



ESTUDO COMPARATIVO DE MINERAÇÃO DE OPINIÕES EM REDE VAREJISTA

Andrea Hecksher¹, Nelson F. F. Ebecken¹

1 Universidade Federal do Rio de Janeiro - UFRJ

RESUMO

O presente artigo compara os resultados obtidos por uma análise realizada a partir de técnicas de Mineração de Textos em conteúdos extraídos da *Fan Page* de uma rede de comércio varejista com os resultados obtidos por uma pesquisa de campo feita para a mesma rede varejista. Este trabalho pretende corroborar a ideia de que as técnicas de Mineração de Textos, em particular a Análise de Sentimentos, apresentam-se cada vez mais como soluções eficientes e vantajosas para conhecimento de opiniões e estudos de percepção e reputação de marcas.

Palavras-chave: Mineração de opiniões, Rede Varejista, Análise de Sentimentos, Satisfação do Cliente.

1. INTRODUÇÃO

Analisar a reputação de marcas, capturar a opinião dos consumidores e adquirir conhecimento sobre suas preferências em relação a produtos e serviços são parte primordial do trabalho de pesquisadores e profissionais de *marketing*. Entretanto, métodos tradicionais de pesquisa de mercado, como grupos focais e entrevistas presenciais, são dispendiosos e demorados.

Já é uma realidade, para qualquer gestor de marca, a importância dos meios digitais e das redes sociais em estratégias de *branding*¹ e nas campanhas publicitárias da empresa. Atualmente, grandes empresas usam as mídias sociais para interagir com os clientes, visando aumento de vendas, melhoria na fidelização, retenção de clientes, criação de uma consciência e reputação da marca. Por meio de geração de conteúdos informativos ou promocionais, buscam aumentar, fidelizar e reter seu público-alvo. O emprego das mídias sociais favorece atividades como: gestão de marcas (por meio de publicidade, *marketing* e distribuição de conteúdo), vendas, atendimento e suporte ao cliente, desenvolvimento de produtos e inovação (Culnan; Mchugh; Zubillaga, 2010; Di Gangi; Wasko; Hooker, 2010 *apud* He; Zha; Li, 2013).

A cada dia aumenta mais o número de empresas que utilizam as mídias sociais, como o *Facebook* e o *Twitter*, para fornecer serviços e interagir com os clientes. Da mesma forma, é crescente o número de clientes que interagem com empresas e marcas por meio desses canais ou que espalham suas opiniões em diferentes *sites* e aplicativos de mídia social. Como resultado, uma grande quantidade de conteúdos gerados espontaneamente por usuários de mídias sociais está disponível gratuitamente. Para efetivamente avaliar seu ambiente de negócios, as empresas precisam monitorar essa massa de informação e transformar dados de mídia social em conhecimento para gestores e profissionais de *marketing*.

A internet facilitou o acesso à informação e à socialização. A disponibilidade gratuita de motores de busca e de *sites* com serviços de comparação de preços e de produtos tem tornado os consumidores cada vez mais exigentes. Os relatos de consumidores e as trocas de ideias nas comunidades virtuais elevam o conteúdo gerado por usuários a uma fonte valorosa para analisar interesses e opiniões dos clientes sobre empresas, produtos e serviços. De uma forma geral, as opiniões e as experiências de consumo relatadas por consumidores comuns têm mais credibilidade para outros pretendidos consumidores do que as propagandas ou qualquer outro formato de estratégia de *marketing*. A discussão virtual tem grande influência na percepção de uma marca ou na escolha de um produto.

1 “*Branding* é uma disciplina híbrida, resultado da associação das áreas de *marketing*, publicidade e *design*, e que trata da gestão, comunicação e forma, respectivamente” (GOMEZ, 2009 *apud* GOMEZ et al., 2012).



O uso inteligente das mídias sociais pelas empresas não se resume em manter um canal de promoção e de comunicação com seu público. Além disso, existe uma função fundamental de prover um panorama sempre atualizado das opiniões dos consumidores sobre seus produtos ou serviços. Sendo assim, *blogs* e *sites* de redes sociais podem ser compreendidos também como medidores e indicadores de reputação de marcas. Por meio de rastreamento e extração de informações relevantes de *posts* e *tweets*, as empresas podem intervir imediatamente na comunicação com clientes insatisfeitos, utilizar comentários de consumidores como *feedback* e identificar preferências e demandas por serviços e produtos.

Com base no fato de que cerca de 20% de *microblogs* mencionam um nome de marca (Jansen *et al.*, 2009), a gestão da percepção da marca em mídias sociais deve ser parte integrante da estratégia de *marketing* da empresa. As redes sociais digitais se tornaram a fonte mais rica e gratuita para conhecimento de opiniões e críticas, de usuários e consumidores, sobre produtos e serviços.

Diante desse cenário, a presente pesquisa exploratória pretende corroborar a ideia de que técnicas de Mineração de Textos, em particular a Análise de Sentimentos, apresentam-se, cada vez mais, como soluções eficientes e vantajosas tanto para conhecimento de opiniões e estudos de percepção e reputação de marcas como para confirmação de suas características de praticidade e de acessibilidade em relação a tempo e custo. A conclusão emerge de uma análise comparativa entre os resultados obtidos por uma pesquisa de campo feita em uma rede de comércio varejista do Rio de Janeiro e os resultados obtidos por uma análise realizada a partir de técnicas de Mineração de Textos em conteúdos extraídos do *Facebook* da rede em questão. Os objetivos, tanto da pesquisa de campo como da mineração e análise da rede social da empresa, consistiram em obter informações em relação às opiniões dos clientes sobre as lojas, o atendimento e o mix de produtos oferecidos pela rede para tomar ciência dos pontos fortes e fracos da empresa.

2. MÉTODOS DA PESQUISA

Para medir a percepção dos clientes da rede varejista, foi realizada uma pesquisa exploratória por meio de técnicas de Mineração de Textos aplicadas em conteúdos retirados da *Fan Page* da rede de comércio varejista.

Para a mineração de opiniões e a análise do conteúdo textual, foram usadas duas técnicas de Mineração de Textos: Extração de Palavras-chave e Análise de Sentimentos.

2.1 Mineração de textos

A Mineração de Textos é uma extensão da Mineração de Dados. A Mineração de Dados refere-se à extração de conhecimento a partir de bases de dados que podem conter vários formatos de informação, como valores numéricos, textos, *hiperlinks* e outros formatos de multimídia. A Mineração de Textos, também conhecida como Mineração de Dados de Texto (Hearst, 1999 *apud* Ranfagni; Guercini; Camiciottoli, 2014), concentra-se na extração e na elaboração da informação a partir estritamente de fontes textuais (Witten, 2005 *apud* Ranfagni; Guercini; Camiciottoli, 2014).

Mineração de Textos é a etapa de análise do processo conhecido como KDT (*Knowledge Discovery in Textual Databases*) e pode ser entendida como um processo de análise de conteúdos textuais que emprega, geralmente, técnicas estatísticas por meio da utilização de algoritmos computacionais para a descoberta de conhecimento em dados textuais não estruturados. As etapas do processo de Mineração de Textos compreendem: pré-processamento, adaptação de algoritmos de Mineração de Dados para aplicações em Mineração de Textos, descoberta de conhecimento na internet e visualização da informação.

De acordo com Tan (1999 *apud* He; Zha; Li, 2013), a Mineração de Textos tem um alto potencial comercial, uma vez que textos são a forma mais natural de armazenamento e troca de informação. Em 1999, Tan já mencionava um estudo o qual indicara que cerca de 80% da informação de uma organização estava contida em documentos de texto, como e-mails, memorandos, correspondência de clientes e relatórios. Desde essa data, o volume de informação textual só aumenta, e, hoje, o maior desafio é o tratamento de toda essa enorme massa de dados textuais. Portanto, para extrair informações úteis a partir de um grande número de documentos textuais, tornou-se imperativa a utilização de técnicas automatizadas de computação (He, 2013a; Liu; Cao; He, 2011 *apud* He; Zha; Li, 2013).

A Mineração de Textos aplica-se, principalmente, à clusterização, sumarização de texto para extração de informações e análise de *links*. A clusterização, ou análise de *cluster*, em particular, tem a vantagem de descobrir padrões, tendências e correlações. Clusterização é a classificação não supervisionada de dados, formando agrupamentos ou *clusters*. Ela representa uma das principais etapas de processos de análise exploratória de dados, denominada Análise de Clusters (Jain; Murty; Flynn, 1999).

Técnicas de Mineração de Textos têm sido aplicadas com sucesso em pesquisas e análises de diversas áreas, tais como negócios, ciências da saúde, educação, busca e recuperação da informação em áreas como biblioteconomia e ciência da informação. Contudo, na maior parte dos casos, a aplicação



da Mineração de Textos tem a finalidade de extrair informações novas e relevantes em grandes quantidades de textos não estruturados.

Exemplos simples de Mineração de Textos que estão se tornando cada vez mais adotados em *websites* são as Nuvens de Palavras ou Nuvens de *Tags*, as quais oferecem um resumo visual do conteúdo temático dos textos já publicados ou visualizações gráficas do ciclo de notícias de determinado tópico ao longo do tempo.

Existe uma grande variedade de aplicativos, comerciais e gratuitos, que podem ser usados para Mineração de Textos: por exemplo, *IBM SPSS Modeler*, *Leximancer*, *SAS Enterprise Miner*, *Polyanalyst* da *Megaputer* e a linguagem e ambiente de desenvolvimento *R*. Por meio de algoritmos estatísticos computacionais sofisticados, essas ferramentas facilitam e automatizam o tratamento de grandes volumes de dados textuais não estruturados para geração de cenários, previsões e *insights*.

De acordo com Ranfagni, Guercini e Camiciottoli (2014), pesquisadores de mercado já começam a mostrar interesse no uso de técnicas de Mineração de Textos em análises de comportamento do consumidor e de percepção de marca. No mesmo artigo, eles propõem um novo conceito metodológico em que a técnica qualitativa da etnografia digital é reforçada pela integração de técnicas quantitativas de Mineração de Textos para análise das relações entre marcas e consumidores em comunidades *on-line*. Para corroborar a relevância da Mineração de Textos em pesquisas recentes, citam exemplos de análises que recorreram a essas técnicas:

- Análise de tendência: Rickman *et al.* (2007 *apud* Ranfagni; Guercini; Camiciottoli, 2014) discutiram a utilização de ferramentas de Mineração de Textos para análise de tendências através do rastreamento de “*buzz*” em *weblogs* de moda.
- Análise de um fórum de reclamação *on-line*: Chen (2009 *apud* Ranfagni; Guercini; Camiciottoli, 2014) utilizou Mineração de Textos para analisar um fórum de reclamação *on-line* a fim de determinar o que os consumidores percebiam como as principais questões e razões para a insatisfação.
- Análise de sentimento: usando *sites* de avaliação do consumidor como fontes de dados, Lee *et al.* (2011 *apud* RANFAGNI; GUERCINI; CAMICIOTTOLI, 2014) aplicaram técnicas de Mineração de Textos para levantamento de avaliações favoráveis e desfavoráveis dos atributos de câmeras digitais para análise de mercado.
- Análise de opiniões de consumidores de produtos eletrônicos: Archak *et al.* (2007 *apud* RANFAGNI;

GUERCINI; CAMICIOTTOLI, 2014) extraíram inteligência de negócios por meio da identificação sistemática de palavras que expressam o nível de satisfação e preferências em relação a características de produtos.

- Análise de opiniões de consumidores do mercado automobilístico: Feldman *et al.* (2008 *apud* RANFAGNI; GUERCINI; CAMICIOTTOLI, 2014), por meio da extração de co-ocorrência de menções de marcas e adjetivos que os consumidores usavam para descrever opiniões, estudaram as preferências dos consumidores para obter uma melhor compreensão do mercado de automóveis.

2.2 Mineração de Opiniões

Mineração de Opiniões é o estudo computacional de opiniões e avaliações de pessoas em relação a entidades como produtos, serviços, organizações, personalidades e eventos. É um campo da pesquisa em Processamento de Linguagem Natural e Mineração da *Web* que, nos últimos anos, vem despertando o interesse de pesquisadores acadêmicos e de mercado.

A Mineração de Opiniões usa a técnica de Análise de Sentimentos, que tem como finalidade extrair informações subjetivas para determinar “polaridade” sobre assuntos específicos a partir de um conjunto de documentos textuais, em sua maioria opiniões *on-line* expressas espontaneamente por usuários em *blogs*, redes sociais e *sites* de reclamação de consumidores.

A Análise de Sentimentos de um segmento de texto envolve a identificação do sujeito e do objeto, bem como a avaliação do sujeito em relação ao objeto. Por exemplo, no caso de “o atendimento foi excelente”, o objeto é “o atendimento”, e a avaliação é “excelente”.

A Análise de Sentimentos classifica sentimentos em uma das três categorias: positiva, negativa ou neutra. Este trabalho considerou apenas sentimentos positivos e negativos. De acordo com Nascimento *et al.* (2012), não existe um consenso na literatura sobre quais as características típicas de textos classificados como “neutro”. Muitos artigos ignoram a categoria “neutro”, porque é difícil encontrar um comentário que não expresse nenhuma opinião e, mesmo que existam, tornam-se irrelevantes para a análise.

Para Kim *et al.* (2015), Extração de Tópicos e Análise de *Buzz* geralmente estão relacionadas à análise de tendências de mercado, enquanto a Análise de Sentimentos é mais utilizada para análise de reputação de produtos, serviços e empresas e de reações dos consumidores às ações de *marketing*.



Mostafa (2013), após uma revisão bibliográfica, conclui que a maioria das aplicações de Análise de Sentimentos tem sido realizada em quatro áreas distintas: avaliação de produtos, resenhas de filmes, extração de informações para orientação política e previsões do mercado de ações.

2.3 Extração de Palavras-chave

A Extração de Palavras-chave ou Extração de Tópicos é uma técnica que proporciona o conhecimento e a ordenação dos termos ou das frases que aparecem com maior frequência em um conteúdo textual. Os resultados podem ser usados para obter *insights*, ou seja, conhecimento prévio sobre a essência de grandes volumes de informações não estruturadas.

Na busca pelas palavras-chave, pode ser habilitada a opção de procurar por frases ou expressões negativas, como “não gostei”. Essa função é de grande valia para análises nas quais se pretende distinguir sentimentos e opiniões em conteúdos textuais, uma vez que, ao utilizar como base apenas a ocorrência de palavras, poderia haver distorções de sentido e problemas de compreensão, podendo, por exemplo, o “não gostei” ser separado e reconhecido como opinião positiva, “gostei”.

Na Extração de Palavras-chave, são aplicados filtros para descartar palavras e expressões indesejadas ou sem relevância para a análise, conhecidas como “*stop words*”. Os programas de Mineração de Textos, em geral, já possuem um dicionário de *stop words*, mas, em sua maioria, disponibilizam também a opção de inclusão de novos termos.

Neste trabalho, para visualização dos resultados da Extração de Palavras, foi usada a Nuvem de Palavras, conhecida também como Nuvem de *Tags*. Na Nuvem de Palavras, o tamanho de um termo tem relação com sua frequência e, conseqüentemente, com sua relevância no conteúdo de dados textuais, ou seja, quanto maior a palavra, pode-se considerar também maior a sua importância no que se refere ao corpo textual.

3. PESQUISA DE CAMPO UTILIZADA PARA COMPARAÇÃO

O objetivo da pesquisa de campo, encomendada pela rede varejista ao IBOPE², consistiu em obter informações em

2 O IBOPE é uma empresa privada de pesquisa que desenvolve pesquisas de opinião. As atividades do IBOPE Inteligência são regidas por padrões éticos da Associação Brasileira das Empresas de Pesquisa (ABEP) e da *European Society of Market Research* (ESOMAR). <http://www.ibope.com.br/pt-br/ibope/quemosomos/Paginas/default.aspx>

relação às opiniões dos clientes sobre as lojas da rede de comércio varejista, bem como entender os pontos fracos e fortes da empresa.

As investigações para coleta de dados da pesquisa de campo foram realizadas por meio de entrevistas pessoais em três lojas da rede localizadas no Rio de Janeiro. O público-alvo abordado nas entrevistas eram pessoas que estivessem fazendo compras nas lojas selecionadas. A pesquisa foi realizada no período de 15 a 21 de maio de 2014. A amostragem adotada foi a intencional e composta por 216 entrevistas no total. Para amostra de 216 casos, a margem de erro é de 7 pontos percentuais, considerando nível de confiança a 95%.

Para a comparação com os resultados das análises feitas a partir das técnicas de Mineração de Textos, foram consideradas mais relevantes as seguintes questões da pesquisa de campo:

- “Quando você vai escolher um estabelecimento deste segmento de comércio para fazer compras, o que leva em consideração em primeiro lugar? E em segundo lugar? E em terceiro lugar?”;
- “Com base na sua experiência com essa loja como um todo, usando uma escala de 0 a 10, em que 0 significa que você com certeza não recomendaria e 10 significa que você com certeza recomendaria essa loja para amigos ou familiares, qual é a chance de você recomendar esta loja para amigos ou familiares? Você pode escolher qualquer valor entre 0 e 10”.

O quadro 3 (na seção 4) mostra os cinco atributos que foram mais mencionados nas entrevistas em resposta a pergunta: “Quando você vai escolher um estabelecimento deste segmento de comércio para fazer compras, o que leva em consideração em primeiro lugar? E em segundo lugar? E em terceiro lugar?” O “preço” foi considerado o atributo de maior importância na escolha dentre as lojas deste segmento.

Para a análise da pergunta “Você nos recomendaria a um amigo?”, o Ibope aplicou a metodologia NPS - *Net Promoter Score*, utilizada como uma medida que expressa o grau de satisfação do cliente com a empresa.

De acordo com o NPS, o entrevistador faz a pergunta de recomendação e o entrevistado deve dar notas de 0 a 10, sendo que 0 significa que certamente não recomendaria e 10 significa que certamente recomendaria.

Conforme a resposta, cada entrevistado é classificado em um dos três segmentos do quadro 1.



Quadro 1. Classificação pela metodologia NPS

Promotores	=	nota 9-10	→	Tendência a falar bem
Passivos	=	nota 7-8	→	Neutros (tendência a não falar)
Detratores	=	nota 0-6	→	Tendência a falar mal

O índice NPS é definido pela diferença entre Promotores e Detratores

Fonte: Elaborado a partir de IBOPE Inteligência (2014)

A figura 1 apresenta o resultado da compilação das pesquisas realizadas nas três lojas a partir da pergunta “Você nos recomendaria a um amigo?”.



Figura 1. Classificação NPS em relação às notas de avaliação dos clientes

Fonte: Elaborado a partir de IBOPE Inteligência (2014)

4. ANÁLISE DOS DADOS: APLICAÇÃO DAS TÉCNICAS DE MINERAÇÃO DE TEXTOS E COMPARAÇÃO COM A PESQUISA DE CAMPO

Os textos extraídos para a Mineração de Opiniões eram postagens feitas apenas por clientes e foram retirados da *Fan Page* da rede de comércio varejista – comentários e avaliações sobre: serviços, estabelecimento, produtos e funcionários da empresa. Foram retiradas todas as postagens de clientes do início de março ao final de agosto de 2014 para confrontar com a pesquisa de campo feita em maio de 2014. O intervalo de tempo foi necessário para que se chegasse a uma amostragem próxima ao número de entrevistas.

Para analisar o conteúdo textual, foram usadas duas técnicas de Mineração de Textos: Extração de Palavras-chave e Análise de Sentimentos. A Mineração dos Textos foi realizada no programa *PolyAnalyst*, basicamente, a partir dos nós *Keyword Extraction* e *Sentiment Analysis*.

Na Extração de Palavras-chave (figura 2), foram extraídos apenas os substantivos, como forma de levantamento dos assuntos mais relevantes. Os adjetivos foram descartados nessa etapa porque a Análise de Sentimentos, para extrair as avaliações mais relevantes dos clientes, já faz a seleção dos adjetivos mais frequentes.

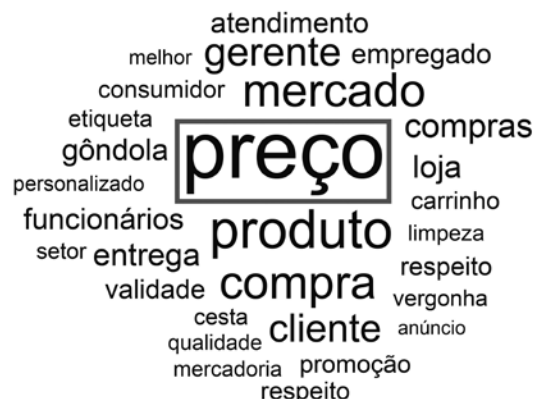


Figura 2. Nuvem de Palavras extraída da *Fan Page* da rede

Fonte: Os próprios autores

Comparando os resultados da Mineração de Textos em conteúdos retirados do *Facebook* da empresa, as maiores queixas dos consumidores foram muito semelhantes aos resultados compilados a partir das respostas dos questionários feitos em três lojas físicas da rede varejista.

Traçando um paralelo entre o quadro 2 da Extração de Palavras-chave da *Fan Page* e o quadro 3 com os atributos que os clientes, entrevistados nas lojas, consideraram mais importantes e que apareceram com maior frequência entre as três primeiras escolhas, percebe-se uma semelhança nos resultados.

Quadro 2. Resultado da Extração de Palavras-chave

Palavras-chave	Relevância
Preço	100%
Produto	63,47%
Mercado	53,28%
Compra	43,98%
Gerente	33,89%
Cliente	30,38%
Entrega	24,58%
Funcionários	23,24%

Fonte: Os próprios autores

Quadro 3. Resultado da pesquisa feita nas lojas

	Mencionados - 1º + 2º + 3º lugar
Preço	85
Variedade de marcas e produtos	64
Localização	54
Ambiente	50
Atendimento	46

Fonte: Elaborado a partir de IBOPE Inteligência (2014)

Na pesquisa feita nas lojas (quadro 3), os clientes, em sua maioria, apontaram o “preço” como fator primordial na es-



colha do estabelecimento para fazer compras. O resultado da Extração de Palavras-chave (quadro 2) mostrou também que o assunto mais mencionado pelos clientes na página do Facebook da empresa era o “preço” dos produtos. Nesta mesma análise da *Fan Page*, “produto” apareceu como segundo assunto mais comentado, e na pesquisa de campo também, só que mais especificamente sobre a variedade de produtos e de marcas. O terceiro e o quarto fatores mais citados na pesquisa feita nas lojas (quadro 3), “localização” e “ambiente”, foram também as observações e as preocupações referentes a “mercado”, terceiro assunto do quadro 2, questões discutidas entre os clientes na rede social da empresa. O “atendimento”, quinto tópico do quadro 3, na prática, engloba do quarto ao oitavo assunto do quadro 2 – “compra”, “gerente”, “cliente”, “entrega” e “funcionários” – todos termos de atividades e atores que integram o atendimento.

É importante frisar que, nas tabelas e nos gráficos resultantes da análise, tanto na Extração de Palavras-chave quanto na Análise de Sentimentos, é possível selecionar cada palavra e visualizá-la na frase. Dessa forma, pode-se verificar sua relevância dentro do contexto e analisar se há concordância com o que o programa está sugerindo.

A Análise de Sentimentos realizada em conteúdo da *Fan Page* da empresa (figura 3), por meio da extração e classificação de comentários positivos e negativos, resultou em uma taxa de aprovação de 53% e de desaprovação de 47%. A pesquisa feita nas lojas, a partir das respostas dos clientes sobre a pergunta “Você nos recomendaria a um amigo?” (figura 4), apontou uma taxa de reprovação de 45%. Comparando os dois métodos de análise de opiniões, a diferença de 2% nos resultados pode ser considerada pequena. Portanto, novamente observa-se uma semelhança nos resultados.



Figura 3. Resultado da Análise de Sentimentos em conteúdo da *Fan Page* da empresa
Fonte: Os próprios autores



Figura 4. Resultado da pesquisa de campo referente à pergunta: “Você nos recomendaria a um amigo?” (extraído da Figura 1 - Classificação NPS)

Fonte: Elaborado a partir de IBOPE Inteligência (2014)

Na Análise de Sentimentos, ainda foi possível selecionar as avaliações que apareceram com mais frequência nos comentários extraídos da *Fan Page* (figura 5). O que apareceu mais em evidência foi o fato de os clientes considerarem os preços dos produtos da rede varejista “muito caros”.



Figura 5. Avaliações que apareceram com mais frequência na Análise de Sentimentos
Fonte: Os próprios autores

5. CONCLUSÃO

Blogs e redes sociais, como *Twitter* e *Facebook*, ocupam atualmente um lugar importante na formação de opinião da sociedade. Dessa forma, são também uma fonte valiosa para Mineração e análise de informações e opiniões sobre diferentes assuntos, como marcas, produtos, serviços e personalidades. Portanto, cada vez mais, o conteúdo gerado espontaneamente por usuários dessas mídias converte-se em objeto de estudo para áreas relacionadas à gestão de relacionamento com clientes, ao controle de opinião pública e às pesquisas de *marketing*. Tanto no campo acadêmico como no mercado, cresce o interesse em pesquisas sobre métodos e técnicas de Mineração de Textos que auxiliem no processo de transformação de um enorme volume de dados textuais não estruturados em informações relevantes para tomadas de decisões e planejamento estratégico.

Diante desse cenário, a presente pesquisa exploratória teve como propósito demonstrar as vantagens da aplicação de técnicas de Mineração de Textos para conhecimento de opiniões de clientes e estudos de percepção e reputação de marcas. A eficácia das técnicas de Mineração de Textos foi comprovada por meio de uma análise comparativa dos resultados obtidos por uma pesquisa de campo feita para uma rede de comércio varejista com os resultados obtidos por uma análise realizada a partir das técnicas de Análise de Sentimentos e Extração de Palavras-chave em conteúdos extraídos do Facebook da rede em questão.

A Análise de Sentimentos e a Extração de Palavras-chave atingiram o objetivo principal: conhecimento das opiniões dos clientes sobre as lojas, sobre o atendimento e sobre os



produtos oferecidos pela rede. Os resultados obtidos por meio da classificação e da ordenação, da mesma forma que as conclusões da pesquisa realizada nas três lojas físicas, além do levantamento dos atributos positivos, apontaram os maiores problemas da empresa percebidos pelos clientes, questões prioritárias para discussões sobre mudança de estratégia e posicionamento de mercado da rede varejista.

O presente trabalho demonstrou conformidade na comparação dos resultados obtidos pela análise realizada a partir de técnicas de Mineração de Textos em conteúdos extraídos da *Fan Page* da rede de comércio varejista com os resultados obtidos pela pesquisa de campo feita para a mesma rede. A pesquisa exploratória realizada na rede social da empresa levantou informações e opiniões de consumidores muito semelhantes às respostas dos questionários da pesquisa de campo.

Os resultados da mineração e da comparação com as conclusões das entrevistas pessoais ratificaram a ideia de que as técnicas de Mineração de Textos apresentam-se como soluções eficientes e vantajosas para conhecimento de opiniões e estudos de percepção e reputação de marcas. Por consequência, corroboraram também as características de praticidade e de acessibilidade em relação a tempo e custo quando comparadas às técnicas tradicionais de pesquisa de mercado.

Apesar de apresentar as vantagens e comprovar a eficiência das técnicas de Mineração de Textos, este trabalho não pretende, de forma alguma, descartar a importância dos métodos tradicionais de pesquisa de mercado. É certo que, para um conhecimento mais objetivo e aprofundado de um determinado assunto ou do que se pretende descobrir, técnicas investigativas tradicionais, como grupos focais e entrevistas presenciais, são mais apropriadas porque permitem a condução do assunto por meio da intervenção do moderador ou do entrevistador.

Como conclusão, recomenda-se o uso das técnicas de Mineração de Textos para coleta e análise de informações em pesquisa de mercado e para análise de opiniões em estudos de percepção de marca, principalmente devido a seus benefícios em relação a tempo e custo. No entanto, propõe-se aqui o uso dessas técnicas para pesquisas exploratórias onde se pretende realizar um estudo preliminar sobre o objeto da pesquisa ou familiarizar-se com o tema que está sendo investigado.

Para trabalhos futuros em que se pretenda aplicar técnicas de Mineração de Textos para extração e análise de opiniões, sugere-se um rastreamento mais amplo na *Web* de comentários sobre a empresa, o produto ou o serviço, como por exemplo, em *sites* de reclamações de consumidores. Além disso, para efetivamente aumentar a vantagem competitiva e avaliar o ambiente de negócios, as empresas precisam monitorar e analisar não só o conteúdo gerado pelo cliente em seus próprios *sites* de mídia social mas também a informação

textual em outros *sites* e em outras redes sociais de seus concorrentes, além de *blogs* e redes sociais que abordem temas relacionados à sua área de atuação no mercado.

REFERÊNCIAS

- Camiciottoli, B.C., Ranfagni, S. & Guercini, S. (2014), "Exploring brand associations: an innovative methodological approach", *European Journal of Marketing*, 48(5), pp.1092–1112, disponível em: <http://www.emeraldinsight.com/10.1108/EJM-12-2011-0770> (Acesso em 12 de junho de 2015).
- Chen, L., Qi, L. & Wang, F. (2012), "Comparison of feature-level learning methods for mining online consumer reviews", *Expert Systems with Applications*, 39(10), pp.9588–9601, disponível em: <http://dx.doi.org/10.1016/j.eswa.2012.02.158> (Acesso em 19 de Junho de 2015).
- Chung, W. & Tseng, T.B. (2012), "Discovering business intelligence from online product reviews: A rule-induction framework", *Expert Systems with Applications*, 39(15), pp.11870–11879, disponível em: <http://dx.doi.org/10.1016/j.eswa.2012.02.059>. (Acesso em 8 de Junho de 2015).
- Gomez, L., Mateus, C., Cardoso, H., & Rosa, C. (2013). "Territorial branding: DNA Alvito". Disponível em: http://www.ideasrevolution.pt/cms/wp-content/uploads/2013/01/IDEASpaper_ALVITO_CHILE.pdf. (Acesso em 20 de junho de 2015).
- He, W., Zha, S. & Li, L. (2013), "Social media competitive analysis and text mining: A case study in the pizza industry", *International Journal of Information Management*, disponível em: <http://doi.org/10.1016/j.ijinfomgt.2013.01.001> (Acesso em 29 de Maio de 2015).
- Hu, N. *et al.* (2012), "Manipulation of online reviews: An analysis of ratings, readability, and sentiments", *Decision Support Systems*, 52(3), pp.674–684, disponível em: <http://dx.doi.org/10.1016/j.dss.2011.11.002> (Acesso em 24 de Julho de 2015).
- IBM SPSS Modeler, disponível em: <http://www-03.ibm.com/software/products/pt/spss-modeler> (Acesso em 20 de agosto de 2015).
- Jain, a. K., Murty, M.N. & Flynn, P.J. (1999), "Data clustering: a review", *ACM Computing Surveys*, 31(3), pp.264–323, disponível em: <http://doi.org/10.1145/331499.331504> (Acesso em 7 de agosto de 2015).
- Jansen, B.J. *et al.* (2009), "Twitter power: Tweets as electronic word of mouth", *Journal of the American Society for Information Science and Technology*, 60(11), pp.2169–2188, disponível em: <http://doi.org/10.1002/asi.21149> (Acesso em 24 de Julho de 2015).
- Khadjeh Nassirtoussi, A. *et al.* (2014), "Text mining for market prediction: A systematic review", *Expert Systems with Applications*, 41(16), pp.7653–7670, disponível em: <http://doi.org/10.1016/j.eswa.2014.06.009> (Acesso em 15 de Maio de 2015).



- Kim, Y. & Jeong, S.R. (2015), "Opinion-Mining Methodology for Social Media Analytics", *KSI Transactions on Internet and Information Systems*, 9(1), pp.391–406
- Leximancer, disponível em: <http://info.leximancer.com/> (Acesso em 20 de agosto de 2015).
- Li, Y. & Li, T. (2013), "Deriving market intelligence from micro-blogs", *Decision Support Systems*, disponível em: <http://dx.doi.org/10.1016/j.dss.2013.01.023> (Acesso em 17 de junho de 2015).
- Liu, H. et al. (2013), "Electronic Commerce Research and Applications Combining user preferences and user opinions for accurate recommendation", *Electronic Commerce Research and Applications*, 12(1), pp.14–23, disponível em: <http://dx.doi.org/10.1016/j.elerap.2012.05.002> (Acesso em 12 de Junho de 2015).
- Loh, S., Wives, L.K. & de Oliveira, J.P.M. (2000), "Concept-based knowledge discovery in texts extracted from the Web", *ACM SIGKDD Explorations Newsletter*, 2(1), pp.29–39, disponível em: <http://dl.acm.org/citation.cfm?id=360414> (Acesso em 20 de Fevereiro de 2015).
- Megaputer - PolyAnalyst, disponível em: <http://www.megaputer.com/site/polyanalyst.php> (Acesso em 20 de agosto de 2015).
- Milolidakis, G. & Kimble, C. (2013), "Digital traces for business intelligence. A case study of mobile telecoms service", *Journal of Enterprise Information Management*, 27(1), pp. 66–98, disponível em: <http://www.emeraldinsight.com/doi/full/10.1108/JEIM-09-2012-0061> (Acesso em 27 de Fevereiro de 2015).
- Mostafa, M.M. (2013), "More than words: Social networks' text mining for consumer brand sentiments", *Expert Systems with Applications*, 40(10), pp.4241–4251, disponível em: <http://dx.doi.org/10.1016/j.eswa.2013.01.019> (Acesso em 20 de Fevereiro de 2015).
- Nascimento, P. et al. (2009), "Análise de sentimento de tweets com foco em notícias", disponível em: <http://www.lbd.dcc.ufmg.br/bdbcomp/servlet/Trabalho?id=11993> (Acesso em 6 de março de 2015).
- Poria, S. et al. (2014), "EmoSenticSpace: A novel framework for affective common-sense reasoning", *Knowledge-Based Systems*, disponível em: <http://dx.doi.org/10.1016/j.knsys.2014.06.011>. (Acesso em 10 de maio de 2015).
- R Project for Statistical Computing, disponível em: <https://www.r-project.org/> (Acesso em 20 de agosto de 2015).
- Ranfagni, S., Guercini, S. & Camiciottoli, B.C. (2014), "An interdisciplinary method for brand association research", *Management Decision*, 52(4), pp.724–736, disponível em: <http://www.emeraldinsight.com/10.1108/MD-04-2012-0284>. (Acesso em 20 de Julho de 2015).
- Reyes, A. & Rosso, P. (2012), "Making objective decisions from subjective data: Detecting irony in customer reviews", *Decision Support Systems*, 53(4), pp.754–760, disponível em: <http://dx.doi.org/10.1016/j.dss.2012.05.027>. (Acesso em 20 de abril de 2015).
- SAS Enterprise Miner, disponível em: http://www.sas.com/en_us/software/analytics/enterprise-miner.html. (Acesso em 20 agosto 2015).
- Seol, H., Lee, S. & Kim, C. (2011), "Identifying new business areas using patent information: A DEA and text mining approach", *Expert Systems with Applications*, 38(4), pp.2933–2941, disponível em: <http://doi.org/10.1016/j.eswa.2010.06.083> (Acesso em 5 de Fevereiro de 2015).
- Sobkowicz, P., Kaschesky, M. & Bouchard, G. (2012), "Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the web", *Government Information Quarterly*, 29(4), pp.470–479, disponível em: <http://dx.doi.org/10.1016/j.giq.2012.06.005> (Acesso em 14 de Março de 2015).
- Thorleuchter, D. & Poel, D. Van Den (2012), "Predicting e-commerce company success by mining the text of its publicly-accessible website", *Expert Systems with Applications*, 39, pp.13026–13034, disponível em: <http://dx.doi.org/10.1016/j.eswa.2012.05.096> (Acesso em 15 de Março de 2015).
- Xu, K., Guo, X., et al. (2012), "Electronic Commerce Research and Applications Discovering target groups in social networking sites: An effective method for maximizing joint influential power", *Electronic Commerce Research and Applications*, 11(4), pp.318–334, disponível em: <http://dx.doi.org/10.1016/j.elerap.2012.01.002> (Acesso em 6 de Junho de 2015).
- Xu, K., Li, J. & Song, Y. (2012), "Identifying valuable customers on social networking sites for profit maximization", *Expert Systems with Applications*, 39(17), pp.13009–13018, disponível em: <http://dx.doi.org/10.1016/j.eswa.2012.05.098> (Acesso em 12 de Julho de 2015).
- Yu, Y., Duan, W. & Cao, Q. (2013), "The impact of social and conventional media on firm equity value: A sentiment analysis approach", *Decision Support Systems*, disponível em: <http://dx.doi.org/10.1016/j.dss.2012.12.028> (Acesso em 3 de Fevereiro de 2015).
- Zerr, S. et al. (2013), "Sentiment Analysis and Opinion Mining in Collections of Qualitative Data", http://www.arcomem.eu/wp-content/uploads/2013/08/ipres_gute_arbeit.pdf (Acesso em 11 de Abril de 2015).
- Zhang, L. & Liu, B. (2014), "Aspect and entity extraction for opinion mining", *Data Mining and Knowledge Discovery for Big Data*, pp.1–40, disponível em: <http://doi.org/10.1007/978-3-642-40837-3> (Acesso em 14 de Maio de 2015).
- Zubrinic, K., Kalpic, D. & Milicevic, M. (2012), "The automatic creation of concept maps from documents written using morphologically rich languages", *Expert Systems with Applications*, 39(16), pp.12709–12718, disponível em: <http://dx.doi.org/10.1016/j.eswa.2012.04.065>. (Acesso em 11 de Abril de 2015).