Using credibility to stop the spread of fake news

Hessel Akkerman

University of Twente

BSc Creative Technology

Supervisor: Andreas Kamilaris Critical observer: Job Zwiers

January 2021

Abstract

One of the biggest problems of today, is the spread of fake news. More and more people are getting exposed to fake news and might make wrong decisions because of it. In this thesis, the goal is to find an automatic way to separate fake news articles from real articles, which could be used to stop this spread. To do this, the credibility of an article will be looked at, in order to determine if an article is real or not. An algorithm was created, to determine the credibility, based on the author, publisher, and listed references. These article attributes would be compared with other datasets on the internet, and an estimation of the credibility would be given. If an article were estimated as credible, it would also mean the content would be true. Through testing, an F1 score of .75 was found, with the true positive being detecting a real article as real. This result shows that there is a potential for using credibility, to verify an article as real. The credibility was also used to detect fake articles as fake, and an even higher F1 score was found, but because testing was done with limited data, and a non-credible source could still speak the truth, this result would be considered as inconclusive. Even though this is the case, being able to detect real news is a big part of the puzzle, in trying to stop the spread of misinformation, and thus the results from this thesis can be considered a step forward in solving the problem of fake news.

Acknowledgement

For this project, I would really like to thank my supervisor, Andreas Kamilaris, for his help and guidance. Our meetings really helped me when trying to come up with good ideas, and made sure I stayed on schedule, which could be quite difficult for me sometimes, with being in COVID-19 lockdown. His feedback also had a big impact on the results found in this report. Without him, I would not have created an algorithm that worked as well as this one.

Contents

1	Intr	Introduction 7					
	1.1	Proble	em statement	. 7			
	1.2	Possible solution					
	1.3	Research question					
	1.4	Outlir	ne	. 8			
2	Stat	State of the art on fake news detection					
	2.1	Defining fake news					
	2.2	2 Fake news detection methods					
		2.2.1	Linguistic analysis	. 10			
		2.2.2	Network analysis	. 10			
		2.2.3	Crowdsourcing	. 11			
		2.2.4	Conclusion	. 11			
		2.2.5	Discussion	. 11			
	2.3	Relate	d work	. 12			
		2.3.1	Fact check websites	. 12			
		2.3.2	Report button	. 13			
		2.3.3	Verification	. 14			
		2.3.4	Teaching programs	. 15			
	2.4	Deter	mining credibility	. 15			
3	Methods and techniques						
4	Idea	ation		18			
	4.1	The al	gorithm	. 18			
		4.1.1	Authors	. 18			
		4.1.2	Publisher	. 18			
		4.1.3	References	. 19			
	4.2	User i	nteraction	. 19			
		4.2.1	Browser extension	. 19			
		4.2.2	API	. 20			
		4.2.3	Website	. 20			
5	Spe	cificatio	on	21			

	5.1	Algorithm specifications	21	
	5.2	User interface specifications	21	
6	Real	lisation		
	6.1	Tools	22	
	6.2	Credibility algorithm	22	
		6.2.1 Author credibility	22	
		6.2.2 Publisher credibility	23	
		6.2.3 Reference credibility	24	
	6.3	Testing	24	
	6.4	Creating a user interface	25	
	6.5	Creating an API	25	
7	Eval	uation	26	
8	Conclusion			
9) Future work			

List of Figures

1	Politifact.com	12
2	Reporting fake news on facebook	13
3	Account verification on twitter	14
4	Website of the news literature project	15
5	Creative Technology design cycle	17
6	A model of the algorithm	19
7	Concept for a news verification mark	20
8	Screenshot of explaining why an article is credible	26

1. INTRODUCTION

1.1. Problem statement

One of the big problems of today is the increase in the spread of fake news. With all our current technology, it has become possible that anybody's voice can be heard. Generally, this is a good thing, and it is benefitting society in a lot of ways. However, it turns out that this also has a downside, namely the fact that also any lie can be heard. People make up stories, because of financial reasons, political reasons, or otherwise, and use these stories to convince others. This telling of lies in itself is not that big of a problem, as long as the target audience is small, but due to the scale of the internet, its audience could be the whole world. Combine this with the fact that most people have difficulty in determining if an article is true or not, and you have a problem. In a psychological experiment about lie detecting [1], it was found that only 54 percent of all participants were able to separate a lie from the truth. This is only a little more than chance, which indicates the difficulty is quite large.

Besides the spread, and being unable to separate truth from convincing lies, there is also the problem of living in a so called bubble. Today, people get most of their information online, and this results in another big problem, namely polarization. The big internet companies try to make as much money as possible, and to do this, they recommend stuff you might like, so they can keep your attention, and show you their ads. The problem with this, is that the user only sees content that he or she agrees with, because the user does often not care about viewpoints different from itself. By repeatedly showing the same type of articles, and thus confirming the user's beliefs, the user starts to believe that what he or she believes is the only truth.

To give an idea of the scale of the problem is, a few examples of the problems that fake news has caused, in response to COVID-19 outbreak, can be looked at. When the first signs of an outbreak of COVID-19 were showing, a fake news article spread around on social media, that stated the cause of the virus were the new 5G cell towers that were recently installed. Because of this, people started setting these towers on fire, because they believed destroying these towers would stop the outbreak. Another example related to COVID-19, was refusing to adhere to the rules that were implemented to slow the spread of the virus. A fake news article had stated that COVID-19 was not real, and was a lie created by the government to control the people. By making people believe this, and thus leading to them not adhering to the rules, it might have caused other people to get infected, and it might even have been the cause of death for some.

1.2. Possible solution

A way of solving the problem of fake news, would be to limit the exposure to it. Due to the importance of the problem, already a lot of research has been done, to detect if an article is fake or not, because if you know an article is fake, you can try to stop the spread. However, since fake news is still a big problem, and seems to become an even bigger problem, it is very important do more research about detecting fake news. Thus, in this thesis, the goal will be to make a step forward in the detection of fake news, so that hopefully one day, websites could have fake news filters build in, and thus limit exposure.

1.3. Research question

When looking at related work, most work focuses on finding patterns in fake news. However, the most effective method today is still people manually fact checking articles. The problem with this method, is that it does not scale. If it would be possible to automate what these fact checkers do, it could have a lot of potential in solving the problem. Fact checkers determine if something is real, by cross-checking facts, and by determining the credibility. Since cross-checking is still quite difficult with today's technology, the goal in this thesis will be to find out if credibility could help solve the problem. This leads to the following research question:

What is the potential of using an articles credibility to determine if an article is fake news or not fake news?

1.4. Outline

In this thesis, first the state of the art will be described. There, a definition of the term fake news will be given, related research will be looked at, and related work will be described. After this, the methods and techniques for answering the research question will be explained. Following the methods and techniques chapter, chapter four will describe the ideation process, chapter five will give a specification, chapter six will be about the realization, chapter seven will be the evaluation, and finally chapter eight will give a conclusion, and an answer to the research question. After this, also some potential future research will be discussed.

2. State of the art on fake news detection

In this section, the goal is to find what the current state of fake news detection is, related to the scope of this thesis. In order to answer this question, the state of the art will be divided into three subsections, where the first section will be about the definition of fake news, the second section will be about what methods are currently being used to identify fake news, and the third section will be about what implementations are currently used.

2.1. Defining fake news

To find a possible solution for fake news, it is important to get a good understanding of the problem. To get a better understanding of the problem, the term fake news should first be better defined. The term fake news does unfortunately not have a universal definition. Fake news has been defined as "a news article that is intentionally and verifiable false" [2], "a news article or message published and propagated through media, carrying false information regardless of the means and motives behind it"[3], "misinformation"[4], "satire news"[5] or "improper stories"[6]. From all these definitions, a general broad definition for fake news can be defined.

Definition 1. (Broad definition): Fake news is false news.

With this definition, it becomes clear that fake news is news that is not factually correct, but this is still too broad to be able to actually address the problem that needs to be solved. With his definition, it could also mean that fake news is an article with an accidental mistake in it. This is unfavorable, but not the big problem that needs to be solved. To make the definition narrower, also something about the intention should be mentioned in the definition. The motivation for creating fake news can be divided into three categories, namely for political reasons, as a way to earn money, or to have fun [7]. From all these motivations can be concluded that the reason for creating fake news is always for the writer's personal gain, and thus always intentional.

Besides looking at intention, also something should be mentioned about where the news comes from. If a satirical site publishes an article that is not actually factually correct, it should not be classified as fake news. Only when a publisher also publishes factual correct information, like a news site, or social media, the term fake news can be used. From these stricter specifications, a more narrow definition of fake news can be defined.

Definition 2. (*Narrow definition*): Fake news is intentionally false news published by a news outlet.

2.2. Fake news detection methods

In the last number of years, there has been already a lot of research done on the identification of fake news, to prevent people from making the wrong choice due to misinformation. However, as fake news is still around today, it shows that there is not a good solution yet [8]. Due to the scale of the problem of fake news, it is of utmost importance to try to find a possible solution, but to do this, it will be necessary to get insight into methods that have shown potential already. Therefore, the main goal of this literature review will be to give insight into the question: what proven methods have been found to detect fake news?

When looking at fake news detection, the ways of detecting fake news can be divided into two groups, human intervention, and the use of algorithms [9]. In this literature review, a few

algorithms that have been successful will be described. This will be divided into two sections, where the first section is about detection based on linguistic features, and the second section is about detection based on network analysis. Due to the importance of automation, because of scale and reliability[?], the section about algorithmic detection will be the main part of the review. However, something will also be said about the use of human intervention, in the third section. After that, a conclusion to the research question will be given.

2.2.1 Linguistic analysis

When looking at the algorithms, they too can be divided into two categories. One of these categories is looking at linguistic features. In a recent research about analyzing micro blogs, the way words were used together, could determine if a post were credible or not credible [10]. To see if something were fake, they looked at many different combinations of words, and when the combination was a bit weird, it would classify as fake. With all these different combinations extracted from the text, the researchers were able to train a machine learning program, which made it possible to run an automatic credibility assessing method, and thus mark articles as fake or not. Besides looking at the combination of different words, also determining the type of words that are used, can be a way to classify if an article is fake or not. It turns out that, if many words are used, that indicate for example exaggeration, falsification, deception, or omission, then there is a good chance that the text contains false information. In a recent research, researchers found that creating a model using these words that indicate the fakeness of a text, they were able to create a model with up to 74 percent accuracy in detecting fake news [11]. Another way of doing a linguistic analysis, is not just looking at the words, but also looking at the sentence they are in, so called deep syntax. The use of deep syntax has also shown a lot of promise in the fight against fake news [6]. With deep syntax, the structure of a text is analyzed, meaning that they look at how a sentence is written, and how it fits in the text. This is done by transforming the sentences into a set of rewritten rules, which in turn are rewritten into a parse tree. From this parse tree, a probability of fakeness can be calculated, and this probability seems relatively accurate. To give an indication, a recent research showed that this method has shown a detection rate of 85-91 percent accuracy [12].

2.2.2 Network analysis

The other category where an algorithm has shown a lot of promise is network analysis. In this case the link between different datasets is analyzed. One implementation of this, is a research about fact checking on Wikipedia [13]. In this research, a graph was created, and each node was a factual statement, extracted from Wikipedia. Statements that were more consistently true, received a higher ranking then statements that were not consistent. With this ranking, an algorithm could run, and determine if an article on Wikipedia was fake or not. For this research, the scope was only the Wikipedia website, which is still relatively small, but due to the success in detecting truthfulness on Wikipedia, it also shows a lot of promise in the real world. Another approach of using network analysis, was research about determining if a message posted on social media was fake, by looking at the users who liked the post [14]. The research used a combination of two classification techniques to determine the likelihood of a user liking fake posts. One technique was creating a model that could make an accurate guess about the likelihood of a user to like fake news, based on the user's previous liked posts. The other technique was based on a Boolean crowd-sourcing algorithm that could help when there was not enough information about the user.

With these classifications, it was possible to create a machine learning program, and even with a relatively small dataset, they were able to obtain a classification accuracy exceeding 99Finally, another way of using network analysis, is to look at the amount of hits an article gets when queried in a search engine [15]. Depending on the amount of hits it gets, it makes assumptions about an article. If the amount of hits does not exceed a threshold value, it gets classified as fake, and if it does, it is assumed the article is real. This method has shown promise and is relatively easy to implement. However, if a fake article has spread too far over the internet, it is less likely that the analysis still works.

2.2.3 Crowdsourcing

Next to text analysis, there is also another method to see if an article is fake. This method is using an algorithm derived from crowdsourcing. Crowdsourcing is an interesting method because it lets the users identify an article as fake. Based on what the users think of the text, an estimation about the correctness of the article is made. In this review, it was earlier mentioned that a type of network analysis could be done by determining who liked a post. This was network analysis, because the text was compared with other sources, namely the type of users. However, this was also a way of crowdsourcing. Using crowdsourcing has shown different potentials. On the one hand, studies show that there is a potential for using crowdsourcing for detecting fake news[16], but when we look at the real world, it will be very difficult to implement, due to the fact it is not completely automated. Crowdsourcing still requires input from a human, and this is the main problem. A human cannot always be certain that an article is correct or not and will probably not be bothered by reporting if an article is fake. In fact, it might even lead to an increase in fake news since people might mark actual facts as fake.

2.2.4 Conclusion

In this review, the research question was about what the proven methods are to detect fake news. It can be concluded that there are already a lot of methods, and they all reach a relative high accuracy. In this review, it was found that the use of algorithms was a way to detect fake news, and these algorithms could detect fake news in two ways, by doing a linguistic analysis or by doing a network analysis. It also showed that the use of crowdsourcing was a method used to detect fake news, but this showed less potential.

2.2.5 Discussion

In this review, it becomes clear that there are a lot of detection methods already, but there are still some problems, since fake news is still a problem in today's world. The problem here lies in the possibility of the identification factors changing. When trying to determine if an article is fake by looking at linguistic cues, it assumes the writer of the article will write in a specific way. But once the fake news writer finds out what makes his article classify as fake, he can change his style of writing, and suddenly the article is not classified as fake anymore. This leads to some sort of cat and mouse game, which means the research on detecting fake news, needs to stay ahead in order to be effective. The same problem also partially holds for network analysis. Again, if the writer of the fake article can find out what makes his article classify as fake, he can adapt, by for example posting the same content on different websites, and the article will now be classified as real. However, the problem here is not as big as with linguistic cues, because it is much more difficult to fool a network analysis. If a majority disagrees that something is fake, it will take much more work to make an article be classified as real.

2.3. Related work

Due to the fact that fake news is a big problem, there is already a lot done, to try to stop it. In this section, several topics related to the detection of fake news and the prevention of fake news spreading, will be discussed.

2.3.1 Fact check websites

Over the last few years, a number of fact check websites have popped up, where fact checkers analyze statements, as another method to prevent the spread of misinformation. To name an example, there is politifact¹, a website that checks statements on social media that are about politics. Besides politifact, there are a lot of other fact checking websites, for example: factcheck.org², the Washington post fact checker³, snopes⁴, Full Fact⁵, truthorfiction⁶ or gossipcop⁷.

MENU Q	PØLITII The Poynter In			
	Jeremy "Spike" Cohen stated on January 22, 2021 in a Facebook post Says the Biden administration labeled Libertarians terrorists.	False Plane a mir-	Editor's Picks Jake Tapper stated on January 13. 2021 in remarks in a CNN breadcast The second presidential impeachment of Donald Trump was "It the most bipartisan impeachment in American history"	
Q	Instagram posts stated on January 21, 2021 in an Instagram post National Guard members turned their backs on Biden's motorcade on Inauguration Day. By Samenthe Putterman - January 22, 2021	FALSE Printed and O	Usual Jacobson - January 13, 2021	
	 Video shows Biden saying "salute the Marines" after hearing the words in an earpiece before his inauguration. By Bill McCarthy - January 22, 2021 	FALSE Participant Control of the second	By Amy Sherman - January 8, 2021	

Figure 1: Politifact.com

¹https://www.politifact.com ²https://www.factcheck.org ³https://www.washingtonpost.com/news/fact-checker/ ⁴https://www.snopes.com ⁵https://fullfact.org ⁶https://www.truthorfiction.com ⁷https://www.gossipcop.com

2.3.2 Report button

One of the methods that is often used on social media sites, is the ability to report something, when an article does not seem right. Next to the text will be a button, where you could notify the company behind the website, that the message you read might potentially be fake. The company then uses an algorithm, hired fact-checkers, or a combination of the two to determine if something is fake news. In the case that the content of the text is not actually true, the text might be removed, or made less likely to be seen by users.

Help Us Understand What's Happening ×				
What's wrong with this post?				
 It's annoying or distasteful Examples: pointless stories, memes or viral images, about someone or something that bothers me 				
 It's pornography Examples: nudity, sexual arousal, sexual acts 				
 It goes against my views Examples: makes fun of my personal values, religion or politics 				
It advocates violence or harm to a person or animal Examples: graphic injury, terrorism, or animal abuse				
 It's a false news story Examples: purposefully fake or deceitful news, a hoax disproved by a reputable source 				
See more options				
Back				

Figure 2: Reporting fake news on facebook

2.3.3 Verification

Another thing that has been implemented by companies, to stop the spread of misinformation, is verifying user accounts. People on social media, that have a lot of followers, will receive a verified mark. Before this was normal to do for these users, it meant that if person A has the same name as person B, it can spread information in the name as person B, in the name of person A. If this person A would be an influential figure, it might lead to the spread of fake news. The verification mark makes it more clear than when person A says something, he is actually the person who said it. Of course, someone who has a verification mark, can still spread misinformation, like Donald Trump for example. However, this verification does make it more difficult for people who are trying to spread fake news.



Figure 3: Account verification on twitter

2.3.4 Teaching programs

Another method, to prevent the spread of fake news, is to try to inform people better, on how to know if something is fake or not. More and more schools see the importance of critical reading in today's world, and with the help of a few nonprofit organizations, they try to prepare their students for the online world. One of these organizations that helps with this is the news literacy project⁸. This organization focuses on helping students to tell the difference between fact and fiction. In the end, trying to teach people how to detect fake news, might be the most promising solution, due to the fact websites would not need to censor articles.



Figure 4: Website of the news literature project

2.4. Determining credibility

In this thesis, the goal is to use credibility to make an assumption about an article being real or not. Therefore, it is also important to define what makes an article credible. From previous research, it was found that an articles credibility depends on two things, the articles content, and the source[17]. As was already mentioned in the introduction, to analyze the content is very difficult, so the choice was made to use the source as a way to determine credibility. The source can be split into two different article attributes, namely the author of the article, and the publisher of the article. To also try to use the article content to determine credibility, also references could be looked at, since listed references can have a real impact on the credibility. For example, an article that states climate change is real, and references something that explains why, is a lot more credible than an article that states climate change is a hoax for example, without any proof.

⁸https://newslit.org

3. Methods and techniques

In order to find a possible answer to the research question, the Creative technology design process was used[18]. In figure 5, an illustration of this process can be found. It starts by asking the design question in the ideation phase, and from here, some potential ideas that could answer this question are thought up. Then, in the specification phase, the best idea is chosen, and will be defined more clearly, in order to be sure, the idea is still good. After being sure the idea is good, the idea is fully realized in the realization phase. Finally, the implemented idea is evaluated to determine if answers the design question. It is also important to mention here that in every phase, it is possible to go back one phase, to reevaluate the problem.



4. Ideation

As already mentioned in the introduction, in this thesis, the goal is to find out if article credibility can help with fake news detection. If the article is credible, the assumption can be made that the article is also factually correct. This assumption could be very useful to only allow credible articles on a website, or to give credible articles a "certified as true" mark.

4.1. The algorithm

In order to determine article credibility, three different article attributes can be looked at. First, there is the author, or authors. Second, there is the publisher, and third, there are the listed references in the article. All these attributes could help to determine credibility, and in this subsection, several ideas of how these attributes could be used, will discussed. In figure 6, a model for the algorithm can be seen, that is using the discussed attributes.

4.1.1 Authors

One of the first ideas was to use LinkedIn, a social media platform that a lot of people also use as their CV these days. The idea was, to automatically search for the authors listed with the article, on LinkedIn, and use a web scraper to get information about the author, to see if he or she was actually a journalist of some sort. If the author were actually a journalist, it would mean the article would probably be much more credible than an article written by just anybody. However, LinkedIn does not want others scraping their data, so they made it really difficult to do this. Luckily, there were also other websites, containing information about journalists. This made it possible to search for the author on these websites, to determine if they were a journalist or not.

Another idea that came up when trying to determine credibility based on the author, was the number of scientific citations the author had. A good way to find the number of citations, is by using google scholar. To implement the number of citations, a solution was found by using googles custom search API. Querying the API with the author would return a JSON object, from which the number of citations could be extracted.

4.1.2 Publisher

To determine the credibility of the publisher, the idea was to see if the publisher belonged to an actual company, and check if this company was in a category associated with publishing news articles. For companies in the Netherlands, it meant the website of the Dutch chamber of commerce could be scraped, and information could be extracted. For companies in the US, an API to an open-source database could be used to check the publisher's credibility.

Another idea that came up to determine credibility based on publisher, would be by checking them against a database with websites that were known to be unreliable. From another research paper about fake news, a list of 1000 websites were found, that were known to be unreliable. By adding these sites to an owned database, it would mean it would be possible to check if the publisher existed in that database.

Another idea to determine publisher credibility was to also have a database containing websites that were known to be credible. These websites would be credible news sources, for example

websites from newspapers. This list of credible news sources could be found on Wikipedia and publishersglobal. By using a web scraper, the domain names of these credible sources could be extracted and added to an owned database. Then, the algorithm could check if the publisher existed in this database and see if the publisher were credible or not.

4.1.3 References

The references listed in an article could also have a big impact on the credibility. One of the ideas regarding reference, was that the more references there are, the higher the probability is that the article is real. Furthermore, with having a way to determine credibility based on publisher, the same method could be used to determine the credibility of the reference. These two methods could potentially be very helpful to determine article credibility. However, it is not perfect, since credible references listed in the article, that have nothing to do with the content of the article would result in a false positive with regards to correctly stating an article is not fake.



Figure 6: A model of the algorithm

4.2. User interaction

To find a way to determine if an article is real or not is the most important aspect of this research, but in order to help people, other than other researchers, solve the problem of fake news, it is also important to create something that regular people can use. If the algorithm shows potential, it makes sense to create a way for anyone to interact with the algorithm, to determine if they should believe the article that they are reading.

4.2.1 Browser extension

An ideal solution to the problem would be if people could install a browser extension, and the extension would scan each article on the website and determine if it were real or not. If the extension believed the article would be real, it could give it some sort of verification mark, as can be seen in figure 7. There are however quite some difficulties with this. It would mean

automatically finding the listed authors and references, and this difficult, since every website is different.



Figure 7: Concept for a news verification mark

4.2.2 API

Another potential solution would be to create an API and make it publicly available. This would mean website owners could use this API to only let articles that were known to be real, onto their website, or it would make it possible to give articles a credibility mark. Using a credibility mark, it could lead to people only believing credible articles, and it could help against the spread of fake news. The problem however with creating an API would also require work from others, which means it becomes a lot more difficult to actually have an impact.

4.2.3 Website

Finally, another potential idea would be to create a website, where people can check if the article they read, is real or not. It would be something similar to politifact, but whereas politifact relies on humans to determine if the article is real or not, which is quite slow, this website could do the same thing, but faster. A user would be able to put information about the article into the website, and the website would tell what discoveries were made about the article and give a relatively accurate approximation about if the content of the article can be trusted or not.

5. Specification

As previously mentioned, this research makes use of the creative technology design process. In this section of the report, the second step in the process will be described, namely the specification. From the state of the art, and from the ideation, several specifications can be listed.

5.1. Algorithm specifications

For the algorithm, the following specifications are:

- The algorithm should be able to find information about an author.
- The algorithm should be able to find information about a publisher.
- The algorithm should be able to find information about a reference.
- The algorithm should be able to make an estimation of the credibility.
- The algorithm should be able to determine if an article is real, based on the estimated credibility.

5.2. User interface specifications

For the user interface, the following specifications are:

- It should be possible to insert information about the author, or authors.
- It should be possible to insert information about the publisher.
- It should be possible to insert information about a reference, or references.
- The UI should give an indication of the credibility of the algorithm.
- The UI should explain why the article is credible.
- The UI should indicate if the article can be assumed as real.
- Interaction with the algorithm should be publicly available.

6. REALISATION

In this section, the third step of the creative technology design process will be described, the realization. First, the tools used to create the project will be discussed, and after this the creation of the algorithm will be explained. Then, a small section will explain how the data used by the algorithm was gathered. After this, it will be explained how the testing was done, and finally something will be said about creating a user interface, and an attempt to create an API.

6.1. Tools

Early on in the project, it was decided to use C sharp, to write the algorithm. The main reason for this was the author being familiar with the C sharp language. Another reason for choosing C sharp, was because of the .NET framework. The .NET framework makes it possible to share code between project, so the algorithm could be used by an API, the testing program and website at the same time.

To create the project using C sharp, visual studio 2019 was used. The reason for using visual studio was the integration with Microsoft's cloud platform, Azure. Azure has some useful services and does not cost a lot of money. For source control, Azure DevOps was used, and for trying to create a website, an Azure storage account was used. When trying to create the API, Azure functions was used to host the API, and postman was used for testing.

6.2. Credibility algorithm

To determine if an article was real or not, it was necessary to find out if the article was credible or not. So, to do this, an algorithm was created. The algorithm works by giving it an articles authors, publisher, and used references. From this, it calculates a score, and if this score exceeds a certain threshold, the article can be classified as real. In the subsections below, the process of getting this score will be further discussed.

To create the algorithm, a solution was created in visual studio, and a project was added to this solution, with the type being a .NET standard Class Library. This means it is a project that does not run on its own but allows other projects in the solution to call functions created in the class library. This makes it possible to share the algorithm code with multiple other projects.

6.2.1 Author credibility

As mentioned before, to determine if the article is credible, a score needs to be calculated. One part of getting this score, is by finding information about the author, or authors. As was mentioned in the ideation section, the first idea that came up to find information about authors of the article, would be by automatically looking them up on LinkedIn⁹. This would be a perfect place to find professional information about people, for example if they were a journalist of some sort. However, in the past, more people had the idea of gathering information of LinkedIn automatically, and since the owners of LinkedIn did not want this, they implemented measures to prevent web scraping. After trying a lot of possible things to get around these measures, it was decided that even though it would be a gold mine of information, it would be better to focus on something else.

⁹https://www.linkedin.com

As an alternative for using LinkedIn, a website was found, with a database of over 85.000 known journalists. This website, called Journa¹⁰, allows users to follow certain journalist, sending a notification to the user when the journalist has written an article. This means it is in the best interest for journalists, to register on this site. To prevent spam, the website requires for the registration, to use LinkedIn, so not just anybody can add its own name.

The algorithm uses this website to calculate a credibility score. This can be done because it generates a profile page for each registered user. Each profile has its own URL, so the algorithm can recreate this URL, pasting the given author into the URL. Then the algorithm sends out a http request, and if it fails, it means the author is not in that database, but if it does not fail, it means the author is in that database. By existing in the data base, it means the author is most likely a professional, and thus it can be assumed he or she has done its research, and the article will not contain false information.

Beside journa.com, also another database with journalists was used to find more information about the given author, namely presshunt¹¹. Presshunt is another website that list journalists, with over 500.000 listed journalists worldwide. unfortunately, the same trick used for journa.com could not be used here. So, another solution had to be found to extract information from the website. This solution was using a web scraper. At first, some attempts at web scraping were done by using .NET functions, and by trying to use an extension called scrapysharp. These methods unfortunately only allowed for the HTML, not generated by JavaScript, to be extracted. This did not contain the required information, so as a solution, Puppeteer was used. Puppeteer is a headless browser, that can be used to do web scraping, with its most important feature being, that information from JavaScript generated HTML can be extracted. From this extracted information, the algorithm could then determine if the author were legitimate.

Finally, as a way to check for credible authors, that were not professional journalists, a function was implemented that could check the number of citations the author had. This is because an author might still be very credible, when it has a certain number of citations. Here the assumption can be made that an academic will not spread lies around, and thus will be credible. To do this, the google custom search API was used. The API made it possible to search for authors on google scholar and see how many citations the author had. Based on the amount, the algorithm could determine how credible the author was.

6.2.2 Publisher credibility

Next to determining if an article is credible, based on authors, credibility can also be determined by the publisher. The first idea to determine if a publisher was credible or not, was to check if the domain name of this publisher belonged to a legitimate company. For Dutch companies, this could be done by visiting the website of the Dutch chamber of commerce. By using Puppeteer, the domain name could be searched for, and if the search results exactly matched the domain name, and if their description contained something about publishing or journalism, it could be assumed that the publisher was legitimate.

To also make the algorithm work for non-Dutch companies, a database with companies based in the united states was used. This was an open-source database, with a publicly available API. By calling this API, information about the publisher could be found, and this information included what sector the company was in. Unfortunately, this database was removed from the internet, a

¹⁰journa.com

¹¹ presshunt.co

few weeks before the end of project, making it unusable. To try to find a solution, another database with us companies was found. An attempt was made to use Puppeteer to extract information about a publisher from that database, but unfortunately, the website had measures build in to prevent web scraping.

With the method of looking if a publisher was a legitimate company not working, a new solution needed to be found. The decision was made to create an own database, with reliable companies, to prevent losing access to a database again. To create the database, a Cosmos DB database was created in the Azure cloud. The reason for this was the database being free to use. This database was then filled by getting the names of every major news source in each country, gathered from Wikipedia and publisersglobal¹². Since the amount of news sources was over 3000, a web scraping program using Puppeteer was created, to automatically extract the information from the websites, and store it in the database. When executing the algorithm, it could check if the given publisher existed in the database, and if so, return a high credibly score.

As a way to speed up the program, and to be certain unreliable publishers would not be classified as credible, also a database containing known unreliable publishers was created. As with the reliable publishers database, this database was also created as a Cosmos DB database. To fill this database, the findings of a previous research were used¹³. This research had found 1001 websites that were known to be unreliable publishers. So, a small program was created to extract all this information from the CSV file provided by that research, and to put it into an own database. This meant the algorithm could before doing all the other checks, check if the given publisher existed in this database, and if it did, it could immediately stop the publisher check, and return a low credibility score.

6.2.3 Reference credibility

In order to have a third way to check for the credibility of an article, there was also something implemented that allowed for the listed references to have an impact on credibility. To determine the credibility of an article, the number of references was looked at. Up to a certain amount, the more listed references there were, the higher the credibility score. Here the assumption was made that the given references, do actually have to do something with the article. Besides looking at the amount, also the publisher of the reference was also checked, by using the same code that checks the publisher of the article.

6.3. Testing

Once the algorithm was created, it was important to also test if the algorithm worked. The test results will be described later in this thesis, this section describes how the program to test the algorithm was written. In order to do the testing, a console application was written in C sharp. At first, it only allowed manual input, but after deciding on what data set to use, also a way to automatically test was implemented. This meant it could read a .CSV file, extract the author or authors, and publisher, run both through the algorithm, and give an indication about the article being real or not. The program would then put this result into a new .CSV file, along with the reasons for getting the result, and article information.

¹²https://www.publishersglobal.com

¹³https://github.com/several27/FakeNewsCorpus

Credible	Dublichor	Credible	Credible	
Authors	Fublisher	References	Credible	
>= 2	Neutral	>= 0	Yes	
>= 3	Not credible	>= 0	Yes	
>= 0	Credible	>= 0	Yes	
>= 0	Neutral	>= 2	Yes	
>= 0	Not credible	>= 4	Yes	

Table 1: Combinations to determine credibility

To improve the algorithm, some fine tuning was done. The algorithm determined if something were credible, based on a score, and since every given attribute would give a score, the score amount of these attributes and the threshold values could easily be changed, to make the algorithm perform better. Thought testing, it was found that an article is credible when two or more authors are found by the algorithm, or if the publisher is known to be a credible source. For the references, an assumption was made that if the article has four or more credible references, it can also be considered as credible. There are also some combinations that can occur, and determine the article is credible. Combinations can be seen in table 1.

6.4. Creating a user interface

Initially, when first creating the algorithm, the only user interface was through typing the article attributes into the console. However, this made it not very accessible for others to use. So, an attempt was made to create a web app, using the ASP.NET, and the Blazor framework. Blazor allows for a web app to be built in C sharp, instead of having to use JavaScript. This means it is faster, and in the authors opinion, also makes it easier to create the web app. Unfortunately, halfway during development of the web app, it turned out that puppeteer could not be called from the browser. The assumption was made that since a web app created with Blazor is a regular program, running virtually in the browser, would work fine with Puppeteer, but it turned out this was not the case. The only way to get a web app that uses the algorithm, would be though an API, but as can be read in the next section, this was also unsuccessfully because of Puppeteer.

6.5. Creating an API

To make it possible for other programs to use the credibility algorithm, an attempt was made to create an API, or application programming interface. The idea was to use Azure functions, a platform running in the cloud, which is so called serverless. This means every time the API is called, a program starts, returns something, and shuts down again. This usually works quite well, and it costs next to nothing. The algorithm did however not work when deployed to Azure functions, due to the fact the algorithm does real time web scraping, which requires puppeteer. The problem here is that puppeteer will not install on an Azure function, which means it does not work. There might be a way around this problem, perhaps by using Azure app services, which is a traditional server. However, due to the limited time span to work on this project, it was not possible to explore this possible solution.

7. Evaluation

To determine the effectiveness, of using credibility as a way to stop the spread of misinformation, a data set with both fake news and real articles was necessary. This data set would also need to contain the authors who wrote this article and should contain information about where the article came from. Most fake news data sets do not contain this information, but luckily, a data set used by a previous research[19], did contain this information.

Besides the earlier research having a data set that could be used, another reason for using that data set was to compare the results found by the credibility algorithm, against the results of the previous research. This could give a good indication about how well the algorithm works. Unfortunately, the used data set did not contain references, which meant the reference part of the algorithm became a bit obsolete and could not be tested.

During testing, the observation was made that explaining how the algorithm would come to its conclusions, would make the algorithm much more legitimate, and the results much more believable. So, a feature was implemented into the algorithm, that would list all the discoveries it had made about the article, and use this to explain why it came to its conclusion. A screenshot of an explanation for why an article is credible, and thus considered real, can be seen in figure 8.

Author: Jack Shafer Author: Erick Trickey Author: Zachary Karabell Author score: 350 Reference score: 0 The following discoveries were made... Author Jack Shafer exists in the Presshunt journalists database Author Jack Shafer exists in the Journa journalists database Author Erick Trickey exists in the Presshunt journalists database Author Zachary Karabell exists in the Presshunt journalists database Author Zachary Karabell exists in the Journa journalists database Author Zachary Karabell exists in the Journa journalists database Author Zachary Karabell exists in the Journa journalists database Author Zachary Karabell exists in the Journa journalists database

Figure 8: Screenshot of explaining why an article is credible

In order to determine the effectiveness of the algorithm, a metric called the F1 score was calculated. The F1 score is a well-known scoring metric, that is often used to determine effectiveness of fake news detection algorithms. It combines precision, a measure of how often something classified as true is actually true, and recall, a measure of how often something that is actually true, is classified as true. The score ranges from zero being the lowest, to one being the highest. To calculate the F1 score, the following formulas were be used:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN}$$
(2)

$$F_{1}score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$
(3)

To do the testing, 422 articles from a data set gathered in previous research[19] were run though the algorithm. Half of the articles were real news, whereas the other half were fake news articles. During the testing, a realization was made about the credibility algorithm. From the beginning, the goal had been to detect a fake news article within a set of real articles, but this started to make less sense when using credibility. It would make the most sense to use credibility as a way to detect real news, as an article not classified as credible, could still be true. For this reason, a true positive in this test was considered correctly determining if the article was real. From the test results, it was found that there were 128 true positives, 2 false positives, and 83 false negatives. This led to the following F1 score:

$$Precision = \frac{128}{128 + 2} = 0.98 \tag{4}$$

$$Recall = \frac{128}{128 + 83} = 0.61\tag{5}$$

$$F_1 score = 2 \cdot \frac{0.98 \cdot 0.61}{0.98 + 0.61} = 0.75 \tag{6}$$

Now if the choice had been made to use correctly identifying fake news as fake, as the true positive, a higher score would be found. It would mean there would be 209 true positives, 2 false negatives and 83 false positives. This would lead to the following F1 score:

$$Precision = \frac{209}{209 + 83} = 0.72 \tag{7}$$

$$Recall = \frac{209}{209 + 2} = 0.99\tag{8}$$

$$F_1 score = 2 \cdot \frac{0.72 \cdot 0.99}{0.72 + 0.99} = 0.83 \tag{9}$$

With a F1 score of 0.83, the argument could be made that it does well in detecting fake news. As a matter of fact, it is working even better that earlier research, that used the same data set, which got as highest score a score of 0.80, when trying to identify fake news. However, it felt like the wrong choice to take this score as the end result. During testing, it would classify a real article as fake, a bit too many times, indicating that this way would be way too harsh. Being not credible should not mean fake, but it should rather mean, to think a bit more about the mentioned facts, and do some more research, since a non-credible source could still be right. Next to this, there was also the fact that the testing data was limited, and results may differ when using other data sets, and thus no real conclusion could be made.

Because of this, the choice was made to look at how well real news can be detected, as an end result of this thesis. After all, this is also a very important part of the puzzle. To get an F1 score of 0.75 is relatively good, but it could be better. The reason for a part of the false negatives, was due

to the fact that an article would not always have listed authors, or the listed authors would not actually be names. Besides this, another reason for false positives were that a name of an author would not always be found. This is partly due to the amount of data sets used, as it could always be more, but it might also have to do with the fact that some authors do not want their name existing in these journalists' databases, due to privacy for example.

Another reason for false negatives, was because a great number of real articles came from websites that existed in the unreliable websites database. However, it makes sense here that the articles were not classified as real, as the algorithm also classifies them as not credible, and the algorithm uses the credibility to determine if an article is real or not. If the real articles were to be published on neutral websites, it would mean less false negatives, which would then positively influence the F1 score.

8. Conclusion

When starting this project, the goal was to make a step forward in finding a way to stop the spread of misinformation. From the state of the art, it was found that already a lot of research was done about this subject, but there seemed to be no good solution for the problem yet[8]. From the state of the art, it also became clear that a lot of this research was focusing on trying to find patterns. The problems with patterns are however that it leads to a cat a mouse game. New patterns need continuously be discovered, as the way of writing fake news articles changes, to avoid being detected by one of these patterns. However, there was also another way found, to determine if an article is real or not, by comparing the article with other data sets. One way of doing this, is by looking at the credibility of an article.

For this thesis, the research question was, "What is the potential of using an articles credibility to determine if an article is fake news or not fake news?" To answer this question, the results from the evaluation can be looked at. As was already mentioned in the evaluation, the choice was made to focus on detecting real articles, instead of focusing on detecting false articles, since a non-credible source might still be correct, and testing was done with limited data. Getting an F1 score of 0.75, this research can be considered a small success, and shows that there is some potential for using credibility in the fight against fake news. When comparing this result with results from another research, that also uses credibility to detect misinformation[19], the scores are relatively similar. The earlier research found a F1 score of 0.80, and even though the F1 score found in this project was lower, it is still quite close, and thus shows potential.

Another reason for having potential, is that being able to detect real news, could help people to think more critically about the article they just read. By marking true articles as real, with some sort of verification mark for example, people would probably be more inclined to believe this article, than an article that is not marked as real. So if an article appears without a real news verification, people might start to think a bit about the mentioned facts, rather than automatically assuming the article is true. This critical thinking might in turn lead to a decrease in the spread of misinformation.

In conclusion, the argument can be made that there is a potential for using credibility, to determine if an article is real. It does needs to be mentioned that testing was done with only one data set, due to the data set requiring information about the article's authors and publisher. Next to this, there is also the fact that not all credible authors and publishers exist in accessible databases yet, and thus not all credible articles will be classified as real. However, even though these problems are there, the algorithm could still be very useful, and have a positive impact on stopping the spread of misinformation.

9. Future work

As mentioned in the conclusion, the use of credibility has shown some potential. However the algorithm is not always certain about the credibility. This is mostly due to the fact that it uses only a limited amount of data sets to determine an articles credibility. A potentially big increase in accurately determining if an article is credible or not, would be if the algorithm would use more data sets to get information from. So, some potential future work could be to find more data sets, and make it possible for the algorithm to use these data sets.

Next to this, as was also already mentioned in the conclusion, the testing that was done, was relatively limited. Some potential future work could be to find more data sets, containing authors, publishers, and references, or perhaps create an own data set by webscraping real and fake articles. This would give a much more certain verdict about how big the potential is.

Besides adding more data sets and doing more testing, it would also be nice to get the API working in the future. As was mentioned in the ideation, being able to tell if an article is real or not, could allow for some sort of verification of the article. By making the the API publicly available, it would allow other sites to get information about the articles they have on their website, and determine if they are definitely real. If this were the case, it could give the article a verified mark, similar as how twitter for example verifies user names. Furthermore, next to the fact that it could be used verify articles, by creating an API, the algorithm could also easily be used in other fake news detection programs.

References

- C. Bond and B. DePaulo, "Accuracy of deception judgments," *Personality and Social Psychology Review*, 2006. [Online]. Available: https://journals.sagepub.com/doi/10.1207/s15327957pspr1003_2
- [2] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," *Journal of Economic Perspectives*, vol. 31, no. 2, pp. 211–36, May 2017. [Online]. Available: https://www.aeaweb.org/articles?id=10.1257/jep.31.2.211
- [3] N. Kshetri and J. Voas, "The economics of "fake news"," *IT Professional*, vol. 19, no. 6, pp. 8–12, 2017.
- [4] A. Kucharski, "Study epidemiology of fake news," *Nature*, vol. 540, no. 7634, pp. 525–525, 2016. [Online]. Available: https://www.nature.com/articles/540525a#citeas
- [5] V. L. Rubin, Y. Chen, and N. K. Conroy, "Deception detection for news: Three types of fakes," *Proceedings of the Association for Information Science and Technology*, vol. 52, no. 1, pp. 1–4, 2015. [Online]. Available: https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/pra2.2015. 145052010083
- [6] J. Golbeck, M. Mauriello, B. Auxier, K. H. Bhanushali, C. Bonk, M. A. Bouzaghrane, C. Buntain, R. Chanduka, P. Cheakalos, J. B. Everett, W. Falak, C. Gieringer, J. Graney, K. M. Hoffman, L. Huth, Z. Ma, M. Jha, M. Khan, V. Kori, E. Lewis, G. Mirano, W. T. Mohn IV, S. Mussenden, T. M. Nelson, S. Mcwillie, A. Pant, P. Shetye, R. Shrestha, A. Steinheimer, A. Subramanian, and G. Visnansky, "Fake news vs satire: A dataset and analysis," 2018. [Online]. Available: https://doi.org/10.1145/3201064.3201100
- Y. Ishida and S. Kuraya, "Fake news and its credibility evaluation by dynamic relational networks: A bottom up approach," *Procedia Computer Science*, vol. 126, pp. 2228 – 2237, 2018.
 [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050918312018
- [8] A. Bondielli and F. Marcelloni, "A survey on fake news and rumour detection techniques," *Information Sciences*, vol. 497, pp. 38 – 55, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0020025519304372
- [9] A. Figueira and L. Oliveira, "The current state of fake news: challenges and opportunities," *Procedia Computer Science*, vol. 121, pp. 817 – 825, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050917323086
- [10] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," 2011. [Online]. Available: https://doi.org/10.1145/1963405.1963500
- [11] D. S. Appling, E. J. Briscoe, and C. J. Hutto, "Discriminative models for predicting deception strategies," 2015. [Online]. Available: https://doi.org/10.1145/2740908.2742575
- S. Feng, R. Banerjee, and Y. Choi, "Syntactic stylometry for deception detection," pp. 171–175, 2012. [Online]. Available: https://www.aclweb.org/anthology/P12-2034.pdf
- [13] G. L. Ciampaglia, P. Shiralkar, L. M. Rocha, J. Bollen, F. Menczer, and A. Flammini, "Computational fact checking from knowledge networks," *PloS one*, vol. 10, no. 6, 2015.
 [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone. 0128193

- [14] E. Tacchini, G. Ballarin, M. L. D. Vedova, S. Moret, and L. de Alfaro, "Some like it hoax: Automated fake news detection in social networks," 2017. [Online]. Available: https://arxiv.org/abs/1704.07506
- [15] D. K. Vishwakarma, D. Varshney, and A. Yadav, "Detection and veracity analysis of fake news via scrapping and authenticating the web search," *Cognitive Systems Research*, vol. 58, pp. 217–229, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/ S1389041719301020
- [16] L. De Alfaro, V. Polychronopoulos, and M. Shavlovsky, "Reliable aggregation of boolean crowdsourced tasks," vol. 3, no. 1, 2015. [Online]. Available: https: //escholarship.org/uc/item/1fz8s2tv
- [17] D. Marchionni, "Journalism-as-a-conversation: An experimental test of sociopsychological/technological dimensions in journalist-citizen collaborations," *Journalism*, vol. 16, no. 2, pp. 218–237, 2015. [Online]. Available: https: //doi.org/10.1177/1464884913509783
- [18] A. Mader and W. Eggink, "A design process for creative technology," 2014. [Online]. Available: https://www.designsociety.org/download-publication/35942/a_design_process_ for_creative_technology
- [19] N. Sitaula, C. K. Mohan, J. Grygiel, X. Zhou, and R. Zafarani, "Credibility-based fake news detection," in *Disinformation*, *Misinformation*, and *Fake News in Social Media*. Springer, 2020, pp. 163–182. [Online]. Available: https://link.springer.com/chapter/10.1007/ 978-3-030-42699-6_9