

## Impacto da inflação nos *tweets* sobre leite e derivados<sup>1</sup>

Anna Leticia Franco Monteiro <sup>2</sup>

Ygor Martins Guimarães <sup>3</sup>

Darlan Henrique Da Costa Silva <sup>4</sup>

Thallys da Silva Nogueira <sup>5</sup>

Kennya Beatriz Siqueira <sup>6</sup>

Priscila Vanessa Zabala Capriles Goliatt <sup>7,8</sup>

**Resumo:** Atualmente o Twitter tem mais de 200 milhões de contas ativas. Como um ambiente com grande volume de informação, ele é propício à mineração de dados para fornecimento de informações importantes sobre tendências de mercado. O Observatório do Consumidor (OC) armazena todos os conteúdos em que são citados os derivados lácteos e os analisa de diferentes formas. Uma delas é a análise de sentimentos. A análise de sentimento é feita através da técnica de aprendizado de máquina supervisionado, em que é necessária a preparação de um conjunto de dados rotulados para que o algoritmo seja treinado a partir deles. Dessa forma, é possível que os *tweets* sejam classificados entre negativos, neutros e positivos. O presente estudo buscou identificar o impacto da inflação no sentimento negativo nos *tweets* sobre leites e derivados, relacionando a quantidade desses *tweets* com a inflação. Notou-se que a inflação não é um fator que altera o sentimento negativo do consumidor de lácteos.

**Palavras-chave:** lácteos, consumidor, polaridade, análise de sentimento, preços

### Impact of inflation on *tweets* about dairy products

**Abstract:** Currently, Twitter has over 200 million active accounts. As an information-intensive environment, it is conducive to data mining to provide important insights into market trends. The Observatório do Consumidor (OC) stores all the content about dairy products and analyzes them in different ways. One of them is sentiment analysis. Sentiment analysis is done through the supervised machine learning technique, in which it is necessary to prepare a set of labeled data so that the algorithm can be trained from them. In this way, it is possible to be classify *tweets* between negative, neutral, and positive content. The present study sought to identify the impact of inflation on negative sentiment in *tweets* about dairy products, relating the amount of these *tweets* to inflation. It was noted that inflation is not a factor that impacts the negative sentiment of dairy consumers.

**Keywords:** dairy, consumer, polarity, sentiment analysis, price

### Introdução

O Observatório do Consumidor (OC), é uma inovação da Empresa Brasileira de Pesquisa

<sup>1</sup> O presente trabalho foi realizado com o apoio do CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brasil: (a) Parte da tese de doutorado do quarto autor, financiada pelo CNPq, (b) Bolsista do CNPq pelo programa de Mestrado e Doutorado em Inovação MAI/DAI.

<sup>2</sup> Graduanda em Engenharia Computacional-UFJF. e-mail: anna.franco@estudante.ufff.br

<sup>3</sup> Graduando em Ciências Econômicas-UFJF. e-mail: ygor\_martinsg@hotmail.com

<sup>4</sup> Graduando em Sistema de Informação-UFJF. e-mail: darlan.silva@ice.ufff.br

<sup>5</sup> Doutorando em Modelagem Computacional, Programa de Pós Graduação em Modelagem Computacional-UFJF/Minas Gerais. e-mail: thallys.nogueira@estudante.ufff.br

<sup>6</sup> Pesquisadora, Embrapa Gado de Leite - Juiz de Fora, MG. e-mail: kennya.siqueira@embrapa.br

<sup>7</sup> Professora-orientadora, Programa de Pós Graduação em Modelagem Computacional-UFJF. e-mail: capriles@ice.ufff.br

<sup>8</sup> Orientadora

Agropecuária - Embrapa Gado de Leite, em parceria com a Universidade Federal de Juiz de Fora (UFJF) e o Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas - Campus Juiz de Fora (IF-Sudeste/MG), criada para ser uma alternativa aos métodos tradicionais de pesquisa de mercado. Baseado na coleta, processamento e análise de dados de redes sociais, o OC usa tecnologias como inteligência artificial, mineração de dados, web-semântica e processamento de linguagem natural para identificar o perfil e as tendências de consumo dos consumidores de lácteos.

Atualmente o OC analisa o conteúdo do Twitter. Após a coleta de todas as postagens sobre leite e derivados, é realizada a análise de sentimentos. Essa análise é feita através do uso de técnicas de inteligência artificial que buscam identificar a polaridade do sentimento que o consumidor emite na frase escrita, podendo ser negativa, neutra ou positiva. Liu (2012) descreve a análise de sentimentos como uma área de estudo que busca entender e analisar as opiniões expressas pelas pessoas sobre um determinado assunto. Dessa forma, com o uso da análise de sentimentos é possível identificar se os consumidores estão satisfeitos ou não com o produto.

Nesse sentido, a hipótese principal do trabalho é que variáveis econômicas que afetam a experiência do consumidor, podem causar impacto na análise de sentimentos, como por exemplo: o desemprego, a renda e a inflação. O recente aumento generalizado dos preços na economia brasileira, sem a contrapartida da renda, gera perda de poder de compra dos indivíduos, o que tende a desencadear em um sentimento negativo do consumidor.

Diante disso, o presente estudo teve a finalidade de analisar se em períodos em que a inflação está alta houve impacto no sentimento geral dos consumidores dos produtos lácteos e como a percepção deles varia em contextos econômicos diversos.

## Material e Métodos

A análise de sentimentos tem como objetivo rotular cada publicação do Twitter em 3 classes: negativa, neutra e positiva. Cada *tweet* passa por uma etapa de tratamento em que são removidas as *stopwords*, as url's de sites, os caracteres especiais, os sinais de pontuação e acentuação com o objetivo de padronizar os textos e facilitar o processamento e classificação dos dados (NOGUEIRA, 2021).

O tratamento dos dados foi realizado com auxílio da biblioteca NLTK (*Natural Language Toolkit*) (BIRD *et al.*, 2019). Já a construção do modelo de análise de sentimentos foi realizada com a utilização da biblioteca *scikit-learn* (PEDREGOSA *et al.*, 2011). Ambas as bibliotecas foram manipuladas por meio da linguagem de programação *Python* (VAN ROSSUM *et al.*, 2009).

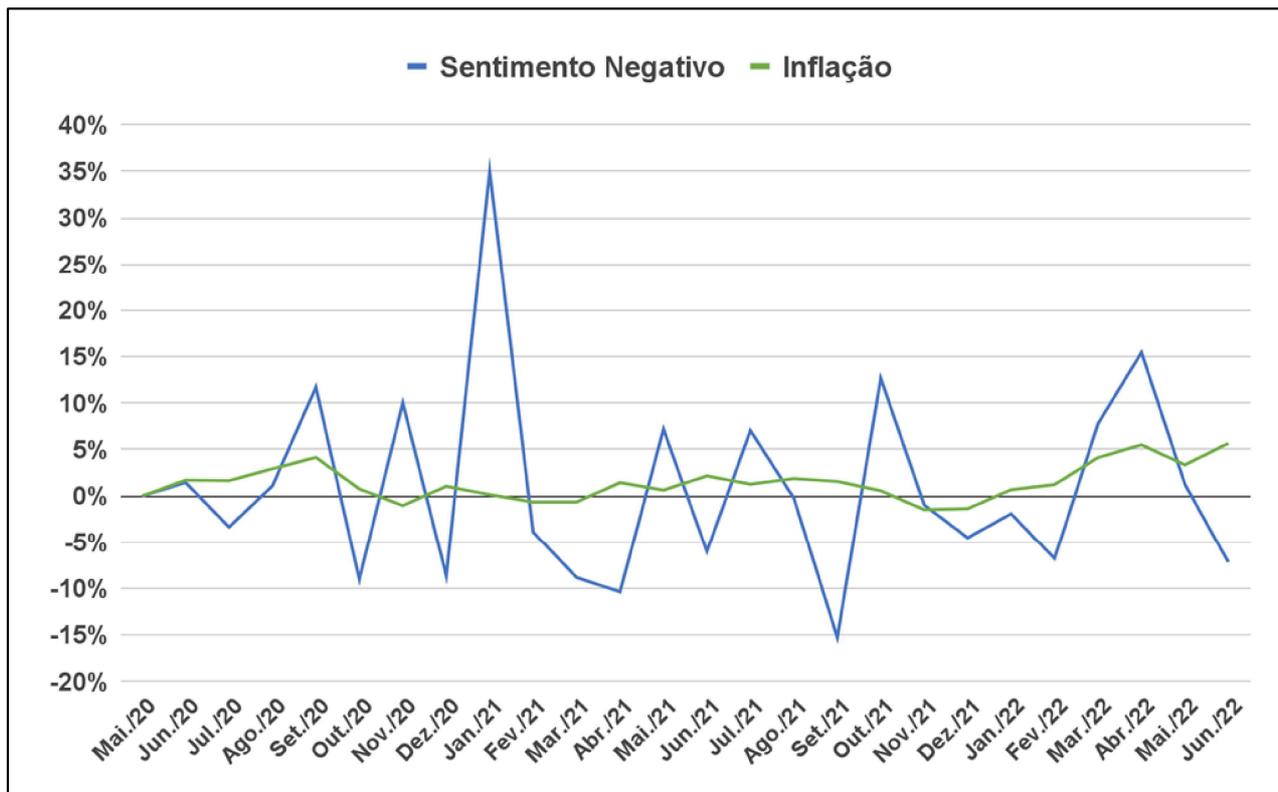
Após a realização da análise de sentimentos, selecionou-se todas as publicações que foram classificadas com o sentimento negativo para análise. Para a inflação, empregou-se os dados do Índice de Preços ao Consumidor Amplo (IPCA) do IBGE (2022).

## Resultados e Discussão

Para avaliar o impacto da inflação sobre o sentimento expresso nos *tweets* sobre leite e derivados, optou-se por usar a variação da porcentagem de postagens com conteúdo negativo (Figura 1).

---

<sup>1</sup>Palavras consideradas "irrelevantes" em um determinado texto.



**Figura 1.** Variação da porcentagem de tweets com sentimento negativo e a inflação.

Fonte: resultados da pesquisa.

De acordo com a Figura 1, não há relação direta entre a variação de preços e o sentimento negativo manifestado nos tweets sobre leite e derivados. Portanto, não se pode afirmar que a inflação afeta o sentimento negativo do consumidor. Apesar da variável preço afetar o consumo de lácteos, o estudo mostrou que esta variável não influencia o sentimento manifestado pelo consumidor. Esse resultado indica que a análise de sentimento reflete mais uma análise de preferências do consumidor do que uma análise da quantidade consumida.

## Conclusões

Os resultados indicaram que a inflação, apesar de ser uma variável importante no consumo dos lácteos, não afeta o sentimento do consumidor em relação ao mesmo. Isso mostrou, portanto, que a análise de sentimento reflete as preferências dos consumidores. Assim, sugere-se que futuros trabalhos analisem outras variáveis, como as demográficas por exemplo, para avaliar melhor os impactos sobre as preferências do consumidor manifestadas na análise de sentimento.

## Agradecimentos

Agradecemos à Universidade Federal de Juiz de Fora e ao Grupo de Modelagem Computacional Aplicada pela concessão da bolsa.

## Referências

BIRD, S., KLEIN, E., LOPER, E. **Natural language processing with Python: analyzing text with the natural language toolkit.** O'Reilly Media, Inc. 2019.

IBGE/SIDRA. Disponível em: <<https://sidra.ibge.gov.br/pesquisa/snipc/ipca/>>. Acesso em: jun. 17, 2022.

LIU, B. Sentiment analysis and opinion mining.: Synthesis lectures on human language technologies. **Morgan Claypool Publishers**. v. 5, p. 1–167, 2012.

NOGUEIRA, T. S. **Mineração de dados em rede social para avaliação de tendências de consumo do queijo artesanal no Brasil**. Juiz de Fora, MG: UFJF, 2021. 112 f. Dissertação (Mestrado em Modelagem Computacional) - Universidade Federal de Juiz de Fora, Juiz de Fora.

PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O. Scikit-learn: Machine learning in python. **Journal of Machine Learning Research**. p. 2825–2830., 2011.

VAN ROSSUM, G., DRAKE, F. L. **Python 3 Reference Manual**. **Scotts Valley, CA: CreateSpace**; 2009.