


Article

# A Reliable Weighting Scheme for the Aggregation of Crowd Intelligence to Detect Fake News

Franklin Tchakounté <sup>1,\*</sup> , Ahmadou Faissal <sup>1,\*</sup>, Marcellin Atemkeng <sup>2</sup> and Achille Ntyam <sup>1</sup>

<sup>1</sup> Department of Mathematics and Computer Science, Faculty of Science, University of Ngaoundéré, Ngaoundéré 640001, Cameroon; ntyam\_achille@yahoo.fr

<sup>2</sup> Department of Mathematics, Rhodes University, Grahamstown 6140, South Africa; M.Atemkeng@ru.ac.za

\* Correspondence: tchafros@gmail.com (F.T.); faissalahmadou@gmail.com (A.F.)

Received: 5 April 2020; Accepted: 29 May 2020; Published: 12 June 2020



**Abstract:** Social networks play an important role in today's society and in our relationships with others. They give the Internet user the opportunity to play an active role, e.g., one can relay certain information via a blog, a comment, or even a vote. The Internet user has the possibility to share any content at any time. However, some malicious Internet users take advantage of this freedom to share fake news to manipulate or mislead an audience, to invade the privacy of others, and also to harm certain institutions. Fake news seeks to resemble traditional media to establish its credibility with the public. Its seriousness pushes the public to share them. As a result, fake news can spread quickly. This fake news can cause enormous difficulties for users and institutions. Several authors have proposed systems to detect fake news in social networks using crowd signals through the process of crowdsourcing. Unfortunately, these authors do not use the expertise of the crowd and the expertise of a third party in an associative way to make decisions. Crowds are useful in indicating whether or not a story should be fact-checked. This work proposes a new method of binary aggregation of opinions of the crowd and the knowledge of a third-party expert. The aggregator is based on majority voting on the crowd side and weighted averaging on the third-party side. An experimentation has been conducted on 25 posts and 50 voters. A quantitative comparison with the majority vote model reveals that our aggregation model provides slightly better results due to weights assigned to accredited users. A qualitative investigation against existing aggregation models shows that the proposed approach meets the requirements or properties expected of a crowdsourcing system and a voting system.

**Keywords:** crowd; crowdsourcing; fake news; detection; aggregation; third party; majority vote; weighted average

## 1. Introduction

Social networks have an important place in today's society, and in our relationships with others. They allow us to remain available on many boards as with family, friends, customers as well as fans, etc. They are ideal tools to send messages, share ideas to a large community of people. The Internet user has the possibility to share any content at any time. The users are free to exchange information on social media. Social networks give the opportunity to the Internet user to have an active role, e.g., a user can relay certain information via a blog, made comments, or even vote. In 2019, the global social penetration rate reached 45% as reported by Statista [1].

However, some malicious Internet users take advantage of this freedom to share fake news to manipulate or mislead an audience, invade each other's privacy, and also to harm certain institutions. In recent years, fake news has gained popularity in social media [2]. Fake news has been one of the most concerns in socio-political topics in recent years reported by Statista [1]. Websites that deliberately

published hoaxes and misleading information appeared on the Internet and were often shared on social media to increase their reach. As a result, people have become suspicious of the information they read on social media. An investigation of 755 individuals concerning fake news in some African countries in 2016 confirmed the previous assertion [2]. Fake news is spread by social medias and fake news sites, which specialize in creating content that attracts attention and the format of reliable sources, but also by politicians or major media outlets with political agenda. Fake news seeks to resemble traditional media to establish its reliability to the public. Its sobriety leads the public to share it and, therefore, be disseminated quickly [3]. Such fake news can cause rivalries between countries, or lure users into a scam, among others.

The rapid spread and dissemination of fake news on the Internet highlights the need for automatic hoax detection systems. Several authors have proposed to overcome this issue by relying on crowd intelligence [4–8]. These authors do not exploit associatively expertise of the crowd and expertise of a third party to make decisions. The crowd is only useful to indicate whether a story should be checked or not. The crowd includes people found on the Internet and third parties include people with certain responsibilities in some institutions. Third parties belong to a higher level of the hierarchy and has a certain importance to give opinions about a post related to their institution. This raises the fundamental question of how to aggregate the opinions of the crowd with expert knowledge to decide on a post.

The general objective of this work is to propose an approach to detect fake news while combining crowd intelligence and third-party expert knowledge. The specific objectives are the following:

- To understand concepts around crowdsourcing, aggregation models and fake news;
- To design the crowd and the third-party elements;
- To propose a model of aggregation of a crowd and third-party intelligence;
- To experiment with the proposal on a system developed online.

In this work, the model assumes that people in the crowd and in the third-party group do not provide fake opinions. A novel binary aggregation method of crowd opinions and third-party expert knowledge is implemented. The aggregator relies on majority voting on the crowd side and weighted average on the third-party side. A model is proposed to compute voting weight on both sides and a comparison between weights is made to take a final decision. The latter concerns opinion with the higher weight.

The rest of the paper is organized as follows: Section 2 presents various works that are similar to our proposal. Section 3 presents concepts about crowdsourcing and fake news which are relevant to our context. Section 4 describes the proposed model for aggregating users' opinions to signal fake news. This section describes the research design and the methodology phases. Section 5 and 6 describe the experiments on real samples of posts, results and discussions. Section 7 concludes and presents future works for further investigation.

## 2. Related Works

This section relies on a recent and consistent survey of research around fake news [9,10]. The research present in this section is well documented in the literature, but warrant that we present them to ease the reader's understanding and further use throughout the document.

### 2.1. Detection of Fake News

The studies concerning fake news can be split into four directions. The first direction includes analyzing and detecting fake news with fact-checking. The manual fact-checking is achieved by known group of highly credible experts to verify the contents [11,12], crowd-sourced individuals [4,7,8,13,14] and crowd-sourced fact-checking websites such as HoaxSlayer [15] and StopBlaBlaCam [16]. Kim et al. [8] propose a leverage crowd-sourced signals to prioritize fact-checking of news articles by capturing the trade-off between the collection of evidence (flags); the harm caused from more users being exposed

to fake news to determine when the news needs to be verified. Tschatschek et al. [4] develop an approach to effectively use the power of the crowd (flagging activity of users) to detect fake news. Their aim was to select a small subset of news, send it to an expert for review, and then block the news which is labeled as fake by the expert. Sethi et al. [13] propose a prototype system that uses crowd-fed social argumentation to check the validity of proposed alternative facts and help detect fake news. Tacchini et al. [14] propose an approach derived from crowdsourcing algorithms to classify posts as hoax or non-hoax. Shabani et al. [7] propose a hybrid machine crowd approach for fake news detection that uses a machine learning and crowdsourcing method. The automatic fact-checking takes into consideration the volume of information generated in social media. It relies on reliable fact retrieval methods for further processing to deal with redundancy, incompleteness, unreliability, and conflicts [17]. Once data is cleaned, knowledge is extracted under a graph form and fed into an exploitable knowledge base [18].

The second direction concerns identifying quantifiable characteristics or features susceptible to discriminate hoaxes from benign ones [19–21]. In this regard, the authors investigate attribute-based language features and structure-based language features that characterize a post. They built machine-learning-based strategies on structured information to derive classification and regression models [22,23].

The third direction refers to the study of how hoaxes propagate and spread on the social graph. For that, one can make qualitative analysis of patterns to recognize fake news propagation [24–26]. Some proposals consisted of mathematical models for fake news propagation based on epidemic diffusion models [27] and game-theoretical models [28]. The detection based on fake news propagation is achieved using supervised learning on cascade (a tree or tree-like structure representing the propagation of a hoax) features [29]. Another option in this direction aims to use graph kernels to compute the similarity between various cascades of posts [30]. This information is then used as features to supervised learning algorithms to detect hoaxes [31].

The fourth direction consists to detect hoaxes by assessing credibility related to hoax headlines [32], hoax sources [33], hoax comments [34], and hoax spreaders [28]. Since headlines consist of luring people to click on the fake links, Pengnate et al. [32] assess the credibility of news headline to detect clickbait. For that, they analyze linguistic features and non-linguistic features along with supervised learning techniques. The credibility of hoax sources is related to the detection of its web page credibility. This is assessed relying on content and link features, which are used to characterize the web page by exploiting machine learning frameworks [33]. Comments about posts are also exploited to retrieve knowledge about fake news [34]. The authors use three options. The first option aims at assessing comment credibility of a sequence of language and style features from user comments. The second option deals with extracting metadata associated and related to user behavior such as burstiness, activity, timeliness, similarity, and extremity to correlate with comment semantics. The third option uses graph-based models to draw relationships among reviewers, comments, and products. Shu et al. [28] model social media as a tri-relationship network among news publishers, news articles, and news spreaders. This structure is then exploited in entity embedding and representation, relation modeling, and semi-supervised learning to detect fake news.

## 2.2. Limitations

Apart, from the first direction, existing approaches have in common that they directly rely on post features to characterize their nature (fake or not fake). Works within the first option involve the participation of people to give opinions based on experiences. They exploit collective intelligence including third-party individuals or websites to take the final decision about the post and exploit the crowd to indicate whether or not a post should be checked. Fact-checking websites rely on professional journalists who have experience in the field of fact-checking, these journalists make field visits, collect information on the post, process the information before disseminating it on the platform fact-checking. However, this method has three limitations: (i) a huge waste of time in the descent, processing, and

dissemination of information facilitating the rapid spread of fake news; (ii) there are not enough professional journalists to cover the entire region of the country; (iii) some websites or individuals involved may be inaccessible and slow down the decision-making process.

### 2.3. Contribution

In view of these limitations, this research proposes a model that associates expertise of the crowd and expertise of a third party to make decisions, based on majority voting and weighted averaging.

## 3. Background

In this section, we present a review of crowdsourcing and fake news that are relevant to this work.

### 3.1. Crowdsourcing

The crowd refers to people involved in crowdsourcing initiatives. According to Kleemann et al. [35], the crowd including users or consumers, is considered to be the essence of crowdsourcing. According to Schenk et al., Guittard et al. [36,37], the nucleus of the crowd as amateurs (e.g., students, young graduates, scientists or simply individuals), although they do not set aside professionals. Thomas et al. [38], Bouncken et al. and Swieszczak et al. [39,40] identify the crowd as web workers. Ghezzi et al. [41] believe that crowdsourcing systems have as main entry the problem or task that must be solved by the crowd. Depending on the structure of the request established by the crowdsourcer and the typology, depending on the skills required of the crowd, the crowdsourcing context will have different characteristics and involve different processes. Each participant is assigned the same problem to solve and the crowdsourcer chooses a single solution or a subset of solutions. Crowdsourcing is also seen as a business practice that means literally to outsource an activity to the crowd [42]. According to Howe and Jeff [42], crowdsourcing describes a new web-based business model that joins the creative solutions of a distributed network of individuals which leads to an open call for proposals. In other words, a company posts a problem online, a vast number of individuals offer solutions to the problem, the winning ideas are awarded some form of a bounty and the company mass produces the idea for its own gain. According to Lin et al. [43], crowdsourcing can be viewed as a method of distributing work to a large number of workers (the crowd) both inside and outside of an organization, for the purpose of improving decision-making, completing cumbersome tasks, or co-creation of designs and other projects.

### 3.2. Fake News

According to Granskogen [44], false information can be both misconducts, false statements, and simple mistakes made when creating information or reports. These forgeries are more difficult to detect, because normally the whole story is not created as a false story, and it is a mixture of true and false information. This may result from the use of outdated sources, biased sources, or from the simple formulation of hypotheses without verification of the facts. The motivations that drive users to generate and disseminate fake news are sometimes for political benefits or to damage corporate reputations or to increase advertising revenues, and to attract attention. Fake news seek to resemble traditional media to establish credibility with the public. Its serious appearance pushes the public to share it. Fake news can thus spread quickly [3]. Kumar et al. [45] classify misinformation into two categories, one based on intent and the other based on knowledge content :

- *Categorization based on intent:* False information based on the author's intention, can be misinformation and disinformation. Misinformation is the act of disseminating information without the intention of misleading users. These actors may then unintentionally distribute incorrect information to others via blogs, articles, comments, tweets, and so on. Readers may sometimes have different interpretations and perceptions of the same true information, resulting in differences on how they communicate their understanding and, in turn, inform others' perception of the facts. Contrary to misinformation, disinformation is the act of disseminating information

with the intention of misleading users. It is very similar to understanding the reasons for the deception. Depending on the intention of the information or the origin of the information, false information may be grouped into several different groups. Fake news can be divided into three categories: hoaxes, satire, and malicious content. A hoax consists of falsity information to resemble the truth information. These may include events such as rumors, urban legends, pseudo-sciences. They can also be practical jokes, April Fool's joke, etc. According to Marsick et al. [46], the hoax is a type of deliberate construction or falsification in the mainstream or social media. Attempts to deceive audiences masquerade as news and may be picked up and mistakenly validated by traditional news outlets. Hoaxes range from good faith, such as jokes, to malicious and dangerous stories, such as pseudoscience and rumors. A satire is a type of information in which the information is ridiculed. For example, a public person ridiculed in good faith while some of its most prominent parts are the following ones taken out of context and made even more visible. Satire can, like hoaxes, be both good faith and humorous, but it can also be used in a malicious way to lower someone's or something's level. Finally, a malicious content is made with the intention of being destructive. This content is designed to destabilize situations, change public opinions, and use false information to spread a message with the aim of damaging institutions, persons, political opinions, or something similar. All the different types can be malicious if incorrect data is entered in the data element. The intention will be completely different.

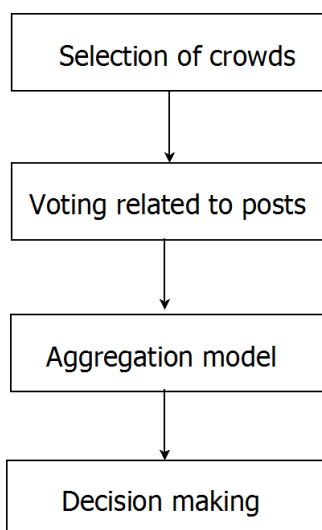
- *Categorization based on knowledge:* According to Kumar et al. [45], knowledge-based misinformation is based on opinions or facts. The opinion-based false information describes cases where there is no absolute fundamental truth and expresses individual opinions. The author, who consciously or unconsciously creates false opinions, aims to influence the opinion or decision of readers. The fact-based false information involves information that contradicts, manufactures, or confuses unique value field truth information. The reason for this type of information is to make it more difficult for the reader to distinguish truth from false information and to make him believe in the false version of the information. This type of misinformation includes fake news, rumors, and fabricated hoaxes.

#### 4. Aggregation Model

The aggregation model consists of adding a binary voting system to the crowdsourcing system and, therefore, the decision is made via the majority vote aggregator, MV and the weighted average, WA. This section presents the research methodology on crowdsourcing, the process by which the crowd was selected, the voting process, and the decision-making on a post.

##### 4.1. Research Methodology

The research methodology includes four layers. The first layer concerns the identification and selecting crowd who give an opinion on a post; the second layer refers to the voting process, the third layer is about to aggregate different opinions to estimate the nature of the post and the last layer takes the aggregated value and gives the final result about the post. Figure 1 presents the methodology adopted in this work. Details of each layer presented in Figure 1 are provided in Section 4.2.



**Figure 1.** Research methodology.

#### 4.2. Selection of Crowd

The process of crowd selection is very important in this model, it is, therefore, necessary to pay attention to this selection to avoid biased and inconsistent results. The selection is made according to the field of the post narrative. That is to say when the post is in the field of education, the potential candidates from the crowd are: students, teachers, school administration, and all the links in the educational chain. If the post is in the health field, the potential candidates from our crowd will be patients, nurses, nurse aides, doctors, specialists, directors and all the links in the health chain. For example, if the post is aimed at a higher education institution, the invitation is issued to the student public, i.e., students will be considered to be simple users and the officials of this institution will be considered to be accredited users. Concerning the simple users, they are considered to be the essence of crowdsourcing, they participated in a voluntary way for the success of the proposed model. We provided them "the Internet connection" to conduct this experiment. The invitation was sent on the student platform of the WhatsApp groups from the different courses of the Department of Mathematics and Computer Science, particularly at the Bachelor 3, Master 1, and Master 2 levels. We had received a total of 85 single users, of which we have 15 female and 70 male users. We had given a guide with instructions on how to connect to the platform, how to use the platform and what should be done by the students.

The instructions given to the participants include logging on the platform and giving their opinions on different posts. Opinions must be binary either "true" or "false". However, participants have the possibility to research the truthfulness of post in case of doubt or in case that they cannot give their opinions immediately.

The observer bias is controlled when selecting the crowd. We had selected students who have a minimum background regarding the posts. We consider that students at Bachelor 3, Master 1 and Master 2 levels have an affordable knowledge of the facts concerning a university institution, on the different problems of the competitive entrance exams, in fact, they are better equipped to give us reliable results unlike students at Bachelor 1 and Bachelor 2 level who have just integrated the university curriculum and therefore likely not to give us reliable results. The crowd is made of two types of users: simple and accredited users.

##### 4.2.1. Simple User

Simple users are people who are considered likely to have a small knowledge base in relation to the facts that are unfolding or the posts that are submitted to them. However, every simple user has a competency attesting that the user has at least a small amount of knowledge about the post and is



assigned a score or weight of 0.25 to this competency. These users are at the last level in the hierarchy shown in Figure 2.

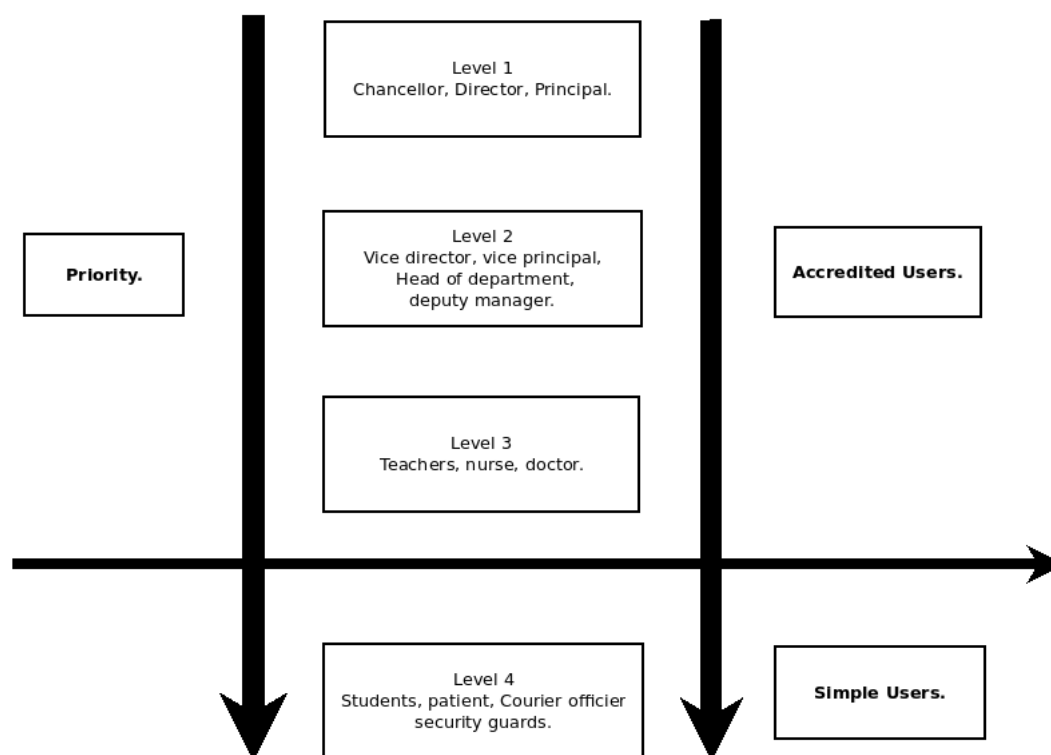


Figure 2. Level of user competence.

#### 4.2.2. Accredited User

Accredited users are people working in an institution whose fake news is likely to reach. These users have a higher degree of veracity than simple users. Each accredited user occupies a position of responsibility in the institution or simply works as a staff member of the institution. The position of responsibility held is a function of the institution's hierarchy, as shown in the university hierarchy as an example in Figure 3. This hierarchy is grouped into three levels:

- **Level 1:** this level concerns users who occupy the highest positions of responsibility in the structure. These users can be Chancellor, Directors, Principals, etc. In this level, we consider that the opinions of these users have a degree of truthfulness above 90% for a weight of 1.
- **Level 2:** this level concerns users who occupy the average positions of responsibility in the structure. These users may be sub directors, vice-principal, deans, head of departments, etc. In this level, we consider that the opinions of these users have a degree of truthfulness of at least 75% for a weight of 0.75.
- **Level 3:** this level concerns users who occupy the weakest positions of responsibility in the structure. These users may be teaching staff, support staff, mail handlers, etc. In this level, we consider that the opinions of these users have a degree of truthfulness of at least 50% for a weight of 0.5. The level of competence of these users is shown in the Figure 2.
- **Justification of degree of truthfulness** We assign degree of truthfulness different from one because we consider external influences in giving opinions. All the three degrees (0.90 in level 1, 0.75 in level 2 and 0.5 in level 3) respect this consideration. Additionally, people in a level from highest to lowest is assigned a degree of truthfulness to reveal difference in terms of credibility related to information. This fact is because people in the highest level is likely to be quickly

informed about a situation in the institution, i.e., the director is the one to be contacted in prior, about a situation in his institution.

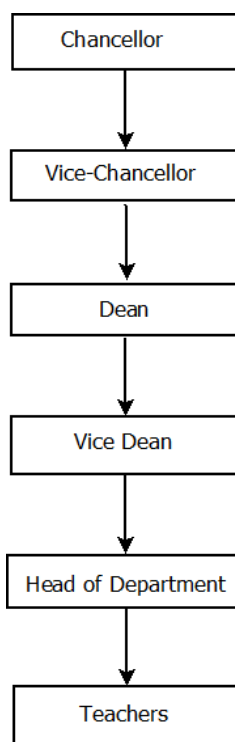


Figure 3. University hierarchy.

Given the difficulty with accredited users, we have simulated these users with all possible cases. The simulation was done as follows: We had 15 accredited users spread over five different institutions. For each institution, we had taken 3 accredited users (one from level 1 with a weight of  $w = 1$ , one from level 2 with a weight of  $w = 0.75$  and one from level 3 with a weight of  $w = 0.5$ ), we had simulated each user with the two possible responses, i.e., “true” and “false”. This means 30 possible responses for whole institutions. In general, for each institution, if we have  $n$  accredited users, we will have a total of  $2^n$  possible responses. In our particular case, each institution has three accredited users, so 8 possibilities of responses, i.e.,  $2^3$ . The possible responses for each institution are shown in Table 1.

Table 1. Simulation Accredited Users.

A1	A2	A3
0	0	0
0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0
1	1	1

Response 1 means true and 0 means false. We recall that the accredited users of an institution can be simple users in another institution.

#### 4.3. Voting Related to Post

This section is about the post-vote where each selected user is asked to give his opinion about a post. It is about voting by True or False just by submitting on the platform radio button available at



hoax.smartedubizness.com [47] that created in order to collect user votes. Both types of users submit the same post for binary aggregation. To vote, each user must first create an account in the platform and specify the type of user he is. Then the authentication page is submitted to them and the post will be displayed after this authentication. This page contains several posts and each user is asked to give his opinion for each post and submit at the end by the 'Save' button of the online platform. When the user is confused about a post, he has the possibility to search for objective opinions on the Internet.

#### 4.4. Aggregation Model

When both types of users submit their votes, through the crowdsourcing platform (available at hoax.smartedubizness.com), it is a matter of gathering all the opinions and aggregating them to make a decision on the post.

##### 4.4.1. Formalization

Consider  $V_S$  the set of simple voters,  $V_a$  the set of accredited voters and  $X$  the set of possible choices reduced into two distinct options:  $X = \{-1, 1\}$ . The corresponding formalization for this situation is that one of the two options is good (true) and the other bad (false), but individual judgments may differ simply because some individuals may make mistakes. The problem is how to aggregate the information? Let  $W_j$  be the competence of an accredited voter  $j$ ,  $W_j \in [\frac{1}{4}, 1]$ . All the simple users have the last level competence  $\frac{1}{4}$ . Assume there are  $M$  workers and  $N$  tasks with binary labels  $\{-1, 1\}$ . Denote by  $z_i \in \{-1, 1\}$ ,  $i \in [N]$  the true label of task  $i$ , where  $[N]$  represents the set of first  $N$  integers;  $N_j$  is the set of tasks labeled by worker  $j$ , and  $M_i$  the workers labeling task  $i$ . The labeling results form a matrix  $L \in \{-1; 1\}^{N \times M}$ , where  $L_{ij} = 1$  denotes the true answer if worker  $j$  labels task  $i$ , and  $L_{ij} = -1$  if otherwise. The goal is to find an optimal estimator  $\tilde{z}$  of the true labels  $z$  given the observation  $L$ . Thus, the matrix  $L \in \{-1; 1\}^{N \times M}$  can be described as:

$$L_{ij} = \begin{cases} 1; & \text{if the users provide a True opinion,} \\ -1; & \text{if the users provide a False opinion.} \end{cases} \quad (1)$$

A naive approach to identify the correct answer from multiple users' responses who have the same competence is to use majority voting. Majority voting simply chooses what most users agree on. This majority vote will only be cast on users with the same competence or on the level among the levels listed in Section 4.2.2. When users do not have the same competence or are not at the same levels, we aggregate these opinions using the weighted average WA. As its name suggests, the majority voting MV method consists of returning the majority of user opinions. Its mathematical formalization is as follows:

$$Z_{i\text{majority}} = \sum_{i \in N} L_{ij}. \quad (2)$$

The simple majority rule: in this rule, a proposal is validated if more than half of the users vote "true" or "false". More formally when we have  $N$  members, and we have two choices  $X$  and  $Y$ , the proposal is validated if and only if the number  $Y$  of "true" satisfies  $Y \geq (N + 1)/2$ . According to [48], the set of voters together with a set of rules constitute a 'voting system'. A convenient mathematical way to formalize the set of rules is to single out which sets of voters can force an affirmative decision by their positive votes. The weighted average is the average of several values assigned coefficients. These coefficients are the skills or competence associated with each user and the assigned values are the opinions of the users. The weighted average of user opinions is expressed mathematically as follows:

$$\tilde{w} = \frac{\sum_j W_j L_{ij}}{\sum W_j}, \quad (3)$$

where  $W_j$  is the competence of users and  $L_{ij}$  the opinion matrix. As there are two types of users, aggregation is done on both sides.

#### 4.4.2. Simple Users Side

When simple users submit their votes, either 1 for true or  $-1$  for false, we store all their opinions in a database, then we apply the aggregator “Majority Vote” to estimate the opinion of simple users. Then we calculate the weight of users who have allowed to have most votes. When there is equality, the model does not decide directly but refers to the opinion of accredited users to give the final opinion. Simple users are weighted to participate in decision-making, we do not prejudice the opinions of simple users, we take them into account in decision-making. If we do not weight these simple users, we will prejudice their opinions in favor of the opinions of the more advanced users.

#### 4.4.3. Accredited Users Side

In this work, an expert is an accredited user who belongs to an institution with a certain level of responsibility. It is a person we know and can trace in an institution. However, it does not mean that his opinion is 100% reliable. If we assign a degree of reliability to the expert opinion at 100% it does mean that we have neglected the external influence that can occur on the expert opinion. We assign a weight of  $w = 1$  to the experts but with a degree of 90%, and we take into account that the expert opinion can be influenced by an external factor. For example, an expert (accredited user) holding a higher position of responsibility in an institution may have his opinion influenced just to keep his position of responsibility, or his opinion may be politicized just because he is a member of some political party and is defended by that party in order to keep his position. Each accredited user is at a skill level of either (L1), (L2) or (L3). When all accredited users submit their opinions through the platform, their opinions are stored in the database and then the aggregator “Weighted Average” is applied to estimate the final opinion of the accredited users.

The weighted average is the sum of the product of the competence and the opinion voted by the user, then normalized to the sum of the competences of all accredited users who voted. The weighted average can be positive or negative in three cases : (i) if the weighted average returns a positive value then the estimated opinion is 1 (true); (ii) otherwise the opinion is  $-1$  (false); (iii) if the weighted average returns 0, then the decision is not taken immediately, the model refers to the opinion of simple users.

#### 4.4.4. Computing Weight Model

The weighting allows the opinions of all users to be taken into account without prejudicing the opinions of others. Weighting is of paramount importance because opinions of accredited users may differ in some cases or experts may have different opinions on a position and this complicates decision-making. We have therefore chosen to use the concept of “weighting” to weight these opinions to make a final decision without prejudicing opinions of all parties.

The hierarchy has 4 levels. In reality, we have 5 levels but the last level is a level where we do not take into account the opinion of these users and their weight is  $w = 0$ . It is a level where the opinion is not taken into account, that is why we do not count this level in the hierarchy of levels. Knowing that the weight,  $W_j \in [0; 1]$  and as we have 4 levels, we divide this interval into 4 and we obtain an amplitude of  $A_{amp} = 0.25$ . In general, for a given interval  $[a;b]$ ,  $W_j$  is calculated as:

$$W_j = b - j \times A_{amp}, \quad (4)$$

where  $j \in [0; b - 1]$  is the position of the level. We thus obtain the weights as a function of the levels respectively 1, 0.75, 0.50, 0.25. The calculation of the user weight is done for the purpose of deciding when the opinion of simple users differs from that of accredited users. For this, we perform the following:

- When the opinion of simple users gives 1 (true) after aggregation, we sum the weights of all users who gave their opinion 1 (true). Similarly, when the opinion gives  $-1$  (false), we sum the weights of users who gave their opinion  $-1$  (false).

- For accredited users, when the weighted average gives a positive result, the opinion will be 1 (true) and we sum the weights of the users who gave their opinion 1. On the other hand, when the weighted average gives a negative result, the opinion will be -1 (false) and we sum the weights of users who gave their opinion -1 (false).

The total weight  $P_i$  for the post  $i$  is calculated using the following mathematical formula:

$$P_i = \sum_{\substack{j \in \Theta \cup \Gamma \\ 1 \leq i \leq N}} W_j L_{ij}, \tag{5}$$

where  $W_j$  is the weight assigned by user  $j$  for the post  $i$ ,  $\Theta$  is the set of opinions "true" and  $\Gamma$  is the set of opinions "false".

The proposed model is depicted in Figure 4. This model categorizes users into two types: simple and accredited. Each type of user gives his or her opinion to aggregate all user opinions. The majority vote aggregator is applied on the simple users's side and the weighted average on the accredited users's side when there is no equality of opinions in each type of user. Then a comparison of opinions between simple and accredited users and a weighting calculation is made to make a final decision on a post.

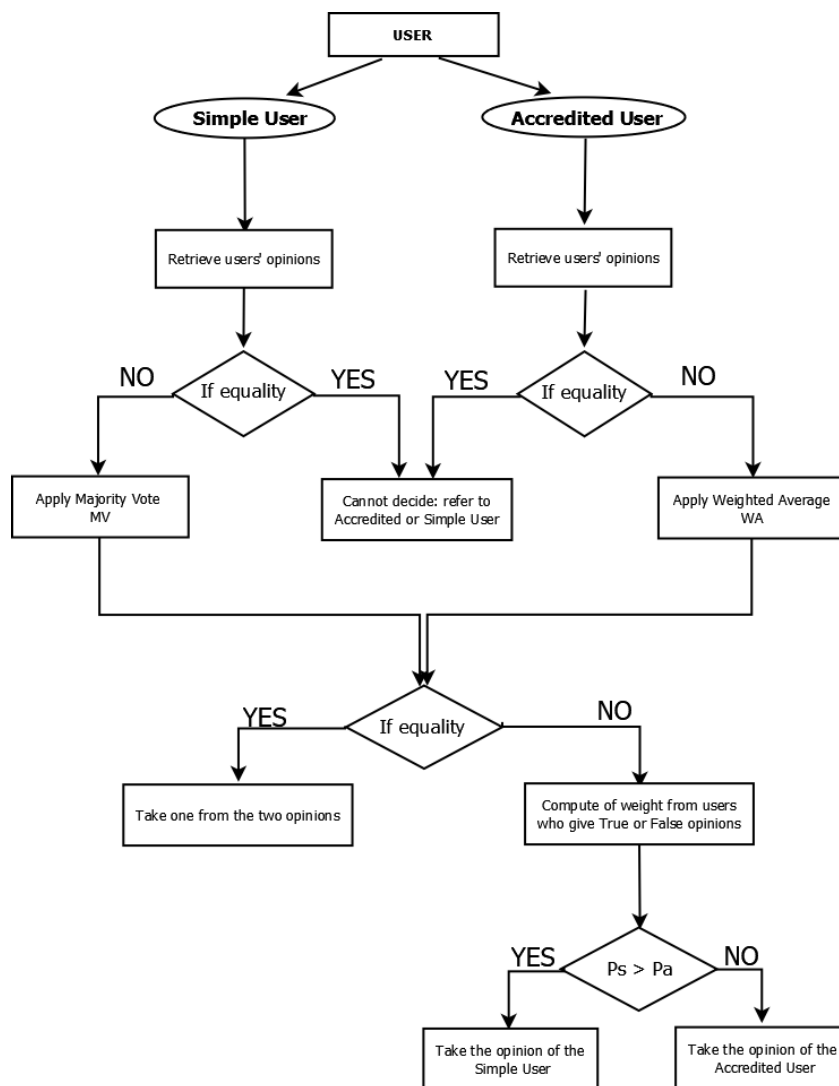


Figure 4. Flowchart of proposed model.

#### 4.5. Decision-Making

To give the final opinion of the position, two comparison cases of opinions are made on both parties. In case, the opinion of simple users gives 1 (true) after aggregation and those of accredited users also gives 1 (true), the final decision of the post will be 1 (true). In case, the opinion of simple users gives  $-1$  (false) after aggregation and those of accredited users gives  $-1$  as well, the final decision of the post will be  $-1$  (false). The challenge lies at the level where the opinions of both sides differ. When opinions differ, the weightings calculated in relation to the voted opinion shall be examined.

1. If  $P_a > P_s$ , then the final opinion is the opinion of accredited users with a weight of  $P_a$
2. Else, then the final opinion is the opinion of simple users with a weight of  $P_s$

This decision is based on Table 2.

**Table 2.** Decision-making.

Simple Users	Accredited Users	Decision
$-1$ (F)	$-1$ (F)	$-1$ (F)
$-1$ (F)	1 (T)	1 (T)
1 (T)	$-1$ (F)	$-1$ (F)
1 (T)	1 (T)	1 (T)

Based on the comparison of user weights, we have two cases. The first case is the case where weights of accredited users are **greater than** that of simple users. Table 3 presents details of this case.

**Table 3.** Decision-making  $P_a > P_s$ .

Simple User	Accredited User	Weight	Decision
$-1$ (F)	$-1$ (F)	—	$-1$ (F)
$-1$ (F)	1 (T)	$P_a > P_s$	1 (T)
1 (T)	$-1$ (F)	$P_a > P_s$	$-1$ (F)
1 (T)	1 (T)	—	1 (T)

The second case is the case where weights of simple users are **greater than that** of accredited users. Table 4 presents details of this case.

**Table 4.** Decision-making  $P_s > P_a$ .

Simple User	Accredited User	Weight	Decision
$-1$ (F)	$-1$ (F)	—	$-1$ (F)
$-1$ (F)	1 (T)	$P_s > P_a$	$-1$ (F)
1 (T)	$-1$ (F)	$P_s > P_a$	1 (T)
1 (T)	1 (T)	—	1 (T)

**Example 1:** Consider a set of 10 simple users and 04 accredited users in an educational institution. The post is in the field of education “Stabbed teacher at Nkolbisson High School got into a fight in class with his student before passing away”. Table 5 presents the opinions of simple users and their associated weights.

**Table 5.** Decision-making Simple User.

Simple User	Competence	Vote
Student 1	0.25	1
Student 2	0.25	1
Student 3	0.25	−1
Student 4	0.25	1
Student 5	0.25	1
Student 6	0.25	1
Student 7	0.25	−1
Student 8	0.25	1
Student 9	0.25	1
Student 10	0.25	1

Table 6 presents the opinions of accredited users and their associated weights.

**Table 6.** Decision-making Accredited User.

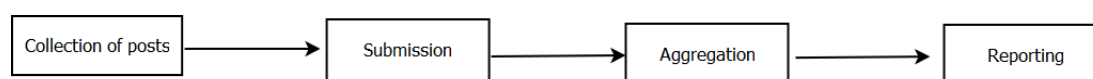
Accredited User	Competence	Vote
Principal	1	−1
Vice-principal	0.75	1
SG	0.75	−1
Teacher	0.5	1

$MV = 1$  because 8 students voted 1 and  $P_s = 2(0.25 \times 8)$ ;  $W_A = -0.16$  means that the  $W_A$  is  $-1$ . Please note that  $P_a = 1.75(1 + 0.75)$  and  $MV \neq W_A$ .  $P_s > P_a$  then the post is 1 (true) with a weight of 2 of the simple users or 80%.

In this work, the minimum of users to make a correct decision is 10 including at least one accredited user and at least nine simple users. The weights assigned to accredited users implies that they are more important. So there must have at least one accredited user in the decision-making.

## 5. Experiments, Results and Discussion

This experimentation process can be summarized in four steps according to the Figure 5.

**Figure 5.** Experimentation process.

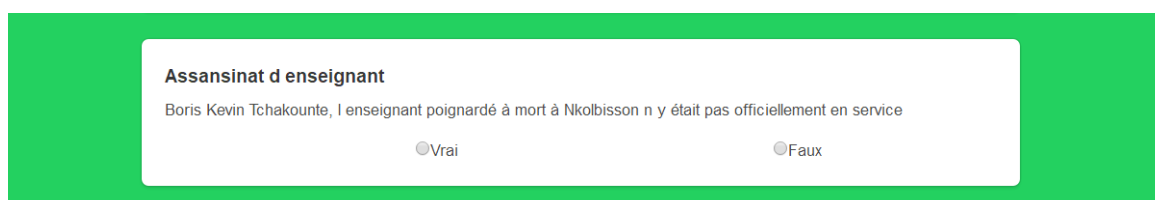
This section presents the experiment of the proposed aggregation model on different use cases. The results are discussed, and limitations are outlined. As shown in Figure 5, the first step consists of collecting fake news samples; the second step submits the hoaxes to user categories requested to give their opinion. The third step applies the aggregation model taking as inputs a different opinion; the last step reports the decision made based on the output from the aggregation model.

### 5.1. Collection of Posts

The dataset of fake news has been collected from December 2019 to January 2020 in stopblablacam [16], which is a web site that aims to assess information vehiculated in social media in Cameroon. The dataset includes 25 posts that were on top headlines across social media in Cameroon. They are related to the education aspect and listed in Appendix A (Original version) and Appendix B (English adapted version). There are posts concerning news about competitions for entrance exams, the announcement of official exams, awards of official diplomas, dates of competitions and exams, defense of master and PhD theses, and other related fake news.

### 5.2. Submission

Every post is submitted to the crowd via a developed online platform [47]. The crowd is made up of 80 students from the University of Ngaoundéré in computer science and 05 students from secondary school GHS Burkina under-graduates and post-graduates recruited through student social media group. The link of submission had been sent to these students and volunteers participated in the assessment process. The crowd is made of 3 accredited users in each institution concerned with the post. Since it was difficult to work with real life individuals in these institutions, we adopted to simulate the opinions of accredited users in the three top levels of hierarchy. A student authenticates and visualizes the set of posts to assess based on experience and intelligence. Figure 6 shows a post stating that “The murdered teacher at Nkolbison was not in his official workplace” in the proposed platform. The student is asked to select the appropriate opinion.



The screenshot shows a voting interface on a green background. The title is "Assassinat d enseignant". The text of the post is "Boris Kevin Tchakounte, l'enseignant poignardé à mort à Nkolbisson n y était pas officiellement en service". Below the text are two radio buttons: "Vrai" (True) and "Faux" (False).

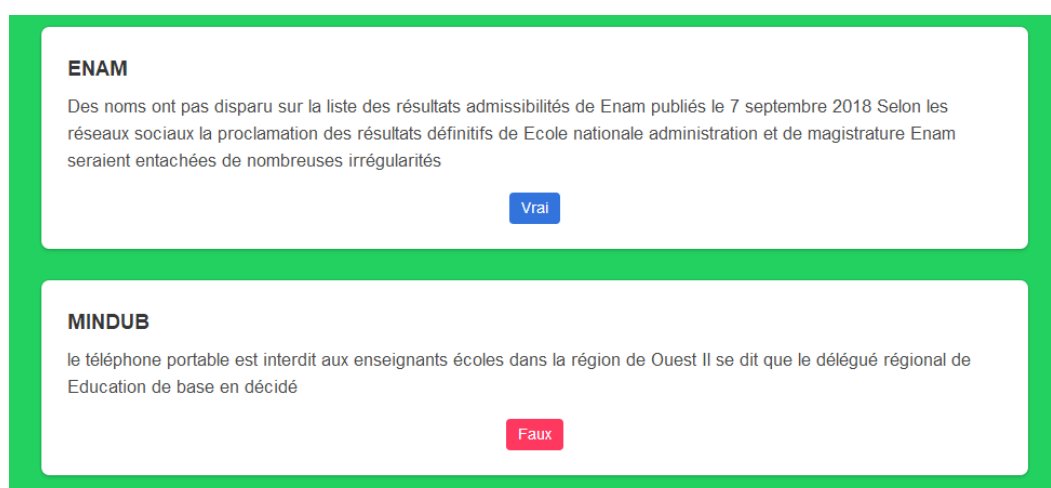
Figure 6. Voting post in the form.

### 5.3. Aggregation

This section applies the proposed aggregation model; this model is applied on both sides of users, on the simple user side the majority vote aggregator is applied and on the accredited user side the weighted average is applied. Weight comparison is made in case of a difference of opinion between simple and accredited users. The final decision is obtained by taking the opinion with the highest weight.

### 5.4. Reporting

This section reports the final decision of the post after applying the proposed model. It gives the final decision of the post i.e., True or False in the platform developed and available online. Figure 7 shows an example of a post checked and aggregated based on the proposed model in the platform.



The screenshot shows two posts with their final aggregated results. The first post is titled "ENAM" and the text is "Des noms ont pas disparu sur la liste des résultats admissibilités de Enam publiés le 7 septembre 2018 Selon les réseaux sociaux la proclamation des résultats définitifs de Ecole nationale administration et de magistrature Enam seraient entachées de nombreuses irrégularités". Below the text is a blue button labeled "Vrai". The second post is titled "MINDUB" and the text is "le téléphone portable est interdit aux enseignants écoles dans la région de Ouest Il se dit que le délégué régional de Education de base en décidé". Below the text is a red button labeled "Faux".

Figure 7. Result via the form.

The results of the various posts applied to the proposed model are depicted in Table 7.

Table 7. Results.

Title Post	Nature of Post	Number of User	$\Theta$	$\Gamma$	$Decision_{Model}$	$Decision_{MV}$
1- ENAM	True	50	26	24	True	True
2- GCE	True	47	30	17	True	True
3-MINDUB	True	45	20	25	False	False
4- MINSEC	True	47	27	20	True	True
5- Concour	False	48	18	30	False	False
6- Bilingusme	True	44	28	16	True	True
7- Philosophie	True	49	41	8	True	True
8- StopBlaBla	True	47	26	21	True	True
9- Week end	True	48	32	16	True	True
10- FMIP Kaelé	True	47	28	19	False	True
11-Conc ENAM	False	48	25	23	False	True
12- UBa	False	46	22	24	False	False
13- Examen	False	48	23	25	True	False
14- Avocat	True	47	24	23	False	True
15-Bacc 2018	True	50	33	17	True	True
16- Corruption	True	50	41	9	True	True
17- Salaire	False	48	18	30	False	False
18- MINFOPRA	False	48	25	23	True	True
19-PBHev	False	49	25	24	False	True
20- Jeux-U	False	49	25	24	True	True
21- Examen FS	False	48	14	34	False	False
22-IUT	False	50	10	40	False	False
23-CDTIC	False	48	16	32	False	False
24-Doctoriales	False	46	21	25	True	False
25- Soutenance FS	False	45	26	19	False	True

### 5.5. Detection Performance

Different criteria has been used to evaluate the performance of the proposed model. The objective is to detect fake news through the crowd. The results of the detection performance are obtained in Table 8.

- **True Positive (TP):** Number of posts correctly identified as fake news;
- **False Positive (FP):** Number of posts wrongly identified as fake news;
- **True Negative (TN):** Number of posts correctly identified as not fake news;
- **False Negative (FN):** Number of posts wrongly identified as not fake news;

**Precision:** Concerning positive samples, the precision is the proportion of true positives from the true and false positives ( $Precision_P$ ). Concerning negative samples, the precision is the proportion of true negatives from the true and false negatives ( $Precision_N$ ). Their formulas are given as follows:

$$Precision_P = \frac{TP}{TP + FP} \quad (6)$$

$$Precision_N = \frac{TN}{TN + FN} \quad (7)$$

**Accuracy:** Accuracy is considered to be the proportion of correctly identified instances among the total number of posts examined. It is used to calculate the score of user's prediction labels when matched to the gold standard ones. Accuracy is calculated according to the following formula.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{TP + TN}{n}, \quad (8)$$

where  $n$  represents several posts or size of the sample. Each metric attempts to describe and quantify



the effectiveness of the proposed model by comparing a list of predictions to a list of correct answers. Table 8 shows the result of the detection performance of the proposed model.

**Table 8.** Results metric.

	Samples	TP	TN	FP	FN	<i>Precision<sub>p</sub></i>	<i>Precision<sub>N</sub></i>	Acc
Majority Vote Model	25	11	8	5	1	0.68	0.88	0.76%
Our Model	25	11	9	4	1	0.73	0.9	0.80%

It appears that the proposed aggregation model has identified fake news with a probability of 80%. The model wrongly recognized four benign posts. It failed to recognize either fake news or true news due to the selection of crowd, relatively to the subject of the post. As an example, the 3rd, 18th, 19th, and 20th posts failed to be recognized as benign because students did not have enough knowledge about this news. For instance, the 20th post is about university games. This news is difficult to verify by students from secondary school. Also, the 24th post concerned PhD student presentation. This post was hardly verifiable by a student at a lower level. The proposed model accuracy is 80%, which shows that it is capable of detecting and recognizing fake news. Our model is more precise in the identification of negative samples (88.88%) than in the identification of positive samples (0.6875%). A positive note concerns the 8th post which was formally declared by StopBlaBlaCam as fake news, the model has misclassified this post as benign. Later, the StopBlaBlaCam journalists recognized that this post was benign. This situation shows that the proposed model is robust in its decision-making.

We have experimented the majority vote model against the proposed model on this study samples. The first, i.e., the model in which each voter has a weight 1, is accurate with a probability of 0.76 to correctly classify an instance, with an average precision rate of 0.78. The proposed model has a slightly better detection performance than the majority vote in terms of accuracy, TN and FP. The difference between the two is on one case. This difference reveals the importance of coupling accredited users and simple users. In fact, Table 8 presents that there is one negative case correctly identified by the proposed model. This case refers to the post 25. The majority vote model gives true for the post 25 instead of false. It means that more users provided true opinions. Since this model totally trusts an opinion, i.e., there is no influence, the final result is true. On the contrary, the proposed model adds external factors influencing an opinion and even, truthfulness related to information sources. Accredited people fit within highest hierarchy and therefore they are informed with update information. Considering truthfulness brings to some reality. Nevertheless, we need experiments with more participants and larger dataset of posts to provide more evidence of the presence of accredited people.

### 5.5.1. Comparison with Related Works

Table 9 shows seven criteria used to compare the proposed model with similar works in the literature. For the first criterion the crowd can be simple, which means that all users have the same skill to assess the post. The second criterion is the decision, which specifies how users are involved in decision-making. The third criterion includes techniques used in various work to determine the nature of the post. The fourth criterion describes the methodology exploited in different works. The fifth criterion describes roles played by members of crowd. The sixth criterion specifies whether it is a classification or detection problem. The last criterion specifies whether the nature of posts is known or unknown at the beginning.

Table 9. Comparison with related works.

Authors	Types of Crowd	Decision	Associated Techniques	Methods	Role of Crowd	Problem	Nature of Post
(Tschatschek et al., 2018)	One type (Simple)	Expert involvement	Hybrid (Crowdsourcing, Artificial Intelligence)	to select a small subset of k news, send them to an expert for review, and then block the news which are labeled as fake by the expert	Selection of facts to be checked	Classification post as fake and not fake	Unknown at start
(Vedova et al., 2018)	One type (Simple)	Use harmonic Boolean label crowdsourcing for making decision logistic regression and harmonic Boolean label crowdsourcing	Define a threshold $\lambda$ and classified the posts	Combining content-based and social-based approaches	Provide the BLC algorithm input data	Classification post as fake or not fake	Known at start
(Tacchini et al. 2017)	One type (Simple)	Use algorithms derived from crowdsourcing (BLC) for making decision	Logistic regression and crowdsourcing	The algorithms compare what people say, correct for the effect of the liars, and reconstruct a consensus truth	Provide the BLC algorithm input data	Classification post as hoax or not	Known at start
(Pérez-Rosas et al., 2017)	One type (Simple)	Machine Learning make decision	Crowdsourcing and Machine Learning	Explore and analyze linguistic properties to detect false content	Uses the crowd just to collect dataset	Identification of fake content in online news	Known at start
(Kim et al. 2018)	One type (Simple)	Expert involvement	Crowdsourcing	Users flag any story in their feed as misinformation and, if a story receives enough flags, it is sent to a third party for fact-checking. Combines the machine learning algorithms with the wisdom of crowds to detect fake news and satire	To determine when the news needs to be verified	Detection fake news	Unknown at start
(Shabani et Shakhn, 2018)	One type (Simple)	Expert involvement	Hybrid (Machine Learning and crowdsourcing)	Detect fake news while combining crowd intelligence and third-party expert knowledge.	Selection of facts to be checked	Classification post	Known at start
Our model	Two types (Simple and Accredited)	Simple user and Accredited user involvement	Crowdsourcing		Involved in decision-making	Detection fake news	Unknown at start

## 6. Discussion

Discussions are based on the following criteria.

- **Type of crowd:** All the authors use only one type of crowd, i.e. they consider that all the people involved in the crowdsourcing process have the same level of competence or skill, these authors ignore that some people may have higher knowledge, experience or skills than others. The proposed model takes into account two categories of users with different knowledge. Accredited users are people with certain administrative responsibilities in an institution. Their social status confers some legitimacy to give opinions about news happened in this institution. The second category includes crowd users who are anonymous. However, they can provide judgment about the news because they also belong to the mentioned institutions. They belong to the lowest level of the hierarchy.
- **Decision-making on a post:** All these authors do not involve the crowd in the decision-making, they just use the crowd either to collect data or to check whether a post should be checked or not. Most of these authors use a third-party (expert) for decision-making. The proposed model does not involve a third party but remains focused on the knowledge and skills of the crowd to decide on a post.
- **Associated techniques :** most authors use hybrid techniques combining crowdsourcing and machine learning. Only the author (Kim et al. 2018) uses only crowdsourcing just to determine when the news should be checked. The proposed model uses crowdsourcing to verify and involves the crowd in the decision-making process of a post.
- **Methodology:** each author has a specific method to detect false information. The proposed model also has its own method to detect fake news, a method combining the intelligence of the simple crowd and third party to detect fake news.
- **Role of the crowd in the crowdsourcing process:** most authors do not fully use the skills of the crowd, they limit the role of the crowd just to verification and rely on an expert for decision-making. However, the crowd has a huge skill that these authors neglect, that of participation in decision-making. The proposed model takes the initiative to involve the crowd in decision-making.

From this discussion, it appears that the proposed model uses a new method of aggregating opinions involving the crowd in decision-making and categorizing the crowd into two types. Each type has a different specific competence. This method only uses crowdsourcing which is easily accessible to the crowd unlike methods using machine learning which require a large amount of data.

## 7. Conclusion and Future Works

In this project, an overview of crowdsourcing, fake news and a model for aggregating users' opinions (simple and accredited) are presented. We conducted literature on the various methodology, followed by the design and implementation of our proposed methods. This work follows a qualitative approach for the collection, aggregation of user opinions, and detection of fake news following our aggregation model. Authors relying on crowd do not use the expertise of the crowd and the expertise of a third-party in an associative way to make decisions. They just use the crowd to indicate whether or not a story should be verified. This work, therefore, combines the expertise of the crowd with that of a third party to make decisions about a position.

This work also proposed a novel binary aggregation method of crowd opinions and third-party expert knowledge. The aggregator relies on majority voting on the crowd side and weighted average on the third-party side. A model is proposed to compute voting weight on both sides. A comparison between weights is made to take a final decision. The latter concerns opinion with the higher weight. Experiments made with 25 posts and 50 participants demonstrate that associating the crowd and third-party expertise provides slightly better performance results compared with the majority vote model. Moreover, our aggregation model is accurate with 80% with an average precision of 81.5%

(precision for positive: 73.33% and for negative: 90%). This work requires more experiments on a larger dataset of heterogeneous posts as well as participants to provide more evidence of the aggregation model.

In future work, the model will be extended to :

- Consider various users and posts of various domains;
- Formulate an optimization problem around the assignment of weights;
- Couple our method with other mechanisms to contribute to the detection of fake news.

**Author Contributions:** Conceptualization, F.T. and A.F.; methodology, F.T. and A.F.; software, A.F.; validation, F.T.; writing—original draft preparation, F.T., A.F. and M.A.; writing—review and editing, A.N. and M.A.; supervision, F.T. All authors have read and agreed to the published version of the manuscript.

**Acknowledgments:** We thank the anonymous referees for numerous valuable comments to improve the manuscript.

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

## Abbreviations

The following abbreviations are used in this manuscript:

MV	Majority Vote
WA	Weighted Average
L1	Level 1
L2	Level 2
L3	Level 3

## Appendix A. Original Version of List of Posts

1. ENAM: Des noms n'ont pas disparu sur la liste des résultats d'admissibilités de l'ENAM publiés le 7 septembre 2018, selon les réseaux sociaux la proclamation des résultats définitifs de l'Ecole Nationale d'Administration et de Magistrature (ENAM) seraient entachées de nombreuses irrégularités.
2. GCE: le gouvernement camerounais n'a pas autorisé la collecte de 3500 FCFA pour le «GCE» blanc.
3. MINDUB: le téléphone portable est interdit aux enseignants d'écoles dans la région de l'Ouest, Il se dit que le délégué régional de Education de base en a décidé.
4. MINSEC: le Ngondo plaide pour 13 anciens proviseurs Sawa démis de leurs fonctions, un compte rendu non officiel relaie les griefs formulés par cette assemblée traditionnelle lors une rencontre tenue le 14 septembre dernier.
5. Concour: les concours des ENS et ENSET ne sont pas encore lancés au Cameroun, des arrêtés lançant ces deux concours auraient été signés par le Premier ministre chef du gouvernement.
6. Commission Bilinguisme: 15 véhicules une valeur de 700 millions ont été commandés par la Commission nationale du Bilinguisme, le document qui circule actuellement attribuant une grosse dépense à ce jeune organe étatique est il authentique.
7. Philosophie: Dès septembre la philosophie sera enseignée en classe de la seconde, il se dit que cette matière n'est plus seulement réservée à la classe de Terminale.
8. StopBlaBla: Stopblablacam s'est trompé au sujet d'une vidéo sur armée camerounaise dans l'Extrême Nord, Stopblablacam a commis une erreur en assurant il y'a quelques jours que la vidéo impliquant l'armée camerounaise était un fake.
9. Week end: Des fonctionnaires camerounais prennent leur weekend dès jeudi soir le 19 décembre 2019. Certains agents de l'Etat sont absents de leur lieu de service sans permission pendant des semaines.
10. FMIP kaelé: il y'a eu un mouvement humeur à la faculté des mines et des industries pétrolières de l'Université de Maroua à Kaelé dans la soirée du 23 juillet 2018 des images et des messages

informant une grève des étudiants à la faculté des mines et des industries pétrolières de Kaelé FMIP ont été diffusés sur les réseaux sociaux.

11. Concour ENAM: la Fonction publique camerounaise a mandaté aucun groupe pour préparer les candidats au concours de l'Enam, des groupes de préparation aux concours administratifs prétendent qu'ils connaissent le réseau sûr permettant être reçu avec succès.
12. UBa: cinq étudiants de l'université de Bamenda dans le Nord ouest du Cameroun n'ont pas été tués le 11 juillet 2018, ces tueries selon certains médias seraient la conséquence des brutalités policières qui opèrent des arrestations arbitraires dans les mini cités de l'université de Bamenda.
13. Examen: à l'examen, écrire avec deux stylos à bille encre bleue de teintes différentes n'est pas une fraude, il existe pourtant une idée reçue assez répandue sur le sujet.
14. Avocat: huit avocats stagiaires ont été radiés du barreau pour faux diplômes, cette décision a été prise lors de la session du Conseil de ordre des avocats qui s'est tenue le 30 juin dernier.
15. Bacc 2018: la délibération des jurys au bac 2018 au Cameroun s'est faite autour de la note 8 sur 20. Qui dit délibérations aux examens officiels au Cameroun dit ajouts de points et repêchage des candidats.
16. Corruption: La corruption existe au ministère de la Fonction publique du Cameroun, des expressions comme Bière, Carburant, Gombo constituent tout un champ lexical utilisé par les corrompus et les corrupteurs.
17. Salaire: le gouvernement n'exige pas 50 000 Fcfa à chaque fonctionnaire pour financer le plan humanitaire dans les régions anglophones, selon la rumeur la somme de 50 000 FCFA sera prélevée à la source des salaires des agents publics.
18. Fonction publique: le ministère de la Fonction publique camerounaise ne recrute pas via les réseaux sociaux, des recruteurs tentent de convaincre les chercheurs emploi que l'insertion dans administration publique passe par internet.
19. PBHev: il n'existe pas d'entreprise PBHev sur la toile, il se dit pourtant qu'il s'agit d'une entreprise qui emploie un nombre important de Camerounais.
20. Jeux Universitaire ndere: Les jeux universitaire ngaoundere 2020 initialement prévu a ngaoundere va connaitre un glissement de date pour juin 2020 car le recteur dit que l'Université de ngaoundere ne sera pas prêt pour avril 2020, les travaux ne sont pas encore achevé.
21. Examen FS: Le doyen de la faculté des sciences informe a tous les étudiants de la faculté que les examens de la faculté des sciences sont prévu pour le 11fevrier 2020.
22. IUT: le Directeur de l'IUT de ndere a invalider la soutence d'une étudiante de Génie Biologique pour avoir refuser de sortir avec lui.
23. CDTIC: Le Centre de developpement de Technologies de l'information et de la communication n'est plus a la responsabilité de l'université de ngaoundéré, mais plutot au chinois.
24. Doctoriales: Les doctoriales de la faculté des sciences s'est tenu en janvier sous la supervision de madame le recteur, pour ce faire plusieurs doctorants sont appelés a reprendre leur travaux a zero.
25. Soutenance de la FS: Il est prevu des soutenance le 15 decembre 2019.
26. Greve des enseignants: les enseignants des universités d'Etat sont en grève depuis le 26 novembre 2018 Il se rapporte que les enseignants du supérieur sont en grève
27. Recrutement Docteur PhD: des docteurs PhD ont manifesté ce mercredi au ministère de l'Enseignement supérieur, il se dit qu'ils ont une fois de plus réclamé leur recrutement au sein de la Fonction publique.
28. Descente des élèves dans la rue: des élèves du Lycée bilingue de Mbouda sont descendus dans la rue ce mardi 24 avril, il se dit qu'ils ont bruyamment manifesté pour revendiquer le remboursement de frais supplémentaires exigés par administration de leur établissement.

29. Faux diplôme: le phénomène des faux diplômes est en baisse au Cameroun, le ministère de l'Enseignement supérieur renseigne qu'on atteint une moyenne de 40 faux diplômes sur un ensemble de 600.
30. Trucage des âges: Des Camerounais procèdent souvent au trucage de leur âge, le ministre de la Jeunesse Mounouna Foutsou présenté des cas de falsifications de date de naissance lors du dernier renouvellement du bureau du Conseil national de la Jeunesse.
31. Interpellation d'enseignants: 300 enseignants camerounais n'ont pas été interpellés les 27 et 28 février 2018 à Yaoundé selon le Nouveau collectif des enseignants indignés du Cameroun le chiffre 300 ne peut être qu'erroné puisque le mouvement compte juste 56 membres soit un grossissement du nombre interpellés par cinq.
32. Report d'examen: les examens officiels ont pas été reportés à cause des élections, il se murmure qu'en vue de la présidentielle des municipales législatives et sénatoriales les dates des épreuves sanctionnant enseignement secondaire ont été chamboulées.
33. Assassinat d'enseignant: Boris Kevin Tchakounte, l'enseignant poignardé à mort à Nkolbisson n'y était pas officiellement en service.
34. Lycée d'obala: Un lycéen n'a pas eu la main coupée à Obala.
35. MINJUSTICE: les résultats du concours pour le recrutement de 200 secrétaires au ministère de la Justice ont été publiés.
36. Doctorat professionnel: le doctorat professionnel n'est pas reconnu dans l'enseignement supérieur au Cameroun.
37. MINAD: le ministre de l'Agriculture a proscrit, pour l'année 2019, des transferts de candidats admis dans les écoles agricoles.
38. MINFOPRA: le ministère camerounais de la Fonction publique dispose désormais d'une page Facebook certifiée.
39. Ecole de Nyangono: Nyangono du Sud n'est pas retourné sur les bancs de l'école.
40. CAMRAIL: Camrail a lancé un concours pour le recrutement de 40 jeunes camerounais et l'inscription coûte 13000 fcfa
41. Report date concours: la date du concours pour le recrutement de 200 secrétaires au ministère de la Justice a été reportée.
42. MINFI: Le ministre camerounais des finances ne recrute pas 300 employés.

#### **Appendix B. English Version of List of Posts**

1. NSAM: Names have not disappeared from the list of NSAM admissibility results published on 7 September 2018, according to social networks the proclamation of the final results of the National School of Administration and Magistracy (NSAM) would be marred by numerous irregularities.
2. GCE: the Cameroonian government has not authorized the collection of 3500 FCFA for the white "GCE".
3. MINDUB: Mobile phones are forbidden to teachers in schools in the Western Region, it is said that the Regional Delegate for Basic Education has decided on this.
4. MINSEC: Ngondo pleads for 13 former Sawa headmasters dismissed, an unofficial report relays the grievances formulated by this traditional assembly during a meeting held last September 14.
5. Competition: The ENS and ENSET competitions have not yet been launched in Cameroon, decrees launching these two competitions would have been signed by the Prime Minister, Head of Government.
6. Bilingualism Commission: 15 vehicles worth 700 million have been ordered by the National Bilingualism Commission. The document currently in circulation attributing a large expenditure to this young state body is authentic.
7. Philosophy: As of September, philosophy will be taught in the second year of secondary school. It is thought that this subject is no longer reserved only for the Terminale class.

8. StopBlaBla: Stopblablacam was wrong about a video about the Cameroonian army in the Far North, StopBlaBlaCam made a mistake a few days ago by assuring that the video involving the Cameroonian army was a fake.
9. Weekend: Cameroonian officials are taking their weekend starting Thursday night, December 19, 2019. Some government officials are absent from their place of duty without permission for weeks.
10. FMIP Kaelé: there was a mood movement at the Faculty of Mines and Petroleum Industries of the University of Maroua in Kaelé in the evening of July 23, 2018 images and messages informing of a strike by students at the Faculty of Mines and Petroleum Industries of Kaelé FMIP were broadcast on social networks.
11. ENAM competition: the Cameroonian civil service has not mandated any group to prepare candidates for the ENAM competition, some groups preparing for administrative competitions claim that they know the secure network to be received successfully.
12. UBa: Five students of the University of Bamenda in north-west Cameroon were not killed on 11 July 2018. According to some media reports, these killings are the consequence of police brutality which carry out arbitrary arrests in the mini cities of the University of Bamenda.
13. Examination: on the exam, writing with two blue ink ballpoint pens of different shades is not a fraud, yet there is a common misconception on the subject.
14. Lawyer: eight trainee lawyers were disbarred for false diplomas. This decision was taken at the session of the Bar Council held on 30 June last.
15. Bacc 2018: the deliberations of the juries for the 2018 Bacc in Cameroon were based on a score of 8 out of 20. Deliberations at the official exams in Cameroon meant adding points and repechage of candidates.
16. Corruption: Corruption exists in Cameroon's Ministry of Civil Service, expressions such as Beer, Fuel, Gombo constitute a whole lexical field used by the corrupt and corrupters.
17. Salary: The government does not require 50,000 FCFA from each civil servant to finance the humanitarian plan in the English-speaking regions, rumor has it that the sum of 50,000 FCFA will be deducted at source from the salaries of civil servants.
18. Civil service: the Cameroonian Ministry of Civil Service does not recruit via social networks; recruiters try to convince job seekers that integration into the public administration is via the Internet.
19. PBHev: There is no PBHev company on the web, yet it is said to be a company that employs a significant number of Cameroonians.
20. University Games : The University Games Ngaoundéré 2020 initially planned in Ngaoundéré will experience a date shift to June 2020 because the rector says that the University of Ngaoundéré will not be ready for April 2020, the work is not yet completed.
21. FS Exam: The Dean of the Faculty of Science informs all students of the Faculty of Science that the Faculty of Science exams are scheduled for February 11, 2020.
22. IUT: the Director of the IUT of Ngaoundéré has invalidated the support of a student of Biological Engineering for refusing to go out with him.
23. CDTIC: The Centre for the Development of Information and Communication Technologies is no longer the responsibility of the University of Ngaoundéré, but rather Chinese.
24. Doctoral: The Faculty of Sciences' doctoral session was held in January under the supervision of Madam Rector, for this purpose several doctoral students are called to resume their work from scratch.
25. FS defense: The defense is scheduled for December 15, 2019.
26. Teachers' strike: State University teachers are on strike since 26 November 2018 It is reported that higher education teachers are on strike.



27. Recruitment PhDs: PhDs demonstrated on Wednesday at the Ministry of Higher Education, it is said that they have once again demanded their recruitment within the civil service.
28. Pupils taking to the streets: Pupils from the Bilingual High School of Mbouda took to the streets on Tuesday 24 April. It is said that they noisily demonstrated to demand the reimbursement of additional fees demanded by the administration of their school.
29. False diplomas: the phenomenon of false diplomas is decreasing in Cameroon, the Ministry of Higher Education informs that an average of 40 false diplomas out of a total of 600 is reached.
30. Trickery of ages: Cameroonians often trick their age, the Minister of Youth Mounouna Foutsou presented cases of falsification of date of birth during the last renewal of the office of the National Youth Council.
31. Interpellation of teachers: 300 Cameroonian teachers were not questioned on 27 and 28 February 2018 in Yaounde according to the New Collective of Indignant Teachers of Cameroon, the figure 300 can only be erroneous since the movement has just 56 members, i.e., an increase in the number of teachers questioned by five.
32. Postponement of exams: the official exams have not been postponed because of the elections, it is whispered that in view of the presidential, legislative and senatorial municipal elections, the dates of the exams sanctioning secondary education have been disrupted.
33. Teacher's assassination: Boris Kevin Tchakounté, the teacher stabbed to death at Nkolbisson was not officially on duty there.
34. Obala High School: A high school student did not have his hand cut off in Obala.
35. MINJUSTICE: The results of the competition for the recruitment of 200 secretaries in the Ministry of Justice have been published.
36. Professional doctorate: The professional doctorate is not recognized in higher education in Cameroon.
37. MINAD: the Minister of Agriculture has banned, for 2019, transfers of candidates admitted to agricultural schools.
38. Cameroon's Ministry of Public Service now has a certified Facebook page.
39. Nyangono School: South Nyangono has not gone back to school.
40. CAMRAIL: Camrail has launched a competition for the recruitment of 40 young Cameroonians and registration costs 13000 fcfa.
41. Postponement of competition date: the competition date for the recruitment of 200 secretaries in the Department of Justice has been postponed.
42. MINFI: Cameroon's finance minister is not recruiting 300 employees.

## References

1. Fake News—Statistics & Facts. Available online: <https://www.statista.com/topics/3251/fake-news/> (accessed on 31 May 2020).
2. Wasserman, H.; Madrid-Morales, D. An Exploratory Study of “Fake News” and Media Trust in Kenya, Nigeria and South Africa. *Afr. J. Stud.* **2019**, *40*, 107–123. [[CrossRef](#)]
3. Harsin, J.; Richet, I. Un guide critique des Fake News: de la comédie à la tragédie. *Pouvoirs* **2018**, *1*, 99–119. [[CrossRef](#)]
4. Tschatschek, S.; Singla, A.; Gomez Rodriguez, M.; Merchant, A.; Krause, A. Fake news detection in social networks via crowd signals. In Proceedings of the Web Conference 2018, Lyon, France, 23–27 April 2018; pp. 517–524.
5. Della Vedova, M.L.; Tacchini, E.; Moret, S.; Ballarin, G.; DiPierro, M.; de Alfaro, L. Automatic online fake news detection combining content and social signals. In Proceedings of the 2018 22nd Conference of Open Innovations Association (FRUCT), Petrozavodsk, Russia, 11–13 April 2018; pp. 272–279.
6. De Alfaro, L.; Di Pierro, M.; Agrawal, R.; Tacchini, E.; Ballarin, G.; Della Vedova, M.L.; Moret, S. Reputation systems for news on twitter: A large-scale study. *arXiv* **2018**, arXiv:1802.08066.

7. Shabani, S.; Sokhn, M. Hybrid machine-crowd approach for fake news detection. In Proceedings of the 2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC), Philadelphia, PA, USA, 18–20 October 2018; pp. 299–306.
8. Kim, J.; Tabibian, B.; Oh, A.; Schölkopf, B.; Gomez-Rodriguez, M. Leveraging the crowd to detect and reduce the spread of fake news and misinformation. In Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, Marina Del Rey, CA, USA, 5–9 February 2018; pp. 324–332.
9. Zhou, X.; Zafarani, R. Fake news: A survey of research, detection methods, and opportunities. *arXiv* **2018**, arXiv:1812.00315.
10. Sharma, K.; Qian, F.; Jiang, H.; Ruchansky, N.; Zhang, M.; Liu, Y. Combating fake news: A survey on identification and mitigation techniques. *ACM Trans. Intell. Syst. Technol.* **2019**, *10*, 1–42. [[CrossRef](#)]
11. Hassan, N.; Arslan, F.; Li, C.; Tremayne, M. Toward automated fact-checking: Detecting check-worthy factual claims by ClaimBuster. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 13–17 August 2017; pp. 1803–1812.
12. Hassan, N.; Li, C.; Tremayne, M. Detecting check-worthy factual claims in presidential debates. In Proceedings of the 24th ACM International Conference on Information and Knowledge Management, Melbourne, Australia, 19–23 October 2015; pp. 1835–1838.
13. Sethi, R.J. Crowdsourcing the verification of fake news and alternative facts. In Proceedings of the 28th ACM Conference on Hypertext and Social Media, Prague, Czech Republic, 4–7 July 2017; pp. 315–316.
14. Tacchini, E.; Ballarin, G.; Della Vedova, M.L.; Moret, S.; de Alfaro, L. Some like it hoax: Automated fake news detection in social networks. *arXiv* **2017**, arXiv:1704.07506.
15. Latest Email and Social Media Hoaxes—Current Internet Scams—Hoax-Slayer. Available online: <https://hoax-slayer.com/> (accessed on 31 May 2020).
16. StopBlaBlaCam. Available online: <https://www.stopblablacam.com/> (accessed on 31 May 2020).
17. Brown-Liburud, H.; Cohen, J.; Zamora, V.L. The Effect of Corporate Social Responsibility Investment, Assurance, and Perceived Fairness on Investors’ Judgments. In Proceedings of the 2011 Academic Conference on CSR, Tacoma, WA, USA, 14–15 July 2011.
18. Hoffart, J.; Suchanek, F.M.; Berberich, K.; Weikum, G. YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia. *Artif. Intell.* **2013**, *194*, 28–61. [[CrossRef](#)]
19. Pisarevskaya, D. Deception detection in news reports in the russian language: Lexics and discourse. In Proceedings of the 2017 EMNLP Workshop: Natural Language Processing Meets Journalism, Copenhagen, Denmark, 7 September 2017; pp. 74–79.
20. Potthast, M.; Kiesel, J.; Reinartz, K.; Bevendorff, J.; Stein, B. A stylometric inquiry into hyperpartisan and fake news. *arXiv* **2017**, arXiv:1702.05638.
21. Volkova, S.; Shaffer, K.; Jang, J.Y.; Hodas, N. Separating facts from fiction: Linguistic models to classify suspicious and trusted news posts on twitter. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), Vancouver, Canada, 30 July–4 August 2017; pp. 647–653.
22. Ren, Y.; Ji, D. Neural networks for deceptive opinion spam detection: An empirical study. *Inf. Sci.* **2017**, *385*, 213–224. [[CrossRef](#)]
23. Wang, Y.; Ma, F.; Jin, Z.; Yuan, Y.; Xun, G.; Jha, K.; Su, L.; Gao, J. Eann: Event adversarial neural networks for multi-modal fake news detection. In Proceedings of the 24th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), London, UK, 19–23 August 2018; pp. 849–857.
24. Du, N.; Liang, Y.; Balcan, M.; Song, L. Influence function learning in information diffusion networks. In Proceedings of the International Conference on Machine Learning, Beijing, China, 21–26 June 2014; pp. 2016–2024.
25. Najari, A.; Denoyer, L.; Gallinari, P. Predicting information diffusion on social networks with partial knowledge. In Proceedings of the 21st International Conference on World Wide Web, Lyon, France, 16–20 April 2012; pp. 1197–1204.
26. Draper, N.R.; Smith, H. *Applied Regression Analysis*; John Wiley & Sons: New York, NY, USA, 1998; Volume 326.
27. Kucharski, A. Study epidemiology of fake news. *Nature* **2016**, *540*, 525. [[CrossRef](#)]
28. Shu, K.; Sliva, A.; Wang, S.; Tang, J.; Liu, H. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explor. Newsl.* **2017**, *19*, 22–36. [[CrossRef](#)]

29. Ma, J.; Gao, W.; Wong, K.F. Rumor detection on twitter with tree-structured recursive neural networks. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, Australia, 15–20 July 2018.
30. Vishwanathan, S.V.N.; Schraudolph, N.N.; Kondor, R.; Borgwardt, K.M. Graph kernels. *J. Mach. Learn. Res.* **2010**, *11*, 1201–1242.
31. Wu, K.; Yang, S.; Zhu, K.Q. False rumors detection on sina weibo by propagation structures. In Proceedings of the 2015 IEEE 31st International Conference on Data Engineering, Seoul, Korea, 13–16 April 2015; pp. 651–662.
32. Petersen, E.E.; Staples, J.E.; Meaney-Delman, D.; Fischer, M.; Ellington, S.R.; Callaghan, W.M.; Jamieson, D.J. Interim guidelines for pregnant women during a Zika virus outbreak—United States, 2016. *Morb. Mortal. Wkly Rep.* **2016**, *65*, 30–33. [[CrossRef](#)] [[PubMed](#)]
33. Esteves, D.; Reddy, A.J.; Chawla, P.; Lehmann, J. Belittling the source: Trustworthiness indicators to obfuscate fake news on the web. *arXiv* **2018**, arXiv:1809.00494.
34. Dungs, S.; Aker, A.; Fuhr, N.; Bontcheva, K. Can rumour stance alone predict veracity? In Proceedings of the 27th International Conference on Computational Linguistics, Santa Fe, NM, USA, 20–26 August 2018; pp. 3360–3370.
35. Kleemann, F.; Voß, G.; Rieder, K. Un (der) paid Innovators. The Commercial Utilization of Consumer Work through Crowdsourcing. *Sci. Technol. Innov. Stud.* **2008**, *4*, 5–26.
36. Schenk, E.; Guittard, C. Towards a characterization of crowdsourcing practices. *J. Innov. Econ. Manag.* **2011**, pp. 93–107. [[CrossRef](#)]
37. Guittard, C.; Schenk, E. Le crowdsourcing: typologie et enjeux d’une externalisation vers la foule. *Document de Travail du Bureau d’économie Théorique et Appliquée* **2011**, *2*, 7522.
38. Thomas, K.; Grier, C.; Song, D.; Paxson, V. Suspended accounts in retrospect: An analysis of twitter spam. In Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference, Berlin, Germany, 2–4 November 2011; pp. 243–258.
39. Bouncken, R.B.; Komorek, M.; Kraus, S. Crowdfunding: The current state of research. *Int. Bus. Econ. Res. J.* **2015**, *14*, 407–416. [[CrossRef](#)]
40. Świeszczyk, M.; Świeszczyk, K. Crowdsourcing—what it is, works and why it involves so many people? *World Sci. News* **2016**, *48*, 32–40.
41. Ghezzi, A.; Gabelloni, D.; Martini, A.; Natalicchio, A. Crowdsourcing: A review and suggestions for future research. *Int. J. Manag. Rev.* **2018**, *20*, 343–363. [[CrossRef](#)]
42. Howe, J. The rise of crowdsourcing. *Wired Mag.* **2006**, *14*, 1–4.
43. Lin, Z.; Murray, S.O. Automaticity of unconscious response inhibition: Comment on Chiu and Aron (2014). *J. Exp. Psychol. Gen.* **2015**, *144*, 244–254. [[CrossRef](#)] [[PubMed](#)]
44. Granskogen, T. Automatic Detection of Fake News in Social Media using Contextual Information. Master’s Thesis, NTNU, Trondheim, Norway, 2018.
45. Kumar, S.; Shah, N. False information on web and social media: A survey. *arXiv* **2018**, arXiv:1804.08559
46. Marsick, V.J.; Watkins, K. *Informal and Incidental Learning in the Workplace (Routledge Revivals)*; Routledge: Abingdon, UK, 2015.
47. Ahmadou, F. Welcome to Our Crowdsourcing Platform. Available online: [hoax.smartedubizness.com](http://hoax.smartedubizness.com) (accessed on 1 June 2020).
48. Kirsch, S. Sustainable mining. *Dialect. Anthropol.* **2010**, *34*, 87–93. [[CrossRef](#)]

