A New Framework for Analyzing News in the Financial Markets to Enhance the Investor's Perception

Issam Aattouchi and Mounir Ait Kerroum
Ibn Tofail University, Kenitra, Morocco
Email: {aattouchi.issam, aitkerroum.mounir}@uit.ac.ma

***Abstract*—In finance, the flow of news is constantly updated in a way that changes the investors' understanding, influences their sentiment, and hence, shapes the financial markets (Asset returns, Volatility, Interest rates, etc.). Thus, one of the most important challenges that an investor has to overcome is to extract useful information from text data (News) in order to be used in decision making after reviewing all the literature dealing with the analysis of financial news. In this paper, we introduce an automated methodology to extract relevant words and topics from the FOMC (The Federal Open Market Committee) Minutes reports. Using the Latent Dirichlet Allocation (LDA) algorithm to track the main topics in the FOMC reports especially the "inflation" topic, we made a dimension reduction of our corpus by considering only the first 250 words belonging to this topic. Using the reduced corpus, we applied filters to clean non discriminative words and we employ deep learning in detecting only sentiment-charged words. By the end, we obtained a list of the most relevant words (dictionary) in the FOMC minutes and a prediction (with an accuracy of 98.76% and a correlation of 77%) of the Fed Funds rates.**

***Index Terms*—FOMC, classification, LDA, CNN, prediction, news, finance, market, interest rates, neural networks**

I. Introduction

The development of the financial markets has a positive influence on economic growth. One way of looking at the role of finance in economic growth is the key role that financial markets play in providing the necessary conditions for technological innovation. Indeed, the development of financial markets is a by-product of economic expansion, which creates opportunities that give impetus to the development of the financial system.

The latest advances in the field of Artificial Intelligence (AI) and machine learning have provided a very wide range of opportunities to explore, computational intelligence in finance has been a very popular topic for both financial industry and academia in the last few decades, the financial sector is undergoing a transformation in view of the new emerging artificial intelligence technologies and the increasing competitiveness within the markets, the aim of this study is to address the challenges of extracting useful information from textual data to facilitate decision making in the financial sector. This research focuses on developing a new framework for analyzing financial market news in order to improve investor perception.

The Federal Reserve, officially named Federal Reserve System [1] and often referred to the "Fed", is responsible for making decisions about the target federal funds rate, and it maintains detailed meeting records, which is called FOMC minutes. It was created in 1913 by the Federal Reserve in order to ensure the stability of the monetary and financial system, after a series of banking panics, notably in 1907, when Americans had withdrawn their liquidities in masse from the banks. Today, the Fed fulfills three main missions, as do other central banks, such as the European Central Bank (ECB).

The Fed defines the monetary policy of the United States: it sets the mandatory reserve rate, the discount rate, and directs open market operations (interventions on the financial markets to influence the level of interest rates). The decisions taken by the Fed on interest rates, as well as the opinions on the American economic situation expressed publicly by its President, have a decisive announcement effect on the direction of the world's financial markets.

In particular, the FOMC (The Federal Open Market Committee) meets eight times a year and publishes statements on the day of the meetings as well as the minutes with a three-week time lag. In this paper, we will analyze in depth the FOMC minutes which contain the analyses and opinions of FOMC members on the economic situation and financial around the world. In this analysis, we have collected the FOMC minutes from 1968 to 2019, which gave us a total of 484 documents. Although the format of the minutes has changed over time, they essentially contain sections such as the list of participants, briefing recent developments of financial, analyses and opinions on the situation of financial and the political actions of the FOMC. These sections are not standard and are not distributed accordingly in the minutes. Thus, we will proceed with the automatic cleaning and preparation of our data.

In our analysis, we begin our study by a description of our approach from the data disposal to the data preparation. Next, we describe our method of extracting the useful

Manuscript received September 17, 2021; revised January 10, 2022.

doi: 10.12720/jait.13.2.125-131

information. In particular, we will introduce a dual algorithm filtering (using both LDA and TF-IDF representations) that determines the most powerful and discriminating words that the Fed tends to use when introducing an increase/decrease in fund rates. Finally, we merge the filtering model with the CNN model to obtain accurate predictions of U.S. interest rate movements.

II. Related Works

Natural Language Processing (NLP) is one of the most promising fields in Machine Learning and maybe its hottest area that has managed to resolve many tasks that seemed to be extremely difficult in the past, including: Information Retrieval, Information extraction, Machine translation, Spam filter, Sentiment analysis, etc. The main difficulty when applying NLP is the high dimensionality of text data. Hence, many methods were developed to quantify text into numerical vectors.

In finance, one of the most important challenges that an investor has to overcome is to extract useful information from text data (News) in order to use it in decision making. Thus, obtaining accurate forecast of stock market movement is the engine of financial prediction. Many algorithms have been used for this purpose, studies on Algo-trading have focused on forecasting stock prices or indexes as this is the main focus of this field of finance. The LSTM model is the most widely used in implementations.

In this section, about 20 articles on the reactions of financial markets are reviewed. We present some existing studies and related work that have examined news analytics in Finance.

Junfeng Jiang and Jiahao Li [2], built a finance-specific sentiment of financial markets in the special case of the Chinese Market. They began by collecting several news and comments from many influential Chinese financial websites automatically, then they applied word embedding techniques to represent numerically this news, and finally constructing a senti-score to news and verify its reliability.

In this algorithm, they begin by using a well pretrained model of Word2vec in order to have a developed representation of English words. In addition, they used the WordNet representation in order to define two complementary distances: Morphological Similarity and Semantics Similarity, using theses similarities, a relevant senti-score can be computed. In this algorithm, the authors have built a lexicon using 100 words that appear in financial news frequently and have specific sentiment (50 positive words and 50 negative words). Although this model has many advantages, it still has some points to develop.

For example, computing sentimental factors for every financial product.

Bao W, Yue J, Rao Y [3], presents a deep learning framework where Wavelet Transforms (WT), Stacked Autoencoders (SAEs), and Long-Short Term Memory (LSTM) are combined for stock price forecasting. The SAEs for hierarchically extracted deep features are introduced into stock price forecasting for the first time. The deep learning framework comprises three stages. First, the stock price time series is decomposed by WT to eliminate noise. Second, SAEs are applied to generate deep high-level features for predicting the stock price. Third, high-level denoising features are fed into LSTM to forecast the next day's closing price.

Although the proposed integrated system has satisfactory predictive performance, it still has some insufficiencies. For instance, a more advanced hyper-parameters selection scheme might be embedded in the system to further optimize the proposed deep learning framework.

Andrew Han [4], proposes the use of neural network models to predict relative factor returns. Using a rolling window of relevant news articles, the author utilizes three models: a baseline two-layer neural network that uses doc2vec-produced document embeddings, a Gated Recurrent Unit (GRU) model with attention, and a convolutional neural network, doc2vec that represents words using a large Tree structure. That is, words are linked only according to the meaning or semantic lexical correlation. Thus, distances between words may accurately represent semantic differences. In fact, this measure can be used only as an index of how words are mutually semantically different, which affects the quality of the results.

Christopher Rohlfs *et al.* [5] present the effects of Federal Open Market Committee text content on the direction of short- and medium-term interest rate movements. They used the medLDA model on a set of 146 documents and obtain accuracies of 64% in predicting the Federal Funds Target Rate and the Effective Rate. Since the words relevant to short- and medium-term interest rates differ, they apply a supervised approach to learn distinct sets of topics for each dependent variable being examined. And they generate predictions with and without controlling for factors relevant to interest rate movements. The authors have obtained lower but comparable accuracies, because, when applying LDA to a corpus of text, the algorithm gives just a set of words for each topic without determining exactly the nature of the topic. Thus, finding the number and the subjects of the optimal topics remains the user's work

A Convolutional Neural Network (CNN) is employed to categorize moods at the phrase level and capture local correlations of spatial or temporal features efficiently [6]-[10]. However, CNN's long-term reliance on consecutive financial text input is tough to manage. The constraint is then overcome using RNN and LSTM. RNN (The Recurrent Neural Network) is a neural network with directed connections [11], whereas the LSTM Long-term and Short-term Memory solves the problem of disappearing gradient operations and stores more information than the RNN [12], [13]. In financial news analysis, RNN-based models have been regularly employed [14]. To learn sequential correlations and extract features in simultaneously, CNN and Bi-LSTM [15] can be combined.

Many hybrid models combining LSTM and CNN architectures have been developed to capture the local and global aspects of documents and texts. Unlike existing

solutions, our model presents a hybrid approach to solve the challenge of proposing a new text-mining methodology to predict asset returns using a purely statistical technique. In fact, all sentiment scores that were presented above did not give importance for the type of prediction they are being used for. We can also say that our suggested method improves the performance of previous work in the sentiment analysis field and introduces innovation compared to existing models by using the notion of mixture between LDA and CNN as well as new metrics like the sentiment index. The idea here is to use a Dataset of text news associated with their generated asset returns to build a strong model of sentiment score. Precisely, we tried to learn the sentiment score model from the joint behaviour of text news and stock returns.

III. Proposed Approach

The architecture of the suggested approach is presented in this section. This model architecture can be separated into three sections, as indicated in Fig. 1:

- Data pre-processing
- Modeling of the subject with double LDA and TF IDF
- Prediction of interest rates with CNN.

The first step of our approach is data collection and filtering. The second step is the reduction of the pre-processed corpus keeping the words that represent more our domain, then the last step is based on the convolutional neural network algorithm to predict interest rates while relying on the prediction of sentiments in FOMC report detected by the LDA algorithm.

The next sections tackle the specifics of each of these steps. Section III.A provides details on the first step, Section III.B provides details on the second, and Section III.C describes the third step.

Figure 1. The proposed architecture.

A. Pre-processing Layer

As demonstrated in Fig. 2, the pre-processing strategy entails presenting the corpus data in a more structured format in order to extract FOMC characteristics and opinion terms more easily.

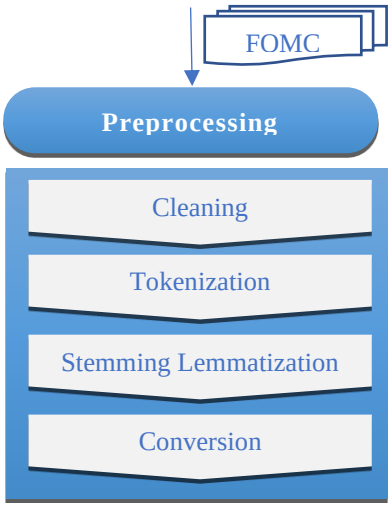

Figure 2. Préprocessing layer.

1) Cleaning

The most common words in a text, such as (is, by, and the), decrease the accuracy rate of sentiment classification. The URLs in the corpus do not contain much information about the sentiment of the document. Therefore, before feature extraction, the proposed system eliminates frequently occurring URLs and words.

In addition, the proposed system uses a keyword manager to remove content that does not contribute to sentiment. This includes articles (a, an, the), symbols (@, date, #, etc.) and punctuation.

2) Tokenization

The tokenization technique divides the corpus's composite text into tiny tokens. The composite text can contain a word space and delimiters. As a result, the suggested method eliminates delimiters and word spaces by employing an N-gram token.

3) Stemming and lemmatization

The process of reducing a word into its basic form is in text analysis known as stemming. Lemmatization plays an important role in sentiment analysis. It determines the lemma of words used in a sentence using a lexicon.

The suggested approach obtains lexical information about each word after lemmatization. As a result, the suggested system analyzes the stem and lemma terms.

4) Character conversion and lowercase setting

Since we are mainly interested in Fed policy, we have noticed that every decision made by the Fed has started with the phrase "unanimously" in every minute since 1968. Thus, we have simply cut off every minute in this instance first from the word "unanimous". This allowed us to remove the most irrefutable information from the documents and to have a text containing essentially the political decisions of the Fed.

B. Subject Modeling Layer

After processing our corpus and cleaning our data, we applied the LDA algorithm.

1) Latent Dirichlet Allocation (LDA)

According to [16], the LDA is a hierarchical Bayesian model that projects a text document into a small latent space covered by a set of automatically learned topical databases. More precisely, LDA is a three-level model in which each element is considered as a finite mixture of latent subjects. This algorithm can be used for different purposes, such as topic extraction, size reduction, novelty detection, summarization, similarity and relevance judgments, etc. The objective of the algorithm is to represent short text descriptions (or any other data collection) that allow the processing of textual corpora while preserving essential statistical relationships that are useful for basic tasks.

a) Notations

We will use the language of text collections while explaining the algorithm, even if the LDA is not necessarily related to text and can be applied to other fields such as collaborative filtering, content-based image search and bioinformatics. Formally, we start by defining the following elements:

b) Dictionary

Let us consider D as a dictionary of all possible words and index them by [1,...,V] where V = |D|.

We will use one-hot coding for the representation of the words.

For each document $w \in D$, LDA assumes the generating procedure outlined below. A document is a sequence of N words designated by:

$$\mathrm{w} = (\mathrm{w}_1, \mathrm{w}_2, \dots, \mathrm{w}_\mathrm{n}) \quad (1)$$

where w_n is the nth word in the sequence. A corpus is a set of M documents designated by

$$\mathrm{D} = (\mathrm{w}_1, \mathrm{w}_2, \dots, \mathrm{w}_\mathrm{M}) \quad (2)$$

The main idea behind the approach is that the documents are represented as random mixtures on topics latent, with each subject defined by a word distribution.

The main idea of the algorithm is that the documents are represented as random mixtures on topics latent z_n, where each topic is characterized by a distribution over the words.

Choose $\mathrm{N} \sim \mathrm{Poisson}(\lambda)$ where N is the size of the document and choose $\theta \sim \mathrm{Dir}(\alpha)$

For each of the N words w_n:

- Choose a topic $z_n \sim \mathrm{Multinomial}(\theta)$. α must be estimated.
- Choose a word $\mathrm{w_n}$ from $\mathrm{p}(\mathrm{w_n} \mid \mathrm{z_n}, \beta)$ a multinomial probability conditioned by topic z_n.

2) Filtering using Term Frequency Inverse Document Frequency representation (TF-IDF) [17]

Indeed, some words are not relevant and should be neglected in our study. More precisely, words appearing in almost all documents such as market, reserve, currency... must be intuitively removed. New words that are common in one or a small group of documents have a strong discriminating power and must be counted heavily. This intuition justifies the use of the TF-IDF algorithm to assign a relevance score to words. Let us start by transforming our corpus into a matrix using the Bag of Words representation.

According to [17], the TF-IDF is defined as follows, where *f*(*w*, *d*) is the frequency of the word w in the document d and *f*(*w*,*D*) indicates the percentage of documents in the corpus *D* containing *w*:

$$tfidf(w,d) = f_r(w,d)\log\left(\frac{1}{f(w,C)}\right) \quad (3)$$

3) Predicting US Funds rates using deep learning (CNN)

Sentiment analysis is a typical technique for determining how the documents feel about a subject. Our central idea is to merge the use of the double filter LDA and TF-IDF with in-depth learning to determine expectations about the interest rate and the market as a whole based on document FOMC Minutes. The reason we chose the in-depth learning methodology for sentiment analysis is that it learns characteristics during the learning process. Deep learning methods produce an abstract representation that is insensitive to local changes in the input data. In addition, problems related to textual data, such as data tagging, semantic indexing, and rapid information retrieval, can be well handled using deep learning.

Deep learning enables complicated artificial intelligence tasks to be performed with a simpler model. Although Deep Learning algorithms have been employed in several Big Data areas, such as computer vision [18]-[23] and speech recognition [24]-[29], they are still intact in the context of sentiment analysis. In this article, we evaluate the adoption of deep learning for sentiment analysis of financial data. We will use the Convolutional Neural Network (CNN) [30] which is an example of various deep neural networks.

The convolution layers of this model, which is the most extensively employed for image analysis, may use the internal structure of the data. Duo to the internal structure that exists within textual documents, CNN is employed in a variety of applications, including markup systems, entity search systems, sentence modeling systems, and more [23], [24].

CNNs are utilized in NLP applications for local feature extraction and are made up of numerous convolution layers. Convolution is conducted on the input characteristics via linear filters in these networks. To apply a CNN on a sentence S with 's' words, first, an embedding vector of size e is created.

Then, on the sub-matrices of the input feature matrix, a filter F of size e×h is applied repeatedly. This produces a feature map.

$\mathrm{M} = [m_0, m_1, \dots, m_{\mathrm{s-h}}]$ as follows [31]:

$$m_\mathrm{i} = F.\boldsymbol{S}_{\mathrm{i:i+h-1}} \quad (4)$$

where: $i = 0, 1, ..., s-h$; $S(\mathrm{i:j})$ is a sub-matrix of S from row i to j.

It is a common practice to minimize the dimension of feature maps by feeding them to a pooling or sub-sample layer. Max-pooling is a pooling strategy which selects the most essential feature b of the feature map in the following way:

$$\mathrm{b} = max_{0 \le \mathrm{i} \le \mathrm{s-h}}\{\mathrm{mi}\} \quad (5)$$

The outputs of pooling layer are concatenated and form a pooled feature vector which may then be used as the input of a fully connected network (as shown in Fig. 3).

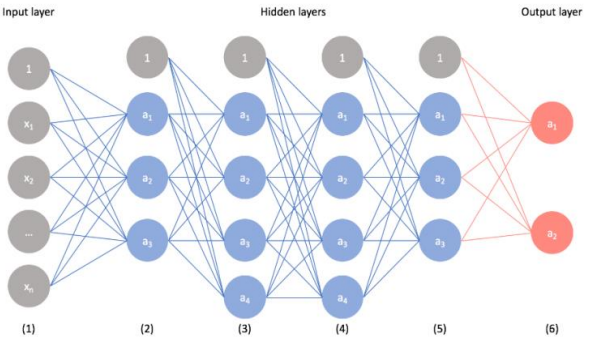

Figure 3. Convolutional neural networks.

IV. RESULTS

A. Experimental Environment

Experiments in this study are performed on a PC with the following hardware configuration: Intel (R) Core (TM) i7- 6500U, CPU @3.0GHZ*8, 16.00GB RAM and 1TB hard disk, 64-bit Operating System. The software environment is identical: Linux operating system (Ubuntu 16.04), Python, anaconda and Jupiter.

B. Results

After processing our corpus and cleaning our data, we applied the double filter LDA / TF-IDF.

1) First filter LDA

In fact, as we saw in the introduction, the minutes generally contain different topics that are not necessarily divided into paragraphs. That is, each paragraph may contain a mixture of different topics. As we have seen in the theoretical description above, the LDA made this exact hypothesis that justifies its usefulness for the minutes. Indeed, the estimation of the parameters of the model α and β (using the corpus) is done using the EM (Expectation-Minimization) type algorithm, an algorithm of variational inference and maximum likelihood.

By cross-validation, we found that the number of topics that give the most efficient results is 8. Table I shows a few words belonging to each topic in our study:

TABLE I. TOPICS PROVIDED BY THE LDA ALGORITHM

Topic 1	price, market, rate, increase, economic, decline, inflation, period, rise
Topic 2	rate, percent, market, growth, reserve, monetary, unanimous, aggregate, reserve
Topic 3	growth, member, price, rate, policy, continue, inflation, business, consumer
Topic 4	bank, treasury, market, credit, liquidity, money, condition, term, facility
Topic 5	duty, testimony, telegram, description, tragic, bulletin, applicability, disclose, write
Topic 6	assistant, president, board, governor, economist, Richmond, vice senior, statistic
Topic 7	Dallas, Chicago, district, city, Francisco, York, activity, Atlanta, Boston
Topic	foreign, market, bank, currency, operation, shall, open, security, transaction

The LDA, through the parameter θ, gives the proportion of each topic in a given document (Of course, the sum of the components of θ is by Definition 1 because they are probabilities). We can see from Fig. 4 below that the LDA manages to detect interesting topics such as inflation and interest rates (T1, T2, T3), places (T7), decision-makers (T6), review of the financial situation (T5). Since our objective is to study the effect of the FOMC minutes on interest rates, we are particularly interested in topics T1, T2 and T3. Our approach is to examine the most common words in each of these topics, merge these words and exclude all other words that do not belong in the merged dictionary. Intuitively, the idea comes from the natural assumptions that these topics are profoundly responsible for interest rate movements.

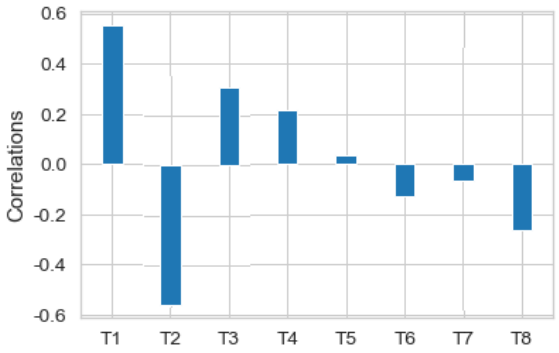

Figure 4. Correlation between proportions given by LDA and interest rates.

Fig. 5 shows the Inter topics Distance Map, where the circles represent different subjects and the distance between them. Similar subjects appear closer together and dissimilar subjects appear farther apart. The relative size of the circle of a subject in the graph corresponds to the relative frequency of the subject in the corpus.

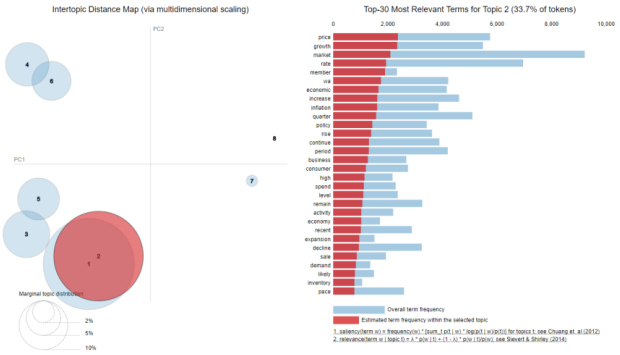

Figure 5. Distribution of topics.

2) Second filter TF-IDF

As discussed below, some words such as: market, reserve, currency, are present in almost all minutes. Their discriminative power is therefore too small. Hence, it seems natural to filter out this kind of words. To do so, we have opted for the TF-IDF measure.

Finally, we have obtained a set of significant words and our dictionary was reduced heavily. As showing in Fig. 6, the word cloud of our corpus when we filter out all the

insignificant words (using the double filter we defined above) appears far more relevant than the non-proceeded cloud.

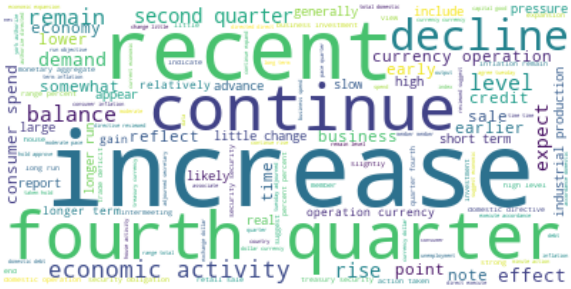

Figure 6. Word cloud of the filtered FOMC minutes.

3) Predicting US Funds rates using CNN

As described in the last section, we have successfully reduced the size of our dictionary so that it contains only UD funds rate related words. This methodology tracks the relevant words in our study.

We plot here the variation of Fed funds rates VS the predicted sentiments. Note that we scaled both graphs as usually done in time series prediction. Fig. 7 shows both the sentiment score computed by our algorithm and the variation of the Fed interest rest during the same period.

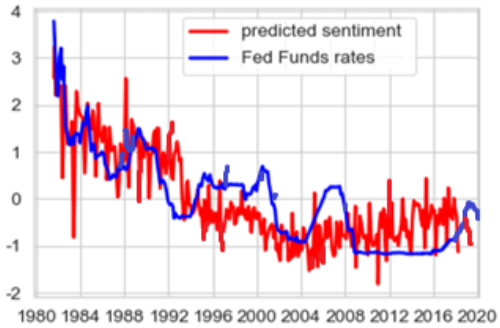

Figure 7. The variation of Fed funds rates VS the predicted sentiments.

The R2 correlation between these two-time series is 77%. The minutes are a strong factor to make predictions on the US Funds rates. Our sentiment seems too noisy, even though it is very correlated to rates variations, which makes it unusable to prediction. We can remark visually that the allure of the two graphs seems to be extremely correlated. Hence, we have implemented an algorithm for smoothing volatile graphs. This gives the following result (as shown in Fig. 8):

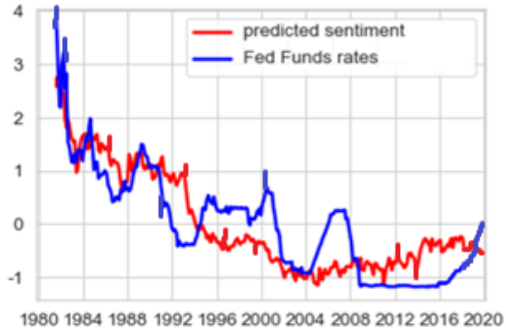

Figure 8. Smoothed sentiment versus the variations of the Fed funds rates.

In order to prove our proposed approach, we compared our system with several algorithms including SVM, RNN, ANN, LSTM, naive bayes, and CNN. The proposed approach is considered the most accurate classifier. Our comparison of the classifications reveals efficiency measures, being readily available for each applied algorithm.

Although different efficiency measures are applied for each algorithm, the most efficient one must be chosen in order to predict interest rates. Our comparison shows that our Double Filter LDA TFIDF and CNN model gives an accuracy of 98.76% for identifying interest rates and recalls in comparison to the other algorithms. This comparison leads to an improvement in the result performance as shown in Table II. Our choice of the regression algorithm revolves around the measures, which provide efficient results.

TABLE II. TOPICS PROVIDED BY THE LDA ALGORITHM

	ACCURACY	RECALL	PRECISION
ANN	0.780246914	0.820224719	0.719211823
CNN	0.864197531	0.885416667	0.837438424
LSTM	0.807407407	0.777777778	0.862068966
NB	0.497530864	0.547911548	0.644508671
RNN	0.765432099	0.754716981	0.751173709
SVM	0.62962963	0.634517766	0.615763547
THE PROPOSED SYSTEM	0.987654321	0.99009901	0.985221675

V. CONCLUSION

In our analysis, we used ML techniques to deeply extract information from financial news. We presented a powerful tool that performs feature extraction on the corpus of minutes by effectively tracking relevant words related to Fed funds rate. We have also merged this extractor with a Deep Learning algorithm to predict Fed Interest rates. Despite the small size of our data, we succeeded to obtain very good results.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Issam Aattouchi conceived the overall approach and implemented the tools presented. The first version of this document was prepared by Issam Aattouchi, but it was further revised and updated by Mounir Ait Kerroum who was responsible for supervising this research work All authors analyzed the results, and approved the final version of this paper.

REFERENCES

[1] Recent developments. [Online]. Available: https://www.federalreserve.gov/
[2] J. Jiang and J. Li, "Constructing financial sentimental factors in Chinese market using natural language processing," arXiv preprint arXiv:1809.08390, 2018.
[3] W. Bao, J. Yue, and Y. Rao, "A deep learning framework for financial time series using stacked autoencoders and long-short term memory," *PLoS ONE*, vol. 12, no. 7, p. e0180944, 2017.
[4] A. P. Han, "Financial news in predicting investment themes," 2019.

[5] C. Rohlfs, S. Chakraborty, and L. Subramanian, "The effects of the content of FOMC communications on US treasury rates," in *Proc. Conference on Empirical Methods in Natural Language Processing*, 2016, pp. 2096-2102.
[6] Y. Kim, "Convolutional neural networks for sentence classification," in *Proc. Conference on Empirical Methods in Natural Language Processing*, Doha, Qatar, 2014, pp. 1746-1751.
[7] D. Tang, B. Qin, and T. Liu, "Document modeling with gated reccurrent neural network for sentiment classification," in *Proc. Conference on Empirical Methods in Natural Language Processing*, Lisbon, Portugal, 2015, pp. 1422-1432.
[8] X. Zhang, J. Zhao, and Y. LeCun, "Character-level convolutional networks for text classification," in *Advances in Neural Information Processing Systems*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, Eds., Curran Associates, Inc., 2015, pp. 649-657.
[9] R. Johnson and T. Zhang, "Deep Pyramid convolutional neural networks for text categorization," in *Proc. the 55th Annual Meeting of the Association for Computational Linguistics*, Vancouver, Canada, 2017, pp. 562-570.
[10] M. Kraus and S. Feuerriegel, "Decision support from financial disclosures with deep neural networks and transfer learning," *Decis. Support Syst.*, vol. 104, pp. 38-48, Dec. 2017.
[11] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," ArXiv180107883 Cs Stat, Jan. 2018.
[12] S. Sohangir, D. Wang, A. Pomeranets, and T. M. Khoshgoftaar, "Big data: Deep learning for financial sentiment analysis," *J. Big Data*, vol. 5, no. 1, p. 3, 2018.
[13] K. S. Tai, R. Socher, and C. D. Manning, "Improved semantic representations from tree-structured long short-term memory networks," in *Proc. the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, Beijing, China, 2015, pp. 1556-1566.
[14] E. Shijia, L. Yang, M. Zhang, and Y. Xiang, "Aspect-based financial sentiment analysis with deep neural networks," in *Companion Proc. The Web Conference 2018, Republic and Canton of Geneva*, Switzerland, 2018, pp. 1951-1954.
[15] T. Chen, R. Xu, Y. He, and X. Wang, "Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN," *Expert Syst. Appl.*, vol. 72, pp. 221-230, 2017.
[16] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993-1022, 2003.
[17] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," *Information Processing and Management*, vol. 24, no. 5, pp. 513-523, 1988.
[18] D. A, Freedman, *Statistical Models: Theory and Practice*, Cambridge: Cambridge University Press, 2009.
[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Image net classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097-1105, 2012.
[20] R. Socher, E. H. Huang, J. Pennin, C. D. Manning, and A. Y. Ng, "Dynamic pooling and unfolding recursive autoencoders for paraphrase detection," *Adv. Neural Inf. Process. Syst.*, vol. 24, pp. 801-809, 2011.
[21] J. Gao, L. Deng, M. Gamon, X. He, and P. Pantel, "Modeling interestingness with deep neural networks," US Patent App. 14/304,863, 2014.
[22] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, "A convolutional neural network for modelling sentences," arXiv preprint arXiv:1404.2188, 2014.
[23] D. Tang, F. Wei, N. Yang, M. Zhou, T. Liu, and B. Qin, "Learning sentiment-specific word embedding for Twitter sentiment classification," *ACL*, vol. 1, pp. 1555-1565, 2014.
[24] K. Pearson, "Notes on regression and inheritance in the case of two parents," *Proc. R. Soc. Lond.*, vol. 58, pp. 240-242, 1895.
[25] A. Graves, A. R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013.
[26] G. Dahl, A. R. Mohamed, and G. E. Hinton, "Phone recognition with the mean-covariance restricted Boltzmann machine," *Adv. Neural Inf. Process. Sytst.*, vol. 23, pp. 469-477, 2010.
[27] E. George, D. Yu, L. Deng, and A. Acero, "Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition," *IEEE Trans. Audio Speech Lang. Process.*, vol. 20, no. 1, pp. 30-42, 2012.
[28] F. Seide, G. Li, and D. Yu, "Conversational speech transcription using context-dependent deep neural networks," in *Proc. Twelfth Annual Conference of the International Speech Communication Association*, 2011.
[29] A. Mohamed, G. E. Dahl, and G. Hinton, "Acoustic modeling using deep belief networks," *IEEE Trans. Audio Speech Lang. Process.*, vol. 20, no. 1, pp. 14-22, 2012.
[30] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.
[31] H. T. Nguyen and M. L. Nguyen, "An ensemble method with sentiment features and clustering support," *Neurocomputing*, vol. 370, pp. 155-165, 2019.

Issam Aattouchi is a Doctoral degree student in Department of Computer Engineering, Faculty of Science, Ibn Tofail University, Morocco. He obtained his state engineer diploma in Computer Engineering from Cadi Ayyad University, Morocco. His research interests include artificial intelligence, machine learning, human machine interfacing, and application of machine learning in finance.

Mounir Ait Kerroum is a qualified Professor in Department of Computer Engineering, National School of Business and Management, Ibn Tofail University, Morocco. He obtained his M.Eng. in Computer Science and Telecommunications, and he gained his Ph.D. in IT and Telecommunications from Mohammed V University of Rabat, Morocco. His research interests include Artificial Intelligence, Pattern Recognition, Deep Learning, Classification of Hyperspectral Images, Classification of Medical Images, Recognition of Arabic Manuscript Text.