO	16
2.1	Análise de Sentimentos	16
2.1.1	Abordagem Léxica para Análise de Sentimentos	17
2.1.1.1	Dicionário Fixo de Palavras	18
2.1.1.2	StopWords	20
2.1.1.3	N-gramas	20
2.1.1.4	Cálculo de Intensidade de Sentimento e Polaridade de uma Frase	20
2.1.2	Análise de Sentimentos por Aprendizagem de Máquina	21
2.1.2.1	Árvores de Decisão	22
2.1.2.1.1	Random Forest	23
2.1.2.2	Redes Neurais Artificiais	23
2.1.2.3	Support Vector Machine	25
2.2	Análise Afetiva em Redes Sociais	25
2.3	Relação entre a Faixa Etária e as Características na Escrita	27
2.4	Sistemas de Recomendação Baseados em Sentimentos	28
2.4.1	Tipos de Sistemas de Recomendação	29
2.4.2	Componentes dos Sistemas de Recomendação	29
2.5	Ferramentas de Análise de Sentimentos	30
2.5.1	ANEW	30
2.5.2	OpinionFinder	30
2.5.3	WordNet	31
2.5.4	SentiWordNet	31
2.5.5	SentiStrength	32
2.5.6	Sentimeter-Br	33
2.5.6.1	Enhanced-Sentimeter (eSM)	33

2.6	Prós e Contras dos Trabalhos Relacionados	34
3	METODOLOGIA	38
3.1	Abordagem Léxica para Análise de Sentimentos	38
3.1.1	Dicionário de Palavras	38
3.1.2	Extração de Frases Atuais das Redes Sociais	38
3.1.3	Determinação do Algoritmo para Análise de Sentimentos Considerando a Polaridade dos Advérbios	39
3.1.3.1	Etapas da Métrica Proposta Baseada na Polaridade dos Advérbios . . .	39
3.1.3.2	Obtenção da Intensidade de Sentimentos Considerando a Polaridade dos Advérbios	39
3.1.3.3	Validação do Método Proposto de Pontuação dos Advérbios por meio de um SR	40
3.2	Parâmetros do Perfil de Usuários na Análise de Sentimentos	41
3.2.1	Tratamento dos Dados Extraídos de Redes Sociais	42
3.2.2	Fase de Classificação de Usuários por Faixa Etária	44
3.2.3	Classificação de Usuários por Faixa Etária Aplicada em uma Métrica de Intensidade de Sentimento	46
3.2.3.1	Testes Subjetivos para Obtenção de Parâmetros do Perfil de Usuários . .	47
3.2.3.2	Testes Subjetivos para Obtenção de Sentimentos Fornecidos pelos Avaliadores	48
4	RESULTADOS	49
4.1	Análise de Sentimentos Considerando a Polaridade dos Advérbios . . .	49
4.2	Classificação de Usuários do Twitter por Faixa Etária	50
4.2.1	Determinação de um Modelo de Classificação Etária por meio de Aprendizagem de Máquina	50
4.2.2	Avaliação da Utilidade do Modelo Proposto nas Métricas de Intensidade de Sentimento	52
5	CONCLUSÃO	55
5.1	Contribuições	56
5.2	Produção Científica	57
	REFERÊNCIAS BIBLIOGRÁFICAS	64

1 INTRODUÇÃO

Este capítulo apresentará as principais motivações e justificativa para o estudo da análise de sentimentos em textos da Internet, além de apresentar os objetivos e contribuições da pesquisa.

1.1 Motivações e Justificativa

Um dos avanços mais visíveis atualmente vem ocorrendo na área de tecnologia da informação (BURKE, 2002). A Internet tem sido considerada indispensável para muitas pessoas, fazendo parte do seu dia a dia. As pessoas vivem a influência do ambiente online até para realização de atividades mais básicas, como solicitar a entrega de comida ou manter contato com a família durante período longo de trabalho ou viagem.

Com o constante uso da Internet, atualmente usuários passam horas navegando em sites de comércio eletrônico, lendo notícias sobre esportes, jornalismo e entretenimento, expressando suas opiniões e sentimentos, deixando comentários nas redes sociais sobre o que ocorre a todo instante.

As pessoas têm interagido entre si, cada vez mais, por meio das redes sociais e aplicativos para celular, que estão disponíveis a todo momento. Entre as redes sociais, pode-se citar o Twitter ¹ fundada em vinte e um de março de dois mil e seis, que possibilita ao usuário se expressar sobre os mais diversos assuntos postando mensagens de até 140 (cento e quarenta) caracteres.

Com o objetivo de tomar decisões mais apropriadas para garantir o sucesso de produtos e serviços, seja no comércio, campanhas de publicidade ou qualquer outra área, analisar os comentários do público para avaliar sua satisfação, já se tornou uma realidade. Uma forma de analisar uma marca, por exemplo, é avaliar as experiências dos consumidores que já adquiriram alguns produtos, que na maioria das vezes comentam sobre suas opiniões de forma positiva, negativa ou neutra (SOUZA, 2012).

Os comentários dos consumidores geram um tipo de marketing conhecido como *word-of-mouth* (WOM), entendido como marketing boca-a-boca, que por muitos anos constatou-se ter mais efeito do que as campanhas tradicionais (RICHINS, 1984). O Twitter pode ser observado como uma ferramenta de disseminação de informações boca-a-boca, pois é um ambiente de livre

¹ <https://twitter.com>

expressão, onde as informações dos 255 milhões de usuários ativos que publicam uma média de 500 milhões de mensagens por dia, estão disponíveis para serem exploradas (TWITTER, 2016). As mensagens publicadas nesta rede social possuem no máximo 140 caracteres e são denominadas *tweets*.

Costa e Farias (2004) mostram estudos já consolidados sobre o comportamento das pessoas fora dos ambientes virtuais, e ressaltam a importância de avaliá-los para ver até que ponto estes estudos são válidos para meios online. Sua pesquisa sobre as emoções dos usuários, demonstrou uma grande dificuldade de se falar sobre a própria emoção em diversos momentos, já Günther (2001) mostra que existem pessoas que têm grande facilidade para fazer o mesmo em ambientes virtuais.

Guilhardi (2002) explica que a partir da conceituação de sentimentos e emoções é possível localizar onde estão os pontos que devem ser alterados para se manter ou modificar um sentimento. De tal forma que, para se conseguir um determinado sentimento, as variáveis dos eventos que o produzem é que necessitam de modificações. Ou seja, se existe um sentimento negativo em relação a algum produto, basta conhecer as variáveis que causam este sentimento, para modificá-las e atingir o sentimento desejado.

Definir uma ferramenta que seja capaz de classificar emoções facilitaria as tomadas de decisões dos gestores de empresas e de especialistas em análises comportamentais. A ferramenta poderia ser capaz de prever o sucesso ou fracasso de uma candidatura política, medir a disseminação de uma doença ou mesmo o grau de criminalidade de uma cidade (ROSA; RODRÍGUEZ; BRESSAN, 2015).

Segundo Nielsen (2011), os trabalhos desenvolvidos para análise de sentimentos em redes sociais realizados nos últimos anos englobam dois tipos de abordagens. A primeira, utiliza aprendizagem de máquina supervisionada, baseado em textos já marcados, para classificar a polaridade de novos textos. Outra abordagem cria um dicionário de sentimentos e marca o texto com base em algum modelo matemático que descreva como as palavras e frases do texto correspondem ao dicionário.

Entre os softwares que trabalham com a técnica de dicionário de dados, o que mais os diferencia são as listas de palavras contidas no dicionário. A inclusão de gírias, palavras obscenas, siglas usadas na Internet e símbolos, podem ser muito importantes para atingir um bom desempenho quando se trabalha com textos curtos e informais. A classificação das palavras, que

pode ser de acordo com o sentimento, ou através da polaridade, positivo e negativo, também é uma característica própria de cada técnica de análise.

Dessa forma, além de um dicionário apropriado, na primeira fase desta pesquisa sugere-se que as classes gramaticais podem interferir diretamente no sentimento de uma frase. Mais especificamente, os advérbios são tratados de forma a incrementar as métricas existentes.

Porém, outros fatores também diferenciam os trabalhos de cada autor em busca do aperfeiçoamento da análise de sentimentos. Bosco, Patti e Bolioli (2013) e Farías, Patti e Rosso (2016), relatam que em uma análise rigorosa do sentimento, é importante ter métricas para a detecção de ironia, pois isso pode reverter completamente o sentimento atribuído a uma frase. Em ambos os trabalhos afirma-se que a idade do usuário pode influenciar na classificação de ironia da frase.

Da mesma forma, Huffaker e Calvert (2005) mostraram que os adolescentes se comportam de forma diferente no ambiente online. É possível observar algumas peculiaridades no estilo de escrita, como os temas que têm impacto em sua realidade particular.

A preferência por assuntos que não fazem referência à vida pessoal e assuntos mais positivos podem ser características dos usuários adultos (PENNEBAKER; STONE, 2003). Enquanto o uso de gírias é mais comum em frases postadas por usuários adolescentes (SCHWARTZ *et al.*, 2013). Além disso, a necessidade de anexar alguma mídia que representam o conteúdo mencionado é também uma característica dos usuários adultos (PEMPEK; YERMOLAYEVA; CALVERT, 2009).

Como pode ser observado, o parâmetro de idade tem sido mencionado várias vezes como um fator de influência na análise do sentimento. Assim, na segunda etapa desta pesquisa pretende-se agrupar as características mencionadas e outras que são propostas, como o uso de pontuações, abreviações, símbolos que expressam emoções, variações na escrita, além de outras informações relativas ao usuário, como histórico de *tweets*, número de seguidores e número de pessoas que ele segue, de forma a desenvolver um modelo que apresente contribuições para a análise de sentimentos.

Para validar o modelo proposto, ele foi implementado em uma métrica de sentimentos denominada *Enhanced Sentiment Metric* (eSM) (ROSA; RODRÍGUEZ; BRESSAN, 2015). Na validação do modelo, serão realizados testes subjetivos e os resultados serão comparados com aqueles obtidos através das seguintes métricas de sentimento, *Sentimeter-Br* (ROSA; RODRI-

GUEZ; BRESSAN, 2013), (ROSA; RODRÍGUEZ; BRESSAN, 2013), eSM sem considerar o modelo proposto, eSM com o modelo proposto, e eSM considerando a faixa etária real.

Uma das aplicações referentes à análise de opiniões mais estudada ultimamente é relacionada a negócios e marketing. As empresas querem saber os sentimentos sobre os seus produtos e os desejos dos consumidores. Tal fato implica que a análise de sentimentos pode ser utilizada em diversas áreas, sejam relacionadas à beleza, moda, esportes, saúde, tecnologia e quaisquer outras que existam. Segundo Rosa, Rodríguez e Bressan (2015), estabelecer e prever uma estrutura dinâmica de conteúdo das redes sociais, ainda é um problema em aberto.

Desta forma, o presente trabalho visa contribuir, tanto na área de comércio eletrônico, quanto em pesquisas que são realizadas para verificar como o perfil dos usuários, como gênero, idade, nível educacional, entre outros parâmetros, afetam os sentimentos das pessoas a respeito de determinados assuntos. Posteriormente, a validação do aperfeiçoamento da análise de sentimentos pode ser constatada por meio de uma aplicação, tal como um Sistema de Recomendação (SR).

Com a quantidade de informação gerada atualmente no meio online sobre os mais diversos temas, e a necessidade de se compreender a opinião dos clientes ou usuários, é válido descobrir se o perfil do usuário, assim como as classes gramaticais, afetam a análise de sentimentos.

1.2 Objetivos

O objetivo principal desta dissertação é definir novos mecanismos agregadores, que visam melhorar os resultados das avaliações de sentimentos de frases extraídas das redes sociais e validar as novas métricas propostas de análise de sentimentos por meio da implementação de um SR ou aplicação da métrica proposta em outra já existente.

1.2.1 Objetivos Específicos

Este trabalho atinge os seguintes objetivos específicos:

1. Implementar um SR, que faça uso da métrica aperfeiçoada com a pontuação dos advérbios, para análise de sentimentos;
2. Apresentar novos atributos para aperfeiçoar as métricas de análise de sentimentos existentes;

3. Classificar a faixa etária de indivíduos da rede social Twitter de forma automática por aprendizagem de máquina;
4. Utilizar o atributo faixa etária em uma métrica de sentimentos existente.

1.3 Contribuições

Este estudo apresenta como contribuições:

- Um mecanismo agregador com advérbios, aperfeiçoando os mecanismos já existentes;
- Demonstração de um SR, que pode ser aplicado na indicação de produtos, utilizando o novo mecanismo implementado;
- Classificação da faixa etária de indivíduos da rede social Twitter para uso em métricas que considerem a idade como parâmetro de entrada.

1.4 Estrutura da Dissertação

Neste Capítulo 1, uma seção introdutória foi apresentada, expondo o contexto e a justificativa pela dedicação em análise de sentimentos, além dos objetivos gerais e específicos aos quais este trabalho é direcionado.

O Capítulo 2 constitui de uma revisão teórica dos assuntos: conceitos básicos envolvidos na análise de sentimentos, classificação de parâmetros de entrada em uma métrica de sentimentos e SR baseado em sentimentos. A apresentação da metodologia empregada na definição da métrica proposta, incluindo os métodos subjetivos para avaliação de sentimentos, são detalhados no Capítulo 3.

Por fim, o Capítulo 4 é apresentado, contendo os resultados obtidos e as limitações do modelo proposto. A discussão dos resultados alcançados é apresentada no Capítulo 5.

2 REFERENCIAL TEÓRICO

Na literatura é possível encontrar diversos trabalhos relacionados à análise de sentimentos de frases extraídas das redes sociais. O objetivo deste capítulo é apresentar os principais aspectos conceituais, discutindo os estudos acadêmicos encontrados nos últimos anos.

São apresentadas as principais metodologias de análise de sentimentos e em seguida, os métodos de detecção automática de algumas características pessoais, tais como gênero e idade. Posteriormente, é apresentado um SR baseado em sentimentos.

2.1 Análise de Sentimentos

No estudo realizado por Pang e Lee (2008), o termo “análise de sentimentos” é descrito com o mesmo significado de mineração de opiniões, que é um ramo da mineração de textos com foco, não na classificação de tópicos, mas na classificação de acordo com os sentimentos, ideias e opiniões das pessoas a respeito de um assunto.

Para contextualizar a análise de sentimento, primeiro é necessário conceituá-lo. Sentimento pode ser definido como ato ou efeito de sentir; aptidão para receber as impressões. Desta forma, possuir sentimentos pode ser dito como possuir delicadeza, sensibilidade ou até consciência íntima para observar o meio (FERREIRA, 1993).

Guilhardi (2002) relata que os sentimentos e comportamentos são ambos causados por histórias genéticas e ambientais em conjunto com a situação presente. Desta forma, dominar emoções não pode ser considerado como questão de força de vontade. Se são produzidas por algo, podem ser tratadas como produtos e, desta forma, só devem ser alteradas a partir da mudança nos eventos que as produzem (SKINNER, 1991).

No dia a dia das pessoas, os sentimentos e emoções anteriores podem ser usados para prever o que está para acontecer. Por exemplo, antes de se tomar uma decisão é comum receber opiniões de pessoas mais experientes (PANG; LEE, 2008), para que ela seja a mais apropriada possível (SOUZA, 2012).

Independente do contexto, as opiniões têm grande influência sobre o comportamento das pessoas ou nas organizações. Porém, com a explosão das mídias sociais este cenário passou por modificações, sendo disponibilizado a todos um conteúdo de opiniões diversificado e em grandes volumes. Becker e Tumitan (2013) afirmam que a disponibilização destes conteúdos aumenta as opções dos indivíduos que agora não tem apenas opiniões de sua rede pessoal de contatos, como familiares e amigos, mas também de todos os usuários do meio virtual.

Também houveram mudanças no pensamento estratégico de marketing das empresas, pois perceberam que as tradicionais pesquisas de satisfação usando telefonemas, pesquisas de campo ou questionários, além de gerarem altos custos, atingirem apenas uma amostra da população e terem um retorno demorado, muitas vezes, obtinham um resultado pouco eficaz (BECKER; TUMITAN, 2013).

Com usuários cada vez mais conectados à Internet, e com a rapidez com que um assunto é disseminado, comentado e ignorado, não é recomendável o uso de pesquisas que não façam análises em tempo real. Desta forma, para produzir informação com o grande volume de dados, são necessárias ferramentas capazes de processar automaticamente o conteúdo das publicações, extraíndo as opiniões e sentimentos nelas expressos.

Existem diferenças entre o comportamento de indivíduos em ambiente real e virtual, considerando a Internet e redes sociais. Costa e Farias (2004) ressaltam a importância de validar os estudos sobre comportamento em ambientes virtuais. De uma forma geral, a análise de sentimentos pode ser realizada pela abordagem léxica e pela aprendizagem de máquina, as quais serão tratadas a seguir.

2.1.1 Abordagem Léxica para Análise de Sentimentos

Léxico é o conjunto ou acervo de palavras que um determinado idioma possui. Portanto, a análise léxica estuda as unidades do vocabulário, ou seja, as palavras portadoras de sentido: substantivos, adjetivos, verbos, advérbios entre outras.

Outra maneira de se analisar um texto seria a análise sintática, que se encarrega de examinar, classificar e reconhecer as estruturas da sintaxe, isto é, os períodos, as orações e os termos das orações. E por fim, a análise de um texto também pode ser feita através da análise semântica, que verifica o seu significado.

Na análise léxica, a intensidade do sentimento pode ser substituída por uma polaridade positiva, negativa ou neutra, como exemplifica a Tabela 2.1, construída por Pang e Lee (2008).

Tabela 2.1 – Palavras positivas e negativas

	Palavras
Estudante 1	positivas: brilhante, fenomenal, excelente, fantástico negativas: terrível, horroroso
Estudante 2	positivas: espetacular, legal, excelente negativas: ruim, estúpido, lerdo

Também é comum a construção de um dicionário de palavras, que neste caso, se diferencia dos dicionários convencionais, pois o significado das palavras é substituído pela classificação numérica, que indica o valor do sentimento.

2.1.1.1 Dicionário Fixo de Palavras

O dicionário é um conjunto de palavras às quais são associados valores de sentimentos. O sentimento de cada palavra é definido pelo valor médio de sentimento dado pelos especialistas ou avaliadores remotos, por meio de uma avaliação subjetiva. É utilizada uma escala de valores de -5 a +5 (NIELSEN, 2011), onde palavras de valor extremamente depreciativos recebem uma pontuação de -5.

Como exemplo, o dicionário de palavras de Nielsen (2011) contém palavras retiradas de um microblog, entre elas gírias e palavras de uso informal, tal dicionário teve início a partir de um conjunto de palavras (SAPOLSKY; SHAFER; KAYE, 2011), (BAUDHUIN, 1973) e foi gradualmente ampliado, examinando postagens cada vez mais recentes.

Cada palavra tem uma pontuação que varia de -5 (muito negativo) a +5 (muito positivo). No geral a pontuação se estabelece entre -2 e +2, porém as palavras que demonstram maior ênfase possuem pontuações mais próximas dos extremos, por exemplo, as pontuações de palavras obscenas variam entre -4 e -5.

A palavra “desapontado”, conforme apresentado na Tabela 2.2 que mostra a classificação sentimental de algumas palavras, possui polaridade negativa e recebe a classificação numérica de -2, conforme a escala de sentimentos de Nielsen (2011). Esta pontuação é validada através do método de avaliação remota.

Tabela 2.2 – Exemplos de palavras e seus respectivos valores de sentimento

Palavras do dicionário	Valor de sentimento e polaridade
mendiga	-4 (polaridade negativa)
horrorizado	-3 (polaridade negativa)
desapontado	-2 (polaridade negativa)
adorável	4 (polaridade positiva)
lealdade	3 (polaridade positiva)
agradável	2 (polaridade positiva)

Segundo Rosa, Rodríguez e Bressan (2015), analisar tanto o significado individual de uma palavra como o significado desta palavra em um determinado contexto é fundamental para

uma boa análise de sentimentos e polaridade. O exemplo a seguir, mostra que uma palavra pode ter um valor positivo ou negativo, dependendo do contexto:

- O risoto está extremamente seco: polaridade negativa
- O tapete está limpo e seco: polaridade positiva

Desta forma é importante destacar a possibilidade de se usar dicionários específicos para regiões, grupos com nível educacional diferentes, gêneros, idades, tempo de convívio com as redes sociais ou até mesmo áreas específicas de estudo, conforme Rosa (2015).

Para o exemplo citado anteriormente, no dicionário de Nielsen (2011), cada palavra foi adicionada manualmente e, portanto de tempos em tempos deverá ser atualizado com novas expressões, caracterizando um dicionário fixo.

Um dicionário também pode ser semiautomático, onde possui uma pequena quantidade de palavras fixas e as demais são inseridas automaticamente. Peng e Park (2011) descrevem uma metodologia para este tipo de coleção de vocábulos, que consiste em obter um conjunto de sementes de palavras, extrair as correlações semânticas entre elas por meio de conjunções, construir uma matriz de restrição com o peso da relação de sinônimos e usar um algoritmo para calcular o valor dos novos adjetivos.

Existe também a possibilidade de construção automática de um dicionário (MIZUMOTO; YANAGIMOTO; YOSHIOKA, 2012). Porém, é necessário um grande número de frases com polaridades desconhecidas para determinar o valor de sentimento.

Peng e Park (2011) mencionam que a maioria das ferramentas atuais de análise de sentimento em redes sociais utilizam dicionários de dados semiautomáticos, pois quando os dados são escassos, é importante contar com recursos mais eficazes. Porém, a aprendizagem de máquina vem sendo amplamente utilizada na construção mais automatizada dos dicionários de dados.

Todos os dicionários de dados, independentemente do tipo, devem permitir a inserção de vários tipos de palavras como gírias, símbolos que expressam alegria ou tristeza, chamados de “*emoticons*”, da combinação *emotion* (emoção) + *icon* (ícone), além das variações na ortografia de uma mesma palavra. Porém, não precisam constar as *stopwords*, que deverão estar em uma lista separada.

2.1.1.2 StopWords

Palavras que não possuem valor agregador de sentimentos em um texto, tais como: são, meu, uma, os, a, dessa, por, para e aquilo, são denominadas *stopwords* (ARMANO; FANNI; GIULIANI, 2015). As *stopwords* não precisam ser consideradas na análise de um texto. Elas não tem um valor que incremente ou diminua a análise de sentimentos (DENECKE, 2008).

Segundo Braga (2009) há benefícios na remoção de *stopwords* de uma frase para a análise sentimental. Para identificar quais são as palavras que não possuem valor agregador de sentimento e polaridade em uma frase, pode-se considerar uma lista existente de *stopwords*¹ ou criá-la de acordo com as necessidades, tanto no caso dos dicionários como nas análises efetuadas por métodos de aprendizagem de máquina.

2.1.1.3 N-gramas

Ao se analisar um texto, pode-se procurar por palavras únicas ou combinações de palavras, estas combinações ou sequências são chamadas de n-grama (MOURA *et al.*, 2010).

Quando é feita a análise de uma palavra única, por exemplo, a palavra “nada” estamos nos referindo a um unigrama. Caso a análise seja feita com duas palavras consecutivas como “nada mal”, por exemplo, estamos nos referindo a um bigrama.

O significado de alguns bigramas e trigramas (três palavras consecutivas) pode ser muito diferente de um unigrama, como no exemplo anterior. Desta forma, Rosa, Rodríguez e Bressan (2015) relatam que é importante considerar unigramas, bigramas e trigramas na análise de sentimentos de um texto.

2.1.1.4 Cálculo de Intensidade de Sentimento e Polaridade de uma Frase

As métricas para determinação de sentimento total consideram os valores de intensidade sentimental das palavras individuais que compõem uma frase.

Como exemplificação, a métrica utilizada no dicionário de Nielsen (2011) é mostrada na Equação 2.1:

$$sentimento(F) = \sum_{i=1}^m \frac{sentimento(P_i)}{\sqrt{comprimento(F)}} \quad (2.1)$$

Onde:

- $comprimento(F)$ é o número total de palavras dentro de uma frase F .

¹ <http://www.icmc.usp.br/~igorab/stopList.txt>

- $\text{sentimento}(P_i)$ é o valor de sentimento da palavra i que varia de i a m .
- $\text{sentimento}(F)$ é o valor total da intensidade de sentimento da frase F .

Assim, na frase “My hair is strange” (Meu cabelo é/está estranho), o cálculo segundo o dicionário de Nielsen (2011) é: Valores das palavras: strange = -1; $\text{comprimento}(F) = 4$; conta-se 4 palavras (my, hair, is, strange); $\text{sentimento}(F) = -1/\sqrt{4}$. Logo, $\text{sentimento}(F) = -0,50$.

Neste trabalho de Nielsen (2011), as *stopwords* não são consideradas no dicionário de dados. Veja que a palavra “My”, por exemplo, poderia ter sido retirada por não atribuir um valor de sentimento na frase.

Hogenboom *et al.* (2011) sugere que deva ser aplicada uma regra diferente para palavras negativas. Para cada palavra de negação identificada, por exemplo, o modificador do sentimento é multiplicado por um fator de inversão, ou seja, é multiplicado por -1. Quando todas as frases negativas forem processadas, a frase é pontuada pela soma das pontuações, o inversor é utilizado se necessário.

2.1.2 Análise de Sentimentos por Aprendizagem de Máquina

A aprendizagem da máquina, que neste contexto também pode ser tratada como reconhecimento de padrões ou mineração de dados (MEULEMAN; SCHERER, 2013), está relacionada à extração de padrões em grandes conjuntos de dados. Frequentemente, o objetivo é prever com precisão uma classe ou atributo contendo valores reais, com base em outros atributos e um modelo com parâmetros pré-estabelecidos.

Existem muitos modelos de aprendizagem de máquina (LU, 2009), que são construídos por meio de métodos básicos, como regressão linear e esquemas de árvore, bem como métodos mais sofisticados, como Redes Neurais Artificiais (RNA) ou *Support Vector Machine* (SVM). Normalmente, o aprendizado de máquina não restringe a análise de dados a apenas um modelo, mas compara muitos modelos e escolhe aquele que consegue a melhor precisão preditiva.

A aprendizagem de máquina utiliza algoritmos que recebem padrões já definidos, e a partir destes padrões tem capacidade de classificar novas entradas. Os tipos de aprendizagem são:

- Supervisionada: são utilizados exemplos que indicam ao algoritmo as respostas desejadas para certas amostras de entrada.

- Não supervisionada: quando não existem indicações de resposta desejada para amostras de entrada.

O inconveniente do aprendizado de máquina supervisionado, no contexto de análise de sentimentos, é que devem haver vários exemplos de textos já classificados para formar um corpo confiável de treinamento (SILVA; LIMA; BARROS, 2012), (QIU *et al.*, 2009). Portanto, muitas vezes é necessário construir exemplos manualmente para auxiliar na detecção mais precisa dos sentimentos e polaridade.

A análise de sentimentos realizada pela técnica de aprendizagem de máquina supervisionada utiliza textos já classificados que servem como base de treinamento. Com o uso de modelos já treinados, tenta-se prever o conteúdo emocional de um texto desconhecido.

No entanto, abordagens qualitativas geralmente não são consideradas em pesquisas sobre análise de sentimento, pois seria necessário filtrar ou remover frases fora do padrão que são consideradas como ruído, principalmente no corpus fornecido por Twitter (FRAISSE; PAROUBEK, 2014b), (QADIR; RILOFF, 2013), (FRAISSE; PAROUBEK, 2014a), que contém informações abundantes.

Muitas vezes, na primeira fase da análise de dados, é necessário identificar as principais características ou padrões das amostras, esta tarefa deve ser realizada manualmente por especialistas. Neste contexto, a análise é realizada com volumes de dados amplos e filtrados, abrangendo um amplo perfil demográfico, sem considerar pessoas específicas. Até mesmo mensagens com efeitos mais pessoais, como: “Eu estava com tanta saudade de minha casa” ou “Vou começar a reduzir carboidrato”, não expõem informações privadas dos usuários nos resultados obtidos. Afinal, o objetivo é obter padrões, e não classificar casos isolados.

A fim de conseguir trabalhar com grande quantidade de informações e alcançar classificações desejadas, pode-se utilizar os algoritmos de aprendizado de máquina para fornecer resultados com alta precisão (NEVIAROUSKAYA; AONO, 2013). Podem ser utilizados algoritmos baseados em Árvores de Decisão (J48), *Random Forest*, SVM ou RNA.

2.1.2.1 Árvores de Decisão

Testes utilizando algoritmos de Árvore de Decisão são comuns no contexto de análise de sentimentos (BRIONES; AMARASINGHE; MCINNES, 2016). O princípio da Árvore de Decisão é agrupar instâncias de forma recursiva, minimizando a variabilidade das classes. Para

isso, o algoritmo utiliza os valores dos atributos em cada nó não folha e deposita as instâncias nos nós folha.

Dessa forma, cada nó interno corresponde a uma decisão de classificação (SCHULLER *et al.*, 2011). O algoritmo J48 é uma versão do tradicional algoritmo C4.5. A árvore não é capaz de realizar regressão e desempenha somente o papel de classificador, atendendo perfeitamente à análise de sentimentos.

2.1.2.1.1 Random Forest

Em testes realizados com o algoritmo *Random Forest* (BREIMAN, 2001), todas as árvores de decisão são treinadas com os mesmos parâmetros, mas usam diferentes conjuntos de treinamento. Estes conjuntos são gerados a partir do conjunto de treinamento original usando o procedimento de inicialização: para cada conjunto de treinamento, são selecionados aleatoriamente o mesmo número de vetores como no conjunto original, que são escolhidos com a substituição. Isto é, pode ocorrer de vetores serem utilizados mais de uma vez e outros estarem ausentes (BREIMAN, 2001). Nem todas as variáveis são utilizadas para encontrar a melhor separação, mas um subconjunto aleatório delas.

2.1.2.2 Redes Neurais Artificiais

Redes Neurais Artificiais (RNA) possuem grande capacidade de generalização, e podem aproximar funções usadas tanto para regressão quanto para classificação. O princípio desta rede é a propagação de informações através dos neurônios artificiais, usando os pesos sinápticos para modificar o sinal de entrada e gerar um sinal de saída. O aprendizado dos pesos sinápticos é feito por ajuste proporcional a propagação do erro na direção inversa (BISHOP, 1995).

Nos últimos anos, tem ocorrido o crescente uso do *Deep Learning* em várias áreas, como classificação de imagens (GOODFELLOW *et al.*, 2015), (SHIN *et al.*, 2013) e reconhecimento de fala. Já constam excelentes resultados em classificação de textos (SOCHER *et al.*, 2013), (ZHENG; CHEN; XU, 2013), (GLOROT; BORDES; BENGIO, 2011), (ZHANG; ZHAO; LECUN, 2015). O algoritmo *Deep Learning* permite modelos computacionais compostos de múltiplas camadas de processamento aprenderem representações de dados com múltiplos níveis de abstração.

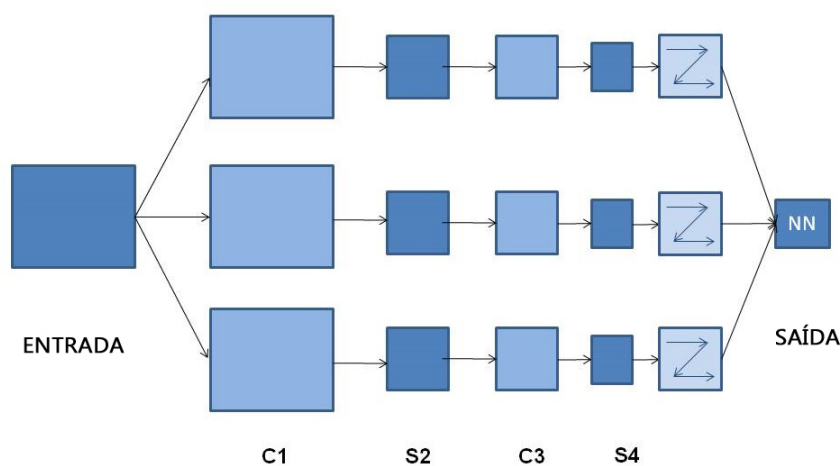
As Redes Neurais Convolucionais Profundas ou *Deep Convolutional Neural Network* (DCNNs) são um dos métodos de classificação utilizados com sucesso em diferentes tarefas

de classificação e reconhecimento (CUN *et al.*, 1990). A aprendizagem profunda é geralmente interpretada em termos do teorema de aproximação universal (HORNÍK; STINCHCOMBE; WHITE, 1989) ou inferência probabilística (MURPHY, 2013).

O teorema de aproximação define uma classe de aproximadores universais, que se refere à capacidade das redes neurais de alimentação direta com uma única camada oculta, de tamanho finito, para aproximar funções contínuas. A interpretação probabilística deriva da aprendizagem de máquina que inclui inferência, bem como conceitos de otimização, como treinamento e teste, relacionados à adaptação e generalização.

O DCNN pode executar tarefas de classificação, e é composto de múltiplas camadas, cada uma calcula transformações convolucionais (CUN *et al.*, 1990). A Figura 2.1 mostra a topologia do DCNN, em que as variáveis *C* representam as camadas de convolução e as variáveis *S* representam as camadas de amostras agrupadas. De *C1* a *S2*, a camada de convolução está presente. De *S2* a *C3* a subamostragem está presente. De *C3* a *S4*, há outra camada de convolução. E, de *S4* para o *Multilayer Perceptron* (MLP) está representada uma saída totalmente conectada.

Figura 2.1 – Topologia de Redes Neurais Convolucionais



Conforme ilustrado na Figura 2.1, é possível estabelecer modelos complexos, sem dificultar necessariamente a interpretação dos resultados, além de ser amplamente utilizado na modelagem de características do comportamento humano (DELAHERCHE *et al.*, 2012), (ZENG *et al.*, 2009), (PANTIC; ROTHKRANTZ, 2003), (SCHULLER *et al.*, 2011).

2.1.2.3 Support Vector Machine

A classificação com Máquina de Vetor de Suporte ou *Support Vector Machine* (SVM), que também é utilizado em propostas de aprendizagem de máquina (GATTI; GUERINI; TURCHI, 2016), tem sido amplamente utilizada em tarefas para classificação de emoção (ZENG *et al.*, 2009). SVMs possuem boas propriedades de generalização (SCHULLER *et al.*, 2011).

SVM utiliza métodos discriminatórios que aprendem limites entre as classes (ALGHOWI-NEM *et al.*, 2016), realizando uma classificação binária com base na separação de hiperplanos; um separador é escolhido de forma a maximizar a distância entre um hiperplano e os vetores de dados mais próximos, os quais são chamados de vetores de suporte (BURGES, 1998).

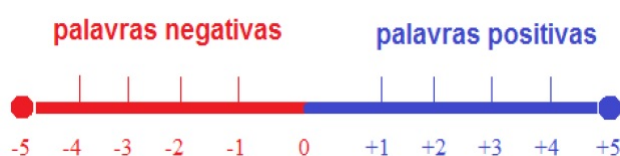
Ao contrário de outras técnicas, as funções do modelo de probabilidade não precisam ser conhecidas anteriormente (SAKR; ELHAJJ; HUIJER, 2010). Isso é muito importante para fins de generalização, pois, em situações práticas, podem não haver informações sobre as funções as distribuições e correlação entre entradas e saídas.

2.2 Análise Afetiva em Redes Sociais

A classificação de opiniões em textos extraídos das redes sociais não tem como objetivo saber os temas dos textos abordados, mas sim as opiniões neles expressadas, portanto, não é tida como uma tarefa trivial devido ao formato livre da escrita. Além do conteúdo postado não obedecer nenhum tipo de padronização, é possível encontrar ironias, erros gramaticais, gírias, símbolos e desenhos que representam emoções (TORTELLA; COELLO, 2015), (ROSA; RODRÍGUEZ; BRESSAN, 2015). Ou seja, os textos possuem um vocabulário com caráter extremamente informal, onde até mesmo abreviações e variações na escrita de uma palavra representam o mesmo significado (JIANG *et al.*, 2011).

Uma forma de expressar o grau de satisfação ou insatisfação de um usuário perante um produto, serviço ou evento é expressá-lo por uma escala, representando um grau de intensidade positiva ou negativa (ROSA; RODRÍGUEZ; BRESSAN, 2015). A Figura 2.2 apresenta uma escala de intensidade de sentimentos que varia de -5 a +5 (do mais negativo ao mais positivo).

Figura 2.2 – Escala de Intensidade Sentimental Positiva ou Negativa de uma Palavra.



Na Figura 2.3 é apresentada uma adaptação das emoções extraídas do trabalho de Posner, Russell e Peterson (2005). Do lado direito da figura estão as emoções que representam um significado positivo, do lado oposto estão as emoções contendo sentimento negativo. Da mesma forma, o eixo horizontal representa a valência e o vertical a excitação ou ativação.

Figura 2.3 – Emoções Contendo Sentimentos Positivos e Negativos.



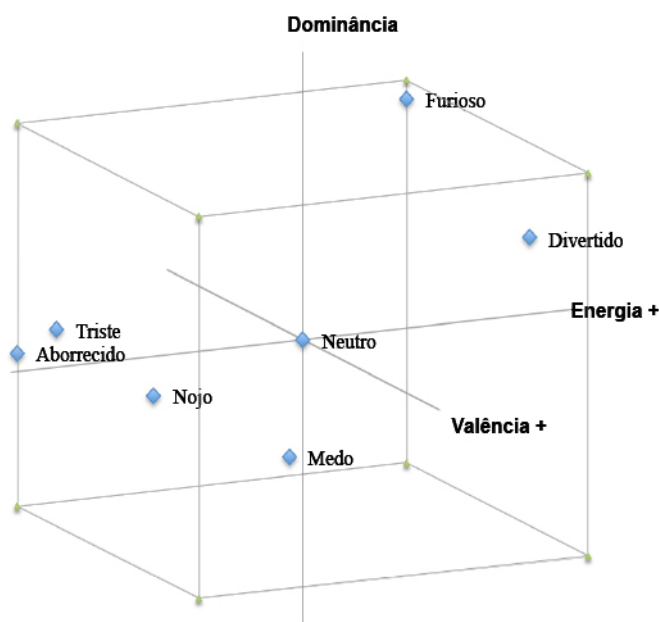
Outro exemplo de classificação é dado por Calix *et al.* (2010), que menciona sobre o modelo “*Big Six*”, onde as emoções do modelo são: felicidade, tristeza, raiva, medo, nojo e surpresa. Apesar deste modelo ter surgido a partir de muitos estudos para demonstrar a universalidade de certas expressões faciais humanas de emoção, ele é também utilizado na abordagem léxica.

Rosa, Rodríguez e Bressan (2015) buscando classificar sentimentos relacionados à música, encontraram a necessidade de alterar as emoções para irritado, feliz, calmo e triste. Mostrando que mesmo utilizando o modelo inicial é possível ajustá-lo de acordo com a análise a ser realizada.

O modelo emocional apresentado no estudo de Lui (2013) tem os valores médios de sete emoções. Diferencia-se do modelo *Big Six* somente na classificação da emoção surpresa, que é substituída pela emoção aborrecida. A Figura 2.4 apresenta estas emoções claramente distantes umas das outras, com exceção das emoções tristeza e aborrecida que se encontram mais próximas.

Pode-se observar que a classificação dos sentimentos pode variar de acordo com o ramo escolhido para realizar a mineração de opiniões.

Figura 2.4 – Gráfico das Sete Emoções de Lui (2013).



2.3 Relação entre a Faixa Etária e as Características na Escrita

É possível encontrar justificativas para a classificação de usuários por faixa etária. Pesquisas voltadas para a área de psicologia, mostram diferenças de comportamento entre as pessoas com idades diferentes (SAWYER *et al.*, 2012), (JANG *et al.*, 2015). No entanto, a rede social Twitter não disponibiliza o parâmetro que contém informações sobre idade.

As pesquisas de Peersman, Daelemans e Vaerenbergh (2011), Nguyen *et al.* (2014) e Zheng *et al.* (2013), definiram que existe a necessidade de se estudar o quanto a idade pode interferir nas análises de sentimentos. Logo, dedicaram esforços a fim de tentar predizê-la.

A estratégia utilizada por Filho, Carvalho e Pappa (2014), foi buscar por descrições no perfil dos usuários que continham as expressões “X anos”, “tenho X anos” ou “fiz X anos”, onde X representaria a idade do usuário. Porém constatou-se que no Twitter, informar a idade na descrição do perfil não é um hábito comum (FILHO; PASTI; CASTRO, 2016) e por isso, a forma de prever a idade conforme Filho, Carvalho e Pappa (2014), não apresenta bons resultados.

Segundo a pesquisa de Sawyer *et al.* (2012), são considerados adolescentes pessoas de 13 a 19 anos. É comum entre os usuários adolescentes de redes sociais, discutir assuntos que ocorrem no seu dia a dia (ARGAMON *et al.*, 2007), assuntos como relacionamentos, escola e amigos são mais frequentes nesta faixa etária (HUFFAKER; CALVERT, 2005).

Os usuários adultos se preocupam com o que escrevem pensando em quem possa ler (PENNEBAKER; STONE, 2003). Dessa forma é possível encontrar mais textos com emoção positiva, deixando de lado a autorreferência, fazendo menos uso da negação (SCHLER *et al.*, 2006) e, conseqüentemente, o uso de gírias também se torna menos frequente (SCHWARTZ *et al.*, 2013).

O fato de adultos falarem menos sobre si próprios, pode ser justificado também pelo menor tempo que os usuários se encontram disponíveis em ambiente online. Na fase adulta, os usuários têm mais ocupações ao longo do dia, já os adolescentes passam mais horas expostos aos meios eletrônicos, de modo que as redes sociais tornam-se uma maneira de expressar ao mundo suas opiniões e sua personalidade (SHAPIRO; MARGOLIN, 2014).

Além dos marcadores de identidade clássicos na idade adulta, como religião, ideologia, política e trabalho, os adultos também usam preferências de mídias para expressar sua identidade. Sendo comum anexarem fotos, vídeos ou um link direcionando para outra página que complementar a aquela informação que foi iniciada na mensagem (PEMPEK; YERMO-LAYEVA; CALVERT, 2009).

2.4 Sistemas de Recomendação Baseados em Sentimentos

Um SR tem como objetivo coletar dados para formar um banco de dados personalizado (ROSA; RODRÍGUEZ; BRESSAN, 2015). De acordo com este banco de dados, itens específicos são recomendados para os usuários.

Um usuário tem suas características, como idade, gênero, nível de escolaridade e facilidade de acesso. Estas características auxiliam na personalização de um sistema de busca, publicidade ou indicação de conteúdo, que pode ser, por exemplo, texto ou multimídia (CHEN; CHEN, 2001), (ZHOU *et al.*, 2011), (BERTINI *et al.*, 2013),

Rodriguez, Perez e Lozano (2010) demonstram uma operação básica de um SR, que é formalmente descrita pelas equações 2.2 e 2.3. O conjunto de todos os usuários do sistema é representado pela variável C , e o conjunto de todo o conteúdo recomendado pela variável S ; R é a classificação ordenada, ou conjunto ordenado de recomendações, mostrando os benefícios do conteúdo $s \in S$ para o usuário $c \in C$.

A utilidade ou relevância do conteúdo é dada por:

$$u = C \times S \rightarrow R \quad (2.2)$$

Já o cálculo da relevância dos resultados de pesquisa no SR é expressa por:

$$\forall c \in C, s'_c = \arg_{s \in S}^{\max} u(c, s) \quad (2.3)$$

A análise de sentimento e afetividade tem sido explorada em SR para recomendar conteúdos de acordo como o estado emocional das pessoas. Um conteúdo pode ser recomendado, por exemplo, com base no humor ou no nível de satisfação do usuário.

2.4.1 Tipos de Sistemas de Recomendação

Existem basicamente três tipos de SR (FAN; SHEN; MAI, 2008), estes são:

- Baseado em Conteúdo: leva em consideração somente o perfil do usuário e seu histórico. Não utiliza dados de outros usuários para efetuar uma recomendação.
- Colaborativo: baseia-se na relação entre itens de preferência de todos os usuários do sistema. Um item escolhido por um usuário pode ser recomendado a outro usuário com mesmo perfil.
- Híbrido: faz a associação entre os sistemas baseados em conteúdo e colaborativos.

2.4.2 Componentes dos Sistemas de Recomendação

Em geral, um SR possui três componentes:

- Dados de fundo, que são informações que o SR tem antes de recomendar um conteúdo;
- Dados de entrada, que são informações passadas pelo usuário ao SR para ser gerada uma recomendação;
- Algoritmos que combinam os dados de fundo e de entrada para gerar a recomendação.

No tipo colaborativo, os dados de fundo são os históricos de avaliação dos usuários sobre um conjunto de itens. Já os dados de entrada são os históricos de avaliação de um usuário específico.

Assim, um SR tem como componentes principais as informações sobre itens e sobre os usuários do sistema, sendo que o objetivo é determinar o grau de relevância entre eles, por meio de algoritmos.

2.5 Ferramentas de Análise de Sentimentos

Na literatura atual é possível encontrar uma grande quantidade de trabalhos acadêmicos relacionados à análise de sentimentos de frases extraídas das redes sociais.

Nesta seção serão apresentadas as principais métricas de análise de sentimentos e as ferramentas de busca e extração de textos da Internet. Além de efetuar uma comparação com o intuito de descobrir novos parâmetros agregadores de sentimentos.

2.5.1 ANEW

Antes da Internet e das redes sociais se tornarem tão comuns em nosso meio, Bradley e Lang (1999) já se preocupavam em desenvolver normas de afetividade para palavras inglesas a fim de padronizar a análise de sentimentos que, naquele momento, seria realizado por pesquisadores no estudo da emoção e atenção.

Posteriormente Nielsen (2011) desenvolveu uma comparação entre a análise de sentimentos utilizando a lista de palavras classificadas de acordo com a afetividade desenvolvida anteriormente, com uma nova lista desenvolvida diretamente para o meio online.

O dicionário de Nielsen (2011), embora possua palavras retiradas de microblogs tais como gírias e palavras obscenas, ainda não trata a detecção de negação, manipulação de *emojis* e variações na ortografia.

ANEW é um estudo que considerou relevante utilizar as informações do usuário para a análise de sentimentos (BRADLEY; LANG, 1999). Foi estudado como a presença ou ausência da informação de gênero afetaria o resultado final da análise de sentimentos.

2.5.2 OpinionFinder

As frases podem ter diferentes sentimentos e são classificadas em polaridades. Uma palavra pode possuir algum sentimento sendo chamada de polar, ou ser neutra, quando não expressa nenhum sentimento, por exemplo alguma informação objetiva, sem valor sentimental ou de emoção.

Wilson, Wiebe e Hoffmann (2005) sugerem uma ferramenta que além de realizar esta separação entre polar e neutra, ainda tenha uma fase de desambiguação, ou seja, classificar as palavras que quando sozinhas expressam um significado, e quando lidas dentro de algum contexto podem expressar outro significado.

No dicionário OpinionFinder, primeiro é realizada a classificação em polar ou neutro. Depois as instâncias classificadas como polar são avaliadas como positivas ou negativas. Para as duas etapas é utilizado algoritmo de aprendizagem de máquina, possibilitando identificar automaticamente a polaridade contextual para um grande subconjunto de expressões de sentimentos.

Um dicionário expandido e também palavras da lista de General Inquirer (GILMAN, 1968), foram utilizadas no OpinionFinder. Na época em que foi desenvolvido, Gilman (1968) não conseguia prever a grande variedade de problemas teóricos e práticos, aos quais sua pesquisa seria útil. Embora não tenha discutido estudos em detalhes, os resultados até aquele momento eram encorajadores e serviram de referência para novas pesquisas na área.

2.5.3 WordNet

O Wordnet² é uma base de dados de palavras em inglês, contendo verbos, advérbios, substantivos, adjetivos e seus sinônimos. Esta base de dados é de maneira simplificada, um enorme dicionário sem nenhuma pontuação por palavra (MILLER, 1995), (WORDNET, 2010).

Existe uma relação entre palavras sinônimas, como carro e automóvel, e não se limita a palavras com a mesma sequência de letras, mas sim aos significados específicos de cada uma, rotulando as relações semânticas entre elas. Diferencia-se de um dicionário convencional, já que o último não segue nenhum padrão explícito em relação aos significados.

Trabalhos como o de Peng e Park (2011) utilizam dicionários de dados que usam o *WordNet*, cujas palavras do dicionário estão na língua inglesa e é necessário efetuar a adição da intensidade de sentimentos ou polaridade em cada palavra.

2.5.4 SentiWordNet

A fim de auxiliar na extração de opiniões, Esuli e Sebastiani (2006) aprofundaram as pesquisas de forma a automatizar a classificação das polaridades. Desenvolveram então o SentiWordNet³, que utiliza um método baseado na formação de um conjunto de classificadores ternários, cada um deles capaz de classificar entre subjetivo (positivo ou negativo) ou objetivo (neutro, sem polaridade). Sendo que cada um deles é treinado de forma distinta, produzindo diferentes resultados. A classificação final é obtida em relação aos resultados apontados por cada

² wordnet.princeton.edu

³ sentiwordnet.isti.cnr.it

um, que podem apontar para a mesma resposta, ou ser realizada uma contagem proporcional ao número que os classificadores atribuíram.

Esuli e Sebastiani (2006) analisaram a importância de considerar o perfil do usuário e, posteriormente, o foco do estudo foi modificado para cobrir a análise de diferentes idiomas.

Denecke (2008) relata a importância de se obter uma ferramenta de análise de sentimentos em outros idiomas, além do Inglês. Propõe o agrupamento de técnicas como um tradutor e o SentiWordNet (ESULI; SEBASTIANI, 2006) para analisar frases alemãs. O resultado obtido alcançou bom desempenho na classificação de textos. Desta forma considerou-se que a técnica utilizada pode ser uma fonte confiável para a análise de sentimentos em um contexto multilíngue, com precisão de 66%.

O SentiWordNet é baseado no dicionário WordNet e possui mais de 110 mil palavras inglesas. Porém ambos não possuem um vocabulário informal. Como os textos extraídos das redes sociais são em sua maioria formados por este tipo de texto, vê-se a necessidade de adicionar e modificar estes dicionários.

2.5.5 SentiStrength

Das ferramentas de análise citadas até agora, somente o SentiStrength tem suporte para o idioma Português (do Brasil – Br), porém utiliza um dicionário genérico. O mais recomendável é trabalhar com dicionários específicos para cada área estudada, pois de acordo com Rosa, Rodríguez e Bressan (2015), as mesmas palavras podem receber diferentes classificações e pontuações, dependendo do tema da frase analisada.

As principais contribuições do trabalho de Thelwall *et al.* (2011) foram a abordagem de aprendizagem de máquina para otimizar os valores finais de sentimentos, os métodos para extrair sentimentos de textos com ortografia fora do padrão, e um método de correção ortográfica relacionado.

Um conjunto de recursos estendidos foi utilizado para comparação, incluindo n-gramas, *emoticons*, palavras com ortografia corrigida (quando apropriado), pontuação repetida, pontos de interrogação e exclamação. Estas funcionalidades incorporaram a maioria dos textos analisados pelo SentiStrength.

Na pesquisa de Thelwall *et al.* (2011), ele se sente otimista em divulgar o percentual de acertos do SentiStrength, relatando que as frases positivas atingiam 60,6% de acurácia, nas

frases negativas este valor chegava a 72,8%. Hoje sabe-se que é necessário obter maiores percentuais de acerto, a fim de tornar os resultados cada vez mais confiáveis e realistas.

2.5.6 Sentimeter-Br

Existem muitas técnicas para criação de um dicionário de dados. Entre elas Rosa, Rodriguez e Bressan (2013) denominam sua proposta como *Sentimeter-Br*. Esta métrica considera n-gramas e diferencia os valores do sentimento dependendo do tempo verbal, nos quais os verbos no passado têm um menor valor de sentimento do que os verbos no tempo presente.

Entre as palavras contidas neste dicionário estão adjetivos positivos e negativos, gírias, emoções e palavras negativas, como por exemplo: não, nunca mais. A *Sentimeter-Br* contém um total de 2.596 palavras, que são utilizadas para classificar a polaridade de uma frase com uma representação numérica do valor.

Apesar do bom resultado, esta ferramenta ainda precisa ser testada em outras áreas, já que os testes foram realizados com foco direcionado ao comércio. Ainda não se sabe se o resultado obtido é válido para assuntos como negócios, tecnologias ou saúde (ROSA; RODRIGUEZ; BRESSAN, 2013).

Esta métrica necessita de várias fases de cálculo até obter o sentimento final da frase, o que a torna inconveniente para aplicações que necessitem de um menor custo de processamento, como, por exemplo, aplicações para celulares.

2.5.6.1 Enhanced-Sentimeter (eSM)

O eSM é outra métrica de sentimento, que considerou as características do perfil do usuário, entre elas, o gênero, nível de educação, localização geográfica e idade. O eSM é a associação de uma métrica de sentimento baseada em léxico, o *Sentimeter-Br*, com um fator de correção baseado nas informações de perfil do usuário. O modelo eSM de uma frase F_i é dado por:

$$eSM(F_i) = Sentimeter_Br(F_i) * C * exp(a_1 * A_1 + a_2 * A_2 + \dots + a_n * A_n + g_1 * M + g_2 * F + e_1 * G + e_2 * nG + t_1 * T_2 + \dots + t_m * T_m) \quad (2.4)$$

Observe que presume-se que as informações de idade estão disponíveis no perfil do usuário.

Onde:

- C é uma constante de escala, obtida por testes subjetivos (ROSA, 2015);
- $a_1 \dots a_n$ são fatores binários relacionados a grupos etários;
- $A_1 \dots A_n$ são os fatores de peso de cada faixa etária, foram considerados quatro grupos;
- g_1 e g_2 são fatores binários relacionados ao gênero;
- M e F são os fatores de peso do gênero, homem ou mulher, respectivamente;
- e_1 e e_2 são fatores binários relacionados ao nível educacional (ensino superior ou não);
- G e nG são os fatores de peso do nível educacional, ensino superior ou não, respectivamente.

2.6 Prós e Contras dos Trabalhos Relacionados

Günther (2001) ressalta, que se tratando de análise de sentimentos em textos escritos na Internet, é importante levar em consideração a época em cada usuário nasceu pois, de acordo com sua pesquisa, existe uma grande diferença entre os nativos digitais que são os usuários que nasceram depois do advento da Internet e os imigrantes digitais.

As pessoas que nasceram antes da difusão da Internet vivem em um meio de adaptação e não conseguem se expressar tão naturalmente em ambientes virtuais (GÜNTHER, 2001), já que presenciaram outras épocas em que a comunicação era realizada através de reuniões entre amigos, cartas, bilhetes entre outros.

Os estudos de Pang e Lee (2008) e Nielsen (2011) abordam um domínio genérico. Porém, a análise de sentimentos ou polaridade deve ser feita em domínios específicos, pois deste modo aumenta-se o desempenho da análise.

A pesquisa de Braga (2009) mostra resultados benéficos na remoção de palavras não agregadoras de sentimentos (*stopwords*). Desta forma, elas podem ser removidas dos dicionários e também das análises de sentimentos efetuadas pela aprendizagem de máquina.

Um dos grandes problemas da análise de sentimentos é estudar a subjetividade e os termos presentes em um texto que auxiliem no momento de classificar uma frase como positiva, negativa ou neutra. Trabalhos como o de Peng e Park (2011), consideram somente palavras da

classe gramatical “adjetivo”, que são as características dos produtos, descartando o restante das palavras.

Trabalhos de análises de textos por meio de dicionários como Bradley e Lang (1999), Wilson, Wiebe e Hoffmann (2005), Qiu *et al.* (2009) e Nielsen (2011) apresentam algumas limitações, no sentido de não considerar a maioria das classes gramaticais, deixando de fazer uma análise detalhada dos advérbios seguidos de adjetivos.

No trabalho realizado por Rosa, Rodríguez e Bressan (2015), os resultados deixam concluir que para aplicação em um SR, é necessário considerar na análise de sentimentos o perfil do usuário. O estudo (ROSA; RODRÍGUEZ; BRESSAN, 2015) considera que é necessário levar em consideração a localização geográfica, além de simplificar os cálculos, que são realizados em várias fases.

De acordo com os trabalhos citados, é essencial que a análise de sentimentos considere as classes gramaticais (advérbios, adjetivos, diferentes tipos verbais), expressões como gírias e palavras obscenas, além de se obter um cálculo apropriado, que seja simples no que diz respeito a processamento, mas ainda assim eficaz.

De acordo com o que foi mostrado, ainda é possível aperfeiçoar em alguns pontos a análise de sentimentos em textos da Internet. Neste trabalho foi realizado um estudo para comparar as principais métricas existentes, a fim de propor novos parâmetros de análise ou simplesmente agrupá-los em uma única ferramenta, de modo que o valor de sentimento calculado esteja cada vez mais próximo da realidade.

Desta forma os parâmetros analisados são os seguintes:

1. Método de implementação – é avaliado se a ferramenta utiliza um modelo matemático (Mod. M.) para classificar as frases ou faz uso da técnica de aprendizagem de máquina (I. A.).
2. Dicionário – pode ser usado um dicionário fixo. Caso seja necessária alguma alteração, esta deverá ser feita manualmente, ou semiautomaticamente, onde as palavras são incluídas de acordo com a necessidade.
3. Aplicação – em quais redes sociais foram aplicados os testes de validação.
4. Gramática – cada trabalho considera uma forma distinta de pontuar as classes gramaticais.

5. N-gramas – existem análises realizadas através de unigramas ou de conjuntos compostos por mais de uma palavra.
6. StopWords – em alguns trabalhos essas palavras são removidas, em outros não.
7. Perfil do usuário – neste item, existem vários tópicos para serem analisados: idade, gênero, nível de escolaridade.
8. Localização geográfica – em cada região ou país, as pessoas podem se expressar de maneiras diferentes.
9. Suporte para outros idiomas – considerando se a ferramenta foi desenvolvida para um idioma específico ou se existe suporte para outros.

A Tabela 2.3 mostra a análise dos parâmetros citados em cada ferramenta mostrada anteriormente. Considere cada uma das ferramentas sendo uma coluna, e cada um dos parâmetros citados, uma linha. Onde foi marcado com *, significa que aquela ferramenta mencionou como trabalho futuro ou considerou importante analisar o parâmetro para atingir melhores resultados.

Tabela 2.3 – Comparação entre as ferramentas

—	Anew	Opinion Finder	WordNet	Senti-WordNet	Senti-Strength	Sentimeter-Br
Método	Mod. M.	I.A.		I.A.	I.A.	Mod.M.
Dicionário	fixo	fixo	fixo	fixo	fixo	fixo
Aplicação	Twitter		Digg		MySpace	Twitter Facebook
Gramática	subst. adjetivos	subst. verbos advérbios	subst. adjetivos verbos advérbios	subst. adjetivos verbos advérbios		subst. adjetivos tempos- verbais
N-gramas	unigramas	det. neg.	não	não	n-gramas det. neg.	n-gramas
StopWords	não	não	não			X
Perfil	gênero					*
Localização	não			*		*
Idiomas	não	não	não	X	X	*

Os campos em branco não significam necessariamente que as métricas não abordam estes parâmetros. Simplesmente nos trabalhos relacionados não foi mencionado de forma direta, a fim de que se pudesse entender que levaram em consideração estes aspectos.

Na classificação gramatical é comum incluírem os advérbios. Porém tratá-los de maneira específica, considerando que eles podem intensificar ou inverter a polaridade de uma frase, ainda não é bem difundido.

No parâmetro de n-gramas das métricas OpinionFinder e SentiStrength, é encontrada a detecção de negação. Proposta que avalia palavras em conjunto, com o objetivo de inverter o sentimento quando existe uma negação.

O uso de *stopwords*, apesar de apontar vantagens segundo Braga (2009), nesta relação de métricas só é mencionado em *Sentimeter-Br*.

O perfil do usuário no que diz respeito a gênero, idade, nível educacional, além da característica de imigrante ou nativo digital, e da localização geográfica ainda são temas novos no que diz respeito à análise de sentimentos. Principalmente a consideração de que existem diferenças na maneira de se expressar daqueles que usam as redes sociais diariamente e aqueles que às usam ocasionalmente.

Além dos parâmetros citados na Tabela 2.3, ainda têm-se que levar em consideração a análise de sentimentos em áreas específicas, pois as pessoas podem se manifestar de forma distinta em determinados assuntos.

Apesar de se tratar de textos informais, também é importante avaliar as pontuações, ainda que atualmente o uso de *emoticons* tenha se expandido, é comum em situações que precisam de ênfase, encontrar os sinais de pontuação.

3 METODOLOGIA

Neste capítulo é apresentada a metodologia utilizada para o cumprimento das duas etapas da pesquisa.

Inicialmente, a proposta para melhoria da análise de sentimentos considerando a polaridade dos advérbios é mostrada. São detalhados os experimentos realizados e a implementação da métrica sugerida.

Em seguida, é apresentado um modelo proposto para classificar grupos de usuários por faixa etária, considerando duas fases, o tratamento dos dados extraídos das redes sociais e a fase de classificação dos usuários.

3.1 Abordagem Léxica para Análise de Sentimentos

A seguir serão apresentados os requisitos para orientação no desenvolvimento do sistema de análise de sentimentos e polaridade.

3.1.1 Dicionário de Palavras

O dicionário deve permitir a inserção de vários tipos de palavras como gírias, expressões de alegrias e tristeza por meio de *emoticons* e variações na ortografia de uma mesma palavra.

Não devem constar *stopwords*, que deverão estar em uma lista separada, nem as palavras classificadas como advérbios, que terão um dicionário específico. Os valores das palavras devem ser escolhidos conforme a média de opinião de especialistas ou avaliadores remotos.

3.1.2 Extração de Frases Atuais das Redes Sociais

As frases a serem extraídas das redes sociais devem possuir um tempo curto de postagem de no máximo um mês, a fim de representarem os sentimentos atuais. Após extrair diversas frases na Internet, os sentimentos das palavras serão contabilizados, de forma que o resultado poderá ser utilizado em aplicações como, por exemplo, um SR interligando as polaridades e sentimentos obtidos. Dependendo de cada sentimento encontrado, um determinado conteúdo será recomendado.

3.1.3 Determinação do Algoritmo para Análise de Sentimentos Considerando a Polaridade dos Advérbios

Nesta seção, primeiramente, é tratada uma visão geral da métrica de análise de sentimentos. Depois, são mostrados os testes preliminares considerando os diferentes valores atribuídos aos advérbios. Por fim, é efetuada a validação da métrica proposta de advérbios por meio de um SR.

3.1.3.1 Etapas da Métrica Proposta Baseada na Polaridade dos Advérbios

Para realização deste trabalho foram utilizadas as seguintes etapas:

- Acesso ao dicionário que contém as palavras com os valores de sentimentos, a fim de retirar os conjuntos de palavras (bigramas) que em sua composição possuem algum advérbio, além dos advérbios propriamente ditos;
- Determinação da intensidade inicial do sentimento;
- Consolidação do valor de sentimento da frase realizando uma segunda leitura em busca de advérbios, que podem inverter ou intensificar o valor final de sentimento da frase.

3.1.3.2 Obtenção da Intensidade de Sentimentos Considerando a Polaridade dos Advérbios

Esta etapa foi desenvolvida em duas fases. Na primeira, a sentença é analisada de modo que um valor inicial de sentimento é obtido. Na segunda fase, é introduzido o valor dos advérbios, conforme a Tabela 3.1.

Tabela 3.1 – Exemplos de alguns advérbios e seus respectivos valores

Advérbios	Valores de sentimento no dicionário de advérbios
muito	+2
pouco	-2
menos	-1,5
mais	+1,5
tão	+1,9
completamente	+2,2
não	-0,8
certamente	+0,9

O algoritmo representado na Figura 3.1 sugere que se uma frase é classificada com polaridade negativa e contém um advérbio com valor negativo, a polaridade final de sentimento da frase deverá ser invertida.

Figura 3.1 – Algoritmo Utilizado para Obtenção do Valor Final de Sentimento da Frase.

```

If polaridade_inicial_sentença = negativo
    If polaridade_advérbio = negativo then
        polaridade_final_sentença = polaridade_sentença_invertida
    If polaridade_advérbio = positivo then
        polaridade_final_sentença = polaridade_sentença_intensificada

If polaridade_inicial_sentença = positivo
    If polaridade_advérbio = negativo then
        polaridade_final_sentença = polaridade_sentença_invertida
    If polaridade_advérbio = positivo then
        polaridade_final_sentença = polaridade_sentença_intensificada

```

Por exemplo, a frase “Ela é bonita” tem um valor de sentimento inicial que corresponde a $(3/\sqrt{3}) = 1,73$. Se mudarmos a frase para “Ela não é bonita”, o valor de sentimento corresponderia a $(3/\sqrt{4}) * (-0,8) = -1,20$. E ainda, se o exemplo fosse alterado para “Ela é muito bonita”, seria $(3/\sqrt{4}) * 2 = 3,00$.

No primeiro exemplo, como a frase não continha advérbios, é realizado apenas o primeiro cálculo. Nos outros dois exemplos que possuem advérbios de polaridade negativa e positiva, respectivamente, é realizada a segunda fase logo após o cálculo inicial.

3.1.3.3 Validação do Método Proposto de Pontuação dos Advérbios por meio de um SR

Os testes foram realizados em ambiente de laboratório para descobrir a pontuação dos diferentes advérbios e escolher as respectivas mensagens de recomendação. Em seguida, um método remoto subjetivo foi usado para validar o SR proposto.

Os estudos foram realizados por 25 avaliadores. Cada um deles analisou uma lista de frases marcando-as com uma respectiva polaridade. Preencheram um questionário com a preferência por uma mensagem positiva ou negativa, dependendo da polaridade da frase. Posteriormente, cada avaliador postou frases na rede social, que foram armazenadas e classificadas em um banco de dados.

A Figura 3.2 apresenta a arquitetura de rede para implementar o SR em um dispositivo móvel. A frase extraída da rede social é analisada e, em seguida, é determinado o valor de

sentimento final. O SR envia mensagens para o dispositivo eletrônico em que o conteúdo dessa mensagem pode ser positivo ou neutro.

Figura 3.2 – Implementação do SR em um Dispositivo Eletrônico.



3.2 Parâmetros do Perfil de Usuários na Análise de Sentimentos

Como já foi citado, parâmetros do perfil de usuários são considerados importantes para realizar a análise de sentimento de uma frase. Desta forma, esta seção apresenta a metodologia para obter uma previsão da faixa etária, considerando algumas informações extraídas diretamente da rede social e alguns parâmetros que foram considerados importantes durante os testes desta pesquisa.

Usuários do Twitter com idade inferior a treze anos não são o público-alvo desta pesquisa. Adotou-se a classificação dos usuários apenas em duas faixas etárias, a primeira composta por todos os usuários menores de vinte anos, aqui denominados adolescentes e a segunda por todos os usuários com idade igual ou superior a vinte anos, denominados adultos.

Para conseguir uma predição mais exata da faixa etária, foram considerados algumas informações extraídas diretamente da rede social, como o uso de gírias e mídias em anexo. Foram utilizados também, outros parâmetros que são propostos por este trabalho.

Entre eles está a pontuação, onde foi considerado se o usuário havia redigido algum tipo de pontuação na mensagem, sendo que as vírgulas e os pontos finais foram desconsiderados, por serem mais comuns em qualquer tipo de texto. Os símbolos que expressam emoções (*emo-*

ticons) já haviam sido considerados na detecção de gênero (HUFFAKER; CALVERT, 2005), e aqui serão tratados juntamente com as pontuações.

Durante a coleta das mensagens que seriam as amostras, pôde-se perceber também que é comum entre os adolescentes estender a grafia de algumas palavras, “O teu é lindooooo-ooo!!!!”. Em contraste, é possível verificar também que a média de caracteres por palavras é menor do que em textos escritos por adultos resultando, na maioria das vezes, em *tweets* mais curtos.

O uso de gírias foi incrementado com as abreviações predefinidas em um dicionário além das variações na grafia das palavras. Se o *tweet* possui ou não um *link* direcionando para outra página, ou algum tipo de anexo como fotos ou vídeos, compõe a entrada que refere-se às mídias anexas. Também foi considerado se a mensagem possuía marcadores como “#” que destaca algum assunto em evidência e “@” que é utilizado para mencionar o nome de outro usuário.

Os últimos parâmetros, número de pessoas que o usuário segue, número de seguidores e total de *tweets* publicados, são extraídos diretamente do perfil do usuário e fazem parte do seu histórico na rede social.

3.2.1 Tratamento dos Dados Extraídos de Redes Sociais

Para classificar os usuários como adolescentes ou adultos foram extraídas, em Outubro de 2016, via API do Twitter, 1000 mensagens de cada assunto selecionado. São eles: “convênio médico”, “alimentação saudável”, “carboidrato”, “PEC 241”, “novena”, “trabalho profissional” e “mãe”.

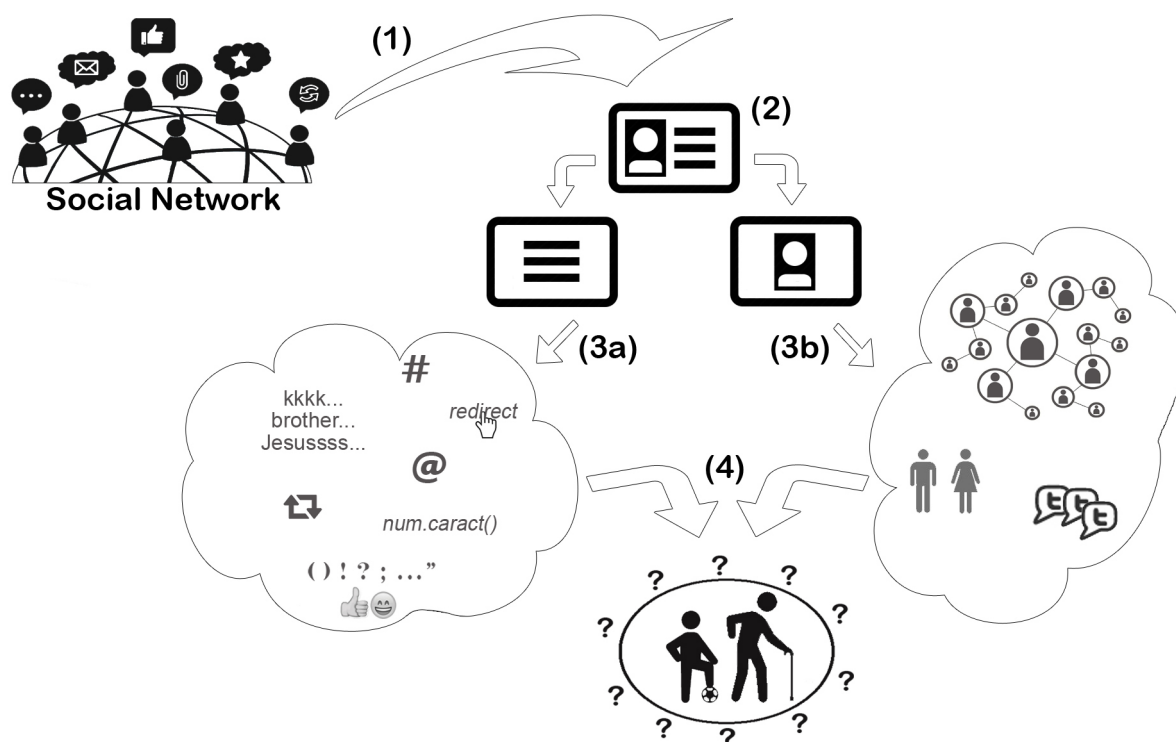
As palavras-chave utilizadas para coleta das mensagens tinham o propósito de abranger um grande número de usuários diferentes e remetem à temas como responsabilidades, esporte, saúde, política, religião, trabalho e família. O dicionário de palavras utilizado neste trabalho considerou diferentes tópicos, cada qual com suas respectivas palavras e suas pontuações sentimentais.

Foi obtido um número total de 7000 *tweets* ou amostras, em que 60 continham mensagens repetidas. Outras 660 foram consideradas *outliers* por apresentarem características muito incomuns em relação às demais, resultando nas 6280 amostras válidas que foram utilizadas para análise.

80% das amostras foram utilizadas como modelagem e 20% foram utilizadas como teste. É importante ressaltar que a análise do conteúdo das amostras foi realizada manualmente e que em todas estas amostras, informações do perfil do usuário estavam disponíveis.

A Figura 3.3 mostra como cada uma das amostras foram utilizadas. Primeiramente (1), a partir de um dos assuntos selecionados, os *tweets* foram extraídos da rede social contendo a mensagem escrita e informações do usuário (2). A partir do conteúdo escrito (3a) foram analisados alguns parâmetros e os demais foram obtidos através do histórico e perfil do usuário (3b). A combinação entre os parâmetros encontrados (4), foi utilizada para prever a faixa etária que melhor classificaria cada um dos usuários.

Figura 3.3 – Representação da Metodologia Utilizada.



Uma vez definidos os parâmetros mais relevantes para prever a faixa etária dos usuários, utilizaram-se os algoritmos de aprendizagem de máquina, em que cada um dos parâmetros se tornou uma entrada.

A primeira entrada considera se a mensagem se trata de um *retweet*, ou seja, se a mensagem não foi escrita por aquele usuário, mas sim compartilhada. Em seguida é observado a presença de marcadores como @ e #. O uso de gírias, abreviações e variações na escrita é o critério da quarta entrada. As pontuações, considerando caracteres como “ ” ? ! ... : () / ; e os

símbolos que expressam emoções, são demonstrados como quinto atributo na aprendizagem de máquina.

Se o *tweet* possui ou não, um link direcionando para outra página, ou algum tipo de anexo como fotos ou vídeos, compõe a sexta entrada, que nesta pesquisa foi chamada de URL (*Uniform Resource Locator*). URL pode ser entendido como um Localizador Uniforme de Recursos, é uma forma padronizada de representação de diferentes documentos, mídias e serviços de rede, capaz de fornecer a cada documento um endereço único.

As próximas entradas são campos numéricos que representam a quantidade de caracteres do *tweet*, obedecendo a política de contagem do Twitter, o número de pessoas que o usuário segue, o número de seguidores que ele possui, total de *tweets* que já publicou em seu perfil, além de um número que representa o assunto ou tema que foi utilizado na coleta da mensagem.

Para validar os resultados que seriam obtidos posteriormente, foi acrescentada a informação de gênero, compondo a última entrada que não pode ser extraída pela API mas pode ser constatada avaliando o perfil do usuário que postou cada um dos *tweets*, além do seu histórico de mensagens e fotos.

O parâmetro adolescente representa a saída do algoritmo de aprendizagem de máquina e é representado como adolescente ou não adolescente (adulto).

3.2.2 Fase de Classificação de Usuários por Faixa Etária

A normalização dos dados considera os tipos de parâmetros como binários ou numéricos. Os parâmetros como: *rt*, *arropa*, *hashtag*, *pontuação*, *gírias*, *url* e a definição se o usuário é adolescente ou não são binários, porque possuem apenas a resposta YES ou NO, no caso da resposta ser positiva ou negativa, respectivamente.

Para o campo gênero foi atribuído F para feminino ou M para masculino. E o restante das entradas: *assunto*, *caracteres*, *segue*, *seguidores* e *tweets* são campos numéricos que representam o valor real extraído, como pode ser visto na Figura 3.4.

Na aprendizagem de máquina, os testes foram realizados utilizando os seguintes algoritmos: *Multilayer Perceptron*, *DCNN*, *Árvores de Decisão*, *Random Forest* e *SVM*, conforme mencionado no Capítulo 2.

Ambos os testes foram implementados pelo pacote de análise de dados *Waikato Environment for Knowledge Analysis* (WEKA) que é muito popular em pesquisas similares com um

Figura 3.4 – Representação dos Dados Normalizados

```

@relation twitter

@attribute rt {YES,NO}
@attribute arroba {YES,NO}
@attribute hashtag {YES,NO}
@attribute girias {YES,NO}
@attribute pontuacao {YES,NO}
@attribute url {YES,NO}
@attribute caracteres numeric
@attribute segue numeric
@attribute seguido numeric
@attribute tweets numeric
@attribute assunto numeric
@attribute sex {F,M}
@attribute adolesc {YES,NO}

@data

NO, YES, NO, NO, NO, NO, 116, 310, 947, 21876, 5, M, NO
NO, YES, NO, NO, YES, NO, 130, 193, 11, 56, 5, M, NO
NO, NO, NO, NO, NO, NO, 140, 509, 4562, 84457, 5, M, NO
NO, NO, NO, NO, NO, NO, 88, 221, 136, 9116, 1, F, YES
NO, NO, NO, NO, NO, NO, 33, 286, 823, 36194, 1, F, YES
NO, NO, NO, YES, YES, NO, 132, 62, 124, 1488, 1, F, YES

```

código desenvolvido na linguagem de programação Python para executar algoritmos de Aprendizagem Profunda (KIM, 2014).

Vale ressaltar que a rede neural recursiva utilizada em alguns estudos (SOCHER *et al.*, 2013) obteve bons resultados em termos de construção de representações de sentenças. Porém, ela usa uma estrutura de árvore e, em uma frase, as palavras posteriores têm mais valor dominante do que palavras anteriores, isto pode reduzir a eficácia para capturar a semântica de uma frase completa. Em vez disso, o DCNN captura melhor a semântica de textos em comparação com a rede neural recursiva (LAI *et al.*, 2015). Por esse motivo, neste trabalho utilizou-se o DCNN.

A ferramenta de aprendizado da máquina, utilizada neste trabalho, permite o ajuste de todos os parâmetros característicos da rede neural, como taxa de aprendizado, taxa de momento, número de épocas, método de treinamento, número de camadas e neurônios, entre outros parâmetros (DERIU *et al.*, 2016). Assim, após realizar vários testes usando validação cruzada (*Cross-Validation*), foram definidas 100 épocas de treinamento, taxa de aprendizado de 0,0004 e taxa de momento de 0,9.

Os testes também foram realizados utilizando o algoritmo de Árvore de Decisão que é comum no contexto da análise de sentimento (BRIONES; AMARASINGHE; MCINNES, 2016). O algoritmo usado pelo WEKA é o J48, uma versão do tradicional algoritmo de C4.5. A árvore não é capaz de realizar regressão e desempenha somente o papel de classificador, atendendo perfeitamente ao problema sugerido.

No teste realizado pelo algoritmo *Random Forest*, todas as árvores são treinadas usando os mesmos parâmetros, mas em diferentes conjuntos de treinamento, buscando encontrar a melhor separação para um subconjunto de variáveis.

A classificação com SVM é muito utilizada para fins de generalização (SAKR; ELHAJJ; HUIJER, 2010), pois em situações práticas, pode não haver informações suficientes sobre as funções e distribuições entre as entradas e as saídas.

Nos testes experimentais realizados neste trabalho, foram utilizadas as mesmas amostras para todos os algoritmos de aprendizado de máquina para avaliar a relevância de cada parâmetro. Posteriormente, o modelo gerado pelo aprendizado da máquina foi aplicado nos *tweets* extraídos da rede social. A faixa etária foi classificada e cada sentença foi utilizada na métrica de sentimento eSM para avaliar o desempenho do modelo proposto.

3.2.3 Classificação de Usuários por Faixa Etária Aplicada em uma Métrica de Intensidade de Sentimento

A métrica do sentimento eSM foi utilizada para avaliar a utilidade do modelo proposto para a classificação de grupos etários. O eSM trabalha com o perfil do usuário, incluindo a faixa etária fornecida, no caso de a rede social oferecer esses dados e/ou o usuário conceder permissão.

Vale ressaltar que o eSM usa quatro grupos etários, porém, não atinge seu melhor desempenho se a faixa etária não estiver disponível. Com uma modificação no sistema, o modelo proposto de classificação é utilizado para obter a informação faltante, classificando o usuário apenas em dois grupos, adolescentes e adultos.

Além disso, a métrica *Sentimeter-Br* também foi usada nos testes de desempenho. Esta métrica não considera informações de perfil de usuário. Ela será usada para demonstrar a relevância das informações do perfil na precisão de determinação da intensidade do sentimento.

Os testes subjetivos foram realizados em duas fases. Na primeira fase, o teste foi realizado em um ambiente de laboratório para obter os parâmetros corretos do perfil do usuário. Na segunda fase, testes subjetivos remotos foram realizados para extrair mais mensagens de uma rede social.

3.2.3.1 Testes Subjetivos para Obtenção de Parâmetros do Perfil de Usuários

Foram realizados testes subjetivos em ambiente de laboratório para obter os parâmetros corretos do perfil do usuário. Esses testes abordaram grande diversidade de perfis de avaliadores, desde a região de nascimento (Norte, Nordeste, Centro-Oeste, Sudeste e Sul do Brasil), idade, nível educacional e gênero. Além disso, os participantes preencheram algumas informações, como seu nome de usuário na rede social Twitter, que foi gravado em um banco de dados para continuar monitorando remotamente os avaliadores.

Em seguida, os avaliadores começaram a escrever mensagens na rede social e pontuaram suas próprias mensagens “@”, “#” e *retweet* com base em uma escala de intensidade de sentimento, de -5 a +5 representando os valores mais negativos e os mais positivos, respectivamente.

Para realizar estes testes, os avaliadores usaram um dispositivo móvel com um aplicativo instalado. A primeira captura de tela dessa aplicação é apresentada na Figura 3.5. Este aplicativo é usado para armazenar as informações do perfil do usuário, como idade e gênero.

Figura 3.5 – Testes de Aplicação em um Dispositivo Móvel.



As informações dos avaliadores foram salvas em um banco de dados e o campo correspondente à idade foi separado em dois grupos:

- Adolescentes, usuários com idade inferior a 20 anos;
- Adultos, usuários com idade igual ou superior a 20 anos.

No total, 76 avaliadores participaram dos testes, constituídos por 38 mulheres e 38 homens. Posteriormente, cada avaliador publicou sentenças remotamente na rede social Twitter, cujas mensagens foram capturadas por um script.

As mensagens foram analisadas pelo mesmo avaliador que publicou as sentenças, pelo *Sentimeter -Br*, e pela métrica eSM com e sem o modelo proposto para prever a faixa etária. Nesta fase, cada avaliador publicou 14 mensagens, duas para cada assunto, atingindo um total de 1064 mensagens.

O tempo médio gasto por usuário para atender às instruções do teste, completar as informações necessárias e escrever as mensagens foi, aproximadamente, 42 minutos.

3.2.3.2 Testes Subjetivos para Obtenção de Sentimentos Fornecidos pelos Avaliadores

Os avaliadores que participaram dos exames presenciais foram monitorados remotamente durante um período de teste de 2 semanas. No total, 5.323 mensagens foram extraídas da rede social. É importante notar que os avaliadores autorizados a serem monitorados forneceram as informações necessárias da rede social.

Nestes testes, o número de mensagens escritas por avaliador não foi limitado, e eles também forneceram seu sentimento real para cada sentença.

O objetivo dos testes foi calcular o valor da intensidade do sentimento com base na métrica eSM, que foi alimentada com dados reais do perfil do usuário, e com a métrica *Sentimeter-Br* que não considera o perfil do usuário.

Os testes subjetivos foram realizados para extrair o valor real dos grupos etários e os sentimentos reais obtidos pelos avaliadores. Assim, o modelo proposto para classificação de grupos etários poderia ser aplicado nas frases coletadas, substituindo a idade pela faixa etária.

Os resultados de desempenho na fase de validação consideram as seguintes funções estatísticas, a raiz quadrada do erro quadrado médio (RMSE), o erro máximo absoluto, erro médio absoluto e o fator de correlação de Pearson (PCC), todos eles baseados nos valores de sentimento obtidos pelas métricas e pelos avaliadores.

4 RESULTADOS

Nesta seção serão apresentados os resultados obtidos através da análise de sentimentos considerando a polaridade dos advérbios, e também da classificação de usuários do Twitter por faixa etária.

4.1 Análise de Sentimentos Considerando a Polaridade dos Advérbios

Nesta etapa da pesquisa, foi possível obter alguns resultados, conforme descrito a seguir.

As frases coletadas do Twitter foram analisadas pela ferramenta SentiStrenght, pela ferramenta *Sentimeter-Br* e também pela métrica proposta, considerando os diferentes valores atribuídos aos advérbios.

De acordo com os testes que foram realizados, as frases que em sua composição, não possuíam advérbios, continuaram com pontuações equivalentes. No entanto, frases que possuíam advérbios já classificados, mostraram resultados mais próximos da realidade.

A Tabela 4.1, apresenta a pontuação de sentimento atribuída às seguintes frases:

1. “Estou muito feliz”
2. “Você está completamente errado”

Tabela 4.1 – Pontuação de sentimento atribuída pelas ferramentas

	SentiStrenght	Sentimeter-Br	Solução Proposta
Frase (1)	3,00	2,30	3,50
Frase (2)	-3,00	-1,50	-3,30

A Tabela 4.2 exibe uma comparação entre o erro máximo considerando o algoritmo proposto e a ferramenta SentiStrenght.

Tabela 4.2 – Avaliação de desempenho das ferramentas de análise de sentimento

Sentimento	Erro Máximo Absoluto
Solução Proposta	0,21
SentiStrenght	0,72

Como pode ser observado, ainda possuindo pontuações relativamente próximas, o algoritmo proposto apresenta o erro máximo absoluto bem inferior ao da ferramenta SentiStrenght, para estes casos.

A solução proposta mostra a importância de se estudar o valor de sentimento dos advérbios em uma frase. Além disso, o SR utilizado como parâmetro para calcular o erro máximo, tem baixa complexidade e apresenta baixos impactos na análise do consumo de energia.

Os resultados encontrados até aqui foram suficientes para encorajar novas pesquisas a fim de verificar novos parâmetros que ainda não são considerados na análise de sentimentos, e que, se utilizados da maneira mais adequada, terão capacidade de melhorar os resultados atuais.

4.2 Classificação de Usuários do Twitter por Faixa Etária

Nesta seção serão apresentados os resultados da classificação por aprendizagem de máquina para gerar o modelo proposto e sua validação em conjunto com o eSM através de testes subjetivos.

4.2.1 Determinação de um Modelo de Classificação Etária por meio de Aprendizagem de Máquina

Foram coletadas 7.000 mensagens de redes sociais, sendo que, 6280 foram consideradas válidas e utilizadas para determinar o modelo de classificação por faixa etária. A partir deste número, 80% das mensagens (5024) foram utilizadas para treinamento, e 20% (1256) para avaliar seu desempenho.

Na fase de treinamento do modelo, com a metodologia utilizada, o DCNN obteve uma acurácia de 0,910 para usuários classificados como Adolescentes e 0,956 para usuários classificados como Adultos. Além disso, os valores de *Recall* atingiram 0,919 e 0,931, respectivamente.

Por meio destes dados, é possível calcular a medida-F, média harmônica entre as informações obtidas, que nesta pesquisa atingiu o valor de 0,930 para a classificação geral, demonstrado na Tabela 4.3, além dos resultados obtidos pelos outros algoritmos de aprendizagem de máquina.

A fim de esclarecer a relevância de alguns parâmetros, a Figura 4.1 demonstra parcialmente o comportamento do teste realizado pela Árvore de Decisão, algoritmo J48, durante a classificação dos usuários, exibindo os pesos de cada parâmetro.

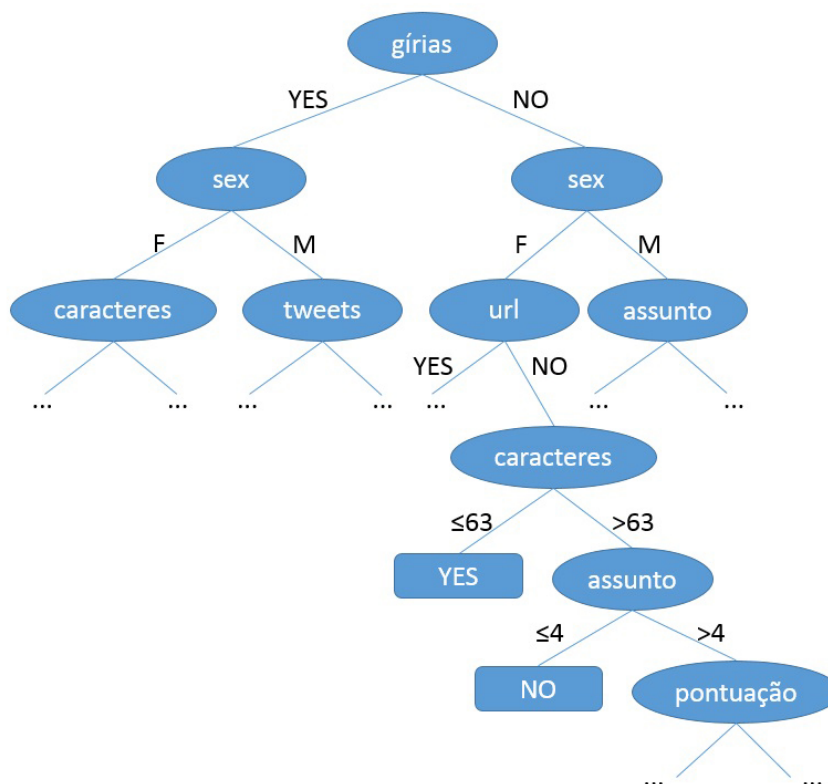
Observando o peso atribuído a cada um dos parâmetros pela Rede Neural e também o posicionamento dos nós pela Árvore de Decisão, foi verificado que alguns parâmetros como,

os marcadores “@ ” e “# ”, e também o parâmetro conhecido como *retweet* (rt), poderiam ser retirados sem perda de precisão na classificação final.

Tabela 4.3 – Resultados da aprendizagem de máquina para classificação da faixa etária, considerando todos os parâmetros – fase de treinamento

Algoritmo	Acurácia	Recall	Medida-F	Classe
Multilayer Perceptron	0,799	0,814	0,810	Adolescente
	0,867	0,854	0,860	Adulto
	0,836	0,836	0,840	Geral
DCNN	0,910	0,919	0,910	Adolescente
	0,956	0,931	0,940	Adulto
	0,938	0,916	0,930	Geral
Árvore de Decisão	0,792	0,811	0,800	Adolescente
	0,858	0,841	0,850	Adulto
	0,829	0,830	0,830	Geral
Random Forest	0,792	0,863	0,830	Adolescente
	0,891	0,837	0,860	Adulto
	0,849	0,849	0,850	Geral
SVM	0,810	0,819	0,810	Adolescente
	0,869	0,849	0,860	Adulto
	0,840	0,838	0,840	Geral

Figura 4.1 – Representação da Árvore de Decisão



Os resultados obtidos no segundo teste, usando as mesmas estruturas e amostras anteriores, podem ser vistos na Tabela 4.4.

Tabela 4.4 – Resultados da aprendizagem de máquina para classificação da faixa etária, descartando parâmetros irrelevantes - fase de treinamento

Métrica	Acurácia	Recall	Medida-F	Classe
Multilayer Perceptron	0,819	0,838	0,830	Adolescente
	0,879	0,862	0,871	Adulto
	0,882	0,869	0,880	Geral
DCNN	0,929	0,936	0,930	Adolescente
	0,967	0,965	0,970	Adulto
	0,952	0,951	0,950	Geral
Árvore de Decisão	0,822	0,818	0,820	Adolescente
	0,871	0,873	0,870	Adulto
	0,842	0,851	0,850	Geral
Random Forest	0,798	0,869	0,830	Adolescente
	0,893	0,847	0,870	Adulto
	0,854	0,853	0,850	Geral
SVM	0,816	0,814	0,810	Adolescente
	0,869	0,871	0,870	Adulto
	0,851	0,850	0,850	Geral

Como pode ser observado, os parâmetros “@”, “# ” e *retweet* não tiveram grande influência na classificação.

Após a conclusão da fase de treinamento da rede, o modelo pode ser usado para testar outras amostras com pesos e taxas corrigidas. Então, as 1256 amostras restantes foram utilizadas para avaliar o desempenho do modelo.

A Tabela 4.5 mostra os resultados obtidos na fase de validação, na qual pode-se observar que o DCNN atingiu um valor geral de medida-F de 0,94 que é muito semelhante ao valor da medida-F alcançado na fase de treinamento.

4.2.2 Avaliação da Utilidade do Modelo Proposto nas Métricas de Intensidade de Sentimento

Um número total de 6387 mensagens foram utilizadas na avaliação da utilidade do modelo proposto em métricas de intensidade de sentimento, foram coletadas 1064 mensagens nos exames presenciais realizados em laboratório e 5323 mensagens coletadas nos testes remotos.

Tabela 4.5 – Resultados da aprendizagem de máquina para classificação por faixa etária - fase de validação

Métrica	Acurácia	Recall	Medida-F	Classe
Multilayer Perceptron	0,801	0,829	0,810	Adolescente
	0,873	0,854	0,860	Adulto
	0,848	0,849	0,850	Geral
DCNN	0,903	0,912	0,910	Adolescente
	0,956	0,958	0,960	Adulto
	0,937	0,937	0,940	Geral
Árvore de Decisão	0,812	0,813	0,810	Adolescente
	0,864	0,863	0,860	Adulto
	0,833	0,842	0,840	Geral
Random Forest	0,784	0,858	0,820	Adolescente
	0,888	0,841	0,860	Adulto
	0,845	0,849	0,850	Geral
SVM	0,800	0,809	0,800	Adolescente
	0,853	0,866	0,860	Adulto
	0,833	0,843	0,840	Geral

A Tabela 4.6 apresenta a avaliação de desempenho e a utilidade do modelo proposto considerando o RMSE, o erro máximo absoluto, o erro médio absoluto e o PCC.

Tabela 4.6 – Avaliação do desempenho da métrica de sentimento, que considera o modelo proposto para prever a faixa etária

	Sentimeter-Br	eSM sem o modelo proposto	eSM com o modelo proposto	eSM com a idade real do grupo
RMSE	0,34	0,29	0,25	0,23
Erro Máximo Absoluto	0,26	0,19	0,15	0,12
Erro Médio Absoluto	0,19	0,15	0,10	0,06
PCC	0,88	0,90	0,94	0,96

A avaliação considerou os seguintes cenários: o *Sentimeter-Br*, que não considera nenhum parâmetro de perfil; o eSM sem a informação de idade e o modelo proposto; o eSM considerando a faixa etária classificada pelo modelo proposto; e finalmente, o eSM considerando o grupo de idade real preenchido por avaliadores a partir dos testes subjetivos presenciais.

A intensidade do sentimento marcada pelos avaliadores representa os valores mais próximos da realidade.

Como pode ser observado na Tabela 4.6, o *Sentimeter-Br* apresenta o pior desempenho porque é uma métrica de sentimento que não leva em consideração quaisquer dados do perfil do usuário, e como é de se esperar, o eSM com a real classificação de faixa etária (quatro grupos) atingiu o melhor desempenho.

Neste trabalho, a comparação principal é entre o eSM com e sem o modelo proposto, no qual o modelo proposto ajudou a obter um melhor desempenho. Isso é muito importante, porque em muitos casos a idade dos usuários não está disponível nas redes sociais.

5 CONCLUSÃO

O gênero já foi considerado como uma informação útil para prever a idade dos usuários (MARQUARDT *et al.*, 2014), e como pode ser visto na representação da Figura 4.1, ocupa um dos nós superiores da Árvore de Decisão. Isso significa que a informação de gênero tem grande importância no desempenho das métricas de sentimento.

Embora estudos (FILHO; PASTI; CASTRO, 2016) mostrem que os critérios para a classificação de gênero, como a distinção de nomes ou a frequência de palavras, não são eficientes, outros trabalhos conseguiram prever o gênero com uma precisão de aproximadamente 80% (KOPPEL; ARGAMON; SHIMONI, 2002), analisando a frequência de algumas palavras específicas.

Desta forma, os esforços para automatizar este processo são válidos para incluí-lo em métricas de sentimento. Nesta pesquisa, nos concentramos na previsão da informação de idade porque esta informação, na maioria das vezes, não está disponível, e a análise de seu impacto na precisão de uma métrica de sentimento também foi estudada.

Nossos resultados experimentais demonstraram que o atributo de URL, que se refere ao usuário mencionando outras páginas, anexando fotos e vídeos ao tweet, é um recurso relevante para prever a faixa etária do usuário. Isso significa que será mais comum encontrar adultos comentando notícias ou até anúncios em seu perfil, diferenciando-os dos adolescentes.

Atributos como o uso de gírias e o número de caracteres da mensagem também são muito específicos em cada faixa etária e, por serem facilmente detectados, garantem maior confiabilidade para os resultados.

De acordo com os resultados, verificamos que existem assuntos sobre os quais adultos ou adolescentes se expressam com mais frequência. Este parâmetro também foi de grande importância porque as palavras-chave foram previamente definidas. Portanto, além de alcançar uma ampla abrangência de temas e pessoas, eles ainda equilibraram a quantidade de amostras em cada faixa etária. É claro que, em certos assuntos, é possível prever quem será o público-alvo, facilitando a classificação e comprovando a eficácia do modelo de classificação do grupo etário proposto.

Além disso, os parâmetros “@”, “# ” e *retweet*, que foram removidos não afetaram o desempenho do modelo proposto. Então, o processamento desta informação não deve ser considerado ou aplicado, diminuindo o tempo necessário para a análise e previsão dos grupos etários.

Dos algoritmos de aprendizado de máquina utilizados, o mais bem aceito foi o DCNN que obteve os melhores resultados atingindo um valor de medida-F igual a 0.940 na fase de validação do modelo de classificação do grupo etário.

Na classificação de usuários por faixa etária, foi visto que nos casos em que a informação da idade do usuário não está disponível, o modelo pode prevê-la, ajudando a melhorar a precisão das métricas de intensidade do sentimento. Além disso, o modelo proposto para classificação de usuários por faixa etária pode ser usado em SR para sugerir o conteúdo mais adequado para cada usuário. Embora seja relatado que pode haver diferenças entre a chamado “idade social” e “idade biológica” (NGUYEN *et al.*, 2014), aqui o objetivo é saber em qual classificação cada usuário se identificaria melhor.

Finalmente, é importante ressaltar que a métrica eSM considera quatro grupos etários e o modelo de classificação por idade proposto considera apenas dois grupos etários. No entanto, os resultados apresentados na Tabela 4.6 mostram que o desempenho do “eSM com modelo proposto” e do “eSM com a idade real do grupo” atingiram valores semelhantes de PCC e RMSE. Este fato demonstrou que os grupos adolescentes e adultos são os grupos etários mais relevantes a serem considerados.

5.1 Contribuições

Na primeira etapa desta pesquisa, na análise léxica das frases extraídas das redes sociais, pode-se perceber a importância de considerar as classes gramaticais da língua portuguesa. Mesmo se tratando de textos informais, a ênfase atribuída aos advérbios resultou em análises mais confiáveis e realistas.

De acordo com os trabalhos relacionados, constatou-se que outros parâmetros deveriam ser estudados. Para obter os parâmetros mais relevantes, um extenso número de frases foram analisadas qualitativamente, para determinar as características dos grupos adolescentes e adultos, considerando o estilo de escrita, o histórico e o perfil dos usuários. Os resultados experimentais mostram que os parâmetros utilizados nesta pesquisa podem atingir uma alta acurácia para determinar a faixa etária dos usuários do Twitter.

A importância de considerar o perfil dos usuários em uma métrica de intensidade de sentimento é reivindicada em vários estudos. O fato da métrica *Sentimeter-Br*, que não considera o perfil de usuário, apresentar o pior desempenho em comparação com a métrica eSM, pode justificar o uso do perfil do usuário.

As redes sociais nem sempre fornecem informações dos usuários, ou eles restringem suas informações pessoais. Nestes casos, o modelo proposto para prever a faixa etária é muito importante para melhorar o desempenho das métricas de intensidade do sentimento.

Os resultados apresentados no Capítulo 4 demonstraram que, considerando grupos etários apenas para adolescentes e adultos, as métricas de sentimento podem obter resultados semelhantes em relação aos casos em que a idade real é considerada.

É importante ressaltar que o modelo proposto também é capaz de trabalhar com outras métricas de análise de sentimentos.

5.2 Produção Científica

As produções científicas são listadas à seguir:

- Artigo completo publicado em Periódico:

GUIMARAES, R. G.; ROSA, R. L.; GAETANO, D. D.; RODRIGUEZ, D. Z.; BRESSAN, G. Age groups classifications in social network using deep learning. *IEEE Access*, IEEE, 2017.

- Artigo curto publicado em Congresso Internacional:

GUIMARAES, R. G.; RODRIGUEZ, D. Z.; ROSA, R. L.; BRESSAN, G. Recommendation system using sentiment analysis considering the polarity of the adverb. In: *2016 IEEE International Symposium of Consumer Electronics (ISCE)*. [S.l.: s.n], 2016. p. 71-72.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALGHOWINEM, S.; GOECKE, R.; WAGNER, M.; EPPS, J.; HYETT, M.; PARKER, G.; BREAKSPEAR, M. Multimodal depression detection: Fusion analysis of paralinguistic, head pose and eye gaze behaviors. *IEEE Transactions on Affective Computing*, PP, n. 99, p. 1–1, Dez 2016. ISSN 1949-3045.
- ARGAMON, S.; KOPPEL, M.; PENNEBAKER, J. W.; SCHLER, J. Mining the blogosphere: Age, gender and the varieties of self-expression. *First Monday*, San Jose, California, v. 12, n. 9, p. 214–217, Maio 2007.
- ARMANO, G.; FANNI, F.; GIULIANI, A. Stopwords identification by means of characteristic and discriminant analysis. In: *Proceedings of the International Conference on Agents and Artificial Intelligence - Volume 2*. Portugal: SCITEPRESS - Science and Technology Publications, Lda, 2015. (ICAART 2015), p. 353–360. ISBN 978-989-758-074-1. Disponível em: <<http://dx.doi.org/10.5220/0005194303530360>>.
- BAUDHUIN, E. S. Obscene language and evaluative response: An empirical study. *Psychological Reports*, SAGE Publications, v. 32, n. 2, p. 399–402, Abr 1973.
- BECKER, K.; TUMITAN, D. Introdução à mineração de opiniões: Conceitos, aplicações e desafios. *Simpósio Brasileiro de Banco de Dados*, 2013.
- BERTINI, M.; BIMBO, A. D.; FERRACANI, A.; GELLI, F.; MADDALUNO, D.; PEZZATINI, D. A novel framework for collaborative video recommendation, interest discovery and friendship suggestion based on semantic profiling. In: *Proceedings of the 21st ACM International Conference on Multimedia*. New York, NY, USA: ACM, 2013. (MM '13), p. 451–452. ISBN 978-1-4503-2404-5. Disponível em: <<http://doi.acm.org/10.1145/2502081.2502264>>.
- BISHOP, C. M. *Neural Networks for Pattern Recognition*. New York, NY, USA: Oxford University Press, Inc., 1995. ISBN 0198538642.
- BOSCO, C.; PATTI, V.; BOLIOLI, A. Developing corpora for sentiment analysis: The case of irony and senti-tut. *IEEE Intelligent Systems*, Institute of Electrical and Electronics Engineers, Inc., USA United States, v. 28, n. 2, p. 55–63, Mar 2013.
- BRADLEY, M. M.; LANG, P. J. *Affective norms for English words (ANEW): Instruction manual and affective ratings*. [S.l.], 1999.
- BRAGA, I. Avaliação da influência da remoção de stopwords na abordagem estatística de extração automática de termos. In: *7th Brazilian Symposium in Information and Human Language Technology (STIL 2009)*, São Carlos, SP, Brazil. [S.l.: s.n.], 2009. p. 18.
- BREIMAN, L. Random forests. *Machine Learning*, v. 45, n. 1, p. 5–32, Out 2001. ISSN 1573-0565.
- BRIONES, G.; AMARASINGHE, K.; MCINNES, B. T. Sentiment analysis in twitter. In: *Proceedings of Semantic Evaluation*. San Diego, California: [s.n.], 2016. p. 215–219.
- BURGES, C. J. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, v. 2, n. 2, p. 121–167, Jan 1998. ISSN 1573-756X.

- BURKE, R. R. Technology and the customer interface: what consumers want in the physical and virtual store. *Journal of the academy of Marketing Science*, Springer, v. 30, n. 4, p. 411–432, Set 2002.
- CALIX, R. A.; MALLEPUDI, S. A.; CHEN, B.; KNAPP, G. M. Emotion recognition in text for 3-d facial expression rendering. *IEEE Transactions on Multimedia*, IEEE, v. 12, n. 6, p. 544–551, Set 2010.
- CHEN, H.-C.; CHEN, A. L. A music recommendation system based on music data grouping and user interests. In: ACM. *Proceedings of the tenth international conference on Information and knowledge management*. [S.l.], 2001. p. 231–238.
- COSTA, A. C. R.; FARIAS, S. A. Emoções e satisfação em compras on-line: o “ser” é humano em ambientes intermediados por computadores. *ENCONTRO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO*, v. 28, 2004.
- CUN, Y. L.; BOSER, B.; DENKER, J. S.; HOWARD, R. E.; HABBARD, W.; JACKEL, L. D.; HENDERSON, D. *Advances in Neural Information Processing Systems 2*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1990. 396–404 p. ISBN 1-55860-100-7.
- DELAHERCHE, E.; CHETOUANI, M.; MAHDHAOUI, A.; SAINT-GEORGES, C.; VIAUX, S.; COHEN, D. Interpersonal synchrony: A survey of evaluation methods across disciplines. *IEEE Transactions on Affective Computing*, v. 3, n. 3, p. 349–365, Jul 2012. ISSN 1949-3045.
- DENECKE, K. Using sentiwordnet for multilingual sentiment analysis. In: *IEEE 24th International Conference on Data Engineering Workshop*. Cancun, Mexico: [s.n.], 2008. p. 507–512.
- DERIU, J.; GONZENBACH, M.; UZDILLI, F.; LUCCHI, A.; LUCA, V. D.; JAGGI, M. Sentiment classification using an ensemble of convolutional neural networks with distant supervision. In: *Proceedings of Semantic Evaluation*. San Diego, California: [s.n.], 2016. p. 1124–1128.
- ESULI, A.; SEBASTIANI, F. Sentiwordnet: A publicly available lexical resource for opinion mining. In: CITESEER. *International Conference on Language Resources and Evaluation*. Genova, Italy, 2006. v. 6, p. 417–422.
- FAN, Y.; SHEN, Y.; MAI, J. Study of the model of e-commerce personalized recommendation system based on data mining. In: IEEE. *Electronic Commerce and Security, 2008 International Symposium on*. [S.l.], 2008. p. 647–651.
- FARÍAS, D. I. H.; PATTI, V.; ROSSO, P. Irony detection in twitter: The role of affective content. *ACM Transactions on Internet Technology (TOIT)*, ACM, v. 16, n. 3, p. 1–24, Jul 2016.
- FERREIRA, A. d. H. *Dicionário aurélio eletrônico*. [S.l.]: Ed. Nova Fronteira, 1993.
- FILHO, J. A. B. L.; PASTI, R.; CASTRO, L. N. de. Gender classification of twitter data based on textual meta-attributes extraction. *Advances in Intelligent Systems and Computing*, v. 444, p. 1025–1034, Mar 2016.
- FILHO, R. M.; CARVALHO, A. I.; PAPPAS, G. L. Inferência de sexo e idade de usuários no twitter. BraSNAM - III Brazilian Workshop on Social Networks Analysis and Mining, 2014.

- FRAISSE, A.; PAROUBEK. Twitter as a comparable corpus to build multilingual affective lexicons. In: *Workshop on Building and Using Comparable Corpora*. Reykjavik, Iceland: [s.n.], 2014. p. 26–31.
- FRAISSE, A.; PAROUBEK, P. Toward a unifying model for opinion, sentiment and emotion information extraction. In: *International Conference on Language Resources and Evaluation*. Reykjavik, Iceland: [s.n.], 2014. p. 3881–3886.
- GATTI, L.; GUERINI, M.; TURCHI, M. Sentiwords: Deriving a high precision and high coverage lexicon for sentiment analysis. *IEEE Transactions on Affective Computing*, v. 7, n. 4, p. 409–421, Out 2016. ISSN 1949-3045.
- GILMAN, R. C. The general inquirer: A computer approach to content analysis. philip j. stone , dexter c. dunphy , marshall s. smith , daniel m. ogilvie. *American Journal of Sociology*, v. 73, n. 5, p. 634–635, Mar 1968.
- GLOROT, X.; BORDES, A.; BENGIO, Y. Domain adaptation for large-scale sentiment classification: A deep learning approach. In: *Proceedings of the Twenty-eight International Conference on Machine Learning*. Washington, USA: [s.n.], 2011. p. 513–520.
- GOODFELLOW, I. J.; BULATOV, Y.; IBARZ, J.; ARNOUD, S.; SHET, V. Multi-digit number recognition from street view imagery using deep convolutional neural networks. *Computer Vision and Pattern Recognition*, v. 6, n. 2, p. 1–13, Maio 2015.
- GUILHARDI, H. J. Análise comportamental do sentimento de culpa. *Ciência do comportamento: conhecer e avançar*, v. 1, p. 173–200, 2002.
- GÜNTHER, J. *Digital natives & digital immigrants*. [S.l.]: Studienverlag Innsbruck, 2001. v. 9. 1–6 p.
- HOGENBOOM, A.; ITERSÓN, P. van; HEERSCHOP, B.; FRASINCAR, F.; KAYMAK, U. Determining negation scope and strength in sentiment analysis. In: IEEE. *IEEE International Conference on Systems, Man, and Cybernetics*. Anchorage, AK, USA, 2011. p. 2589–2594.
- HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. *Neural Netw.*, Elsevier Science Ltd., Oxford, UK, UK, v. 2, n. 5, p. 359–366, Julho 1989. ISSN 0893-6080.
- HUFFAKER, D. A.; CALVERT, S. L. Gender, identity, and language use in teenage blogs. *Journal of Computer-Mediated Communication*, Wiley Online Library, v. 10, n. 2, p. 01–24, Jun 2005.
- JANG, J. Y.; HAN, K.; SHIH, P. C.; LEE, D. Generation like: Comparative characteristics in instagram. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Seoul, Republic of Korea: ACM, 2015. p. 4039–4042. ISBN 978-1-4503-3145-6.
- JIANG, L.; YU, M.; ZHOU, M.; LIU, X.; ZHAO, T. Target-dependent twitter sentiment classification. In: ASSOCIATION FOR COMPUTATIONAL LINGUISTICS. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. PA, USA, 2011. v. 1, p. 151–160.

KIM, Y. Convolutional neural networks for sentence classification. *CoRR*, abs/1408.5882, Set 2014. Disponível em: <<http://arxiv.org/abs/1408.5882>>.

KOPPEL, M.; ARGAMON, S.; SHIMONI, A. R. Automatically categorizing written texts by author gender. *Literary and Linguistic Computing*, ALLC, v. 17, n. 4, p. 401–412, Nov 2002.

LAI, S.; XU, L.; LIU, K.; ZHAO, J. Recurrent convolutional neural networks for text classification. In: *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*. Austin, Texas: AAAI Press, 2015. p. 2267–2273. ISBN 0-262-51129-0.

LU, Z. J. The elements of statistical learning: data mining, inference, and prediction. *Journal of the Royal Statistical Society: Series A*, Wiley Online Library, v. 173, n. 3, p. 693–694, Feb 2009.

LUI, S. A preliminary analysis of the continuous axis value of the three-dimensional pad speech emotional state model. *The 16th edition of the International Conference on Digital Audio Effects*, Maynooth, Ireland, Set 2013.

MARQUARDT, J.; FARNADI, G.; VASUDEVAN, G.; MOENS, M.-F.; DAVALOS, S.; TEREDESAI, A.; COCK, M. D. Age and gender identification in social media. In: *Proceedings of Evaluation Labs*. Sheffield, UK: [s.n.], 2014. p. 1129–1136.

MEULEMAN, B.; SCHERER, K. R. Nonlinear appraisal modeling: An application of machine learning to the study of emotion production. *IEEE Transactions on Affective Computing*, v. 4, n. 4, p. 398–411, Out 2013. ISSN 1949-3045.

MILLER, G. A. Wordnet: A lexical database for english. *Commun. ACM*, ACM, New York, NY, USA, v. 38, n. 11, p. 39–41, nov. 1995. ISSN 0001-0782.

MIZUMOTO, K.; YANAGIMOTO, H.; YOSHIOKA, M. Sentiment analysis of stock market news with semi-supervised learning. In: *IEEE. IEEE/ACIS International Conference on Computer and Information Science*. Shanghai, China, 2012. p. 325–328.

MOURA, M. F.; NOGUEIRA, B. M.; CONRADO, M. da S.; SANTOS, F. F. dos; REZENDE, S. O. Um modelo para a seleção de n-gramas significativos e não redundantes em tarefas de mineração de textos. *Boletim de Pesquisa e Desenvolvimento. Campinas: EMBRAPA Informática Agropecuária*, v. 23, 2010.

MURPHY, K. P. *Machine learning: a probabilistic perspective*. Cambridge, Mass.: MIT Press, 2013. ISBN 9780262018029 0262018020.

NEVIAROUSKAYA, A.; AONO, M. Sentiment word relations with affect, judgment, and appreciation. *IEEE Transactions on Affective Computing*, v. 4, n. 4, p. 425–438, Out 2013. ISSN 1949-3045.

NGUYEN, D.-P.; TRIESCHNIGG, R. B.; DOĞRUÖZ, A. S.; GRAVEL, R.; THEUNE, M.; MEDER, T.; JONG, F. M. de. Why gender and age prediction from tweets is hard: Lessons from a crowdsourcing experiment. In: ASSOCIATION FOR COMPUTATIONAL LINGUISTICS. *Proceedings of the 25th International Conference on Computational Linguistics*. Dublin, Ireland, 2014. p. 1950–1961.

- NIELSEN, F. A. A new anew: Evaluation of a word list for sentiment analysis in microblogs. *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts*, Heraklion, Crete, p. 93–98, Set 2011.
- PANG, B.; LEE, L. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.*, Now Publishers Inc, Hanover, MA, USA, v. 2, n. 1-2, p. 1–135, Out 2008. Disponível em: <<http://dx.doi.org/10.1561/1500000011>>.
- PANTIC, M.; ROTHKRANTZ, L. J. Toward an affect-sensitive multimodal human-computer interaction. *Proceedings of the IEEE*, IEEE, v. 91, n. 9, p. 1370–1390, Set 2003.
- PEERSMAN, C.; DAELEMANS, W.; VAERENBERGH, L. V. Predicting age and gender in online social networks. In: *Proceedings of the 3rd International Workshop on Search and Mining User-generated Contents*. Glasgow, Scotland, UK: ACM, 2011. p. 37–44. ISBN 978-1-4503-0949-3. Disponível em: <<http://doi.acm.org/10.1145/2065023.2065035>>.
- PEMPEK, T. A.; YERMOLAYEVA, Y. A.; CALVERT, S. L. College students' social networking experiences on facebook. *Journal of applied developmental psychology*, Elsevier, v. 30, n. 3, p. 227–238, Jan 2009.
- PENG, W.; PARK, D. H. Generate adjective sentiment dictionary for social media sentiment analysis using constrained nonnegative matrix factorization. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, Catalonia, Spain, v. 51, p. 61801, Jul 2011.
- PENNEBAKER, J. W.; STONE, L. D. Words of wisdom: language use over the life span. *Journal of personality and social psychology*, American Psychological Association, v. 85, n. 2, p. 291, Ago 2003.
- POSNER, J.; RUSSELL, J. A.; PETERSON, B. S. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, Cambridge Univ Press, v. 17, n. 03, p. 715–734, Set 2005.
- QADIR, A.; RILOFF, E. Bootstrapped learning of emotion hashtags# hashtags4you. In: *Proceedings of the 4th workshop on computational approaches to subjectivity, sentiment and social media analysis*. Georgia, USA: [s.n.], 2013. p. 2–11.
- QIU, G.; ZHANG, F.; BU, J.; CHEN, C. Domain specific opinion retrieval. In: *Proceedings of the 5th Asia Information Retrieval Symposium on Information Retrieval Technology*. Sapporo, Japan: Springer-Verlag, 2009. p. 318–329. ISBN 978-3-642-04768-8.
- RICHINS, M. L. Word of mouth communication as negative information. *ACR North American Advances*, 1984.
- RODRIGUEZ, J. D.; PEREZ, A.; LOZANO, J. A. Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, v. 32, n. 3, p. 569–575, Mar 2010.
- ROSA, R. L. *Análise de sentimentos e afetividade de textos extraídos das redes sociais*. Tese (Doutorado) — Universidade de São Paulo, 2015.

ROSA, R. L.; RODRÍGUEZ, D. Z.; BRESSAN, G. Music recommendation system based on user's sentiments extracted from social networks. *IEEE Transactions on Consumer Electronics*, v. 61, n. 3, p. 359–367, Out 2015.

ROSA, R. L.; RODRIGUEZ, D. Z.; BRESSAN, G. Sentimeter-br: A new social web analysis metric to discover consumers sentiment. In: IEEE. *IEEE International Symposium on Consumer Electronics (ISCE)*. Las Vegas, USA, 2013. p. 153–154.

ROSA, R. L.; RODRÍGUEZ, D. Z.; BRESSAN, G. Sentimeter-br: A social web analysis tool to discover consumers' sentiment. In: *IEEE 14th International Conference on Mobile Data Management*. Milan, Italy: [s.n.], 2013. p. 122–124.

ROSA, R. L.; RODRIGUEZ, D. Z.; BRESSAN, G. Sentimeter-br: Facebook and twitter analysis tool to discover consumers sentiment. In: *The Ninth Advanced International Conference on Telecommunications*. Rome, Italy: [s.n.], 2013. p. 61–66.

ROSA, R. L.; RODRÍGUEZ, D. Z.; BRESSAN, G. Análise afetiva de frases extraídas das redes sociais. *Anais do Simpósio Brasileiro de Tecnologia da Informação*, 2015.

SAKR, G. E.; ELHAJJ, I. H.; HUIJER, H. A. S. Support vector machines to define and detect agitation transition. *IEEE Transactions on Affective Computing*, v. 1, n. 2, p. 98–108, Jul 2010. ISSN 1949-3045.

SAPOLSKY, B. S.; SHAFER, D. M.; KAYE, B. K. Rating offensive words in three television program contexts. *Mass Communication and Society*, Taylor & Francis, v. 14, n. 1, p. 45–70, Jan 2011.

SAWYER, S. M.; AFIFI, R. A.; BEARINGER, L. H.; BLAKEMORE, S.-J.; DICK, B.; EZEH, A. C.; PATTON, G. C. Adolescence: a foundation for future health. *The Lancet*, v. 379, n. 9826, p. 1630–1640, Abr 2012.

SCHLER, J.; KOPPEL, M.; ARGAMON, S.; PENNEBAKER, J. W. Effects of age and gender on blogging. In: *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs*. [S.l.: s.n.], 2006. v. 6, p. 199–205.

SCHULLER, B.; BATLINER, A.; STEIDL, S.; SEPPI, D. Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge. *Speech Communication*, v. 53, n. 9, p. 1062 – 1087, 2011. ISSN 0167-6393. Sensing Emotion and Affect - Facing Realism in Speech Processing. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167639311000185>>.

SCHWARTZ, H. A.; EICHSTAEDT, J. C.; KERN, M. L.; DZIURZYNSKI, L.; RAMONES, S. M.; AGRAWAL, M.; SHAH, A.; KOSINSKI, M.; STILLWELL, D.; SELIGMAN; SELIGMAN, M. E. Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one*, Public Library of Science, v. 8, n. 9, p. 73–79, Nov 2013.

SHAPIRO, L. A. S.; MARGOLIN, G. Growing up wired: Social networking sites and adolescent psychosocial development. *Clinical child and family psychology review*, Springer, v. 17, n. 1, p. 1–18, Mar 2014.

SHIN, H.-C.; ORTON, M. R.; COLLINS, D. J.; DORAN, S. J.; LEACH, M. O. Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study

using 4d patient data. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, v. 35, n. 8, p. 1930–1943, Ago 2013.

SILVA, N. R.; LIMA, D.; BARROS, F. Sapair: Um processo de análise de sentimento no nível de característica. In: *4th International Workshop on Web and Text Intelligence (WTI'12)*, Curitiba. [S.l.: s.n.], 2012.

SKINNER, B. F. *Questões recentes na análise comportamental*. [S.l.]: Papirus, 1991.

SOCHER, R.; PERELYGIN, A.; WU, J.; CHUANG, J.; MANNING, C. D.; NG, A. Y.; POTTS, C. Recursive deep models for semantic compositionality over a sentiment treebank. In: *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. Seattle, WA: Association for Computational Linguistics, 2013. p. 1631–1642.

SOUZA, M. V. d. S. Mineração de opiniões aplicada a mídias sociais. Pontifícia Universidade Católica do Rio Grande do Sul, 2012.

THELWALL, M.; BUCKLEY, K.; PALTOGLOU, G.; CAI, D.; KAPPAS, A. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, Wiley, v. 62, n. 2, p. 2544–2558, Ago 2011.

TORTELLA, P. L.; COELLO, J. M. A. Análise de sentimentos em mídias sociais. *Anais do V Encontro de Iniciação em Desenvolvimento Tecnológico e Inovação*, 2015.

TWITTER. *Quem está no Twitter*. 2016. Disponível em: <<https://business.twitter.com/pt/help/who-s-on-twitter?location=>>. Acesso em: 16 Fev 2016.

WILSON, T.; WIEBE, J.; HOFFMANN, P. Recognizing contextual polarity in phrase-level sentiment analysis. In: ASSOCIATION FOR COMPUTATIONAL LINGUISTICS. *Proceedings of the conference on human language technology and empirical methods in natural language processing*. Vancouver, Canada, 2005. p. 347–354.

WORDNET. *About WordNet*. 2010. Disponível em: <<http://wordnet.princeton.edu/wordnet/>>. Acesso em: 17 Jun 2016.

ZENG, Z.; PANTIC, M.; ROISMAN, G. I.; HUANG, T. S. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 31, n. 1, p. 39–58, Jan 2009. ISSN 0162-8828.

ZHANG, X.; ZHAO, J.; LECUN, Y. Character-level convolutional networks for text classification. In: *Proceedings of the 28th International Conference on Neural Information Processing Systems*. Montreal, Canada: MIT Press, 2015. p. 649–657.

ZHENG, L.; YANG, K.; YU, Y.; JIN, P. Predicting age range of users over microblog dataset. *International Journal of Database Theory and Application*, v. 6, n. 6, p. 85–94, Out 2013.

ZHENG, X.; CHEN, H.; XU, T. Deep learning for chinese word segmentation and pos tagging. In: . Seattle, Washington, USA: ACL, 2013. p. 647–657.

ZHOU, R.; KHEMMARAT, S.; GAO, L.; WANG, H. Boosting video popularity through recommendation systems. In: *Databases and Social Networks*. New York, NY, USA: ACM, 2011. (DBSocial '11), p. 13–18. ISBN 978-1-4503-0650-8. Disponível em: <<http://doi.acm.org/10.1145/1996413.1996416>>.