

# Sentiment analysis bias in 2022 Brazilian election tweets

Rafael Santos Gouveia

Andrei Fokin Teixeira

Affiliations: Engineering and Data Science Department - Coimbra University

rgouveia@student.uc.dei.pt

afokin@student.uc.dei.pt

## Abstract

In the 21th century, democratic elections and social media are merging with consequences for human relations. One of them is the user feelings getting more prominent. Often they are just supporting or opposing candidates in the election process, but with polarization, identifying hate speeches, fake news and bias become interesting things for the data scientists work to support society with good practices. The present study will look for tweet sentiments during the 2022 Brazilian elections in relation to the two main candidates in the dispute, Lula da Silva and Jair Bolsonaro.

## 1. Introduction

Technology impacts all aspects of human life, from infrastructure to communication, from health to entertainment. Its presence is such that it reduces costs, distances and increases people's access and participation to goods and services. Information is one of these goods mainly after the advent of the Internet.

Social networks are examples of open spaces inside the Internet, some used for study, some for photos, some for videos and so on. Sometimes they have many objectives, but they end up specializing in some way. "Many companies and people use them to spread products and services and publish opinions, facts that have turned the social networks into powerful sources of information on many topics" (Cristiani *et al*, 2020). But because the so-called "algorithms", over time "social networks such as Facebook and Twitter have been accused of facilitating the spread of falsehoods, hatred and conspiracy theories" (Elizabeth *et al*, 2018), even even when people begin to feel comfortable and within a group that thinks in a way that confirms their personal beliefs.

Twitter is an open and low-censorship wall, where everyone can express free opinions, some of them about policy. On the readers' side, policy is an aspect of life that

gains strength in a digital arena, where one can express more radical opinions with less social punishment, especially in anonymous profiles. On the side of the candidates, in turn, policy is a competitive game, whose objective is the conquest of power and, in democratic societies, convincing the average voter that their campaign is a good option requires sophisticated tools.

When voters are increasingly connected to new spaces for discussion, the migration of parties and candidates to social networks is expected. At the limit, the candidate submits its opinion to the masses opinion and voters support things that make sense to them, in a loop known as "confirmation bias". Thus, there is a mutual exchange that strengthens a set of ideas that make sense for a particular social group, even after the election, when a government and its opposition are transformed. Therefore, the Internet becomes a space for governance itself, where networks, such as Twitter, are the communication channel and users have the feeling of greater participation.

In Brazil, where society is heterogeneous and where social networks are a novelty for political purposes, the electoral campaign of 2022 has the potential for aggression in the real world. Therefore, the main motivation of this work is to understand how a machine learning model can be used to classify and identify comments with political content. Insights can be used inside companies not to spread hate.

To this end, two are the main hypotheses raised: (i) the ranking of a good estimator should take into account the initial proportion of sentiments, for example, if there are loads of positive content rows, is expected to have a good-feeling estimator; and (ii) there should be no ranking bias when changing proper names from Lula to Bolsonaro or vice versa.

## 2. Related Work

One of the uses of social networks data is to create security systems against fake news, malicious robots or any abuse of democratic or moral rules. Even when "the real-time sentiment classification on general topics on a

real time basis is particularly challenging for data dynamicity and lack of labeled data” (Guerra *et al*, 2011), there are opportunities for the data scientist to improve content classification models. A phenomenon of communication that may help this complex analysis is that “while new terms may arise and old terms may change their meaning, user bias tends to be more consistent and robust over time as a basic property of human behavior” (Guerra *et al*, 2011). Technology must advance faster than human behavior, so the greater the technical capacity to detect patterns, the greater the chance of good tools for protecting data and/or people, according to the interests of society itself.

Cristiani et al (2020) worked in a sentiment analysis of the 2018 election in Brazil. The goal was to relate people's opinions to traditional street opinion polls: “the results show that Twitter is a great source of information, especially as a source of research on the opinion of its users. The numbers show that the amount of positive messages have a strong relationship with the polls (...). Although the studies did not show the relationship between the tweets classified as neutral and negative, this information can be useful to assist in various types of decision making”.

Chaudhry et al (2020) analyzed tweets of the 2020 US elections and compared Trump's performance in his two electionss. A time series analysis was used to also bring dynamics through the year of 2020 in different states and for different subjects. One conclusion is that “for all states where sentiment results did not corroborate with election results, long-term trends before and after the election reveal that there was an increase in the positive sentiment of the winning candidate”. The current paper will not take into account, but it would be interesting for future works to have insights on how former-president Lula da Silva returned to power after 12 years and how to use macro data to predict future election scenarios.

Ricci (2020) brings some insights about preprocessing data. Regarding the use of emojis, by removing then “the Random Forest algorithm had the greatest gain of all”, precisely the algorithm used in this paper. “In the case of stopwords removal, the Logistic Regression and SVM algorithms had a decrease in accuracy, while Naïve Bayes and Random Forest algorithms had a gain. This reveals that also the type of specific treatment affects different algorithms in different ways”.

### 3. Data & Approach

This work will be carried out in the sequence of actions: (i) obtaining data from Twitter; (ii) preprocessing data by removing accentuation, obvious writing errors and

orthographic mistakes; (iii) train a random dataset with two different models (Random Forest and Decision Trees) to help checking accuracy; and (iv) after a good accuracy, do a simple test write some random sentence and changing the name of the two main candidates to see if there is some bias related to the persona name.

#### 3.1 Inicial data information

This project has the Twitter Application Programming Interface (API) as its source of data. Talking about the tweets themselves, 413 comments were collected at all and divided as follows: 20% for training and 80% for testing. From the whole universe, 236 were taken about Lula, 213 about Bolsonaro. There is some intersection between them, because in some tweets, both are mentioned. Figure 1 shows the groups.

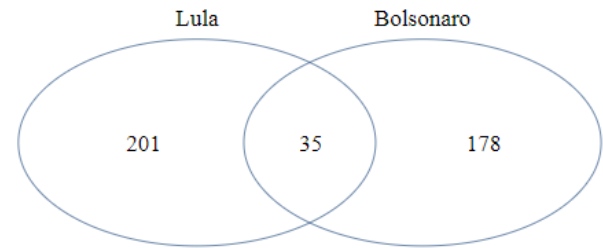


Figure 1: Number of tweets for each candidates

Initially, there were only used "lula" and "bolsonaro" tweets, but “eleicoes” theme was took into consideration to complement the first ones and to help both in the diversity of opinions around elections as well as increase the number of observations. A list was created with new words to better characterize each group and collect even more information without context losses:

Candidates	Key words
Lula da Silva	“lula”, “lulalivre”, “lulaladrao”, “molusco”, “lulapresidente”, “lula13”
Jair Bolsonaro	“bolsonaro”, “bozo”, “bozonaro”, “jair”, “elenao”, “elesim”, “mito”, “messias”, “bolsonaro22”
Eleições 2022	“eleicao”, “eleicao22”, “eleicoes2022”, “eleicoes22”

Table 1: Dictionary of words used to group candidates

The period is about October, after the result of the 1st round, where there was no decision and when the campaign continues until the end of the month. Tweets whose location is identified and within Brazil are sought, as the sense is to see opinions from people who are

experiencing and being impacted by the process. Thus, tweets were collected outside the official accounts of the two candidates, no matter if they were original or retweeted and being indifferent to photos, videos, emojis and any other non-textual artifice.

### 3.2 Programming approach

After the data is defined, the tweets are extracted directly from Twitter's API with the help of Python's Tweepy library, which will link the Python code to the network's social media information. This tool allows to collect tweets from the day the query was performed or up to a certain date. Some filters were used for the queries, one of these filters is the language of the tweets in "pt-br". The result of this query is a json document that has all the data and metadata of a tweet.

After extration, to clean up the data, some steps were took into the following pre-processing steps:

1. removing whitespace before and after sentences;
2. removing accentuation and cedillas;
3. removing all special characters, leaving just letters and numbers;
4. removing "@" from users;
5. removing "#" from hashtags; and
6. removing stop words; and
7. normalizing words from the dataset using lematize: here the lemma of the word is found so that the context of the sentence is taken into account.

On this treated database, Scikit-learn algorithm Count Vectorizer (CV) is used to transform the text into a vector based on the frequency of each word that occurs in the text. Another Scikit-learn algorithm, Tfidf Vectorier (TV), is used to measure how relvant a word is to a document by creating a matrix with this values and comparing how it will affect the prediciotn of the model.

Three Scikit-learn algorithms Random Forest Classifier (RFC), Decision Tree Classifier (DTC) and Support Vector Classification (SVC) are used to train and test vectors and transform them into sentiment classifiers. They are supervised machine learning algorithms used for classification modeling. Their classifiers create a set of decision trees from a randomly selected subset of the training set. "The main distinction between them is that DTC are graphs that illustrate all possible outcomes of a decision using a branching approach. In contrast, RFC outputs are a set of decision trees that work according to the output" (Talari, 2022). SVC "works mapping data points to a high-dimensional space and then finding the optimal hyperplane that divides the data into two classes." (Kumar *et al*, 2022).

Last but not least, after training each model, some metrics are used to see their efficiencies, such as:

- accuracy: the number of hits in our model divided by the total sample
- f1-score: the harmonic mean between precision and recall, a global metric for the model.

## 4. Experimentation

The treated data for each candidate and ranking sentimento (1: positive, 0: neutral, -1: negative) follow:

Candidates	Classification	Quantity
Lula da Silva	1	109
	0	25
	-1	102
Jair Bolsonaro	1	60
	0	28
	-1	125

Table 2: Real values by candidate and emotion

Many tweets have obvious classification, but facing doubtful scenarios must be dangerous for the estimator. It's mandatory to the following some criteria, so here it follow some adopted rules:

- having a journalistic format or a simple report of some event: neutral;
- when the comment is clearly positive for one and clearly negative for another: positive;
- critical but thoughtfully written: negative;
- presence of adversative adverbs for positive and negative content: neutral.

The final database was passed through RFC, DTC and SVC, whose rating targets is in the following list:

```
['neutro' 'negativo' 'positivo']
```

Figure 2: Decision list

Below, the test metrics result:

Model	Accuracy CV	Accuracy Tfidf	F1-Score CV	F1-Score Tfidf
RF	0,67	0,67	0,66	0,66
DT	0,57	0,51	0,57	0,49
SVC	0,67	0,68	0,62	0,67

Table 3: Test metrics results by model

Table 3 shows that even when using two vectorization methods, Count Vectorizer and Tfidf Vectorizer, the results of Accuracy and F1-score are very close when comparing the same models. When comparing the results of the training algorithms DTC falls short of RFC and SVC, which have similar results.

Once the average estimator was known, an experiment was conducted to understand how accurate it is with respect to the candidates' names (bias). In this way, it is possible to understand not only the efficiency of the estimator but also its consistency.

A test dataframe was created with new phrases to swap the names Lula and Bolsonaro in each one. Below in Table 4 are some of the phrases used, where X is the dedicated space to fill with the name:

#	Random created sentence
1	"Gosto muito do presidente X"
2	"X foi um ladrão e merece ir para a cadeia"
3	"X foi um dos piores presidentes do Brasil, roubou e fez o que queria para se manter no poder"
4	"X foi muito bom para o povo brasileiro, deu oportunidades e emprego"

Table 4: Dictionary of sentences to test bias

Results are presented in Table 5 for RFC model, Table 6 for DTC model and Table 7 for SVC. The experiment showed two things: the first, in some sentences, every classifier gave different scores by changing the name. On the other hand, the initial sentiment distribution corroborates a better score for Lula than Bolsonaro.

Treated sentences	RFC TV	RFC CV
gosto presidente lula	-1	-1
gosto presidente bolsonaro	-1	-1
lula ladrao merece ir cadeia	1	1
bolsonaro ladrao merece ir cadeia	1	1
lula piores presidentes brasil roubou fez queria manter poder	1	1
bolsonaro piores presidentes brasil roubou fez queria manter poder	1	1
lula bom povo brasileiro deu oportunidades emprego	1	1
bolsonaro bom povo brasileiro deu oportunidades emprego	-1	0

Table 5: Bias experiment results / Random Forest

Treated sentences	DTC TV	DTC CV
gosto presidente lula	-1	-1
gosto presidente bolsonaro	-1	-1
lula ladrao merece ir cadeia	-1	1
bolsonaro ladrao merece ir cadeia	1	1
lula piores presidentes brasil roubou fez queria manter poder	1	1
bolsonaro piores presidentes brasil roubou fez queria manter poder	-1	1
lula bom povo brasileiro deu oportunidades emprego	1	1
bolsonaro bom povo brasileiro deu oportunidades emprego	-1	0

Table 6: Bias experiment results / Decision Tree

Treated sentences	SVC TV	SVC CV
gosto presidente lula	-1	-1
gosto presidente bolsonaro	-1	1
lula ladrao merece ir cadeia	1	1
bolsonaro ladrao merece ir cadeia	1	1
lula piores presidentes brasil roubou fez queria manter poder	1	1
bolsonaro piores presidentes brasil roubou fez queria manter poder	1	-1
lula bom povo brasileiro deu oportunidades emprego	1	1
bolsonaro bom povo brasileiro deu oportunidades emprego	1	-1

Table 7: Bias experiment results / Support Vector

To summarize the tables above, a simple aggregate mean is calculated for each candidate and model (Table 8):

Candidate	RFC	STC	SVC
Lula da Silva	0,5	0,25	0,5
Jair Bolsonaro	0,125	-0,125	0,25

Table 7: Simple classification mean by model

All of the algorithms are likely to give better sentimental classification for Lula's tweets, each one with its metrics. The less biased classifier is the SV one, while the most polarized (even weak, because is around zero) is the ST one.

Figures 3 to 5 show the evolution of the classification accumulated mean (CAM) by vectorization method for each candidate in a different vectorizer method. So, the first group will look for Lula's CAM evolution with TFDiF Vectorizer, the second group for Lula's CAM progress under Count Vectorizer, the third for Bolsonaro's CAM sequence under TV and the last for Bolsonaro's CAM under CV.

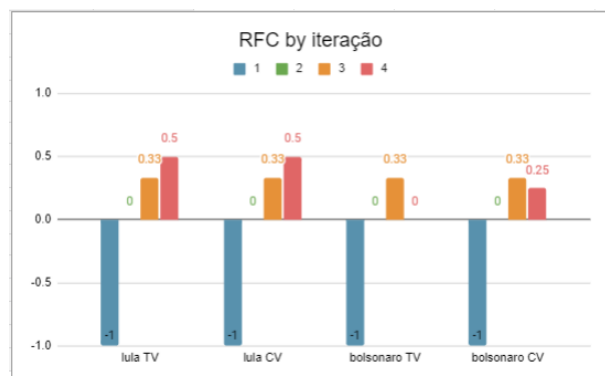


Figure 3: CAM by vectorization / Random Forest

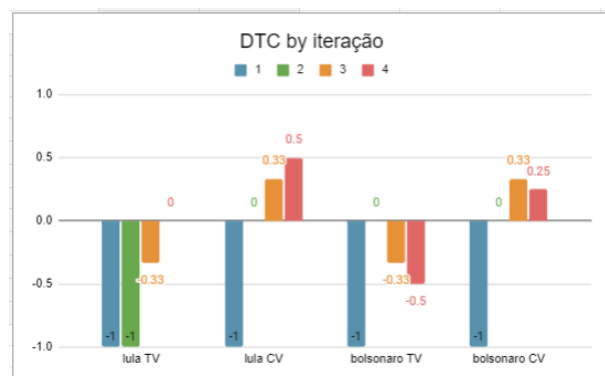


Figure 4: CAM by vectorization / Decision Tree

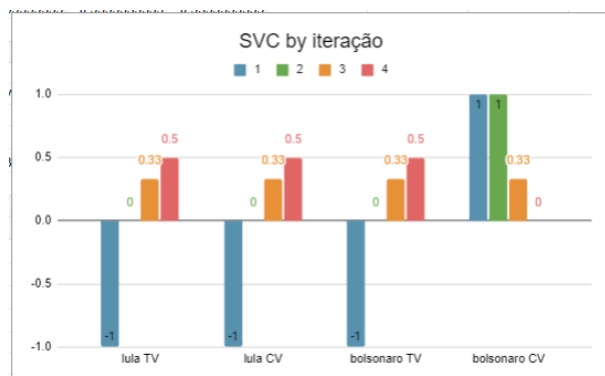


Figure 5: CAM by vectorization / Support Vector

## 5. Conclusions

This work had a objective to check how classification text models works to estimate emotional sentiment by

reading texts. The content were political tweets written in one of the most polarized election in Brazil's history. After cleaning data, three models were used to estimate the sentiment. By collecting almost a hundred comments for training, transforming them into numerical / vectorial representations and doing some initial classification between -1, 0 and 1 under the sma criteria, the results tend to classify sentences with "lula" better than those with "bolsonaro" word.

In all three models, this must happened due to a low data ammount. Also, some tweets mention both of the candidates in the same sentence as presented in Figure 1. A third reason was the difficulty in retrieving this data, first because of the Twitter API that does not allow downloading tweets from periods longer than 7 days from the request date and second, there is no control of which tweets will come, so there is the possibility of having duplicate tweets or data can come unbalanced, for example, i.e., it would be needed a very long time to get enough and balanced data (taking new diferente tweets from 7 to 7 days). A forth reason is that, in pratical terms, there were just possible to extract tweets from an old account with some historic and interactions with users and comments, so this could be led to receive biased data.

By eliminate each of the reason above, results could be less biased. That is why, for future work, a larger ammount of tweets will be used to see if this same set of variables, such as two types of tweet vectorization algorithms and three classification algorithms will maintain the biased behavior or if the bias will decrease. Also, a bigger experimentation test can be put in practice, with, for example, 100 sentences. This would create interesting views of the evolution of the CAM for each model, a valuable output for comercial applications, social feedback and political analytics

The hypothesis raised that the initial distribution of tweets should impact the final results is verified and true. A higher amount of negative tweets for Jair Bolsonaro caused the comments to be more poorly evaluated than those for the newly elected president Lula da Silva.

Finally, the main contribution of this work is to understand the great number of methodologic elements that are able to impact the final result of social media contente analysis: the age of twitter account, the number of observations, criteria to classify the train dataset, a well balanced number of positive and negative feelings, the number of experiments to check bias, tools to vectorize sentences, models to run (and all of this in a time slap objective). In summary, the study was capable of transmitting the importance of text classification and how data scientists have powerful tools to support society against hate and fake news on social media.

## 6. References

A Cristiani, D Lieira, H Camargo., 2020. *A sentiment analysis of brazilian elections tweets*, in: Anais do VIII Symposium on Knowledge Discovery, Mining and Learning, p. 153-160.

Chaudhry, H.N.; Javed, Y.; Kulsoom, F.; Mehmood, Z.; Khan, Z.I.; Shoaib, U.; Janjua, S.H. Sentiment., 2020. *Analysis of before and after Elections: Twitter Data of U.S. Election*. Electronics 2021, 10, 2082. Available on: <[https:// doi.org/10.3390/electronics10172082](https://doi.org/10.3390/electronics10172082)>.

Elisabeth Lex, Mario Wagner, Dominik Kowald: *Mitigating Confirmation Bias on Twitter by Recommending Opposing Views*, 2018. CoRR abs/1809.03901. Available on: <[https://www.researchgate.net/publication/327592064\\_Mitigating\\_Confirmation\\_Bias\\_on\\_Twitter\\_by\\_Recommending\\_Opposing\\_Vi](https://www.researchgate.net/publication/327592064_Mitigating_Confirmation_Bias_on_Twitter_by_Recommending_Opposing_Views)ews>.

Kumar Ajitesh, *SVM Classifier using Sklearn: Code Examples*, 2022. Available on: <<https://vitalflux.com/svm-classifier-scikit-learn-code-examples/>>.

PH Calais Guerra, A Veloso, W Meira Jr, V Almeida, 2011. *From bias to opinion: a transfer-learning approach to real-time sentiment analysis*, in: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, p. 150-158.

Ricci, R. D. *Análise de sentimentos no Twitter sobre a Reforma da Previdência no ano de 2019*, 2020, Monografia, Universidade Federal de Uberlândia, Uberlândia, MG, p. 70.

Talari Sakumar, *Random Forest vs Decision Tree: Key Differences*, 2022. Available on: <<https://www.kdnuggets.com/2022/02/random-forest-decision-tree-key-differences.html>>.