Towards Evaluating the COVID'19 related Fake News Problem: Case of Morocco

Oussama Maakoul
ISO-Lab: Laboratory of Informatics, Systems and Optimization
Faculty of Science, Ibn Tofail University
Kenitra, Morocco
oussama.maakoul1@gmail.com

Sabah Boucht
ISO-Lab: Laboratory of Informatics, Systems and Optimization
Faculty of Science, Ibn Tofail University
Kenitra, Morocco
Sabahboucht47@gmail.com

Karima EL HACHIMI
ISO-Lab: Laboratory of Informatics, Systems and Optimization
Faculty of Science, Ibn Tofail University
Kenitra, Morocco
elhachimi.karima@gmail.com

Salma Azzouzi
ISO-Lab: Laboratory of Informatics, Systems and Optimization
Faculty of Science, Ibn Tofail University
Kenitra, Morocco
salma.azzouzi@gmail.com

2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS) | 978-1-7281-6921-7/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICECOCS50124.2020.9314517

Abstract—Nowadays social media play an important role in our societies. In fact, they remain one of the most commonly used means to easily obtain information online. However, social media can also be a catalyst for the proliferation of fake news around the world, especially in times of crisis where massive misinformation can have serious consequences. In particular, fake news about the COVID'19, as the virus spreads, leads to an infodemic of misinformation, hence the need for news verification mechanisms in social media. In this paper, we approach the problem of fake news related to COVID'19 as a global pandemic with a significant impact on various sectors. We also provide an aggregation system to detect and analyze fake news related to the COVID'19 pandemic in the Moroccan context based on data sets scrapped from Facebook.

Keywords— COVID'19, Fake news, social media, Machine Learning.

I. INTRODUCTION

The Coronavirus (COVID-19) is a humanitarian emergency, which began in early December 2019 in China. The first confirmed case in Morocco was on March 2020 and 17 days later, the Moroccan government declared a state of health emergency and began to take strict measures by closing airports, cafes, schools and mosques to contain the virus as well as prosecuting those who published fake news regarding the situation through social media.

In fact, Fake news and misinformation has been spread around the world since the emergence of the new corona virus disease through social media (in the form of videos, messages, or postings), thus profoundly affecting the social and economic sectors and causing panic in the public health field. In this context, the detection of fake news, which is described as the task of classifying news with a corresponding degree of certainty, is an essential task to ensure not only that social media users get reliable information, but also to help preserve and sustain a trustworthy ecosystem of news.

In the other hand, even if technologies such as artificial intelligence and natural language processing tools are very promising for researchers aiming to create systems capable of automatically detecting false news, the task of detecting false news is still a difficult task as it requires models to summarize the news and compare it to real news[1].

Moreover, the digitization of news has challenged traditional definitions of news. In fact, the new online platforms such as Facebook or Twitter provide space for non-journalists to produce journalistic outputs as well as reaching a mass audience [2]. In this work, we chose the social media network: Facebook as it is the most widely used platform by Moroccan people of different ages and intellectual levels. Our work aims to collect messages from pages or groups -of such platform- linked to COVID'19 and analyze the comments in each message by extracting and measuring the percentage of frequently repeated keywords indicating the veracity of the message.

To this end, the paper is organized as follows: The second section describes some basic concepts. Then, we introduce the problematic statement and some related works respectively in sections 3 and 4. Afterward, we describe materials and methods to implement our model in sections 5 and 6. Finally, section 7 gives some conclusions and identifies future works

II. PRELIMINARIES

A. *Fake News*

Fake news is defined as fabricated information that imitates the content of news media in form but not in organizational process or intention. Fake news sources, in turn, neglect the editorial standards and procedures of the mainstream media to ensure information is reliable and trustworthy. Fake news overlaps with other information disorders, such as misinformation and disinformation[3].

B. *Fake News Typology*

We can identify six ways in which the earlier studies identified and exploited "fake news" term [4]:

		Author's immediate intention to deceive	
		High	*Low*
Level of Facticity	*High*	Advertising	Satire
		Manipulation	
		Propaganda	
	Low	Fabrication	Parody

Table1. Fake News Typology

- **Satire** refers to satirical news programs that typically provide news updates to viewers using humor or exaggeration.
- **Parody** uses a presentation style that imitates mainstream news media. it plays on the absurdity of issues and highlights them by composing purely fictitious news stories.
- **Fabrication** refers to stories that do not have a Factual basis but are released to establish credibility in the style of news articles.
- **Advertising** uses fake news to describe materials using news values to draw attention but misleading many people in the process.
- **Propaganda** refers to news that a political entity creates for the benefit of a government, organization, or public figure in order to influence the public's perception.

C. Natural Language Processing (NLP)

Natural Language Processing (NLP) is a tract of Artificial Intelligence and Linguistics, devoted to make computers understand the statements or words written in human languages. Natural language processing came into existence to ease the user's work and to satisfy the wish to communicate with the computer in natural language [5].

D. Machine Learning Approaches

Machine learning makes use of statistics theory in the construction of mathematical models. The model may be predictive in making future predictions, or descriptive in acquiring knowledge from data, or both. We distinguish two major machine learning types [6]:

1) Unsupervised Machine Learning which aims to find undetected patterns in a dataset with no pre-existing labels and with minimal supervision. In this case, we can use:

- **Clustering** which is the task of grouping a set of objects according to their similarity
- **Association rules** allow links between data objects to be built into large databases. This technique revolves around the discovery of exciting relationships amongst variables in large databases.

2) Supervised Machine Learning which aims to learn a mapping from the input to an output whose correct values are provided by a supervisor. We can use:

- **Regression** is a method to predict continuous variables by identifying the differences between dependent variables and independent ones.
- **Classification algorithm** is the process of finding a function that helps to divide a set of data into different parameter-based classes.

III. PROBLEMATIC STATEMENT

Sharing news through social media has become a common method of disseminating information due to its ease of access, low cost and speed of distribution. However, the use of social media for news consumption is a double-edged sword, allowing the widespread dissemination of fake news.

In this context, we can observe today the increasing amount of fake news about COVID-19 on social media, which has become a very serious concern for our society; it is particularly harmful in the comments section. That is why news verification is an indispensable process to identify intentionally misleading information in the comments area.

However, the identification of fake news is far from trivial and a major challenge lies in the interpretation of different dialects that people use in comments. To this end, our strategy of data collection uses a tool for collecting comments from Facebook groups, and a machine-learning-based solution (logistic regression) to detect fake news.

IV. RELATED WORKS

In recent years, the detection of fake news has drawn research attention and many methods have been suggested. Authors in [7] have built a model for detecting false news from the datasets by using machine learning classifiers K-Nearest Neighbors, SVM, Decision Tree, Random Forest.

In [8], the authors suggest a new model for early detection of fake news on social media by classifying pathways for news propagation. First, they model each news story's propagation path as a multivariate time series in which each tuple is a numerical vector that represents the characteristics of a user involved in spreading the news. Then they create a time series classifier that combines both recurrent and convolutional networks capturing global and local user characteristic variations along the propagation path respectively to detect fake news.

In addition, the work [9] presents a multi-stage intervention system that incorporates reinforcement learning with a point process network activity model to counter false news in social networks. In the network, the spread of fake news and mitigation events is modeled 0on a multivariate Hawkes process with additional exogenous control terms. Moreover, authors in [10] developed a novel algorithm, "Detective" which performs Bayesian inference to detect fake news and learns about the flagging accuracy of users over time. The algorithm uses subsequent sampling to deliberately trade-off exploitation (selecting news that maximizes objective value at a given epoch) and exploration (selecting news that maximizes the value of information towards learning about users' flagging accuracy).

Furthermore, the FakeNewsTracker solution proposed in [11] is a system for comprehension and identification of false news. FakeNewsTracker can gather data automatically for news pieces and social context, which would help more research on recognizing and predicting fake news using powerful visualization techniques. Another work [12] is proposed by authors to build a real-world data sets that measure user confidence levels on fake news, and select representative groups of both "experienced" users who can recognize fake news items as fake and "naïve" users who are more likely to believe fake news. The research in [13] proposes a Tri-relationship embedding TriFN system, which simultaneously models publisher-news relationships and user-news interactions to identify fake news. The SpotFake system in [14] is a multimodal framework for fake news detection. This method detects fake news without taking any subtasks into account. The system exploits an article's textual as well as visual characteristics. Specifically, it used language models (such as BERT) to learn text features, and image features are therefore learned from VGG-19 pre-trained on ImageNet dataset. All experiments are conducted on two datasets available to the public (Twitter and Weibo).Finally, this paper is inspired from the method proposed in [15] which aim to combine between Machine Learning algorithms that subdivide into supervised learning techniques, and natural language processing methods. In the next section, we introduce our model representing the system to detect fake news from a dataset of comments.

V. FAKE NEWS PROPOSED MODEL

A. Architecture

The proposed system aims to analyze the veracity of information spreading in Facebook. The diagram in Fig 1 gives an overview of our model.

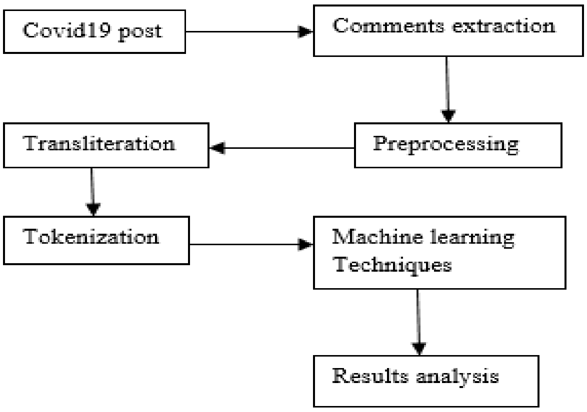

Fig. 1. The architecture of the proposed system

As described in figure Fig.1, the system consists of extracting, processing, categorizing public comments on the post, and calculating the percentage of each category then determine whether the post is Fake news or not. The structure is composed of the following main units: data extraction, data preprocessing, transliteration, tokenization, Machine learning algorithm and results analysis.

B. Data Collection

In this section, we describe our strategy for collecting data from Facebook pages/groups.

1) Data extraction is the act of analyzing and crawling data to get information from data sources in different structures and forms. We sourced our dataset from Facebook, a popular social networking website where users can communicate, show information about themselves, and post news on various domains.

We tried to extract posts and comments using several tools such as “instant Data Scraper” and “comments exporter.com”, we got the type of export, page name, timestamp, source URL, Profile name, post content, comments, likes and so many other information that we downloaded as an Excel file.

2) Data preprocessing: We first remove manually unnecessary columns like Profile name, timestamp, comments source. Regarding the comments column, we transliterate Arabic characters into Latin ones, convert letters from uppercase to lowercase, gather stop words in a file and remove them from comments in order to keep only the meaningful words. The second step was to label 1 all comments meaning that the post is fake and 0 otherwise. We can represent those steps by the following algorithm:

Algorithm1. (Data preprocessing)

1. For each comment
2. Comment is a sticker/URL/ads/names?
   → Remove comment
3. Comment contains Arabic letters?
   →Transliterate (letters)
4. Comment contains stop words?
   → Keep only keywords (fake or real)
5. Comment contains uppercase letters?
   → Lowercase (letters)
6. Keyword==fake? →Label it 1
7. Keyword==real? →Label it 0
8. End for

3) Transliteration*:* In order to normalize the language used in comments, we transliterate Arabic letters into Latin using a python function. The example below shows the result obtained after transliteration of an Arabic post

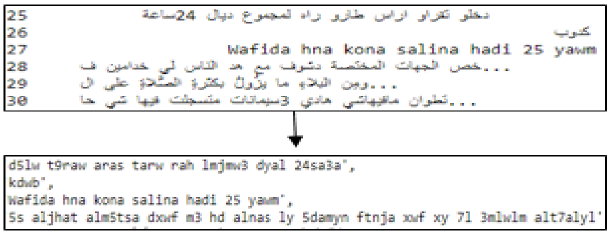

Fig. 2. An Arabic post transliteration

4) Tokenization is the act of splitting up a sequence of strings into words, phrases, or symbols. So after retrieving the comments column and transliterated all characters, a python script tokenized comments into words using the natural language toolkit (NLTK) (Fig.3)

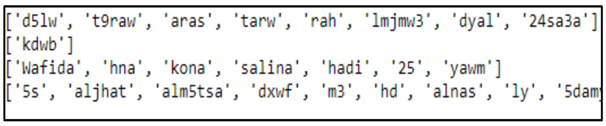

Fig. 3. A comment tokenized using NTLK

C. Our Machine Learning Approach

1) Feature extraction*:* A feature is the pattern in data set that is used to extract relevant data in order to train and increase the accuracy of models. The process of feature extraction consists of reducing the number of variables in dataset to get informative and non-redundant values and transform it into a form more suitable for subsequent classification and that require few computing resources to process.

2) Logistic regression*:* Logistic Regression refers to a machine learning technique that is used for binary classification problems where the value of the target variable is categorical. The method is commonly used when the dependent variable belongs to one class or another or can be coded as (1 or 0) to refer (yes or no) respectively.

This algorithm uses a linear equation with independent predictors called sigmoid function to predict a value that should be squashed into a range of [0, 1] by the following function:

$$g(x) = \frac{1}{1+e^{-x}} \quad (1)$$

3) Train, validation and test: The training, validating and testing of data are three important concepts in Machine Learning; the first concept is used to make sure that the machine recognizes patterns in the data, the second one is used to evaluate and ensure better accuracy and efficiency of the algorithm for the train. The third concept is used to evaluate the ability of predicting new answers based on the training step.

Post id	Percentage of keyword fake	Percentage of keyword real	State of the post
1	60%	40%	Fake
2	20%	80%	Real

Table 2. Example of posts classification

In the next section, we suggest our algorithm to determine whether the content of a post is fake or not by calculating the occurrence percentage of keywords like fake and real.

D. Results Analysis

1) Algorithm *:*

Algorithm2. (Fake news classifier)

Input: Posts, comments,
Output: Verdict

Begin Algorithm

1. // browsing posts
2. **For each** post (PostContent, CommentsContent, Comments)
3. // browsing comments and calculating the number of keywords occurrences
4. **For each** comment in Comments
5. **If** CommentsContent=='fake'
6. PostContent[fake]= PostContent[fake]+1
7. **If** CommentsContent=='real'
8. PostContent[real]= PostContent[real]+1
9. **End if**
10. // determine the veracity of the post according to the percentage of occurrences
11. **If** (PostContent[fake]/count(Comments)*100) > (PostContent[real]/count(Comments)*100)
12. determine PostConetent as fake
13. **else**
14. determine as real
15. **End if**
16. **End for**
17. **End For**

End Algorithm

Fig. 4. The proposed Algorithm to predict Fake posts

2) Process Description*:* We can get the predictive results by following the steps given in the previous algorithm. In this context, the idea is to browse all posts in the dataset, then browse and apply treatments on each post's comments to deduce its veracity.

In our case study, we worked on just one post, however, the same processing performed on one post can be generalized and applied to all posts in the dataset, so this is what we described by our algorithm.

To do this, we denote the arguments taken as input by:

- **Posts:** refers to the identifiers of posts to be classified
- **Comments:** refers t**o** the list of comments for each post.

The output of the algorithm is denoted by **Verdict** which refers to the fact that the input of the post is fake or real. We also define two variables: **CommentsContent** and **PostContent**, the first one is a string that takes as value the keyword "fake" or "real" depending on the content of each browsed comment (**line 5 & line 7**). The second is a list of length 2. Instead of using a counter going from 0 to 1 as index, we chose to reference them as **PostContent [fake]** instead of **PostContent [0]** and **PostContent [real]** instead of **PostContent [1]** to make the algorithm more understandable. The list items are incremented according to the value of the **CommentContent** variable (**line 6 & line 8**).

We also defined two functions:

- **count ():** which is used to count the total number of comments browsed for each post (**line 11**).
- **determine ():** which identifies the post as fake if the percentage of comments saying that the post is fake is greater than the percentage of those saying that the post is real. Otherwise, it identifies it as real news (**line 12 & line 14**).

E. Case Study: COVID'19 relatedFake News

1) Context*:* In this work, we tried to detect and analyze COVID'19 related fake news in Morocco. To this end, we illustrate our approach through a case study using this post **"22 new cases infected with Coronavirus in the Tangier and Tetouan region on May 28, 2020"**. While the ministry of health announces daily at 6 p.m., the total number of newly infected cases in every region, this post has been republished at 11 p.m. as a new detection of infected cases, therefore the subscribers of the page were confused and their comments were either to confirm or deny if such news is true.

2) Data Collection*:* As there are no datasets available for the detection and classification of COVID'19 related fake news in Morocco, we build our dataset by collecting data from comments. Therefore, our dataset contains more than 80 comments. After extracting data, we proceed then to the next steps including data preprocessing, transliteration, and tokenization in order to prepare the classification phase.

3) Classification: By classifying different comments, we obtained 50 keywords meaning that the news is fake while 30 keywords mean that the news is real (Fig.5).

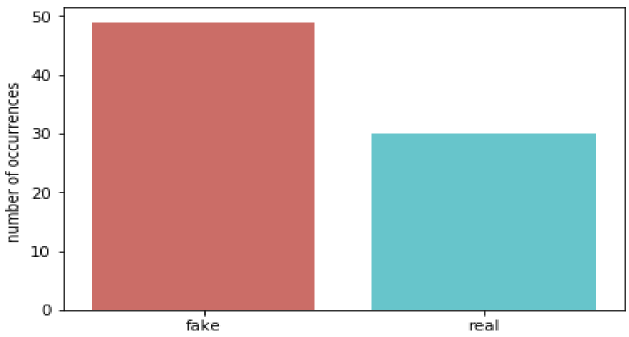

Fig. 5. A comment tokenized using NTLK

As results, we obtained the percentage of keywords as Fake (rep. Real) as follows: 62% (resp. 37%)

```
percentage of fake is 62.0253164556962
percentage of real 37.9746835443038
```

Fig 6. Percentage of keywords

*4) **Prediction**:* After having trained our model we moved on to the prediction phase, until now in our case study we worked with only one post, the method is then generalized for around twenty posts which are all related to the same COVID'19 news, this integration step enables our method to form a more accurate prediction. Yet, since we have encountered several problems and limitations that we're going to cite in the "discussion" section, it is obvious that some predictions are going to be correct and some aren't. Indeed, some of them are classified as "real" posts despite that they are fake information and this is due to several factors such as the equality of the percentages of the fake comments and real ones, and the absence of an additional criterion which we can consider to have a correct prediction.

VI. DISCUSSION

In this section, we discuss some issues encountered while conducting this project. The major problem we faced during data collection is the fact that Facebook restricts data collection and impose a specific process to be followed to obtain authorization. Although we have followed the process and fulfilled the requested criteria, we couldn't collect the data. To face this situation, we used alternatives techniques such as "Instant Data Scraper" and "Comments extractor". However, the results are still not very satisfactory.

Moreover, logistic regression is a multivariable model that is widely used in real-world applications in binary classification. Our choice stems from the fact that this algorithm does not need to know in advance the theoretical rules linking the input of the model to its output, it just learns them based on the data. Furthermore, while we compare "regression logistics" to other algorithms such as Naive Bayes, RandomTree in terms of precision, the ROC (Receiver Operating Characteristic) curve or the recall curve, we have found that "regression logistics" gives -in many cases of study- a better accuracy compared to others, and the same Recall as RandomTree. As for the naive Bayes algorithm, it displays the area closest to the ROC curve.

However, the algorithm presents some limitations. In fact, as we used "open source" tools during the extraction phase, we can't customize the recovered comments. By default, the extraction starts from the first comment up to a limited number while eliminating the remaining "responses". Also, it may be that half of the comments retrieved are either "douaa", advertisements or others while those eliminated carry more value and can increase the performance of the model. In such situation, some unavailable data as well as tests on the parameters of the machine learning model could have an impact on the performance of the model.

In this context, several improvements can be added to evolve more our model and to make it more efficient. Indeed, we can have very satisfactory results by working on well collected and cleaned data either by using special software which is often proprietary or by increasing the time devoted to the collection, cleaning, and loading of data, which constitute the most important phase of a Machine Learning project. We can also improve our system by adding the ability to verify the information published in the image form by working on text extraction and recognition from images.

VII. CONCLUSION AND PERSPECTIVES

In this paper, we suggest analyzing comments on Facebook posts in order to detect fake news using a machine learning approach.

As we chose to use "logistic regression" algorithm in this work, we plan to combine other methods to detect fake news in social media with high accuracy using SVM and the naïve Bayes classifier in our future work and then to compare these different approaches.

Furthermore, we plan to improve the process of collecting and processing data for the Arabic text, especially when it comes to working on fake news in videos often called DeepFake.

REFERENCES

[1] A. Thota, P. Tilak, S. Ahluwalia, N. Lohia, S. Ahluwalia, and N. Lohia, "Fake News Detection: A Deep Learning Approach," *SMU Data Sci. Rev.*, vol. 1, no. 3, p. 10, 2018.

[2] S. Robinson and C. Deshano, "'Anyone can know': Citizen journalism and the interpretive community of the mainstream press," *Journalism*, vol. 12, no. 8, pp. 963–982, Nov. 2011, doi: 10.1177/1464884911415973.

[3] D. M. J. Lazer *et al.*, "The science of fake news: Addressing fake news requires a multidisciplinary effort," *Science (80-.).*, vol. 359, no. 6380, pp. 1094–1096, Mar. 2018, doi: 10.1126/science.aao2998.

[4] E. C. Tandoc, Z. W. Lim, and R. Ling, "Defining 'Fake News': A typology of scholarly definitions," *Digital Journalism*, vol. 6, no. 2. Routledge, pp. 137–153, 07-Feb-2018.

[5] D. Khurana, A. Koli, K. Khatter, and S. Singh, "Natural Language Processing: State of The Art, Current Trends and Challenges," no. Figure 1, 2017.

[6] E. Alpaydin, *Introduction to Machine Learning*, Second Edi. The MIT Press, 2010.

[7] A. Lakshmanarao, Y. Swathi, and T. Srinivasa Ravi Kiran, "An effecient fake news detection system using machine learning," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 10, pp. 3125–3129, 2019, doi: 10.35940/ijitee.J9453.0881019.

[8] Y. Liu and Y. F. B. Wu, "Early detection of fake news on social

media through propagation path classification with recurrent and convolutional networks," in *32nd AAAI Conference on Artificial Intelligence, AAAI 2018*, 2018, pp. 354–361.

[9] M. Farajtabar *et al.*, "Fake news mitigation via point process based intervention," *34th Int. Conf. Mach. Learn. ICML 2017*, vol. 3, no. Icml, pp. 1823–1836, 2017.

[10] S. Tschiatschek, A. Singla, M. Gomez Rodriguez, A. Merchant, and A. Krause, "Fake News Detection in Social Networks via Crowd Signals," pp. 517–524, 2018, doi: 10.1145/3184558.3188722.

[11] K. Shu, D. Mahudeswaran, and H. Liu, "FakeNewsTracker: a tool for fake news collection, detection, and visualization," *Comput. Math. Organ. Theory*, vol. 25, no. 1, pp. 60–71, Mar. 2019, doi: 10.1007/s10588-018-09280-3.

[12] K. Shu, S. Wang, and H. Liu, "Understanding User Profiles on Social Media for Fake News Detection," in *Proceedings - IEEE 1st Conference on Multimedia Information Processing and Retrieval, MIPR 2018*, 2018, pp. 430–435, doi: 10.1109/MIPR.2018.00092.

[13] K. Shu, S. Wang, and H. Liu, "Beyond news contents: The role of social context for fake news detection," in *WSDM 2019 - Proceedings of the 12th ACM International Conference on Web Search and Data Mining*, 2019, pp. 312–320, doi: 10.1145/3289600.3290994.

[14] S. Singhal, R. R. Shah, T. Chakraborty, P. Kumaraguru, and S. Satoh, "SpotFake: A multi-modal framework for fake news detection," in *Proceedings - 2019 IEEE 5th International Conference on Multimedia Big Data, BigMM 2019*, 2019, pp. 39–47, doi: 10.1109/BigMM.2019.00-44.

[15] K. Stahl, "Fake news detection in online social media Problem Statement," no. May, p. 6, 2018.