Applied Artificial Intelligence
An International Journal

ISSN: 0883-9514 (Print) 1087-6545 (Online) Journal homepage: https://www.tandfonline.com/loi/uaai20

Analyzing the Brazilian Financial Market through Portuguese Sentiment Analysis in Social Media

A. E. O. Carosia, G. P. Coelho & A. E. A. Silva

APPLIED ARTIFICIAL INTELLIGENCE
2020, VOL. 34, NO. 1, 1–19
https://doi.org/10.1080/08839514.2019.1673037

Taylor & Francis
Taylor & Francis Group

Check for updates

Analyzing the Brazilian Financial Market through Portuguese Sentiment Analysis in Social Media

A. E. O. Carosia[a,b], G. P. Coelho[b], and A. E. A. Silva[b]

[a]Federal Institute of São Paulo (IFSP), São João da Boa Vista, Brazil; [b]School of Technology (FT), University of Campinas (UNICAMP), Limeira, Brazil

ABSTRACT
According to the Efficient Market Hypothesis, financial market movements are dependent on news and external events that have a significant impact on the market value of companies. Thus, a great amount of applications has arisen to explore this knowledge through automatic sentiment and opinion extraction. The technique known as Sentiment Analysis (SA) aims to analyze opinions, sentiments, and emotions present in unstructured data, leading many papers to address the impact of news and social media publications on the financial market. However, the literature lacks works considering the effects of sentiment available on social media and its impacts on the Brazilian stock market. This work aims to conduct a study of the Brazilian stock market movement through SA in Twitter considering tree perspectives: (i) absolute number of tweet sentiments; (ii) tweets sentiments weighted by favorites; and (iii) tweets sentiments weighted by retweets. The analyzed period was the Brazilian electoral period of 2018 (01-Oct-2018 to 31-Dec-2018). In this paper, we first developed a comparison study with SA Machine Learning techniques (Naive Bayes, Support Vector Machines, Maximum Entropy, and Multilayer Perceptron) and then applied the best algorithm to establish the relations between sentiments and the Brazilian stock market movement considering different time frames (windows sizes). Results indicate that Multilayer Perceptron was the best technique to perform SA in Portuguese. In addition, we observed that the predominant sentiment in social media relates to the stock market movement, improving accuracy as long as windows sizes are increased.

Introduction

There is a great research interest about the viability of predicting the behavior of the financial market. For a long time, investors have accepted the Efficient Market Hypothesis, which says that high returns in the financial market can not be obtained by studying the past value of stock prices (Malkiel and Fama 1970). However, there are economists who claim that

CONTACT A. E. O. Carosia arthuremanuel.carosia@ifsp.edu.br Federal Institute of São Paulo (IFSP), São João da Boa Vista, Brazil

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/uaai.

the stock prices in the financial market may be at least partially predictable (Malkiel 2003). In fact, there are many studies in the scientific literature that present algorithms to predict the financial market, as indicated by Atsalakis and Valavanis (2009), which reviewed over 100 articles published in the area.

Financial market forecasting methods are based on the following techniques: (i) Fundamental Analysis; (ii) Technical Analysis; (iii) Time Series Forecasting with Traditional Models; and (iv) Machine Learning. Fundamental Analysis is based on statistics, projections, conditions of supply and demand of goods and services, and fundamentals of economics and business. Technical Analysis, also known as graphical analysis, attempts to predict market behavior by deriving patterns based on the study of graphs that describe historical data (Francis and Kirzner 1991). Forecasting time series with traditional models is an attempt to design forecasting models to track patterns in historical data. Finally, the Machine Learning approach aims at using a set of data to learn patterns that approximate the generating function of the data.

In addition, studies based on financial behavior show that emotions can influence investment decisions (De Long et al. 1990). Schumaker et al. (2012) show that sentiments contained in news or financial disclosures affect the market. Thus, particularly due to the growth of social media, it is also possible to notice great interest in Sentiment Analysis (SA) to predict stock market movements (Bollen, Mao, and Zeng 2011). This technique has the purpose of analyzing texts computationally and assigning them to a given class. Classes can be positive or negative, according to the text content. Finally, some techniques are applied with the objective of associating momentary sentiment with the financial market movement. Examples of techniques used for SA are: lexical analysis (Li et al. 2014), Machine Learning, and, more recently, Deep Learning, which is the application of Artificial Neural Networks (ANN) with several layers (Kraus and Feuerriegel 2017).

Although the literature is increasing, the studies present certain limitations: Assis et al. (2018) reviewed over 100 papers and verified that only 3 studies used data from the Brazilian market. In addition, as far as we were able to find, there are no studies that used SA in Portuguese and verified its relation with the Brazilian stock market.

Thus, this study aims to analyze the Brazilian stock market movement through SA in Twitter, during the Brazilian electoral period of 2018 (01-Oct -2018 to 31-Dec-2018). During that period, the country presented great political and economic instability in its first presidential race after a process of impeachment of the former president, Dilma Roussef, in addition to the discovery of several reports of corruption cases in the public sector. The results from this period have brought uncertainty to the market (Econômico 2018), which was evidenced by a steep valorization of the USD (Economia 2018f,e), which reached a historic high of BRL 4,00, in addition to a recurring decline in the Bovespa index (Ibovespa) (Economia 2018b,c). On the other hand, the period also showed abrupt USD devaluations, as well as records of transactions and points on the stock market (Economia 2018a,d; Globo 2018). Facts such as these show the economic complexity that Brazil had been through during this period.

In summary, this paper presents the following contributions:

- Study of different Machine Learning techniques for SA in Portuguese texts.
- Study of the influence of tweets on the movement of the Brazilian stock market according to three perspectives: (i) absolute number of tweets sentiments; (ii) tweets sentiment weighted by number of *favorites* (FAVs) obtained from social media users; and (iii) tweets sentiment weighted by number of *retweets* obtained from social media users.
- Study of different time frames (daily windows) to verify the impact of the predominant sentiment in Twitter and its impacts on the financial market. Each daily window corresponds to an average of predominant sentiment in the market calculated through a certain number of days.

This paper is structured as follows. Section 2 presents the theoretical basis, Section 3 presents the proposed methodology and Section 4 presents the results and discussions. Finally, Section 5 concludes this paper.

Theoretical Foundation

Nowadays, the World Wide Web (WWW), through news websites, blogs, or social networks, is a great source of information. Thus, several applications have arisen to explore this knowledge through the extraction of opinions and sentiments with the Sentiment Analysis (SA) technique (Martins, Pereira, and Benevenuto 2015).

This section presents the theoretical basis of this study and is divided into two subsections: Section 2.1 presents the SA technique, while Section 2.2 presents the studies that involve SA and the financial market and highlights the differences between our study and related studies in the literature.

Sentiment Analysis

Sentiment Analysis is a Natural Language Processing (NLP) technique that aims to analyze opinions, sentiments, and emotions present in unstructured data (Hussein 2018; Pang and Lee et al. 2008). Thus, a commonly performed task is the identification of three classes of sentiments within a document: positive, neutral, and negative. Thereby, the construction of sentiment analyzers (i.e., classifiers) usually uses two techniques: lexical methods and Machine Learning algorithms.

Lexical approaches depend on linguistic knowledge and are robust when ported to other domains (Taboada et al. 2011). This approach uses a list containing a set of word pairs and the associated polarity of sentiments. Although there are techniques to increase simpler lexical methods, as presented by Taboada et al. (2011), it is difficult to achieve results that overcome Machine Learning techniques (Avanço and Volpe Nunes 2014).

The most commonly used Machine Learning techniques for SA, according to Sun, Luo, and Chen (2017), are: Naive Bayes, Maximum Entropy (ME) and Support Vector Machines (SVM) (Mocherla, Danehy, and Impey 2017). Both Naive Bayes and ME calculate the association probability of document *d* to a certain class *c*. However, while the first uses Bayes' Theorem to calculate the probabilities, the second technique uses a function called maximum entropy to perform the classification. On the other hand, the SVM technique searches for a hyperplane that separates the documents in different classes, maximizing the distance between them. These three techniques have shown great accuracy in text classification, according to Pang, Lee, and Vaithyanathan (2002).

Artificial Neural Networks (ANNs) have been attracting considerable attention in the scientific literature due to promising results in classification and prediction tasks (Kraus and Feuerriegel 2017). An ANN is composed of interconnected artificial neurons, which are mathematical processing units that receive several inputs and generate outputs. These artificial neurons simulate the behavior of biological neurons (Haykin 1994). Their results in text classification have been promising, especially with Deep Neural Networks (Kraus and Feuerriegel 2017).

It is worth highlighting the innovative aspect of this study in the sense that, so far, there is few works on opinion mining for Brazilian Portuguese, although there are some initiatives in this sense through lexical analyzers. Avanço and Volpe Nunes (2014) present a lexical approach to SA in Portuguese through a dictionary that shows the polarity of the words according to their sentiment, as well as using linguistic knowledge about the context in which the word is included in the sentence. On the other hand, Martins, Pereira, and Benevenuto (2015) carry out the task of SA in Portuguese through a methodology called SentiPipe, which combines a set of lexical methods adapted to Portuguese.

Related Studies

Great interest can be found in the literature to predict the financial market through SA. The study developed by Bollen, Mao, and Zeng (2011) aims to measure the sentiment of Twitter posts and correlate it with the financial market movement. For this, two sentiment monitoring tools were used: OpinionFinder, which measures the feeling as positive and negative, and

Google-Profile of Mood States (GPOMS) which measures humor in 6 dimensions (Calm, Alert, Sure, Vital, Kind, and Happy). In their study, public Tweets from February 28 to December 19, 2008, were used and the results showed a precision of 86.7% in the correlation with the daily changes closing values of the Dow Jones Industrial Average (DJIA).

Schumaker et al. (2012) studied the correlation between financial news and financial market movements. Their study used the Arizona Financial Text System enhanced through the OpinionFinder tool. Data from financial and market news of the S&P 500 companies between October 26 and November 28, 2005, were used. The results showed that news articles can predict the financial market with a 59% accuracy and lead to a 3.3% profit in the stock market.

The study developed by Arias, Arratia, and Xuriguera (2013) investigated whether an indicator of public sentiment extracted from Twitter's daily messages can improve the prediction of social, economic, or commercial indicators. The technique used in this study was the summary-tree, which is based on decision trees. The data source used was Twitter, from March 2011 to 2013. The results showed that nonlinear models leverage Twitter data by predicting financial trends, while linear ones fail systematically to predict any type of financial time series. Concerning the forecast of box office revenue trends, the results have shown that SVM presented the best results.

Makrehchi, Shah, and Liao (2013) developed a study that considers the sentiment of daily Tweets and makes predictions about the next movement of the financial market. The study used a lexical SA approach by counting specific words or phrases. In addition, it also presented the automatic extraction of labels, positive or negative, associated with the financial market for use in classification tasks through supervised learning. The data source used was the S&P 500 index, from March to June 2012. As a result, the authors showed that the aggregated sentiment of the day correlates with market performance. The developed proposal was tested as an investment strategy in the stock market and obtained 20% of the return in four months.

Yoshihara et al. (2014) presented a SA approach with the use of Deep Learning through a combination of Recurrent Neural Networks (RNN) and Deep Belief Networks (DBN) to predict the financial market. For this, SA was applied to newspaper articles from ten companies of the Nikkei index, from 1999 to 2008. The financial market movement predicted through SA showed that the error rate was reduced from 47.30% to 40.05%, on average, when compared to SVM and DBN approaches.

Li et al. (2014) developed a quantitative trading strategy through the impact of the media on stock markets. The proposed methodology extract nouns with a Part of Speech tagger and, to perform SA, words that indicate sentiments are previously labeled as positive or negative according to stock market information. The data source used came from several financial websites from China. The study showed that the main information from news articles about a particular company can enrich the investors' knowledge and affect their trading activities. Also, the study found that public sentiment causes emotional fluctuations in investors and can impact their decision-making process. Finally, the paper demonstrated that the impact of the media on companies varies according to the company's characteristics and the news content.

Nofer and Hinz (2015) measured the sentiment in Twitter posts and verified if they correlated with the financial market. They used the German version of the Profile of Mood States (POMS), which consists of 19 adjectives that belong to 5 dimensions of humor (sadness, hopelessness, fatigue, anger, and positive mood). The data source consisted of German Tweets from January 1, 2011, to March 17, 2012. The study showed that the correlation between tweets and the financial market exists when considering the number of Twitter followers in each publication. With the proposed approach, the stock portfolio obtained up to 36% of the return within a six-month period in the German financial market.

Zhao et al. (2016) presented a prediction technique for financial market movements through SA of the social network Sina Weibo. The technique used for SA was lexical analysis through selection of characteristics and words in conjunction with SVM. In addition to the textual data, the authors also used historical data to predict the market value of the Shanghai Composite Index from September to December 2015. On average, the SA presented a precision of 62% to 68%, while the market forecast showed 53% to 60% accuracy, on average.

Feuerriegel and Prendinger (2016) developed negotiation strategies that use textual news for decision-making based on new information entering the market. To analyze the news, a rule and dictionary-based approach was proposed. As a source of data, regulated announcements of companies in English were used from January 2004 to June 2011, in addition to data from the CDAX financial index. Their study also presented a proposal to a news-based investment decision support system.

The study carried out by Yan et al. (2016) aimed to establish a relationship between the Chinese stock market and Chinese microblogging services. It used two techniques of SA focused on the Chinese language: ROST Content Mining and C-POMS. In addition, it used a Probabilistic Neural Network for financial market prediction. The textual data were collected from Chinese microblogs from March to June 2014. As a result, SVMs technique proved to be better in the Chinese financial market prediction than the Probabilistic Neural Networks.

Kraus and Feuerriegel (2017) showed the use of deep learning in financial disclosures in addition to transfer learning with a corpus of different contexts. As data source, the study used 13,135 regulated advertisements of German companies. The results reveal a greater precision in the prediction of stock market movement using deep learning techniques. As a highlight, it showed that deep learning produces a 6.8% improvement in the results, while deep learning with transfer learning produces an improvement of 7.1%.

Vargas, de Lima, and Evsukoff (2017) used deep learning to predict the movement of the S&P 500 index through news headlines. For SA, the study used deep learning with a composition of the Convolutional Neural Network (CNN) and RNN architectures. Thus, 106,494 news items from the Reuters site were collected from October 20, 2006, to November 21, 2013, along with the historical series of the S&P 500 index. The results showed that CNN may be better than RNN to analyze the semantics of the texts and the RNN was better than CNN in capturing context information and modeling the temporal characteristics of the market.

Hájek (2018) combined the Bag of Words and SA techniques applied to annual reports of companies with financial indicators to predict abnormal stock returns, i.e., accumulated returns above the return of the market portfolio. The SA was developed through two dictionaries, while the market forecast was made by deep learning with Multilayer Perceptrons (MLP). The data used came from news of 1,402 US companies listed on the New York Stock Exchange (NYSE) and on Nasdaq with a stock price of at least USD 3.00 in 2013. The experiments have shown that this method has results similar to the Naive Bayes algorithm, but outperforms other traditional Machine Learning algorithms in predicting abnormal stock returns.

As previously mentioned, there are few studies about SA and its correlation with the Brazilian market. Santos, Laender, and Pereira (2015) show a correlation between financial news reports and the stock market. However, the study presents only statistical correlations, without developing SA activities. The study carried out by Lima et al. (2016) describes the use of SVM for SA on tweets about Petrobras. However, it fails to present a correlation between the collective sentiment identified about the company and the financial market movement of the company's shares.

In general, we found that there are few proposals for SA in Portuguese using Machine Learning. Moreover, the existing methods for English produce inferior results when applied to contents in Portuguese (Martins, Pereira, and Benevenuto 2015) approach due to particular characteristics of the Portuguese language, such as: irregular verbs, typical expressions of the language, double entendres, metaphors, among others. In addition, the process of automatic translation from one language to another is an activity prone to errors. Therefore, this study focuses on building a Machine Learning SA approach to Portuguese.

This study contributes to the field in the following ways:

- It presents the application of MLPs to SA in Portuguese tweets, which presents superior performance compared to the most used techniques of Machine Learning for SA: Naive Bayes, ME and SVM;

- It establishes a correlation between collective sentiment on Twitter in Brazil and financial market movements;
- It presents three perspectives to correlate SA and the financial market: (i) daily sentiment in absolute number of tweets; (ii) daily sentiment weighted by Favorites (FAVs); and (iii) daily sentiment weighted by Retweets (RTs);
- It presents an analysis of different time frames (windows) to verify the period in which the average sentiment in the social media has greater impact in the financial market;
- It performs SA of 16,352 unlabeled tweets over a politically and economically troubled period in Brazil: the 2018 electoral period. This contribution allows the identification of the predominant sentiment throughout the period in the country, besides allowing its correlation with the financial market movement.

Methodology

This study aims to predict the Brazilian financial market movement through Sentiment Analysis (SA) of Twitter social media during the period of the presidential elections of 2018 (01-Oct-2018 to 31-Dec-2018). For this, the following steps were developed: (i) implementation of an SA module for the Portuguese language; (ii) comparison of Machine Learning techniques (Naive Bayes, ME, SVM, and MLP) to determine the most appropriate SA for the financial domain; (iii) application of the SA module in unlabeled Tweets collected from 01-Oct-2018 to 12-Dec-2018; (iv) analysis of the predominant sentiment through three proposals (daily sentiment number, daily sentiment number weighted by Favorites (FAVs), and daily sentiment number weighted by Retweets (RTs)); and (v) study of the association between the predominant sentiment in social media and the financial market movement.

The proposed architecture to predict stock market using SA is presented in Figure 1. There are three main modules: Preprocessor, Sentiment Classifier, and Sentiment Average. The Preprocessor module prepares the input document to be classified by the sentiment classifier. This task consists of three steps: (i) data cleaning (i.e., removal of links, special symbols, emoticons, etc.) and removal of stop words using the Natural Language Toolkit framework (NLTK, available at http://www.nltk.org/), which are words that do not add useful information to the sentence; (ii) application of the stemming technique, which transforms each word into its stem (Ravi and Ravi, 2015); and (iii) generation of the vector of characteristics using the bag of words technique (Loughran and Mcdonald, 2011). The output of the Preprocessor module is used by the Sentiment Classifier module, which aims to classify the sentiment of each

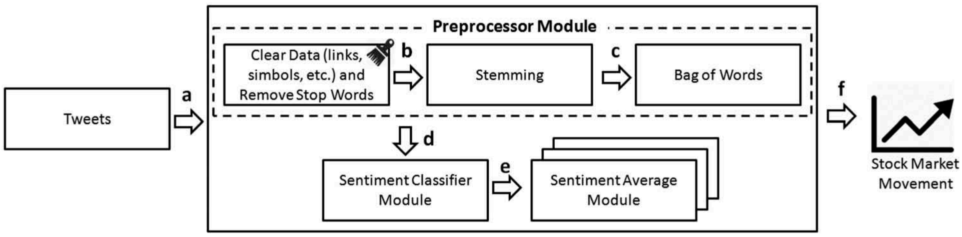

Figure 1. Proposed architecture to predict stock market using SA.

document individually through a Machine Learning technique (Naive Bayes, ME, SVM or MLP, as presented in Section 4.1).

Finally, the Sentiment Average module calculates the daily sentiment of a certain topic through a defined period (i.e., days, weeks, months, etc.). This module considers the predominant sentiment using three perspectives, as follows:

(1) Absolute number of sentiments: calculated considering that the predominant sentiment (positive or negative) is the one which presents highest number of tweets. The calculation is performed according to Eq. (1), in which *nPos* means the number of positive tweets and *nNeg* means the number of negative tweets.

$$Sentiment = \frac{(nPos - nNeg)}{(nPos + nNeg)} \tag{1}$$

(2) Number of sentiments weighted by Favorites (FAVs): calculated considering the number of likes of each tweet, according to Eq. (2), in which $fav(tweetPos_i)$ means the number of FAVs in the positive $tweet_i$ ($tweetPos_i$), $fav(tweetNeg_i)$ means the number of FAVs in the negative $tweet_i$ ($tweetNeg_i$), *nPos* means the number of positive tweets, and *nNeg* means the number of negative tweets.

$$Sentiment = \frac{\sum_{i=0}^{nPos} fav(tweetPos_i) - \sum_{i=0}^{nNeg} fav(tweetNeg_i)}{\sum_{i=0}^{nPos} fav(tweetPos_i) + \sum_{i=0}^{nNeg} fav(tweetNeg_i)} \tag{2}$$

(3) Number of sentiments weighted by retweets (RTs): calculated considering the number of reproductions of each tweet, according to Eq. (3), in which $rt(tweetPos_i)$ means the number of RTs in the positive $tweet_i$ ($tweetPos_i$), $rt(tweetNeg_i)$ means the number of RTs in the negative $tweet_i$ ($tweetNeg_i$), *nPos* means the number of positive tweets and *nNeg* means the number of negative tweets.

$$Sentiment = \frac{\sum_{i=0}^{nPos} rt(tweetPos_i) - \sum_{i=0}^{nNeg} rt(tweetNeg_i)}{\sum_{i=0}^{nPos} rt(tweetPos_i) + \sum_{i=0}^{nNeg} rt(tweetNeg_i)} \tag{3}$$

For each of these three perspectives, the final result will be a value between −1 and 1, in which values between −1 and 0 mean that the predominant sentiment is negative and between 0 and 1 means that the predominant sentiment is positive. As a result, it is possible to have a perception of the general sentiment about a certain topic.

Consider, for example, Table 1. There are 150 tweets classified as positive and 160 tweets classified as negative. First, considering the perspective of the absolute number of tweets sentiment, the predominant sentiment is negative, as presented in the following equation:

$$\frac{+150 - 160}{150 + 160} = -0.03 \tag{4}$$

On the other hand, consider the perspective of tweets sentiment weighted by FAVs. Suppose 2 FAVs for each positive tweet and 1 FAV for each negative tweet, this will result in a predominantly positive sentiment, as follows:

$$\frac{150 * 2 - 160 * 1}{150 * 2 + 160 * 1} = +0.30 \tag{5}$$

The same idea can be applied considering the perspective of the tweets sentiment weighted by RTs. Suppose 3 RTs for each positive tweet and 1 RT for each negative tweet, and the predominant sentiment will be also positive, according to the following equation:

$$\frac{150 * 3 - 160 * 1}{150 * 3 + 160 * 1} = +0.47 \tag{6}$$

Results

This section presents the results obtained in this study according to the following steps: (i) identification of the best Machine Learning technique to carry out SA in Portuguese (considering the Naive Bayes, SVM, Maximum Entropy, and MLP); and (ii) analysis of the predominant sentiments over different time frames (windows) and their respective association with the stock market movement.

Table 1. Example of SA, considering three perspectives: absolute number of tweets, FAV-weighted tweets, and RT-weighted tweets.

Perspective	Sentiment	Number of Tweets	Weight	Predominant Sentiment
Absolute number of tweets sentiments	Positive	150		**Negative**
	Negative	160		
Tweets sentiments weighted by FAVs	Positive	150	2	**Positive**
	Negative	160	1	
Tweets sentiments weighted by RTs	Positive	150	3	**Positive**
	Negative	160	1	

Sentiment Analysis Module

The results presented in this section show the procedure to identify the best Machine Learning technique for SA in Portuguese. For this, data were collected from tweets and news, in Portuguese, related to a company of the automobile industry and classified by specialists adopting the investor's point of view. Two classes were considered: positive and negative. The database is composed of 2,132 financial news from various news websites and 11,027 tweets (Martins, Pereira, and Benevenuto 2015).

The following Machine Learning techniques were compared: Naive Bayes, ME, SVM, and MLP. For this, the framework Scikit-Learn (https://scikit-learn.org) was used. The Naive Bayes algorithm was used in its Multinomial version (Kibriya et al. 2004), suitable for classification tasks with discrete attributes. The ME algorithm was implemented using the Logistic Regression algorithm. The Multiclass version of the SVM algorithm was also used, which is suitable for classification tasks with several classes. Finally, the MLP was configured with 1 hidden layer and 100 neurons with ReLU activation function in the hidden layer and sigmoid activation function in the output layer, and learning rate of 0.001 for 25 epochs. The parameters for the algorithms were defined empirically, evaluated through the validation set according to the configuration that brought better classification results. For this, a grid-search was performed to investigate the best MLP configuration considering the following values: (i) 1, 3, 5, and 10 hidden layers; (ii) 1, 5, 10, 50, and 100 neurons per layer; (iii) 0.1, 0.01, and 0.001 for learning rate; and (iv) 1, 10, 50, and 100 epochs.

The tests were performed according to k-fold cross-validation ($k = 10$) and the results of the average of the executions are summarized in Table 2. To analyze the results of SA, we used the following metrics: F1-Score and accuracy. *F1-Score* is defined as a harmonic average between precision (p) and recall (r), given by Eq. (7).

$$F1 - Score = 2 * \frac{(p * r)}{(p + r)} \tag{7}$$

Precision shows the relationship of properly classified instances in a class and all instances classified as belonging to the same class; recall measures the

Table 2. Results presented by the Naive Bayes, ME, SVM, and MLP techniques.

Technique	Precision	Recall	F1-Score	Accuracy
Naive Bayes	80.4%	55.9%	55.3%	80.6%
ME	77.8%	62.4%	65.3%	84.1%
SVM	77.0%	55.3%	54.9%	79.7%
MLP	75.3%	71.3%	72.7%	84.8%

relationship between instances classified as pertaining to a given class and the total number of instances that are actually part of this class. Finally, accuracy is the number of correctly classified instances divided by the total number of instances.

Tables 3 and 4 present the Mann–Whitney U Test (Corder and Foreman 2014), used to verify the statistical difference between classifiers, for the F1-Score and accuracy measures, respectively. It is possible to see that the observed differences are significant considering a tolerance of 5%. In the tables, for statistically different results, the value 1 was used, while 0 was used otherwise.

The results show that, considering the MLP, F1-Score is statistically different from the other classifiers. However, accuracy of the MLP and ME techniques do not present significant statistical difference. In addition, it is worth noting that the MLP technique was able to obtain better values of F1-Score and accuracy in SA activity, being superior to Naive Bayes, ME and SVM techniques. With these results, this technique was chosen for the Sentiment Analysis Module, to carry out the classification of unlabeled tweets. The results of this task are presented in Section 4.2, along with the association between sentiments and the Brazilian stock market.

Classification of Unlabeled Tweets by the Sentiment Analysis Module

To evaluate the Sentiment Analysis Module, developed with the MLP technique, unlabeled tweets (i.e., without indication of positive or negative sentiments) were collected from Twitter between 01-Oct-2018 and 31-Dec-2018, totalizing 16,352 tweets. Thus, it was possible to verify the evolution of users' sentiments over time regarding the automobile company according to three perspectives: (i) daily sentiment number; (ii) daily sentiment number weighted by Favorites (FAVs);

Table 3. Mann–Whitney statistical test for the F1-Score metric. Zero means no statistically significant difference, whereas one means there is a significant statistical difference (5% tolerance).

Technique	Naive Bayes	ME	SVM	MLP
Naive Bayes		1	0	1
Maximum Entropy	1		1	1
SVM	0	1		1
MLP	1	1	1	

Table 4. Mann–Whitney statistical test for metric accuracy. Zero means no statistically significant difference, whereas one means there is a statistically significant difference (5% tolerance).

Technique	Naive Bayes	ME	SVM	MLP
Naive Bayes		1	1	1
Maximum Entropy	1		1	0
SVM	1	1		1
MLP	1	0	1	

and (iii) daily sentiment number weighted by Retweets (RTs). These perspectives allow the analysis not only of the number of messages published in the social network, but also the relevance of these messages to the users, which is characterized both by the number of followers of a user and by the number of reactions to its messages. In the indicated period, the daily opening and closing stock prices of the studied company varied according to the graph shown in Figure 2.

Figures 3–5 present the evolution of the users' sentiments (positive or negative) regarding the company during the period according to the three perspectives considered in this study. It can be verified that, during the entire period, the predominant sentiment is negative in the three perspectives, thus impacting on the financial market, since it is also possible to verify that the value of the stock prices (Figure 2) fell from US$ 18.14 to US$ 14.6, considering the opening prices, and from US$ 17.99 to US$ 14.46, considering the closing prices.

The results obtained in this analysis are summarized in Table 5, considering different daily time frames (windows) and the three perspectives already mentioned. Each daily window corresponds to a number of day intervals in which the SA is carried out to calculate the predominant sentiment in the

Figure 2. Opening and closing stock prices.

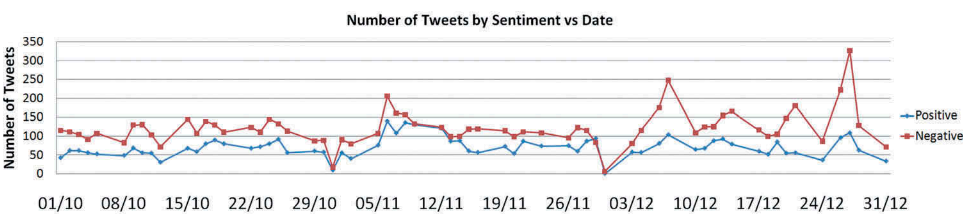

Figure 3. Evolution of users' sentiments (positive or negative) over the period considering the number of tweets.

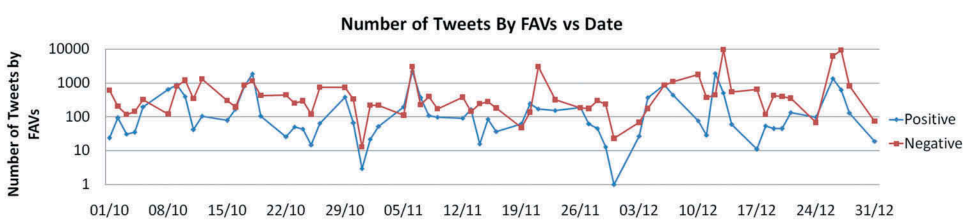

Figure 4. Evolution of users' sentiments (positive or negative) over the period considering the number of FAVs.

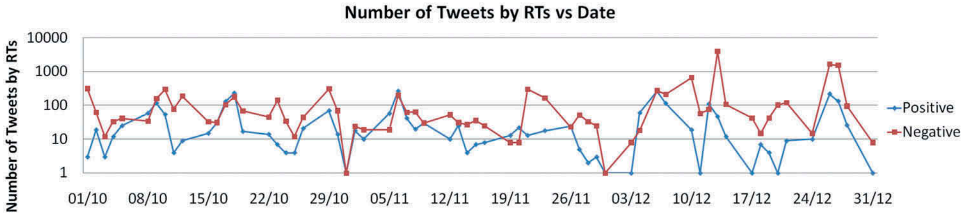

Figure 5. Evolution of users' sentiments (positive or negative) over the period considering the number of RTs.

Table 5. Average result of the movements of the financial market considering different strategies: absolute sentiment number, weighted by FAVs, weighted by RTs.

Window	Analysis	Opening Prices	Closing Prices	Volume
1 Day	# Tweets	58%	50%	55%
	FAVs	55%	47%	55%
	RTs	56%	48%	60%
3 Days	# Tweets	62%	58%	43%
	FAVs	55%	48%	50%
	RTs	58%	55%	47%
5 Days	# Tweets	67%	66%	50%
	FAVs	62%	60%	55%
	RTs	67%	66%	47%
10 Days	# Tweets	64%	64%	53%
	FAVs	64%	64%	53%
	RTs	64%	64%	53%
1 Month	# Tweets	67%	67%	58%
	FAVs	67%	67%	58%
	RTs	67%	67%	58%
2 Months	# Tweets	96%	91%	57%
	FAVs	96%	91%	57%
	RTs	96%	91%	57%
3 Months	# Tweets	100%	100%	0%
	FAVs	100%	100%	0%
	RTs	100%	100%	0%

period. This step allows to verify the number of days that better correlates to the financial market movement. In the analysis, the predominant sentiment in a given period (positive or negative) was compared with the respective movement of the financial market (up or down) and, finally, the average of each correspondence between them was calculated. That is, it is considered a success if the market was high and the preponderant sentiment was positive and vice versa. In detail, the calculation for the period considering different window values was performed as described below.

(1) The calculation of predominant sentiment is performed daily, considering the perspectives determined by Equations 1, 2, or 3, presented in Section 3.
(2) The average of the predominant sentiment is calculated considering a window size of D days.

(3) A comparison of the result calculated in the previous step with the movement of the financial market is made considering the difference in the stock value between day 1 and day D of the window.
(4) If the sentiment is positive and the market has risen, it is considered a hit, as well as if the sentiment is negative and the market is down.
(5) This calculation is repeated day by day, considering the window of size D, until the end of the period analyzed in this study.

Thus, it is possible to observe that, when considering the 1-day window, i.e., the effect on the stock market 1 day after the average sentiment is calculated, the predominant sentiment considering the number of tweets corresponds to the movement of the opening price of the stock market by 58%, the movement of the closing price by 50%, and the movement of transaction volume by 55%. On the other hand, it is possible to observe that the predominant sentiment considering the weighted by FAVs perspective is able to correspond to the movement of the opening price of the financial market by 55%, to the movement of the closing price by 47% and to the transaction volume movement by 55%. Considering the weighted by RTs perspective, it is possible to verify that it follows the movement of the opening price of the financial market by 56%, the movement of the closing price by 48%, and the movement of transaction volume by 60%.

To analyze window sizes greater than 1 day, the following time frames were considered: 3, 5 and 10 days, as well as 1, 2 and 3 months. For this, we considered the average of the sentiments in each period, as well as the variation between the final value and the initial value of the stock prices for analysis of the market movement. In the 3-day window, all perspectives of this study show improvement in their results when compared to the 1-day window. Likewise, it is possible to observe an improvement of the results considering the 5-day window when compared to the 1-day and 3-day windows.

In addition, it is possible to observe that the 10-day window shows a slight reduction in the quality of the results when compared to the 5-day window and it is possible to verify that from the 10-day window the three perspectives started to present similar values in the subsequent windows (1 month, 2 months, and 3 months). This is probably because the predominant sentiment is already spread through the Social Network, having little or no change in its average when considering factors such as number of FAVs and RTs. Finally, the 3-month window, i.e., the complete period, showed: 100% of the hits on the opening price, 100% accuracy on the closing price, and 0% of the accuracy on transaction volume. In this specific case, this result was obtained because the predominant sentiment obtained in the 3-month period was negative and the values of both opening and closing shares fell. On the other hand, the transaction volume increased, which showed an inverse correlation with the predominant sentiment in the period.

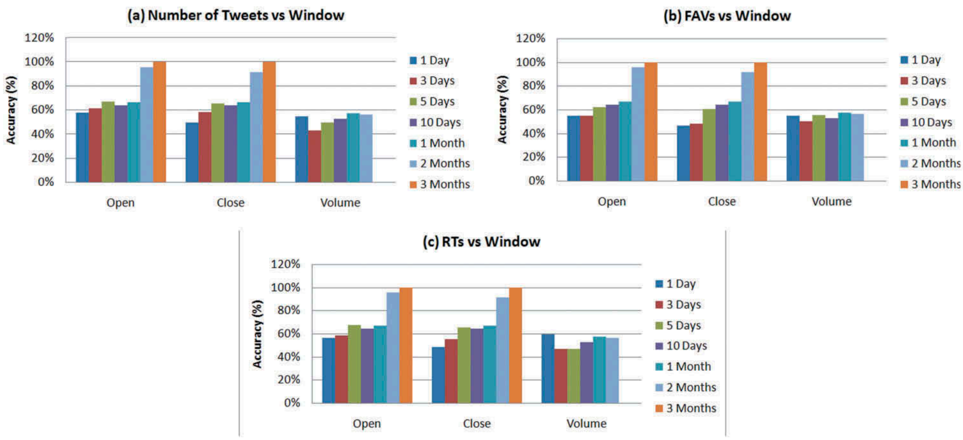

Figure 6. Results considering: a. Number of Tweets sentiments; b. sentiments weighted by FAVs; and c. sentiments weighted by RTs.

The graphs shown in Figure 6a–c illustrate the results obtained in this study considering the perspective of each SA. It is possible to verify that, as the interval of days comprising each window grows, the accuracy of the correspondence between the predominant sentiment and the movement of the financial market increases considering the opening and closing prices.

Conclusion

This paper presented the impacts of the predominant sentiment in the social media Twitter in the Brazilian stock market movements. Among the Machine Learning techniques evaluated, the best technique for SA in Portuguese was the MLPs, which presented better results regarding accuracy and F1-score. This study also showed that it is possible to verify an association between the predominant sentiment in social media and stock market movements within three perspectives: absolute number of tweets sentiments, sentiments weighted by FAVs and sentiments weighted by number of RTs.

In addition, it is worth noticing that the results showed that the accuracy between SA and the opening and closing stock prices becomes more precise as the size of the daily time frames (windows) increases. We believe this happens because, as more days are added to the analysis, it is possible to verify the predominant sentiment in the social media with greater precision, as well as to see its impacts on the financial market. However, we did not observe the improvement of the accuracy, as the window was increased, of the transaction volume prediction.

The differential of this study is the analysis of different Machine Learning techniques for SA in Brazilian Portuguese, the choice of a politically and economically troubled period of the country, as well as the proposal to analyze the predominant sentiment with three perspectives (absolute number of tweets sentiments, tweets sentiments weighted by FAVs, and tweets sentiments weighted by RTs), according to an analysis involving different time frame windows (1 day, 3, 5, and 10 days, 1 month, 2, and 3 months).

In addition, this study analyzed a period before the elections until a later period, being possible to verify some of its impacts, considering as previous information the available sentiment in the social media. The political and economic instability experienced in the country was evidenced by the rise of the dollar in the period and great oscillations in the stock market, due to the period of the first and second rounds of the presidential elections (01-Oct-2018 to 31-Dec-2018) and the uncertainty about the choice of the next president (Economia 2018f,b). In addition, we can see this movement in the financial market by the devaluation of the stock prices of the company analyzed in this study, as well as its predominant negative sentiment throughout the period. These results show that there is an association between the predominant sentiment in social media and the movement of stock prices in the financial market.

Future activities include the analysis of other ANN architectures for the Sentiment Analysis module, considering, in particular, RNN, CNN, and LSTM architectures, as well as investigating the correlation between predominant sentiment and the stock market behavior in a less troubled period of Brazil.

Acknowledgments

The authors thank Espaço da Escrita - Pró-Reitoria de Pesquisa - UNICAMP - for the language services provided. We also thank MSc. Renato Martins (CEFET-MG), PhD Adriano Pereira (UFMG), and PhD Fabrício Benevenuto (UFMG) for providing the databases used for the experiments of this study.

References

Arias, M., A. Arratia, and R. Xuriguera. 2013. Forecasting with twitter data. *ACM Transactions on Intelligent Systems and Technology* 5 (1):1–24. doi:10.1145/2542182.

Assis, C. A. S., E. J. Machado, A. C. M. Pereira, and E. G. Carrano. 2018. Hybrid deep learning approach for financial time series classification. *Revista Brasileira De Computação Aplicada* 10 (2):54–63. doi:10.5335/rbca.v10i2.7904.

Atsalakis, G. S., and K. P. Valavanis. 2009. Surveying stock market forecasting techniques–Part II: Soft computing methods. *Expert Systems with Applications* 36 (3):5932–41. doi:10.1016/j.eswa.2008.07.006.

Avanço, L. V., and M. D. G. Volpe Nunes. 2014. Lexicon-based sentiment analysis for reviews of products in Brazilian Portuguese. Intelligent Systems (BRACIS), 2014 Brazilian Conference on, 277–81. IEEE, São Carlos, SP, Brazil. doi:10.1094/PDIS-04-13-0439-PDN.

Bollen, J., H. Mao, and X. Zeng. 2011. Twitter mood predicts the stock market. *Journal of Computational Science* 2 (1):1–8. doi:10.1016/j.jocs.2010.12.007.

Corder, G. W., and D. I. Foreman. 2014. *Nonparametric statistics: A step-by-step approach*. Hoboken, NJ: John Wiley & Sons.

Economia, G. 1. G. 2018a. Bovespa bate recorde após eleição de Bolsonaro, mas fecha em queda à espera de detalhes do novo governo. Accessed April 16, 2019. https://g1.globo.com/economia/noticia/2018/10/29/bovespa-29102018.ghtml.

Economia, G. 1. G. 2018b. Bovespa fecha em queda de 2,8% com cenário eleitoral; Eletrobras recua forte. Accessed April 16, 2019. https://g1.globo.com/economia/noticia/2018/10/10/bovespa-10102018.ghtml.

Economia, U. O. L. 2018f. Dólar passa de R$ 4 após pesquisas eleitorais; até onde vai a moeda? Accessed April 16, 2019. https://economia.uol.com.br/cotacoes/noticias/redacao/2018/08/21/cotacao-dolar-alta-eleicoes-pesquisas-limite-maxima.htm.

Econômico, V. 2018. Jogo eleitoral pode levar Ibovespa a 45 mil ou a 170 mil pontos. Accessed April 16, 2019. https://www.valor.com.br/financas/5576153/jogo-eleitoral-pode-levar-ibovespa-45-mil-ou-170-mil-pontos.

Feuerriegel, S., and H. Prendinger. 2016. News-based trading strategies. *Decision Support Systems* 90:65–74. doi:10.1016/j.dss.2016.06.020.

Francis, J. C., and E. Kirzner. 1991. *Investments: Analysis and management*. New York, NY: McGraw-Hill.

Globo, O. 2018. Após 1° turno, Bolsa registra maior volume financeiro da história. Accessed April 16, 2019. https://oglobo.globo.com/economia/apos-1-turno-bolsa-registra-maior-volume-financeiro-da-historia-23139170.

Hájek, P. 2018. Combining bag-of-words and sentiment features of annual reports to predict abnormal stock returns. *Neural Computing and Applications* 29 (7):343–58. doi:10.1007/s00521-017-3194-2.

Haykin, S. 1994. *Neural networks: A comprehensive foundation*. Upper Saddle River, NJ: Prentice Hall PTR.

Hussein, D. M. E.-D. M. 2018. A survey on sentiment analysis challenges. *Journal of King Saud University-Engineering Sciences* 30 (4):330–38. doi:10.1016/j.jksues.2016.04.002.

Kibriya, A. M., E. Frank, B. Pfahringer, and G. Holmes. 2004. Multinomial naive bayes for text categorization revisited. *Australasian Joint Conference on Artificial Intelligence*, 488–99. Berlin, Heidelberg: Springer.

Kraus, M., and S. Feuerriegel. 2017. Decision support from financial disclosures with deep neural networks and transfer learning. *Decision Support Systems* 104:38–48. doi:10.1016/j.dss.2017.10.001.

Li, Q., T. Wang, L. Ping, L. Liu, Q. Gong, and Y. Chen. 2014. The effect of news and public mood on stock movements. *Information Sciences* 278:826–40. doi:10.1016/j.ins.2014.03.096.

Lima, M. L., T. P. Nascimento, S. Labidi, N. S. Timbó, M. V. L. Batista, G. N. Neto, E. A. M. Costa, and S. R. S. Sousa. 2016. Using sentiment analysis for stock exchange prediction. *International Journal of Artificial Intelligence & Applications (IJAIA)* 7 (1): 59–67.

Long, D., J. Bradford, A. Shleifer, L. H. Summers, and R. J. Waldmann. 1990. Noise trader risk in financial markets. *Journal of Political Economy* 98 (4):703–38. doi:10.1086/261703.

Loughran, T., and B. Mcdonald. 2011. Combining bag-of-words and sentiment features of annual reports to predict abnormal stock returns. *Journal of Finance* 66 (1):35–65. doi:10.1111/j.1540-6261.2010.01625.x.

Makrehchi, M., S. Shah, and W. Liao. 2013. Stock prediction using event-based sentiment analysis. 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 337–42. IEEE, Atlanta, GA, USA, November.

Malkiel, B. G. 2003. The efficient market hypothesis and its critics. *Journal of Economic Perspectives* 17 (1):59–82. doi:10.1257/089533003321164958.

Malkiel, B. G., and E. F. Fama. 1970. Efficient capital markets: A review of theory and empirical work. *The Journal of Finance* 25 (2):383–417. doi:10.1111/j.1540-6261.1970.tb00518.x.

Martins, R. F., A. Pereira, and F. Benevenuto. 2015. An approach to sentiment analysis of web applications in portuguese. Proceedings of the 21st Brazilian Symposium on Multimedia and the Web, 105–12. Manaus, AM: ACM.

Mocherla, S., A. Danehy, and C. Impey. 2017. Evaluation of naive bayes and support vector machines for wikipedia. *Applied Artificial Intelligence* 31 (9–10):733–44. doi:10.1080/08839514.2018.1440907.

Nofer, M., and O. Hinz. 2015. Using Twitter to predict the stock market. *Business & Information Systems Engineering* 57 (4):229–42. doi:10.1007/s12599-015-0390-4.

Pang, B., L. Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval* 2 (1–2):1–135. doi:10.1561/1500000011.

Pang, B., L. Lee, and S. Vaithyanathan. 2002. Thumbs up?: Sentiment classification using machine learning techniques. Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10, EMNLP '02, Stroudsburg, PA, USA, 79–86. Association for Computational Linguistics.

Ravi, K., and V. Ravi. 2015. A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Systems* 89:14–46. doi:10.1016/j.knosys.2015.06.015.

Santos, H. S., A. H. Laender, and A. C. Pereira. 2015. A twitter view of the Brazilian stock exchange market. *Lecture Notes in Business Information Processing* 239:112–23.

Schumaker, R. P., Y. Zhang, C.-N. Huang, and H. Chen. 2012. Evaluating sentiment in financial news articles. *Decision Support Systems* 53 (3):458–64. doi:10.1016/j.dss.2012.03.001.

Sun, S., C. Luo, and J. Chen. 2017. A review of natural language processing techniques for opinion mining systems. *Information Fusion* 36:10–25. doi:10.1016/j.inffus.2016.10.004.

Taboada, M., J. Brooke, M. Tofiloski, K. Voll, and M. Stede. 2011. Lexicon-based methods for sentiment analysis. Technical Report 2.

Vargas, M. R., B. S. L. P. de Lima, and A. G. Evsukoff. 2017. Deep learning for stock market prediction from financial news articles. 2017 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), 60–65. IEEE, Annecy, France, June.

Yan, D., G. Zhou, X. Zhao, Y. Tian, and F. Yang. 2016. Predicting stock using microblog moods. *China Communications* 13 (8):244–57. doi:10.1109/CC.2016.7563727.

Yoshihara, A., K. Fujikawa, K. Seki, and K. Uehara. 2014. Predicting stock market trends by recurrent deep neural networks. In *Springer International Publishing*, ed. by D. N. Pham and S. P. Park, 759–69. Cham: Springer.

Zhao, B., H. Yongji, C. Yuan, and Y. Huang. 2016. Stock market prediction exploiting microblog sentiment analysis. 2016 International Joint Conference on Neural Networks (IJCNN), 4482–88. IEEE, Vancouver, BC, Canada, July.