

Article

Towards the Detection of Fake News on Social Networks Contributing to the Improvement of Trust and Transparency in Recommendation Systems: Trends and Challenges

Oumaima Stitini *, Soulaymane Kaloun and Omar Bencharef

Computer and System Engineering Laboratory, Faculty of Science and Technology, Cadi Ayyad University, Marrakesh 40000, Morocco; so.kaloun@uca.com (S.K.); o.bencharef@uca.com (O.B.)

* Correspondence: oumaima.stitini@ced.uca.ma; Tel.: +212-64-282238

Abstract: In the age of the digital revolution and the widespread usage of social networks, the modalities of information consumption and production were disrupted by the shift to instantaneous transmission. Sometimes the scoop and exclusivity are just for a few minutes. Information spreads like wildfire throughout the world, with little regard for context or critical thought, resulting in the proliferation of fake news. As a result, it is preferable to have a system that allows consumers to obtain balanced news information. Some researchers attempted to detect false and authentic news using tagged data and had some success. Online social groups propagate digital false news or fake news material in the form of shares, reshares, and repostings. This work aims to detect fake news forms dispatched on social networks to enhance the quality of trust and transparency in the social network recommendation system. It provides an overview of traditional techniques used to detect fake news and modern approaches used for multiclassification using unlabeled data. Many researchers are focusing on detecting fake news, but fewer works highlight this detection's role in improving the quality of trust in social network recommendation systems. In this research paper, we take an improved approach to assisting users in deciding which information to read by alerting them about the degree of inaccuracy of the news items they are seeing and recommending the many types of fake news that the material represents.

Keywords: multiclass classification; unlabeled data; semisupervised learning; self-training; recommendation system; fake news



Citation: Stitini, O.; Kaloun, S.; Bencharef, O. Towards the Detection of Fake News on Social Networks Contributing to the Improvement of Trust and Transparency in Recommendation Systems: Trends and Challenges. *Information* **2022**, *13*, 128. <https://doi.org/10.3390/info13030128>

Academic Editor: Arkaitz Zubiaga

Received: 29 November 2021

Accepted: 13 December 2021

Published: 3 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The widespread impact of fake news has opened up new academic directions to conduct challenging studies to counter the problem. Digital misinformation or fake news content is spread through these social communities in the form of shares, re-share, and repost [1]. Sometimes it is obligatory to extract meaningful information [2]. The spread of this misinformation through social networks has the same evolution as the transmission of infectious diseases [3]. Therefore, insights about fake news spreads could be provided from the analysis of the dynamics of transmission. For example, the recent pandemic Coronavirus, causing the illness COVID-19, can evolve and compete in a host population shaped by social contacts [4], much like rumors and fake news. Propagation of information on social networks is inundated with fake news, which can be of different forms. Some express humor, while some are serious and create doubt in public [5,6]. 40% of people living with HIV trusted sources of information about health and its associations with COVID-19 disruptions. For example, people believe that eating garlic offers protection against the virus (available at <https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-021-10856-z>, accessed on 28 November 2021). The vast majority of people, approximately 69.5%, moved to consult mainstream media, they used television networks news outlets such as CNN (24.0%), Fox News (19.3%), and other local or national stations (35.2%) due to the high fake

information spread on social networks (available at <https://www.cidrap.umn.edu/news-perspective/2020/10/trust-covid-info-sources-varies-demographics-beliefs>, accessed on 28 November 2021).

Identifying misleading information determines the news truthfulness by examining news content and related information, such as dissemination patterns [7]. Various perspectives have received much interest in addressing this issue, where fake news identification based on supervised learning dominates this domain and has succeeded. Several kinds of research aim at detecting fake news with labeled data. Different researchers have different studies aiming at different types of fake news to classify them on social networks [8]. Most studies focus on detecting fake or real news [5]. Other studies focus on a single type of fake news [9,10]. Some work concentrates upon classifying two or three types of rumors, for example, early detection of rumors [11], curbing the spread of rumors [12]. In the age of information overload, recommender systems play an essential role in identifying orders in the face of a flood of data [13,14]. It is undoubtedly critical to identify and mitigate fake news [5], representing a challenging, socially relevant problem.

In this work, we attempt to fill several gaps. We present several contributions, including the following:

- We provide an overview of traditional techniques used to detect fake news and modern approaches used for multiclassification using unlabeled data.
- We focus on detecting forms of fake news dispatched on social networks when there is a lack of labeled data.
- We aim to demonstrate how this detection can help in improving and enhancing the quality of trust and transparency in the social network recommendation system.

The key contributions of this paper are as follows: Section 2 contains our outline objectives. We describe the Literature Review about digital environments Psychology and its impact on information sharing Trust in Section 3. Then, in Section 4, we will elaborate on the state-of-the-art. In Section 5, we show our proposed methodology. We mention the results of our work on Section 6. At the end of the work, we discuss and conclude all the work in Sections 7 and 8.

2. Outline of Objectives

Our work's aims are defined as three challenges:

RQ1: How to differentiate between different fake news forms?

We describe how many types or forms exist in fake news, and we describe the difference between each other. Research on fake news is at an early stage and requires deep analysis to choose the relevant features precisely. In general, we could categorize fake news into two levels:

- High level:
 - Manufacturing: represents false information published in a newspaper to gain credibility.
 - Manipulation: decontextualization of image or video content to make false news.
 - Propaganda: whose aim is to influence public opinion and modify its perception of events.
- Low level:
 - Satire or false satirical information whose primary purpose is to provide humor to readers.
 - Clickbait: the primary goal of content is to bring attention and encourage users to click on a link to a certain web page.

False information can be categorized based on whether it is intended to do harm or has an ulterior motive. They are referred to as fake news. Figure 1 shows the main types of false information from low level to high level. The disinformation category is found at

the low level, and it refers to inaccurate information that is released without the purpose to damage. The high level is represented by the misinformation category. It is a type of incorrect information that is intended to mislead and hurt the reader.

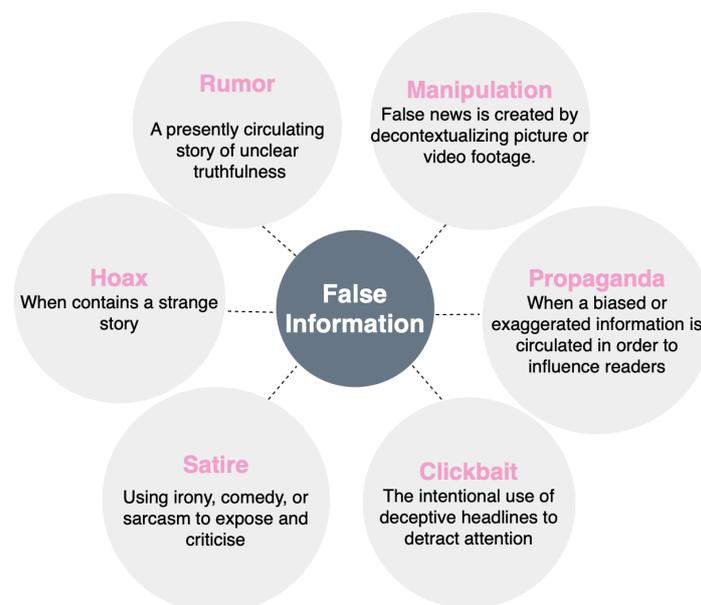


Figure 1. False information types.

RQ2: How to perform multiclass classification using unlabeled data?

We examine the limits of multiclass classification utilizing unlabeled data to identify false news and the possibilities of combining self-training algorithms with the majority voting to enhance multiclass classification performance.

Ref. [15] focuses on semisupervised learning. We present a novel multiclass loss function and new codewords to tackle the multiclass semisupervised classification issue. The suggested loss function's main purpose is to decrease the disagreement between classifier predictions and pairwise similarity. They develop a multiclass semisupervised boosting technique called GMSB using a small batch of labeled data and a vast pool of unlabeled data.

Graphs are widely utilized in numerous fields of computer science because they are a universal modeling tool that allows for the representation of structured data. The objects handled and their relationships are specified in a single, human-readable formalism. Ref. [16] This study proposes an efficient method for solving the graph matching issue in a classification setting. Ref. [17] present a novel online learning method based on the prediction margin for the partial feedback situation, inspired by previous work on learning from complementary labels, where a complementary label identifies a class to which an instance does not belong. In the context of online classification, our method focuses on the prediction margin and learning from complimentary labels.

RQ3: Mitigating infodemics, how does fake news on social networks affect user trust and transparency?

We aim to examine the relationship between exposure to and trust in fake news and study the trust and transparency in social network recommender systems.

Trust is a multidisciplinary term that was applied in a variety of contexts. Aside from being interdisciplinary, trust is a complicated term with varied connotations in different contexts. In recent years, trust and social influence patterns have attracted more attention in regard to virtual worlds. The importance of trust in social relationships cannot be overstated. Reputation may be viewed as a collective phenomena and a product of social processes, which extends well beyond any single individual's thoughts or experiences.

We might consider reputation as a natural evolution product that provides human societies with more collective cognitive potential.

3. Digital Environments Psychology and Its Impact on Information Sharing Trust

Trust is vital in an age where computers influence many of our most important decisions and machines work alongside humans to serve consumers. The epidemic compelled fast digitization all across the world: schools were modified to enable online learning, many occupations became completely remote, and automation increased in a wide range of industries. Trust is a multidisciplinary term that was applied in a variety of contexts. Aside from being interdisciplinary, trust is a complicated term with varied connotations in different contexts. Furthermore, several questions should be asked:

1. How can virtual reality influence trust, and thus, impact user decision-making in advance?
2. How does reputation within a virtual environment affect trust in information exchange?
3. Why is there a tendency to trust strangers more than they deserve, and why do strangers have a good reputation?

The idea of reputation is strongly connected to trust and may be utilized in virtual environments contexts to gauge members' confidence in cloud providers. As a result, reputation is described as a collection of feedback on an item, character, or entity's traits (reliability, capability, and usability) Understanding virtual communities can give useful insights into the digital economy. Virtual communities are sites where people with similar interests may share information. The community's knowledge base, as well as the members themselves, may be immensely helpful to businesses. The knowledge base, which is often made public through the community's dialogue, gives insight into the members' likes, dislikes, demographics, habits, and issues.

Reputation evolved into a critical metric for determining the reliability of sources. For example, reputation has an impact on and influences customer decision-making. Individuals who are not directly recognized by reputation also have an impact on their decision-making. Indeed, in e-markets, similarly trustworthy persons earned varying trade volumes based on their reputation. Trusting someone demands analyzing them to determine their trustworthiness. One approach of determining trustworthiness is to examine a stranger's intentions. The decision to trust someone might be equated with the conclusion that the individual is trustworthy. In turn, determining whether or not someone is trustworthy is dependent in part on determining that person's intentions based on the information provided.

4. State of the Art

With the rapid growth of the internet and the information era, which opened up new areas of freedom due to the capabilities of the new tools it gives, there is a large amount of misleading content circulating through social networks.

Related work can broadly be divided into the following categories: (Section 4.1) Exploratory analysis of the characteristics of fake news, (Section 4.2) Traditional machine learning-based detection, (Section 4.4), (Section 4.3) deep learning-based detection, (Section 4.4) The effect of digital environments on trustworthiness in information exchange, and (Section 4.5) Social Network-Based Recommendation.

4.1. Exploratory Analysis of the Characteristics of Fake News

Several studies were conducted throughout the years on the features of false news and how to recognize it. Ref. [9] focus on a single type of fake news which is a social media rumor. They describe rumor early detection as identifying a rumor at an early stage before it spreads on social media, so that relevant measures may be taken earlier. Early detection is especially critical for a real-time system since the longer a rumor spreads, the more damage it does and the more likely people are to believe it.

With the advent of social media, all users now have quick access to information. Ref. [1] aimed at developing systems to identify rumors originated as a result of the widespread and quick dissemination of information and the lack of ways to assure the veracity of such information.

4.2. Traditional Machine Learning Based Detection

The rapid dissemination of disinformation is a developing global problem because it has the potential to significantly alter individual reputation and social behavior. The implications of unregulated disinformation dissemination might range from political to financial, but they can also have a long-term impact on world opinion. Ref. [18] offer a novel technique of multilevel, multiclass false news detection based on dataset relabeling and iterative learning, they propose a multilevel supervised learning-based false news detecting technique. They achieved 66.29% accuracy. The input dataset is made up of multilabels and speaker profiles (name, designation, party affiliation, credit history, etc.). Machine learning techniques are used to train multilevel models, which are then evaluated with SVM and decision tree classifiers. Fake news is misinformation presented as a credible news story and used to deceive people's views. Ref. [19] describe a very basic method for encoding texts and how the existence of words in general influences the categorization of texts as authentic or fraudulent. They investigate the topic of natural language processing, which is the wide study of how computers and machines can interpret human-to-human communication and how machines assess texts based on contextual information. They achieved 94.88% accuracy. One of the primary elements encouraging the change of characteristics in modern fake news is the rise of communications mediated by social media. Ref. [5] define the characteristics, and the method of propagating fake news were discussed. They spoke about the usual techniques of spotting bogus news. According to the literature, Natural Language Processing (NLP) was used to detect fake news. Users on social media propagate misinformation quickly and without fact checking. Ref. [20] look at the socially damaging issue of disinformation transmitted via the eyes of the users. They use a language model to categorize people as either false news spreaders or true news checkers. They offer a language model that analyzes user-written social media posts, transforming them into high-dimensional arrays through a transformer-based encoder and max pooling them to generate a user-related high level embedding. They achieved 80.42% in the precision metric and they found that is more suitable than accuracy. The second major contribution of their study is the development and dissemination of a gold standard for recognizing false news spreaders in the context of COVID-19 news.

4.3. Deep Learning-Based Detection

Detecting incorrect language is difficult owing to natural language phenomena such as spelling errors and changes, polysemy, contextual ambiguity, and semantic variants. In this research, [2] present a unique deep learning-based approach named "Convolutional Bi-Directional LSTM (C-BiLSTM)" for automatically detecting such unsuitable language. They particularly interested in resolving this issue in two application scenarios: (a) query completion recommendations in search engines and (b) user chats in messengers. They achieved 93.50% in the precision. Ref. [11] propose a hybrid deep model to represent text semantics of information with context and capture sentiment semantics characteristics for false information detection in this study. Finally, they apply the model to a benchmark dataset and a Weibo dataset, demonstrating that the model performs well. They offer a unique technique for classifying false information for evaluating information trustworthiness on social media. They obtained 43.30% accuracy.

4.4. The Effect of Digital Environments on Trustworthiness in Information Exchange

Reputation has evolved into a critical criterion [21] for determining the reliability of sources. Ref. [22] examined how reputation within a virtual environment affects fairness in material allocations and trust in information exchange. These findings suggest that

reputational effects increase fairness and trust even in a noisy, ambiguous and uncertain environment, but this effect is modulated by age and gender. Ref. [23] Reputation systems are widely used in a high number of web-based services to enhance cooperation among users, as well as to ensure they function well. Table 1 summarizes the reviewed contributions regarding the reputation in digital environments.

Table 1. Summarization of contributions under the reputation concept.

Contribution	Definition of Problem Handled
[23]	Good reputation subjects tend to attract other positive. The reputation inertia effect feedback, regardless of personal/actual experience. Conversely, bad reputation similarity subjects are more likely to attract negative evaluations.
[22]	Fairness and trust are two important aspects Reputation Effects on prosocial behaviors in the virtual environment of social interactions.
[21]	There is evidence that on the internet we tend to trust strangers more than we reasonably should because we implicitly represent/treat them as having a good reputation

4.5. Social Network-Based Recommendation

When contemplating how to improve recommendation systems, it is critical to keep social media in mind. Social networks, in particular, are an instantiation of the new social network-based recommendation techniques, taking into account the vast amount of information and interactions. The word “SRS” refers to the initial name for customized recommendation approaches. Ref. [6] propose a model-based approach recommender system based on ratings, reviews, and social relationships to address the role of social networks on recommendation and to increase suggestion accuracy. In general, SRS may be divided into two types: friend- and trust-based recommendation methods as mentioned in Figure 2.

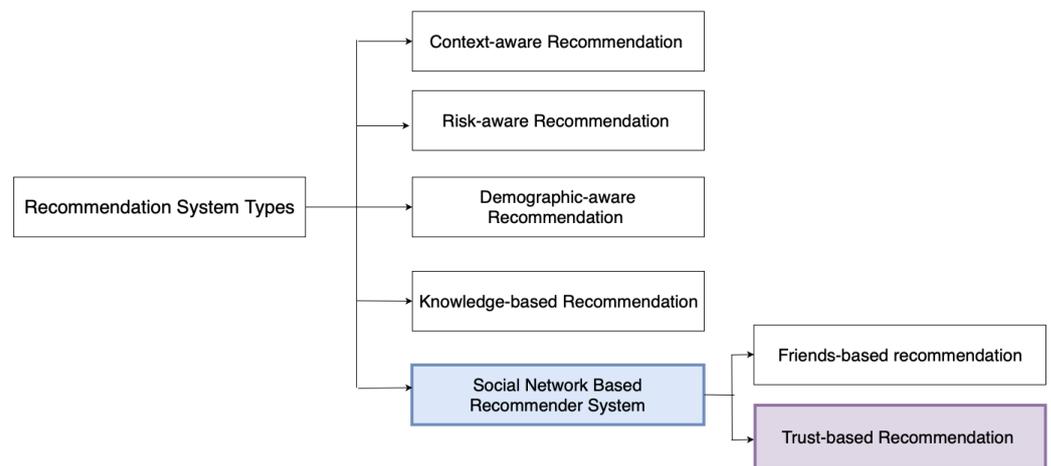


Figure 2. Modern Recommendation Approaches.

4.5.1. Friend-Based Recommendation

The concept of SRS is based on social friendship relationships among network members [24], in which the relationship is exchanged among the participants in the social network. Friendship relationships show how concerned people engage in mutual contact on social networks instead of trusting relationships, which might have a one-sided trust value. Some academics have used collaborative filtering skills to propose things to users, and the benefits of social relationships between users should be included in the recommendation process.

4.5.2. Trust-Based Recommendation

People prefer to obtain information from trustworthy sources such as parents, friends, and relatives in real life. The social network is an important aspect of our daily lives, and it contains a plethora of information that can assist the recommendation system forecast more accurately. The trust connection is used with a trust-aware recommendation system to mitigate the collaborative filter approach's limitations, such as sparsity and cold start. More specifically, the trust viewpoint in recommendation systems plays a significant role in overcoming various limits and problems. There are two distinct applications of trust in the recommendation domain:

- **Trust in user relationships**
Existing recommender systems still face a significant problem due to a lack of explanation or inaccurate recommendation results. As a result, adopting a reliable recommender system becomes crucial. Ref. [25] provides a systemic summary of three categories of trust-aware recommender systems: social-aware recommender systems that leverage users' social relationships; robust recommender systems that filter untruthful noises (e.g., spammers and fake information) enhance attack resistance; and explainable recommender systems that provide explanations of recommended items. They describe how deep learning methods work for the trust-aware recommendation in representation, predictive, and generative learning. Recommendation systems, also known as recommender systems, are one of the most popular topics these days since they are frequently used to anticipate an item for the end-user based on their preferences. The goal of [26] is to gather proof that using social network information between users can improve the quality of standard recommendation systems. They describe the recommendation system approaches and the role of the trust relationship in a social network to overcome the limitation of the traditional approaches. A trust-aware recommendation system provides active users with the flavor he/she like based on his/her direct or indirect trust sources. At the first stage, these words start with the traditional recommendation system, move to the new modern approaches, and focus on the trust recommendation system that has more attention in the current stage. Social media news consumption is growing increasingly common these days. Users gain from social media because of its inherent characteristics of rapid transmission, low cost, and ease of access. Because user engagements on social media can aid in the detection of fake news, Ref. [27] explore the relationship between user profiles and fake/real news. They create real-world datasets that measure users' trust in fake news and choose representative groups of "experienced" users who can spot false news and "naive" users who are more prone to believe it. They perform a comparative analysis of explicit and implicit profile features between these user groups, revealing their potential to differentiate fake news. Recommender systems are responsible for providing the users with a series of personalized suggestions for specific items. Trust is a concept that recently takes much attention and was considered in online social networks.
- **Trust in the systems recommendations**
The authors in [28] provide a recommended reliability quality measure, present an RPI for reliability quality prediction, and an RQR for reliability quality recommendation (RRI). Both quality metrics are predicated on the idea that the stronger a dependability measure is, the better the accuracy findings will be. Users of the internet can share experiences and knowledge and connect with other users using social networks. The activities of these users center on social network services, and they create collaborative content via social networks. Many recent studies on CRMs capable of searching and recommending accurate and necessary information for users amid the sea of virtually infinite information generated were conducted. However, traditional CRMs fail to reflect interactions among users. They also are not capable of reflecting the status or reputation of users in a social network service. Consequently, we proposed a new CRM that is capable of reflecting the status of the content creators, thus overcoming problems associated with the traditional recommendation methods.

4.6. Our Proposed Work

In this research study, we use the third definition of trust in our trust-aware recommender system. We are concentrating on the trust in system recommendations. Many researchers introduce the trust relationship to enhance the traditional recommendation system such as, we define trust as indicating the quality of the system's recommendations.

Table 2 compares our work with the previous benchmark-based studies along three themes: (1) classic detection, (2) new different detection, and (3) trust in social networks. The notation ✓ indicates the use of the method, while the other notation demonstrate the absence of the method in the study. Classic detection theme means the traditional way used for classification. New different detection theme means the use of a new way for classification like the use of vectorization used in our proposed approach. Trust in social network theme means if the fake news detection method used establish the trust and transparency in social networks.

Table 2. Comparison between our benchmark study and prior benchmark studies.

Contribution	Fake News Detection		
	Classic Detection	New Different Detection	Trust in Social Networks
[29]	✓	×	×
[30]	×	✓	✓
[5]	×	✓	✓
[31]	×	✓	✓
[11]	✓	×	×
[9]	✓	×	✓
[2]	✓	×	×
[10]	×	✓	✓
[27]	×	✓	✓
[32]	✓	×	×
Our Approach	✓	✓	✓

5. Methodology

5.1. Aim of the Study

It is necessary to specify the meaning and type of information to classify an item as «fake» but it is better to specify also the form of fake news. As we have already indicated, the types of fake news vary from low to high level. Any news analysis must be based on a formal classification of incorrect information (propaganda, satirical information, manipulation, and manufacturing). However, our interest is not to find a general classification procedure but rather to build an automatic algorithm that will multiclassify any news with the specific percentage of each form and recommend to users an alerting about truthful content. There are three keys challenges to achieve the goal of our proposed approach:

1. Difference between fake news forms
2. Multiclass classification using unlabeled data:
 - Label estimation based on similarity between unlabeled item and the whole labeled dataset to enhance and improve the self-training algorithm.
 - Comparison between new estimated labels using similarity and new labels predicted by the voting majority.
3. Social network recommendation system
Nevertheless, there is an important lack of research into reliability and trust in news dispatched on social networks. Trust in social networks is a crucial part of our daily lives, and we should concentrate more on the trust in system recommendations.

5.2. Methodology and Approach

In this section, the architectures of the proposed approach we evaluated are described. The goal of our experiments is to detect fake news and demonstrate how this detection can contribute to the improvement of trust and transparency in social network recommendation systems. Figure 3 demonstrate the workflow of the proposed approach. The model consists of 6 steps:

1. Step 1: Vectorization
2. Step 2: Trust Network Construction
3. Step 3: Initial Prediction
4. Step 4: Trust Network Reconstruction
5. Step 5: Final Prediction
6. Step 6: Recommendation

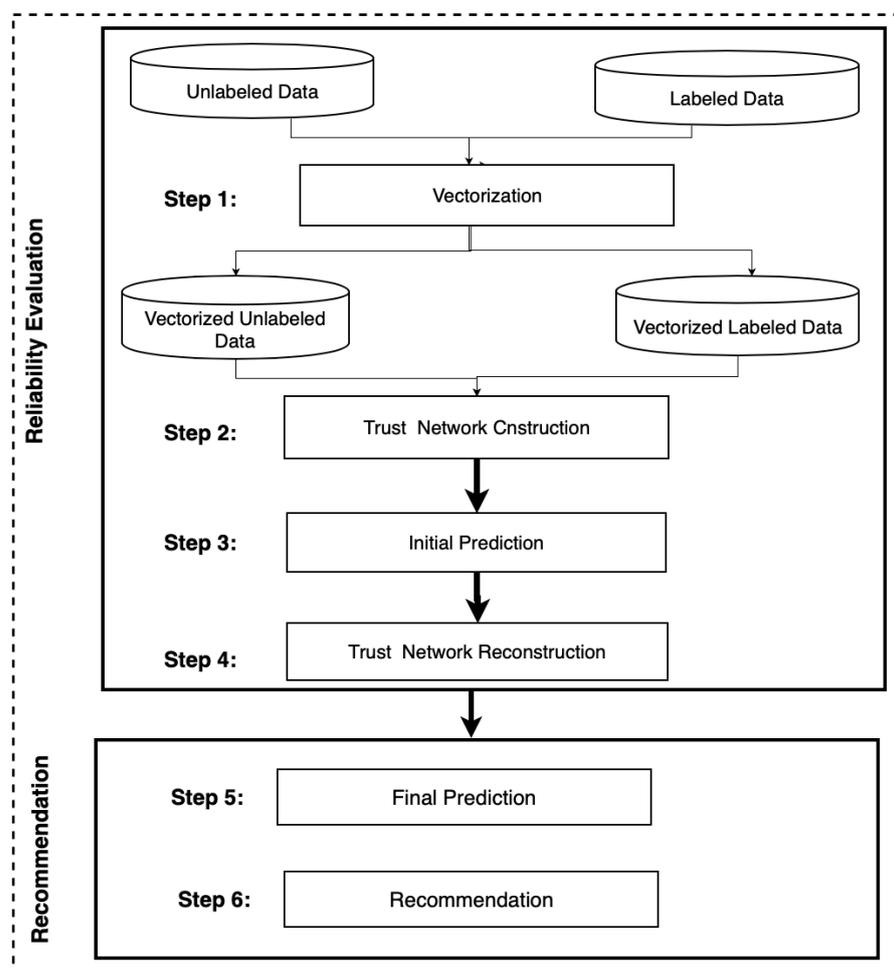


Figure 3. Overview of proposed method.

5.2.1. Vectorization

The vectorization step is used for labeled and unlabeled data. Each news text is vectorized into vectors. We use the modern contextual approach named Bert as embedding technique because it present several advantages than others. Table 3 show the comparison between all contextual approaches.

Table 3. Comparison between contextual approach.

	Free Contextual Approach	Modern Contextual Approach	
	Word2Vec	ULMFiT	BERT
Description	Word2Vec, a neural network-based approach for learning word embeddings utilized in different NLP applications. The basic principle underlying word2vec is that the vectors created are learned by comprehending the words context.	ULMFiT, which stands for Universal Language Model Finetuning, is a transfer learning method that uses a standard 3-layer LSTM architecture for pretraining and fine-tuning tasks and Language modeling (LM) as the source task due to its ability to capture general language features and provide a large amount of data that can be fed to other downstream NLP tasks.	BERT is another language representation learning approach that encodes context using attention transformers rather than bidirectional LSTMs.
Why We Use	-	-	<ul style="list-style-type: none"> • BERT models generate embeddings that are context-dependent. • Truly Bidirectional: BERT is deeply bidirectional due to its novel masked language modeling technique.
Why We Not Use	Word2Vec models generate embeddings that are context-independent	ULMFiT uses a concatenation of right-to-left and left-to-right LSTMs and ULMFiT uses a unidirectional LSTM.	-

5.2.2. Trust Network Construction

Once the vectorization of the labeled and unlabeled data is done successfully, we proceed to a similarity study between each unlabeled data and the labeled data set. A trust network can be constructed based on the similarity recommendation. Each unlabeled text is predicted by calculating the similarity between this text and the whole labeled data. The max value found represents the similarity percentage which mean the specified unlabeled item will has the same label as the item who has the max value of similarity. The predicted label should be the same as the text having the max value of similarity. After calculating the similarity weights between unlabeled news and all labeled news data. Finally we will have a set of data that are labeled using similarity processes, i.e., we construct the initial trust network with those that have a higher similarity score

5.2.3. Initial Prediction

The trust network can predict the initial predictions of the unsimilar news by using the voting majority recommendation. Firstly, the model is trained with the labeled data (the result from the trust network construction). We use this model to predict labels of unlabeled data using voting majority (the most frequent values of all five algorithms used).

5.2.4. Trust Network Reconstruction

In other words, this step attempts to form a new trust network with higher quality than the initial one. This step is necessary to improve the accuracy of the predictions. The voting majority (the most frequent values of all five methods utilized) is required to rejoin the dependable network, which improves the self-training algorithm.

5.2.5. Final Prediction

The recommendation phase consists of two steps, including final prediction and recommendation. The final prediction represents the predicted classes of the given news

text. Self-training is an efficient learning method for learning with few labeled data and many unlabeled data. The weak initial classifier can introduce incorrectly mislabeled data which is then used to train the final classifier. The result is a drop in the precision of the semisupervised self-training classifier. We propose a novel approach to ensure the new labeled data will be used to train the final classifier.

5.2.6. Recommendation

This recommendation step is the last in our proposed approach. In the Final prediction step, the final label of the unlabeled news is calculated based on the newly constructed trust network. Furthermore, in the recommendation step, the algorithm predicts the rates of the unlabeled news and then recommends to the user on the social network a list of types of news.

6. Experiments and Results

6.1. Data Collection Process

Because the term fake news encompasses a wide range of subcategories, there are few common benchmark datasets for the field of false news detection. We used two kinds of fake news identification datasets. The first one contains three types: satire, propaganda and manufacturing [33]. The second one [34] contains a manipulation or bias form. We group both data into one in a balanced manner. Each article contains the title text and label. The corpus has 443 news of each label satire propaganda manufacturing and manipulation.

6.2. Dataset

The absence of manually labeled fake news datasets is a barrier to advance computationally expensive, text-based models that cover a wide variety of topics. For our purpose, we need a set of news articles that are directly classified into news types (satire, propaganda, manufacturing, and manipulation). We researched the available datasets containing these news categories. We found two datasets, combining them to create full sets with the four categories for multiclass classification.

6.3. Data Preprocessing

Preprocessing data is a typical first step that precedes training, and evaluating data using machine learning algorithms can be just as effective as the data you feed. It is crucial that the data is appropriately formatted and those meaningful elements are integrated to be reasonably accurate to result in the best possible outcomes. We applied data preprocessing steps to our existing data to reduce the scale of the real data. News texts posted on social networks are an unstructured source of data that contains noisy information. The raw text must be preprocessed before the functionality of the templates is disabled. There are different ways of converting the text into a modeling-ready form. Next, we omitted the punctuation. Next, in the lower cases in the text, we translated the capital letter detail. When used as text classification features, we eliminated the stop words that are meaningless in a language and produce noise. The next move is to turn the phrases into their original form. We applied the preprocessing data measures to the news on social networks to decrease the real data size. Raw news texts are an unstructured source of information and might contain noisy content. Until removing the functionality for the models, the raw text has to be preprocessed. There are various means of transforming the text into a shape ready for modeling. Table 4 depicts the transformation of the raw dataset into a usable format utilizing the processes previously outlined.

6.4. Methodology

This section discusses the machine learning classifier we used to determine our goals.

Logistic Regression Classifier: the logistic regression approach is the first strategy we are employing to ensure that this model works properly. Logistic regression in machine learning states that it can identify a relationship between the highlights (probability) and

likelihood (outcome) of a certain result. The simplicity of use of logistic regression demonstrates its selection. It is easy to import it from the sklearn linear model.

Naïve Bayes Classifier: Naive Bayes classifiers are a type of basic machine learning classifier in machine learning. Using multinomial NB and pipelining ideas, Naive Bayes is a common technique for determining if news is true or false [35].

Decision Tree Classifier: this classifier is one of the greatest in machine learning. It operates in a model-like manner. Decision trees produce good outcomes. A decision tree may also be useful for identifying fake news [36].

Linear SVM: is used interchangeably with a support vector network (SVN). SVMs are trained using data that were previously divided into two groups.

All these algorithms were designed for binary classification and do not natively support classification tasks with more than two classes. As long as our study does multiclass classification, i.e., classification with more than two classes. We adopt the One-vs-Rest strategy which divides a multiclass classification into a binary classification problem per class. It is a heuristic method to use binary classification algorithms for multiclass classification. It consists in dividing the multiclass dataset into several binary classification problems. Given their reliable prediction results on binary classification we based our choice to use these classifiers.

Table 4. Transition steps in data preprocessing.

Before Preprocessing	After Preprocessing
<p>Red State : News Sunday reported this morning that Anthony Weiner is cooperating with the FBI, which has re-opened (yes, lefties: “re-opened”) the investigation into Hillary Clinton’s classified emails. Watch as Chris Wallace reports the breaking news during the panel segment near the end of the show: the news is breaking while we’re on the air. Our colleague Bret Baier has just sent us an e-mail saying he has two sources who say that Anthony Weiner, who also had co-ownership of that laptop with his estranged wife Huma Abedin, is cooperating with the FBI investigation, had given them the laptop, so therefore they didn’t need a warrant to get in to see the contents of said laptop. Pretty interesting development of federal investigations will often cooperate, hoping that they will get consideration from a judge at sentencing. Given Weiner’s well-known penchant for lying, it’s hard to believe that a prosecutor would give Weiner a deal based on an agreement to testify, unless his testimony were very strongly corroborated by hard evidence. But cooperation can take many forms—and, as Wallace indicated on this morning’s show, one of those forms could be signing a consent form to allow the contents of devices that they could probably get a warrant for anyway. We’ll see if Weiner’s cooperation extends beyond that. More Related.</p>	<p>state news Sunday reported this morning that Anthony Weiner is cooperating with which lefties opened investigation into Hillary Clinton classified emails. Watch Chris Wallace reports breaking news during panel segment near show news breaking while colleague Bret Baier just sent mail saying sources that Anthony Weiner also ownership that laptop with estranged wife Huma Abedin cooperating with investigation given them laptop therefore they didn’t need warrant contents said laptop pretty interesting development targets federal investigations will often cooperate hoping that they will consideration from judge sentencing given Weiner well known penchant lying hard believe that prosecutor would give Weiner deal based agreement testify unless testimony were very strongly corroborated hard evidence cooperation take many forms Wallace indicated this morning show those forms could signing consent form allow contents devices that they could probably warrant anyway Weiner cooperation extends beyond that more related</p>

6.5. Experiment Results

6.5.1. Compared Fake News Detection Results

Table 5 compares the results of our suggested technique, with those of other approaches.

Table 5. Performance comparison for fake news detection.

Metric	[29]	[20]	[11]	Our Approach
Accuracy	0.96	-	0.33	0.96
Precision	0.98	0.80	-	0.96
Recall	0.98	0.81	0.59	0.96
F1	0.98	0.80	0.43	0.96

6.5.2. Experimental Results

Table 6 shows the results for Trust Network Construction. The reader can observe that the first step has remarkable results as the beginning. Table 7 describes the results for Trust Network Reconstruction. The reader can observe that the second step has remarkable improvements when compared with the precedent step. Table 8 mentions the results of the recommendation phase. The authors can conclude that the performance of all recommendations is improved from the initial to the last given results.

Table 6. Experimental results for “Trust Network Construction”.

Different Models	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.76	0.76	0.76	0.76
Naïve Bayes	0.44	0.69	0.44	0.39
Decision Tree	0.41	0.35	0.41	0.35
Linear SVM	0.73	0.74	0.73	0.73

Table 7. Experimental results for “Trust Network Reconstruction”.

Different Models	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.93	0.93	0.93	0.91
Naïve Bayes	0.71	0.73	0.71	0.66
Decision Tree	0.74	0.62	0.74	0.67
Linear SVM	0.94	0.94	0.94	0.91

Table 8. Experimental results for recommendation phase.

Different Models	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.96	0.96	0.96	0.96
Naïve Bayes	0.64	0.67	0.67	0.55
Decision Tree	0.64	0.60	0.64	0.60
Linear SVM	0.95	0.95	0.95	0.94

7. Discussion and Limitations

The fundamental idea behind this study is to analyze potential fake news published in social networks rather than detecting them and classifying each one. As a result, our proposed approach looks for a suggestion model that fits three key features:

- Difference between fake news forms posted on social networks.
- Multiclass classification using unlabeled data.
- Improving the trust and transparency in the social network recommendation system.

The key obstacle in social networks is the growth spread of misinformation, and users need assistance to make decisions on which information to read. In other words, they need truthful content. As a result, our approach intends to gather between detecting fake news forms and trust in the social network recommendation system to increase suggestion quality and RS accuracy. Our methodology is tested on both datasets [33,34].

The proposed technique has some limitations that can be addressed in the future work. The suggested method does not take into account the relationship between users in social networks. The correlation between users and their shared news stories as feature engineering can help determine who can share fake news and then deduce the credible user who share the real information which will improve trusting in friendship inside social networks. The suggested method may also be modified to incorporate sophisticated deep learning techniques such as convolutional neural networks, LSTM. The suggested system is currently a sequential pipeline, with news passing through each stage one by one.

8. Conclusions

In our everyday lives, social networks are one of the most significant sources of information. It is a network of nodes that are linked together to communicate information. In reality, most individuals prefer to obtain information from reliable sources. Trust is a notion that has lately received much attention and was taken into account in online social networks. In this research work, we provide an overview of approaches used to detect fake news inside social networks, for that we present a multiclass, semisupervised strategy for self-training based on a small set of classified data and a large amount of unlabeled data for improving the accuracy of trust-aware recommender systems, and we demonstrate how this detection can help in improving and enhancing the quality of trust and transparency in the social network recommendation system. The proposed semisupervised method efficiency was tested on two benchmark datasets regarding the precision of classification using the most commonly available simple learners logistic regression, decision tree, naive Bayes, and linear SVM. Furthermore, the research studied and compared the four approaches accuracies. Logistic regression is the model that achieves the best accuracy, with a score of 96%. In our model, Bert is used to transform review texts into vectors, and the Logistic Regression model is the suitable one to predict ratings among others. Our numerical findings confirm the effectiveness and robustness of the proposed approach. Therefore, the approach contributes to more effective, reliable, and robust predictive models of multiclass fake news classification.

Author Contributions: Conceptualization, O.S. and O.B.; Formal analysis, O.S.; Funding acquisition, O.S.; Investigation, O.S.; Methodology, O.S.; Project administration, S.K. and O.B.; Supervision, S.K. and O.B.; Validation, S.K. and O.B.; Visualization, O.S., S.K. and O.B.; Writing—original draft, O.S.; Writing—review & editing, O.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not Applicable, the study does not report any data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Alzanin, S.M.; Azmi, A.M. Detecting rumors in social media: A survey. *Procedia Comput. Sci.* **2018**, *142*, 294–300. [CrossRef]
2. Yenala, H.; Jhanwar, A.; Chinnakotla, M.K.; Goyal, J. Deep learning for detecting inappropriate content in text. *Int. J. Data Sci. Anal.* **2018**, *6*, 273–286. [CrossRef]
3. Oumaima, S.; Soulaïmane, K.; Omar, B. Artificial Intelligence in Predicting the Spread of Coronavirus to Ensure Healthy Living for All Age Groups. In *Emerging Trends in ICT for Sustainable Development*; Ahmed, M.B., Mellouli, S., Braganca, L., Abdelhakim, B.A., Bernadetta, K.A., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2021; pp. 11–18.
4. Oumaima, S. How Can We Analyse Emotions on Twitter during an Epidemic Situation? A Features Engineering Approach to Evaluate People's Emotions during The COVID-19 Pandemic. Available online: <https://doi.org/10.17605/OSF.IO/U9H52> (accessed on 23 August 2021).
5. de Oliveira, N.R.; Pisa, P.S.; Lopez, M.A.; de Medeiros, D.S.V.; Mattos, D.M.F. Identifying Fake News on Social Networks Based on Natural Language Processing: Trends and Challenges. *Information* **2021**, *12*, 38. [CrossRef]

6. Ji, Z.; Pi, H.; Wei, W.; Xiong, B.; Wpzniak, M.; Dama, R. Recommendation Based on Review Texts and Social Communities: A Hybrid Model. *IEEE Access* **2019**, *7*, 40416–40427. [[CrossRef](#)]
7. Hassan, T.; McCrickard, D.S. Trust and Trustworthiness in Social Recommender Systems. In Proceedings of the 2019 World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019.
8. Collins, B.; Hoang, D.T.; Nguyen, N.T.; Hwang, D. Trends in combating fake news on social media—A survey. *J. Inf. Telecommun.* **2020**, *1*–20. [[CrossRef](#)]
9. Li, Q.; Zhang, Q.; Si, L.; Liu, Y. Rumor Detection on Social Media: Datasets, Methods and Opportunities. *arXiv* **2019**, arXiv:1911.07199.
10. Heuer, H.; Breiter, A. Trust in news on social media. In Proceedings of the 10th Nordic Conference on Human-Computer Interaction, Oslo, Norway, 29 September–3 October 2018.
11. Wu, L.; Rao, Y.; Yu, H.; Wang, Y.; Nazir, A. False Information Detection on Social Media via a Hybrid Deep Model. In Proceedings of the International Conference on Social Informatics, Saint-Petersburg, Russia, 25–28 September 2018.
12. Imran, M.; Castillo, C.; Diaz, F.; Vieweg, S. Processing Social Media Messages in Mass Emergency: Survey Summary. In Proceedings of the The Web Conference 2018, Lyon, France, 23–27 April 2018.
13. Stitini, O.; Kaloun, S.; Bencharef, O. The Recommendation of a Practical Guide for Doctoral Students Using Recommendation System Algorithms in the Education Field. In *Innovations in Smart Cities Applications*; Ahmed, M.B., RakKaraş, İ., Santos, D., Sergeyeva, O., Boudhir, A.A., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2021; Volume 4, pp. 240–254.
14. Oumaima, S.; Soulimane, K.; Omar, B. Latest Trends in Recommender Systems Applied in the Medical Domain: A Systematic Review. In Proceedings of the 3rd International Conference on Networking, Information Systems & Security, Marrakech, Morocco, 31 March–2 April 2020. [[CrossRef](#)]
15. Tanha, J. A multiclass boosting algorithm to labeled and unlabeled data. *Int. J. Mach. Learn. Cybern.* **2019**, *10*, 3647–3665. [[CrossRef](#)]
16. Martineau, M.; Raveaux, R.; Conte, D.; Venturini, G. Learning error-correcting graph matching with a multiclass neural network. *Pattern Recognit. Lett.* **2020**, *134*, 68–76. [[CrossRef](#)]
17. Kaneko, T.; Sato, I.; Sugiyama, M. Online Multiclass Classification Based on Prediction Margin for Partial Feedback. *arXiv* **2019**, arXiv:1902.01056.
18. Tayyaba, R.; Wasi, H.B.; Arslan, S.; Usman, A.M. Multi-Label Fake News Detection using Multi-layered Supervised Learning. In Proceedings of the 2019 11th International Conference on Computer and Automation Engineering (ICCAE 2019), New York, NY, USA, 23–25 February 2019; pp. 73–77. [[CrossRef](#)]
19. Vijayaraghavan, S.; Wang, Y.; Guo, Z.; Voong, J.; Xu, W.; Nasser, A.; Cai, J.; Li, L.; Vuong, K.; Wadhwa, E. Fake News Detection with Different Models. *arXiv* **2020**, arXiv:2003.04978.
20. Leonardi, S.; Rizzo, G.; Morisio, M. Automated Classification of Fake News Spreaders to Break the Misinformation Chain. *Information* **2021**, *12*, 248. [[CrossRef](#)]
21. Duradoni, M.; Collodi, S.; Perfumi, S.C.; Guazzini, A. Reviewing Stranger on the Internet: The Role of Identifiability through “Reputation” in Online Decision Making. *Future Int.* **2021**, *13*, 110. [[CrossRef](#)]
22. Duradoni, M.; Paolucci, M.; Bagnoli, F.; Guazzini, A. Fairness and Trust in Virtual Environments: The Effects of Reputation. *Future Int.* **2018**, *10*, 50. [[CrossRef](#)]
23. Duradoni, M. Reputation Matters the Most: The Reputation Inertia Effect. *Hum. Behav. Emerg. Technol.* **2020**, *2*, 71–81. [[CrossRef](#)]
24. Gao, P.; Baras, J.; Golbeck, J. Trust-aware Social Recommender System Design. In *Doctor Consortium of 2015 International Conference On Information Systems Security and Privacy*; Science and Technology Publications, Lda: Setúbal, Portugal, 2018.
25. Dong, M.; Yuan, F.; Yao, L.; Wang, X.; Xu, X.; Zhu, L. Trust in Recommender Systems: A Deep Learning Perspective. *arXiv* **2020**, arXiv:2004.03774.
26. Tharwat, M.E.A.A.; Jacob, D.W.; Fudzee, M.F.M.; Kasim, S.; Ramli, A.A.; Lubis, M. The Role of Trust to Enhance the Recommendation System Based on Social Network. *Int. J. Adv. Sci. Eng. Inf. Technol.* **2020**, *10*, 1387–1395. [[CrossRef](#)]
27. Shu, K.; Wang, S.; Liu, H. Understanding User Profiles on Social Media for Fake News Detection. In Proceedings of the 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Miami, FL, USA, 10–12 April 2018; pp. 430–435.
28. Bobadilla, J.; Gutiérrez, A.; Ortega, F.; Zhu, B. Reliability quality measures for recommender systems. *Inf. Sci.* **2018**, *442–443*, 145–157. [[CrossRef](#)]
29. Gereme, F.; Zhu, W.; Ayall, T.; Alemu, D. Combating Fake News in “Low-Resource” Languages: Amharic Fake News Detection Accompanied by Resource Crafting. *Information* **2021**, *12*, 20. [[CrossRef](#)]
30. Kasnesis, P.; Toumanidis, L.; Patrikakis, C.Z. Combating Fake News with Transformers: A Comparative Analysis of Stance Detection and Subjectivity Analysis. *Information* **2021**, *12*, 409. [[CrossRef](#)]
31. Galal, S.; Nagy, N.; El-Sharkawi, M.E. CNMF: A Community-Based Fake News Mitigation Framework. *Information* **2021**, *12*, 376. [[CrossRef](#)]
32. Qian, F.; Gong, C.; Sharma, K.; Liu, Y. Neural user response generator: Fake news detection with collective user intelligence. In Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI’18), Stockholm, Sweden, 13–19 July 2018; pp. 3834–3840. [[CrossRef](#)]
33. Fact Checking. Available online: <https://hrashkin.github.io/factcheck.html> (accessed on 24 February 2021).

34. Getting Real about Fake News. Available online: <https://www.kaggle.com/mrisdal/fake-news/data> (accessed on 24 February 2021).
35. Granik, M.; Mesyura, V. Fake news detection using naive Bayes classifier. In Proceedings of the 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kyiv, Ukraine, 29 May–2 June 2017; pp. 900–903. [[CrossRef](#)]
36. Lyu, S.; Lo, D.C.T. Fake News Detection by Decision Tree. In Proceedings of the 2020 Southeast Con, Raleigh, NC, USA, 28–29 March 2020; pp. 1–2.