Check for updates

Natural Language Processing of Messages from a Social Network for Inflation Analysis

Leonardo Silva Vianna(✉), Lizyane Rosa Antunes, Rafael Maia Pinto, and Alexandre Leopoldo Gonçalves

Graduate Program in Knowledge Engineering and Management, Federal University of Santa Catarina, Florianópolis, Brazil
vianna@rocketmail.com

Abstract. Inflation is a progressive increase in the average price of goods and services. It can be measured using different inflation rates, which can vary according to the time, but also according to the region or country. Depending on the variation of product and service prices, users of social networks could discuss and express their private opinions about inflation. We questioned if there is a correlation between the messages transmitted by Twitter© users and the monthly variation of a specific inflation rate. Consequently, this research aims to examine Twitter© messages with content about inflation in Brazil through natural language processing and network analysis. The Twitter© messages from users in Brazil, obtained through the Application Programming Interface of this social network platform, were analyzed. The steps performed included querying the API for data acquisition, processing the messages using Natural Language Processing techniques, and executing a network analysis. We concluded that inflation influences the behavior of social network users and, additionally, natural language processing of Twitter© messages can reveal relevant knowledge for inflation analysis and have potential for prediction purposes.

Keywords: Inflation · Natural language processing · Network analysis

1 Introduction

Inflation is a progressive increase in the average price of goods and services. It can be measured using different inflation rates, which can vary according to the time, but also according to the region or country [1]. Although most economists consider that some inflation is beneficial, it can cause several harmful effects on the economy when it gets out of control. Among these effects, there is the loss of purchasing power, damage in the long-term planning, and the higher cost of public debt, compensating creditors for the value loss of money over time [2].

In Brazil, one of the main objectives of the Central Bank is to maintain inflation under control [3]. In this context, the agency acts through the monetary policy and its efforts are mainly focused in controlling the amount of currency circulating in the economy. This

A. L. Pinto and R. Arencibia-Jorge (Eds.): DIONE 2022, LNICST 452, pp. 218–229, 2022.
https://doi.org/10.1007/978-3-031-22324-2_17

action can generate direct impacts on interactions between economic agents, particularly families, companies, and government. The monetary policy actions result in an effect on market expectations, both short and long term [2]. The standard economic theory state that a tightening in the monetary policy can reduce inflation most of the time, but only for a moment and after a period of time. Pursuing this objective by increasing the public bonds rates leads to a decrease in consumption and, consequently, to a reduction in inflation [4]. Therefore, inflation is a key factor associated with monetary policy, and its control and management require considerable efforts from all related stakeholders.

There are several indexes used to measure the speed at which overall prices change. The national consumer price index (INPC) and the broad national consumer price index (IPCA) – in Portuguese, respectively, *Índice Nacional de Preços ao Consumidor* e *Índice Nacional de Preços ao Consumidor Amplo* – are quantified by the Instituto Brasileiro de Geografia e Estatística (IBGE) in the main metropolitan regions. Likewise, Fundação Getúlio Vargas (FGV) analyzes the prices of different economic sectors (such as the agricultural and industrial sectors), as also collects data in some capitals of the country to generate different general price indexes. Finally, the Fundação Instituto de Pesquisas Econômicas (FIPE) analyzes prices in the city of São Paulo to build a consumer price index [5].

Depending on the variation of product and service prices, users of social networks could discuss and express their private opinions about inflation. Thus, we questioned if there is a correlation between the messages transmitted by Twitter© users and the monthly variation of inflation measured through the IPCA. Consequently, this research aims to examine Twitter© messages with content about inflation in Brazil, with the purpose to identify a correlation between messages and inflation fluctuation.

The evaluation of the aspects related to the price fluctuation linked to inflation could provide understanding of the behavior change in social network users. This knowledge would support data modeling used to develop predictions in different applications, affecting the predictability of the economy behavior. Using data models could also help to improve economic projections, traditionally made by evaluating expert opinion through a heuristic approach.

2 Related Research

Aromi and Martin [6] analyzed the existence of information in Twitter messages that could be applied to predict inflation in Argentina, considering the exchange rate variation between the Argentine peso and the US dollar. For this purpose, they used a corpus of Twitter© messages (tweets) from users located in Argentina to produce an attention indicator that calculated the frequency of the words 'inflation' and 'inflationary' in a specific month, providing a monthly variation rate. The results obtained from a test dataset demonstrated that the models had greater accuracy than a baseline, supporting the claim that social media content could provide valuable information about future inflation levels.

To analyze how the investor behavior characteristics obtained from Twitter© impact the gold price, Kumar, Rao and Srivastava [7] extracted and processed sentiments from tweets and used them for prediction. Likewise, they also applied data from Google

Insights for Search© to compare predictability with macroeconomic analysis and other factors related to economic stability and indicators of investor fear. Using the data from Twitter©, the sentiments were weekly computed by a Bayesian classifier to calculate a measure of optimism, applying a correlation analysis with the gold price at different lag intervals.

According to Alex [8], Twitter© messages are dynamic and reflect the opinion of a large sample of the population, allowing to extract useful information about the expected inflation rate. In his research, the author selected some relevant keywords to identify tweets related to current and expected prices of goods and services. Next, he applied the Latent Dirichlet Allocation (LDA) algorithm and a dictionary of bigrams and trigrams to filter the data obtained from Twitter©. Finally, the author computed the daily number of tweets that represented the increase and decrease of inflation expectations, considering them as directional indicators of the inflationary trend.

Kim, Cha and Lee [9] collected real-time information from Twitter© messages, with an interval of a single day, for evaluation of the daily prediction of price fluctuation of some foods in Indonesia, justified by the fact that governments of developing countries act on food price fluctuations. The authors also evaluated if an additional timestamp for modeling the data led to better results and found that the model was more accurate with a single day lapse. The results obtained by the authors brought to the conclusion that the proposed model accurately predicted food prices and is also resilient, despite the small amount of data. They also stated that the model obtained great short-term accuracy, as there was a tendency for mentions in tweets to increase on the days of the greatest changes in food prices.

Bastida et al. [10] used a hybrid method to analyze the regional Gross Domestic Product (GDP) prediction, applying autoencoder algorithms to learn an intermediate representation associated with tweets and GDP, through three deep learning neural networks: Vanilla autoencoder, convolutional and recurrent. Then, the authors applied Support Vector Regression (SVR) and Random Forest Regressor algorithms in the intermediate representation for GDP prediction. The results of the hybrid method proposed by the authors surpassed those obtained by a baseline (persistence model), a linear regression only on GDP data, or the application of only one autoencoder algorithm.

Techniques for text sentiment analysis were designed primarily for the social media domain, but not for the Economy field. Shapiro, Sudhof and Wilson [11] developed a model for economic sentiment analysis, considering its temporal aspect. The analysis of economic and financial journal articles allowed building a model based on lexical sentiment analysis, combined with some existing lexical models. The authors also performed predictions to estimate the impulse response in some macroeconomic indicators. As a result, they concluded that positive sentiment was related to the increase in consumption, output, and interest rates, as also to the decrease in inflation.

Although the behavior of users in social networks has been analyzed and compared with different inflation rates, to the best of our knowledge, studies about inflation that performed natural language processing to obtain a corpus of text and, thereafter, fulfilled a network analysis have not been published yet.

3 Method

This research consists of the analysis of Twitter© messages (tweets) from users in Brazil, obtained through the Application Programming Interface (API) of this social network platform. The steps performed included querying the API for data acquisition, processing the messages using Natural Language Processing techniques, and executing a network analysis.

The research period was determined by the existence of messages on Twitter© that contained the search terms established for extracting the messages. Although message extraction was defined on the first date available in the API (October 1, 2006), the first extracted tweet had the creation date on October 12, 2010, as shown in the Result Analysis section. Consequently, the first year considered for analysis in the research was 2011.

3.1 Data Acquisition

Tweets were extracted through the Twitter© API using the 'id', 'text', 'created_at', and 'geo.place_id' response fields, as shown in Table 1. A query was built to obtain the messages through the interface, using the standardized operators of the second version of the Twitter© API (Twitter API v2).

Table 1. Response fields applied in the Twitter© API with the data used in the research.

Field	Type	Description
id	String	Unique identifier of each tweet
text	String	Text with the content of the posted tweet
created_at	Date	Tweet creation date in ISO 8601 standard
geo.place_id	String	Tweet postage location identifier, if cataloged by Twitter (despite being collected, the information was not applied in the present research)

The query process used to obtain the tweets was executed in Python language, as also other processes performed in the present research. The search terms instructed in the query were: 'inflacao lang:pt place_country:BR'. Using these settings, the query obtained tweets that contained the term 'inflation' and extracted only messages in Portuguese posted by users located in Brazil. Procedures to query the interface also required to set the start and end dates for data collection, which were repeated on a daily basis. The dataset obtained was stored in a standard comma-separated values (CSV) file for further processing, containing the fields presented in Table 1.

In addition, the IPCA index (considered the official inflation index by the federal government) was acquired from the website of the IBGE [12]. A file containing a historical series with the monthly variation of the IPCA between July 1994 and September 2021 was obtained from the website, further processed and stored in a suitable format for use in the research. In this process, the data with a monthly frequency were restricted to the defined study period.

Thereafter, graphs were generated to assist the interpretation of results obtained with the search conducted in the Twitter© API. These charts provided the comparison between the number of tweets obtained and the IPCA measured in each month, as shown in the Result Analysis section.

3.2 Natural Language Processing

Different techniques were applied for text preprocessing with the purpose to suppress some expressions commonly found in tweets and recurrent in several messages. Therefore, expressions identifying names of users mentioned in the text (preceded by the character '@'), links to other websites (URLs), reposted message tags (retweets), and line break marks (line feeds) were removed.

In the next step, the spaCy© library, available in Python©, was used to construct a corpus of texts. A set of the standard library components was separately performed to each message. The components included functions used to remove stop words and to apply a lemmatization process on the obtained terms, with the objective of inflecting the different forms of each word. In addition to the set of stop words provided by the library, other empty words identified were also suppressed after text processing. Then, the obtained terms were filtered again to eliminate punctuation and spaces, as also keeping only those with alphabetic characters. Finally, the corpus was also normalized to lowercase characters without accents (frequent in the Portuguese language) and then appended to a dataset.

Finally, bigrams of each message were obtained from the processed corpus and stored in a CSV file. The bigrams comprised the connected words from each message, combining a first word with its successor. The file also contained timestamps of the tweet creation dates (field 'created_at'). A structure with the 'source', 'target', and 'time' fields was specifically defined in the file for the network analysis.

3.3 Network Analysis

The Gephi© program was used for the network analysis of the corpus obtained by the natural language processing of the tweets. The previously stored bigrams were interpreted as nodes and edges of a network by the program, additionally with the provided timestamp. When processing this data, the program calculated the weight of each edge, which is the number of existing connections between all the terms of the corpus. Because data only contained bidirectional connections, the program defined the edges as non-directed.

Due to the substantial number of terms and aiming to allow an adequate interpretation of the corpus, only the most significant nodes were used for the network construction. The determined baseline was 5% of the weight of the edge with the highest value – ('inflation', 'high'), with a calculated weight of 2,576. Consequently, only edges with a weight equal to or greater than 128.8 were kept for analysis. Therefore, the edges were grouped by year, enabling an analysis of the temporal evolution of the tweets and identifying the most relevant terms in the posted messages.

Gephi© automatically calculated the centrality measures for the network of words obtained for each year (from 2011 to 2020), which were stored in CSV files. The betweenness centralities were used to analyze the temporal evolution of the network [13], considering only nodes with values greater than zero. A table was built with the concatenated values to present the results and an interactive graph was produced.

Although the terms in the messages posted in 2021 were not used for the analysis of the temporal evolution, due to the lack of complete data at the moment it was extracted, the tweets posted in this year were used for the construction of a graph of words. Therefore, this chart was applied for a non-temporal analysis of the term correlations.

4 Results Analysis

The messages extracting process through the Twitter© API resulted in 60,292 tweets, as shown in Table 2. Figure 1, in turn, presents the number of tweets grouped by the day that the social network users posted the message, according to the search terms applied in the query executed in the interface. It is important to emphasize that this graph contains the results of the entire message acquisition process, including the first messages posted in 2010.

Table 2. Summary of message extraction results through the Twitter© API.

Data characteristic	Result
Number of tweets	60,292
Days with tweets	3,886
First day	2010-10-12
Last day	2021-09-30
Unique geolocations	2,066

Figure 1 represents the number of tweets related to price increases over the months, since 2011. The highest peak of tweets happened in October 2014, the year of federal elections in Brazil. Specifically in these elections, there was a second round of voting on the last Sunday in October. Therefore, possibly, the theme "inflation" was a significant factor of political discussion in electoral campaigns.

The joint representation of the tweets and the inflation rates that actually occurred, expressed in Fig. 2, allows a cross-examination of the issue. In October 2014, the inflation rate was not significantly higher than the average. However, in the months following the elections, resulting in the reelection of the president, inflation rates rose considerably, which may suggest that there was some form of price control by the government for electoral purposes. Nevertheless, it can be noticed that tweets about inflation decreased, because the peak of discussions in October 2014 was not surpassed in other moments. Although, in subsequent months, the average of tweets related to inflation grew when compared to previous periods, which suggests a greater concern of society with post-2014 election inflation.

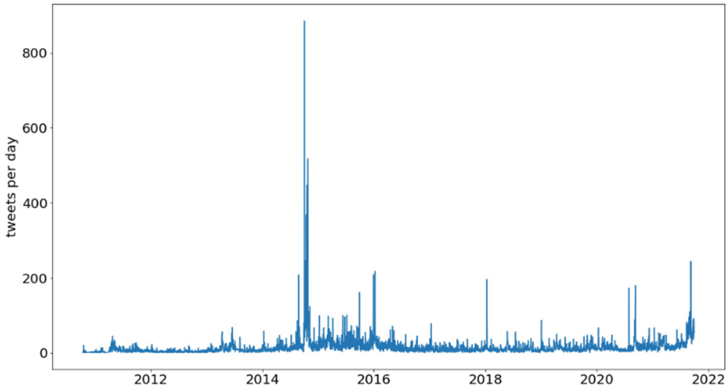

Fig. 1. Number of tweets per day, obtained by the query executed in the Twitter© API.

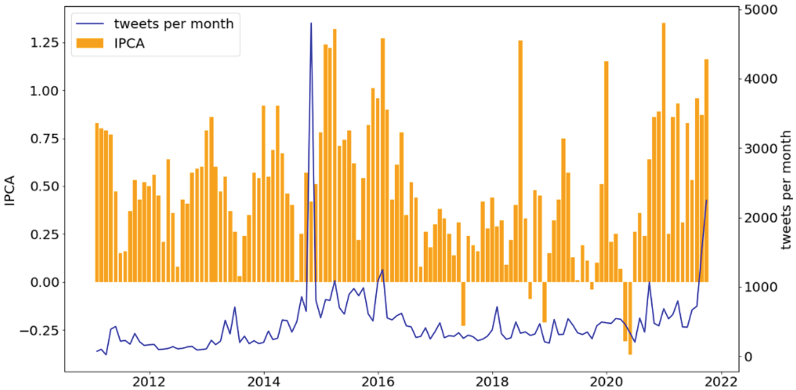

Fig. 2. Number of tweets per month and the monthly variation of the IPCA between 2011 and 2021.

Therefore, Gephi© provided a set of terms that presented at least some betweenness centrality, allowing the identification of the most relevant terms, as shown in Table 3. In the context of temporal analysis, the evolution of the betweenness centrality was examined from 2011 until 2020 because the data was available in all months in these years. We also produced an interactive graph in order to analyze the relevance temporality of the terms contained in the messages, which can be accessed through the electronic address https://public.flourish.studio/visualisation/7900429/.

Table 3 enhances that the term 'inflacao' (inflation) had a high centrality due to its application as a search term. Otherwise, in the first years of the analysis, we identified the terms with the highest centrality related to the demand of maintaining price stability. Among these, the terms 'manter' (to keep), 'real' (Brazilian currency) and 'preco' (price) repeatedly appeared as the most relevant until 2013. After the first years, starting in 2014, other terms acquired greater relevance, demonstrating the persistence of a situation that resulted in an economic crisis. The terms 'alta' (high), 'contas' (accounts), 'desempregado' (unemployed) and 'crise' (crisis) emerged with greater centrality. Thereafter, in

2015 and 2016, the inflation – which increased the price of important products – could be identified by the greater centrality of the terms ‘gasolina’ (gasoline) and ‘gas’ (cooking gas).

Table 3. Time evolution of the betweenness centrality from 2011 until 2020.

Term	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
inflacao	7383.85	7408.67	9862.33	10132.92	11109.95	10138.50	10410.93	11120.92	11554.08	11270.92
real	243.00	243.00	281.00	285.00	299.00	285.00	289.00	299.00	305.00	301.00
manter	243.00	243.00	281.00	285.00	299.00	285.00	289.00	150.00	305.00	301.00
preco	125.95	126.33	284.83	288.83	304.05	147.50	151.03	155.33	310.33	156.33
dilma	122.00	–	141.00	143.00	150.00	143.00	145.00	150.00	153.00	151.00
crise	122.00	122.00	141.00	143.00	150.00	143.00	145.00	150.00	153.00	151.00
aumento	75.93	86.00	86.75	93.58	107.25	94.00	100.77	102.33	108.08	102.33
acima	39.12	2.50	45.00	45.25	46.75	45.25	45.62	46.00	47.83	47.50
salario	36.12	54.17	42.00	42.25	43.75	42.25	42.62	44.50	44.83	44.50
juro	5.20	3.75	5.08	5.08	5.08	2.83	3.42	5.08	5.08	5.08
alto	3.50	3.83	6.92	9.25	8.95	7.33	7.83	7.33	9.00	7.33
subir	2.33	1.17	2.92	2.92	2.78	2.00	2.17	2.33	2.83	2.33
ano	2.00	2.17	3.25	2.00	2.00	2.00	2.00	2.00	2.00	2.00
dolar	1.25	0.33	2.25	1.25	7.92	0.92	2.33	3.75	6.25	1.25
governo	1.00	0.50	2.00	1.00	1.00	1.00	1.00	2.00	1.00	2.00
baixo	0.50	0.50	0.50	0.50	0.50	–	0.50	0.50	0.50	0.50
taxa	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
gasolina	0.42	0.58	0.42	0.42	299.87	144.67	288.53	299.67	305.92	301.67
minimo	0.33	–	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
contas	–	122.00	141.00	143.00	150.00	143.00	145.00	150.00	153.00	151.00
passar	–	–	141.00	0.50	0.50	0.50	0.50	0.50	0.50	0.50
fhc	–	–	141.00	0.50	0.50	0.50	–	150.00	–	151.00
alta	–	–	141.00	–	290.28	–	145.00	–	153.00	–
familia	–	–	141.00	143.00	150.00	143.00	145.00	150.00	153.00	151.00
pib	–	–	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
desempregado	–	–	0.42	146.42	153.42	146.42	148.42	153.42	156.42	154.42
dinheiro	–	–	–	143.00	150.00	143.00	145.00	150.00	153.00	151.00
fome	–	–	–	1.00	1.00	1.00	1.00	1.00	1.00	1.00
coxinhas	–	–	–	–	150.00	–	–	–	–	–
gas	–	–	–	–	150.00	–	145.00	150.00	153.00	151.00
pao	–	–	–	–	5.92	–	–	–	–	–
estavel	–	–	–	–	3.20	–	–	–	–	–
democracia	–	–	–	–	–	–	145.00	1008.00	1029.00	1015.00
deputado	–	–	–	–	–	–	4.00	298.00	304.00	300.00
senador	–	–	–	–	–	–	3.00	150.00	153.00	151.00
federal	–	–	–	–	–	–	3.00	444.00	453.00	447.00

(*continued*)

Table 3. (*continued*)

Term	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
vereador	–	–	–	–	–	–	–	870.00	888.00	876.00
distrito	–	–	–	–	–	–	–	730.00	745.00	735.00
estadual	–	–	–	–	–	–	–	588.00	600.00	592.00

Through the temporal evolution of the network, the term 'democracia' (democracy) emerged as one of the greatest centralities from 2017 onwards. The incidence of this word in the tweets remained high throughout the rest of the analyzed period, demonstrating the concern regarding this sensitive topic in discussions held by the society. In 2018, due to the state and federal elections, the words 'vereador' (city councilor), 'deputado' (congressman), 'estado' (state), 'federal' (federal) and 'distrito' (district) began a period in which they had greater relevance in the corpus of texts, remaining between the most important ones through 2020.

The network of words obtained in 2021 enabled to build a graph containing the most relevant edges, as shown in Fig. 3. The terms connected to inflation have red edges, allowing identification of the word correlations, calculated by the edge weight through Gephi©. The graph indicated that the edges with greater weight (characterized by the thickness of the connection curve) correlate 'inflation' to terms that denotes its increase: ('inflacao', 'alto'): 2,576; ('aumentar', 'inflacao'): 1,704; ('acima', 'inflacao'): 1,508; and ('inflacao', 'subir'): 1,376. Likewise, the edge ('inflacao', 'desempregado') – in English, ('inflation', 'unemployed') –, with a weight of 1,676, enabled to infer that there was a correlation between inflation rate growth and the rise in unemployment in this year.

Furthermore, excluding the bigrams with 'inflation', the edge with the highest weight was ('salario', 'minimo') – in English, ('salary', 'minimum') –, with a weight value of 1,011. Other bigram with the same characteristics were ('banco', 'central') and ('juro', 'alto') – respectively, ('bank', 'central') with 546, and ('interest rate', 'high') with 483 –, indicating the importance of the central bank to carry out the right decisions for controlling high inflation rates.

The monetary policy of a country seems to be a technical aspect restricted to Economists and stock traders. Nevertheless, the achievements brought forward that it was a major concern for users of social networks, even though the words applied in the messages were not specific technical terms used in the Economy. Therefore, the results reflected that social network environments are responsive to the influence of price increase, as it directly impacts everyday life.

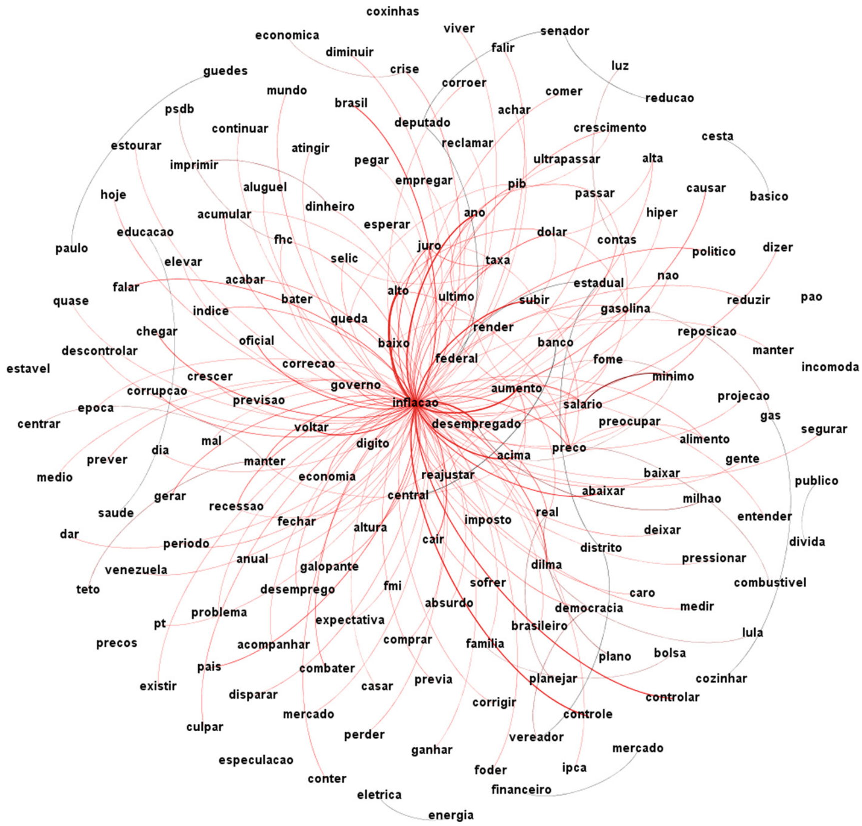

Fig. 3. Graph of the most relevant terms of the network, from the messages of 2021.

5 Final Remarks

Social networks influence people's behavior and, consequently, the economy, in both positive and negative aspects. Considering that inflation is directly linked to the loss of purchasing power, as also to long-term planning and the cost of public debt, deepening studies in this area becomes a central concern to macroeconomics. The use of artificial intelligence as an ally to conduct analysis of the economic scenario is important and can bring the opportunity to apply natural language processing techniques for better understanding the inflationary phenomenon in Brazilian society.

In this study, we suggested the existence of a correlation between the messages transmitted by Twitter© users and the monthly variation of inflation measured through the broad national consumer price index (IPCA) in Brazil. When the inflation rate is high, Twitter© users could discuss and express their private opinions about the general price level more frequently. The use of Twitter© messages to make predictions is well known in the world academy, since the social networks reflect the opinion of a significant sample of the social stratum, and is used to predict inflation [5], to predict the price of goods [7, 8], and to analyze feelings expressed in social network messages [6, 10].

Considering the methods applied in this research, natural language processing provides the ability to process and understand unstructured text, by parsing the meaning of a user content message in an automatized manner [14]. Its main objective is to allow computers to understand human language, which is inherently complex and ambiguous. The task of splitting the words from a corpus of text and rendering in an *n*-gram format enable the text representation in a low dimension format for a better machine understanding. However, as context-specific interpretation and domain knowledge are some challenges when performing NLP [15].

The network analysis identified prominent bigrams that reflected the price increase with a straight connection with inflation. Although this correlation was expected, terms related with economics difficulties associated with extended periods of inflation and the degradation of the economic scenario were also discovered in the tweets. Likewise, the bigrams without the term 'inflation' also included words associated with the worse circumstances caused by the price increase. Consequently, we concluded that inflation influences the behavior of social network users and, additionally, natural language processing of Twitter© messages can reveal relevant knowledge for inflation analysis and have potential for prediction purposes.

The research results contributed to increase the knowledge for building good data models for predictions in the Economy field. These capabilities should be examined in future research using tweets and could be confronted with other inflationary indices, such as the IGP-M and the INPC, as well as the use of other tools to compile data for analysis.

Acknowledgement. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. It was carried out under the Graduate Program in Knowledge Engineering and Management (PPGEGC), at the Federal University of Santa Catarina (UFSC).

References

1. Mankiw, N.G.: Macroeconomics, 3rd edn. Worth Publishers, New York (1996)
2. Akerlof, G.A., Dickens, W.T., Perry, G.L.: Low inflation or no inflation: should the federal reserve pursue complete price stability? Challenge **39**(5), 11–17 (1996)
3. Metas para a inflação: https://www.bcb.gov.br/controleinflacao/metainflacao. Accessed 22 Nov 2021
4. Nakamura, E., Steinsson, J.: High-frequency identification of monetary non-neutrality: the information effect. Q. J. Econ. **133**(3), 1283–1330 (2018)
5. Índices de preços: https://www.bcb.gov.br/controleinflacao/indicepreco. Accessed 09 Sep 2021
6. Aromi, J.D., Martin, L.: Forecasting inflation with twitter. Asociación Argentina de Economía Política (2020)
7. Kumar, J., Rao, T., Srivastava, S.: Economics of gold price movement-forecasting analysis using macro-economic, investor fear and investor behavior features. In: Srinivasa, S., Bhatnagar, V. (eds.) BDA 2012. LNCS, vol. 7678, pp. 111–121. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-35542-4_10
8. Alex, D.: Anchoring of inflation expectations in large emerging economies. J. Econ. Asymmetries **23**, e00202 (2021)

9. Kim, J., Cha, M., Lee, J.G.: Nowcasting commodity prices using social media. PeerJ Comput. Sci. **3**, e126 (2017)
10. Bastida, J.O., Gallego, A.J., Juan, J.R.R., Albarrán, P.: Regional gross domestic product predictions using Twitter deep learning representations. In: 17th International Conference on Applied Computing, pp. 89–96. Lisboa (2020)
11. Shapiro, A.H., Sudhof, M., Wilson, D.J.: Measuring news sentiment. J. Econom. **228**(2), 221–243 (2022)
12. Índice Nacional de Preços ao Consumidor Amplo: https://www.ibge.gov.br/estatisticas/economicas/precos-e-custos/9256-indice-nacional-de-precos-ao-consumidor-amplo.html. Accessed: 10 Oct 2021
13. Freeman, L.C.: A set of measures of centrality based on betweenness. Sociometry **40**(1), 35–41 (1977)
14. Chen, X., Ding, R., Xu, K., Wang, S., Hao, T., Zhou, Y.: A bibliometric review of natural language processing empowered mobile computing. Wirel. Commun. Mob. Comput. **18**, 1–21 (2018)
15. Pouyanfar, S., et al.: A survey on deep learning: algorithms, techniques, and applications. ACM Comput. Surv. **51**(5), 1–36 (2018)