Using Supervised Classification to Detect Political Tweets with Political Content

Brunna de Sousa Pereira Amorim
Federal University of Campina Grande
Campina Grande, Paraíba, Brazil
brunnasousa@copin.ufcg.edu.br

André Luiz Firmino Alves
Federal Institute of Education, Science and Technology of Ceará
Iguatu, Ceará, Brazil
andre.alves@ifce.edu.br

Maxwell Guimarães de Oliveira
Federal University of Cariri
Juazeiro do Noirte, Ceará, Brazil
maxwell.oliveira@ufca.edu.br

Cláudio de Souza Baptista
Federal University of Campina Grande
Campina Grande, Paraíba, Brazil
baptista@computacao.ufcg.edu.br

ABSTRACT

Social media platforms have been increasingly used by modern society. In most platforms, users usually share content on various subjects and, in particular, politics is a favorite one. There are many interests in detecting and analyzing such a political content. However, there is a challenge in the process of detecting specific subjects from social media data mainly due to its informality. In this paper, we propose and compare two techniques, based on supervised classification, for the detection of tweets with political content. The results obtained by our approach have demonstrated satisfactory performance, which motivates further research to be undertaken.

CCS CONCEPTS

• **Information systems → Web and social media search**; • **Computing methodologies → Machine learning approaches**;

KEYWORDS

Machine learning, Deep learning, Logistic regression, Social networks analysis, Text mining

ACM Reference Format:
Brunna de Sousa Pereira Amorim, André Luiz Firmino Alves, Maxwell Guimarães de Oliveira, and Cláudio de Souza Baptista. 2018. Using Supervised Classification to Detect Political Tweets with Political Content. In *Proceedings of the 24th Brazillian Symposium on Multimedia and the Web (WebMedia'18).* ACM, New York, NY, USA, Article 4, 8 pages. https://doi.org/10.1145/3243082.3243113

1 INTRODUCTION

Nowadays we see a massive use of social networks, with networks like Facebook[1] that already have more than 2 billion users worldwide. This massive amount of users frequently posts on various subjects, among which politics is one of the most frequently used topics.

The Brazilian electoral legislation and political parties are investigating ways to detect a political post and analyze it to ascertain various characteristics, whether it is an out-of-term political advertisement and therefore illegal; if it is fake news; if it is offensive and attentive to the honor of another, and may cause harm to a particular candidacy; among other electoral crimes.

On the other hand, candidates or citizens in general may be interested in reading political posts of a particular candidate, a political party or even an ideology (e.g. left-wing or right-wing, conservative, liberal, democrat, Christian, etc.).

Monitoring social networks online enables a broad spectrum analysis of opinions, both on municipal, state and federal levels, and political topics of international interest. These analyzes are useful to many society sectors such as political parties, politicians, citizens, business groups, social movements, NGOs (Non-Government Organizations), or any other form of civil organization. Moreover, these analyzes subsidize the behavioral strategies of politicians as well as the strategies of political parties in the conduct of electoral campaigns. Such analyzes, including sentiment analysis techniques, provide these people with an instant and up-to-date reading of the reality regarding politics.

Giménez emphasizes that during the electoral period, politicians and voters engage in diversified conversations from the thematic point of view [14]. In addition, posts about government programs, details about the candidate's curriculum, as well as attacks on opposing candidates are disclosed through political marketers. There is, therefore, a very powerful means of communication with wide penetration in society, which can have significant effects on the election results.

Therefore, developing computer techniques that can carry out such political analyzes from detection to discourse analysis, through the identification of false news, the use of sarcasm and irony, among others, constitutes an important and challenging area of research.

This paper presents a solution that addresses one of the complex problems discussed above: the detection of political posts in social networks. For such, machine learning techniques were used, more precisely, supervised classification. We used Twitter posts (called tweets) that were classified using logistic regression and

[1]https://www.facebook.com/

WebMedia'18, October 16 - 19, Salvador-BA, Brazil

ACM ISBN 978-1-4503-5867-5/18/10.
https://doi.org/10.1145/3243082.3243113

deep learning models. At the end, a performance comparison of these algorithms is made and analyzed.

The remaining of this paper is structured as follows. Section 2 presents a theoretical reference on text mining and learning techniques used. Section 3 discusses related work. Section 4 focuses on the methodology used in the study. Finally, section 5 concludes the article and points out directions for future work.

2 THEORETICAL BACKGROUND

The automated detection of specific content in tweets has been based on several well-known techniques in the literature [8], such as Information Retrieval (IR) [9] , Natural Language Processing (NLP) [19] , Machine Learning (ML) [23] and Information Extraction (IE) and text mining [2].

In this context, text mining is a variation of data mining that attempts to find interesting patterns and information previously unknown from various text sources. The difference between pure data mining and text mining is that, in text mining, the patterns are extracted from natural language instead of a structured database [17]. This feature adds complexity to the text mining approaches, since it is a challenge working with texts in natural language. Moreover, texts of social media nature, such as tweets, come with an additional complexity as texts tends to be shorter, more scattered and with imprecise and informal language.

Text mining involves NLP, IR, IE and ML in the process of discovering and classifying texts automatically. Generally in text mining, the extraction and selection of features, the classifying method and pre-processing have a substantial influence in the text classification success [24]. For this reason, this work focuses on the classification method.

Text classification is a task that can be defined as the assignment of predefined categories to text documents, where these documents could be news, technical reports, web pages, tweets, etc. Formally, text categorization is finding a function that approaches the classification function $F : T \rightarrow C, f(t_i) = c_i$ that describes how texts are related with the classes, assigning a text $t_i \in T$ to a category $c_i \in C$, where T represents the text domain and $C = \{c_1, ..., c_n\}$ is the set of n predefined categories.

In general, a task of text classification begins with a training set $T = (t_1, ..., t_n)$ of texts that are already labeled with a category $c_i \in C$ (for example, Political or Non-Political). Then, the task is to determine a classification model (F function) that is able to correctly assign a class c_i to a new text t_i of the T domain.

To measure the performance of a classification model, a fraction of the labeled texts is reserved for training (find the F function). Another fraction of labeled texts is used to test, where the built model is evaluated and the results of the model classification is compared to the true text label.

In this work, we used two classification models and performed a comparative study using logistic regression and deep learning algorithms, which were applied to a process of text classification on social media texts.

Logistic Regression is one of the most popular algorithms used in machine learning to try solving classification problems that the independent variables (features) are categorical [29]. In statistics, logistic regression is a non-linear regression model that solves problems of supervised learning for a binary dependent variable [15]. The logistic model can be used to estimate the probability of a binary response y, based on a set of n independent or predictor variables $x = [x_1, x_2, x_3, ..., x_n]$, for which $y|x$ follows the Bernoulli distribution with a $p(x)$ probability of success [11]. Recently, the use of logistic regression in researches that involves social data has been very widespread, as in [3, 12, 31], , .

Deep Learning is a special approach in machine learning that covers both supervised learning and semi or unsupervised learning. This approach seeks solving artificial intelligence problems that generally are not solved with traditional algorithms, as an example of knowledge representation, reasoning and planning [27]. Deep learning imitates the human brain capacity to learn with experience to identify objects [21] , being able to process raw data through a deep neural network hierarchy and classify objects that were already trained or that it has experience to do it, just as our brain has the ability to process raw neural inputs to learn high-level features autonomously. Consequently, this technique provides a more precise and fast processing due to its capacity of improvement and self-learning.

More specifically, in natural language processing, the use of the Bag of Words model is known as one of the main ways to convert natural language text into a format suitable to be used by machine learning algorithms. In this context, Skansi states that only deep learning has good alternatives to this model, such as one-hot encoding and word embeddings, while other machine learning techniques use Bag of Words variations almost exclusively. In spite of offering alternatives, deep learning also presents good results with Bag of Words [27].

Applications of computer vision, image processing, or speech recognition are traditional applications of the deep learning algorithms [22]. However, deep learning architectures such as Deep Neural Networks (DNN), Deep Belief Networks (DBN), Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNNs) have been adopted in several other applications such as pattern recognition and applications of natural language processing. In addition, deep learning is currently state-of-the-art in several of these application areas, due to training process stability and generalization, and the scalability to large volumes of data [1].

In general, deep learning techniques implement extraction and feature classification techniques to perform the text classification. The texts are represented as a sequence of words in a feature vector. Then, every word in the sequence is projected in a continuous vector space, where the vectors are obtained by multiplying matrices considering word weights. The word vectors feed the networks, which process the vectors in several (deep) predicting layers by calculating probabilities [16]. In other words, given a document D, with the word sequence $[w_1, w_2, ..., w_n]$, the network output is the class in which the document was classified. $p(k|D, \theta)$ is used to denote the probability of a document to belong to class k, where θ are the networks parameters.

The RNNs improve the time complexity and analyzes texts word by word, preserving the semantics of all previous texts in a hidden layer, however words that were added more recently are more meaningful than the other words. Network models such as Long

Short-Term Memory (LSTM) were proposed to overcome this problem [16].

3 RELATED WORK

The emergence of various online platforms such as social media has posed challenges regarding the huge amount of data, which requires the usage of suitable techniques in order to extract relevant information. Text mining has been used in social media data in order to provide many kinds of data analysis such as sentiment analysis. Many studies in the area of sentiment analysis obtain the sentiment polarity of a tweet just based on its textual information [5, 7]. Anchiêta and Moura [6] used a machine learning approach to generate a summary of the main bugs and new features pointed by users in Mobile stores apps, assisting the developers to improve the quality of their app reviews. Focusing on sentiment analysis in the domain of electoral processes, this section provides a brief review of the state-of-the-art.

In [13] the authors proposed two methods for distinguishing political and non-political Dutch tweets, being one a rule-based approach and the other one a supervised learning approach. The rule-based approach consists of classifying tweets as politically relevant according to the presence or absence of certain keywords, which were composed by political parties names and their leaders. To implement the supervised learning approach, the authors trained and compared the results of eight machine learning algorithms so the five best performing models (Logistic Regression, Classification and Regression Trees, Support Vector Machines - SVM, Random Forest and Artificial Neural Network) were combined to make a classifier that uses a voting system. The rule-based approach had an accuracy of 62%, while the supervised learning approach went up to 97%.

In [4] the authors investigated the benefit of utilizing hashtags to determine sentiment polarity of tweets in the political domain. They proposed an effective way to consider hashtags in classification. For such, they proposed rules to automatically annotate the dataset based on hashtags and provide an idea on how to enrich terms in tweet by adding hashtag terms. The rules consist of altering the tweet text to express the sentiment polarity of the hashtags used, which can be negative, positive, neutral or irrelevant. The results presented an accuracy of 95% in sentiment classification of tweets using hashtags, and it outperformed the approach based on unigram feature combined with Naive Bayes, SVM or Logistic Regression.

In [26] the authors focused on discursive practices used by politicians and journalists on social media. For such, they created a data set of manually annotated tweets written in four different languages, which is a costly time-consuming task. In order to expand this data set to perform data analysis, they examined the FastText [10, 18], a library for fast text representation and classification, to automatically annotate unknown tweets. The FastText had an accuracy of only 51%. To improve such results, they used three different methods: weakly supervised learning, external skipgram language models and active learning to boost the performance, increasing the accuracy by 3%.

With the growing interest in using social media data for predicting political elections, the authors in [30] studied the capability of Twitter data as an ex-ante indicator of event outcomes by modeling the momentum of political campaigns. They proposed three momentum indicators with predictive capability and analyzed them with USA's 2014 midterm election and 2016 Presidential election data. The Sentiment Indicator (SI) compares the collective sentiments of the tweets of the competing candidates running for the same office within a time period, while the Curvature Indicator (CI) measures the direction of the momentum daily or during a period. The Growth Indicator (GI) is a derivative of a fitted model to measure the momentum in a specified time period. These indicators were used in the proposed model for predictive analysis. The authors conducted an empirical study to validate these indicators, concluding their usefulness to calculate the momentum in a selected period and evaluate the trend.

Other related works also relied on machine learning for classifying social media data in the domain of politics. Nwulu [25] compared Multi Layer Perceptron, SVM and Voted Perceptron algorithms in order to classify texts according to a political positioning (e.g. right-wing versus left-wing, or conservative versus liberal). In [20] the authors analyzed tweets about the 2015 Spanish elections using SVM, Random Forest and Extra Tree supervised techniques to perform a multiclass classification of social media posts.

Amid the diversity of political parties and the consequently diversity of political positionings in the Brazilian scenario, this work is distinguished by evaluating techniques for classifying texts that enable the identification of political content in Brazilian social media and provide a multi classification according to this specific and complex scenario. Our classifier can be embedded in an information system with the purpose of filtering social media posts regarding politics. Such a filtering can be applied to at least two cases: for the people who do not wish to be notified about such kind of content (hiding filter); or for the ones who may be very interested on this kind of content (highlighting filter). People may be motivated by a number of different purposes including electoral infractions, defamation, fake news or even to follow an electoral process on social media with targeted content.

4 METHODOLOGY

The methodology used in this work was based on several phases detailed and demonstrated in the flow chart presented in Figure 1.

The first phase consisted in planning the experiments. We decided to collect data from Twitter given its relevance in the scientific community and, especially, given its openness to collect information. In the experiments, a computer with the following configuration was used: 16GB of RAM, Intel Core i7 3.4 GHz processor and 1 Terabyte of Hard Drive. The following software infrastructures were used: Windows 10 operating system, Comma Separated Values (CSV) files, Python and Java programming languages, and the statistical package R.

The second phase consisted in data acquisition. In order to classify the logistic regression model and the deep learning model implemented in this work, we collected both data with political content and data without political content. We used the Figure 2: Number of collected tweets. Twitter API for R (TwitteR[2]) to collect

[2]https://www.rdocumentation.org/packages/twitteR/versions/1.1.9

about 13,651 tweets in Portuguese during the second semester of 2017.

From those collect tweets, approximately 6,851 were classified as political (because it references to political entities or has a political context), and 6,800 were classified as non-political, thus maintaining our database balanced, a very relevant feature when training machine learning algorithms.

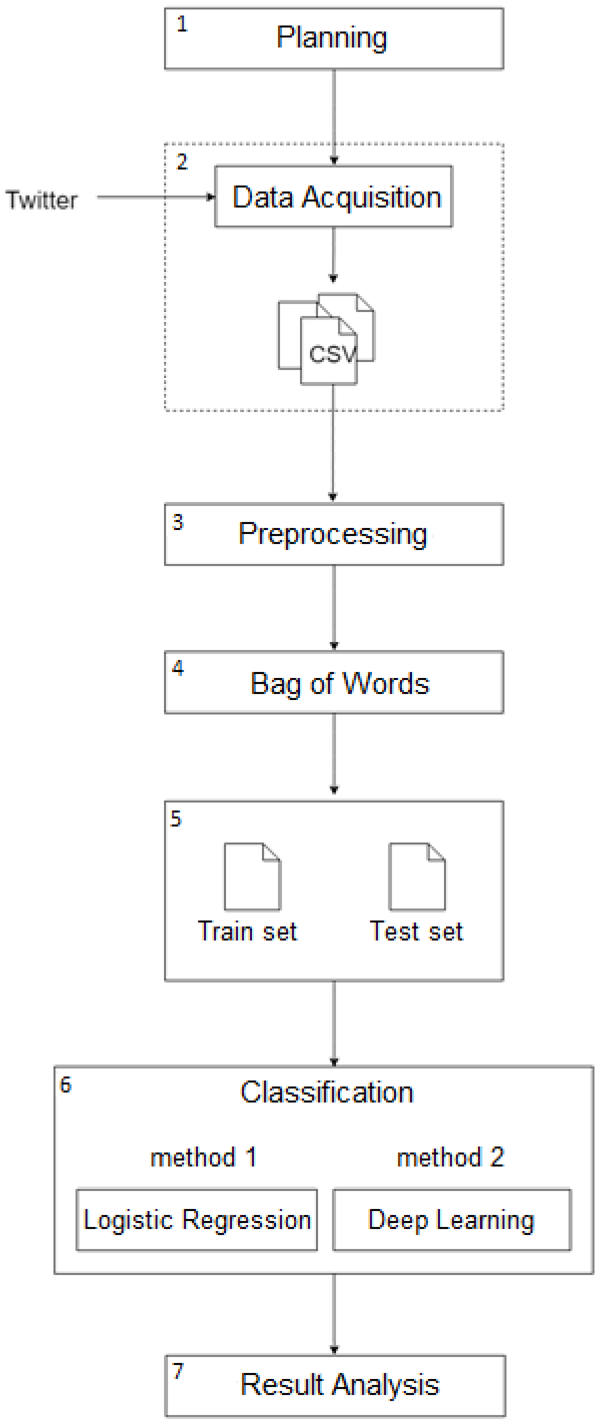


Figure 1: Flow chart of the methodology used.

Figure 2 shows the amount of tweets collected in this study, accordingly to its binary classification (political or non-political). The classification between political or non-political was performed based on collecting tweets that had a set of pre-defined keywords and tweets from political parties official Twitter accounts.

The tweets labeled as political were collected from official Twitter accounts of some Brazilian political parties that had the most significant amount of tweets, such as *PCdoB* (@PCdoB_Oficial), *PMDB* (@PMDB_Nacional), *PSB* (@PSBNacional40), *PSDB* (@Rede45), *PSOL* (@PSOLOficial) and *PT* (@ptbrasil), and by searching for a set of keywords[3]. For data considered non-political, we did a search excluding keywords related to politics.

The third phase was to perform text preprocessing using NLP techniques, aiming to identify and eliminate terms that do not contribute in the training of the machine learning model and tweet classification. As a result, stopwords[4] (i.e. irrelevant words in a language that alone have no relevant meaning), hyperlinks (or URLs) and mentions to other users have been deleted from the textual data. The mentions to other users were removed to preserve their privacy and because this information does not have textual relevance for our purpose.

Considering all selected tweets to this study, we created a Bag of Words (BoW) that stores all words present in the texts written by Twitter users. To build the BoW, we implemented a Java application that read the tweets from the database, selecting only the text of these tweets to create the dictionary, which consists of all the words present in all tweets. Then, each collected tweet is represented by

[3]https://github.com/brunnaam/TweetClassification/blob/master/palavras-chave-politico.txt
[4]https://github.com/brunnaam/TweetClassification/blob/master/stop-words-pt.txt

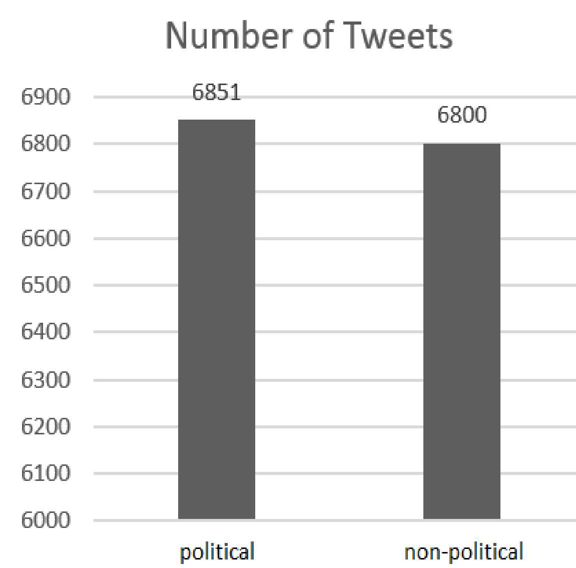


Figure 2: Number of collected tweets.

a histogram vector that measures, through a numeric statistic, how important each word of the tweet is to the dictionary. To measure this importance, the term frequency-inverse document frequency technique was used (TF-IDF) [28].

The TF-IDF is a widely used technique in Text Mining and Information Retrieval because it reflects how important a word is for a document in a corpus (a document set). Therefore, each word of each tweet text is represented by this statistic, making it possible classifying tweets in political or non-political.

Another feature present in a tweet is its metadata. With it, we can extract other information about these tweets (not necessarily text), such as the number of likes the tweet have and the number of retweets (how many times other users shared the tweet). These data were also taken into consideration in our work, being added to the Bag of Words in order to provide information that could be relevant to train the models. The creation of the Bag of Words consisted in the fourth phase of the experiment.

In the fifth phase, the collected data was split in a training set (70%) and a test set (30%). In the sixth phase, we built a logistic regression model with a 0.05 value for the inverse of regularization strength parameter, which was trained using the out-of-sample train set and a 10-fold cross-validation. A deep learning model was also implemented containing one input layer and two convolution layers, and was trained using the out-of-sample train set and a 10-fold cross-validation.

Finally, the seventh and last phase consisted in analyzing the results though metrics that could evaluate the models performance, and perform a comparison of their results. The metrics used to evaluate the results of the polarity detection algorithms were the metrics frequently used in the literature to evaluate Information Retrieval (IR) systems, such as accuracy, precision, recall and F1-Score, respectively defined by Equations 1, 2, 3 and 4.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1_Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4}$$

In the Equations 1, 2, and 3, TP indicates the true positives, which means the number of tweet-political-classification pairs that the system correctly identifies as positives. TN indicates the true negatives, which corresponds to the number of tweet-political-classification pairs that the system correctly identifies as negatives. FP indicates the false positives, which means the number of tweet-political-classification pairs that the system falsely identifies as positives. Lastly, FN indicates the false negatives, which corresponds to the number of tweet-political-classification pairs that the system falsely identifies as negatives.

5 RESULTS

The text classification of tweets in political or non-political content by using logistic regression and deep learning models provided metrics such as accuracy, precision, recall and F1-Score.

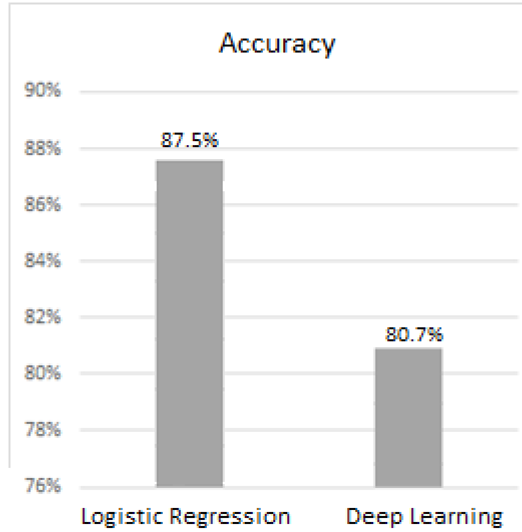


Figure 3: Results for the Accuracy metric.

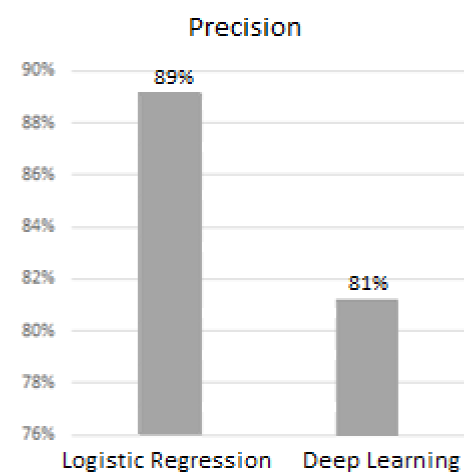


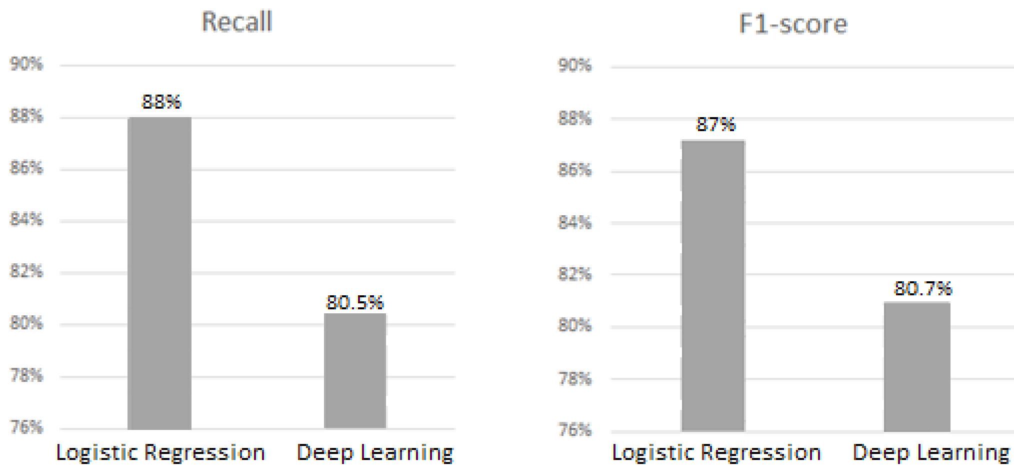
Figure 4: Results for the Precision metric.

Figure 3 shows the values for the accuracy metric obtained with the experiment execution using training and test datasets. As it can be seen, the logistic regression model had a hit rate of 87.5%, while the deep learning model has 80.7%.

Figure 4 shows the experiment results for the precision metric, where the logistic regression model obtained a rate of 89% and the deep learning model 81%. In Figure 5a, we can see the results for the Recall metric, in which the logistic regression model obtained a better result, 88%, against 80.5% of deep learning.

Figure 5b shows the execution results using the training data for the F1-score metric, in which the logistic regression model achieved 87% and the deep learning model had 80.7%.

In addition to performing the experiments with the training and test sets, we also classified the models using a 10-fold cross-validation. Figure 6 shows the results of the metrics evaluated for the



(a) Recall **(b) F1-Score**

Figure 5: Metric's Results

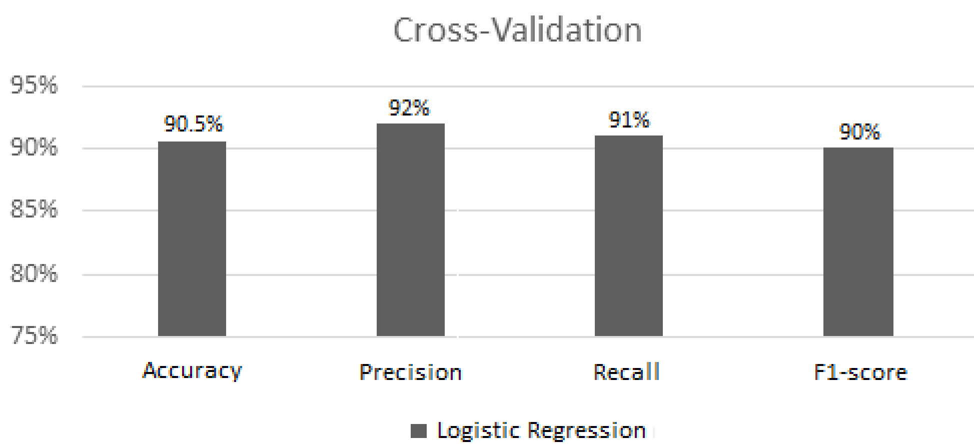


Figure 6: 10-fold cross-validation results for the Logistic Regression model.

logistic regression model, while Figure 7 shows the cross-validation results for the deep learning model.

We can observe that the logistic regression model also demonstrated better results in all the metrics evaluated on the 10-folds cross-validation. For the accuracy metric, the logistic regression model obtained a rate of 90.5%, while the deep learning model obtained 83.8%. For the precision metric, the logistic regression model obtained a result of 92%, against 85% of the deep learning. For the Recall metric, we obtained a 91% rate for logistic regression and an 84% rate for deep learning, while the F1-score was 90% for logistic regression and 83.6% for deep learning.

The two models presented good results for the textual classification, with a hit rate higher than 80%. However, the logistic regression model proved to be a better solution to classify our database than the deep learning model. Logistic regression had a superior result for all the evaluated metrics. In addition to obtaining better results on the classification accuracy rate, the logistic regression model was more fast and efficient, consuming about 40% of the computer's memory during its execution, which lasted approximately thirty minutes. The deep learning model took about 240 minutes to classify the data, consuming approximately 80% of the memory during its execution.

6 CONCLUSION

The textual mining area has evolved a lot, but still has many motivating challenges to overcome. In this article, supervised classification techniques were applied in tweets to identify those ones that have a political context.

The results obtained were satisfactory, since the metrics that were used to measure the performance of the two built models had results above 80%. Such results motivate to deepen this research,

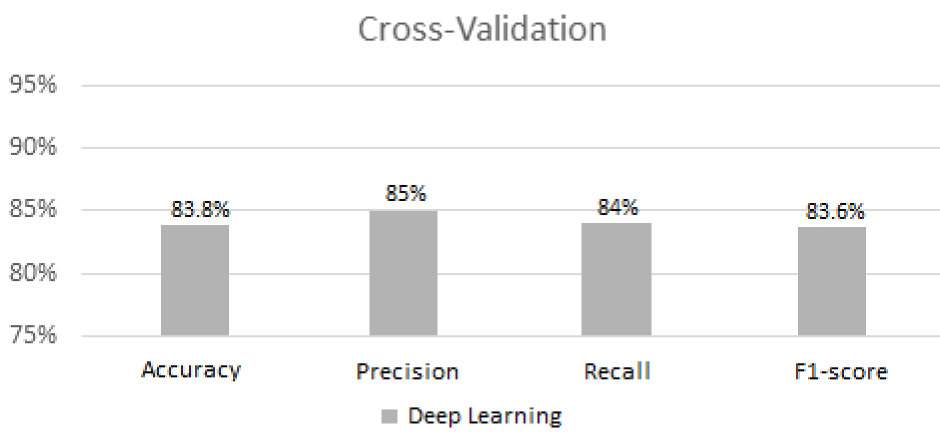


Figure 7: 10-fold cross-validation results for the Deep Learning model.

testing other techniques of supervised classification, as well as adding new features to the models created. When comparing the two supervised classification algorithms we used (logistic regression and deep learning), we realized that the performance of the first one, based on the metrics we used, was higher than the second one. We could conclude that for a corpus with non-bulky data and non-parallel processing, deep learning is not the most appropriate technique to be employed.

As future work, the vast field of textual mining applied to social networks presents several challenges, such as exploring social networks relationships to try better identifying the political bias of a group that interacts with political matters. For example, we can determine groups by employing unsupervised classification and invest on the detection of fake news, sarcasm and irony. In addition, we intend to incorporate techniques of sentiment analysis to identify the sentiment polarity in a political post. Sentiment polarity is an overall content classification in positive, negative or neutral sentiment. Finally, we intend to use other supervised classification algorithms such as Naive Bayes, decision trees, SVM, and compare them to the results obtained so far.

ACKNOWLEDGMENTS

The authors thank the CNPq - Brazilian Research Council for partially funding this research.

REFERENCES

[1] Abebe Abeshu and Naveen Chilamkurti. 2018. Deep learning: the frontier for distributed attack detection in Fog-to-Things computing. *IEEE Communications Magazine* 56, 2 (2018), 169–175.

[2] Charu C. Aggarwal and ChengXiang Zhai. 2012. *A Survey of Text Clustering Algorithms*. Springer US, Boston, MA, 77–128. https://doi.org/10.1007/978-1-4614-3223-4_4

[3] Liliya Akhtyamova, John Cardiff, and Andrey Ignatov. 2017. Twitter Author Profiling Using Word Embeddings and Logistic Regression. *Proceedings of Conference and Labs of the Evaluation Forum - CLEF 2017.*

[4] Ika Alfina, Dinda Sigmawaty, Fitriasari Nurhidayati, and Achmad Nizar Hidayanto. 2017. Utilizing Hashtags for Sentiment Analysis of Tweets in The Political Domain. In *Proceedings of the 9th International Conference on Machine Learning and Computing*. ACM, 43–47.

[5] Andre Luiz Firmino Alves, Claudio De Souza Baptista, Anderson Almeida Firmino, Maxwell Guimaraes De Oliveira, and Anselmo Cardoso De Paiva. 2014. A Comparison of SVM Versus Naive-Bayes Techniques for Sentiment Analysis in Tweets. In *Proceedings of the 20th Brazilian Symposium on Multimedia and the Web - WebMedia 14.* https://doi.org/10.1145/2664551.2664561

[6] Rafael T. Anchieta and Raimundo S. Moura. 2017. Exploring Unsupervised Learning Towards Extractive Summarization of User Reviews. In *Proceedings of the 23rd Brazillian Symposium on Multimedia and the Web - WebMedia 17.* https://doi.org/10.1145/3126858.3131583

[7] Matheus Araujo, Julio Reis, Adriano Pereira, and Fabricio Benevenuto. 2016. An evaluation of machine translation for multilingual sentence-level sentiment analysis. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing - SAC 16.* https://doi.org/10.1145/2851613.2851817

[8] Farzindar Atefeh and Wael Khreich. 2015. A survey of techniques for event detection in twitter. *Computational Intelligence* 31, 1 (2015), 132–164.

[9] Ricardo Baeza-Yates, Berthier Ribeiro-Neto, et al. 1999. *Modern information retrieval*. Vol. 463. ACM press New York.

[10] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2016. Enriching word vectors with subword information. *arXiv preprint arXiv:1607.04606* (2016).

[11] Alberto Carreras Mesa, Mari Carmen Aguayo-Torres, Francisco J. Martin-Vega, Gerardo Gómez, Francisco Blanquez-Casado, Isabel M. Delgado-Luque, and Jose Entrambasaguas. 2018. Link abstraction models for multicarrier systems: A logistic regression approach. *International Journal of Communication Systems* 31, 1 (2018).

[12] Moon-tong Chan, Dalei Yu, and Kelvin K. W. Yau. 2015. Multilevel cumulative logistic regression model with random effects: Application to British social attitudes panel survey data. *Computational Statistics & Data Analysis* 88 (2015), 173–186.

[13] Eric Fernandes de Mello Araújo and Dave Ebbelaar. 2018. Detecting Dutch political tweets: A classifier based on voting system using supervised learning. In *10th International Conference on Agents and Artificial Intelligence, ICAART 2018.* SciTePress.

[14] Maite Giménez, Tomás Baviera, Germán Llorca, José Gámir, Dafne Calvo, Paolo Rosso, and Francisco Rangel. 2017. Overview of the 1st classification of spanish election tweets task at ibereval 2017. In *Notebook Papers of 2nd SEPLN Workshop on Evaluation of Human Language Technologies for Iberian Languages (IBEREVAL), Murcia, Spain, September*, Vol. 19.

[15] Frank E. Harrell. 2001. Regression modeling strategies, with applications to linear models, survival analysis and logistic regression. In *Springer Series in Statistics*. Springer.

[16] Abdalraouf Hassan and Ausif Mahmood. 2018. Convolutional Recurrent Deep Learning Model for Sentence Classification. *IEEE Access* 6 (2018), 13949–13957.

[17] Marti Hearst. 2003. What is text mining. *SIMS, UC Berkeley* (2003).

[18] Armand Joulin, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. 2016. Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759* (2016).

[19] Dan Jurafsky and James H. Martin. 2009. Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition. In *Prentice Hall series in artificial intelligence*. Prentice Hall, Pearson Education International, 1–1024.

[20] Ankush Khandelwal, Sahil Swami, Syed Sarfaraz Akhtar, and Manish Shrivastava. 2017. Classification Of Spanish Election Tweets (COSET) 2017: Classifying Tweets Using Character and Word Level Features. In *IberEval@ SEPLN*. 49–54.

[21] Yuancheng Li, Rong Ma, and Runhai Jiao. 2015. A hybrid malicious code detection method based on Deep Learning. *International Journal of Software Engineering*

and its Applications 9, 5 (2015), 205–216.

[22] George Loukas, Tuan Vuong, Ryan Heartfield, Georgia Sakellari, Yongpil Yoon, and Diane Gan. 2018. Cloud-based cyber-physical intrusion detection for vehicles using Deep Learning. *IEEE Access* 6 (2018), 3491–3508.

[23] Kevin P. Murphy. 2012. Machine Learning: A Probabilistic Perspective. Adaptive Computation and Machine Learning. In *Adaptive Computation and Machine Learning series*. MIT press.

[24] Arman Khadjeh Nassirtoussi, Saeed Aghabozorgi, Teh Ying Wah, and David Chek Ling Ngo. 2015. Text mining of news-headlines for FOREX market prediction: A Multi-layer Dimension Reduction Algorithm with semantics and sentiment. *Expert Systems with Applications* 42, 1 (2015), 306–324.

[25] Nnamdi I. Nwulu. 2017. Evaluation of machine learning classification algorithms & missing data imputation techniques. In *International Artificial Intelligence and Data Processing Symposium (IDAP)*. IEEE, 1–5.

[26] Erik Tjong Kim Sang, Herbert Kruitbosch, Marcel Broersma, and Marc Esteve Del Valle. 2017. Determining the function of political tweets. In *IEEE 13th International Conference on e-Science*. IEEE, 438–439.

[27] Sandro Skansi. 2018. *Introduction to Deep Learning: From Logical Calculus to Artificial Intelligence*. Springer.

[28] Karen Sparck Jones. 1972. A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation* 28, 1 (1972), 11–21.

[29] Hastie Trevor, Robert Tibshirani, and Jerome H. Friedman. 2009. *The elements of statistical learning: Data Mining, Inference, and Prediction*. New York, NY: Springer.

[30] David Watts, K. M. George, T. K. Ashwin Kumar, and Zenia Arora. 2016. Tweet sentiment as proxy for political campaign momentum. In *IEEE International Conference on Big Data*. IEEE, 2475–2484.

[31] Xiang Zhu, Yuanping Nie, Songchang Jin, Aiping Li, and Yan Jia. 2015. Spammer detection on online social networks based on logistic regression. In *International Conference on Web-Age Information Management*. Springer, 29–40.