Fake news detection by Machine Learning in Latin America: A Systematic Review

Jean Gabriel Nguema Ngomo
Department of Applied Informatics
UNIRIO
mvojgnn@edu.unirio.br

Raquel Torres de Paiva
Department of Applied Informatics
UNIRIO
raquel.paiva@edu.unirio.br

Ana Cristina Bicharra Garcia
Department of Applied Informatics
UNIRIO
cristina.bicharra@uniriotec.br

Abstract

The growing spread of fake news on social network is a major scourge in society. To combat this problem, many studies have focused on different aspects to automatically detect fake news on social networks using artificial intelligence, especially Machine Learning. In the present work, to understand the current state of existing proposals, we conducted a systematic literature review. We propose to organize this literature in the light of a taxonomy of approaches.

Keywords: Fake news; Misinformation; Disinformation; Machine Learning; Latin America; Social Media; Social Network; Cultural aspects

1. Introduction

In recent years, social network have consolidated themselves as the main source of information for the general population, replacing traditional reliable media such as televisions, radios, newspapers, and magazines. While they bring many benefits, such as speed and agility of communication with interconnected people, social network can also lead to serious problems such as the dissemination of fake news, rumors, hate speech, among other problems.

Fake news can have dire consequences for society, leading individuals to make the wrong decisions. In this sense, there are several examples of these consequences. We can mention the invasion of the headquarters of the three powers of Brazil (executive, legislative, and judicial), by vandals, in January 2023, a few days after the beginning of the mandate of the elected president Araújo [6]; the episode that bears resemblance to the invasion of the United States Capitol in January 2020[9]. Fake News covers any domain, distorting or inventing information about the pandemic, vaccines, personalities, for instance. This scourge threatens a fundamental element of our society: the truth. Furthermore, all regions of the world face this kind of problem, and Latin America does not escape this sad reality, as the example above illustrates.

Several researchers have been investigating solutions to combat the spread of fake news on social network, identifying them using Artificial Intelligence, specifically Machine Learning. This type of solution would allow users of social network to discover that certain news is false before it is shared with their contacts, breaking the chain at the root.

During the present work, we found systematic reviews of such researches [27], [35], [3], [34], [4], [26], [11], [39] and [7]. However, there are few of these reviews from the perspective of Latin America, where cultural aspects may present new challenges for Machine Learning. Then in the present study, we conduct a systematic review of approaches directed to fake news detection in social network using Machine Learning with emphasis on this region, considering the last 5 years. Our main motivation is to answer the following research questions:

- Which approaches and Machine Learning techniques are used to detect and combat fake news in Latin America?
- How does Machine Learning help fighting fake news through different types of social network, specifically in Latin America?
- Which conceptual characteristics and limiting factors do Machine Learning applications seeking to identify fake news in social network have, specifically in Latin America?

The rest of this paper is organized as follows. In section

URI: https://hdl.handle.net/10125/106688
978-0-9981331-7-1

2, we describe and characterize some concepts, mainly fake news terminology and Machine Learning concepts. We present our methodology in section 3, followed by our study results and discussion in section 4. In section 5, we present some limitations and challenges about cultural aspects in Latin America with relation to Fake news detection on social networks. We then conclude in section 6.

2. Background

2.1. Fake news

Fake news is one of the most popular terms today. There is worldwide concern over false news and the possibility that it can influence political, economic, health system and social well-being[1]. In the literature and real life, the term fake news is used in different situations and contexts, which causes some confusion about its true meaning. Several terms such as misinformation and malinformation are considered as fake news. Despite the similarity in terminology, they are actually different categories of what is called **information disorder** by S. and D. [35]. These categories are: misinformation, disinformation, and malinformation. The differences between these categories lie in the degree of reliability of the information and its source and in the intended use.

By definition, **malinformation** is the deliberate publication of private information for private rather than public interest, such as revenge porn. Malinformation is often created from change of context, date, or time of the genuine content.

On the other hand, **misinformation** is spread without the intent to deceive someone. According toKumar and Shah [22], its common causes include misrepresentation or distortion of an original piece of true information by an actor, due to lack of understanding, attention or even cognitive biases. People are exposed to misinformation through social network posts like tweets or Facebook posts and comments, website articles, and so on. Then an actor can share them involuntarily believing in their veracity.

While misinformation refers to false or misleading information shared without harmful intent **disinformation**, in contrast, refers to false information deliberately created and shared to deceive or mislead others. Although they have different meanings concerning the intent, the consequence is similar: fostering untruthful information and causing public confusion or harm. Consequently, as far as fake news is concerned, we consider them the same in our study. So, the focus of this work resides on misinformation and disinformation, which encompasses fabricated and false stories, rumors, hoaxes, and satire. Rumors are stories whose veracity cannot be confirmed; hoaxes, on the other hand, are stories with true and false parts; and satire are messages with a mixture of irony and humor.

2.2. Machine learning

Machine learning (ML) is a field of Artificial Intelligence (AI) whose goal is to make the machine solve complex problems intelligently, imitating human behavior. Just as human beings, based on their accumulated experience, adjust their conduct to complex situations, ML algorithms also called techniques are trained to learn and accumulate experience from some information, so that they can make predictions and solve complex problems when submitted to new information. In this sense, once they were properly trained, these algorithms would be able to attest that a given information content is fake news.

There are many type of ML, two of which: supervised learning and unsupervised learning. **Supervisioned Learning** is a subset of ML where the algorithm takes a set of data labeled as training data and makes predictions for unknown data. This is the most common scenario associated with classification(CLA) or regression problems. A regression is a prediction problem in which the output variable has a continuous numeric value, as opposed to classification where the expected output is one item among a finite set of classes. In literature, researchers on fake news use classification to decide if a given content belongs to two classes relative to their veracity: true or false.

On the other hand, **Unsupervised Learning** is the kind of ML where the algorithm input data is not labeled. While labeled data allow the algorithm to understand the meaning of data in a supervised learning, unlabelled data lacks this meaning in an unsupervised learning. The main objective of an unsupervised learning is to group unlabelled data or interpret patterns, associations, and relationships between them. We present a brief description of some ML algorithms in table 1.

Another important sub-area of ML is **Deep Learning (DL)**, which is considered an important branch of unsupervised learning. More specifically, Deep Learning (DL) seeks to carry out its learning simulating the behavior of the human brain, through the so-called **Artificial Neural Networks (ANN)**. In this case, an ANN models a neural network through a set of layers of neurons that are connected[23]. These layers are hidden, and every hidden layer

acts as the input to the next layer until the layers connected to input or ouput of the ANN. ANNs can be used both for complex regression or classification problems[23]. Neural networks are the heart of deep learning algorithms[17]. A **Convolutional Neural Network (CNN)** is a kind of Neural Network, which have good performance for classification and computer vision tasks, enabling computers and systems to derive meaningful information from digital images, videos, and other visual inputs, and based on those inputs. Another Neural Network technique is **Recurrent Neural Network**, commonly used for natural language processing and speech recognition.

Another important key concept in ML is **features**. Given any ML algorithm, they designate data used as its input to train it and making it able to learn on a complex phenomenon. After training, when fed with new data with the same characteristics of these features, the algorithm is able to make predictions about this studied phenomenon to solve a complex problem. Before using any algorithm to solve a specific problem, an important preliminary step is to use the domain knowledge to create features that will make the ML algorithms work correctly. The better will be the features generation, the more efficient will be the algorithm performance. Features come from datasets.

3. Research Methodology

To answer our research questions, we carried out our study following the method proposed by [8] for Systematic Literature Review. Our aim was to analyze researches that studied fake news detection in social network using ML from a Latin American perspective. To this end, we divided our methodology into three phases: planning, conducting, and reporting. The first two phases were performed using the Parsifal tool[1], and the last one is the main purpose of this paper.

3.1. Planning

We defined our keywords and their possible synonyms using the PICOC protocol (Population, Intervention, Comparison, Outcomes, and Context), in the Parsifal tool, as described below.

- Population as the social network, Twitter or Facebook,
- Intervention as machine learning detection or identification,
- Comparison was not applicable,
- Outcome as fake news or misinformation, and

[1] https://parsif.al/

Table 1. Algorithms and their descriptions

TECHNIQUE	DESCRIPTION
Random Forest (RF)	Supervised learning algorithm whose principle is to build a ”forest” by merging a set of decision trees, normally trained with the bagging method.
XGBoost (XGB)	Supervised learning algorithm technique that can be adopted in both classification and regression analysis schemes.
Support Vector Machine - SVM	Supervised learning algorithm used both for classification and regression analysis, essentially ranking data into categories
Logistic Regression (LR)	Supervised learning algorithm, which performs regression analysis estimating the probability of an event occurring based on the previous data.
Naive Bayes (NB)	Supervised learning algorithm, which perform classification applying statistically the Bayes theorem.
Clustering (CLU)	Unsupervised learning method which involves grouping sets of similar data.
k-Nearest Neighbors (KNN)	Supervised learning algorithm that classifies data points based on their proximity.
K-Means (KME)	Unsupervised learning algorithm used to cluster data.
Principal Component Analysis (PCA)	Unsupervised learning algorithm are used to reduce the dimension of data by its linear transformation into a new set of coordinates.
Convolutional neural network (CNN)	Neural network model where neural connections flow in a single direction. Suitable for classification and computer vision tasks.
Recurrent neural network (RNN)	Model of neural networks where the output of processed neurons are saved and fed back into the model with the result.
Long Short Term Memory (LSTM)	A type of RNN able to learn long-term dependencies.

- Context as Latin America, Argentina, Brazil, Chile or Mexico

The final query was composed of: (”fake news” OR ”misinformation”) AND (”detection” OR ”identification”) AND (”machine learning”) AND (”social network” OR ”social media”) AND (”Brazil” OR ”Latin America” OR ”Argentina” OR ”Chile” OR ”Mexico”). We chose three databases commonly used in the computer science area: IEEE Xplore, ACM DL and Springer Link.

3.2. Conducting The Review

To proceed with the process of selection and acceptance of articles, the following inclusion and exclusion criteria were adopted. For inclusion, the criteria adopted were articles dealing with fake news or synonyms and detection techniques; For exclusion, the criteria of publication in low-impact forums, workshops, book chapters or annals, articles that were not in English, publications older than 5 years, articles whose scope was not related to Computer Science were adopted.

The initial selection research started by filtering by content type Article, discipline Computer Science and date published directly on the database, when available on that. After that, the selection research was filtering keywords related to Population in the abstract, to assure articles were specific for fake news detection. Then, we included publications, written in English, from 2018 to 2023, to select the most recent results in the literature. We also excluded duplicates, papers published in low-impact factor media, citations, patents, proceedings, workshops, and book chapters.

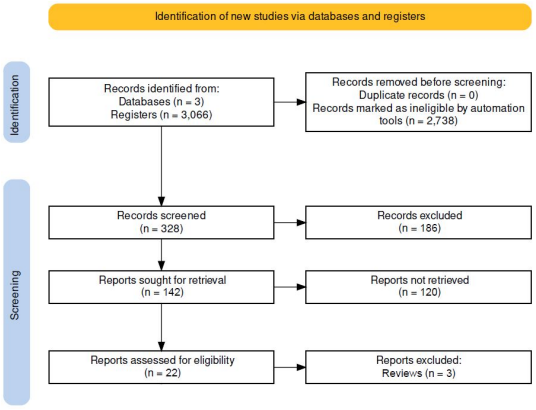

Figure 1. Research flow.

After applying the inclusion and exclusion criteria, we obtained a set of 142 papers. Then we read abstract and introduction of each paper and excluded the ones that did not fit our objectives. We then selected 22 papers for full reading in pair.

To better illustrate the process described in this phase, we produced a PRISMA flow diagram, using a tool available on the web [4] as shown in figure 1.

This process started on the 20th of April 2023. From a total of 22 papers, 03 are fake news detection reviews and 19 are studies that propose the usage of Artificial Intelligence techniques to detect fake news.

4. Results and Discussion

This section presents our analysis of the works selected to answer our research questions.

4.1. General findings

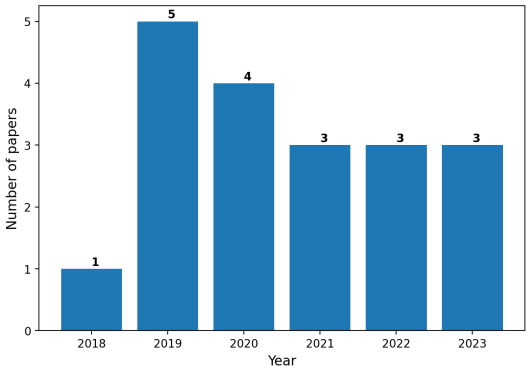

Figure 2. Evolution of number of papers per year.

The figure 2 shows the number of papers published per year in the literature, considering the period investigated, that is, the last 5 years. During this period, we observed that the annual number of published papers reached a peak of 5 papers in 2019. Since then, this number has dropped slightly. However, this number has remained constant above 3 articles in this period. This constancy indicates that this area of investigation of fake news detection mechanisms in social network in Latin America, using ML, still has many questions to be resolved. Regarding the year 2019, all its papers were related to detection of fake news about politics' domain, what can be explained as the following. Soon after the episode of the massive use of Fakes News in the 2016 U.S. presidential election[10], it was not thought that this phenomenon would be repeated in Latin America, two years later in the presidential elections in the two largest economies of the region, namely Brazil and Mexico. But that is exactly what happened in 2018[32], when the spreading of fake news had again a

significant impact on the results of these elections. This impact aroused the interest of researchers to investigate solutions addressing this problem, mainly for politics domain.

Knowledge Domains. From our analysis of the literature, we noticed that the works that investigate fake news in social network in Latin America using ML are concentrated around six major domains: politics, health, economics, culture, general, and sport. We elaborated a bubble chart (Figure 3) to show the importance of each domain. More specifically, in this chart the colors are used only to differentiate the domains, while the size of each bubble shows its significance compared to other bubbles. By far, politics and health are the two domains that have attracted the attention of researchers. There are studies on presidential elections in American countries, particularly in Brazil[21] [40], and the fight against violence, especially in Mexico[24] [38]. In the field of health, studies focus on combating fake news about sars Covid-19[42][29], which is explained by the recent global pandemic context, which caused a proliferation of fake news about false treatments, invented stories about the origin of Covid, and even false statistics about the pandemic. Being a region where people are very passionate about sports, especially soccer, some researchers also seek to identify fake news in this domain[20]. There are also researchers who investigate generic solutions, which would detect fake news in any area.

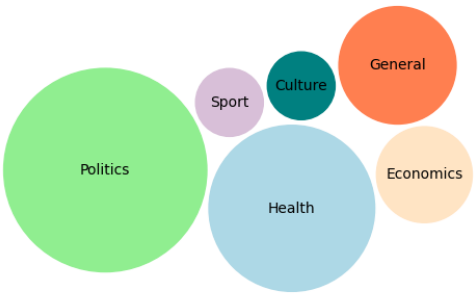

Figure 3. Domains of fake news.

Dataset Sources. In almost all the papers, researchers need to construct datasets and features for their studies, from social network or combining it with online portals. To this aim they use **Natural Language Processing (NPL)**, generating features that capture a better understanding of data of domain being studied.

On the other hand, exploring the literature about the primary source of data used to generate datasets and consequently features of the studies, we observed that these sources are diverse. In fact, as expected sources include social network platforms, from the most popular like Twitter, Facebook, and Google+, but also Sina Weibo. Figure 4 shows all primary sources for the selected papers in this study, as well as their proportion usage. Our study noticed that by far, Twitter is the most used platform most used in literature, followed by Facebook. Many social networks provide ways to extract its data for research proposed by means of API(Application Programming Interface). This facility allows researchers to extract their own datasets as they want. In this sense, researchers can extract tweets on a given domain of knowledge from Twitter like in[5], [25], and [21]. Similarly, posts and comments can be extracted from Facebook as in [33] and [31].

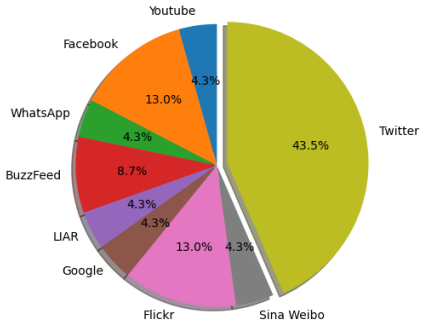

Figure 4. Dataset source and their proportion.

Our study also identifies sources also specific datasets like LLAR, a large dataset of claims from politifact.com[2] created by [2]; BuzzFeed[3][33]. Moreover, some researcher also use online portals about news as a source for datasets and features construction.

4.2. Content formats for features

Social networks have become the preferred environment for spreading fake news. The detection of Fake News by ML seeks to detect and combat this proliferation of fake news, predicting the veracity of the published content on Social Network.

In our literature analyses, in Latin America perspective, we found that the works use all the content format of social networks including textual content and multimedia content, such as images and videos. Despite the multimedia content allure users more than text [35], presenting a high likelihood of spreading fake new, the text format is the most used in research in this field. Some works also combine text content and multimedia content (multimodal content).

[2]https://www.politifact.com/
[3]https://buzzfeed.com.

Table 2. Authors and their techniques

REF	ALGORITHMS
[33]	BERT, KME
[24]	CNN
[14]	CNN, RNN
[18]	RF, XGB
[25]	RF, SVM, LR, CLU
[28]	RF, SVM, KNN
[36]	DL, LSTM
[30]	SVM, PCA
[37]	RF
[21]	CLA, CLU.
[31]	DL, NLP
[5]	RF,LR,XGB,KNN,RNN
[20]	CLU,CNN
[40]	RF, SVM
[19]	RF, SVM, LR,
[41]	PCA, CLA
[38]	DL
[29]	DL
[42]	SVM, KME, CLA

4.3. Machine Learning Applications for fake news detection

Analyzing the approaches from the literature, we found that researchers used many techniques to tackle this big issue of the spreading of fake news on different social network in Latin America. These techniques are located in the spectrum of both supervised and unsupervised learning. Table 2 presents the references of authors and the techniques used by each one. As shown in this table, there is a trend to combine several algorithms in the same study in a kind of pipeline with several steps, to obtain better prediction results. In this sense, many studies may apply a learning technique in a step named feature learning step, before the effective process of detecting fake news. For instance, Reis et al. [33] used **K-Means (KME)** in this step. On the other hand, there are also studies that apply one technique in isolation way as shown in table 2.

Based on our analysis of selected works, we noticed that **RF** is the most used ML technique for fake news detection on social network from selected papers of our review, been specifically used by [5], [25], [2], [29], [18], [28] and [43]. In the literature, some researchers combine it with other ML techniques like SVM [24] and [28], XGB[18]. Using RF and XGB, Jeronimo et al. [18][18] proposed a model based on news subjectivity to detect false and legitimate news.

The RF is followed by SVM, the second most used algorithm for those works directed to fake news identification in America Latina. Researchers in [19], [29], [28], [30] and [42] use it for this purpose in conjunction with other techniques.

Furthermore, our study identified that some researchers also use Deep Learning in their works. Deep Learning which we present in the Background section is responsible for many of the recent advances in the field of AI such as the emergence of the popular chatGPT from openAI, and similar tools. Then it is natural that researchers targeted to identifying fake news on social network seek to take advantage of this field. Some researchers as Oliveira and Albuquerque [29] apply it to seek to understand the dynamics of false message propagation on Twitter, using tweets published in Brazil about Covid-19 during the pandemic period. Similarly, Lukasik et al. [25] proposed a supervised classifier combining RF, SVM, LR, and clustering techniques to classify the stance expressed in each individual tweet in a conversation around a rumor. Another usage of DL for fake news detection is made by Sengupta et al. [38], but combined with blockchains.

For these works applying DL for fake news detection, our analyses identified two kinds of ANN: CNN or RNN (described in section Background). In this case, CNN techniques are used to detect fake news in images and videos based on content on social network by [20] and [24]. On the other hand RNN are used for example in [14] and [5].

The different works found in the literature can be organized into groups of approaches for fake news detection on social network by Machine Learning, according to common characteristics shared in each group. We present these groups in section 4.4.

4.4. Taxonomy for fake news detection on social network

From our analysis of works selected from the literature, we propose to organize the different approaches for detecting Fake News in social network using Machine Learning into a taxonomy. This taxonomy is an extension of the taxonomy proposed by Rastogi and Bansal[35], which is composed of 4 dimensions namely: domain, type of fake news features, data types and social network platforms. Therefore, our taxonomy adds a fifth dimension to this initial taxonomy of Rastogi and Bansal[35], and covers the kind of approaches aimed at identifying Fake News in by Machine Learning on social network. This taxonomy is composed of four categories. Figure 5 shows this taxonomy.

The first category includes approaches which focus only on the analysis of content published on social

networks, applying one or more Machine Learning techniques, to detect Fake News. Most of the works investigated in these reviews fall into this category, including [28], [33], [18], [14], [36],[30],[31],[40],[19] and [41]. As example, [36] apply DL and LSTM for detection of fake news relative to digital manipulation of images, such as images tampering or repurposing. Such iamges has been recurrent on social network, even many meme images that go viral may contain fake news. As another example, Morais et al.[28] combined SVM and RF, two supervised learning techniques, to identify fake news related to politics in social networks, extracting training features in news portals. Its approach allows classifying any information into 4 classes, namely: objectivity/satire and false/legitimate.

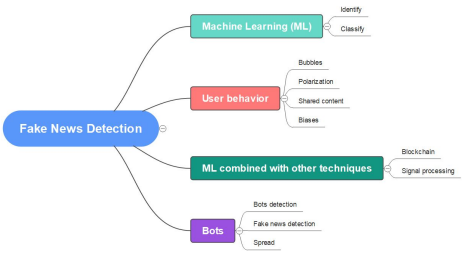

Figure 5. Taxonomy on approach for Fake News detection With Machine Learning.

Since user behavior can help identify fake news[21], another category of approaches that we identified include works that focus on user behavior to detect and reduce fake news on social network. These approaches consider features like subject polarization, bubbles, social isolation, and biases and they combine supervised and unsupervised learning techniques, as well as Deep Learning techniques. This category includes for instance [25],[29], [21] and [24]. As an example, in [24] an approach is proposed for detecting Fake News in social networks using CNN, a deep learning technique, studying characteristics of users of social networks. This work shows that users propagating fake news have characteristics which are significantly different than those of other users. In turn, applying Deep Learning, Oliveira and Albuquerque[29] used a theory of Cultural Evolution (CE) to detect fake news related to the Covid-19 pandemic on Twitter, with a specific focus on socially isolated users. With this, they came to the conclusion that social isolation reduces social cultural interactions. The latter, surprisingly, tend to reduce possible distortions of information regarding CE. Thus, according to them, such social isolation results in the formation of social bubbles and the proliferation of fake news on social networks, as a result of distorted information.

We also find in the literature approaches that, in addition to applying Machine Learning in their investigations, also seek to take advantage of other computational techniques that are not necessarily AI, such as blockchains[38], signal processing[20] or sentiment analysis [42][5]. As an example of this category, [38] proposed a solution that applies Deep Learning to detect fake news on social networks, where news is previously categorized by reviewers, through a secure voting system. With voting, this system guarantees the authenticity of the veracity of the information through a blockchain-based model, called ProBlock.

Less frequently, in for the fourth category we found approaches which consider the direct relationship between the detection of fake news and the detection of social bots. They seek to detect fake news by identifying bots on social network and investigating the characteristics of these bots' posts. In fact, due to the popularization of artificial intelligence and related areas of cognitive computing, the number of bots has exploded in social networks[39]. Some of them are used merely to boost advertisements, but others are used to spread fake news. In this case, when identifying bots, a major challenge for approaches in this category is to identify fake news, separating them from advertisements. Works by [37] and [16] belong to this last category. For instance, based on the premise that bot detection can improve the detection of fake news on social network, [16] seek to detect such bots and their respective fake news using machine learning, focusing on fake news related to the Covid-19 pandemic on Twitter.

5. Limitations and Challenges in cultural aspects

According to sociologists, culture consists of the values, beliefs, language, communication systems and practices that people share and that can be used to define them as a collective[13]. In Latin America cultural elements are diverse and rich, but for the purpose of this paper we focus only on a subset of cultural aspects, which are likely to influence the Machine Learning-based detection of fake news on social network.

First, works aimed at this region face the cultural language barrier, since the local contents of the social networks are mainly in Portuguese or Spanish. The application of machine techniques requires that the content published on the studied social network be treated. This goes through a series of steps through natural language processing. However, the tools

which perform such treatments have limited support for languages other than English. The direct using in local languages would bring gain by capturing the real meaning of terms involved, meaning which could be lost in translation.

In addition, another predominant cultural aspect in the region is the use of figures of speech such as sarcasm and satire in communications, whether oral or written. Due to linguistic characteristics between English and local languages or even the regionalization of the same language, the way of expressing sarcasm is different from one country to another. For example, the way of telling a joke in Mexico is different from the way of telling it in England. This difference is, in a way, the so-called confounding variable. Sarcasm is very common to find on social networks, often as criticisms of society and politics[15]. Few works examined addressed this cultural aspect to identify fake news behind sarcasm or satire. Only [28] and [18] tried to identify satire news, but there is still a need to improve in this field. In the literature [15] proposed a way to find fake news based on sarcasm, using machine learning and sentiment analysis. This paper was not returned by our search query. A challenge remains to find fake news hidden in figures of speech from social networks by Machine Learning.

On the other hand, been economics in development, Latin America countries have high social inequality, suffer from serious problems of violence, public security and high levels of corruption. All these ingredients are great facilitators of political polarization, resulting in fake news in many domains. Moreover, their population has a great cultural and religious mix, combining mainly European, African and Indigenous heritages. Popular beliefs, syncretism and superstition are very expressive for a large part of this population, favoring the emergence of many popular legend, fantasist stories (a well-known maxim in this region is ”people love a story”). No work examined considered these specific cultural aspects. It could bring gains, for example, to build datasets on fallacious legends, to be confronted with what is on the social network.

We also faced others general limitations that may have impacted our work. The main datasets found in our research were mostly related to Brazil and North America, which included Mexico. Many Latin American countries were not identified in the papers selected from our research query, not allowing our analyses to extend or evaluate the influence of cultural aspects in the propagation and detection of fake news with respect to Latin American countries, as initially proposed. It is also possible that some studies are in local languages other than English. In fact, altering the research query including terms in Spanish and Portuguese, 8 additional papers were returned in these languages, indicating the existence of local works on the field. For instance, to tackle the issue of lack of datasets in Portuguese language, [12] proposed an approach to create a corpus for helping fake news detection in Portuguese language, extracting news from online and Brazilian portals. They successfully tested it to detect fake news relative to politics and health domains applying LSTM model.

Another limitation of our review is that some selected papers do not only address the detection of fake news specifically for Latin American countries, but also consider countries from others regions of the world.

6. Conclusion

In this study, we saw that the detection of fake news in social networks is a topic of great importance and concern nowadays, since their dissemination has taken great proportions and cause huge impacts. We identified that most of the studies focused on Machine Learning models for detecting fake news mainly on the Twitter social network. These approaches mostly used text-type news and so their models require natural language processing, and most of their datasets are in English and Portuguese.

From our analysis, we identified that most of papers addressed detection of fake news using different Machine Learning techniques including supervised, unsupervised and deep learning ones. They used these techniques in isolation way or combining them, and studying the effectiveness of different techniques according to datasets adopted as object of their researches. However we identified a growing use of deep learning techniques, following the trend observed in many areas to benefit from the advances and gains that this AI branch has provided to society.

Our study showed that the most predominant domains subject to fake news detection were politics and health. During the period of most of these studies, the world experienced a pandemic, therefore the dissemination of fake news associated with this theme was globalized. In the political sphere, fake news resulted from events such as the presidential elections in Brazil and the fight against violence in Mexico, direct reflections of events in these two countries.

Based on our analysis, we propose a taxonomy of approaches using Machine Learning for fake news detection on social networks in Latin America, composed of four main categories, according to shared characteristics: only Machine Learning on content, considering user behavior, associating machine

learning with techniques that are not necessarily AI or considering the action of the social bot.

We observe that expressive cultural aspects of Latin America, such as syncretism, superstition, sarcasm, and satire are likely to favor the emergence and proliferation of fake news on social networks. However, in most of the works investigated, these aspects were not considered. We only observed the use of satire by two works, which presented limitations. Thus, the investigation of these influences of cultural aspects remains open research.

For other future work, one direction would be analyzing and comparing our proposed taxonomy with studies of regions different from Latin America. More investigations are also needed with respect to user behavior in spreading fake news on different social network, considering cultural characteristics in different regions of the world.

References

[1] D. Roy . Vosoughi and S. Ara. *Lies spread faster than the truth*. 2018. URL: https://www.science.org/doi/full/10.1126/science.aap9559. (accessed: 15.05.2023).

[2] Esma A., Amri S., and Brassard G. "Fake news, disinformation and misinformation in social media: a review". In: *Soc. Netw. Anal. Min.* (2023).

[3] Ahmed A.A.A. et al. "Detecting Fake News Using Machine Learning: A Systematic Literature Review". In: *arXiv e-prints* (2021), arXiv–2102.

[4] Capuano N. et al. "Content-Based Fake News Detection With Machine and Deep Learning: a Systematic Review". In: *Neurocomputing* 530 (2023), pp. 91–103.

[5] Wafa Alotaibi, Faye Alomary, and Raouia Mokni. "COVID-19 vaccine rejection causes based on Twitter people's opinions analysis using deep learning". In: *Soc. Netw. Anal. Min.* (2023).

[6] J. Araújo. "Manifestação pela democracia marca um mês da invasão às sedes dos três Poderes". In: (2023). URL: https://www12.senado.leg.br/radio/1/noticia/2023/02/08/manifestacao-pela-democracia-marca-um-mes-da-invasao-as-sedes-dos-tres-poderes.

[7] Mustafa A. Al-Asadi and Sakir Tasdemir. *Using Artificial Intelligence Against the Phenomenon of Fake News: A Systematic Literature Review*. Springer International Publishing, 2022.

[8] Kitchenham B. and Brereton P. "A systematic review of systematic review process research in software engineering". In: *Information and Software Technology* 55.12 (2013), pp. 2049–2075.

[9] BBC. *US Capitol riots: What happened and who has been punished?* 2022. URL: https://www.bbc.com/news/world-60265900. (accessed: 04.05.2023).

[10] Aaron Blake. *A new study suggests fake news might have won Donald Trump the 2016 election*. 2018. URL: https://www.washingtonpost.com/news/the-fix/wp/2018/04/03/a-new-study-suggests-fake-news-might-have-won-donald-trump-the-2016-election/. (accessed:06.08.2023).

[11] Murari et al. C. "A Review of Fake News Detection Methods using Machine Learning". In: *2021 2nd International Conference for Emerging Technology (INCET)* (2021), pp. 1–5.

[12] A. C. Charles, L. Ruback, and J. Oliveira. "Fakepedia Corpus: A Flexible Fake News Corpus in Portuguese". In: Springer-Verlag, 2022, pp. 37–45.

[13] Nicki Lisa Cole. *So What Is Culture, Exactly?* 2019. URL: https://www.thoughtco.com/culture-definition-4135409#:~:text=According%20to%20sociologists%2C%20culture%20consists,to%20that%20group%20or%20society.. (accessed:06.10.2023).

[14] Jia Ding, Yongjun Hu, and Huiyou Chang. "BERT-Based Mental Model, a Better Fake News Detector". In: (2020), pp. 396–400.

[15] Silva F. C. D., Garcia A. C. B., and Sean W. M. S. "Sentiment Gradient - Improving Sentiment Analysis with Entropy Increase". In: *Inteligencia Artificial* 26.71 (2023), pp. 114–130.

[16] M. Heidari et al. "BERT Model for Fake News Detection Based on Social Bot Activities in the COVID-19 Pandemic". In: (2021).

[17] IBM. *What are convolutional neural networks?* 2023. URL: https://www.ibm.com/topics/convolutional-neural-networks. (accessed:06.10.2023).

[18] Caio Libanio Melo Jeronimo et al. "Fake News Classification Based on Subjective Language". In: (2020).

[19] Manmeet Kaur et al. “Detection of fake images on whatsApp using socio-temporal features”. In: *Soc. Netw. Anal. Min.* (2022).

[20] Spencer Lee Kirn and Mark K. Hinders. “Dynamic wavelet fingerprint for differentiation of tweet storm types”. In: *Soc. Netw. Anal. Min.* (2020).

[21] J. K. Kobellarz et al. “Reaching the bubble may not be enough: news media role in online political polarization”. In: *EPJ Data Sci.* (2022).

[22] S. Kumar and N. Shah. *False Information on Web and Social Media: A Survey*. 2018. URL: https://arxiv.org/abs/1804.08559. (accessed: 20.05.2023).

[23] Severin Landolt, Thiemo Wambsganß, and Matthias Söllner. “A Taxonomy for Deep Learning in Natural Language Processing”. In: *Hawaii International Conference on System Sciences(HICSS-54)*. 2021.

[24] Yang Liu and Yi-Fang Brook Wu. “FNED: A Deep Network for Fake News Early Detection on Social Media”. In: *ACM Trans. Inf. Syst.* (2020).

[25] Michal Lukasik et al. “Gaussian Processes for Rumour Stance Classification in Social Media”. In: *ACM Trans. Inf. Syst.* (2019).

[26] S.I. Manzoor, J. Singla, and Nikita. “Fake News Detection Using Machine Learning approaches: A systematic Review”. In: (2019), pp. 230–234.

[27] Francisco D.C. Medeiros and Reinaldo Bezerra Braga. “Fake News Detection in Social Media: A Systematic Review”. In: (2020).

[28] J. I. Morais et al. “Deciding among Fake, Satirical, Objective and Legitimate News: A Multi-Label Classification System”. In: *Proceedings of the XV Brazilian Symposium on Information Systems*. Association for Computing Machinery, 2019.

[29] D. V. B. Oliveira and U. P. Albuquerque. “Cultural Evolution and Digital Media: Diffusion of Fake News About COVID-19 on Twitter”. In: *SN COMPUT. SCI.* (2021).

[30] N. R. Oliveira, D. S. V. Medeiros, and D. M. F. Mattos. “A Sensitive Stylistic Approach to Identify Fake News on SocialNetworking”. In: *IEEE Signal Processing Letters* (2020), pp. 1250–1254.

[31] Oladapo Oyebode et al. “COVID-19 Pandemic: Identifying Key Issues Using Social Media and Natural Language Processing”. In: *J Healthc Inform Res* (2022).

[32] Sintia Radu. *How Digital Disinformation Spread Through the 2018 Latin American Elections*. 2019. URL: https://www.usnews.com/news/best-countries/articles/2019-08-02/mexico-colombia-and-brazil-faced-heavy-amounts-of-disinformation-in-2018-elections. (accessed:06.08.2023).

[33] Julio C. S. Reis et al. “Explainable Machine Learning for Fake News Detection”. In: (2019), pp. 17–26. DOI: 10.1145/3292522.3326027. URL: https://doi-org.ez39.periodicos.capes.gov.br/10.1145/3292522.3326027.

[34] Kumar S. and Bhavna A. “A Review of Fake News Detection Using Machine Learning Techniques”. In: (2021), pp. 1–8.

[35] Rastogi S. and Bansal D. “review on fake news detection 3T’s: typology, time of detection, taxonomies”. In: *International Journal of Information Security* (2023), pp. 177–212.

[36] E. Sabir et al. “Deep Multimodal Image-Repurposing Detection”. In: *Proceedings of the 26th ACM international conference on Multimedia*. ACM, Oct. 2018.

[37] R. Schuchard et al. “Bot stamina: examining the influence and staying power of bots in online social networks”. In: *Appl Netw Sci* (2019).

[38] E. Sengupta et al. “ProBlock: a novel approach for fake news detection”. In: *Cluster Comput* (2021), pp. 3779–3795.

[39] F. C. D. Silva, R. Vieira, and A. C. B. Garcia. “Can Machines Learn to Detect Fake News? A Survey Focused on Social Media”. In: *Hawaii International Conference on System Sciences(HICSS-52)*. 2019.

[40] J. V. Souza et al. “A systematic mapping on automatic classification of fake news in social media”. In: *Soc. Netw. Anal. Min.* (2020).

[41] M. C. Souza et al. “A network-based positive and unlabeled learning approach for fake news detection”. In: *Machine Learning* (2022), pp. 3549–3592.

[42] Minyu Wan et al. “Data-driven analytics of COVID-19 ‘infodemic’”. In: *International Journal Data Science Anal* (2023), pp. 313–327.

[43] William Yang Wang. *”Liar, Liar Pants on Fire”: A New Benchmark Dataset for Fake News Detection*. 2017.