

XXV Workshop de Iniciação Científica da Embrapa Gado de Leite

Juiz de Fora, MG – 5 de março de 2020

Análise dos sentimentos expressos na rede social Twitter em relação ao queijo¹

Emerson Campos², Nedson D. Soares², Thallys Nogueira³, Kennya Siqueira³, Emerson Moraes^{4,5}

¹Este trabalho foi apresentado com o apoio do CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil. Parte do projeto/ação gerencial “Residência Zootécnica Digital”.

²Bolsista Embrapa Gado de Leite e-mail: emersonwendelim@gmail.com; thallysnogueira@ice.ufjf.br; nedson@ice.ufjf.br

³Pesquisadora – Embrapa Gado de Leite – Juiz de Fora, MG. e-mail: kennya.siqueira@embrapa.br

⁴Professor – Núcleo de Informática do Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas Gerais – Juiz de Fora, MG. e-mail: emerson.moraes@ifsudestemg.edu.br

⁵Orientador

Resumo: É notório que o queijo é um produto alimentício que cada vez mais está presente na mesa dos brasileiros. A diversidade de sabores é um dos fatores que estimularam o consumo. No entanto, pouco se sabe sobre a opinião e satisfação do consumidor brasileiro de queijo. Diante disso, este artigo objetiva avaliar a satisfação dos usuários do Twitter via análise de sentimentos quanto a expressões relacionadas ao assunto queijo. No mês de fevereiro de 2020, foram coletadas informações que contribuíram para a construção e análise de uma base de dados composta por 6.500 *twittes* em arquivo de texto provenientes de todo o Brasil. A base de dados contém informações sobre os usuários e as respectivas mensagens postadas relacionadas à palavra queijo. Utilizando técnicas de inteligência artificial foi possível observar a aceitação dos usuários sobre o queijo, de um modo geral, como também deixar evidente a clara predileção pelo tradicional pão de queijo.

Palavras-chave: consumo, queijos, redes sociais, twitter, aceitação

Analysis of the feelings expressed on the social network Twitter regarding cheese

Abstract: Cheese is well known as a food product that is more and more available in the Brazilians meals. The diversity of flavors is one of the factors that stimulated its consumption. However, little is known about the opinion and satisfaction of the Brazilian cheese consumer. Therefore, this article aims to assess the satisfaction of Twitter users via sentiment analysis regarding cheese. In February 2020, 6500 tweets were collected in text files from all over Brazil. The database contains information about users and their respective posted messages related to the word cheese. Using artificial intelligence techniques, it was possible to observe the users' acceptance regarding cheese, as well as to preference for the traditional cheese bread.

Keywords: consumption, cheese, social networks, twitter, acceptance

Introdução

O queijo é um dos derivados lácteos mais consumidos no mundo. Importante fonte de proteínas de alta digestibilidade, os queijos também são ricos em nutrientes. Devido as suas grandes vantagens nutricionais, o queijo possui grande importância nos hábitos de consumo da população brasileira, sendo parte integrante da dieta nacional.

Graças à popularização da internet, o consumidor se encontra cada vez mais informado, mais exigente e mais consciente a respeito dos diversos tipos de queijos presentes no mercado.

XXV Workshop de Iniciação Científica da Embrapa Gado de Leite
Juiz de Fora, MG – 5 de março de 2020

O acesso a essas informações contribui de fato para tal conhecimento e a ascensão das redes sociais possibilitou ao mercado novos modelos de negócios baseados em análise de dados extraídos das redes sociais. Portanto, as redes sociais tem se tornado importante para as grandes empresas devido ao enorme fluxo de dados trocados diariamente, o que facilita a compreensão sobre sentimentos e opiniões expressos por seus usuários.

Estas postagens são utilizadas para descrever as experiências pessoais e memórias que os usuários tiveram na utilização de um dado produto ou serviço, fornecendo às empresas uma base de dados importante para a tomada de decisão. Porém, nem sempre esses comentários são de todo favoráveis à empresa, o que pode comprometer a sua reputação e sobrevivência no mercado competitivo (TARANNUM *et. al.*, 2015).

Uma plataforma popular nos dias de hoje é o Twitter que, sozinho, possui mais de 200 milhões de usuários, que compartilham cerca de 400 milhões de *tweets* por dia (CHA *et. al.*, 2010). Nessa rede, pesquisadores e empresas conseguem coletar dados para análises de conteúdo em grande escala (CELEBRATING, 2020). Por meio desses dados, pode-se acompanhar um determinado fluxo de *posts* para identificar oportunidades de vendas de um produto baseados nas opiniões de seus consumidores.

Métodos para análise de sentimentos estão cada vez mais populares visando desenvolver aplicações que tem como objetivo apresentar informações sobre as necessidades de diversos tipos de consumidores no mercado. Porém, no segmento de queijos do Brasil não se tem conhecimento de pesquisas que se aprofundam em avaliar a opinião dos usuários do Twitter. Portanto, nesse trabalho, objetiva-se elucidar esse contexto, aplicando técnicas de inteligência artificial para analisar os sentimentos de raiva, desgosto, medo, alegria, tristeza e surpresa dos usuários do Twitter em relação ao queijo.

Material e Métodos

Para a realização do estudo, fez se o uso do *software* Orange Canvas. O Orange Canvas, é um *software* desenvolvido na Linguagem Phyton para visualização e análise de dados. Após a coleta das informações, foi realizado um pré-processamento desses dados para melhor seleciona-los, eliminando nessa etapa informações irrelevantes para o estudo. Em seguida, as informações obtidas foram mineradas com o uso de algoritmos de agrupamento, criando clusterização das mensagens de texto em grupos específicos por ordem de relevância ao que cada usuário escreveu. Com as informações tratadas, fez se o uso de padrões de filtragens para melhor compreender e visualizar o sentimento dos usuários quando se referiam a assuntos relacionados ao tema queijo.

Resultados e Discussão

A pesquisa analisou informações expressas em mensagens de texto no Twitter sobre a palavra queijo, com a finalidade de descobrir a existência de algum padrão entre as informações contidas no mesmo. A aplicação Orange Canvas nos permite também trazer outras análises, como o uso do *widget* WordCloud, como já descrito, para visualização de dados. A WordCloud

XXV Workshop de Iniciação Científica da Embrapa Gado de Leite
Juiz de Fora, MG – 5 de março de 2020

ou Nuvem de Palavras é formada com base na periodicidade das palavras no conjunto de *tweets* que após processada nos mostra a maior incidência delas escritas pelos usuários.

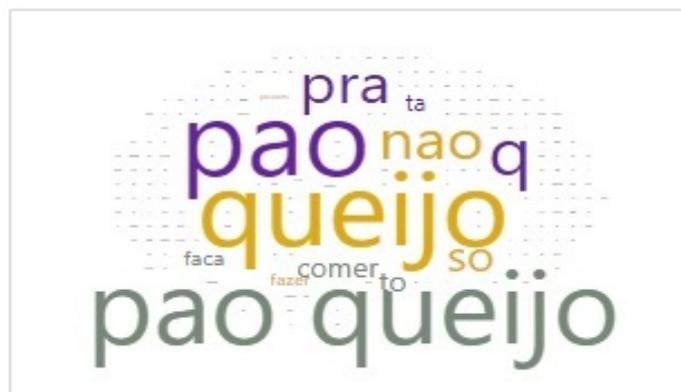


Figura 1. Nuvem de palavras apresentando os itens que mais se relacionaram com o tema queijo no Twitter.

Quando analisados os termos mais citados dentre os *tweets*, percebem-se quais são as palavras que estabelecem maior conexão com o assunto, tais como pão de queijo. Também é possível apresentar a análise de sentimentos dos *tweets*, baseado na busca pelo reconhecimento do sentimento do usuário em concordância com palavras usadas pelo mesmo. Para esse estudo foram classificadas seis diferentes categorias: raiva, desgosto, tristeza e medo como polaridade negativa. Alegria e surpresa como polaridade positiva.

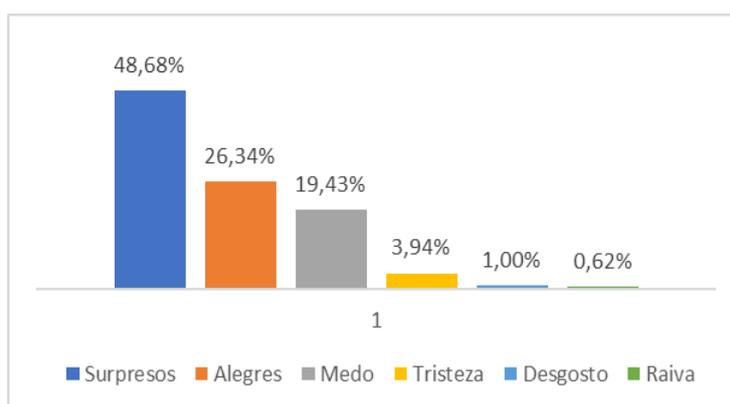


Figura 2. Análise de sentimento dos usuários do Twitter sobre queijo.

Os resultados indicam que a maior parte dos usuários que escreveram sobre queijo no Twitter demonstraram sentimentos positivos (75%). Cabe salientar que resultados pontuais podem ser afetados por recursos computacionais e o seu poder de processamento das informações. Tais resultados devem ser analisados apenas como um indicador de emoções em constante transição e não como parâmetro decisivo sobre o tema estudado.

XXV Workshop de Iniciação Científica da Embrapa Gado de Leite
Juiz de Fora, MG – 5 de março de 2020

Conclusões

Neste trabalho foi feita a análise de sentimentos dos usuários do Twitter quando pesquisados pelo tema queijo. Como conclusão foi possível identificar a importância da mineração de dados junto ao mercado atual no apoio à tomada de decisão. A análise observou grande aceitação dos usuários quando se referiam a palavra queijo, porém é importante ressaltar que as informações pós processadas não só remetiam propriamente ao queijo como um todo, mas também como componente base para outro alimento de grande destaque pelos usuários do Twitter, como o pão de queijo.

Como trabalho futuro, sugere-se uma análise mais aprofundada utilizando outras técnicas de mineração, com amostras maiores de dados afim de poder comparar e discutir os métodos mais adequados para evidenciar o tema abordado neste estudo.

Agradecimentos

Ao Programa Residência Zootécnica Digital da Embrapa Gado de Leite pela concessão da bolsa.

Referências

CELEBRATING. **Blog** **Twitter** 2020. Available at: <https://blog.twitter.com/official/en_us/a/2013/celebrating-twitter7.html>. Accessed on: Feb 28 2020.

TARANNUM, N., RIZVI, S.H., KEOLE, R.R. A Preliminary Review of Web-Page Recommendation in Information Retrieval Using Domain Knowledge and Web Usage Mining. **International Journal of Advance Research in Computer Science and Management Studies**. Volume 3, Issue 1. ISSN: 2321 – 7782. Research Article / Survey Paper / Case Study, 2015. Disponível em: www.ijarcsms.com. Acesso em: 17 fev. 2020.

M. Cha, H. Haddadi, F. Benevenuto, and K. P. Gummadi. **Measuring User Influence in Twitter: The Million Follower Fallacy**. In Int'l AAAI Conference on Weblogs and Social Media (ICWSM), 2010