Topic-based Models with Fact Checking for Fake News Identification

Vlad Cristian Dumitru
University Politehnica of Bucharest
Bucharest, Romania
d.vladk@gmail.com

Traian Rebedea
University Politehnica of Bucharest
Bucharest, Romania
traian.rebedea@cs.pub.ro

ABSTRACT

The phenomenon of fake news has expanded more and more over the last period of time. This expansion is strongly related to the evolution of technology and social networks, which facilitates the distribution of information more than ever. In the context of this phenomenon growing up so fast and bringing unjust influences amongst people, studies have emerged on the automatic solutions proposed to combat fake news in the online environment. The study started with collecting a dataset that also contains more current information. As a first step the paper proposes a division of news according to topics using clustering algorithms. Subsequently, the system involves the training of a detection model within each cluster obtained, as well as the training of a global model. The study examines the possibility of using these models to form a mixed detection system. In the second part of the study, the focus is on conducting experiments based on the facts that appear in the news and the possibility of filtering the articles according to the relevant facts using an information news related database.

Author Keywords
Fake news detection, Linguistic features, Natural language processing, Topic-based model

ACM Classification Keywords
I.2.7 Natural Language Processing

DOI: 10.37789/rochi.2021.1.1.28

INTRODUCTION

Various forms of fake news have existed since the earliest times of history. Examples attesting to this can be found since antiquity, but with the invention of the printing press in the 16th-17th centuries, all sorts of unusual stories began to circulate to the public. It all led to mass distribution and sales of printed materials, regardless of their credibility and sources [1]. Propaganda, misinformation, deception or manipulation, all have in common the fact that they pursue various hidden purposes by masking the truth or by inducing misconceptions in society. These hidden purposes may include: political gain, financial gain or denigration of the image of a person or entity. The phenomenon has grown quite abruptly in recent years, with the US presidential election in 2016, where the term "*fake news*" has become a real trend. Five years later, the situation continues at the same pace, with even more advanced methods of creating fake content as close to reality as possible, such as "*deep fake*". Reasons for discussion and propaganda arose in the current context of the pandemic generated by the Covid19 virus, with all kinds of articles whose statements are difficult to be verified at a first sight.

Although there are more subcategories related to fake news such as satire, parody or clickbait, a general definition of the term could be the following: fake news represent a way to spread false information in order to mislead the public, damage the reputation of an entity or have a political or financial gain. The idea of misleading and influencing the public is also linked to the notion of hyper-partisan news that have the role of presenting extremist or conspiratorial opinions with intentional misconceptions. Given the fact that online fake news has been growing sharply lately, especially with the US presidential election campaign in 2016, attempts are being made to find automated solutions to identify such news articles circulating online. Approaches aim in most cases to train machine learning models that process language indicators of texts.

An important aspect to mention is that the spread of fake news in recent years has been facilitated by the current environment. People spend a lot of their time online, always connected to technology and Internet. In this context of a ubiquitous Internet, social networks have become part of people's lives. It is this connection of people to social networks that has become an excellent medium for distributing false news. Social networks facilitate the formation of common interest groups, in which it is much easier for a certain idea to take shape in an online community. Subconsciously, when a person is repeatedly exposed to the same content, even if that content is made up of false information, he will tend to give in and assimilate the information he frequently encounters. Thus, the very advantages of social networks can make them a double-edged sword, which can be used in a negative way to spread fake news [2]. Even if people do not want to be part of the fake news expansion on social networks, they could be part of it without their knowledge because not so many are able to distinguish between real and fake news.

Having as a starting point the context previously described, the problem to be solved is to create a system that aims to detect fake news circulating online. The general approach is based on training machine learning models to identify such items. In the current study, two datasets were analyzed in the experiments. As a starting point, we wanted to divide a dataset consisting of various news items into topics, using clustering as a working method. The idea behind this division is to delimit the news into similar topics in order to be able to train more specific detection models, per topics. Thus, for each cluster formed after clustering, a supervised classification model was trained. Also, for the entire dataset, such a model was trained in order to be able to later make a comparison between the global model and the model per topic for each cluster. The study analyzed the results obtained using attributes generated using Term Frequency – Inverse Document Frequency (TF-IDF) technique. In the last part of the study, an indexing of verified facts was used to filter the most important sentences in an article. The idea behind the last mentioned experiments was to try to extract the relevant facts from a news story, in order to train models that are not affected by irrelevant information.

Considering that a general solution for detecting fake news in the online environment is very difficult to achieve at the moment, the main goal of the current paper is to analyze several methods that can be used in the process of identifying fake news and proposing some techniques that could help in the overall end goal of studies in this area.

The next section of the paper describes the current state of the art for the fake news detection systems. The following sections present the proposed system for this study and the obtained results. The last section of the paper summarizes the conclusions of the survey.

RELATED WORK

An extremely important step in solving any problem is that its initial data is well known. In this case, the hypothesis can be considered as consisting of attributes that can be extracted from fake news datasets. Cardoso et al. [3] describe the main attributes considered in the analysis of fake news in the online environment. According to the mentioned study, the main categories of attributes that can be extracted for this problem are the linguistic ones and those based on the social content. Linguistic features refer to the attributes resulting from the processing of natural language. Many of the scientific papers in this field have extracted elements such as: N-grams, word embeddings, parts of speech, scores depending on the polarity of the feelings transmitted. Social network features refer to the number of likes of a post that contains a news story, the number of redistributions for a news story, the number of friends viewing or sharing the news, the average number of posts for a user in a given time frame, and so on. Within a network-based approach, both the aspects related to the profiles of users who interact with the news and the way in which the information is disseminated in the network are analyzed, considering the links that appear.

Within the field of fake news detection, different views on the problem can be found, depending on the solutions and systems proposed by researchers. Such a system that proposes the identification of fake news is found in the work of Shu et al. [4], where the social news ecosystem is seen as made up of three entities: publishers, the news itself and social media users. Therefore, a system is modeled by the three elements mentioned and the relationships that are formed between them.

The architecture of the solution proposed in the above-mentioned article consists of five components. The first of these refers to the attributes that can be extracted from the linguistic content of the news. To use these attributes, the authors used Nonnegative Matrix Factorization (NMF) algorithms. The second component of the architecture targets attributes that can be extracted from users. An adjacency matrix A is formed between users, in which it is marked between which users there is a friendship relationship at the level of the social network. As with attributes extracted from news content, NMF is used to learn user-level attributes. The next component addresses the relationships given by the links between user attributes and the news tags they distribute. In this case, a user credibility score is calculated. The central idea in calculating credibility is that less credible users tend to coordinate better with others, forming larger clusters, while credible users tend to form smaller clusters. The fourth part of the system is the relationship between publishers and news. Thus, the relationship between the vector and the partisan inclinations of publishers and the publisher-news matrix (which represents an adjacency matrix meaning news written by publishers) is used to add new attributes to the proposed system. The vector with the partisan inclinations of the authors was obtained using the MBFC website. The last component of the proposed system is the linear classifier, used to learn the attributes modeled by all the other components described above.

Another approach to the problem of detecting fake news can be found in the article Monti et al. [5]. In this paper the emphasis is on learning the patterns under which fake news spreads within a social network, using geometric deep learning. A large part of deep learning comes down to Euclidean data, for example data from 1-dimensional or 2-dimensional space. Non-euclidean data can represent more complex objects and concepts with greater accuracy than 1D or 2D representation. Geometrical deep learning is a subdomain of deep learning that aims to create neural networks that can learn from non-euclidean data. An example of non-euclidean data is the graph [6]. The authors of the article consider that modeling the problem in the form of a graph and exploiting it using geometric deep learning facilitates the construction of a system that incorporates heterogeneous data such as: content, activity and profile of users, social graph, news dissemination.

The attributes extracted to create the fake news detection model [5] can be divided into four categories depending on what they describe: user profile (location, profile settings, whether or not the account has been verified, etc.), user activity, network (connections between users, number of followers, friends, etc.), content (word embeddings for the linguistic content of the news). The system architecture consists of a Convolutional Neural Network (CNN) graph containing four layers: two convolutional layers, both of size 64 at the output and two fully connected layers, the first producing 32 attributes at the output, and the second 2 attributes. After the two convolutional layers, mean-pooling was used to reduce the size, and after the last completely connected layer in the architecture, a SoftMax layer was used. Finally, the model predicts the probability that the news will be false or true. In the study, two settings were used to train the models: a setting in which one tries to predict the type of a news starting from all the routes in which it appears in the network and a setting in which the detection is made starting from only one cascade (the term "cascade" was defined to describe the sequence of redistributions of a URL, starting from its source). The results are around 90% accuracy for both settings, with a small plus for the first option. It was observed that the length of the cascades influences the operation of the models, especially for the second setting, which is somewhat intuitive.

Wang et al. [7] also emphasized that other types of attributes are needed in addition to content-based ones to create a model that detects fake news in a real environment. The argument for this statement is that the events that are described in the news circulating online change over time, and models trained on certain datasets using linguistic attributes may not have the same performance correlated with other types of articles news describing other events. To get rid of this impediment, the authors of the article propose to train a model that learns the common attributes between events and to eliminate from the attributes used those related to specific events. Thus, to put this theory into practice, the first step is to identify those non-transferable attributes between events. This problem can be addressed by measuring the difference between the attribute vectors of posts that describe different events. The greater the difference, the more it results that the vectors contain attributes that uniquely describe events, so they represent non-transferable attributes.

The proposed system is called Event Adversarial Neural Networks (EANN) and, as its name suggests, is inspired by adversarial networks. EANN consists of three major components: a multi-modal feature extractor, the fake news identifier, and an event discriminator that aims to learn the invariant attributes between events.

PROPOSED SOLUTION

The main difficulty in creating a general fake news detection system is that the topics of the news are constantly changing, being very diverse and closely related to events happening around the world. Thus, it becomes very difficult to train a general model to detect fake news because solutions and models should be updated regularly in order to adapt to changes in news topics and to continuously learn new attributes relevant to the current context. Based on this idea, that news topics are constantly changing and that there are a large number of news categories, we tried to analyze the problem by dividing the dataset by topic and building more specific detection models to see how this distribution influences models training. The following subsections will describe the datasets used in the study, the proposed system, the attribute processing technique, as well as the indexing of statements and filtering of news according to these indexed facts. We considered that the diversity of data helps to make models work better in a practical situation and facilitates the generalization of their ability to detect fake news. The novelty of the current paper over other related articles like the ones described in related work section consists in the fact that the most important facts are filtered from news articles using indexed statements and, moreover, news are separated in categories in order to obtain more specific classification models.

Datasets

Buzzfeed [8] is a dataset that appears quite often in articles on this topic. It is generally used in conjunction with other datasets because it is not very generous in terms of the number of news items it contains. It consists of news published on Facebook over the course of a week in the US presidential election in the fall of 2016. The news was tagged by accredited journalists. Although the original division is into unbiased articles, left-wing articles and right-wing articles, there is also a classification of them into true and fake news, even if this division is not balanced.

The datasets used in this field are quite recent, but they focus mainly on the period until or during the US presidential election in 2016. So we found it useful to collect a dataset that also contains news closer to current topics, such as the latest US presidential election or coronavirus. *Politifact.com* is a site that deals with the labeling of statements that appear in the online environment depending on the degree of truth. This dataset does not contain the whole news, but only a statement about a certain topic. Often articles are built on a single statement, so focusing on identifying the veracity of key statements is a necessary point in the process of combating fake news.

True	True	2250
Mostly true/Half true	Partly true	6437
Mostly false	Partly false	2991
False/Pants on fire	False	6439
Total:		18027

Table 1. Politifact dataset statistics.

There are six possible labels for the statements, but in this study we merged labels "mostly true" and "half true", as well as labels "false" and "pants on fire", resulting in four categories totaling 18027 statements, as shown in Table 1.

System architecture

The system proposed for analysis aims in a first stage to divide the news into different topics based on clustering algorithms. To obtain these clustering models we used K-Means and Agglomerative Clustering. Thus, an attempt was made to separate the news according to their topics. Then, in each cluster obtained, a classification model was trained using Support Vector Machine (SVM), these being further referred to in the text as topical models or cluster-based models. Also, for a clustering algorithm and a certain number of clusters provided at the input, a global detection model was trained, which contains the entire dataset. It was necessary to calculate different models in order to make a correct comparison between the results obtained. The comparison with a single global model would not have been correct because parts of the test sets of the topical models could have been part of the training set of the global model.

Regarding the procedure for calculating the datasets for topical and global models, the steps were as follows: for each cluster obtained with a clustering algorithm and with a number of clusters given at the input we divided into fake news and real news; for fake news we kept 75% for training and 25% for testing; we did the same for the real news in that cluster; subsequently, we concatenated the 2 sets for training and the 2 sets for testing, obtaining per cluster a set of 75% training data and 25% test data, thus keeping the news ratio from the initial set (fake news vs. real news) in the new datasets obtained per cluster. These datasets were used both to train local models and the global model. Thus, the data do not overlap and the comparison is made in a fair way. The entire model calculation process can be seen in Figure 1.

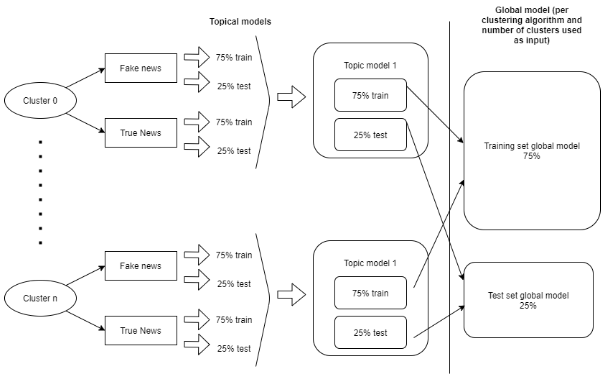

Figure 1. Procedure for obtaining the datasets for topical models and global models.

Text processing

Attributes for clustering and for training SVM models were calculated using TF-IDF. To obtain the number vectors that represent the attributes, a text processing was first performed. For the content of a news story, abbreviations were removed, the letters were all transformed into lowercase letters, punctuation marks, numbers, stop words were removed and a lemmatization was done. After this processing, TF-IDF was applied and 2500 attribute number vectors were obtained for each word. TF-IDF is a mechanism through which the most relevant words can be selected for each entry from a dataset, their importance being related to their appearance in the entire corpus of documents. The same attribute vectors were used both for clustering datasets by topic and for training per-topic and global models.

Indexing Politifact dataset

Another procedure required in the experiments for this study was the indexing of the Politifact dataset, which was done using Lucene[19], a Java library that facilitates indexing and retrieval for texts and provides a multitude of classes and functions to help implement it. Each entry in the Politifact dataset was saved in a separate document so that the indexing of the statements in the Politifact involved the indexing of these documents. ClassicAnalyzer was used as a text analyzer, which divides text into words according to punctuation and eliminates punctuation, normalizes words, turns everything into lowercase letters, and uses a list of English stop words to remove those words that are too common to be useful in searches.

Buzzfeed news filtering

The purpose of indexing the statements in the Politifact is to use these facts to filter the most relevant facts from an article that contains a whole story. The dataset proposed for analysis to establish the efficiency of using the relevant facts from the articles in the training of detection models is Buzzfeed. Although it is not a large dataset and is not balanced at all, containing only 1604 news items, of which 1313 are true and only 291 false articles, it is suitable to see if a filter is beneficial because it contains news facts that can be found in Politifact. If the filtering system works well on these smaller datasets, then it can be generalized to larger datasets, where indexing can be done using many more fact-checking sites, thus broadening the fact news database. The larger the information base, the more efficiently filtered news will be.

In order to filter a news story, in the first step it was delimited in sentences. Then, for each sentence longer than 3 words, queries were formed for groups of 3 words that

[19] https://lucene.apache.org/

can be at a maximum editing distance of 12 positions. If the sentence is shorter than 3 words, the entire sentence was searched. Those sentences that had at least one result found in the indexed documents of the Politifact have been preserved. Table 2 contains some statistics on applied filtering. As it can be seen, there are also some news that have been completely reduced, which means that in these cases there was no search that was successfully executed in the indexed documents. However, this percentage is quite low, about 4% of the total news in the dataset. As statistics that can be found in the table there are: the length of the texts before and after the filtering, the number of words in the initial text and in the filtered text, as well as the number of initial and final sentences. On average, the text after filtering is about 30% the size of the original text, which means that keeping the relevant facts shortens the articles by a fairly high percentage.

Total news (real news/fake news):			1604 (1313/291)
Completely reduced:	68/1604	Percent:	4.24%
Completely reduced (real):	54/1313	Percent:	4.11%
Completely reduced (fake):	14/291	Percent:	4.81%
Filtered averages:			
Length before filtering:	3399.76	Length after filtering:	1171.04
Percent length after filtering compared to the initial length:			33.44%
Number of words before filtering:	558.46	No. of words after filtering:	190.98
Percent number of words after filtering compared to the initial no. of words:			34.20
Number of sentences before filtering:	24.45	No. of sentences after filtering:	6.99
Percent number of sentences after filtering compared to the initial no. of senteces:			28.61

Table 2. Statistics regarding news filtering for Buzzfeed dataset.

Following these selections of sentences based on the facts found in Politifact, we performed a manual analysis to see if the facts found for a particular news item are really relevant to summarize it. We discovered that there are articles where the most relevant facts are not all kept, because they were not treated in the Politifact and therefore do not remain after filtering. For example, one of the articles talks about a lawsuit related to water pollution in Flint, a city in America. However, in Politifact you can find information about the fact that in Flint the water was polluted, but nothing is found related to any trial. Thus, the abstract does not even contain the main ideas, in the sense that certain sentences that could be more relevant to what the article wants to convey are removed. From what we have noticed, there are some articles for which there are no related facts in Politifact, and for these articles are kept sentences that contain more common word groups such as: *New York City*, *official told*, *united arab emirate*, *school new jersey*, *government way* etc. However, the ideas kept are quite appropriate to summarize the news. Although there are also articles where the most relevant facts are not found or where the filtering does not even find adjacent facts in Politifact, we have discovered many articles in which keeping the essential sentences seems to work quite well.

Based on the above observations, another decision stage was added in the news filtering. Where the remaining news text has either the length of the characters, the number of words or the number of sentences less than 10% of the original news, a different filtering is applied. This new approach involves keeping 40% of the initial news using the first (20%) and last (20%) phrases. We found this change useful because a low percentage may mean that those facts were not found in Politifact, and the first and last sentences may be the summary for an article because it generally represents the introduction and conclusions, in which the main ideas are presented. Table 3 contains the statistics on the news that fall into this type of filtering based on the first and last sentences. As it can be seen, about 13% of the total news needed a recalibration of the filter according to the first and last sentences of the article.

Total number of articles:	1604 (1313 real/291 fake)
Number of articles with percent > 10%	1390 (~87%)
Number of articles with percent <= 10%	214 (170 real/44 fake) (~13%)

Table 3. Number of news by filtering type.

The whole decision-making process to keep the facts considered relevant in an article can be seen in Figure 2.

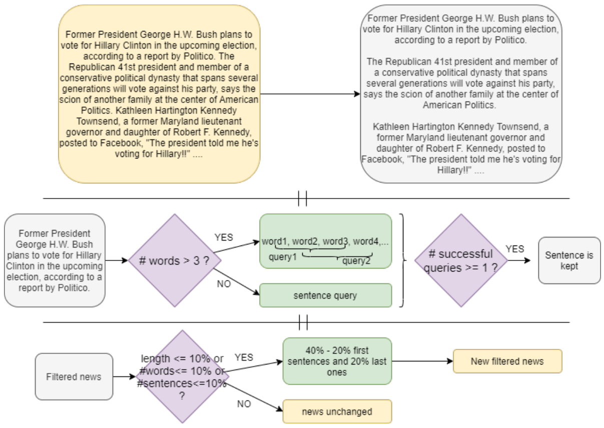

Figure 2. News processing by relevant facts.

RESULTS

Politifact two labels

In order for the Politifact dataset to fit the system architecture described above, for the first set of experiments we restricted the classes in the dataset to only two labels: real (8687) and fake (9340). For the clustering algorithms we chose 5, 10 and 20 clusters for each. On average, the results for the global models trained for each type of clustering are found in Table 4.

Accuracy:	0.64057
Recall:	0.64826
Precision	0.65471
F1 score:	0.65147

Table 4. Average results for global models for Politifact dataset using two classes.

Below are the results for Agglomerative Clustering with 10 clusters (Figure 3). The graphs contain the obtained accuracy. The columns in turn represent: the number of news (statements) in a cluster, the number of training data, the number of testing data, the number of real news in the test set, the number of fake news in the test set, the number of news items correctly identified by the topic model (cluster-based model), the number of news items correctly identified by the corresponding global model (baseline model), the number of news items correctly identified only by the topic model, the number of news items correctly identified only by the global model, the number of news items not correctly identified by any model and the number of news items correctly identified by a set of votes between the two models. For the voting system, a news item is labeled as fake if either the topical or the global model identified it as fake, and real if both types of models consider it to be true.

	articles	train	test	test_r	test_f	cor_topic	corr_big	topic_only	big_only	neither	ens
0	315	236	79	36	43	49	50	5	6	24	48
1	15484	11613	3871	1819	2052	2452	2458	109	115	1304	2445
2	546	409	137	78	59	81	82	13	14	42	82
3	635	476	159	100	59	96	98	13	15	48	98
4	366	274	92	51	41	66	56	11	1	25	59
5	90	67	23	4	19	19	20	0	1	3	19
6	209	156	53	38	15	38	35	7	4	11	35
7	104	77	27	11	16	15	10	6	1	11	15
8	99	74	25	16	9	16	14	6	4	5	14
9	179	133	46	22	24	33	33	4	4	9	33

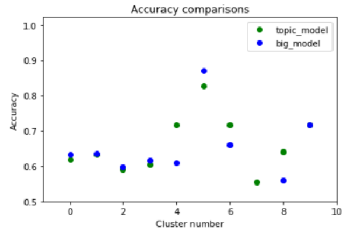

Figure 3. Results for Agglomerative Clustering (10 clusters).

Politifact multiclass classification

Considering that the statements on Politifact can be classified into six different categories in terms of trust, we considered useful a set of tests in which to use more possible labels for the news. Thus, we have kept the division in Table 1, where a statement can be: true, partly true, partly false and false. This time the classification models need to be trained for multiple classes. In the figure below, 0 is a true statement, 1 is partially true, 2 is partially false, and 3 is false. The average results for the global models corresponding to each clustering algorithm can be seen in Figure 4.

```
Accuracy:  0.4838280903854674
Recall:  0.4838280903854674
Precision:  0.4838280903854674
F1_score:  0.4838280903854674
              precision    recall  f1-score   support

           0      0.375     0.005     0.010       565
           1      0.450     0.709     0.550      1610
           2      0.333     0.004     0.008       749
           3      0.529     0.652     0.584      1590

   micro avg      0.484     0.484     0.484      4514
   macro avg      0.422     0.343     0.288      4514
weighted avg      0.449     0.484     0.405      4514
```

Figure 4. Average results for global models for multiclass.

We again extracted the results for Agglomerative Clustering (Figure 5) for 10 clusters, and the results can be tracked below. The columns represent in turn: number of news (statements) in the cluster (art), the number of news in the training set (tr), the number of true news in the test set (te_r), the number of partially true news in the test set (te_ar), the number of partially fake news in the test set (te_af), the number partially fake news in the test set (te_f), number of fake news in the test set (te_f), number of news correctly identified by the topic model (cor_topic), number of news correctly identified by the global model (corr_big), the number of news correctly identified only by the topic model (to_only), the number of news items correctly identified only by the global model (big_only) and the number of news that were not correctly identified by either the topic model or the global one (neith).

	art	tr	te_r	te_ar	te_af	te_f	cor_topic	corr_big	to_only	big_only	neith
0	13499	10123	407	1170	569	1230	1643	1645	81	83	1650
1	3013	2259	125	326	110	193	368	366	31	29	357
2	207	154	4	19	11	19	27	29	2	4	22
3	499	373	13	48	28	37	51	58	5	12	63
4	150	112	2	10	4	22	22	28	1	7	9
5	98	71	1	2	3	21	27	21	6	0	0
6	245	182	8	17	10	28	28	29	8	9	26
7	102	75	3	8	7	9	12	10	4	2	13
8	87	63	1	1	2	20	24	20	4	0	0
9	127	93	3	11	6	14	15	19	0	4	15

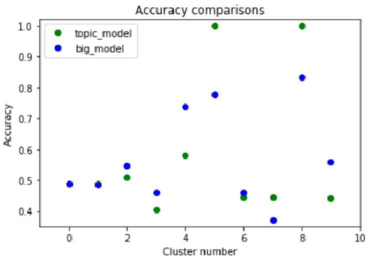

Figure 5. Results for Agglomerative Clustering (10 clusters) for multiclass.

For some clusters we could not train local models because there was too little data in a certain category. For example, in Figure 5, for cluster 8 it was just one real news and we could not do the division for training and testing. In contrast, in such clusters where there is very little news in a category, even the initial clustering can be used to determine the dominant type of news in that cluster. For example, also in cluster 8 in Figure 25 there are: 1 true news, 2 partially true news, 7 partially fake news and 77 fake news, so we can say that a news distributed in that cluster has a very high chance of being fake. That is why we set the accuracy to 1 for such clusters where we could not train the topical model, but a certain type of news is quite clear.

Filtering news from Buzzfeed

The last category of experiments aimed at using Buzzfeed dataset, which contains entire news articles, in correlation with Politifact dataset. The possibility of selecting relevant facts from a news item using a knowledge base was analyzed. In this case, the indexed information came from Politifact dataset, and the Buzzfeed news articles were filtered based on the indexed documents. In order to establish the efficiency of the selection of the relevant facts within the articles, the results for running the system with the entire news dataset and for running the system with the filtered news dataset were compared. Regarding the number of clusters that had been chosen in order to train the clustering algorithms, models with 5, 7 and 10 clusters were trained.

Clustering algorithm	Initial dataset	Filtered dataset
K-Means – 5 clusters		
Mean accuracy topical models:	0.83	0.80
Mean accuracy global model	0.83	0.80
K-Means – 7 clusters		
Mean accuracy topical models:	0.81	0.82
Mean accuracy global model	0.82	0.82
K-Means – 10 clusters		
Mean accuracy topical models:	0.83	0.85
Mean accuracy global model	0.84	0.85
Agglomerative clustering – 5 clusters		
Mean accuracy topical models:	0.82	0.84
Mean accuracy global model	0.83	0.81
Agglomerative clustering – 7 clusters		
Mean accuracy topical models:	0.85	0.86
Mean accuracy global model	0.85	0.84
Agglomerative clustering – 10 clusters		
Mean accuracy topical models:	0.85	0.86
Mean accuracy global model	0.86	0.84

Table 5. Results for models trained with initial dataset and filtered dataset.

In this analysis, the emphasis is not on identifying cases where topical models have better results than the global model, as in previous studies, but on comparing the average results for the initial and filtered dataset. Table 5 shows the average results of the classification models corresponding to all involved clustering models. As it can be seen, the K-Means case with 5 clusters is the only one in which the average accuracy of the models per topic decreases, the rest having an increase of 1-2 percent. The fact that the global model trained by the filtered news dataset has slightly lower results is not an impediment because the emphasis must be placed especially on topical models. Another interesting observation is that as the number of clusters increases, so does the accuracy. This is a rule that tends to be kept for both the original and the filtered dataset. We also noticed that the scores for the results of models trained with Agglomerative Clustering are a few percent higher than those of models trained with K-Means for the same number of clusters.

Table 6 contains an example of news that is correctly predicted by filtering using Politifact and incorrect when using the full text. The news is part of cluster 0 if it is filtered and of cluster 1 when it comes to trained models with entire text news. The extracted article was for Agglomerative Clustering with 7 clusters as input. The label for the selected news is fake, being correctly predicted by the topical model trained with news where the facts are extracted and incorrectly by the topical model corresponding to the training with entire text news.

Entire news article:
Washington, D.C. (CNN) The Pentagon is considering a proposal that could send an additional 500 troops to Iraq to help Iraqi and Kurdish forces take Mosul from ISIS in the coming weeks, a US defense official told CNN. The possibility of sending some number of additional forces has been reported for several weeks, including by CNN. The Wall Street Journal was the first to report the proposal centered around the possibility of up to 500 additional forces. No final decision has been made, the official said. If the proposal is approved, it is anticipated that would be the final wave of additional US forces for fighting ISIS, since Mosul remains the last major ISIS stronghold in Iraq, the official added. If the proposal is okayed, it will raise the ceiling on the number of US troops in Iraq from 4,647 to over 5,000. Several hundred additional troops move in and out of Iraq on temporary deployment orders. Currently, there are 4,470 there under the approved ceiling. The Pentagon is not precisely revealing what the troops might be called upon to do. Some number are expected to be military advisors and trainers for Iraqi and Kurdish forces, whiles others could be involved in logistics such as moving supplies and troops around the battlefield, the official said. Arizona Republican John McCain, chairman of the Armed Services Committee, criticized the move Thursday. "Indeed, we read this morning of plans for yet another incremental increase of 500 troops in Iraq, one more step down the road of gradual escalation," he said.
Filtered news article:
Washington, D.C. (CNN) The Pentagon is considering a proposal that could send an additional 500 troops to Iraq to help Iraqi and Kurdish forces take Mosul from ISIS in the coming weeks, a US defense official told CNN. The Wall Street Journal was the first to report the proposal centered around the possibility of up to 500 additional forces. If the proposal is okayed, it will raise the ceiling on the number of US troops in Iraq from 4,647 to over 5,000. Arizona Republican John McCain, chairman of the Armed Services Committee, criticized the move Thursday.

Table 6. Example of fake news correctly predicted using filtration with Politifact and incorrectly otherwise.

CONCLUSIONS

The current study aimed at an analysis of the phenomenon of fake news and how it can be combated using automatic detection systems. Starting from the idea that the news is constantly changing and that the topics differ depending on what is happening around us, the study proposes a system in which data are first divided by topics using clustering algorithms. Furthermore, local detection models were trained for these clusters to establish the ability for a more specific system to provide better identifications than a global model. The mixed detection system assumes that in a first stage the predominant class in the cluster can be consulted. If a cluster contains only fake news or only real news, then most likely a tag can already be assigned to a new news that reaches that cluster. Given that many of the news stories are built on a few important facts that can summarize a news article, the study involved an analysis of Politifact dataset, collected with newer data and containing statements categorized by degree of truth. In this step, classification models for multiple labels were also involved, so as to cover several categories in which the statements on the Politifact can be divided.

As a last step, an analysis was carried out on the effectiveness of identifying important facts in a news story in order to train models that are not affected by irrelevant information. In this sense, the statements on *Politifact.com* were first indexed, which span several years and include various topics, from the 2016 US presidential election to the current situation with news about coronavirus. Then, the news from another dataset, Buzzfeed, went through a filtering process that involved keeping those relevant phrases to summarize the news. The extraction of the relevant facts was done depending on whether or not they were found in the indexed information base using Politifact. Subsequently, a comparison was made between the results of the trained models with the initial Buzzfeed dataset and the same dataset passed through the mentioned filter.

Regarding future approaches, one of the possible directions would be to index more fact-checking trusted sites as for facts database. Given that the dataset for which the filtering system (Buzzfeed) was tested is quite small, also having the disadvantage that it contains few fake news compared to the real ones, the system should be tested with a larger dataset. At the same time, a possible future direction may be to train the models using embeddings generated by a transformer.

REFERENCES

1. Commonsense.org. 2017. [online] Available at: <https://d1e2bohyu2u2w9.cloudfront.net/education/sites/default/files/tlr-asset/newsmedialit_fakenewstimeline_8.5x11.pdf> [Accessed 3 June 2021].
2. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. ACM SIGKDD explorations newsletter, 19(1), 22-36.
3. Cardoso Durier da Silva, F., Vieira, R., & Garcia, A. C. (2019, January). Can machines learn to detect fake news? a survey focused on social media. In Proceedings of the 52nd Hawaii International Conference on System Sciences.
4. Shu, K., Wang, S., & Liu, H. (2019, January). Beyond news contents: The role of social context for fake news detection. In Proceedings of the twelfth ACM international conference on web search and data mining (pp. 312-320).
5. Monti, F., Frasca, F., Eynard, D., Mannion, D., & Bronstein, M. M. (2019). Fake news detection on social media using geometric deep learning. arXiv preprint arXiv:1902.06673.
6. Medium. 2021. What is Geometric Deep Learning?. [online] Available at: <https://flawnsontong.medium.com/what-is-geometric-deep-learning-b2adb662d91d> [Accessed 3 June 2021].
7. Wang, Y., Ma, F., Jin, Z., Yuan, Y., Xun, G., Jha, K., ... & Gao, J. (2018, July). Eann: Event adversarial neural networks for multi-modal fake news detection. In Proceedings of the 24th acm sigkdd international conference on knowledge discovery & data mining (pp. 849-857).
8. Shu, K., Mahudeswaran, D., Wang, S., Lee, D., & Liu, H. (2018). Fakenewsnet: A data repository with news content, social context and spatialtemporal information for studying fake news on social media. arXiv preprint arXiv:1809.01286