Fake News Detection on social media: Review of Literature

Malek Al Gabri1*, Ebtsam Nasser Ali Abu Huliqah1 Mossa Ghurab1 Abdualmajed A. G. Al-Khulaidi1, Ghaleb H. Al Gaphari1

1 Department of Computer Science, Faculty of Computer and Information Technology, Sana’a University, Sana’a, Yemen.
*Corresponding author: malekye@su.edu.ye

ARTICLE INFO

Article history:
Received: Aug 5, 2023
Accepted: Nov 18, 2023
Published: Jan, 2024

KEYWORDS

1. Fake news
2. Social media
3. Machine learning
4. Deep learning
5. False Information

ABSTRACT

In today's world, social media has become one of the most accessible sources of news for people worldwide due to its low cost, easy accessibility, and rapid dissemination. However, the harmful effects of fake news on individuals and society have been evident for a long time. Despite the challenges in detecting false information, numerous scholars are working to understand the issue and its characteristics.

The purpose of this paper is to analyze contemporary frameworks that make use of different machine learning techniques in order to further our understanding of false news identification. The efficiency of these frameworks in recognizing and halting the spread of false information may be compared by examining various methods. A particular technique was used in the review process to provide a thorough and instructive assessment.

Even though a lot of research has been done to identify false news, more can be done in the future. These studies have many limitations such as biased data, adaptability, credibility analysis and the news is only categorized as legitimate or unauthentic by their solutions. Nevertheless, a system of ratings or scores to assess the reliability of news is necessary for a workable approach.

CONTENTS

1. Introduction
2. Fake news repercussions
3. Results discussion
4. Related work
5. Limitation
6. References

1. Introduction:

Because fake news has so many ramifications, it has always been difficult to identify and counteract. Its roots are in the 17th-century dissemination of propaganda, which developed into disinformation throughout the course of the Cold War [6]. The introduction of social media platforms in the contemporary period has made matters worse by accelerating the dissemination of misleading information.

Social media platforms like Facebook, Instagram, and Twitter have evolved into means of quickly disseminating and retrieving information in recent years. It has been established those rumors and fake news are closely related. Fake news is defined as material that has been purposefully manufactured or
misrepresented, whereas rumors are unreliable and dubious information that is disseminated without the intention of misleading. Both have the power to greatly influence public opinion [28]. It is essential to confirm the veracity of news that is offered online.

The immediate reach of social media to millions. Of individuals requires dynamic observing to avoid the quick dispersal of untrue data. In a few cases, people abuse this circumstance to assist their claim motivation or goals, underscoring the require for compelling direction and tending to this issue suitably. Fake news information, comprising of verifiably wrong and intentioned wrong data, has the potential to deceive anybody.

This paper presents a comparative investigation of state-of-the-art models for rumor location, utilizing machine learning, profound learning, and cross breed approaches. Furthermore, it examines potential future headings within the field of rumor discovery. The paper is organized as takes after: Segment 2 gives an outline of the impacts of fake news, Section 3 traces the basic steps in discovery, Segment 4 examines confinements in existing considers, Section 5 recommendations for future research directions and at long last, Area 6 concludes by emphasizing the require for assist investigate within the region of fake news distinguishing proof.

2. Fake news repercussions
On social media, bogus news is more widely disseminated. Fake news may have significant negative effects on the social, political, and economic contexts. Fraudulent news and fake information come in many forms. Fake news has a significant impact because information shapes our perspective of the world. Based on this information, we make important judgments. We build an opinion on a circumstance or a group of individuals based on the facts we learn. If we encounter contrived, misleading, twisted, or phony information online, we cannot make wise selections.

• Fake news and rumors are similar in that they both entail the dissemination of false or unreliable information. Nonetheless, there are a few significant variations between the two:
  - Verification: Unverified facts or tales that spread among people via social media or word-of-mouth are known as rumors. They could result from unfinished business, rumors, or conjecture. On the other side, information that has been purposefully falsified or misrepresented and presented as news is known as fake news. Fake news purposefully contains incorrect information; however, it frequently imitates the format and style of real news stories.
  - Goal: The deliberate creation of fake news aims to mislead or control readers. It could be made with the intention of misleading people, advancing a particular goal, or getting clicks and ad money. On the other hand, someone speculating or spreading untrusted information without malice can accidentally give rise to rumors.
  - Source: Untrustworthy or disreputable sources are frequently the source of fake news. It could be produced by people or groups who want to disseminate false information or who have an agenda. Conversely, rumors might originate from people, miscommunications, or incorrect interpretations of what happened.
  - Disseminate: Through internet forums, websites, or social media accounts, fake news can be purposefully disseminated. Bots or concerted efforts to reach a larger audience may amplify it. Conversely, rumors typically disseminate via social media sharing, word-of-mouth, and encounters with other people.
  - Impact: Because fake news is created and distributed on purpose, it is more likely to have a wider effect. It has the power to sway public opinion, mold narratives, and have a big impact on people's lives, communities, or even political processes. Though they can also be harmful, rumors may only affect a small area or a larger area depending on their context and reach.

The key effects of false news are impact on:

• Health Effects:
A growing number of individuals are using the Internet to look for news about their health.
People’s lives might be impacted by fake news in the health sector [36]. Therefore, this is one of the major challenges today. The recent year has seen a significant increase in false information concerning health. As a result of pressure from physicians, legislators, and health activists, social media companies have altered their policies to prohibit or restrict the dissemination of health misinformation.

- **Financial Impact:**
  Fake news is presently a major issue in all sectors of the economy. In order to increase their earnings, dishonest businesspeople circulated false information. Stock prices may drop as a result of false information and Misinformation can lead to market volatility and possible economic repercussions by causing uncertainty, warping investor perceptions, and influencing investor actions [26].

- **Democratic processes and public trust in institutions impact:**
  Fake news weakens public confidence in institutions and has a detrimental effect on democratic processes. It sways elections and the formulation of public policy, manipulates public opinion, and falsifies information. It weakens confidence in the government, the media, and other reliable sources. Fake news divides people, generates echo chambers, and stifles fruitful conversation. It erodes the credibility of the media and makes it harder to distinguish fact from fiction in journalism. Furthermore, fake news undermines social cohesiveness by instigating violence and taking advantage of delicate subjects. Promoting media literacy, fact-checking, and tightening laws are essential to addressing these effects.

### 3. Related work

The fake news detection frameworks suggested usually have the following fundamental steps for detection:

- Getting of the dataset
- Data preprocessing: (Cleaning, Tokenization, stop words Removal, Stemming, Normalization...etc.)
- Data spilling (Training and Testing)
- Features Extraction: (TF-IDF... etc.)
- Model training
- Classification within a Proposed model (Fake or Real)

This review’s main objective is showing and comparing approaches of fake news detection on social media.

### Table 1: relevant papers in the area

<table>
<thead>
<tr>
<th>Ref</th>
<th>Year</th>
<th>Methodology</th>
<th>Contributions</th>
<th>Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>[16]</td>
<td>2022</td>
<td>The study used eight Arabic contextualized embedding models, the majority of which had not been previously used for Arabic fake news detection. Additionally, the study included an analysis of the state-of-the-art Arabic contextualized embedding models and a comparison with similar fake news detection systems.</td>
<td>Created a huge and varied dataset of Arabic fake news. And developed and evaluated transformer-based classifiers to identify fake news.</td>
<td>There are two news datasets: the first is a translated dataset, and the other is an original Arabic dataset.</td>
</tr>
<tr>
<td>[10]</td>
<td>2021</td>
<td>The focus of this study is to introduce a method for extracting information from Arabic social media text, which includes four stages: collecting data, cleaning it, enriching it, and making it available.</td>
<td>The proposed approach is successful in addressing the challenges of pre-processing and NLP, and provides a well-structured, well-curated, and organized Arabic dataset that has proven</td>
<td>Data of Arabic text are gathered from social media servers and stored in a local database file.</td>
</tr>
<tr>
<td>[22]</td>
<td>2022</td>
<td>Framework that incorporates the automated feature extraction of entities.</td>
<td>The framework outperforms all other methods tested in terms of three key assessment metrics: F1, precision, recall and accuracy score.</td>
<td>There are Four real world Arabic fake news datasets used: Covid19Fakes, Ara, News , Satirical News and ANS dataset.</td>
</tr>
<tr>
<td>[23]</td>
<td>2022</td>
<td>This study used hybrid deep learning (DL) model. This model includes CNN-LSTM modalities. Word vectors, pretrained on Arabic news, were included with the corrected words at a given word length. The JSO optimization algorithm was combined with the framework.</td>
<td>With a precision of 81.6%, the suggested CNN-LSTM provides perfect performance.</td>
<td>One dataset (ANS) that collect Arabic news headlines.</td>
</tr>
<tr>
<td>[5]</td>
<td>2022</td>
<td>A five-category classification system for textual tweets in the Arabic language based on linguistic traits and content. And employed two distinct textual representations: Stemmed text with term frequency inverse document frequency and word embedding using Word2vec (tf-idf) and classification with many supervised machine learning algorithms.</td>
<td>Tops the current state-of-the-art score of 92.95% using a deep learning approach, RNN-GRU.</td>
<td>A dataset of about 35,600 Arabic tweets was gathered manually and carefully annotated.</td>
</tr>
<tr>
<td>[14]</td>
<td>2022</td>
<td>Model based on DNN networks trained with the application of the Flair library for detecting true and false information, as well as the application of the remote, cloud computing platform, GPU cards.</td>
<td>The proposed model allows analysis of texts with high accuracy of 99.8%.</td>
<td>Two freely dataset available in Kaggle one called “ISOT Fake News Dataset”, the other called “Getting real about fake news” (GRaFN).</td>
</tr>
<tr>
<td>[19]</td>
<td>2022</td>
<td>A deep learning approach that combines pre-trained word embedding, CNN LSTM, and ResNet architectures with four separate training datasets. To lessen data imbalances between classes, each data is put through a data augmentation process utilizing the back translation technique.</td>
<td>Choosing the right word embeddings improves the ability of deep learning to identify fake news. On every test dataset, LSTM architecture fared better</td>
<td>Four datasets the first is: SOT Fake News Dataset [2], the second: Fake News Dataset [20] the third is- Fake or Real News Dataset</td>
</tr>
<tr>
<td>[1]</td>
<td>2020</td>
<td>Identify linguistic patterns that distinguish false news from accurate reporting, and then extraction various textual characteristics from the articles</td>
<td>High performance of ensemble learners.</td>
<td>Three freely dataset available in kaggle includes both fake and truthful news articles from multiple domains.</td>
</tr>
<tr>
<td>Reference</td>
<td>Year</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[13]</td>
<td>2021</td>
<td>A hybrid model using various kernel sizes and filters is employed in multiple convolutional neural network (CNN) branches, each with Long Short-Term Memory (LSTM) layers. The proposed model achieved a high accuracy of 91.88% with the PHHEME dataset and 98.62% with the FN-COV dataset. Two datasets were used during the COVID-19 epidemic. One of them, FN-COV, gathered 69976 false and genuine news stories with tags such as &quot;social-distancing,&quot; &quot;covid19,&quot; and &quot;quarantine.&quot; The other dataset is called PHHEME.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[25]</td>
<td>2022</td>
<td>Dense neural network design to precisely anticipate the relationship between a given title and article body. The proposed model outperforms existing model architectures by 2.5%. The specific dataset used for the FNC-I challenge is Emergent, a digital journalism project for rumor debunking.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[12]</td>
<td>2021</td>
<td>Merging several parallel blocks of the single-layer deep Convolutional Neural Network (CNN) with BERT and varied kernel sizes, and filters, we propose a BERT-based (Bidirectional Encoder Representations from Transformers) deep learning technique (FakeBERT). With an accuracy of 98.90%, the suggested model (FakeBERT) performs better than the current models. One dataset real-world fake news dataset.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[4]</td>
<td>2022</td>
<td>Four machine learning models were integrated to forecast the characteristics that influence how Arabic internet users share and interact with news content. The Term Frequency-Inverse Document Frequency method (TF-IDF) was used to discover these characteristics. The high accuracy of proposed models. One dataset is &quot;The Ara News&quot; dataset, which was made publicly available (El-madany et al., 2020).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[27]</td>
<td>2021</td>
<td>Semantic features were inspired from the discourse parse and the rhetoric relations in Arabic by using semi-supervised SVM. The proposed approach increased the accuracy of the system to 85.99%. There are many translated datasets.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[9]</td>
<td>2022</td>
<td>A content-based approach using a CNN to predict whether a given text is true or false. The development of Arabic Language Processing studies. When used on the same Arabic dataset, the model performs better than the FNC models. The publicly available Arabic dataset is &quot;The Arabic Fact-Checking and Stance Detection Corpus.&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[24]</td>
<td>2021</td>
<td>A clever classification model using Natural Language Processing (NLP), Machine Learning (ML), and Harris Hawks Optimizer (HHO) as a wrapper-based feature investigated the impact of feature selection on fake news classification and high excellent results of the proposed model. An Arabic Twitter corpus was used to assess the efficiency of the proposed approach.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>Methodology</td>
<td>Performance</td>
<td>Dataset/Details</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>Construct a classifier using RNN approach models (vanilla, GRU, and LSTMs)</td>
<td>High performance GRU (0.217), LSTM (0.2166) and vanilla (0.215).</td>
<td>LAIR dataset.</td>
<td></td>
</tr>
<tr>
<td>2022</td>
<td>An efficient method for enhancing Arabic rumor identification is presented, which uses a mix of extracted characteristics from the contents and accounts of Arabic tweets. The method is based on an XGBoost classifier that can minimize overfitting and unbalanced class issues during the training process.</td>
<td>The methods employed in An Effective Strategy for Rumor Detection of Arabic Tweets have not performed as well as the suggested XGB-based approach.</td>
<td>The dataset was gathered from several Twitter tweets, including both rumors and non-rumors, using the Search API.</td>
<td></td>
</tr>
<tr>
<td>2021</td>
<td>A novel fake news detection method, Graph-aware Co-Attention Networks (GCAN). that, given the order of its retweeters, can determine whether a short-text tweet is false.</td>
<td>There are two well-known fake news datasets, Twitter15 and Twitter16.</td>
<td>Two well-known fake news datasets, Twitter15 and Twitter16.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Methodology</th>
<th>Performance</th>
<th>Dataset/Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>2020</td>
<td>Following a linked matrix-tensor factorization approach, the news content is combined with the tensor to get a latent representation of both the news content and the social environment.</td>
<td>The matrix is represented as N, where N is the total number of news items in the dataset and V is the total number of words in the vocabulary. This matrix is used to count the word sequences in news articles.</td>
<td>There are two datasets commonly used for fake news detection: BuzzFeed and PolitiFact. These datasets contain labeled news articles and are often used to train and evaluate machine learning models for fake news detection.</td>
</tr>
<tr>
<td>2021</td>
<td>Using a stack ensemble approach comprising three algorithms to boost the detection accuracy, a classification model for false news detection on social media was developed.</td>
<td>Based on viewpoints on news content and social content, the suggested methodology reliably identifies more instances of false news and more accurately.</td>
<td>A multidimensional dataset with attributes that are unimportant for proposed model’s formulation. The most effective features from the dataset were chosen using a mutual information gain features selection approach.</td>
</tr>
</tbody>
</table>
4. Limitation
Current research may face some limitations, including:
- **Data bias**: The models' capacity to handle a variety of false information may be impacted by biases included in the datasets utilized in research. It's possible that some false news stories aren't fairly represented or that some legitimate news stories are favored over others.
- **Adaptability** to novel forms of fake news: As fake news has changed and so have the ways in which it is disseminated, models may find it difficult to adjust to these novel trends. This can necessitate regular model changes and the provision of a test environment that takes into account the most recent advancements in the fake news industry.
- **Determining** the credibility of news can be difficult since it involves examining a number of variables, including the news source, the veracity of the content, and fact-checking. This can make it necessary to create sophisticated models that can evaluate and analyze these elements in an efficient manner.
- **The capacity** of the models employed in the current research to handle other languages, particularly non-English languages, may be restricted. This might necessitate creating multilingual models and enhancing these languages' capacities to identify false news.
- **Evaluating** the credibility of news instead of categorizing it: although many studies have been conducted to identify fake news, there is still room for future investigation. These studies have limitations in that their solutions only classify news as either authentic or unauthentic. However, an effective solution requires a score or rating system that determines the credibility of news.

5. Recommendations for future research directions
- Creation of real-time detection systems: With the speed at which false information circulates, real-time detection systems that can promptly recognize and flag false information are required. In order to enable timely action, future research could concentrate on creating effective algorithms and techniques that can recognize and react to fake news in real-time.
- Investigation of multimodal fake news detection: Textual material is not the only source of fake news. Audio, video, and image files may also be included. Subsequent investigations may examine the amalgamation of other modalities, including text, image, and audio analysis, in order to enhance the precision and resilience of false information identification. This can entail utilizing methods from audio processing, natural language processing, and computer vision to create thorough multimodal detection models.
- Social impact and ethical issues: Fake news has a big impact on society, influencing elections, public opinion, and social cohesiveness, among other things. Subsequent investigations ought to explore the moral implications associated with identifying false news, including but not limited to guaranteeing openness, impartiality, and responsibility in the creation and implementation of detection mechanisms.
Furthermore, researching the wider societal implications of false news and comprehending how it affects people, communities, and democracy may be very insightful for stakeholders and policymakers.

• User-centric approaches: User-centric solutions might result from involving end users in the research process, such as consumers of news and social media. Subsequent investigations may concentrate on comprehending the viewpoints, requirements, and actions of consumers about the ingestion of news and false information. This can help with the creation of approachable tools, educational initiatives, and public awareness campaigns that enable people to recognize and lessen the effects of false information.

• Collaborations across disciplines: To detect fake news, one needs knowledge from a variety of fields, including journalism, computer science, psychology, and the social sciences. Cross-disciplinary cooperation should be promoted in future research in order to take advantage of different viewpoints and approaches. Working together, academics, professionals, legislators, and media professionals may promote creativity and develop comprehensive strategies for countering false information.

6. Conclusion
Utilizing deep learning techniques provides the advantage of automatically learning feature extraction from data, rather than relying on manual feature extraction. As social media continues to expand, the issue of fake news is becoming increasingly severe. To combat this, researchers are working diligently to develop methods to protect society from the spread of fake news.

This review article presents the latest techniques utilized to detect fake news on social media platforms. We provide a taxonomy of approaches for identifying false news, categorizing and summarizing the various methods used in current research. With the introduction of cutting-edge deep learning network architectures, fake news detection will continue to be a hot topic for research for some time. This review will help academics better understand the issues, solutions, and potential paths in the field of false news identification.

7. References


