Using a Rule-based Model to Detect Arabic Fake News Propagation during Covid-19

Fatimah L. Alotaibi, Muna M. Alhammad

Collage of Business Administration

King Saud University

Riyadh, Saudi Arabia

Abstract—Since the emergence of the Covid-19, both factual and false information about the new virus has been disseminated. Fake news harms societies and must be combated. This research aims to identify Arabic fake news tweets and classify them into six categories: entertainment, health, politics, religious, social, and sports. The study also aims to uncover patterns in the spread of Arabic fake news associated with the Covid-19 pandemic. The researchers created an Arabic dictionary and used text classification based on a rule-based system to detect and categorize fake news. A dataset consisting of 5 million tweets was analyzed. The developed model achieves an overall accuracy of 78.1% with 70% precision and 98%recall. The model detected more than 26006 fake news tweets. Interestingly we found an association between the number of fake news tweets and dates. The result demonstrates that as more information and knowledge about Covid-19 become available over time, people's awareness increase, while the number of fake news tweets decreases. The categorization of false news indicates that the social category was highest in all Arab countries except Palestine, Qatar, Yemen, and Algeria. Conversely, fake news related to the entertainment category was the weakest dissemination in most Arab countries.

Keywords—Fake news; Covid-19; text classification; rule-based system; trends

I. INTRODUCTION

Social media networks are increasingly being used as a source of information. Its low cost, easy access, and speed of information delivery encourage people to use it as a searching tool to obtain information. The widespread use of social media around the world provides a perfect setting for the dissemination of fake news. Fake news is described as low-quality news [1] or false news purposely broadcast to divert people away from true news and facts [2]. It is considered to be one of the most dangerous weapons capable of bringing harm to both society and people [3]. Its prevalence increases during crises as crises represent a perfect time for spreading fake news and rumors, especially when knowledge of a problem is limited and ambiguous.

Due to the breakout of the novel coronavirus "Covid-19" in 2020, the world faced health and economic concerns. The virus was first discovered in China in November of 2019 then formally declared a pandemic by WHO [4] on the 11th of March 2020, making it the worst global crisis of the 21st century. Many people are afraid and anxious because of the virus's secrecy and uncertainty. Therefore, both true and false news began to spread about this virus. Social network websites have become essential for obtaining news and information

about Covid-19 [5]. With curfew and lockdown being implemented in many countries, social network websites were full of endless discussion about the Covid-19, with much fake news being exchanged. Many individuals intentionally propagated social media with fake news to gain personal benefits, such as having a lot of likes and followers. In addition, some companies and organizations have benefited from the spread of false and misleading information on social media to promote and advertise products or services to increase their sales profit [6]. The propagation of fake news had a negative influence on public health throughout this crisis, increasing tension, rage, anxiety, panic, and depression [2]. It has been documented, for example, that some consumers in the United Kingdom and Australia experienced panic-buying while purchasing a specific product, such as toilet paper [7].As a result, the negative impact of the propagation of fake news may lead to more severe problems within society as well as the Covid-19 problem.

As a response to control the spread of fake news during the pandemic, many countries and governments have taken the issue seriously and introduced new laws to prevent the spread of such information. For example, Twitter announced that it would remove any misleading and unspecified content about Covid-19. Moreover, social network websites, such as Facebook, Google, WhatsApp, and Microsoft, have pledged to work with governments to fight the spread of fake news. In the middle east, Saudi Arabia is one of the countries that is seriously fighting the spread of misleading information about the Covid-19 pandemic. The Public Prosecution in Saudi Arabia applied high penalties against fake news propagators with fines of up to 3 million SAR and imprisonment of up to 5 years.

Although the media paid considerable attention to studying the spread of fake news [8], most of these studies were conducted in western countries, with the English being the primary language of the studied samples. There are few empirical investigations of the diffusion of fake news in the Arab region. Cultural disparities are evident, highlighting the importance of researching the dissemination of fake news in many cultural contexts. In addition, studies of the diffusion of fake news during the time of pandemics and crises are currently limited. Therefore, this research aims to answer the following two questions: how to detect Arabic fake news about Covid-19 on Twitter through the use of rule-based systems, and what are the current trends of fake news diffusion during the pandemic in the Arab region?

This paper starts with reviewing the literature, followed by introducing the methodology used in this research. Data analysis results and discussion will then be presented. Finally, the conclusion, limitations, and future work will be presented.

II. LITERATURE REVIEW

A. Fake News

Since the inception of fake news, there has been no single definition of fake news [3]. According to [9], fake news is also defined as a news article intentionally written to deliver false information for a different purpose. Therefore, fake news is described in this study as manipulated false or misleading information to appear like real news for several purposes.

Several researchers were interested in studying the spread of fake news. For example, Allcott et al.[10] did a study to measure trends in the diffusion of misinformation circulating on social media from January 2015 to July 2018. The result from their research shows users' interactions with incorrect information rose steadily on both Facebook and Twitter up to the end of 2016. However, after one month, the interactions with false information dropped sharply on Facebook while rising on Twitter. This may result from a change in the Facebook platform after the 2016 elections to combat fake news. Another study was conducted by [2] to develop a method to overcome the spread of fake news in health during the current outbreak. The study focused on determining the type of false health information and used social impact in social media (SISM) methodology to analyze data. In addition, they selected Facebook, Twitter, and Reddit as social media channels for analysis. The study found that posts focused on fake health information are most aggressive. The spread of fake news during the current pandemic, i.e., Covid-19, made many individuals fearful and panicked. According to [2], psychological and neurological problems increased during the current pandemic, with fake news playing a significant role. In a recent study, researchers designed a dashboard to track misinformation on popular social media news sharing platforms (i.e., Twitter). To do so, they collected data from the platform beginning on March 1, 2020, and up to date [5]. This dashboard aims to fight false information and increase awareness about Covid-19.

Fake news can be detected by using three main methods. The first method relies on analyzing the news context where linguistic features are extracted and analyzed to identify fake news based on the writing styles that are commonly used in fake news [11, 12]. In addition, visual features can be used and analyzed to identify fake images [3]. The second method relies on analyzing social context where user-based, post-based, and networks-based segments are analyzed to determine fake news. User-based features such as users' profiles and characteristics can be analyzed to detect fake news through identifying the source of the fake news [12]. Post-based features mainly identify fake news by extracting and analyzing people's responses toward fake news [3]. Networks-based featured can remove by building specific networks among the users who published related posts [3]. The third and most recent one is using a knowledge-based context model, which aims to use external sources to check facts and identify fake news [13]. This last method is currently widely used because of its accurate results and the increasing number of available high-quality fact-checking websites.

B. Text Mining and Text Classification

With the massive volume of data, companies and organizations started to use text mining techniques to improve their services, monitor brand reputation, gain a competitive advantage, and understand customers' behavior [14]. Text mining is used to extract hidden meaningful information from text [15, 16]. Moreover, text mining deals with unstructured data, while data mining deals with structured data [15, 17]. The process of mining text includes several stages: data gathering, data preparation, text transformation, feature selection, pattern selection, and evaluation [17]. Text mining involves a large set of algorithms and techniques for analyzing text, such as information retrieval, natural language processing, text summarization, classification (supervised learning), and clustering (unsupervised learning) [15]. These algorithms and techniques are being used in several contexts such as Business Intelligence, Customer care services, Knowledge management, Bioinformatics, Web Search Enhancement, and Risk Management [17].

In detecting fake news, the study [18] used semantic features and text mining to detect the spread of fake news in online articles. The used dataset was obtained from Kaggle about real-or-fake news. They applied five semantic features, including term frequency (TF), term frequency-inverse document frequency (TFIDF), bigrams, trigrams, quad-grams, and glove word embeddings along with Naive Bayes, random forest, and recurrent neural networks (RNN) classifiers. The study points out the bigram features with the random forest classifier achieved the best accuracy of 95.66%. So, text mining techniques are beneficial for finding misinformation. Another study was conducted by [19] to understand the impact of Covid-19 in Mexican society by using the text mining approach. The study extracted Twitter tweets about Covid-19 from the 13th to the 20th of March 2020, and the geo-localization of the retrieved tweets was Mexico City. The study found a positive correlation between the number of these tweets each day and the number of positive Covid-10 cases reported by the government on the same day. In addition, the people were fearful of health risks and economic crises that Covid-19 may cause. This required developing strategies to comfort the fear and panic associated with Covid-19.

Text classification is one of the essential methods in supervised learning [17]. It aims to assign classes or labels to texts, and it is used in many applications like image processing, document organization, etc. [15]. There are many learning algorithms used for text classification, such as Naïve Bayes Classifier, decision trees, Neural networks, and rule-based classifier. A team of researchers has developed a text-based algorithm, i.e., a supervised decision tree model, for analyzing customers' comments about a famous food brand on Twitter. The developed model predicted about 85% negative comments and 15% positive comments after analyzing 500 tweets [14]. A recent study applied text classification on Twitter data to analyze public fear sentiment during the Covid-19 pandemic [20]. Two machine learning (ML) classification methods were used, i.e., Naïve Bayes and logistic regression. Over nine hundred thousand tweets from February to March of 2020 were

analyzed using R software. The Naïve Bayes was able to classify public fear sentiment with a 91% accuracy rate, while the logistic regression was able to classify public fear sentiment with only a 74% accuracy rate. It is noted that the text classification helped in understanding the feeling of the public. Hence, it can be used to study social phenomena such as the spread of fake news.

1) Rule-based classifier: The rule-based classifier is one approach to text classification which uses a set of rules to separate the text into distinct groups [21]. Each of these rules contains an antecedent part, and a consequent part uses a series of "if-then" to represent them [22]. In [23] used a rule-based classifier with supervised learning to perform a Twitter sentiment analysis. Rules were built based on the occurrences of words related to emotion and opinion within the text. The study found that a rule-based classifier can improve the support vector machine SVM's predictions.

In summary, studies concerned with detecting fake news used either rule-based models or machine learning techniques to detect fake news. In this research, a rule-based classifier with a developed dictionary will be used to classify and detect fake news.

III. RESEARCH METHODOLOGY

Twitter data analysis has occupied a large volume of research in recent years, with many researchers relying on text mining techniques such as classification and clustering [16]. This study will develop an Arabic dictionary to detect fake news on social media (i.e., Twitter) and study the propagation of fake news in the Arab region. For our Arabic dictionary, we gathered fake news about the Covid-19 pandemic from various fact-checking websites, then built our dictionary around it. Our detection model is based on text classification that depends on rule-based systems to classify the tweets as fake or not fake news. The method of this study was divided into three steps: data collection, data preparation, and model building. Besides, we used python language (Jupyter notebook) to preprocess and analyze the dataset, Microsoft Excel to build the dictionary, and Power BI for data visualization.

A. Data Collection

We used a dataset of Arabic tweets on Covid-19 released on GitHub by Dr. Sultan Almujaiwel[1] after getting his consent to use the data. The data were gathered between March 1, 2020, and April 30, 2020. To retrieve the dataset from Twitter, he used Arabic hashtags related to coronavirus disease to retrieve relevant tweets. Those hashtags are "كورونا#, #covid-19", "فيروس_كورونا#, #coronaviruse", "منع_التجول#, #curfew", "كورونا_المستجد#, #novel_coronaviruse", "حظر_التجول#, #curfew", "العالم_في_مواجهة_كورونا#, and #world_facing_covid-19". The number of tweets in this dataset was 5015111 tweets.

B. Data Preparation

To clean and prepare the dataset for analysis, a series of processes were conducted. Noised tweets containing ads, coupons, and other irrelevant tweets were removed, as were old

tweets tweeted on a date other than the study's covered period. Arabic and English punctuation, numbers, hashtags, emoji, and empty tweets were removed too. Additionally, English letters, @username, and website links were replaced with an empty string. We also normalized the tweets in Arabic. Furthermore, missing values in the Location and Username columns were replaced with "undefine" values. As the data preparation finishes, the total number of tweets eligible for analysis was lowered to 4643425.

C. Model Building

To analyze the data, we performed two steps to building the model:

1) The development of fake news dictionary: The main goal of this study's dictionary development is to categorize Arabic tweets that contain fake news. Tweets that have spread in Arab communities are included in the dictionary. Initially, fake news was gathered from fact-checking websites as well as the fake news set published by [24] in GitLab[2]. The fact-checking websites that were used in this study are Misbar[3], Norumors[4], AFPfact check[5], Fatabyano[6], Google fact check tools[7]. Duplicate news was checked and removed as it was gathered from various websites. A total of 212 pieces of fake news were discovered. Each piece of news was then tokenized, and the seven most important words were taken into account. One of these words has been identified as the primary key to fake news, with the remaining words serving as secondary keys. Table I shows two examples of fake news and how to extract the crucial words, where the "Fake" column represents the fake news, the "Key" column represents the fake news's primary key, and the reminder columns represent the fake news's secondary keys[8]. Following that, the fake news in the dictionary was classified into six categories, which corresponded to the categories identified by [24]. Table II. represents the categories and the number of relevant fake news categorized under each category in our dictionary. An example for each of the categories is presented in Table III.

2) Rule-based model development: The rule-based model was performed using two steps. First, the rule-based model will determine whether the primary key in the dictionary of fake news can be found in the text of the tweet. If the first condition is met, the model will check to see if at least two words from the dictionary's secondary words are found in the tweet's text. As a result, if the two conditions are met, the model will verify that the text does not contain any vocabularies that reject fake news, such as: ('false claim- ادعاء المعلومات زائف'; 'this is not true-لا صحه لذلك '; 'misinformation- المعلومات المغلوطة'). This is significant because some accounts on Twitter

---

[1] https://github.com/salmujaiwel

[2] https://gitlab.com/bigirqu/ArCOV-19
[3] MISBAR. 2020. Available: https://misbar.com/
[4] RUMORS, N. 2020. Available: http://norumors.net/
[5] ORGANIZATION, F. C. 2020. Available: https://factcheck.afp.com/
[6] FATABYYANO. 2020. Available: https://fatabyyano.net/
[7] TOOLS, G. F. C. 2020. Available:
https://toolbox.google.com/factcheck/explorer
[8] The fake news dictionary we have developed is available on GitHub: https://github.com/FatimahLAlotaibi/fake-news-dictionary

fight and reject fake news. Therefore, by checking the existence of these vocabularies, one can verify that the tweet categorization as fake news is accurate. If the preceding conditions are met, the fake news category will be retrieved from a dictionary and assigned to the relevant tweet. Otherwise, the tweet will be categorized as not-fake news. Fig. 1 depicts the code used to determine whether a tweet contains fake news and to assign tweets to the appropriate category. The second step aims to label tweets based on the category, with label =1 indicating that the tweet contains fake news and label =0 indicating that the tweet does not contain fake news. To do so, we look to see if the category contains a "not" value, which indicates that the label is equal to zero; otherwise, the label is equal to one. Fig. 2 depicts the code that was used to assign the tweets to the label.

TABLE I. EXAMPLES OF FAKE NEWS IN THE DICTIONARY

Fake	Key	S1	S2	S3	S4	S5	S6
ثلاثة ادوية تسبب الموت السريع عند الإصابة بفيروس كورونا أو الاشتباه به	ادويه	ثلاث	أصابه	السريع	سبب	تؤدي	اشتباه
Three medicines cause rapid death when infected with coronavirus or suspected of it.	Medicines	Three	Infection	Rapid	Causes	lead to	Suspected
رمي جثث موتى كورونا في البحر بالمكسيك	جثث	بحر	حذف	رمي	مكسيك	قتلى	موتى
Throwing corpses of people dead by Corona in the sea in Mexico.	Corpses	Sea	Disposal	Throw	Mexico	Dead	Dead

TABLE II. CATEGORIES OF FAKE NEWS IN THE DICTIONARY

Category	Entertainment	Health	Politics	Religious	Social	Sports	Total
Number of fake news	13	41	29	33	89	7	212

TABLE III. EXAMPLE OF EACH CATEGORY OF FAKE NEWS

Fake (Arabic)	Fake (English)	Category
الممثلة الإسبانية ألبا فلوريس التي اشتهرت بدور نيروبي توفيت بسبب فيروس كورونا المستجدّ	Spanish actress "Alba Floris" who is known as "Nairobi" was death by Covid-19.	Entertainment
ثلاثة ادوية تسبب الموت السريع عند الإصابة بفيروس كورونا أو الاشتباه به	Three medicines cause rapid death when infected with coronavirus or suspected of it.	Health
سياسيّ فرنسيّ يرمي علم الإتحاد الأوروبي بسبب تخلّيه عن الدول الأعضاء خلال أزمة كورونا المستجدّ	French politician throws the European Union flag due to its abandonment of the member states during the Corona pandemic.	Politics
أول صلاة جمعة في الصين بعد إغلاقها اعوام واعوام	The first Aljumuah prayer in China after being close for years	Religious
وزارة الداخلية تحذر فيه من سرقات عديدة حصلت بسبب عصابات تدعي أن هدفهم تعقيم البيوت من فيروس كورونا	The interior ministry warns of numerous thefts that happened due to gangs claiming their goal is home sterilization from coronavirus.	Social
إصابة حارس مرمى نادي ريال مدريد ومنتخب بلجيكا، تيبو كورتوا، بفايروس كورونا.	The goalkeeper of Real Madrid and the Belgian national team "Tibo Coroto" has been infected with the coronavirus.	Sports

```
category='not'
    count=0
    i=0
    while i<len(s):
      if s['key'][i] in (text):# if basic word from dictionary in tweet
        if s['s1'][i] in (text):# if secondary word from dictionary in tweet
          count=count+1
        if s['s2'][i] in (text):
          count=count+1
        if s['s3'][i] in (text):
          count=count+1
        if s['s4'][i] in (text):
          count=count+1
        if s['s5'][i] in (text):
          count=count+1
        if s['s6'][i] in (text):
          count=count+1
        if count>1:
          category=s['Category'][i] # assign category to tweet
          for x in R:  # R is list of word that reject fake news
            if x in (text):
              category="not"
              break
          break
      count=0
      i=i+1
```

Fig. 1. Code to Check if the Tweets are Fake News or Not.

```
data_df1["model-label"]=0
i=0
while i<len(data_df1):
  if data_df1["Category"][i]!= "not":
      data_df1["model-label"][i]=1
  i=i+1
```

Fig. 2.   Code Assigning the Tweets to the Matching Label.

IV. RULE-BASED MODEL EVALUATION

To evaluate and confirm the proposed model, a balanced sample of 2000 tweets was extracted at random. The tweets in the sample were labeled manually, with label=1 indicating fake news and label=0 indicating not-fake news (i.e., real news or others). In this sample, there were 997 tweets labeled as fake news and 1003 tweets labeled as not fake news. The sample was then divided into 70 percent (1400 tweets) for training our model and the remaining 30 percent (600 tweets) for testing. The performance of the proposed model was assessed using accuracy, precision, and recall. These three metrics are measured by calculating the number of correct positive predictions (TP), the number of correct negative predictions (TN), the number of incorrect positive predictions (FP), and the number of incorrect negative predictions (FN) [9], as shown respectively in equations 1, 2, and 3.

$$\text{Accuracy} = \left(\frac{(|TP|+|TN|)}{|TP|+|TN|+|FP|+|FN|} \right) \quad (1)$$

$$\text{Precision} = \left(\frac{(|TP|)}{|TP|+|FP|} \right) \quad (2)$$

$$\text{Recall} = \left(\frac{(|TP|)}{|TP|+|FN|} \right) \quad (3)$$

Table IV displays the precision, recall, and accuracy of our model when applied to training and testing data in this sample. Our model's prediction level is adequate based on these results. As a result, we may use the model to examine our data.

TABLE IV.   ACCURACY, PRECISION, AND RECALL FOR VALIDATING THE RULE-BASED MODEL

	Training Data	Testing Data
Accuracy	79.7%	78.1%
Precision	71.2%	70%
Recall	99%	98%

V. MODEL EVALUATION USING FUTURE DATASET

To validate the proposed model, the researchers obtained 545 tweets at random as future data. The tweets were then manually labeled. When the model was applied to these data, the accuracy was 95.9 percent, indicating that the model was acceptable and applicable.

VI. RESULTS AND DISCUSSION

A. Fake news Propagation Trends across the Arab Region

After ensuring that our model had an acceptable level of accuracy, it was applied to the dataset. 26006 tweets were labeled as fake news by the model. According to an analysis of the source of these tweets, only 13491 fake news tweets have location details. This is because location data on Twitter is provided only if the user has a geo-enabled feature or mentions a valid location in his public profile [5]. The analysis of the location of fake news tweets reveals that the majority of fake news tweets originated in Saudi Arabia, Egypt, Kuwait, Qatar, Yemen, United Arab Emirates, Iraq, Palestine, Lebanon, Jorden, Oman, Algeria, Morocco, Libya, Sudan, Syria, Bahrain, and Tunisia respectively (see Fig. 3). Given that the majority of the analyzed tweets originated from Saudi Arabia and Egypt, it was expected that a high number of fake news tweets are being exchanged in these countries. This is also consistent with recent statistics showing that Saudi Arabia and Egypt are ranked 8th and 18th in the world in terms of the number of Twitter users, respectively [9]. However, the high number of fake news tweets from Kuwait, which is so close to Egypt, is surprising. Kuwait is a small country with a population of only 4.4 million [10] people. Nevertheless, about 99.5% [11] of its population are current Internet users, and more than half of them are Twitter users. Therefore, this may have affected the spread of fake news in Kuwait.

As previously stated, the model was trained to categorize fake news tweets into six categories based on the general theme of the tweet. Table V shows the number of tweets in each category of fake news.

B. Fake News Word Cloud

After extracting the fake news tweets from the dataset, we used a word cloud to analyze the frequency of the words in these tweets. However, it was decided to exclude two words (corona "كورونا"; and virus "فيروس") because they were found to harm the word cloud's results. Fig. 4 depicts the word cloud results for the most common words found in fake news tweets. The most frequent tweeted fake news was related to the daily use of hot steam inhalation to 'kill' the Covid-19 coronavirus.
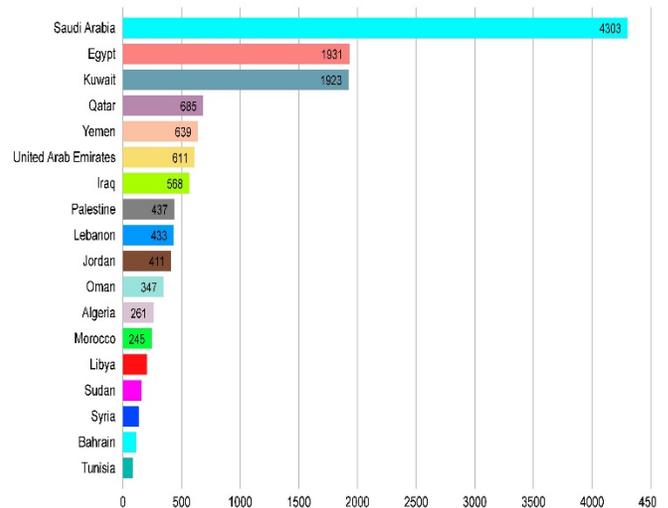

Fig. 3.   Number of Fake News in Arab Countries.

[9] https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/

[10] Central Statistical Bureau, https://www.csb.gov.kw

[11]The World Bank Open Data, https://data.worldbank.org/

TABLE V.   CATEGORIES OF FAKE NEWS TWEETS AND NUMBER OF RELEVANT TWEETS

Category	Number
Entertainment	629
Health	3911
Politics	4032
Religious	6338
Social	9781
Sports	1315
Total	26006

Fig. 4.   Word Cloud for Fake News.

C. Analysis of the Fake News across the Dates

The dataset for our study was obtained from Twitter between March and April. After analyzing the data, we discovered that the spread of fake news tweets was much higher in March than in April. Fig. 5 shows that approximately 63.06% of fake news tweets were posted in March, while only 36.94% were posted in April indicating a decrease in the number of fake news tweets in April. These findings show that ambiguity surrounding the coronavirus and people's feeling of fear may have increased fake news propagation at the beginning of this crisis. Therefore, as users' awareness and knowledge of Covid-19 increases, the spread of fake news tweets decreases.
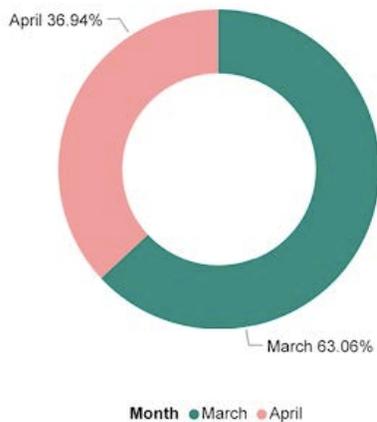

Fig. 5.   Distribution of Fake News Tweets across Categories.

Because the number of fake news tweets in March is higher than in April, we attempted to improve our understanding of the spread of fake news tweets in March by analyzing the distribution of fake news during the month's days. Fig. 6 shows that the 14th, 15th, 16th, and 25th of March saw the most incredible spread of fake news tweets in the Arab region. On or around March 15th, most Arab countries suspended in-class study and switched to online study for both schools and universities. Egypt, for example, temporarily suspended traditional education on March 15th, which explains why the 14th, 15th, and 16th of March have the highest number of fake news tweets. Furthermore, the partial curfew in Saudi Arabia began on March 23rd and in Egypt on March 25th, which could explain the rise in fake news on March 25th. This demonstrates that as more information and knowledge about Covid-19 become available over time, user awareness of the current pandemic grows, while the number of spread fake news tweets decreases.

As previously observed, Saudi Arabia, Egypt, and Kuwait had the highest number of fake news tweets. As a result, we attempted to comprehend the distribution of fake news during March and April. Fig. 7 shows that the number of fake news tweets in Egypt and Kuwait decreased by roughly half between March and April, but not significantly in Saudi Arabia. It is plausible that the total curfew imposed by Saudi Arabia in major cities on April 6th contributed to the continuation of fake news.
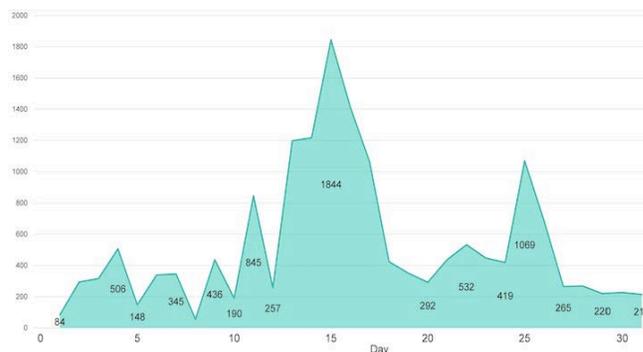

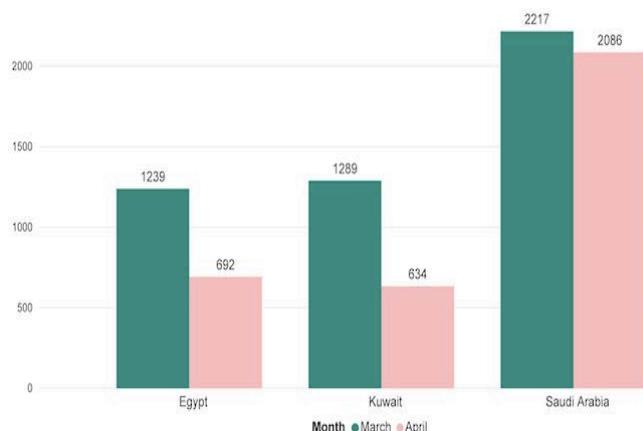Fig. 6.   Spread Fake News Tweets during March.

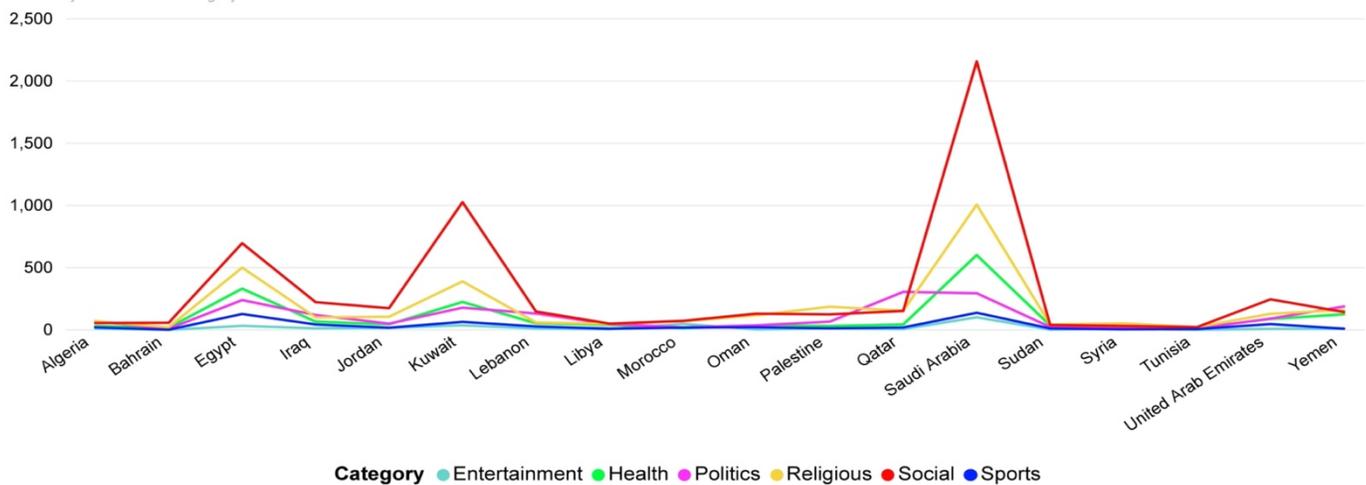Fig. 7.   Spread of Fake News Tweets by Months in the Top Three Arabic Countries

Fig. 8.   Categories of Fake News across Arabic Countries.

D. Distribution of Fake News Categories across Arab Countries

Fig. 8 depicts the various categories of fake news tweets in Arab countries. Fake news tweets about social issues are the most prevalent in all Arab countries except Palestine, Qatar, Yemen, and Algeria. In Palestine, fake news tweets were mostly about religion, whereas in Qatar and Yemen, fake news tweets were mainly about politics. In Algeria, fake news tweets were mostly about religious or political issues. Fake news related to entertainment, on the other hand, was the lowest across all Arab countries, except Morocco, which has a deficient number of fake news tweets related to health topics.

VII. CONCLUSION

This study was carried out to understand better the spread of the fake news phenomenon in the Arab region during Covid-19. We created an Arabic dictionary and used text classification, i.e., a rule-based system, to analyze the spread of fake news in the Arab region using a secondary dataset to detect fake news. We discovered that the number of fake news spreads on Twitter was much higher in March than in April and that dates with significant response measures taken to control the spread of Covid-19 had the highest number of fake news spreads. This figure usually falls a few days after the measurement is taken. This demonstrates that at the start of new changes, people panic, mainly due to a lack of information about what will happen next, and that as more information becomes available, users adapt to the new norms and stop spreading such fake news.

VIII. LIMITATION AND FUTURE RESEARCH

This study has some limitations. The primary limitation is the difficulty of processing Arabic. The Arabic language, unlike other languages, has distinct writing principles such as writing from the right to the left side, the absence of capital letters, and distinct grammatical rules for detecting entities, acronyms, and abbreviations. The presence of many dialects in the Arab region, the use of slang words and colloquial terms, and many spelling errors are also significant challenges when analyzing the Arabic language. The tweets dataset used in this study came from various Arabic countries, and the tweets were written in a variety of Arabic dialects, making it difficult to process and analyze texts using stemming. To improve the efficiency of the analysis, future work should focus on developing the steaming process for the Arabic language. Additionally, because new fake news is created periodically, the developed fake news dictionary must be updated regularly. In addition, the result of the proposed model in this study will be compared in the future with the results of the machine learning models to extra validate and assess the performance of the proposed model. Also, an additional dataset from other social media platforms, such as Facebook, will be used in the future to see if the same trends of fake news dissemination are observed across platforms.

REFERENCE

[1] A. Al-Rawi, Gatekeeping fake news discourses on mainstream media versus social media. Social Science Computer Review, 2019. 37(6): p. 687-704.

[2] C.M. Pulido, L. Ruiz-Eugenio, G. Redondo-Sama, and B. Villarejo-Carballido, A New Application of Social Impact in Social Media for Overcoming Fake News in Health. International journal of environmental research and public health, 2020. 17(7): p. 2430.

[3] K. Shu, et al., Fake news detection on social media: A data mining perspective. ACM SIGKDD explorations newsletter, 2017. 19(1): p. 22-36.

[4] A. Spinelli and G. Pellino, COVID‐19 pandemic: perspectives on an unfolding crisis. The British journal of surgery, 2020.

[5] K. Sharma, et al., Covid-19 on social media: Analyzing misinformation in twitter conversations. arXiv preprint arXiv:2003.12309, 2020.

[6] M.K. Elhadad, K.F. Li, and F. Gebali. COVID-19-FAKES: A twitter (Arabic/English) dataset for detecting misleading information on COVID-19. in International Conference on Intelligent Networking and Collaborative Systems. 2020. Springer.

[7] Y. Chen, et al., A discussion of irrational stockpiling behaviour during crisis. Journal of Safety Science and Resilience, 2020. 1(1): p. 57-58.

[8] C. Silverman, This analysis shows how viral fake election news stories outperformed real news on facebook, 2016. URL https://www. buzzfeed.

com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook, 2016.

[9] S. Yang, et al. Unsupervised fake news detection on social media: A generative approach. in Proceedings of the AAAI Conference on Artificial Intelligence. 2019.

[10] H. Allcott, M. Gentzkow, and C. Yu, Trends in the diffusion of misinformation on social media. Research & Politics, 2019. 6(2): p. 2053168019848554.

[11] Y. Chen, N.J. Conroy, and V.L. Rubin, Misleading Online Content: Recognizing Clickbait as "False News", in Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection. 2015, Association for Computing Machinery: Seattle, Washington, USA. p. 15–19.

[12] C. Castillo, M. Mendoza, and B. Poblete, Information credibility on twitter, in Proceedings of the 20th international conference on World wide web. 2011, Association for Computing Machinery: Hyderabad, India. p. 675–684.

[13] A. Vlachos and S. Riedel. Fact checking: Task definition and dataset construction. in Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science. 2014.

[14] A.S. Halibas, A.S. Shaffi, and M.A.K.V. Mohamed. Application of text classification and clustering of Twitter data for business analytics. in 2018 Majan International Conference (MIC). 2018. IEEE.

[15] M. Allahyari, et al., A brief survey of text mining: Classification, clustering and extraction techniques. arXiv preprint arXiv:1707.02919, 2017.

[16] S.A. Salloum, M. Al-Emran, A.A. Monem, and K. Shaalan, A survey of text mining in social media: facebook and twitter perspectives. Adv. Sci. Technol. Eng. Syst. J, 2017. 2(1): p. 127-133.

[17] S. Sheela and T. Bharathi, Analyzing Different Approaches of Text Mining Techniques and Applications. International Journal of Computer Science Trends and Technology, 2018. 6(4): p. 23-29.

[18] P. Bharadwaj and Z. Shao, Fake news detection with semantic features and text mining. International Journal on Natural Language Computing (IJNLC) Vol, 2019. 8.

[19] J.E.C. Saire and A. Pineda-Briseno, Text mining approach to analyze coronavirus impact: Mexico city as case of study. medRxiv, 2020.

[20] J. Samuel, et al., Covid-19 public sentiment insights and machine learning for tweets classification. Information, 2020. 11(6): p. 314.

[21] M. Thangaraj and C. Vijayalakshmi, Performance study on rule-based classification techniques across multiple database relations. International Journal of Applied Information Systems, 2013. 5(4): p. 1-7.

[22] U.A. Siddiqua, T. Ahsan, and A.N. Chy. Combining a rule-based classifier with ensemble of feature sets and machine learning techniques for sentiment analysis on microblog. in 2016 19th International Conference on Computer and Information Technology (ICCIT). 2016. IEEE.

[23] P. Chikersal, S. Poria, and E. Cambria. SeNTU: sentiment analysis of tweets by combining a rule-based classifier with supervised learning. in Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015). 2015.

[24] F. Haouari, M. Hasanain, R. Suwaileh, and T. Elsayed, ArCov-19: The First Arabic COVID-19 Twitter Database with Propagation Networks. 2020.