



Universidade Federal do Piauí
Centro de Ciências da Natureza
Programa de Pós-Graduação em Ciência da Computação

Abordagem para definir a Reputação do Autor em comentários de produtos na Web utilizando Redes Neurais Artificiais

Carlos Augusto de Sá

Teresina-PI, Março de 2017

Carlos Augusto de Sá

**Abordagem para definir a Reputação do Autor em
comentários de produtos na Web utilizando Redes
Neurais Artificiais**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da UFPI (área de concentração: Sistemas de Computação), como parte dos requisitos para a obtenção do grau de Mestre em Ciência da Computação.

Universidade Federal do Piauí – UFPI

Centro de Ciências da Natureza

Programa de Pós-Graduação em Ciência da Computação

Orientador: Raimundo Santos Moura

Teresina-PI

Março de 2017

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca Setorial do CCN

S111a Sá, Carlos Augusto de.

Abordagem para definir a reputação do autor em comentários de produtos na web utilizando redes neurais artificiais / Carlos Augusto de Sá. – Teresina, 2017.
101f.: il.

Dissertação (Mestrado) – Universidade Federal do Piauí, Centro de Ciências da Natureza, Pós-Graduação em Ciência da Computação, 2017.

Orientador: Prof. Dr. Raimundo Santos Moura.

1. Sistemas de Computação. 2. Sistema Fuzzy. 3. Redes Neurais Artificiais. Título.


CDD 004.21

**Abordagem para Definir a Reputação do Autor em Comentários de Produtos
na Web Utilizando Redes Neurais Artificiais**

CARLOS AUGUSTO DE SÁ

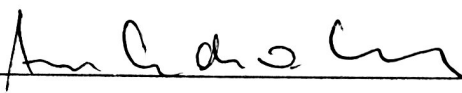
Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Natureza da Universidade Federal do Piauí, como parte integrante dos requisitos necessários para obtenção do grau de Mestre em Ciência da Computação.

Aprovada por:



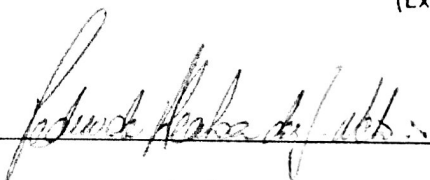
Prof. Raimundo Santos Moura

(Presidente da Banca Examinadora)



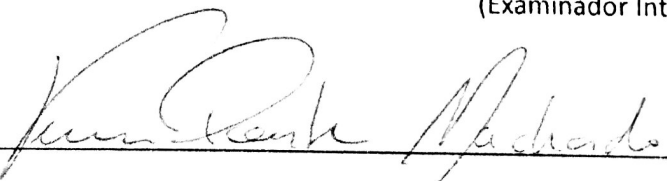
Prof. Anderson Luiz de Oliveira Cavalcanti

(Examinador Externo)



Prof. Pedro de Alcântara dos Santos Neto

(Examinador Interno)



Prof. Vinícius Ponte Machado

(Examinador Interno)

Teresina, 22 de março de 2017.

Ao meu Deus e à minha família.

Agradecimentos

Ao meu Deus, que sempre me dá forças para eu conseguir meus objetivos, independente das dificuldades. A minha mãe Antônia, pela mãe maravilhosa e presente, minha irmã Amanda pelo apoio nas revisões do texto.

A minha bela namorada Celecina pelo apoio e sempre acreditar em mim.

Ao professor e orientador Raimundo Santos Moura pelo suporte dado, sempre com alto conhecimento e paciência.

Ao meus amigos do grupo de pesquisa, em especial Roney Lira, que iniciou a caminhada junto comigo e deu um suporte imensurável ao meu trabalho. Cito também os amigos Antônio Luís, Jardeson, Allan, Gilvan e todos que fizeram parte do nosso grupo de pesquisa em PLN.

No Colégio Técnico de Teresina, agradeço em especial aos professores Sinimbu, Zé Bento, Ricardo Prado, Valdemir Junior, Jossivaldo por terem dado todo o apoio para eu poder realizar este curso de Mestrado.

Ao professor Gilvan, da COPESE, por ter me liberado das atividades e permitir que eu me qualificasse.

“Com ele há sabedoria e potência.”
(Jó 12:13)

Resumo

Com a popularização da Web 2.0, os serviços de comércio eletrônico permitiram uma interação dos seus usuários de forma bastante intensa. Estes serviços disponibilizam, aos seus usuários, meios de opinar e, também, diversos recursos para avaliar as opiniões de outros usuários sobre determinados produtos ou serviços. Atualmente, estas avaliações podem ser feitas através de votos (positivos ou negativos), ranking de estrelas, compartilhamentos ou curtidas, por exemplo. Com um alto volume de opiniões e recursos avaliativos, se torna difícil para os novos consumidores lerem tantos comentários para tomarem decisões a cerca de uma compra. Além disso, conhecer a reputação do autor dos comentários é de suma importância para não ser enganado em suas buscas por uma nova aquisição. Neste sentido, a reputação de um autor é considerada uma variável muito importante para avaliar um comentário na Web. No entanto, não existe uma definição formal sobre como calcular essa variável. Para inferir os melhores comentários sobre produtos ou serviços, Sousa (2015) propôs a abordagem Top(x) que utiliza um Sistema *Fuzzy* com três variáveis de entrada: reputação do autor, número de tuplas $\langle \textit{característica}, \textit{palavra opinativa} \rangle$ e riqueza de vocabulário; e uma variável de saída: grau de importância do comentário. Este trabalho apresenta uma adaptação enfatizando a medida de reputação do autor, que consiste em utilizar seis medidas fornecidas como entradas em uma Rede Neural Artificial *Perceptron* de múltiplas camadas. Em uma avaliação preliminar, a Rede Neural apresentou a acurácia de 91,01% no processo de classificação do autor. Adicionalmente, foi realizado um experimento para comparar as duas abordagens e os resultados mostraram que a abordagem adaptada obteve um ganho na classificação da importância dos comentários.

Palavras-chaves: Reputação do Autor, Redes Neurais Artificiais, Sistemas Fuzzy

Abstract

With the popularization of Web 2.0, the e-commerce services allowed a very intense interaction of its users. These services provide their users with a means of giving their opinions and also various resources to evaluate the opinions of other users about certain products or services. Currently, these ratings can be made by votes (positive or negative), stars ranking, shares or likes, for example. With a high volume of opinions and evaluative resources, it becomes difficult for new consumers to read so many comments to make decisions about a purchase. Also, knowing the author reputation for the reviews is of the utmost importance so as not to be misled in their quest for a new acquisition. In this context, author reputation is considered a very important variable for evaluating a comment on the Web. However, there is no formal definition on how to calculate this variable. In order to infer the best comments about products or services, Sousa(2015) proposed the Top(x) approach that uses a Fuzzy System with three input variables: author reputation, number of tuples <feature, quality word> and richness of vocabulary; and an output variable: degree of importance of the comment. This work presents an adaptation emphasizing the author reputation measure, which consists of using six measures provided as entries in a Artificial Neural Network Multilayer Perceptron. In a preliminary evaluation, the Neural Network presented the accuracy of 91.01% in the author's classification process. In addition, an experiment was carried out to compare the two approaches and the results showed that the adapted approach obtained a gain in the classification of the importance of the comments.

Keywords: Author Reputation, Artificial Neural Networks, Fuzzy Systems

Lista de ilustrações

Figura 1 – <i>Review</i> negativo no <i>site Amazon</i>	3
Figura 2 – Utilidade do <i>review</i> no <i>site Amazon</i>	3
Figura 3 – Abordagem TOP(X) proposta por Sousa (2015)	5
Figura 4 – Opinião a respeito de um celular	12
Figura 5 – Exemplo de opinião comparativa	14
Figura 6 – Neurônio artificial (SILVA; SPATTI; FLAUZINO, 2010)	19
Figura 7 – Arquitetura <i>feedforward</i> de camadas múltiplas (SILVA; SPATTI; FLAUZINO, 2010)	22
Figura 8 – Funções de pertinência para a variável temperatura (TANSCHKEIT, 2004)	24
Figura 9 – Formas de funções de pertinência	25
Figura 10 – Representação de um Sistema de Inferência Fuzzy (TANSCHKEIT, 2004)	25
Figura 11 – Tela de Configuração da Rede Neural na SPSS	28
Figura 12 – Gráfico da função de pertinência da variável Reputação do Autor	45
Figura 13 – Gráfico da função de pertinência da variável Quantidade de Tuplas	45
Figura 14 – Gráfico da função de pertinência da variável Corretude	46
Figura 15 – Gráfico da função de pertinência da saída Importância	46
Figura 16 – Abordagem proposta	47
Figura 17 – Topologia da RNA	51
Figura 18 – Topologia da Rede RBF	53
Figura 19 – Gráfico comparativo entre os comentários positivos	57
Figura 20 – Gráfico comparativo entre os comentários negativos	58
Figura 21 – Análise de erros	59

Lista de tabelas

Tabela 1 – Tabela de Avaliação Geral da medida AR	17
Tabela 2 – Módulos e Tarefas do NLTK	27
Tabela 3 – Córpus anotado quanto a orientação	42
Tabela 4 – Resultado da avaliação realizada por seres humanos quanto a Reputação do Autor	42
Tabela 5 – Avaliação realizada por seres humanos da importância dos comentários	43
Tabela 6 – Base de regras	45
Tabela 7 – Informação da Rede Neural	49
Tabela 8 – Importância das Variáveis de Entrada	51
Tabela 9 – Informação da Rede RBF	52
Tabela 10 – Importância das Variáveis de Entrada na Rede RBF	54
Tabela 11 – Matriz de confusão: abordagem TOP(X) original - Positivos	55
Tabela 12 – Matriz de confusão: abordagem TOP(X) com a RNA MLP - Positivos	55
Tabela 13 – Matriz de confusão: abordagem TOP(X) original - Negativos	56
Tabela 14 – Matriz de confusão: abordagem TOP(X) com a RNA MLP - Negativos	56
Tabela 15 – Comparação dos comentários positivos nas duas abordagens	56
Tabela 16 – Comparação dos comentários negativos nas duas abordagens	57

Lista de abreviaturas e siglas

API	<i>Application Programming Interface</i>
CC	<i>Collective Classification</i>
HITS	<i>Hyperlink-Induced Topic Search</i>
JSON	<i>JavaScript Object Notation</i>
NLTK	<i>Natural Language Toolkit</i>
PLN	Processamento de Linguagem Natural
RBF	<i>Radial Basis Functions</i>
REST	Representational State Transfer
RNA MLP	Rede Neural Artificial <i>Multilayer Perceptron</i>
RSO	Redes Sociais <i>Online</i>
SBC	Sociedade Brasileira de Computação
UFPI	Universidade Federal do Piauí
XML	<i>eXtensible Markup Language</i>

Lista de símbolos

\in	Pertence
β	Letra grega minúscula beta
θ	Letra grega minúscula teta
μ	Letra grega minúscula mi
Σ	Letra grega Sigma, somatório de números

Sumário

Introdução	1
Contexto e Motivação	1
Objetivos	5
Contribuições	6
Estrutura desta Dissertação	7
1 REFERENCIAL TEÓRICO	9
1.1 Processamento de Linguagem Natural	9
1.1.1 Análise de Sentimentos	11
1.2 Reputação do Autor	15
1.3 Aprendizagem de Máquina	17
1.3.1 Redes Neurais Artificiais	19
1.3.1.1 Métodos de Treinamento	22
1.3.2 Sistemas de Inferência Fuzzy	23
1.4 Ferramentas e Recursos	26
1.4.1 NLTK: Natural Language ToolKit	26
1.4.2 Etiquetador Mac-Morpho	27
1.4.3 SPSS	27
1.5 Considerações Finais	29
2 TRABALHOS RELACIONADOS	31
2.1 Ambientes Wiki	31
2.1.1 Histórico das edições	31
2.1.2 Contexto Social	33
2.1.3 Mecanismos de Recompensa	34
2.2 Twitter	34
2.3 Outros Sistemas Web	37
2.4 Considerações Finais	39
3 ABORDAGEM PROPOSTA	41
3.1 Coleta e Preparação do Córpus	41
3.1.1 Subcórpus I	42
3.1.2 Subcórpus II	43
3.2 Abordagem Top(x)	43
3.2.1 Variáveis de Entrada	43
3.2.2 Sistema Fuzzy	44

3.3	Abordagem proposta	46
3.4	Considerações Finais	48
4	EXPERIMENTOS E RESULTADOS	49
4.1	Metodologia	49
4.2	RNA MLP	49
4.3	Redes RBF	52
4.4	Comparação da abordagem original com a abordagem proposta	54
4.5	Análise de erros	58
4.6	Considerações Finais	59
	Conclusão e Trabalhos Futuros	61
	REFERÊNCIAS	63
	ANEXOS	71
	ANEXO A – ETIQUETAS MAC-MORPHO	73
	ANEXO B – BASE DE CARACTERÍSTICAS	75

Introdução

Neste Capítulo serão apresentados o contexto e motivação, bem como os objetivos pretendidos na realização deste trabalho e, por fim, a organização do presente texto.

Contexto e Motivação

Com o advento da Web 2.0, a prática de interagir na grande rede de computadores cresceu muito nos últimos anos. As Redes Sociais *Online* (RSO) como *Twitter*¹ e *Facebook*², bem como serviços de troca de mensagens pelo celular como o *Whatsapp*³, permitem grande interação social entre os usuários da grande rede. Além das RSOs, *sites* de comércio eletrônico, *sites* de hotéis, bem como outros *sites* permitem aos usuários deixarem as suas opiniões e comentários sobre a compra de um produto ou serviço.

A grande quantidade de informação que precisa ser tratada pelos usuários é um problema que surgiu com o crescimento da Web 2.0 e as mídias sociais *online*. Com este crescimento e com o avanço da Computação, a área de Processamento de Linguagem Natural (PLN) ganhou bastante destaque por permitir analisar grandes volumes de dados e realizar processamento de maneira mais eficiente, com o desenvolvimento de novas soluções para trabalhar com a complexidade dessas redes.

A tendência das pessoas formarem grupos e interagir é uma característica de qualquer sociedade (CASTELLS, 2000). A Web 2.0 possibilitou uma nova forma de interação e também influenciou o modo de pensar das pessoas. As interações se tornaram mais fáceis, permitindo um maior tráfego de informações, mais agilidade e um alcance mais amplo do que as mídias convencionais, como rádio e televisão.

Tendo em vista tal cenário, um dos tópicos de pesquisa da Ciência da Computação explorado nos últimos anos é o desafio de transformar textos não estruturados em conhecimento estruturado, de forma que um usuário possa se beneficiar melhor desse volume de informações (SBC, 2006). Neste contexto surgiu a área de pesquisa denominada Análise de Sentimentos ou Mineração de Opiniões, que se preocupa em extrair o sentimento a partir de descrições textuais (LIU, 2010a).

A Análise de Sentimentos é uma subárea de Processamento de Linguagem Natural (PLN), que envolve Ciência da Computação, Linguística e Inteligência Artificial e tem atacado o problema de manipular grandes volumes de informações através de técnicas que

¹ <http://twitter.com>

² <http://facebook.com>

³ <http://whatsapp.com>

analisam a linguagem escrita ou falada (JACKSON; MOULINIER, 2007). Um desafio da área se encontra na filtragem ou pré-processamento de comentários Web, já que existe uma tendência dos usuários de sistemas *online* a escrever com muitas gírias, o que dificulta o trabalho das ferramentas tradicionais de PLN. Por exemplo, no português temos a palavra "vc" indicando o pronome "você", "naum" e "nao" são usados como o advérbio "não". Além disso, a pouca preocupação com a grafia correta das palavras, ironias e sarcasmos dificultam a atuação dos sistemas baseados em PLN.

Dentro do contexto de comércio eletrônico, a comunicação boca-a-boca (do inglês "*word-of-mouth*" - WOM) tem uma grande influência no processo de aquisição de produtos e serviços no cotidiano dos consumidores. Anteriormente, as pessoas buscavam opiniões a respeito de um determinado produto ou serviço questionando amigos próximos ou parentes. Este tipo de comunicação envolve todas as formas de troca de informações entre os consumidores independente das características e utilidade de um produto ou serviço em particular (GOLDENBERG; LIBAI; MULLER, 2001).

No entanto, com o avanço da *Internet*, os serviços como redes sociais e *e-commerce* permitiram uma nova dimensão no processo de escolha por parte de novos consumidores. Com advento destes serviços, diversos usuários podem deixar os seus comentários sobre produtos das mais diversas formas, por exemplo, escrevendo, votando, comentando ou compartilhando opiniões (KIM; SRIVASTAVA, 2007). Desta forma, tem-se uma enorme base de dados para que os novos consumidores possam pesquisar e se beneficiar das opiniões de outros.

Segue que, os usuários de *sites* de *e-commerce*, por exemplo, passaram a comentar mais, devido aos diversos recursos de avaliação disponibilizados. Com mais comentários escritos, tem-se a necessidade de mecanismos automáticos para inferir os comentários mais úteis, vez que é praticamente impossível alguém ler todos os comentários disponíveis. Em uma pesquisa realizada pelo *site Trip Advisor*⁴, observou-se que 80% dos usuários leem de 6 a 12 *reviews* antes de reservarem um hotel e os mesmos estão interessados nos *reviews* mais recentes (TRIPADVISOR, 2013). Porém, é importante notar que nem sempre esta quantidade *reviews* será realmente a mais significativa no processo de escolha do consumidor.

Assim, diante de tamanha disponibilidade de informações presentes nesses ambientes, o novo consumidor tem uma árdua tarefa de buscar os *reviews* mais significativos para auxiliá-lo no seu processo de escolha, tornando assim uma tarefa desgastante (VOHS *et al.*, 2008). Diante deste cenário, um consumidor determinado a comprar um telefone celular, por exemplo, pode simplesmente desistir de realizar a compra devido a sobrecarga de informações presentes nesses sistemas *online* de avaliação.

⁴ Disponível em <http://tripadvisor.com>

Outro detalhe a ser considerado ao verificar o tamanho dessas base de dados é a quantidade de *spam*, textos de baixa qualidade, erros ortográficos, *emoticons*, "internetês"⁵ e informações falsas (LIU, 2011). Além destas considerações, algumas vezes os *reviews* apresentam sarcasmos e ironias, que são difíceis de serem captados pelas técnicas atuais de PLN. No entanto, existem esforços no sentido de resolver esses problemas, como os trabalhos de Hartmann et al. (2014), Carvalho et al. (2009) e Gonçalves et al. (2015).

De um lado, as avaliações positivas a respeito de um produto ou serviço trazem ao novo consumidor mais segurança no processo de compra, por outro lado, as avaliações negativas também podem auxiliar na escolha, gerando um impacto positivo na vendas (HAMILTON; VOHS; MCGILL, 2014). Geralmente, os *reviews* negativos são escritos de forma mais crítica e apresentando boa legibilidade, algumas vezes, até melhor que os *reviews* positivos. A Figura 1 apresenta um *review* negativo retirado do *site* da Amazon⁶.

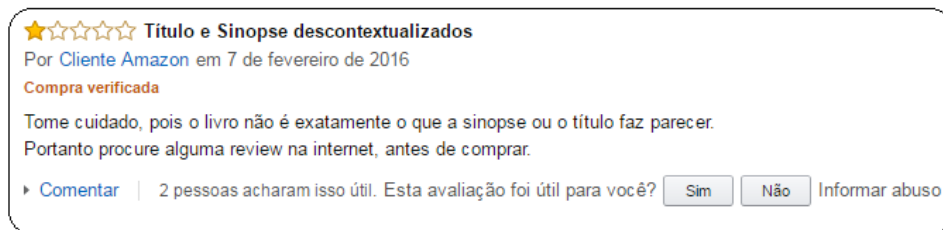


Figura 1 – *Review* negativo no *site* Amazon

Na Figura 2 é apresentado um *review* positivo contendo um recurso bastante interessante para avaliar um comentário, conhecido como utilidade do *review*. Com este recurso, o usuário pode, ao terminar de ler, marcar a opção indicando se o comentário foi útil para ele, desta maneira quanto mais votos "Sim" um *review* possuir, melhor ranqueado ele será. Porém, uma desvantagem dessa técnica é que comentários recentes, e que teriam significância ao novo consumidor, podem ser ignorados por terem poucos votos (LI et al., 2013).

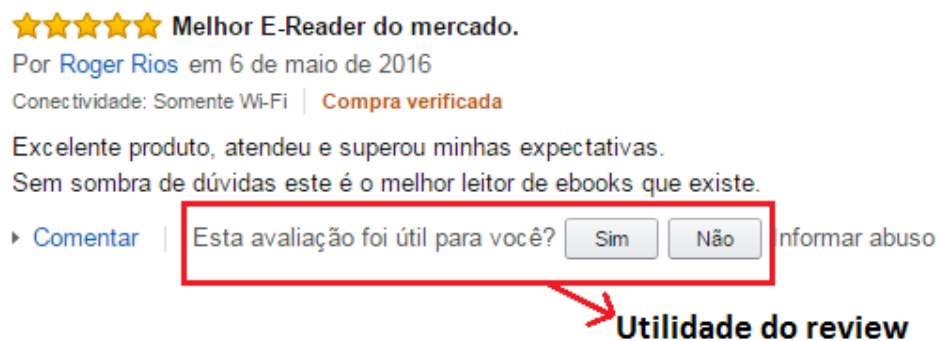


Figura 2 – Utilidade do *review* no *site* Amazon

⁵ Neologismo (de: Internet + sufixo ês) que designa a linguagem utilizada no meio virtual

⁶ <http://amazon.com>

Desde o início de suas operações como *site* de *e-commerce*, a *Amazon* sempre permitiu a introdução de *reviews* por parte dos usuários do serviço (inclusive alguns usuários que leem o comentário no próprio *site* da *Amazon* podem realizar a compra do produto desejado em outra loja). No entanto, com a introdução desta simples questão junto ao *review*: "Este review foi útil para você?", a *Amazon* conseguiu um aumento nos seus lucros de 2,7 bilhões de dólares anuais (SPOOL, 2009).

Alguns *sites* de *e-commerce* possuem sistemas de recomendação automáticos que, de acordo com os hábitos de navegação dos usuários, apresentam páginas personalizadas aos mesmos. Apesar de ser uma funcionalidade poderosa para o *marketing* das empresas de *e-commerce*, estes sistemas são menos eficientes do que as recomendações de amigos próximos ou parentes. Assim, é mais provável que os novos consumidores aceitem as recomendações de pessoas conhecidas do que dos sistemas de recomendações automáticos (SINHA et al., 2001).

A reputação do autor, em geral, implica na confiança que outras pessoas atribuem a um indivíduo em particular baseado nas contribuições intelectuais dentro de um grupo de interesse. Em diversas áreas, quando um autor tem alta credibilidade, as suas opiniões são altamente respeitadas pelos membros ao redor, e podem influenciar o modo de agir e pensar dos outros. Este trabalho foca em considerar diversas medidas sobre reputação do autor.

Dentro do contexto de reputação, os sistemas *online* apresentam diversos recursos para avaliação de um comentário, a saber: quantidade de comentários de um usuário, número de favoritos, quantidade de estrelas atribuídas, pontuação positiva e negativa, número de seguidores e amigos, entre outros. A partir de tais recursos um usuário pode considerar a reputação de um certo autor melhor do que a de outro. O valor dessa influência pode ser calculado baseado em recursos disponibilizados por esses sistemas *online*, que é um dos focos deste trabalho.

Com o objetivo de identificar os comentários mais relevantes, Sousa (2015) implementou a abordagem Top(x) para inferir os melhores comentários sobre produtos ou serviços. A abordagem Top(x) utiliza um Sistema *Fuzzy* com três variáveis de entrada: reputação do autor, número de tuplas $\langle \textit{característica}, \textit{palavra opinativa} \rangle$ e riqueza de vocabulário; e uma variável de saída: grau de importância do comentário, representado pela variável "k" (Figura 3). No entanto, para definir a reputação do autor, Sousa (2015) considerou somente a quantidade de comentários emitidos, ou seja, quanto mais comentários emitidos, melhor a reputação do autor. Destaca-se que essa hipótese é considerada fraca e pode ser refutada pois um *spammer*⁷ é considerado um bom autor. Portanto, este trabalho tem como objetivo analisar outras medidas para inferir a reputação de um autor de comentários na Web.

⁷ Usuários que postam muitas propagandas sem a permissão dos demais usuários

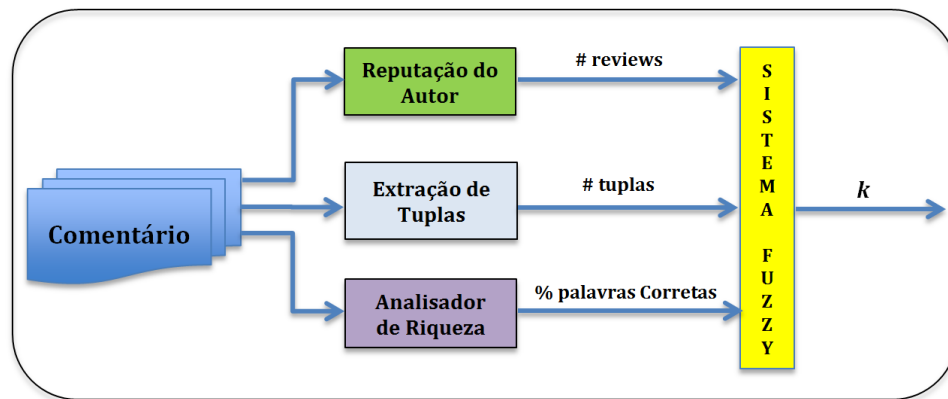


Figura 3 – Abordagem TOP(X) proposta por Sousa (2015)

Um entrave que os *sites* de *e-commerce* apresentam é o fato de possuírem cadastros independentes para os seus usuários e o acesso a estes dados não é permitido, dificultando o trabalho de coleta de informações adicionais. Desta forma, esta pesquisa sugere que tais *sites* possam rever as suas políticas de privacidade no futuro. Uma forma de resolver o entrave é ligar o perfil dos usuários com a suas contas em RSOs populares como *Twitter* e *Facebook*.

As RSOs levam vantagem sobre os *sites* de comentários no que se refere as informações sobre os autores das postagens. O *Facebook*, por exemplo, permite que os participantes possam curtir um conteúdo ou mesmo comentar. Destaca-se que as empresas tentam explorar ao máximo esse novo tipo *marketing*. Este fato se confirma com a grande quantidade de perfis nessa rede social. Outra rede que surgiu recentemente é a *Disqus*⁸, que possibilita debates intensos em *blogs* ou *sites* que possuem o seu *plugin*.

Visando melhorar a análise da variável reputação do autor, este trabalho apresenta um estudo utilizando uma Rede Neural Artificial *Multilayer Perceptron* (RNA MLP) para analisar um conjunto de medidas e definir quais são as mais relevantes no processo de avaliação da reputação do autor. Outro aspecto mostrado é uma comparação entre a abordagem Top(x) original com uma nova abordagem utilizando a Top(x) com uma RNA MLP na dimensão da reputação do autor. A hipótese é que aplicando esta RNA MLP com um conjunto de medidas eficazes para a reputação do autor se obtenha um desempenho superior na classificação da importância dos comentários, auxiliando os usuários em suas pesquisas para a aquisição de produtos ou serviços.

Objetivos

O objetivo principal deste trabalho é analisar medidas para definir a reputação do autor em comentários na Web, utilizando RNA. Foi realizado um experimento em um

⁸ <http://disqus.com>

Córpus⁹ coletado no *site* do Buscapé¹⁰ para extrair um conjunto de medidas que possam auxiliar na definição da reputação de um autor.

Este objetivo principal pode ser dividido nos seguintes objetivos específicos:

- Examinar um conjunto de medidas definidas em outros sistemas web, como os ambientes colaborativos *Wiki*, RSOs e foruns *online*;
- Discutir a relevância das medidas de reputação do autor no Córpus do *site* de *e-commerce* Buscapé;
- Analisar quais são as medidas mais importantes para avaliar a reputação do autor;
- Realizar uma comparação entre a abordagem original proposta por Sousa (2015) e a abordagem proposta nesta pesquisa.

É importante destacar que este trabalho está inserido em um projeto maior desenvolvido no Departamento de Computação da Universidade Federal do Piauí (UFPI) que explora Processamento de Linguagem Natural, Análise de Sentimentos e Aprendizagem de Máquina. A dissertação de Sousa (2015) representa o ponto de partida para a execução desta pesquisa, conforme detalhado no decorrer do desenvolvimento desta dissertação.

Contribuições

As principais contribuições desta pesquisa são listadas abaixo:

- Proposta de um modelo adaptado para inferir os comentários mais importantes sobre vendas de produtos e serviços na Web baseado na abordagem de Sousa (2015), tendo como enfoque a definição da reputação do autor. Para definir a reputação, foi analisado um conjunto de medidas consideradas relevantes no Córpus coletado. Feito isso, as medidas foram usadas como entrada em uma RNA MLP responsável por inferir a reputação dos autores de cada comentário analisado, e, por fim, a saída da rede neural foi usada como entrada em um Sistema de Inferência *Fuzzy* que forneceu o grau de importância dos comentários;
- Criação de um Córpus de pesquisa sobre celulares extraído do *site* do Buscapé¹¹. O Córpus possui 2000 comentários que foram anotados por voluntários do nosso grupo de pesquisa sobre PLN da Universidade Federal do Piauí. Inicialmente, o Córpus

⁹ Conjunto de documentos textuais escritos em um determinado idioma que serve como base para análise. O termo na forma apresentada substitui os termos em latim, *corpus/corpora*, nos grupos de pesquisa de PLN no Brasil.

¹⁰ <http://www.buscape.com.br>

¹¹ Disponível em <https://goo.gl/g5nrwJ>

inteiro foi anotado quanto a polaridade, os voluntários separaram os comentários em positivos, negativos, neutros ou lixo. Após isso, a anotação foi realizada quanto a reputação do autor, para tal, uma amostra de 356 comentários com os seus respectivos autores foi separada e um voluntário atribuiu notas de 0 a 10 a cada autor utilizando informações como: votos positivos, votos negativos e data do *review*, entre outras. Por fim, foi separada uma amostra de 271 comentários e uma avaliação realizada por especialista informava a importância de cada comentário utilizando como base a quantidade informações dos comentários, a riqueza do vocabulário e também informações não-textuais referentes a reputação do autor;

- Estudo comparativo realizado entre o modelo proposto nesta pesquisa e o modelo proposto por [Sousa \(2015\)](#) utilizando como critério de avaliação a média harmônica entre Precisão e Cobertura, a Medida-F.

Estrutura desta Dissertação

Além desta introdução, a estruturação desta dissertação contempla os seguintes Capítulos detalhados a seguir:

O Capítulo 1 apresenta o Referencial Teórico sobre as áreas de Processamento de Linguagem Natural, Análise de Sentimentos, Reputação do autor, Aprendizagem de Máquina, além dos recursos e ferramentas utilizados, preparando o leitor para os demais Capítulos.

O Capítulo 2 discute os Trabalhos Relacionados com esta pesquisa.

O Capítulo 3 descreve a Proposta de pesquisa deste trabalho com discussão acerca das medidas de reputação do autor.

O Capítulo 4 apresenta e discute os Resultados obtidos da abordagem proposta a partir de um experimento realizado no *Córpus do Buscapé*.

Por fim, no Capítulo Conclusão e Trabalhos Futuros são apresentadas as conclusões obtidas e propostas de trabalhos futuros.

1 Referencial Teórico

Neste Capítulo são apresentados os principais conceitos referentes a Processamento de Linguagem Natural, Análise de Sentimentos, Reputação do Autor, Aprendizagem de Máquina, além de ferramentas e recursos importantes para o entendimento geral deste trabalho.

1.1 Processamento de Linguagem Natural

Segundo [Jackson e Moulinier \(2007\)](#), o termo Processamento de Linguagem Natural (PLN) é usado para descrever as funções de componentes de *software* ou *hardware* responsáveis por analisar ou sintetizar a linguagem falada ou escrita. O fato de ser "natural" implica na distinção do discurso humano de linguagens formais, tais como as linguagens de programação, por exemplo Java, LISP e C.

Com a ascensão da *Web* e a quantidade de informações textuais sendo enviadas pela *Internet* diariamente resultou em um interesse na área de PLN já que a mesma é capaz de analisar um grande volume de informações, extrair os mais diversos padrões e sumarizar os resultados de forma a auxiliar, por exemplo, em tomadas de decisões pelas empresas e corporações.

Um dos problemas que o PLN deve tratar é a ambiguidade ([JACKSON; MOULINIER, 2007](#)). Por exemplo, a palavra "banco" pode se referir a uma instituição financeira ou a um local onde as pessoas sentam. De fato, isso é bem entendido pelo ser humano devido ao contexto em que a palavra está inserida, porém para o computador pode haver complicações. Este problema citado se relaciona a uma simples ambiguidade léxica, porém quando se deve tratar ambiguidades a nível sintático, mais complexo será para o computador realizar o processamento. Por exemplo na frase "João viu o homem no parque com o telescópio", o telescópio pertence a quem? A João? Ao parque? Ao homem no parque? Ou seja, uma frase que o ser humano poderia interpretar corretamente com certa facilidade, o computador teria muitos problemas para processar de forma correta. Fora esses problemas citados, ainda existem as figuras de linguagem que muitas vezes até para o ser humano causa confusão como ironias e sarcasmos.

Segundo [Manning e Schütze \(1999\)](#), as principais tarefas tratadas pela PLN são listadas a seguir:

- **Sumarização automática:** Produz um resumo legível de uma parte do texto. Muitas vezes usada para fornecer resumos de texto de um tipo conhecido, como artigos na seção financeira de um jornal;

- **Resolução de correferência:** Dada uma frase ou um pedaço maior de texto, determina quais palavras se referem aos mesmos objetos, ou entidades, por exemplo: "João é professor. Ele ensina informática", no caso o pronome "Ele" faz correferência a "João";
- **Análise do Discurso:** Identifica a natureza das relações discursivas entre sentenças, além de reconhecer e classificar os atos de fala em um pedaço de texto;
- **Tradução automática:** Responsável por traduzir automaticamente texto de uma linguagem humana para outra;
- **Segmentação morfológica:** Separa palavras em morfemas individuais e identifica classes de morfemas;
- **Reconhecimento de entidade nomeada:** Procura extrair e classificar as entidades mencionadas em um texto escrito em linguagem natural;
- **Geração de linguagem natural:** Converte informações de bancos de dados de computador ou intenções semânticas em linguagem humana legível;
- **Compreensão da linguagem natural:** Converte pedaços de texto em representações mais formais, como estruturas de lógica de primeira ordem, que são mais fáceis de manipular pelos programas de computador;
- **Reconhecimento óptico de caracteres (OCR):** Dada uma imagem que representa o texto impresso, determina o texto correspondente.
- **Marcação de classe gramatical:** Dada uma sentença, determina a classe gramatical de cada palavra;
- **Análise sintática (*Parsing*):** Determina a árvore de análise sintática de uma frase;
- **Perguntas e Respostas:** Dada uma questão de linguagem humana, determina sua resposta. Exemplos dessas aplicações são os *chatbots*, que são programas de computadores que tentam simular um ser humano na conversação com as pessoas;
- **Extração de relacionamento:** Identifica as relações entre entidades nomeadas (por exemplo, quem é casado com quem) com base em textos;
- **Quebra de frases:** Encontra os limites da frase em um pedaço de texto. Os limites de frases são normalmente marcadas por pontos ou outros sinais de pontuação;
- **Reconhecimento de fala:** Dado um trecho de áudio de uma pessoa ou pessoas falando, determina a representação textual do discurso;

- **Segmentação de fala:** Dado um trecho de áudio de uma pessoa ou pessoas falando, separa-o em palavras;
- **Análise morfológica e reconhecimento de tópicos:** Dado um pedaço de texto, separa-o em segmentos, cada um dos quais é dedicado a um tópico e identifica o tópico do segmento;
- **Análise morfológica e segmentação de palavras:** Separa um pedaço de texto contínuo em palavras separadas;
- **Desambiguação:** Muitas palavras têm mais de um significado, esta tarefa busca o significado que faz mais sentido no contexto;
- **Recuperação de informação (IR):** Trata-se de armazenar, pesquisar e recuperar informações;
- **Extração de informação (IE):** Trata-se da extração de informação semântica a partir do texto.

Destaca-se que o nosso trabalho não se enquadra em nenhuma das tarefas citadas acima. Na verdade, contempla ideias de diversas tarefas, como: extração de informações, sumarização e classificação de comentários.

Uma subárea da PLN é a Análise de Sentimentos, também conhecida como Mineração de Opiniões, que consiste no estudo computacional das opiniões, avaliações, atitudes e emoções das pessoas referenciando entidades, indivíduos, questões, eventos, tópicos e seus atributos. O estudo é realizado sobre a emoção do usuário em relação ao item, classificando a emoção em positiva, negativa ou neutra. Mais detalhes são discutidos a seguir.

1.1.1 Análise de Sentimentos

De acordo com [Liu \(2010b\)](#), as informações textuais podem ser divididas em dois tipos principais: fatos e opiniões, em que:

- **Fatos:** se referem a expressões objetivas a respeito de entidades, eventos e suas propriedades;
- **Opiniões:** são expressões caracterizadas pela subjetividade, externadas pelo sentimento dos usuários a respeito de entidades, eventos e suas propriedades.

De modo geral, a área de Análise de Sentimentos determina se um trecho de texto possui polaridade positiva, negativa ou neutra. Adicionalmente, pode-se apresentar a polaridade em escalas, como muito positivo ou moderadamente positivo, entre outros

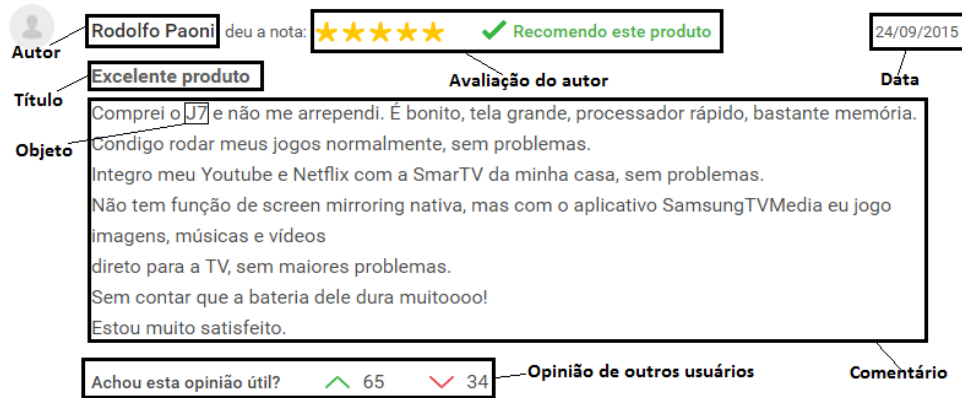


Figura 4 – Opinião a respeito de um celular

(THET; NA; KHOO, 2010). A Análise de Sentimentos utiliza técnicas das áreas de PLN e linguística computacional.

Os conceitos básicos da Análise de Sentimentos são:

- **Objeto:** é uma entidade que pode ser um produto, serviço, pessoa, evento, organização ou tópico. Por exemplo, na Figura 4, o objeto é o aparelho celular "Samsung Galaxy J7".
- **Atributo:** é um aspecto de um objeto que pode ser referenciado no texto. Por exemplo, na Figura 4, os atributos que podem ser identificados são: "tela", "processador", "bateria", entre outros.
- **Autor:** é a pessoa ou organização que emite a opinião. Por exemplo, na Figura 4, o autor identificado é "Rodolfo Paoni".
- **Opinião:** é a avaliação que um autor expressa sobre um objeto e seus atributos. Por exemplo, na Figura 4, tem-se a seguinte opinião: "tela grande, processador rápido, bastante memória", referindo-se a alguns atributos do aparelho celular em questão.
- **Orientação da Opinião:** é a polaridade da opinião: se positiva, negativa ou neutra. Por exemplo, na Figura 4, no fim do texto, o autor expressa: "Estou muito satisfeito", orientando positivamente a sua opinião acerca do produto.
- **Palavras opinativas:** são as palavras que qualificam o objeto e seus atributos. Por exemplo, na Figura 4, o autor expressa como palavras opinativas: "bonito", "rápido" e "bastante". Geralmente as palavras opinativas pertencem a classe gramatical dos adjetivos ou advérbios.
- **Tempo:** instante no qual a opinião foi expressa. Por exemplo, na Figura 4, a data "24/09/2015" indica o dia, mês e o ano que a opinião foi escrita.

- **Avaliação do autor:** avaliação geral dada pelo autor, normalmente representada por estrelas. Por exemplo, na Figura 4, o autor deu nota máxima ao produto marcando as cinco estrelas na avaliação.
- **Opinião de outros usuários:** quantidade de pessoas que acharam o *review* útil e quantidade que marcaram como não útil. Por exemplo, na Figura 4, tem-se 65 usuários considerando o *review* útil e 34 como não útil.

Liu (2010b) define dois tipos de opinião e alguns subtipos, conforme apresentado abaixo:

- **Opinião regular:** opinião que faz referência a somente um objeto. Pode possuir dois subtipos:
 - **Direta:** a opinião é expressa diretamente sobre a entidade ou algum aspecto da mesma. A Figura 4 expressa opinião regular direta.
 - **Indireta:** a opinião é expressa indiretamente sobre a entidade ou algum aspecto da mesma. Por exemplo, na frase "Após tomar o remédio, me senti pior" pode descrever um efeito indesejado do remédio "me senti pior", que indiretamente atribui um sentimento negativo ao alvo da opinião, no caso o "remédio".
- **opinião comparativa:** expressam relações de similaridade ou diferenças entre dois ou mais objetos. A Figura 5 expressa opinião comparativa.

Além desses tipo citados, Liu (2010b) também apresenta outros dois tipos:

- **Opinião explícita:** opinião regular ou comparativa que apresenta uma declaração subjetiva e explícita com o auxílio de palavras de sentimento, por exemplo: "O preço do J7 é alto" (opinião regular) ou "O preço J7 é maior do que o Iphone 6s" (opinião comparativa). Nos exemplos citados, é possível observar claramente o aspecto em que a opinião foi dada, "preço".
- **Opinião implícita:** opinião regular ou comparativa que apresenta uma declaração objetiva, frequentemente expressa sobre um fato desejado ou indesejado, por exemplo: "Esse celular descarrega rapidamente" (opinião regular) ou "A vida útil da bateria dos celulares da Samsung é maior do que dos celulares da Nokia" (opinião comparativa). Nos dois exemplos citados, é exigido do leitor um certo conhecimento do mundo para o mesmo perceber que a bateria de um celular com pouco duração é algo negativo.

Uma opinião é definida formalmente por Liu (2010b) como uma quintupla $(e_i, a_{ij}, s_{ijkl}, h_k, t_l)$, onde:

Excepcional

Excepcional é como eu descreveria esse Smartphone dos aparelhos que já tive como Gran Prime da Samsung e Moto G 2ª geração esse dá um banho, super rápido o Android roda liso sem travamentos a única ressalva é a bateria que não tem autonomia do Gran Prime, mas é melhor do que a do Moto G com certeza podem comprar sem medo de ser feliz.



Achou esta opinião útil?  38  24

Figura 5 – Exemplo de opinião comparativa

- e_i : é o nome de uma entidade;
- a_{ij} : é um aspecto da entidade e_i ;
- s_{ijkl} : é o sentimento sobre o aspecto a_{ij} da entidade e_i , emitido por h_k no instante t_l ;
- h_k : é o autor do sentimento;
- t_l : é o instante no qual a opinião foi expressa por h_k .

É importante observar que os aspectos da entidade são opcionais, pois as vezes os usuários podem simplesmente expressar o sentimento sobre o produto (entidade) ou sobre algum aspecto específico do mesmo, por exemplo: "*O telefone celular é muito bom*". Neste caso o sentimento está sendo emitido sobre a entidade, já em "*O display do telefone é muito resistente*", o sentimento recai sobre um aspecto do produto, a saber: "*O display do telefone*".

No caso da opinião comparativa, a mesma pode expressar uma relação de similaridades ou diferenças entre dois ou mais objetos, ou mesmo entre os aspectos compartilhados destes objetos. Apesar de existir um relacionamento com a opinião direta, estes dois tipos de opiniões apresentam diferenças, tanto semânticas quanto sintáticas. Por exemplo, uma opinião regular direta comum "*A qualidade da foto da câmera é ótima*", enquanto uma opinião comparativa "*A qualidade da foto da câmera X é melhor do que a da câmera Y*".

O emissor da opinião (do inglês: "*opinion holder*") é também conhecido como fonte do discurso ou fonte da opinião (WIEBE; CARDIE, 2005a). Ele pode ser um indivíduo que emite a opinião de forma pessoal a respeito de uma entidade ou mesmo um porta-voz de uma empresa e tem papel fundamental no processo de avaliar o sentimento de uma opinião. Neste trabalho, o emissor da opinião é objeto essencial para a realização desta pesquisa.

Um detalhe importante no que diz respeito a opiniões é que elas podem apresentar diferentes níveis de força (WILSON; WIEBE; HWA, 2004). Por exemplo, em "*Este telefone é uma sucata*" tem força maior do que "*Este telefone é ruim*". Esta força pode ser interpretada de forma gradativa, por exemplo, uma expressão positiva pode ser expressa com um

sentimento *contente*, *feliz*, *jubiloso* ou *maravilhado*, partindo de um valor mensurável de baixa intensidade (*contente*) até um valor de máxima intensidade (*maravilhado*). Uma aplicação prática poderia dividir estes sentimentos em dois grupos: *contente* e *feliz* em um nível de força, e *jubiloso* e *maravilhado* em outro (LIU, 2010b).

Liu (2010b) distingue a análise de opiniões em diferentes níveis de granularidade:

- **Documento:** visa trabalhar um texto coletado como um todo, categorizando-o com seu respectivo sentimento. Cada texto extraído representa uma opinião.
- **Sentença:** a partir de um dado texto (i.e., documento), este nível é responsável por extrair as sentenças contidas no mesmo e verificar se possuem polaridade positiva, negativa ou neutra. Pode-se também apresentar a polaridade em níveis de escalas, por exemplo, uma sentença poderia ter uma polaridade positiva em um nível mais elevado do que outra.
- **Aspecto ou Característica:** é o nível de granularidade mais fino no processo de análise de opiniões. Baseia-se no princípio de que uma opinião consiste basicamente de um sentimento e um alvo.

A importância da área de Análise de Sentimentos para este trabalho está na aplicação de suas técnicas na abordagem Top(x) de Sousa (2015). Na proposta de continuidade desta pesquisa, será realizado um experimento similar, no entanto com o foco na avaliação da reputação do autor do comentário.

1.2 Reputação do Autor

Atualmente, não existe uma definição formal para reputação de autor na Web consolidada pela comunidade. Porém, algumas medidas têm sido propostas em sistemas de avaliação. Segundo Li et al. (2014), a reputação de um usuário em um sistema de avaliação pode ser medida de acordo com as informações postadas por tal usuário, desta forma, quanto mais avaliações justas e confiáveis forem escritas, melhor a reputação do autor. Adicionalmente, os autores alertam para a importância das abordagens minimizarem ao máximo a atuação dos *spammers*, que são usuários que postam propagandas sem a permissão dos demais usuários.

Segue que, aspectos não-textuais como a reputação do autor são de grande importância no processo de escolha de um produto ou serviço por parte dos usuários. De fato, os consumidores estão bastante atentos a qualidade e exposição do autor dos comentários (HU; LIU; ZHANG, 2008). Desta forma, os *reviews* de um autor com alta reputação apresentam mais possibilidades de serem levados em conta no momento das escolhas dos novos consumidores.

Para Jones, Hesterly e Borgatti (1997), a reputação envolve uma estimativa do caráter, habilidades e confiabilidade de um indivíduo. Segundo os autores, a reputação reduz o comportamento de incerteza por prover informações a respeito da confiabilidade e boa vontade dos outros. Na área de governança corporativa, uma boa reputação traz consequências econômicas para as empresas.

Deephouse (2000) define reputação como a avaliação de uma empresa por seus interessados em termos de seus efeitos, estima e conhecimento. Hall (1993) descreve reputação como ativo intangível de uma empresa. Já para Caudron (1997), a reputação é construída não somente por palavras, mas principalmente por meio de atos. Segundo Heinonen (2011), uma medida bem definida do grau de reputação não existe.

Wöhner, Köhler e Peters (2011) citam os Sistemas de Reputação como alternativas para encontrar problemas de qualidade em artigos postados na *Wikipedia*. Tais sistemas são capazes de encontrar e analisar ações passadas de um autor e estimar o comportamento futuro de seus participantes. No ambiente colaborativo da *Wikipedia*, a qualidade geral de um artigo pode expressar a reputação de todos os autores envolvidos no desenvolvimento do mesmo.

Considerando a rede social *Twitter*, Weitzel, Oliveira e Quaresma (2013) indicam que a reputação de um usuário pode ser considerada credível, de acordo com a qualidade da informação disseminada na sua rede de amigos. Os pesquisadores observaram que qualidade das informações postadas são de grande preocupação, especialmente nos domínios da medicina e assistência médica.

Para Sabater e Sierra (2002), a crença na reputação de terceiros pode ser compartilhada e comunicada pelos membros de uma sociedade. Os autores definiram "reputação evidenciada" como a reputação que um agente constrói a respeito de outro agente baseado nos conselhos adquiridos por outros membros da sociedade. Em um mundo ideal, onde os agentes são homogêneos e confiáveis, "a reputação evidenciada" é tão relevante quanto uma informação direta entre os interessados.

Raub e Weesie (1990) argumentam que a reputação pode emergir em relações contínuas, especialmente se as informações a respeito do comportamento de um determinado agente, em uma de suas relações, se disseminam para outros parceiros através de uma rede de informações. De fato, as ações que elevam a reputação de um agente podem fazer com que o mesmo se abstenha de atos ilícitos e má conduta em suas relações sociais.

O *site* Reclame Aqui¹ define uma medida de reputação para avaliar as empresas que são alvos de reclamações por parte dos seus clientes, conforme a fórmula abaixo:

$$AR = ((IR * 2) + (MA * 10 * 3) + (IS * 3) + (IN * 2)) / 100 \quad (1.1)$$

Onde:

¹ <http://www.reclameaqui.com.br/como-funciona/reputacao/>

- **Índice de Resposta (IR):** Porcentagem de reclamações respondidas, sendo que apenas a primeira resposta é considerada;
- **Média das Avaliações (Nota do Consumidor) (MA)** - Leva em consideração apenas reclamações finalizadas e avaliadas. Corresponde à média aritmética das notas (variando de 0 a 10) concedidas pelos clientes para avaliar o atendimento recebido;
- **Índice de Solução (IS)** - Leva em consideração apenas reclamações finalizadas e avaliadas. Corresponde à porcentagem de reclamações onde os consumidores, ao finalizar, consideraram que o problema que originou a reclamação foi resolvido;
- **Índice de Novos Negócios (Voltaria a fazer negócios?) (IN)** - Leva em consideração apenas reclamações finalizadas e avaliadas. Corresponde à porcentagem de reclamações onde os consumidores, ao finalizar, informaram que, sim, voltariam a fazer negócios com a empresa.

Para a medida **Avaliação do Reclame Aqui (AR)**, as empresas são avaliadas conforme os intervalos apresentados na Tabela 1.

Tabela 1 – Tabela de Avaliação Geral da medida AR

Avaliação Geral	Intervalo
Ótimo	AR entre 8 e 10
Bom	AR entre 7 e 7,9
Regular	AR entre 6 e 6,9
Ruim	AR entre 5 e 5,9
Não recomendada	AR menor que 5

Este trabalho considera que a reputação do autor pode ser calculada a partir de um conjunto de seis medidas, abordadas com mais detalhes na Capítulo 3. Como proposta desta pesquisa, uma Rede Neural Artificial foi implementada com o objetivo de apresentar uma medida final para reputação, ao mesmo tempo que destaca as medidas mais importantes.

1.3 Aprendizagem de Máquina

Segundo Samuel (1959), a Aprendizagem de Máquina é uma área de estudos que dá aos computadores a habilidade de aprender sem ser explicitamente programados. Neste contexto, Luger (2005) explica que a capacidade de computadorizar o aprendizado é uma tarefa bastante complexa, além do tratamento da Linguagem Natural.

O processo de aprendizado envolve a generalização a partir da experiência, onde o agente inteligente melhora a sua performance não somente com a repetição de uma tarefa

mas também com outras tarefas similares no domínio. A habilidade de aprender deve ser parte de qualquer sistema dito inteligente e tais sistemas devem ser capazes de se adaptar através de suas interações no ambiente (LUGER, 2005).

A aprendizagem de máquina pode ser dividida em três categorias básicas:

- **Aprendizagem Supervisionada:** A aprendizagem é realizada a partir de exemplos através da definição de classes e dos exemplos em cada classe;
- **Aprendizagem não Supervisionada:** Realizada com base em observação e descoberta. Resulta em um conjunto de descrições de classes, uma para cada classe descoberta no ambiente;
- **Aprendizagem por Reforço:** O sistema inteligente deve transformar situações do mundo em ações com o objetivo de maximizar as medidas de recompensa.

Entre as principais técnicas de inteligência computacional ou sistemas inteligentes, têm-se:

- **Redes Neurais Artificiais:** modelos computacionais inspirados pelo sistema nervoso central de um animal que são capazes de realizar o aprendizado de máquina bem como o reconhecimento de padrões;
- **Sistemas Fuzzy:** modelos computacionais que utilizam a lógica *Fuzzy* e os conjuntos *Fuzzy* com o objetivo de modelar conceitos imprecisos ou aproximados que podem ser expressos em termos linguísticos;
- **Computação Evolutiva:** compreende um conjunto de técnicas de busca e otimização inspiradas na evolução natural das espécies de Darwin (1859);
- **Inteligência Coletiva:** um princípio onde as inteligências individuais são somadas e compartilhadas. Análises de dados complexas podem ser gerados a partir deste conceito;
- **Sistemas Imunológicos Artificiais:** compostos por metodologias inteligentes inspiradas no sistema imunológico humano para a solução de problemas do mundo real;
- **Agentes Inteligentes:** capazes de perceberem seu ambiente por meio de sensores e de agir sobre esse ambiente por meio de atuadores.

As próximas duas subseções abordam com mais detalhes as duas técnicas de aprendizagem de máquina que foram aplicadas neste trabalho.

1.3.1 Redes Neurais Artificiais

Segundo [Silva, Spatti e Flauzino \(2010\)](#), as Redes Neurais Artificiais (RNA) são técnicas de inteligência computacional utilizadas nas mais diversas áreas para solucionar diversos problemas. A ideia central desta técnica é simular um sistema nervoso como forma de inferir uma saída desejada baseado em variáveis de entrada. Entre as principais aplicações das RNA, os autores destacam:

- Avaliação de imagens captadas por satélites;
- Classificação de padrões de escrita e fala;
- Reconhecimento de faces em visão computacional;
- Controle de trens de grande velocidade;
- Previsão de ações no mercado financeiro;
- Identificação de anomalias em imagens médicas;
- Identificação automática de perfis de crédito para clientes de instituições financeiras;
- Controle de aparelhos eletrônicos e eletrodomésticos, como máquinas de lavar roupas, fornos de microondas, geladeiras, cafeteiras, fritadeiras, filmadoras, etc.

As RNAs possuem capacidade de aquisição e manutenção do conhecimento, que podem ser definidas por um conjunto de neurônios artificiais, interligados por diversas interconexões, as sinapses artificiais. Tais interconexões são representadas por vetores e matrizes de pesos sinápticos.

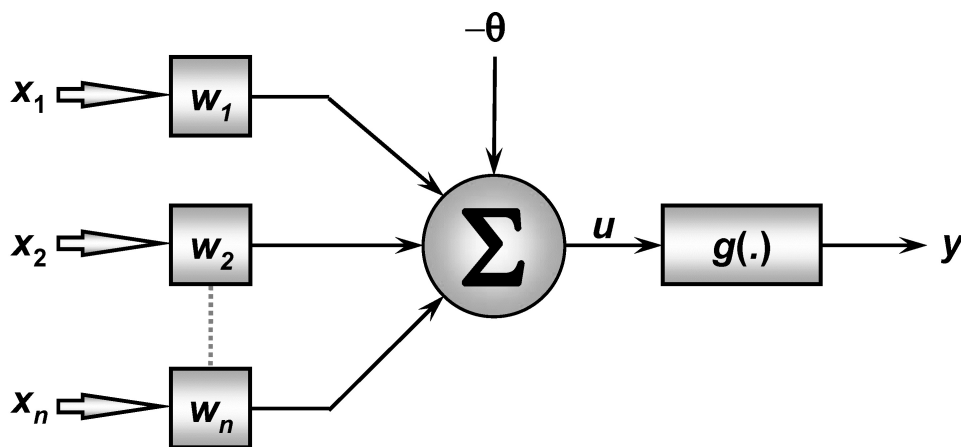


Figura 6 – Neurônio artificial ([SILVA; SPATTI; FLAUZINO, 2010](#))

O modelo de neurônio artificial mais utilizado é o proposto por [McCulloch e Pitts \(1943\)](#) e é apresentado na Figura 6. Em que:

- x_1, x_2, \dots, x_n : conjunto de variáveis de entrada no neurônio artificial;
- w_1, w_2, \dots, w_n : conjunto de pesos sinápticos que pondera os sinais de entrada;
- θ : limiar de ativação;
- Σ : realiza a soma ponderada dos sinais de entrada subtraindo-se do θ ;
- u : potencial de ativação;
- $g(\cdot)$: função de ativação;
- y : saída da rede neural.

Silva, Spatti e Flauzino (2010) citam as principais características das RNA, que seguem:

- **Adaptação por experiência:** as adaptações dos pesos sinápticos da rede são ajustados a partir da apresentação sucessiva de exemplos, tais como padrões, amostras, medidas, os quais são relacionados ao comportamento do processo, possibilitando a aquisição do conhecimento por aquisição;
- **Capacidade de aprendizado:** através de um método de treinamento, a rede consegue extrair o relacionamento existente entre as diversas variáveis que compõem a aplicação;
- **Habilidade de generalização:** após o processo de treinamento, a rede consegue estimar soluções até então desconhecidas;
- **Organização de dados:** a rede é capaz de realizar a sua organização interna visando possibilitar o agrupamento de padrões que apresentam particularidades em comum;
- **Tolerância a falhas:** a rede torna-se tolerante a falhas mesmo se parte da sua estrutura interna for corrompida;
- **Armazenamento distribuído:** o processamento realizado pela rede é distribuído entre as suas diversas sinapses artificiais, permitindo um incremento da robustez da arquitetura frente a eventuais neurônios que se tornam inoperantes;
- **Facilidade de prototipagem:** a implementação da maioria das arquiteturas das RNA podem ser facilmente prototipadas em *hardware* ou em *software*.

Silva, Spatti e Flauzino (2010) destacam também as potenciais áreas de aplicabilidade das RNAs:

- **Aproximador universal de funções:** o objetivo consiste em mapear o relacionamento funcional entre as variáveis de um sistema a partir de um conjunto conhecido de seus valores representativos;
- **Controle de processos:** o objetivo consiste em identificar ações de controle que permitam o alcance de requisitos de qualidade, eficiência e segurança do processos;
- **Reconhecimento/classificação de padrões:** o objetivo consiste em associar um padrão de entrada para uma das classes previamente definidas;
- **Agrupamento de dados:** o objetivo nesta circunstância consiste da identificação e detecção de similaridades entre os diversos padrões de entrada a fim de possibilitar o seu agrupamento;
- **Sistemas de previsão:** o objetivo consiste em estimar valores futuros de um processo levando-se em consideração diversas medidas prévias observadas em seu domínio;
- **Otimização de sistemas:** o alvo consiste em minimizar ou maximizar uma função custo obedecendo também eventuais restrições que são impostas para o correto mapeamento do problema;
- **Memórias associativas:** o objetivo consiste em recuperar padrões corretos mesmo se os seus elementos constituintes forem apresentados de forma incorreta ou imprecisa.

Silva, Spatti e Flauzino (2010) explicam que uma rede neural pode ser dividida em três ou mais camadas, que são nomeadas da seguinte forma:

1. **Camada de entrada:** responsável pelo recebimento das informações;
2. **Camadas escondidas ou intermediárias:** compostas de neurônios artificiais que possuem a responsabilidade de extrair as características associadas ao processo;
3. **Camada de saída:** também constituída de neurônios artificiais responsáveis pela produção a apresentação dos resultados finais da rede.

Com relação a arquitetura, as RNAs podem ser: redes *feedforward* (alimentação à frente) de camada simples, redes *feedforward* de camadas múltiplas, redes recorrentes e redes reticuladas. Neste trabalho, a arquitetura que melhor se adequa ao problema é a *feedforward* de camadas múltiplas, que pode ser observada na Figura 7, pois ela resolve problemas de reconhecimento e classificação de padrões, além de ser considerada uma das arquiteturas mais versáteis quanto à aplicabilidade.

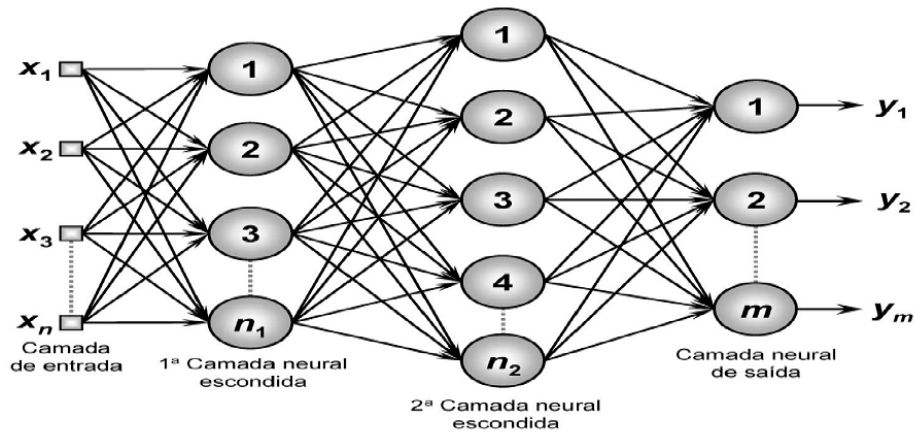


Figura 7 – Arquitetura *feedforward* de camadas múltiplas (SILVA; SPATTI; FLAUZINO, 2010)

A arquitetura *feedforward* de camadas múltiplas é constituída por uma ou mais camadas escondidas. As RNAs são empregadas na solução de diversos tipos de problemas, tais como aqueles relacionados a aproximação de funções, classificação de padrões, identificação de sistemas, otimização, robótica, ao controle de processos, etc.

1.3.1.1 Métodos de Treinamento

Nesta subseção serão apresentados alguns métodos de treinamento para Redes Neurais Artificiais *Perceptron* de Múltiplas Camadas.

Inicialmente, tem-se o algoritmo **Backpropagation** padrão, em que a minimização é geralmente feita com um passo a fixo. A determinação do passo α é essencial, pois para valores baixos, o tempo de treinamento pode ficar muito elevado, e para valores altos os parâmetros podem divergir (HAYKIN, 1994). A velocidade de convergência geralmente é melhorada adicionando-se um termo de momento (RUMELHART; HINTON; WILLIAMS, 1986).

$$\theta_{i+1} = \theta_0 + \alpha_i d_i + \beta_i \Delta \theta_{i-1}, i \geq 0. \quad (1.2)$$

Este termo adicional geralmente evita oscilações no comportamento do erro, pois pode ser interpretado como a inclusão de uma aproximação da informação de segunda ordem (HAYKIN, 1994).

Já no **Método do gradiente**, dentre os que utilizam diferenciação e busca, é considerado o mais simples para obtenção da direção d_i , pois utiliza apenas informações de primeira ordem. Na i -ésima iteração, a direção d_i é definida como a direção de módulo

unitário de maior decrescimento da função J .

$$d = -\frac{\nabla J(\theta)}{\|\nabla J(\theta)\|} \quad (1.3)$$

A lei de ajuste do método é dada por:

$$\theta_{i+1} = \theta_i - \alpha_i \frac{\nabla J(\theta_i)}{\|\nabla J(\theta_i)\|} \quad (1.4)$$

Outro método a se destacar é o de **Levenberg-Marquardt**, eficiente para redes que não possuem mais do que algumas centenas de conexões a serem ajustadas (HAGAN; MENHAJ, 1994). Isto ocorre por que estes algoritmos necessitam armazenar uma matriz quadrada cuja dimensão é da ordem do número de conexões da rede. Segue a sua definição matemática:

$$J(\theta) = \sum_{i=1}^N \sum_{j=1}^m (g_{ij} - \hat{g}_{ij}(x, \theta))^2 = \sum_{k=1}^q r_l^2 \quad (1.5)$$

onde N é o número de amostras, l o número de unidades intermediárias, r o erro residual, m o número de saídas, e q o produto $N \times m$.

1.3.2 Sistemas de Inferência Fuzzy

Os seres humanos lidam diariamente com informações imprecisas ou aproximadas que podem ser expressas em termos linguísticos, por exemplo: avaliar a temperatura do ambiente (frio, morno, quente), a altura de uma pessoa (baixo, alto, mediano) ou o estado sentimental (radiante, feliz, triste). Estes termos linguísticos ou conceitos não-quantificáveis podem ser implementados computacionalmente utilizando a Teoria dos Conjuntos *Fuzzy* e os Conceitos de Lógica *Fuzzy* (ZADEH, 1965).

Os Sistemas de Inferência *Fuzzy* têm sido aplicados em diversas áreas, tais como: problemas de classificação, mineração de dados, previsão de séries, planejamento e otimização (TANSCHKEIT, 2004). A abordagem original apresentada nesta dissertação também faz uso de um Sistema de Inferência *Fuzzy* aplicado para avaliar a importância de comentários postados em *sites* de *e-commerce* (SOUSA, 2015).

Os Conjuntos *Fuzzy* são modelados a partir variáveis linguísticas. Por exemplo, a temperatura pode ser uma variável linguística assumindo os valores baixa, média e alta, que podem ser representados por funções de pertinência, conforme observado na Figura 8.

Tanscheit (2004) destaca que, de forma geral, os valores assumidos pelas variáveis linguísticas podem ser construídos por termos primários de uma linguagem especificada (alto, baixo, médio, grande, pequeno, por exemplo), por conectores lógicos (e, ou e não),

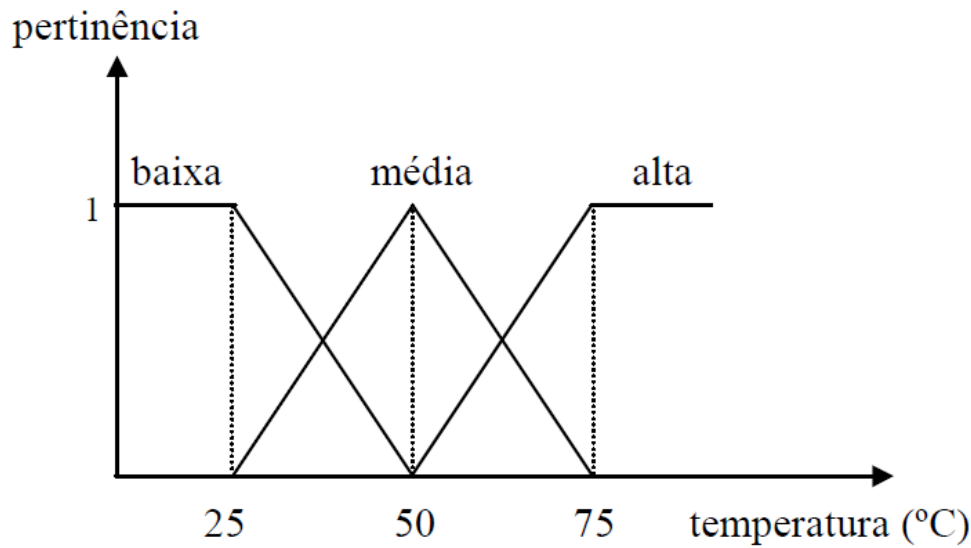


Figura 8 – Funções de pertinência para a variável temperatura (TANSCHKEIT, 2004)

por modificadores (muito, pouco, levemente, por exemplo) e por delimitadores (parênteses). Formalmente, uma variável linguística é caracterizada por uma quintupla $(N, T(N), X, G, M)$, onde:

- N : nome da variável;
- $T(N)$: conjunto de termos de N , ou seja, o conjunto de nomes dos valores linguísticos de N ;
- X : universo de discurso;
- G : regra sintática para gerar os valores de N como uma composição de termos de $T(N)$, conectivos lógicos, modificadores e delimitadores;
- M : regra semântica, para associar a cada valor gerado por G um conjunto *Fuzzy* em X .

No caso da variável temperatura da Figura 8, ter-se-ia:

- N : temperatura;
- $T(N)$: baixa, média, alta;
- X : 0 a 100 °C (por exemplo);
- G : temperatura não baixa e não muito alta, por exemplo;
- M : associa o valor acima a um conjunto *Fuzzy* cuja função de pertinência exprime o seu significado.

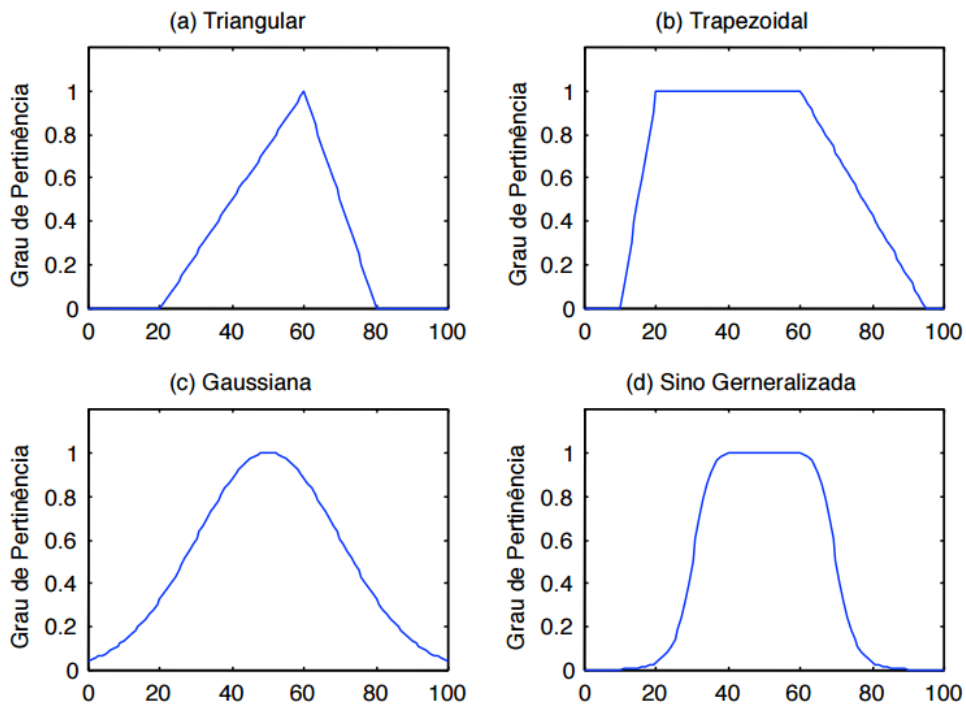


Figura 9 – Formas de funções de pertinência

Dependendo do conceito que se deseja modelar e do contexto que serão utilizadas, as funções de pertinência podem ter diferentes formas, que podem ser definidas por um especialista no domínio ou representadas por formas padrões, por exemplo: Triangular, Trapezoidal, Gaussiana e Sino Generalizada, conforme representado na Figura 9.

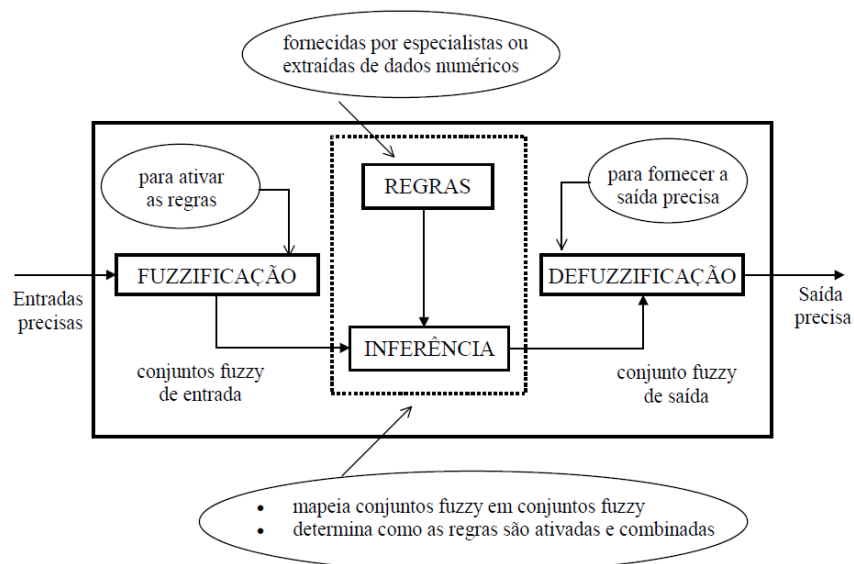


Figura 10 – Representação de um Sistema de Inferência Fuzzy (TANSCHKEIT, 2004)

A Figura 10 mostra uma representação geral de um Sistema de Inferência *Fuzzy*. No estágio inicial, existem as entradas precisas, que são resultantes de medições ou observações, comuns na maioria das aplicações práticas; o processo de **fuzzificação**, responsável por

realizar o mapeamento entre as entradas precisas e os conjuntos *Fuzzy* e a ativação das regras relevantes.

As regras podem ser fornecidas por especialistas ou através de métodos de extração de regras de dados numéricos, estes são particularmente úteis em problemas de classificação e previsão de séries temporais. No caso das regras extraídas por especialistas, embora estes sejam conhecedores do problema em questão, extrair regras na forma "se...então" não é uma tarefa fácil.

Por fim, o estágio de **defuzzificação** recebe a saída através do processo de inferência e efetua a interpretação dessa informação. Isto ocorre porque nas aplicações práticas, geralmente se faz necessário ter saídas precisas. Entre os modelos de inferência mais utilizados nesta etapa estão:

- O modelo de Mamdani (MAMDANI; ASSILIAN, 1975): foi um dos primeiros sistemas de controle a serem desenvolvidos tendo por base a teoria de conjuntos *Fuzzy* e a lógica *Fuzzy*. Neste modelo, variáveis linguísticas são utilizadas nos consequentes;
- O modelo de Takagi-Sugeno-Kang (TSK) (TAKAGI; SUGENO, 1985): similar ao modelo de Mamdani. Porém, a principal diferença entre os modelos é com relação aos consequentes das regras *Fuzzy*. No modelo TSK são usadas funções polinomiais nos consequentes, ao invés de variáveis linguísticas.

1.4 Ferramentas e Recursos

Nesta seção são apresentadas as ferramentas e recursos computacionais usados neste trabalho.

1.4.1 NLTK: Natural Language ToolKit

O NLTK, ou *Natural Language ToolKit*, é uma biblioteca de código aberto escrito na linguagem de programação *Python* que permite a aplicação das mais diversas técnicas de Processamento de Linguagem Natural. A biblioteca possui distribuição para as plataformas: *Windows*, *Macintosh* e *Unix*.

Segundo Bird, Klein e Loper (2012), o NLTK foi desenvolvido em 2001 como parte do curso de linguística computacional do Departamento de Computação e Ciência da Informação na Universidade da Pensilvânia (EUA). Após o seu desenvolvimento, vários colaboradores têm contribuído para a expansão do projeto, além disso, o NLTK vem sendo adotado em cursos de graduação e pós-graduação em várias universidades como uma ferramenta para aplicação de técnicas de PLN.

O NLTK apresenta uma estrutura completa para criar aplicações de PLN, fornecendo tarefas como etiquetagem de textos, análise sintática, classificação de textos e outras, que combinadas podem resolver problemas complexos.

Tabela 2 – Módulos e Tarefas do NLTK

Módulo NLTK	Tarefa de PLN
<code>nltk.corpus</code>	Acessar Córpus
<code>nltk.tokenize</code> , <code>nltk.stem</code>	Processamento de strings
<code>nltk.tag</code>	Etiquetagem
<code>nltk.classify</code>	Classificação

A Tabela 2 apresenta os principais módulos e respectivas tarefas implementadas por cada um.

1.4.2 Etiquetador Mac-Morpho

Um etiquetador é responsável pelo processo de classificação das palavras em suas respectivas classes gramaticais, de acordo com as funções sintáticas (BIRD; KLEIN; LOPER, 2012). O NLTK implementa um módulo responsável pela etiquetagem que define as seguintes classes gramaticais em português: substantivos, adjetivos, advérbios, verbos, pronomes, preposições, conjunções, interjeições e numerais.

O NLTK possui etiquetadores para diversos idiomas. Para o Português há o *Mac-Morpho*, proposto por Aluísio, Pelizzoni e Marchi (2003). O *Mac-Morpho* apresenta 1,1 milhão de palavras anotadas manualmente quanto a sua morfossintaxe. O Córpus utilizado por este etiquetador foi coletado do jornal Folha de São Paulo devido ao mesmo possuir textos de alta qualidade e englobar diferentes domínios e autores. As etiquetas usadas pelo *Mac-Morpho* são descritas no Anexo A.

Segue um exemplo de uma entrada e sua respectiva saída para o etiquetador: na frase "O celular é ótimo", após o NLTK fazer a separação das palavras, o *Mac-Morpho* aplica as classes gramaticais resultando em [(‘O’, ‘ART’), (‘celular’, ‘N’), (‘é’, ‘V’), (‘ótimo’, ‘ADJ’)]. As *tags* ou tuplas resultantes podem ser manipuladas facilmente para extrair características e palavras opinativas.

1.4.3 SPSS

O SPSS é um *software* utilizado para realizar análises estatísticas e preditivas (SPSS, 2007). Ele pode ser utilizado em diversas áreas de pesquisa científica, a saber: aplicação analítica, mineração de dados, mineração de textos e estatística que transformam os dados em informações importantes.

A ferramenta permite realizar diversos testes estatísticos, tais como: correlação, multicolinearidade, hipóteses, contagens de frequência, ordenação de dados, reorganização da informação, entre outras. O SPSS também permite a criação de análises preditivas com RNAs MLP e Redes RBF, que foram usadas neste trabalho.

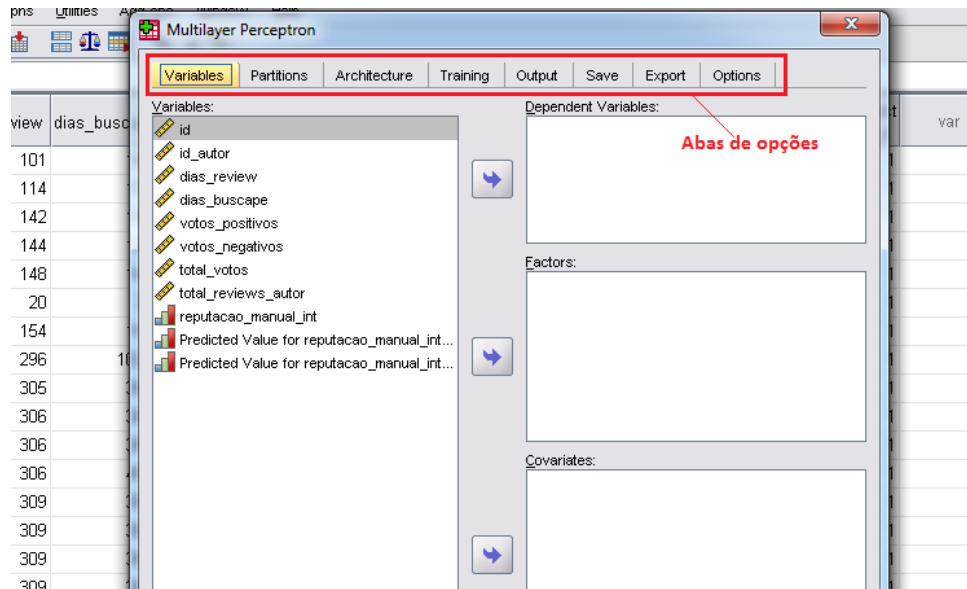


Figura 11 – Tela de Configuração da Rede Neural na SPSS

A Figura 11 mostra a janela de contexto para configuração da Rede Neural *Multi-layer Perceptron* na SPSS. Na área destacada como "Abas de opções", tem-se:

- **Variables:** aba responsável por definir quais serão as variáveis de entrada da Rede Neural e também a variável dependente, ou seja, a variável que será utilizada no treinamento/teste da rede;
- **Partitions:** permite definir as partições de treinamento/teste. Nesta pesquisa, o valor definido foi de 90% para treinamento e 10% para teste;
- **Architecture:** define a arquitetura da rede. É possível definir um intervalo para que o *software* teste automaticamente quantos neurônios deverá conter na camada escondida, bem como a melhor função de ativação. Porém, o usuário tem a liberdade de definir os parâmetros manualmente;
- **Training:** define o tipo de treinamento a ser aplicado, em lotes ou tempo real;
- **Output:** aba responsável por definir quais serão os dados de saída da rede através de diversos relatórios, inclui também a importância das variáveis que define quais são as mais importantes variáveis de entrada;
- **Save:** permite salvar o valor predito da Rede Neural em uma nova variável no *Data View*;

- **Export:** configura a exportação dos resultados em arquivos XML;
- **Options:** define diversos parâmetros opcionais da rede neural, por exemplo, a condição de parada da execução.

1.5 Considerações Finais

Neste Capítulo foram apresentados os principais conceitos sobre Processamento de Linguagem Natural, Análise de Sentimentos, Reputação do autor e as técnicas de Aprendizagem de Máquina, conceitos necessários para o entendimento geral da abordagem proposta.

A área de Processamento de Linguagem Natural é uma área multidisciplinar que envolve, por exemplo, Linguística e Inteligência Artificial, e vem sendo bastante explorada no sentido de extrair conhecimento relevante de grandes volumes de informações, especialmente na era da *Web 2.0*. Seja em *sites* de comércio eletrônico ou Redes Sociais *Online*, as técnicas de PLN podem ser aplicadas com sucesso no desenvolvimento das mais variadas linhas de pesquisa.

A Análise de Sentimentos, ou Mineração de Opiniões, é uma subárea da PLN que permite extrair e definir a polaridade ou orientação semântica de um trecho de texto. Como resultado da sua aplicação pode-se definir, por exemplo, se um comentário é positivo, negativo ou neutro, além disso pode-se definir a polaridade por escalas, como muito positivo ou moderadamente positivo de acordo com a perspectiva do adjetivo utilizado ou de advérbios intensificadores, por exemplo, em uma aplicação "muito bom" poderia apresentar uma polaridade positiva mais elevado que "bom".

Este trabalho faz uso de um *Córpus* extraído do *site* do Buscapé, que é um serviço de *e-commerce* brasileiro que vem apresentando boa visibilidade, e apesar de não ser uma RSO em sua essência, ele carrega muitos aspectos que permitem compará-lo com uma rede social, como avaliações de comentários por outros usuários através de curtidas, sejam positivas ou negativas, além de informações públicas de perfil simplificado. Esta pesquisa sugere que esses serviços de comércio eletrônico possam implementar políticas de integração com RSOs no sentido de melhorar os seus serviços bem como a experiência dos seus usuários, além de auxiliar na implementação de pesquisas científicas mais aprofundadas em seus domínios.

O conceito de reputação do autor também foi explorado neste Capítulo como o principal conceito a ser abordado nesta pesquisa. Apesar de ser um conceito difícil de definir, alguns autores tentam delimitá-lo em termos gerais ou dentro de um domínio específico. Este trabalho discute reputação do autor através de diversas medidas a serem apresentadas no Capítulo 3.

Foram apresentados os fundamentos básicos sobre RNA. A sua aplicação neste trabalho se dá no sentido de inferir a reputação do autor baseado em diversas informações de entrada, além de expor quais destas informações, ou variáveis, possuem mais importância dentro do contexto.

Logo após, os Sistemas de Inferência *Fuzzy* também foram abordados neste Capítulo por terem sido utilizados tanto na abordagem Top(x) original de Sousa (2015), quanto na abordagem adaptada proposta neste trabalho. Estes sistemas se baseiam em regras, nas quais a Teoria de Conjuntos *Fuzzy* e Lógica *Fuzzy* fornecem o ferramental matemático para se lidar com tais regras linguísticas. Eles têm como objetivo central modelar computacionalmente o raciocínio humano, impreciso, ambíguo e vago.

Por fim, as ferramentas que foram utilizadas neste pesquisa foram expostas para auxiliar no entendimento. O NLTK fornece um pacote completo para trabalhar e aplicar as mais diversas técnicas de PLN sem precisar recorrer a outros *softwares*. Dentro do NLTK existe o etiquetador *Mac-Morpho* que foi aplicado neste trabalho com o intuito de atribuir as classes gramaticais a cada palavra coletada do Córpus do Buscapé. O SPSS é uma ferramenta estatística que permite aplicar diversas técnicas de análise de dados, incluindo análises preditivas com RNA, técnica fundamental para a realização da predição da reputação do autor em cada comentário coletado.

2 Trabalhos Relacionados

Neste Capítulo são apresentados os trabalhos relacionados com esta pesquisa. Na literatura especializada, diversos autores têm contribuído para a classificação de usuários em comentários na Web e nas redes sociais, com destaque para o *Twitter* e os ambientes *Wiki*.

O presente Capítulo é separado em seções para melhorar o entendimento, a saber: na seção 2.1 são apresentadas as pesquisas que relacionam medidas de reputação do autor nos ambientes *Wiki*, logo após, a seção 2.2 aborda os trabalhos relacionados na RSO *Twitter*, a seção 2.3 destaca as pesquisas sobre reputação do autor em outros tipos de redes, e por fim, na seção ?? são apresentados trabalhos que aplicam as RNA para calcular reputação do autor.

2.1 Ambientes Wiki

Os ambientes *Wiki* se caracterizam por permitir a colaboração mútua entre os usuários na produção de artigos dos mais variados temas. Nestes ambientes, existe uma liberdade muito grande no sentido de permitir a qualquer usuário, cadastrado ou anônimo, colaborar da forma que achar conveniente. Um problema inerente desta liberdade é a possibilidade de se ter artigos de baixa qualidade, especialmente pela atuação de vândalos¹.

Com relação a qualidade geral dos artigos publicados nos ambientes *Wikis*, estudos mostram que a qualidade dos artigos do *site Wikipedia*² são comparados aos da enciclopédia britânica, apesar do vandalismo (GILES, 2005). Atualmente, este *site* é um dos principais sistemas de *Wiki* da *Internet*.

Deste modo, os trabalhos apresentados nesta seção abordam diversas formas de avaliar a reputação do autor nestes ambientes. Entre as principais formas, destacam-se: histórico das edições, contexto social e mecanismos de recompensa, que serão abordados nas três subseções a seguir.

2.1.1 Histórico das edições

Esta subseção aborda o histórico das edições nos artigos sobre ambientes *Wiki* em diversos trabalhos. Para tal, os autores observaram os históricos das páginas em busca de padrões de *edits* por parte dos usuários deste sistema colaborativo para determinar medidas de reputação.

¹ Usuários que editam os artigos com informações fora do contexto

² <http://wikipedia.org>

Como destaque inicial, [Wöhner, Köhler e Peters \(2011\)](#) realizaram um trabalho no *site* de conteúdo colaborativo *Wikipedia*. Os autores utilizaram o termo "conhecimento das multidões" (do inglês: "*Windsow of the Crowds*") como uma tendência nos sistemas *online* atuais. Destaca-se que *sites* como *Wikipedia* permitem a pesquisa dos mais variados temas criados por usuários colaborando para uma qualidade no mínimo aceitável. O artigo destaca que as contribuições persistentes de usuários na *Wikipedia* duram em média 14 dias sem sofrer modificações. A abordagem alcançou uma precisão de 86% na classificação dos usuários. As classes abordadas são duas:

- **Autores bloqueados:** usuários que tiveram as suas contas desabilitadas pelos administradores por terem cometido vandalismo ou violado o regulamento da *Wikipedia*;
- **Autores regulares:** usuários que colaboram regularmente sem violar as regras da *Wikipedia*.

No trabalho de [Halfaker et al. \(2009\)](#) foram examinados o papel que a qualidade das contribuições da *Wikipedia*, a experiência dos contribuintes e a propriedade do conteúdo foca nas decisões sobre quais contribuições formam parte da Wiki e quais são rejeitadas pela comunidade. Os autores criaram uma medida para automaticamente definir a qualidade de uma contribuição e, curiosamente, eles acharam pouca evidência de que experiência ajuda os contribuintes a evitar rejeições de seus artigos.

[Javanmardi, Lopes e Baldi \(2010\)](#) utilizaram o histórico das páginas para calcular reputação baseado em padrões de *edits*. Na pesquisa, os autores utilizaram todas as páginas do histórico de edições da *Wikipedia* inglesa com o intuito de extrair estatísticas detalhadas sobre o comportamento dos usuários em suas edições. A partir desses dados, os autores definiram três modelos computacionais de reputação de autor em ambientes *Wiki*, que seguem:

1. No primeiro modelo, a reputação é simplesmente uma fração dos *tokens* (i.e., palavras) inseridos;
2. No segundo modelo, o tempo é levado em consideração para penalizar os *tokens* que são deletados em pouco tempo;
3. O terceiro modelo é uma variação do segundo e acrescenta uma pontuação maior caso o usuário tenha uma alta reputação (*admins*) e uma baixa pontuação caso o mesmo seja de baixa reputação (vândalos).

A validação do modelo foi testada comparando os usuários com os vândalos e administradores, que são os dois extremos desse grupo de usuários. Os resultados apresentados

pelos autores demonstram que os três modelos podem ser utilizados para atribuir precisamente valores de reputação para os usuários administradores e vândalos na *Wikipedia*.

Adler e Alfaro (2007) também realizaram um trabalho sobre o padrão de *edits* das páginas da *Wikipedia* para calcularem a reputação do autor. Segundo a pesquisa, os autores ganham reputação quando os seus *edits* são preservados pelos autores subsequentes no mesmo artigo da *Wikipedia* e eles perdem reputação quando os seus *edits* são desfeitos em um período curto de tempo. Para os autores, um sistema de reputação poderia incentivar as contribuições de alta qualidade. Como resultado, eles mostraram que sua abordagem tem um bom valor de predição ao identificar autores de baixa reputação.

Em Adler et al. (2008a), os pesquisadores analisaram a qualidade das contribuições, em adição a quantidade. Como efeito do trabalho realizado, os autores adotaram a "longevidade total de um *edit*" como medida da contribuição de um autor. Utilizando esta abordagem, eles concluíram que algoritmo consegue distinguir pessoas que contribuem pouco de *spammers* ou vândalos, que tem as suas contribuições com valores negativos.

No trabalho de Adler et al. (2008b), os autores desenvolveram um algoritmo que computa os valores quantitativos a respeito da confiança no texto. Os autores utilizaram o histórico de revisão dos artigos da *Wikipedia*, juntamente com a reputação dos autores que colaboraram com a criação do artigo. A confiança de uma palavra no texto é computada com base na reputação do autor original daquela palavra, bem como a reputação de todos os autores que editaram texto nas proximidades da palavra. Como resultado, os autores demonstraram que texto marcado com baixo nível de confiança tem uma alta probabilidade de ser editado no futuro, quando comparado a um texto com alto nível de confiança.

Por fim, Adler, Alfaro e Pye (2010) definiram um sistema de reputação para a *Wikipedia* denominado *Wikitrust*, que computa três características principais em um sistema de reputação:

- **Qualidade do *edit*:** avalia cada mudança introduzida na revisão;
- **Reputação do autor:** os autores que realizam *edits* de boa qualidade ganham reputação;
- **Reputação do conteúdo:** conteúdos que são revisados por autores com uma alta reputação ganham reputação.

2.1.2 Contexto Social

Nesta subseção, são abordados trabalhos relacionados que levam em consideração o contexto social dos usuários envolvidos na edição de um artigo, ou seja, interesses em comum entre eles.

Na pesquisa de [Zhao et al. \(2010\)](#), os autores afirmam que manter a qualidade das contribuições em um sistema colaborativo como a *Wikipedia* é bastante complexo, especialmente pela sua grande expansão. Para solucionar este problema, os pesquisadores desenvolveram um protótipo denominado *SocialWiki*, que considera as contribuições anteriores dos usuários bem como as avaliações que eles recebem de outros usuários. O protótipo considera o "Contexto Social" para a escolha dos usuários que podem editar um artigo (i.e., interesses em comum entre os colaboradores). Como resultado, os autores observaram que delimitando interesses comuns e confiança entre os usuários fez com que fossem selecionados contribuidores de alta qualidade e confiabilidade.

No trabalho realizado por [Anthony, Smith e Williamson \(2009\)](#), os autores observaram que usuários registrados, preocupados com a sua reputação na comunidade, contribuem com conteúdos de alta qualidade. No entanto, os pesquisadores verificaram que vários usuários anônimos fazem muitas contribuições de alta qualidade. Os resultados alcançados sugerem que a contribuição coletiva é de suma importância por permitir que uma massa de contribuidores possam participar.

2.1.3 Mecanismos de Recompensa

Esta subseção aborda mecanismos de recompensa implementados para incentivar as contribuições de qualidade por parte dos colaboradores.

[Hoisl, Aigner e Miksch \(2007\)](#) abordaram mecanismos de recompensa para os autores que mais participam com boas contribuições em sistemas *Wiki*. Como dinheiro não pode ser usado como fator motivacional, os pesquisadores utilizaram outros fatores, por exemplo, aceitação e *status*. Os autores concluíram que a sua abordagem de recompensas baseada em motivação pode produzir artigos de alta qualidade.

2.2 Twitter

Uma contribuição importante que os trabalhos expostos na seção anterior apresentam está na persistência das contribuições (i.e., quanto mais tempo um *edit* persistir, melhor). É importante observar que os ambientes *Wiki* diferem das RSOs, em especial do *Twitter* e dos *sites* de compra/venda de produtos e serviços.

A RSO *Twitter* se tornou uma fonte de coleta de dados e pesquisa bastante requisitada pelos pesquisadores devido a sua grande popularidade e volume elevado de interação entre os usuários em eventos que surgem na rede. Nesta seção, são apresentados os trabalhos relacionados a esta pesquisa que usam o *Twitter* como base.

[Kwak et al. \(2010\)](#) exploraram a rede social *Twitter* para identificar usuários influentes baseado em dados coletados a partir dos "assuntos do momento" (do inglês:

"*trending topics*"). Os autores criaram um *ranking* de usuários de acordo com o número de seguidores. Em seguida, compararam com o algoritmo *PageRank* (PAGE et al., 1998) e notaram que esses dois tipos de ranqueamento são similares. Eles também analisaram os *retweets*, que é o processo de propagar na rede um *tweet* de outro usuário, e criaram um outro *ranking* com essa informação. De acordo com os pesquisadores, após um *tweet* ser retuitado, o mesmo tem um alcance de, no mínimo, 1000 usuários, pois a sua propagação é praticamente instantânea. Os autores concluíram que mais de 85% tópicos classificados em sua abordagem se referem a manchetes de noticiários.

Rabelo, Prudencio e Barros (2012) também exploraram o *Twitter* e definiram uma abordagem para classificar os usuários próximos a um dado usuário "rotulado" através de um grafo e um algoritmo CC (*Collective Classification*). Este algoritmo leva em consideração um subconjunto de nós pré-rotulados dos quais podem derivar a classificação de nós sem rótulos. O algoritmo pode ser definido formalmente como um grafo $G(V,E,C)$, em que:

- $V = \{v_1, \dots, v_n\}$: conjunto de nós representando os usuários;
- E : conjunto de arestas (v_i, v_j) . Uma aresta $(v_i, v_j) \in E$ se os nós v_i e v_j são conectados;
- $C = \{c_1, \dots, c_m\}$: cada nó recebe uma classificação entre m possíveis classes.

Os pesquisadores observaram também o conceito de "*homophily*", ou seja, a tendência de um usuário seguir os seus semelhantes. Eles também analisaram a importância das *hashtags*³ no *Twitter*.

Golbeck et al. (2011) realizaram um experimento no *Twitter* coletando informações dos perfis dos usuários e definiram um método para prever a personalidade dos mesmos. Os autores observaram que os usuários preferem interfaces que sejam mais compatíveis com a sua personalidade, para tal, eles utilizaram um modelo baseado em traços pré-definidos de personalidades concebido por Tupes e Christal (1992). Dessa forma, as empresas podem se beneficiar desse conhecimento e oferecer propagandas direcionadas a essas pessoas baseadas em sua personalidade, tornando a experiência dos usuários ainda mais rica. Os pesquisadores concluíram que compreender o relacionamento entre os usuários e as suas respectivas personalidades abrem um grande espaço para pesquisa.

Weitzel, Oliveira e Quaresma (2014) definiram medidas para calcular a reputação dentro do *Twitter* abordando informações no domínio da medicina. Os autores observaram que a função de retuitar (processo de encaminhar/difundir uma informação dentro de rede) representa bem o nível de interação entre os usuários e criaram medidas utilizando esta função. Uma rede foi criada levando em conta os retuites entre os usuários. Os autores

³ *Hashtag* é uma palavra ou frase após uma cerquilha para identificar mensagens relacionadas a tópicos específicos, por exemplo #jogosolimpicos

concluíram que a maioria dos perfis no *Twitter* são individuais ou de *blogs*, e que a aplicação das suas medidas baseadas em retuites conseguem identificar os nós populares dentro da rede, utilizando como base o domínio da medicina.

No trabalho desenvolvido por [Aggarwal e Kumaraguru \(2014\)](#), os autores identificaram um "mercado negro" que vende/compra contas fraudulentas, curtidas no *Facebook* e até mesmo seguidores no *Twitter* para, artificialmente, melhorarem a reputação social dos usuários. Foram exploradas quatro características com o intuito de detectar usuários suspeitos:

- **Perfil do usuário:** informações do perfil do usuário disponibilizadas pelo *Twitter*;
- **Rede:** descreve o relacionamento do usuário com seus amigos e seguidores;
- **Conteúdo:** natureza dos *tweets* postados pelo usuário;
- **Comportamento do usuário:** define o padrão de postagens, bem como a dinâmica de "seguir" do usuário.

Os autores desenvolveram um mecanismo de aprendizagem de máquina supervisionado utilizando as características supracitadas capaz de prever seguidores suspeitos com uma precisão de 88,2%.

[Wang \(2010\)](#) definiu reputação como uma das características que identifica *spammers* no *Twitter*. Segundo os autores, reputação é a medida da importância relativa de um usuário na rede. Os pesquisadores definiram a reputação como sendo uma relação entre o número de amigos e o número de seguidores. Os resultados obtidos demonstraram que a proposta da medida de reputação consegue detectar comportamentos anormais de usuários (*spammers*).

Na pesquisa realizada por [Chorley et al. \(2012\)](#), os autores realizaram um experimento utilizando metadados do *Twitter* para avaliar a reputação do autor. No estudo, os participantes do experimento deveriam escolher entre dois *tweets* distintos, baseado em informações sobre estes. Os pesquisadores observaram que os participantes preferiam ler *tweets* que possuíam muitos retuites em contrapartida a *tweets* de usuários com muitos seguidores, por exemplo.

[Weng et al. \(2010\)](#) focam no problema de identificar usuários influentes no *Twitter*. Para tal, os pesquisadores observaram os seguidores dos usuários, bem como quem estes seguem. A abordagem proposta pelos pesquisadores, denominada *TwitterRank*, leva em consideração o relacionamento de amizade entre os usuários, por exemplo: dado três usuários, A, B e C, sendo que C segue A e B, que publicam 500 e 1000 *tweets* respectivamente. Neste caso, a influência que B exerce sobre C é duas vezes maior do que A. Eles concluíram

que algoritmo tem uma performance melhor em ranquear os usuários do que o algoritmo do próprio *Twitter*.

Na pesquisa realizada por [Welch et al. \(2011\)](#), os autores também discutem o padrão de "seguir" no *Twitter*. Eles investigaram o relacionamento entre o ato de seguir (relacionamento explícito) e o de retuitar (relacionamento implícito). Os pesquisadores concluíram que retuitar um usuário tem um impacto mais significativo como indicador de interesse do que simplesmente seguir.

Em [Cappelletti e Sastry \(2012\)](#), os autores desenvolveram um algoritmo de ranqueamento, chamado de *IARank*. O algoritmo desenvolvido observa o potencial que um usuário possui de amplificar uma informação dentro do *Twitter*. Eles consideram dois fatores de entrada no algoritmo: a tendência de um usuário ser retuitado ou mencionado e o tamanho da audiência destes retuites ou menções. Os autores concluíram que a abordagem tem um melhor desempenho se aplicado em um Top(5) usuários influentes, quando comparado ao algoritmo de *PageRank* do *Twitter*, porém, para mais de 5 usuários, o *PageRank* é mais preciso.

2.3 Outros Sistemas Web

Apesar dos trabalhos relacionados estarem mais focados nos ambientes *Wikis* e na RSO *Twitter*, outros sistemas web também abordam reputação do autor, por exemplo: fóruns de discussão, sistemas de *blogs*, sistemas de perguntas e respostas, entre outros. Os trabalhos que são considerados de importância para esta pesquisa são abordados a seguir.

No trabalho de [Akay, Dragomir e Erlandsson \(2015\)](#), os autores pesquisaram uma forma de extrair conhecimento de forma inteligente de fóruns *online* sobre câncer no pulmão. O objetivo deles era melhorar o tratamento da doença assim como diminuir os custos em planos de saúde. Segundo o artigo, os usuários mais influentes nesses ambientes são os usuários que intermediam mais transferência de informação dentro da rede, conhecidos como "*information broker*" ou "intermediário de informação".

[Choi e Han \(2013\)](#) afirmam que as pessoas tendem a confiar mais em *sites*/usuários conhecidos (*trustworthy*). Na abordagem proposta, eles desenvolveram um algoritmo para identificar revisores representativos em mídias sociais *online*. Tais revisores possuem alta representatividade devido suas postagens serem avaliadas por muitos usuários. De acordo com os autores, alguns *bloggers*, por serem muito influentes, podem ser utilizados como ferramentas de *marketing* para empresas e mesmo para propaganda política. Para validar o trabalho desenvolvido, os pesquisadores aplicaram o algoritmo em dois *Córpus*, um de músicas e outro de filmes. O algoritmo proposto pelos autores foi capaz de selecionar os 10% ou 50% usuários mais representativos.

Em [Lv et al. \(2013\)](#), os autores introduziram um novo conceito de redes de opiniões e propuseram um algoritmo baseado em *PageRank* chamado *OHRank*. No estudo, os pesquisadores observaram o nível de atividade ou participação dos usuários no ambiente *online*, bem como o tamanho das suas contribuições. Em um dos experimentos, os autores observaram que a força da influência de um determinado usuário em uma rede é diretamente proporcional ao tamanho dessa rede de usuários. Os resultados sugerem que informações intrínsecas do autor (mentalidade pessoal e influência pessoal) e o número de suas postagens têm forte impacto na mineração de opinião.

No trabalho realizado por [Zhou et al. \(2012\)](#), os pesquisadores utilizaram um sistema *online* de perguntas e respostas para identificar os usuários especialistas em determinados assuntos. Os autores propõem uma abordagem baseada em um modelo probabilístico sensível ao tópico, que é uma extensão da abordagem *PageRank*. Por fim, os autores concluíram que essa abordagem é mais eficiente para encontrar os *experts* pois leva em consideração a estrutura de ligação entre os usuários e a similaridade entre os que perguntam e os que respondem.

[Han, Kim e Cha \(2012\)](#) desenvolveram um algoritmo para demonstrar a reputação do autor no *site* de compartilhamento de vídeos *YouTube*⁴. Inicialmente, os pesquisadores criaram uma rede social com os usuários deste *site* a partir dos conteúdos postados, inscrições e favoritos. Com a aplicação desse algoritmo em um experimento, os autores concluíram que a sua técnica é mais eficiente na tarefa de ranquear os usuários em função da sua reputação do que os algoritmos padrões, como o BM25 ([ROBERTSON; ZARAGOZA, 2009](#)). Eles observaram ainda que a reputação do autor está mais relacionada a inscrições nos canais e número de vídeos postados do que qualquer outro fator.

Em [Agarwal et al. \(2008\)](#), os pesquisadores estudaram a respeito da influência de usuários que escrevem em diversos *blogs* na *Internet*. Os autores perceberam que as contribuições de um blogueiro influente podem impactar em seus colegas blogueiros de várias maneiras. Os autores concluíram que o modelo preliminar desenvolvido tem potencial em identificar os colaboradores mais influentes em sistemas de *blogs*.

No trabalho de [Dellarocas \(2000\)](#), o autor discute a respeito dos sistemas *online* de reputação entre vendedores e compradores em *e-commerce*. Segundo o pesquisador, o objetivo desses sistemas é encorajar confiança nas transações baseado em comportamentos passados. Esse trabalho contribui para evitar que os usuários avaliem de forma injusta outros usuários. Como resultados, o autor observou que ao aplicar a combinação filtragem de *clusters* e anonimato controlado, tem-se uma ferramenta poderosa para "imunizar" os atuais sistemas *online* de reputação no que concerne a avaliações injustas.

[McNally, OMahony e Smyth \(2011\)](#) avaliaram modelos de reputação no contexto

⁴ <http://www.youtube.com>

da *HeyStacks*⁵, empresa que personaliza resultados de busca na Web. Para os autores, a reputação é um fator importante a ser considerado ao retornar os resultados em uma busca na Web. Os pesquisadores concluíram que a aplicação do modelo *HITS* (KLEINBERG, 1999) melhora os resultados das buscas em até 45%.

Em Chen et al. (2011), os autores apontam que o sucesso das aplicações em mídias sociais, tais como *blogs* de notícias, fóruns *online* e sistemas de perguntas e respostas, dependem da qualidade das contribuições de usuários regulares. Segundo os pesquisadores, as melhores contribuições são produzidas por um número pequeno de usuários "reputáveis" e consumidas por uma grande parcela da população. Os autores desenvolveram um novo modelo para extrair uma medida de reputação que obteve desempenho melhor do que outros baseados no *Yahoo! News*⁶ e *Yahoo! Buzz* por exemplo.

Na pesquisa de Movshovitz-Attias et al. (2013), os autores abordaram a reputação do autor no *site* de perguntas e respostas *StackOverflow*⁷. Neste *site*, os usuários ganham reputação ao fazer perguntas bem como ao responder. Os pesquisadores analisaram uma base de dados coletada do *site* e observaram que a maioria das perguntas são realizadas por usuários de baixa reputação, no entanto, proporcionalmente, os usuários de alta reputação fazem mais perguntas que os de baixa reputação.

2.4 Considerações Finais

Este Capítulo abordou as pesquisas mais relevantes que se relacionam com o tema deste trabalho, o estado da arte sobre o tema. É importante observar que muitos tipos de redes e sistemas *online* se preocupam em definir reputação do autor, seja as RSOs, *blogs*, *Wikis*, fóruns *online* ou *sites* de perguntas e respostas.

Com relação aos ambientes *Wiki*, a maioria das pesquisas se preocuparam em abordar as edições dos conteúdos postados pelos usuários dessa rede. Quanto mais tempo um *edit* persistir dentro de um artigo, melhor a reputação do autor do mesmo. Aplicando isso a esta pesquisa, pode-se considerar, por exemplo, que um comentário mais antigo poderia ter mais tempo de receber votos de outros usuários, ao contrário de comentários recentes.

Os trabalhos relacionados que focaram em outros sistemas web, tais como: *blogs*, fóruns *online*, sistemas de perguntas e respostas, abordaram diferentes algoritmos de ranqueamento para definir os autores com melhor reputação. O uso do algoritmo *PageRank* é bastante citado nas pesquisas e, aparentemente, tem um bom desempenho ao selecionar as melhores páginas, conseqüentemente os melhores autores. Este algoritmo ordena as

⁵ <https://www.heystack.com>

⁶ news.yahoo.com

⁷ stackoverflow.com

páginas Web resultantes de uma pesquisa baseado na relevância das páginas. Tal relevância é calculada contabilizando a quantidade e qualidade dos *links* que apontam para uma página.

Outro ponto presente em alguns desses trabalhos é a criação de grafos representando os relacionamentos entre os usuários, por exemplo, no trabalho de [Rabelo, Prudencio e Barros \(2012\)](#) esta técnica é aplicada para classificação dos usuários.

No próximo Capítulo é apresentado um conjunto de seis medidas, bem como a abordagem original e a nova abordagem utilizando Redes Neurais para inferir a reputação do autor em *sites* de *e-commerce*.

3 Abordagem proposta

Este Capítulo tem como objetivo descrever o processo de coleta do Córpus de pesquisa, a abordagem Top(x) original e a abordagem proposta que consiste em uma adaptação da abordagem Top(x), explorando novas medidas para definir a reputação do autor.

3.1 Coleta e Preparação do Córpus

Para o entendimento geral da proposta e para a avaliação da abordagem, se fez necessária a construção de um Córpus de comentários. O processo de coleta dos dados foi realizado nos dias 28 e 29 de setembro de 2016 no *site* do Buscapé. Foram coletadas informações referentes a diversos *smartphones* comercializados em que os usuários realizavam comentários escritos acerca dos mesmos, podendo indicar também a quantidade de estrelas do produto. Inicialmente foram coletados 2433 comentários no total, porém foi realizado um ajuste nos mesmos eliminando os duplicados e vazios para evitar inconsistências. Ao final restaram 2000 *reviews*, sendo 1000 de orientação positiva e 1000 negativa. A orientação semântica dos comentários é definida pelo próprio autor e são apresentados aos leitores de forma separada, através de guias com comentários positivos e negativos. Destaca-se que os trabalhos na área normalmente tem utilizado apenas 2000 comentários nos experimentos e avaliações, devido a dificuldade do processo de anotação do Córpus.

No entanto, ao serem analisadas com mais atenção, as orientações dos comentários nem sempre apresentavam a realidade do comentário. Assim, alguns comentários marcados como positivos poderiam ter na verdade orientação negativa ou vice-versa. Além disso, percebeu-se que muitos comentários não apresentavam orientação positiva e nem negativa, sendo neutros com relação ao produto ao qual se destinavam.

Para solucionar esse problema, decidiu-se por criar um Córpus anotado com a intenção de avaliar as abordagens. Neste processo, selecionou-se quatro especialistas da área de Análise de Sentimentos para avaliarem os comentários, sendo 1000 comentários para cada par. Eles avaliaram os comentários de acordo com as seguintes orientações: positivo, negativo, neutro ou lixo, que são comentários sem conteúdo. Ao final desta etapa, um quinto especialista ficou responsável por resolver as discrepâncias entre os quatro primeiros. Todos os especialistas são alunos graduados em Ciência da Computação e atuam como pesquisadores na área de PLN.

A Tabela 3 apresenta o Córpus anotado após o trabalho dos especialistas. Nota-se que muitos comentários foram marcados como lixo. Isto se deve ao fato de vários usuários

Tabela 3 – Córpus anotado quanto a orientação

Orientação	Total
lixo	334
negativo	602
neutro	141
positivo	923

postarem as suas colaborações nos *sites* sem possuir o produto ou mesmo postando algo sem o menor sentido.

3.1.1 Subcórpus I

Em um primeiro momento, foi realizada a avaliação por seres humanos para definir a reputação do autor dos comentários. Dos 2000 comentários do Córpus, separou-se uma amostra para análise e avaliação com 323 comentários, considerando o nível de confiança de 95% e margem de erro de 5%. Adicionou-se 33 comentários, que corresponde a 10% da amostra, totalizando 356, sendo 132 positivos, 131 negativos e 93 neutros, identificada como **Subcórpus I**. Um especialista observou informações referentes ao autor, como: a quantidade de votos positivos em seus comentários, quantidade de votos negativos, total de votos, entre outras medidas, e aplicou uma nota de 0 a 10 para cada um dos autores dos comentários dentro da amostra definida, sendo guiado unicamente pelas variáveis de entrada propostas (Mais detalhes na Seção 3.3). É importante destacar que um mesmo autor pode ter escrito mais de um comentário no Córpus coletado, porém a reputação foi calculada individualmente para o autor do comentário.

Tabela 4 – Resultado da avaliação realizada por seres humanos quanto a Reputação do Autor

Reputação	Quantidade
0	68
1	163
2	50
3	23
4	11
5	16
6	9
7	5
8	1
9	5
10	5

A Tabela 4 mostra o resultado da avaliação realizada por seres humanos quanto a Reputação do Autor dos comentários. As 11 notas aplicadas aos autores dentro da amostra foram generalizadas para o universo completo dos comentários através de uma RNA MLP.

Esta generalização se dá pela rede neural que infere, a partir das medidas de entrada, qual a reputação do autor para qualquer comentário dentro do Córpus.

3.1.2 Subcórpus II

Em um segundo momento, foram selecionados aleatoriamente 271 comentários, sendo 100 positivos, 100 negativos e 71 neutros, para serem avaliados de acordo com a sua importância, identificados como **Subcórpus II**. Foi necessário tal procedimento para poder avaliar a precisão da abordagem original definida por Sousa (2015) e também da nova abordagem que está sendo apresentada neste trabalho. Esse número foi definido tendo em vista a distribuição estatística no Córpus principal em termos das orientações positivas, negativas e neutras, excluindo lixo.

Tabela 5 – Avaliação realizada por seres humanos da importância dos comentários

Importância	Quantidade
excelente	17
bom	24
suficiente	145
insuficiente	85

A Tabela 5 apresenta o resultado da avaliação realizada por um especialista quanto a importância dos comentários. Para realizar o procedimento de avaliação, o especialista levou em consideração três fatores: a quantidade de informações destacadas no comentário, a riqueza do vocabulário e a reputação do autor, esta última apresentada como uma nota de 0 a 10 para o especialista em cada comentário, conforme a reputação inferida pela RNA MLP.

3.2 Abordagem Top(x)

A abordagem Top(x) proposta por Sousa (2015) foi definida para inferir a importância das opiniões de comentários a respeito de produtos e serviços em *sites* de comércio eletrônico. Ela utiliza três variáveis para calcular a importância dos comentários, a saber: reputação do autor, quantidade de padrões linguísticos e o percentual de palavras escritas corretamente, ou riqueza de vocabulário. Estas três variáveis são dadas como entrada em um Sistema *Fuzzy* que fornece como saída a importância calculada do comentário, representada por um valor real definido na escala [0..10].

3.2.1 Variáveis de Entrada

- **Reputação do Autor:** Segundo Sousa (2015), as pesquisas sobre a reputação do autor foram implementadas para resolver os problemas dos *spams* na *Internet*. Ainda

segundo o autor, as opiniões de especialistas sobre questões pertencentes a suas áreas de pesquisa têm um peso maior em comparação a pessoas que não têm o mesmo grau de especialização. A hipótese defendida por [Sousa \(2015\)](#) é que autores que realizam muitos comentários tem melhor reputação do que autores ocasionais;

- **Quantidade de Tuplas:** Segundo [Sousa \(2015\)](#), a identificação de tuplas <característica, palavra opinativa> torna-se relevante para o resultado final do processo de análise de sentimento. A hipótese defendida é que quanto mais características e qualidades forem citadas pelos autores, mais importante é o comentário. O processo de filtragem para identificar tuplas relevantes leva em consideração um trabalho manual que gerou uma base de 40 características relacionadas com o contexto do trabalho, assim as tuplas que possuem característica pertencente à **base de características** são consideradas como relevantes, todas as outras são excluídas. O Anexo B apresenta as características utilizadas na abordagem Top(x);
- **Riqueza de Vocabulário:** De acordo com [Sousa \(2015\)](#), um texto compreensível e bem escrito não pode conter erros ortográficos e gramaticais. Apesar da informalidade dos comentários escritos na web e em redes sociais, palavras escritas incorretamente podem ser consideradas ruídos, em se tratando de análise de sentimentos de comentários. Vários problemas podem ser encontrados em textos escritos em mídias sociais, como: abreviações, substituições fonéticas e *emoticons* ([PALTOGLOU; GIACHANOU, 2014](#)). O autor defende que quanto mais correto um comentário for escrito, mais útil será a opinião expressa.

3.2.2 Sistema Fuzzy

O Sistema de Inferência *Fuzzy* utilizado na abordagem Top(x) para inferir a importância dos comentários foi do tipo *Mamdani* ([MAMDANI; ASSILIAN, 1975](#)). No modelo de *Mamdani*, a definição da saída do sistema é apresentada por meio de variáveis linguísticas, tornando a modelagem mais intuitiva, pois é realizada em um nível de abstração mais próximo do ser humano.

Para as variáveis de entrada, foram utilizados três valores linguísticos, a saber: **Baixo** (B), **Médio** (M) e **Alto** (A). Foram definidas funções de pertinência de forma empírica para cada variável de entrada. As Figuras 12, 13 e 14 apresentam graficamente as funções de pertinência das variáveis reputação do autor, quantidade de tuplas e riqueza de vocabulário, respectivamente. É importante relatar que os domínios das funções foram normalizados de 0 a 100, tomando como base o maior valor em cada uma das variáveis entrada no Sistema *Fuzzy*. A Figura 15 apresenta o gráfico da função de pertinência da importância dos comentários utilizando quatro valores linguísticos: Insuficiente (IF), Suficiente (SF), Bom (BM) e Excelente (EX).

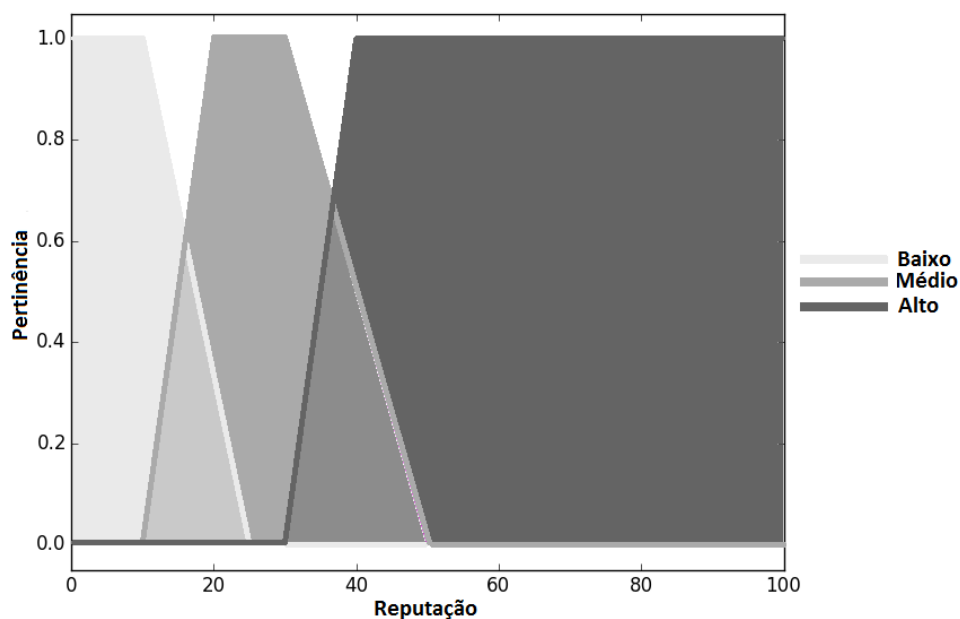


Figura 12 – Gráfico da função de pertinência da variável Reputação do Autor

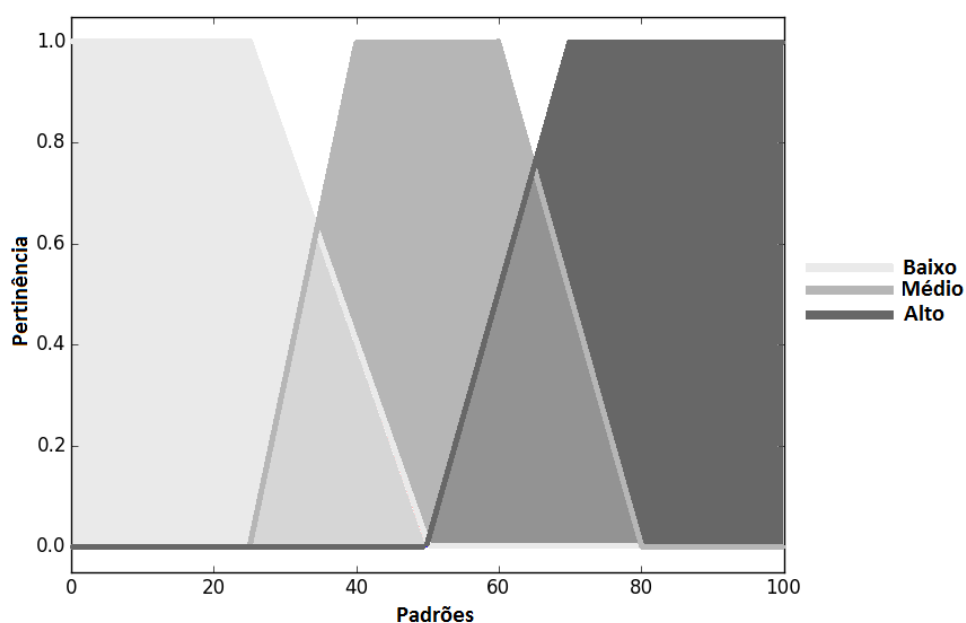


Figura 13 – Gráfico da função de pertinência da variável Quantidade de Tuplas

Tabela 6 – Base de regras

Autor	Padrões/Riqueza								
	B/B	B/M	B/A	M/B	M/M	M/A	A/B	A/M	A/A
B	IF	IF	SF	SF	SF	SF	SF	BM	BM
M	IF	IF	SF	SF	SF	BM	BM	EX	EX
A	SF	SF	BM	BM	BM	BM	BM	EX	EX

A Tabela 6 apresenta a base de regras definida por um especialista que determina a estratégia de tomada de decisão para uma determinada aplicação. O ajuste realizado nestas regras foi feito tendo como base o trabalho de Sousa (2015). Porém, é importante

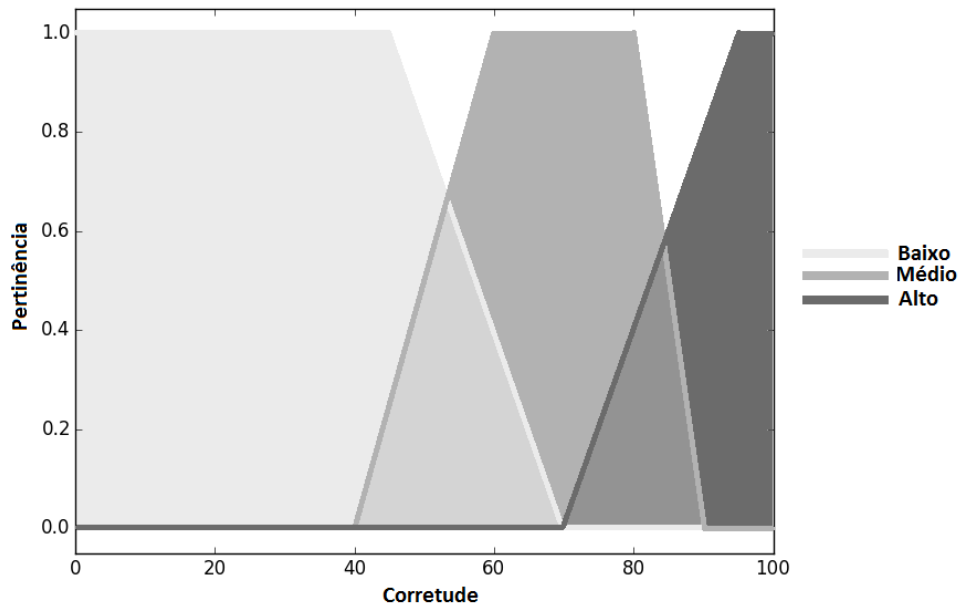


Figura 14 – Gráfico da função de pertinência da variável Corretude

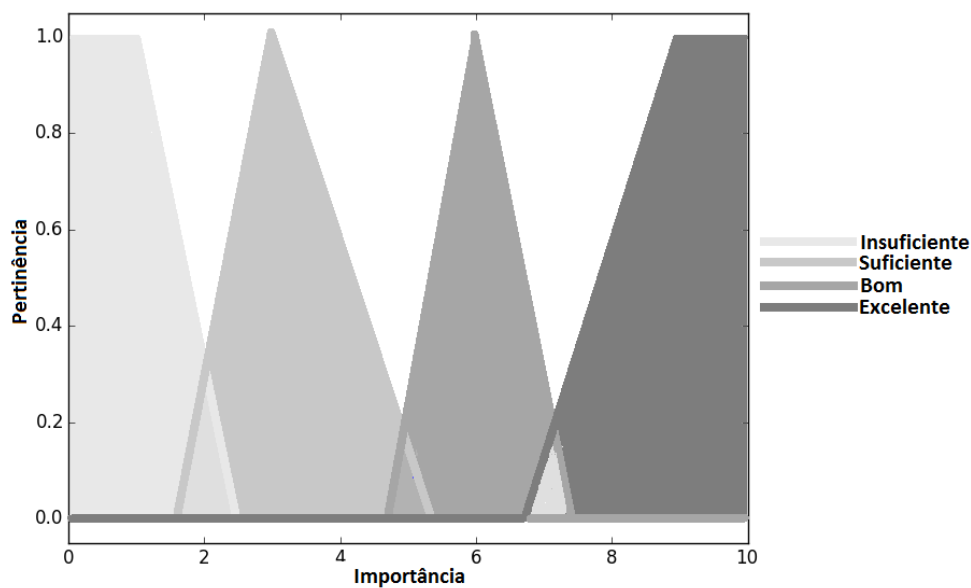


Figura 15 – Gráfico da função de pertinência da saída Importância

destacar que algumas regras de produção foram alteradas em relação a abordagem original com o intuito de melhorar a consistência dos resultados obtidos, após análises críticas feitas por especialistas em Sistemas *Fuzzy*. No entanto, um procedimento de ajustes e alterações na base de regras pode ser realizado para obter melhores resultados.

3.3 Abordagem proposta

A abordagem proposta neste trabalho visa analisar um conjunto de medidas para definir a reputação do autor de comentários em *sites* de vendas de produtos e serviços. Como descrito anteriormente, a proposta representa uma adaptação da abordagem apresentada

por Sousa (2015). Destaca-se que uma RNA MLP foi aplicada com o objetivo de inferir a reputação dos autores dos comentários do Córpus, bem como fornecer a importância de cada medida de entrada. A Figura 16 apresenta a abordagem proposta nesta dissertação. Na figura, a variável 'a' representa a saída da RNA e indica a reputação do autor normalizada para o intervalo de 0 a 10. Já a variável 'k' indica o grau de importância do comentário, como na abordagem original.

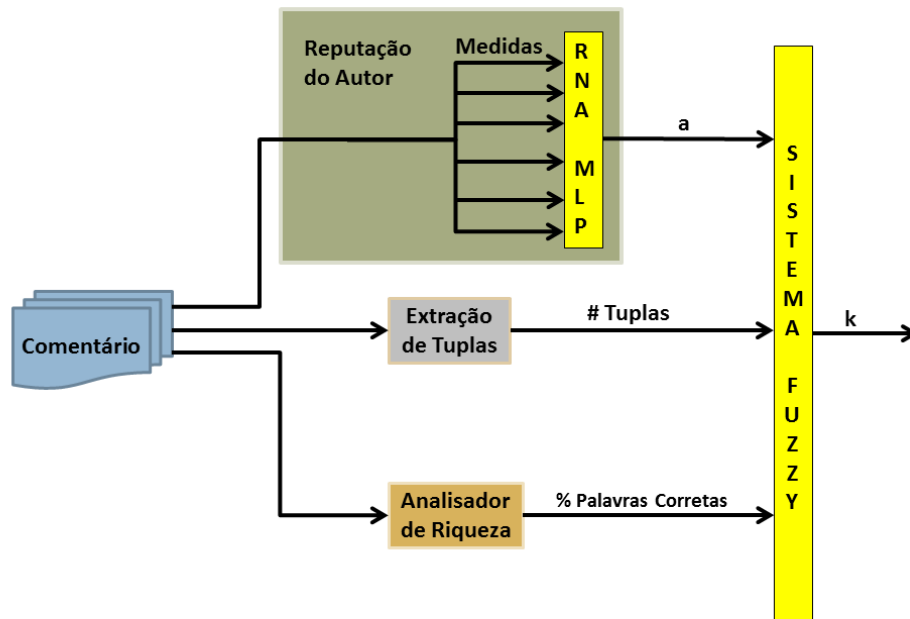


Figura 16 – Abordagem proposta

Definiu-se seis medidas para avaliar a reputação do autor dos comentários de produtos no Córpus coletado. Tais medidas foram extraídas levando em conta o máximo de informações que puderam ser aproveitadas para avaliar a reputação do autor. Seguem as medidas definidas abaixo:

- **Data do *review*:** a data de escrita do *review*, convertida para dias em comparação com a data inicial de coleta do Córpus. Esta informação é importante pois quanto mais recente um *review*, mais atualizado o comentário e, conseqüentemente, deve ser melhor avaliado. No entanto, os *reviews* que são muito recentes podem ser prejudicados no processo de avaliação geral por não ter tempo hábil para leitura pelos consumidores;
- **Data de cadastro no *site*:** a data em que o autor fez o seu cadastro no *site*, convertida para dias em comparação com a data inicial da coleta do Córpus. Esta informação é importante pois considera-se a reputação dos autores experientes melhor do que de autores novos;

- **Votos positivos:** quantidade de votos positivos atribuídos por outros usuários. Quanto mais votos positivos um autor receber de outros usuários em seus comentários, melhor será a sua reputação;
- **Votos negativos:** quantidade de votos negativos atribuídos por outros usuários. A importância dos votos negativos é inversamente proporcional aos votos positivos, pois quanto mais votos negativos o autor receber em seus comentários, pior a sua reputação;
- **Total de votos:** soma dos votos recebidos para o comentário. De forma geral, hipotetiza-se que quanto mais votos o usuário tenha em seus comentários, sejam positivos ou negativos, melhor a sua reputação pois o mesmo está sendo observado;
- **Total de *reviews* do autor do comentário:** quantidade de comentários que o autor realizou no Buscapé. Esta informação é relevante pois indica a participação ativa do usuário dentro do sistema. Destaca-se que [Sousa \(2015\)](#) aplicou essa informação como única medida para avaliar a reputação do autor.

3.4 Considerações Finais

Inicialmente, este Capítulo apresentou os detalhes do processo de coleta e preparação do *Córpus* do *site* do Buscapé. É importante informar a necessidade de especialistas para avaliarem manualmente os comentários, seja no sentido de definir manualmente a polaridade, a reputação do autor ou definir a importância do comentário.

Em seguida, descreveu-se, de forma breve, a abordagem original de [Sousa \(2015\)](#), que é base para esta pesquisa. Essa abordagem utiliza três variáveis de entrada em um Sistema de Inferência *Fuzzy* para inferir a importância dos comentários. Entre as variáveis de entrada está a reputação do autor, que na abordagem original utiliza a quantidade de comentários como valor único de entrada, porém, isso deixa essa medida bastante fraca por ignorar outras informações não-textuais que poderiam ser relevantes para definir a reputação, além de caracterizar *spammers* como bons autores.

Por fim, foi apresentada a abordagem proposta nesta pesquisa que faz uma adaptação na medida de reputação do autor para torná-la mais consistente. Para isto, foram utilizadas seis medidas como entrada em uma Rede Neural Artificial *Multilayer Perceptron* e a saída da RNA representa o valor de entrada da reputação do autor na abordagem original. Os experimentos realizados e os resultados obtidos são apresentados e discutidos no próximo Capítulo.

4 Experimentos e Resultados

Neste Capítulo são apresentados os resultados dos experimentos realizados, bem como a discussão acerca destes resultados.

4.1 Metodologia

Os experimentos detalhados neste Capítulo são divididos em duas etapas, sendo a primeira com a aplicação de uma RNA MLP através da ferramenta SPSS sobre o **Subcórpus I**. Na segunda etapa, o experimento utilizou a saída calculada pela RNA MLP da primeira etapa como entrada para variável reputação do autor em um Sistema de Inferência *Fuzzy* baseado no trabalho de Sousa (2015). Este Sistema *Fuzzy* utilizou a ferramenta NLTK com o objetivo de extrair as tuplas e calcular a riqueza do vocabulário dos comentários coletados no Córpus, tal experimento foi realizado sobre o **Subcórpus II**.

4.2 RNA MLP

Com o objetivo de realizar uma comparação entre a abordagem Top(x) de Sousa (2015) e a nova abordagem proposta neste trabalho, foi executado um experimento aplicando uma RNA MLP para definir a reputação do autor em vez da contagem direta dos comentários. Após este primeiro experimento, um segundo foi realizado calculando a importância dos comentários na abordagem Top(x) original e na nova Top(x) com a RNA MLP. Nesta seção os experimentos são detalhados.

Tabela 7 – Informação da Rede Neural

Camada de Entrada	
Número de neurônios	6
Método de escalabilidade das variáveis	Padronizado
Camada Escondida	
Número de neurônios	8
Função de ativação	Tangente Hiperbólica
Camada de Saída	
Variável dependente	reputacao_manual_int
Número de neurônios	1
Função de ativação	Softmax

A Tabela 7 mostra informações referentes a topologia utilizada na Rede Neural proposta no primeiro experimento. A ferramenta utilizada para concepção e exe-

cução da rede neural foi o *software* de análises estatísticas SPSS. Na camada de entrada tem-se 6 neurônios, em que cada um representa uma variável de entrada, a saber: **dias_review**, **dias_buscape**, **votos_positivos**, **votos_negativos**, **total_votos** e **total_reviews_autor**. Tais medidas foram consideradas importantes para avaliar a reputação do autor dentro do contexto do Córpus coletado. É importante ressaltar que em RSOs, outras medidas podem ser utilizadas para calcular a reputação do autor, por exemplo, número de seguidores e número de retuítes no *Twitter* (KWAK et al., 2010; WENG et al., 2010; AGGARWAL; KUMARAGURU, 2014), no entanto para *sites* de compra e venda elas ainda não estão disponíveis, tornando um grande desafio a tarefa de integrar os perfis de usuários dos *sites* de *e-commerce* e RSOs. Na camada escondida, o melhor ajuste se deu com 8 neurônios e a função de ativação Tangente Hiperbólica. Por fim, a camada de saída utiliza o atributo de supervisão "reputacao_manual_int" como variável dependente para testar a rede, classificando as 11 notas possíveis dos autores (0 a 10) e a função de ativação *Softmax*. É importante relatar que este ajuste é realizado pela ferramenta SPSS, que realiza diversos testes internos para atingir a melhor configuração possível para a performance geral da RNA.

Com relação à função de ativação tangente hiperbólica, a mesma assumirá sempre valores reais entre -1 e 1, tendo sua expressão matemática definida pela equação

$$g(\mu) = \frac{1 - e^{-\beta\mu}}{1 + e^{-\beta\mu}} \quad (4.1)$$

onde β é uma constante real associada ao nível de inclinação da função (parâmetro de inclinação). Já a função *softmax* (BRIDLE, 1990) normaliza entre 0 e 1 as saídas, com o objetivo de definir a probabilidade de uma classe dentro de um problema multi-classes, cuja sua expressão matemática é dada pela equação

$$y_i = \frac{e^{x_i}}{\sum_{j=1}^c e^{x_j}} \quad (4.2)$$

onde c é o número de camadas escondidas da RNA e x_i é o valor resultante do i -ésimo neurônio da camada escondida. O resultado da equação é interpretado como probabilidades, sendo úteis na classificação pois dá uma medida de certeza sobre as classificações do problema.

A Figura 17 mostra a disposição dos neurônios em cada camada na melhor topologia escolhida para a execução da rede neural. O processo de treinamento e teste foi aplicado sobre o **Subcórpus I** utilizando o método de validação cruzada denominado *k-fold*, que consiste em criar k subconjuntos mutuamente exclusivos do mesmo tamanho e, a partir disto, um subconjunto é utilizado para teste e os $k-1$ restantes são utilizados para estimação dos parâmetros. Neste experimento, o valor do k é igual a 10.

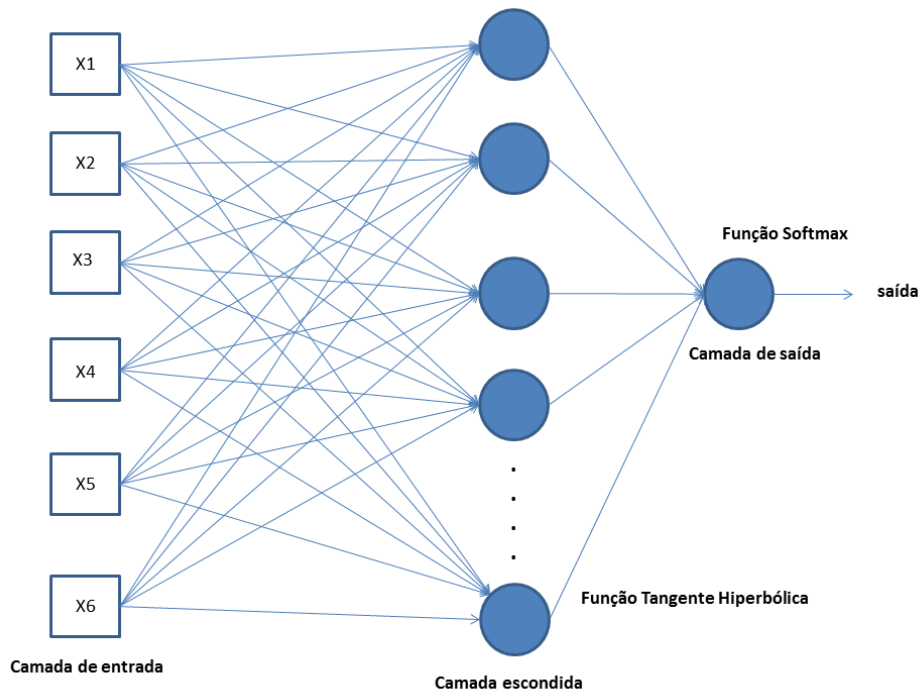


Figura 17 – Topologia da RNA

Tabela 8 – Importância das Variáveis de Entrada

Variável de entrada	Importância	Importância Normalizada
dias_review	0,106	42,10%
dias_buscape	0,086	34,40%
votos_positivos	0,252	100,00%
votos_negativos	0,205	81,70%
total_votos	0,245	97,60%
total_reviews_autor	0,105	41,80%

A importância de uma variável independente é uma medida de quanto o valor predito pelo modelo da rede muda para diferentes valores da variável independente. O valor da importância de uma variável de entrada é calculado levando em consideração o peso das conexões dos neurônios entre as camadas da RNA. A importância normalizada é simplesmente os valores de importância divididos pelos maiores valores de importância e expressos em percentagens. A Tabela 8 apresenta a importância de cada variável de entrada na rede neural proposta. É possível observar que a variável mais importante para avaliar a reputação do autor foi "votos_positivos", seguida de "total_votos" e "votos_negativos". Tal resultado se mostra interessante porque a variável "votos_positivos" é definida por outros usuários da rede, confirmando a boa reputação do autor daquele comentário que recebeu o voto positivo. No entanto, um dos problemas que podem surgir priorizando esta variável é que os *reviews* precisarão de muitos votos positivos para serem lidos deixando de lado os *reviews* mais recentes que não serão lidos por terem poucos ou nenhum voto positivo, mesmo que possuam relevância (LI et al., 2013). Por outro lado, a variável menos

importante foi "dias_buscape" que indica o tempo, em dias, do cadastro do usuário no *site* do Buscapé.

Após o treinamento da rede neural, a reputação inferida foi usada como entrada na abordagem Top(x), substituindo a informação anterior, quantidade de comentários emitidos pelo autor. Com relação a precisão de inferência da RNA MLP dentro do conjunto de treinamento e teste, atingiu-se um valor de **62,08%** no processo de classificação para os valores numéricos de 0 a 10. No entanto, considerando os termos linguísticos usados na abordagem original e descritos a seguir, a precisão da RNA MLP atinge o valor de **91,01%**.

- **Baixo:** 0 a 3;
- **Médio:** 4 a 7;
- **Alto:** 8 a 10.

4.3 Redes RBF

Visando apresentar uma alternativa para a arquitetura MLP, executou-se um experimento com uma rede neural artificial com funções de bases radial para inferir a reputação do autor no Córpus do Buscapé. Este tipo de rede neural pode ser utilizado em diversos tipos de situações, como: aproximador universal de funções, predição de séries temporais, classificação e sistemas de controle.

Tabela 9 – Informação da Rede RBF

Camada de Entrada	
Número de neurônios	6
Método de escalabilidade das variáveis	Padronizado
Camada Escondida	
Número de neurônios	11
Função de ativação	Softmax
Camada de Saída	
Variável dependente	reputacao_manual_int
Número de neurônios	1
Função de ativação	Identidade

A Tabela 9 apresenta informações referentes a topologia da Rede RBF. Utilizou-se também o *software* SPSS para a execução do experimento. Na camada de entrada tem-se 6 neurônios representando as variáveis de entrada, a saber: **dias_review**, **dias_buscape**, **votos_positivos**, **votos_negativos**, **total_votos** e **total_reviews_autor**. Na camada escondida, o melhor ajuste se deu com 11 neurônios e a função de ativação *Softmax*. Por fim, a camada de saída utiliza o atributo de supervisão "reputacao_manual_int" como

variável dependente para testar a rede, classificando as 11 notas possíveis dos autores (0 a 10) e a função de ativação Identidade. É importante ressaltar que o ajuste é realizado pela ferramenta SPSS, assim como o ajuste realizado com a RNA MLP.

A função Identidade retorna sempre o mesmo valor que foi utilizado como argumento, ou seja, os resultados são idênticos aos valores do limiar de ativação μ , na qual sua expressão matemática definida por

$$g(\mu) = \mu \quad (4.3)$$

e tem como uma das suas aplicabilidades em RNAs aproximadoras de universais de funções, visando mapear o comportamento entre as variáveis de entrada e saída de processos (SILVA; SPATTI; FLAUZINO, 2010).

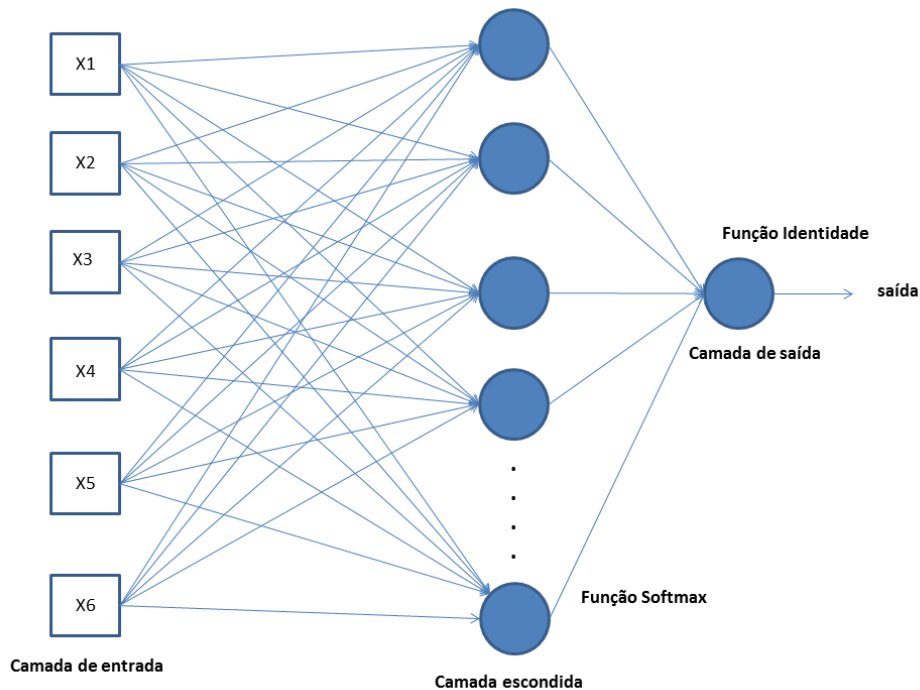


Figura 18 – Topologia da Rede RBF

A Figura 18 mostra a disposição dos neurônios em cada camada na melhor topologia escolhida para a execução da rede neural RBF. Da mesma maneira que a RNA MLP, o processo de treinamento e teste foi aplicado sobre o **Subcórpus I** utilizando o método de validação cruzada *k-fold*, com o *k* igual a 10.

A Tabela 10 apresenta a importância de cada variável de entrada na rede RBF. Observa-se que os resultados são similares aos apresentados pela RNA MLP (Ver Tabela 8), porém existe uma mudança na ordem das variáveis de entrada mais importantes, sendo "total_votos" a variável mais importante, seguida de "votos_positivos" e, por fim,

Tabela 10 – Importância das Variáveis de Entrada na Rede RBF

Variável de entrada	Importância	Importância Normalizada
dias_review	0,127	60,30%
dias_buscape	0,111	52,80%
votos_positivos	0,210	99,60%
votos_negativos	0,210	99,50%
total_votos	0,211	100,00%
total_reviews_autor	0,132	62,60%

"votos_negativos". Por outro lado, a variável menos importante foi "dias_buscape", assim como na RNA MLP.

Com relação a precisão de inferência da Rede RBF dentro do conjunto de treinamento e teste, atingiu-se um valor de **52,25%** no processo de classificação para os valores numéricos de 0 a 10, portanto, inferior ao valor alcançado pela RNA MLP (**62,08%**). Considerando as três categorias (Baixo, Médio e Alto), atingiu-se precisão de **87,36%**, também inferior aos **91,01%** da RNA MLP. Desta forma, devido a vantagem da RNA MLP sobre a Rede RBF, decidiu-se por seguir os experimentos apenas com a primeira arquitetura.

4.4 Comparação da abordagem original com a abordagem proposta

Para avaliação das abordagens, calculou-se as medidas de Precisão (P), Cobertura (R) e Medida-f (F) para cada classe. Essas medidas são normalmente usadas em avaliação de abordagens na área de aprendizagem de máquina ([POWERS, 2011](#)).

É importante destacar que a Precisão (P) é calculada como a porcentagem de exemplos corretamente classificados em cada classe, por meio da seguinte fórmula:

$$P = \frac{TP}{TP + FP} \quad (4.4)$$

onde TP representa os comentários classificados corretamente (do inglês: true positive) e FP representa os comentários classificados incorretamente (do inglês: false positive). Por exemplo, na Tabela 11 a classe EX possui 2 comentários que foram classificados de maneira correta mas 6 de maneira incorreta, sendo a precisão igual a $2/(2+6) = 25\%$.

Já a Cobertura (R) é calculada através da porcentagem de exemplos corretamente classificados em relação ao total de instâncias da classe do Córpus anotado, por meio da fórmula:

$$R = \frac{TP}{TP + FN} \quad (4.5)$$

onde TP representa os comentários classificados corretamente e FN representa os comentários classificados incorretamente em relação ao Córpus (do inglês: false negative). Por exemplo, na Tabela 11 a classe EX possui 2 comentários classificados de maneira correta mas 9 de maneira incorreta em relação ao Córpus, sendo a cobertura igual a $2/(2+9) = 18,18\%$.

Por fim, a Medida-f é uma média harmônica entre Precisão e Cobertura derivada de Rijsbergen (1979) e calculada conforme a equação:

$$F = \frac{2 * P * R}{P + R} \quad (4.6)$$

onde P representa o valor da Precisão e R representa o valor da Cobertura. Por exemplo, com a precisão da classe EX igual 25% e a cobertura igual a 18,18% calculadas anteriormente, tem-se a Medida-f igual a 21,05%.

Tabela 11 – Matriz de confusão: abordagem TOP(X) original - Positivos

Valor Real	Valor Predito				Total
	EXC	BM	SF	ISF	
EXC	2	4	5	0	11
BM	2	1	9	2	14
SF	4	2	30	8	44
ISF	0	2	23	6	31
Total	8	9	67	16	100

Tabela 12 – Matriz de confusão: abordagem TOP(X) com a RNA MLP - Positivos

Valor Real	Valor Predito				Total
	EXC	BM	SF	ISF	
EXC	6	3	2	0	11
BM	2	4	8	0	14
SF	0	4	32	8	44
ISF	0	0	25	6	31
Total	8	11	67	14	100

As Tabelas 11, 12, 13 e 14 apresentam as Matrizes de confusão das duas abordagens, tanto para os comentários positivos quanto para os negativos. Nelas, é possível observar a relação entre os valores preditos pelas duas abordagens e o valor real anotado por um especialista.

A Tabela 15 apresenta os resultados das medidas de Precisão, Cobertura e Medida-F para os comentários positivos na abordagem Top(x) e na nova abordagem Top(x)

Tabela 13 – Matriz de confusão: abordagem TOP(X) original - Negativos

Valor Real	Valor Predito				Total
	EXC	BM	SF	ISF	
EXC	2	0	2	0	4
BM	2	1	7	0	10
SF	4	5	54	0	63
ISF	2	1	18	2	23
Total	10	7	81	2	100

Tabela 14 – Matriz de confusão: abordagem TOP(X) com a RNA MLP - Negativos

Valor Real	Valor Predito				Total
	EXC	BM	SF	ISF	
EXC	2	2	0	0	4
BM	2	8	0	0	10
SF	0	9	53	1	63
ISF	0	1	19	3	23
Total	4	20	72	4	100

Tabela 15 – Comparação dos comentários positivos nas duas abordagens

	TOP(X) Positivos		
	Precisão	Cobertura	Medida-F
EX e BM	17,65%	12,00%	14,29%
IF	37,50%	19,35%	25,53%
TODOS	39,00%	39,00%	39,00%
	TOP(X) RNA Positivos		
	Precisão	Cobertura	Medida-F
EX e BM	52,63%	40,00%	45,45%
IF	42,86%	19,35%	26,67%
TODOS	48,00%	48,00%	48,00%

com a RNA MLP, utilizando o **Subcórpus II**. Levando em conta a Medida-F, média harmônica entre Precisão e Cobertura, observa-se que a abordagem Top(x) com a RNA MLP possui desempenho superior na identificação dos comentários excelentes e bons (EX e BM). Já nos comentários insuficientes (IF) a melhora é pouco significativa e, por fim, quando se agrupam todos os comentários há uma melhora razoável. É importante destacar que os comentários suficientes (SF) não foram apresentados, pois, normalmente, os usuários têm interesse nos melhores comentários ou nos piores (comentários mais críticos). Destaca-se ainda que os resultados baixos (Medida-F < 50%) podem ser justificados pela dificuldade do especialista em avaliar precisamente, seja na avaliação da reputação do autor (**Subcórpus I**), em que o especialista deve evitar o enviesamento da amostra

para não comprometer a capacidade preditiva da RNA, seja na avaliação da importância dos comentários (**Subcorp II**), que pode gerar inconsistências no processo de avaliação final das abordagens.

Tabela 16 – Comparação dos comentários negativos nas duas abordagens

	TOP(X) Negativos		
	Precisão	Cobertura	Medida-F
EX e BM	17,65%	21,43%	19,35%
IF	100,00%	8,70%	16,00%
TODOS	59,00%	59,00%	59,00%
	TOP(X) RNA Negativos		
	Precisão	Cobertura	Medida-F
EX e BM	41,67%	71,43%	52,63%
IF	75,00%	13,04%	22,22%
TODOS	66,00%	66,00%	66,00%

A Tabela 16 apresenta os resultados das medidas de Precisão, Cobertura e Medida-F para os comentários negativos na abordagem Top(x) e na nova abordagem Top(x) com a RNA MLP. A abordagem Top(x) com a RNA MLP mostra-se superior em desempenho em todos os casos, especialmente na identificação dos comentários excelentes e bons. Já na identificação dos comentários insuficientes e todos agrupados, o desempenho é ligeiramente superior ao comparar com a abordagem Top(x) original.

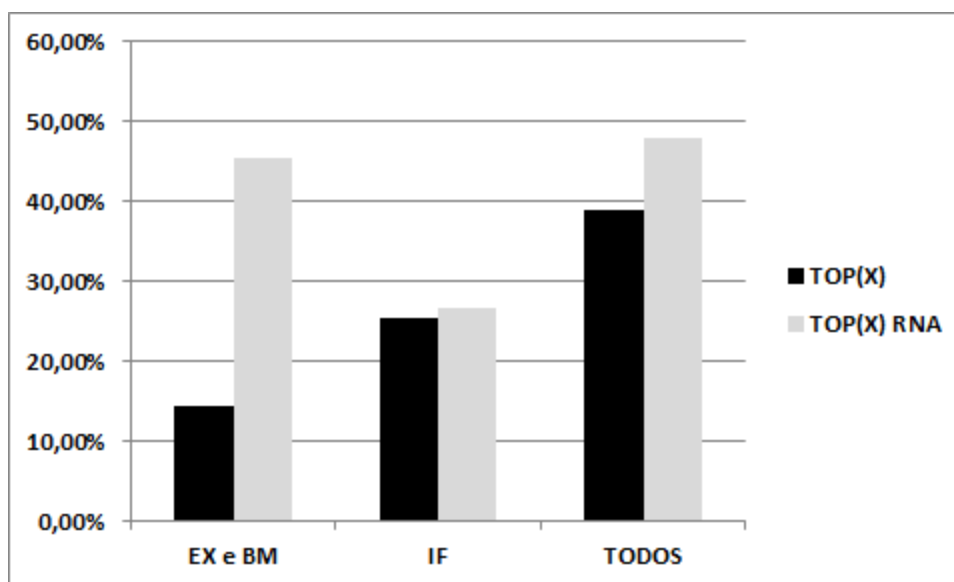


Figura 19 – Gráfico comparativo entre os comentários positivos

A Figura 19 apresenta graficamente a comparação entre as duas abordagens relacionando os comentários positivos em termos de sua importância, utilizando a Medida-F. É possível observar que a abordagem Top(x) com a RNA MLP supera com boa margem a abordagem original nos comentários excelentes e bons, estes são relevantes pois o usuário final vai sempre buscar os melhores comentários para ler e decidir sobre a compra de

um produto ou serviço. Desta forma, o usuário não vai precisar ler um número grande de comentários e poderá focar em um pequeno grupo de comentários selecionados pela abordagem, gerando um ganho de tempo e esforço na pesquisa pelo produto que deseja adquirir.

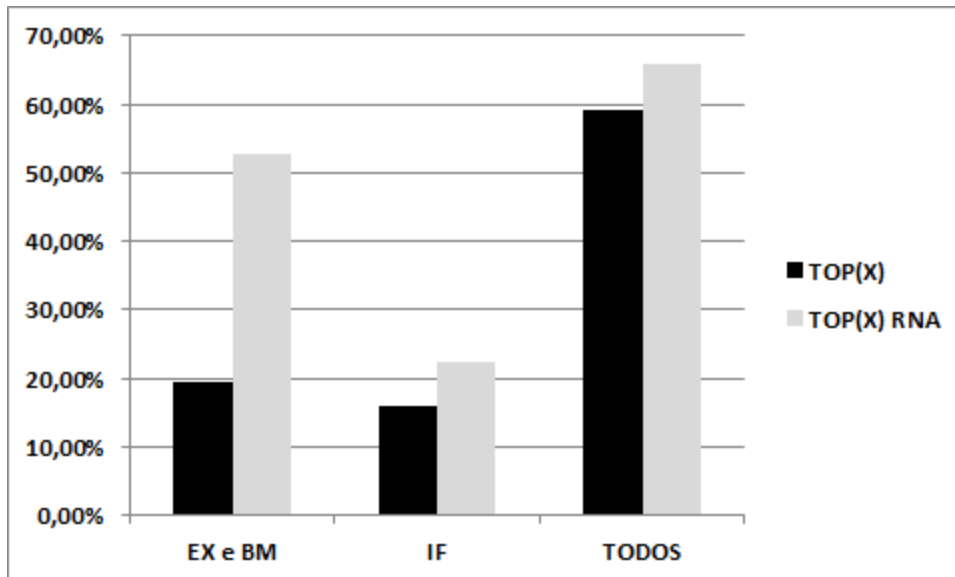


Figura 20 – Gráfico comparativo entre os comentários negativos

A Figura 20 apresenta um gráfico comparando as duas abordagens com relação aos comentários negativos e sua respectiva importância, também com a Medida-F. Novamente é possível observar que abordagem Top(x) com a RNA MLP também supera com boa margem a abordagem original nos comentários excelentes e bons.

4.5 Análise de erros

Esta seção tem como objetivo analisar o motivo de alguns resultados obtidos terem sido abaixo do esperado. Inicialmente, um dos fatores que podem ter contribuído para os resultados considerados insatisfatórios foi o processo de anotação manual do *Cópus*, que nesta pesquisa foi realizado de três formas diferentes: quanto a polaridade, quanto a reputação do autor e quanto a importância. Neste processo de avaliação realizada por seres humanos, os especialistas podem divergir quanto as suas decisões em algum momento devido a natureza subjetiva das opiniões coletadas. Neste sentido, algumas pesquisas demonstram que em anotações feitas por humanos, dificilmente o consenso é maior que 75% (BRUCE; WIEBE, 1999; KU; LIANG; CHEN, 2006; PANG; LEE; VAITHYANATHAN, 2002; WIEBE; CARDIE, 2005b).

A Figura 21 apresenta um diagrama com a relação entre as classes de importâncias anotadas manualmente e as inferidas pelas abordagens Top(x) e Top(x) com a RNA. Pode-se observar que, em muitos casos, as duas abordagens conseguem atingir um bom número

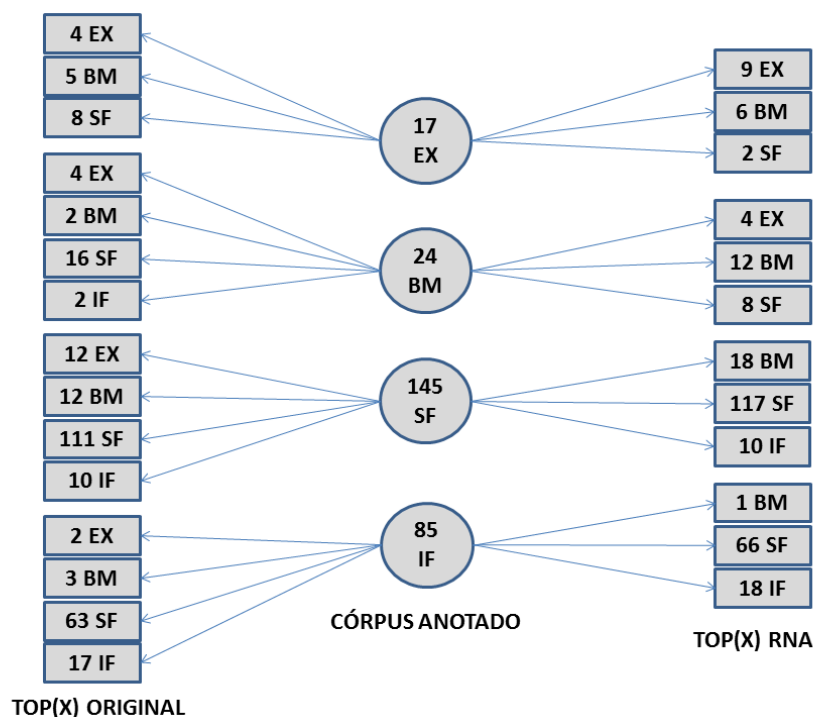


Figura 21 – Análise de erros

de acertos na classificação, no entanto, em outros existem muitos erros na inferência. Por exemplo, as duas abordagens apresentam bom desempenho para inferir os comentários "suficientes", porém para os comentários "insuficientes", o desempenho de ambas é fraco. Outro aspecto a ser analisado se dá quando a abordagem falha ao inferir uma classe corretamente por uma pequena margem, por exemplo, o Sistema *Fuzzy* infere uma saída numérica 6,9 classificando o comentário linguisticamente como "bom", porém a anotação manual o classificou como "excelente", gerando um erro.

Além dos problemas citados acima, suspeita-se que outros aspectos podem causar inconsistências no processo de classificação das abordagens. Por exemplo, imprecisões das ferramentas que implementam as técnicas de aprendizagem de máquina, enviesamento no processo de treinamento/teste da Rede Neural, inconsistências na elaboração das regras de pertinência e variáveis linguísticas do Sistema *Fuzzy*.

Os erros mencionados podem ser corrigidos com a sistematização do processo de anotação do Córpus e a realização de mais experimentos.

4.6 Considerações Finais

Este Capítulo detalhou os experimentos realizados para avaliar a proposta apresentada. Inicialmente, foi aplicada uma RNA MLP sobre o **Subcórpus I** para realizar o treinamento da mesma, após isso, as predições foram generalizadas sobre todos os comentários e seus autores, atribuindo notas de 0 a 10 para cada um. Foi detalhada também a

topologia utilizada pela RNA MLP, incluindo parâmetros internos como: neurônios por camadas e funções de ativação por camada. Como resultado relevante desta etapa tem-se a importância calculada para cada variável de entrada, que era um dos objetivos iniciais da pesquisa.

Este Capítulo apresentou também um experimento com uma Rede Neural com Funções de Base Radial. O objetivo deste experimento foi realizar uma comparação com a RNA MLP em termos de performance no processo de inferência na classificação da reputação do autor no Córpus coletado. Os resultados obtidos apresentaram similaridades com os resultados da RNA MLP quanto a indicação da importância das variáveis de entrada, pouca variação foi observada entre as duas arquiteturas. Porém, no que tange a precisão obtida pela Rede RBF, o valor alcançado foi inferior ao da RNA MLP, ficando assim, esta última como técnica de predição escolhida para o prosseguimento dos experimentos.

A segunda etapa do experimento teve como meta avaliar os resultados obtidos pela abordagem nova com a RNA MLP. Os resultados foram apresentados e comparados com a abordagem original de [Sousa \(2015\)](#) sobre o mesmo Córpus, demonstrando a superioridade em termos das medidas de avaliação, a saber: Precisão, Cobertura e Medida-F.

Por fim, uma seção com análise de erros discutiu os fatores que podem ter contribuído para o baixo desempenho das abordagens no processo de classificação.

Conclusão e Trabalhos Futuros

Este Capítulo apresenta a conclusão da pesquisa e as propostas de trabalhos futuros.

Conclusão

Neste trabalho foi apresentada uma nova abordagem para avaliar comentários de produtos na Web baseado na abordagem Top(x) de Sousa (2015). A nova abordagem propõe uma melhoria na dimensão da Reputação do Autor, que na abordagem original considera somente a quantidade de comentários para inferir a sua reputação, podendo beneficiar os *spammers*. Para resolver tal empecilho, este trabalho apresentou um conjunto de seis medidas extraídas de um Córpus coletado no *site* Buscapé. O Córpus foi anotado quanto a orientação dos comentários, a importância dos mesmos e da reputação do autor, resultado em um conjunto de 271 comentários de um universo de 2000, denominado **Subcórpus II**.

Segue que, com as medidas extraídas relacionadas a reputação do autor, foi aplicada uma RNA MLP para generalizar os resultados da amostragem para o universo inteiro. Tal procedimento obteve uma precisão de **91,01%** na classificação. O experimento com a RNA MLP se deu sobre 356 comentários, denominado **Subcórpus I**.

Além de analisar as medidas para definir a reputação do autor nos *sites* de *e-commerce*, esta pesquisa examinou um conjunto de medidas utilizadas em outros sistemas web, como: ambientes *Wiki*, redes sociais, fóruns *online* e blogs. O estudo sobre esses sistemas web apresentou diversos recursos para a avaliação de autores que foram aplicados com sucesso e que poderiam ser utilizados em *sites* de *e-commerce*. De fato, este trabalho sugere que os portais de comércio eletrônico possam expandir os seus métodos de avaliação de comentários no sentido de reforçar a importância do autor e sua reputação. Uma maneira de realizar essa expansão é permitir a integração dos perfis de usuários dos *sites* com os seus respectivos perfis em redes sociais *online*.

Com relação a relevância das medidas exploradas, esta pesquisa observou que a quantidade de votos positivos que um autor recebe tem um peso significativo em sua reputação, sendo considerada a principal medida para avaliar a reputação do autor no contexto trabalhado. Tal fato demonstra que informações atribuídas por outros usuários possuem mais peso na reputação do que informações intrínsecas ao usuários, como a quantidade de comentários postados pelo mesmo ou a sua data de cadastro no *site*. Porém, um problema inerente ao utilizar os votos para avaliar a reputação é que usuários novos na rede serão prejudicados, neste caso, informações textuais devem ser investigadas em

pesquisas futuras.

Por fim, foram calculadas e comparadas as medidas de avaliação das abordagens, a saber, Precisão, Cobertura e Medida-F. Utilizando como base a média harmônica Medida-F, a nova abordagem obteve melhores resultados para classificar a importância dos comentários excelentes e bons, insuficientes e ao agrupar todos os comentários. Com foco nos comentários excelente e bons, a nova abordagem apresentou resultados significativamente superiores, auxiliando assim aos usuários finais na sua pesquisa por produtos ou serviços, reduzindo o tempo e esforço no processo. No entanto, considera-se ainda abaixo do esperado, pois os resultados na classificação ficaram em torno dos 50%.

Trabalho Futuros

Como trabalhos futuros espera-se aplicar a nova abordagem em um Córpus maior, realizando um processo mais extenso de avaliação realizada por seres humanos, e explorar medidas mais eficientes para definir a reputação do autor. Métodos de avaliação automatizado também poderão ser estudados e aplicados. De fato, pesquisadores do nosso grupos estão realizando trabalhos em outras dimensões da abordagem $\text{Top}(x)$, por exemplo, explorando melhor a riqueza do vocabulário e também há um trabalho que substitui o Sistema *Fuzzy* por uma RNA que identifica melhor os comentários mais importantes (SANTOS et al., 2016).

Referências

- ADLER, B. T.; ALFARO, L. de. A content-driven reputation system for the wikipedia. In: *Proceedings of the 16th International Conference on World Wide Web*. New York, NY, USA: ACM, 2007. (WWW '07), p. 261–270. Citado na página 33.
- ADLER, B. T.; ALFARO, L. de; PYE, I. Detecting wikipedia vandalism using wikitrust. In: *PAN 2010 proceedings*. [S.l.: s.n.], 2010. Citado na página 33.
- ADLER, B. T. et al. Measuring author contributions to the wikipedia. In: *Proceedings of the 4th International Symposium on Wikis*. New York, NY, USA: ACM, 2008. (WikiSym '08), p. 15:1–15:10. Citado na página 33.
- ADLER, B. T. et al. Assigning trust to wikipedia content. In: *Proceedings of the 4th International Symposium on Wikis*. New York, NY, USA: ACM, 2008. (WikiSym '08), p. 26:1–26:12. Citado na página 33.
- AGARWAL, N. et al. Identifying the influential bloggers in a community. In: *Proceedings of the 2008 International Conference on Web Search and Data Mining*. New York, NY, USA: ACM, 2008. (WSDM '08), p. 207–218. Citado na página 38.
- AGGARWAL, A.; KUMARAGURU, P. Followers or phantoms? an anatomy of purchased twitter followers. *CoRR*, 2014. Citado 2 vezes nas páginas 36 e 50.
- AKAY, A.; DRAGOMIR, A.; ERLANDSSON, B.-E. Network-based modeling and intelligent data mining of social media for improving care. *IEEE Journal of Biomedical and Health Informatics*, v. 19, n. 1, p. 210–218, Jan 2015. ISSN 2168-2194. Citado na página 37.
- ALUÍSIO, S.; PELIZZONI, J.; MARCHI, A. R. An account of the challenge of tagging a reference corpus for brazilian portuguese. In: MAMEDE, N. J. et al. (Ed.). *Computational Processing of the Portuguese Language: 6th International Workshop, PROPOR 2003 Faro, Portugal, June 26–27, 2003 Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003. Citado na página 27.
- ANTHONY, D. L.; SMITH, S. W.; WILLIAMSON, T. Reputation and reliability in collective goods: the case of the online encyclopedia wikipedia. *Rationality and Society*, p. 283–306, 2009. Citado na página 34.
- BIRD, S.; KLEIN, E.; LOPER, E. *Natural Language Processing with Python*. [S.l.]: O'Reilly Media, Inc., 2012. Citado 2 vezes nas páginas 26 e 27.
- BRIDLE, J. S. Advances in neural information processing systems 2. In: TOURETZKY, D. S. (Ed.). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1990. cap. Training Stochastic Model Recognition Algorithms As Networks Can Lead to Maximum Mutual Information Estimation of Parameters. Citado na página 50.
- BRUCE, R. F.; WIEBE, J. M. Recognizing subjectivity: A case study of manual tagging. *Natural Language Engineering*, v. 5, p. 187–205, 1999. Citado na página 58.

- CAPPELLETTI, R.; SASTRY, N. Iarank: Ranking users on twitter in near real-time, based on their information amplification potential. In: *SocialInformatics*. [S.l.]: IEEE Computer Society, 2012. p. 70–77. Citado na página 37.
- CARVALHO, P. et al. Clues for detecting irony in user-generated contents: Oh...!! it's "so easy";-). In: *Proceedings of the 1st International CIKM Workshop on Topic-sentiment Analysis for Mass Opinion*. New York, NY, USA: ACM, 2009. p. 53–56. Citado na página 3.
- CASTELLS, M. *The Rise of the Network Society*. 2nd. ed. Cambridge, MA, USA: Blackwell Publishers, Inc., 2000. ISBN 0631221409. Citado na página 1.
- CAUDRON, S. *The Search for Meaning at Work*. [S.l.]: Training and Development, 1997. Citado na página 16.
- CHEN, B.-C. et al. User reputation in a comment rating environment. In: *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, 2011. (KDD '11), p. 159–167. Citado na página 39.
- CHOI, S.-M.; HAN, Y.-S. Representative reviewers for internet social media. *Expert Syst. Appl.*, Pergamon Press, Inc., Tarrytown, NY, USA, v. 40, n. 4, p. 1274–1282, mar. 2013. Citado na página 37.
- CHORLEY, M. J. et al. Better the tweeter you know: Social signals on twitter. In: *2012 International Conference on Privacy, Security, Risk and Trust*. [S.l.: s.n.], 2012. p. 277–282. Citado na página 36.
- DARWIN, C. *On the Origin of Species by Means of Natural Selection, Or, The Preservation of Favoured Races in the Struggle for Life*. [S.l.]: J. Murray, 1859. Citado na página 18.
- DEEPHOUSE, D. L. Media reputation as a strategic resource: An integration of mass communication and resource-based theories. *Journal of management*, Sage Publications, v. 26, n. 6, p. 1091–1112, 2000. Citado na página 16.
- DELLAROCAS, C. Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. In: *Proceedings of the 2Nd ACM Conference on Electronic Commerce*. New York, NY, USA: ACM, 2000. (EC '00), p. 150–157. ISBN 1-58113-272-7. Citado na página 38.
- GILES, G. Internet encyclopedias go head to head. *Nature*, December 2005. Citado na página 31.
- GOLBECK, J. et al. Predicting personality from twitter. In: *2011 International Conference on Privacy, Security, Risk and Trust*. [S.l.: s.n.], 2011. p. 149–156. Citado na página 35.
- GOLDENBERG, J.; LIBAI, B.; MULLER, E. Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Letters*, v. 12, n. 3, p. 211–223, 2001. Citado na página 2.
- GONÇALVES, P. et al. Bazinga! caracterizando e detectando sarcasmo e ironia no twitter. In: *In Proceedings of the Proceedings of the Brazilian Workshop on Social Network Analysis and Mining (BraSNAM)*. [S.l.: s.n.], 2015. Citado na página 3.

- HAGAN, M. T.; MENHAJ, M. B. Training feedforward networks with the marquardt algorithm. *Trans. Neur. Netw.*, IEEE Press, Piscataway, NJ, USA, v. 5, n. 6, p. 989–993, nov. 1994. ISSN 1045-9227. Citado na página 23.
- HALFAKER, A. et al. A jury of your peers: Quality, experience and ownership in wikipedia. In: *Proceedings of the 5th International Symposium on Wikis and Open Collaboration*. New York, NY, USA: ACM, 2009. (WikiSym '09), p. 15:1–15:10. Citado na página 32.
- HALL, R. A Framework Linking Intangible Resources and Capabilities to Sustainable Competitive Advantage. *Strategic Management Journal*, v. 14, p. 607–618, 1993. Citado na página 16.
- HAMILTON, R.; VOHS, K. D.; MCGILL, A. L. We'll be honest, this won't be the best article you'll ever read: The use of dispreferred markers in word-of-mouth communication. *Journal of Consumer Research*, v. 41, n. 1, p. 197 – 212, 2014. Citado na página 3.
- HAN, Y.; KIM, L.; CHA, J. Computing user reputation in a social network of web 2.0. *Computing and Informatics*, v. 31, n. 2, p. 447–462, 2012. Citado na página 38.
- HARTMANN, N. et al. A large corpus of product reviews in portuguese: Tackling out-of-vocabulary words. In: CHAIR), N. C. C. et al. (Ed.). *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. Reykjavik, Iceland: European Language Resources Association (ELRA), 2014. ISBN 978-2-9517408-8-4. Citado na página 3.
- HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. 1st. ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1994. ISBN 0023527617. Citado na página 22.
- HEINONEN, K. Consumer activity in social media: Managerial approaches to consumers' social media behavior. *Journal of Consumer Behaviour*, Wiley Online Library, v. 10, n. 6, p. 356–364, 2011. Citado na página 16.
- HOISL, B.; AIGNER, W.; MIKSCH, S. Online communities and social computing: Second international conference, ocsc 2007, held as part of hci international 2007, beijing, china, july 22-27, 2007. proceedings. In: _____. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. cap. Social Rewarding in Wiki Systems – Motivating the Community, p. 362–371. Citado na página 34.
- HU, N.; LIU, L.; ZHANG, J. J. Do online reviews affect product sales? the role of reviewer characteristics and temporal effects. *Inf. Technol. and Management*, Kluwer Academic Publishers, Hingham, MA, USA, v. 9, n. 3, p. 201–214, set. 2008. Citado na página 15.
- JACKSON, P.; MOULINIER, I. *Natural language processing for online applications: Text retrieval, extraction and categorization*. Amsterdam: John Benjamins, 2007. Citado 2 vezes nas páginas 2 e 9.
- JAVANMARDI, S.; LOPES, C.; BALDI, P. Modeling user reputation in wikis. *Statistical Analysis and Data Mining*, John Wiley & Sons, Inc., New York, NY, USA, v. 3, n. 2, p. 126–139, abr. 2010. Citado na página 32.
- JONES, C.; HESTERLY, W. S.; BORGATTI, S. P. A general theory of network governance: Exchange conditions and social mechanisms. *The Academy of Management Review*, Academy of Management, v. 22, n. 4, p. 911–945, 1997. Citado na página 16.

- KIM, Y. A.; SRIVASTAVA, J. Impact of social influence in e-commerce decision making. In: *Proceedings of the Ninth International Conference on Electronic Commerce*. New York, NY, USA: ACM, 2007. (ICEC '07), p. 293–302. Citado na página 2.
- KLEINBERG, J. M. Authoritative sources in a hyperlinked environment. *J. ACM*, ACM, New York, NY, USA, 1999. Citado na página 39.
- KU, L.-W.; LIANG, Y.-T.; CHEN, H.-H. Opinion extraction, summarization and tracking in news and blog corpora. In: *Proceedings of AAAI-CAAW-06, the Spring Symposia on Computational Approaches to Analyzing Weblogs*. Stanford, US: [s.n.], 2006. Citado na página 58.
- KWAK, H. et al. What is twitter, a social network or a news media? In: *Proceedings of the 19th International Conference on World Wide Web*. New York, NY, USA: ACM, 2010. (WWW '10), p. 591–600. Citado 2 vezes nas páginas 34 e 50.
- LI, B. et al. A topic-biased user reputation model in rating systems. *Knowledge and Information Systems*, 2014. Citado na página 15.
- LI, M. et al. Helpfulness of online product reviews as seen by consumers: Source and content features. *Int. J. Electronic Commerce*, v. 17, n. 4, p. 101–136, 2013. Citado 2 vezes nas páginas 3 e 51.
- LIU, B. Sentiment analysis and subjectivity. In: *Handbook of Natural Language Processing, Second Edition*. Taylor and Francis Group, Boca. [S.l.: s.n.], 2010. Citado na página 1.
- LIU, B. Sentiment analysis and subjectivity. In: INDURKHAYA, N.; DAMERAU, F. J. (Ed.). *Handbook of Natural Language Processing, Second Edition*. Boca Raton, FL: CRC Press, Taylor and Francis Group, 2010. ISBN 978-1420085921. Citado 3 vezes nas páginas 11, 13 e 15.
- LIU, B. *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2011. Citado na página 3.
- LUGER, G. *Artificial intelligence: Structures and strategies for complex problem solving*. 5nd. ed. [S.l.]: Addison-Wesley Longman., 2005. Citado 2 vezes nas páginas 17 e 18.
- LV, P. et al. Ohrank: An algorithm integrating mentality and influence of opinion holder for opinion mining. *Chinese Journal of Electronics*, P.O. Box 165, Beijing, 100036, China, v. 22, n. 4, p. 655 – 660, 2013. ISSN 10224653. Citado na página 38.
- MAMDANI, E. H.; ASSILIAN, S. An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies*, n. 1, p. 1–13, 1975. Citado 2 vezes nas páginas 26 e 44.
- MANNING, C. D.; SCHÜTZ, H. *Foundations of Statistical Natural Language Processing*. Cambridge, MA, USA: MIT Press, 1999. ISBN 0-262-13360-1. Citado na página 9.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 1943. Citado na página 19.
- MCNALLY, K.; OMAHONY, M. P.; SMYTH, B. Evaluating user reputation in collaborative web search. In: . [S.l.]: ACM, 2011. Citado na página 38.

- MOVSHOVITZ-ATTIAS, D. et al. Analysis of the reputation system and user contributions on a question answering website: Stackoverflow. In: *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. New York, NY, USA: ACM, 2013. (ASONAM '13), p. 886–893. Citado na página 39.
- PAGE, L. et al. The pagerank citation ranking: Bringing order to the web. In: *Proceedings of the 7th International World Wide Web Conference*. Brisbane, Australia: [s.n.], 1998. p. 161–172. Citado na página 35.
- PALTOGLOU, G.; GIACHANOU, A. Opinion retrieval: Searching for opinions in social media. In: _____. *Professional Search in the Modern World: COST Action IC1002 on Multilingual and Multifaceted Interactive Information Access*. Cham: Springer International Publishing, 2014. Citado na página 44.
- PANG, B.; LEE, L.; VAITHYANATHAN, S. Thumbs up?: Sentiment classification using machine learning techniques. In: *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10*. Stroudsburg, PA, USA: Association for Computational Linguistics, 2002. (EMNLP '02), p. 79–86. Citado na página 58.
- POWERS, D. M. W. Evaluation: From precision, recall and f-measure to roc., informedness, markedness & correlation. *Journal of Machine Learning Technologies*, v. 2, n. 1, p. 37–63, 2011. Citado na página 54.
- RABELO, J.; PRUDENCIO, R. B. C.; BARROS, F. Collective classification for sentiment analysis in social networks. In: *Proceedings of the 2012 IEEE 24th International Conference on Tools with Artificial Intelligence - Volume 01*. Washington, DC, USA: IEEE Computer Society, 2012. (ICTAI '12), p. 958–963. ISBN 978-0-7695-4915-6. Citado 2 vezes nas páginas 35 e 40.
- RAUB, W.; WEESIE, J. Reputation and efficiency in social interactions: An example of network effects,. *American Journal of Sociology*, v. 96, p. 626–654, 1990. Citado na página 16.
- RIJSBERGEN, C. J. V. *Information Retrieval*. 2nd. ed. Newton, MA, USA: Butterworth-Heinemann, 1979. ISBN 0408709294. Citado na página 55.
- ROBERTSON, S.; ZARAGOZA, H. The probabilistic relevance framework: Bm25 and beyond. *Found. Trends Inf. Retr.*, Now Publishers Inc., Hanover, MA, USA, p. 333–389, 2009. Citado na página 38.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1. In: RUMELHART, D. E.; MCCLELLAND, J. L.; GROUP, C. P. R. (Ed.). Cambridge, MA, USA: MIT Press, 1986. cap. Learning Internal Representations by Error Propagation, p. 318–362. ISBN 0-262-68053-X. Citado na página 22.
- SABATER, J.; SIERRA, C. Reputation and social network analysis in multi-agent systems. In: *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems: Part 1*. New York, NY, USA: ACM, 2002. (AAMAS '02), p. 475–482. Citado na página 16.

- SAMUEL, A. L. Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 1959. Citado na página 17.
- SANTOS, R. L. de S. et al. An experimental study based on fuzzy systems and artificial neural networks to estimate the importance of reviews about product and services. In: *2016 International Joint Conference on Neural Networks (IJCNN)*. [S.l.: s.n.], 2016. Citado na página 62.
- SBC. Grandes desafios da pesquisa em computacao no brasil 2006 - 2016. *Sociedade Brasileira de Computacao, Porto Alegre, RS, Brasil, 2006*, 2006. Citado na página 1.
- SILVA, I. da; SPATTI, D.; FLAUZINO, R. *Redes Neurais Artificiais para Engenharia e Ciências Aplicadas - Curso Prático*. ARTLIBER, 2010. ISBN 9788588098534. Disponível em: <<https://books.google.com.br/books?id=w2VHbwAACAAJ>>. Citado 6 vezes nas páginas 15, 19, 20, 21, 22 e 53.
- SINHA, R. et al. Comparing recommendations made by online systems and friends. In: *In Proceedings of the DELOS-NSF Workshop on Personalization and Recommender Systems in Digital Libraries*. [S.l.: s.n.], 2001. Citado na página 4.
- SOUSA, R. F. de. *Abordagem TOP(X) para Inferir os Comentários mais Importantes sobre Produtos e Serviços*. Dissertação (Mestrado) — Universidade Federal do Piauí, Teresina-PI, 2015. Citado 15 vezes nas páginas 15, 4, 5, 6, 7, 23, 30, 43, 44, 45, 47, 48, 49, 60 e 61.
- SPOOL, J. *The Magic Behind Amazon's 2.7 Billion Dollar Question*. 2009. <<https://articles.ue.com/magicbehindamazon/>>. Acessado em 05/04/2016. Citado na página 4.
- SPSS. *SPSS for Windows, Version 16.0*. Chicago USA: [s.n.], 2007. Citado na página 27.
- TAKAGI, T.; SUGENO, M. Fuzzy Identification of Systems and Its Applications to Modeling and Control. *IEEE Transactions on Systems, Man, and Cybernetics*, v. 15, n. 1, p. 116–132, fev. 1985. Citado na página 26.
- TANSCHKEIT, R. Sistemas fuzzy. *Inteligência computacional: aplicada a administração, economia e engenharia em Matlab*, p. 229–264, 2004. Citado 4 vezes nas páginas 15, 23, 24 e 25.
- THET, T. T.; NA, J.-C.; KHOO, C. S. Aspect-based sentiment analysis of movie reviews on discussion boards. *Journal of Information Science*, Sage Publications, Inc., Thousand Oaks, CA, USA, v. 36, n. 6, dez. 2010. Citado na página 12.
- TRIPADVISOR. *Independent Study of more than 12,000 global respondents reveals online travel review trends*. 2013. <https://www.tripadvisor.com/PressCenter-i6603-c1-Press_Releases.html>. Acessado em 04/04/2016. Citado na página 2.
- TUPES, E.; CHRISTAL, R. Recurrent personality factors based on trait ratings. *Journal of Personality*, 1992. Citado na página 35.
- VOHS, K. D. et al. Making Choices Impairs Subsequent Self-Control: A Limited-Resource Account of Decision Making, Self-Regulation, and Active Initiative. *Journal of Personality and Social Psychology*, v. 94, n. 5, p. 883–898, 2008. Citado na página 2.

- WANG, A. H. Don't follow me: Spam detection in twitter. In: *Proceedings of the 2010 International Conference on Security and Cryptography (SECRYPT)*. [S.l.: s.n.], 2010. p. 1–10. Citado na página 36.
- WEITZEL, L.; OLIVEIRA, J. P. M. de; QUARESMA, P. Measuring the reputation in user-generated-content systems based on health information. *Procedia Computer Science*, v. 29, p. 364 – 378, 2014. 2014 International Conference on Computational Science. Citado na página 35.
- WEITZEL, L. b.; OLIVEIRA, J. D.; QUARESMA, P. Exploring trust to rank reputation in microblogging. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 8056 LNCS, n. PART 2, p. 434–441, 2013. Citado na página 16.
- WELCH, M. J. et al. Topical semantics of twitter links. In: *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*. New York, NY, USA: ACM, 2011. (WSDM '11), p. 327–336. Citado na página 37.
- WENG, J. et al. Twiterrank: Finding topic-sensitive influential twitterers. In: *Proceedings of the Third ACM International Conference on Web Search and Data Mining*. New York, NY, USA: ACM, 2010. (WSDM '10), p. 261–270. Citado 2 vezes nas páginas 36 e 50.
- WIEBE, J.; CARDIE, C. Annotating expressions of opinions and emotions in language. language resources and evaluation. In: *Language Resources and Evaluation (formerly Computers and the Humanities)*. [S.l.: s.n.], 2005. p. 2005. Citado na página 14.
- WIEBE, J.; CARDIE, C. Annotating expressions of opinions and emotions in language. language resources and evaluation. In: *Language Resources and Evaluation (formerly Computers and the Humanities)*. [S.l.: s.n.], 2005. p. 2005. Citado na página 58.
- WILSON, T.; WIEBE, J.; HWA, R. Just how mad are you? finding strong and weak opinion clauses. In: *Proceedings of the 19th National Conference on Artificial Intelligence*. [S.l.]: AAAI Press, 2004. (AAAI'04), p. 761–767. Citado na página 14.
- WÖHNER, T.; KÖHLER, S.; PETERS, R. Automatic reputation assessment in wikipedia. *Thirty Second International Conference on Information Systems, Shanghai 2011*, 2011. Citado 2 vezes nas páginas 16 e 32.
- ZADEH, L. Fuzzy sets. *Information Control*, v. 8, p. 338–353, 1965. Citado na página 23.
- ZHAO, H. et al. Socialwiki: Bring order to wiki systems with social context. In: *Social Informatics - Second International Conference, SocInfo 2010, Laxenburg, Austria*. [S.l.: s.n.], 2010. p. 232–247. Citado na página 34.
- ZHOU, G. et al. Topic-sensitive probabilistic model for expert finding in question answer communities. In: *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*. Maui, HI, United states: [s.n.], 2012. p. 1662 – 1666. Citado na página 38.

Anexos

ANEXO A – Etiquetas Mac-Morpho

Classe Gramatical	Etiqueta	Exemplos
Adjetivo	ADJ	bom, ruim, ótimo, péssimo
Advérbio	ADV	muito, pouco, normalmente
Advérbio Conectivo Subordinativo	ADV-KS	Sei <i>onde</i> mora
Advérbio Relativo Subordinativo	ADV-KS-REL	onde, como, quando
Artigo	ART	o, a, os, as
Conjunção Coordenativa	KC	e, nem, mas, ou, pois
Conjunção Subordinativa	KS	que, porque, assim
Interjeição	IN	ufa!, viva!, ai!, oi!
Numeral	NUM	três, quatro, 3, 4
Palavra Denotativa	PDEN	até, apenas, eis, cá
Particípio	PCP	dormido, espalhado, tido
Pronome Adjetivo	PROADJ	meu, nosso, este, algum
Pronome Conectivo Subordinativo	PRO-KS	Sei <i>quem</i> chegou
Pronome Conectivo Subord. Relativo	PRO-KS-REL	o qual, cujo
Pronome Pessoal	PROPESS	eu, me, Vossa Alteza
Pronome Substantivo	PROSUB	isto, isso, aquilo, alguém
Símbolo de Moeda Corrente	CUR	R\$, US\$
Substantivo	N	hotel, quarto, atendimento
Substantivo Próprio	NPROP	Maria, Vinícius, Globo
Verbo	V	é, foi, gostar, ir
Verbo Auxiliar	VAUX	ter, haver

ANEXO B – Base de Características

Características	
qualidade	preço
produto	custo
som	música
câmera	vídeo
memória	áudio
interface	display
design	celular
tela	aparelho
conexão	design
antena	desempenho
caneta	resolução
fone de ouvido	tamanho
processador	smartphone
teclado	espessura
imagem	screen
foto	fotografias
ligação	valor
sinal	característica
durabilidade	armazenamento
uso	bateria