

Gooi, Hayden (2025) Fake news detection using perceptual hashing algorithms and multimodal logistic regression within a blockchain system. Masters thesis, York St John University.

Downloaded from: <https://ray.yorks.ac.uk/id/eprint/12548/>

Research at York St John (RaY) is an institutional repository. It supports the principles of open access by making the research outputs of the University available in digital form. Copyright of the items stored in RaY reside with the authors and/or other copyright owners. Users may access full text items free of charge, and may download a copy for private study or non-commercial research. For further reuse terms, see licence terms governing individual outputs. [Institutional Repository Policy Statement](#)

RaY

Research at the University of York St John

For more information please contact RaY at ray@yorks.ac.uk

Fake news detection using perceptual hashing algorithms and multimodal logistic regression within a blockchain system



Hayden Gooi 199031519

**A dissertation submitted to York St John University in
accordance with the requirements for the MSc degree by
research in Computer Science**

August 2025

Contents

List of Tables.....	5
List of Figures	6
Abstract	7
Keywords.....	7
Declaration	8
Copyright	9
Acknowledgements.....	10
Chapter 1: Introduction of the research	11
1.1 Background of Fake news	11
1.2 Problem statement and the research objectives	12
1.3 Research questions	13
1.4 Significance of the study	13
1.5 Scope.....	13
1.6 Structure overview of dissertation	14
Chapter 2: Literature Review	16
2.1 Online News Media	16
2.2 Artificial intelligence.....	18
2.3 Fake News	19
2.4 Characteristics of Fake News	21
2.5 Sources of Fake News.....	24
2.5.1 Fake News made by Machines	24
2.5.2 Fake News made by Humans	25
2.6 Detection of Fake News	26

2.6.1 Detection using manual checking	26
2.6.2 Detection using NLP for content analysis.....	27
2.6.3 Detection using Blockchain	29
2.7 Perceptual hashing.....	35
2.8 Regression algorithms	36
2.9 Summary of Literature	40
Chapter 3: Methodology	42
3.1 Introduction.....	42
3.2 Hypothesis.....	42
3.3 Research Approach	42
3.3.1 Type of research.....	42
3.3.2 Dataset preparation	43
3.3.3 Sampling method	44
3.3.4 Variables.....	44
3.3.5 Research stages.....	45
3.4 Methodology Design	46
3.4.1 Collection and definition of data.....	46
3.4.2 Perceptual hash generation	48
3.4.3 Blockchain implementation	50
3.4.4 Regression model classification	53
3.5 Evaluation metrics.....	56
3.6 Ethical considerations	58
3.7 Discussions and Limitations	59
3.8 Potential threats and attacks	61

Chapter 4: Experimental and Validation Results.....	63
4.1 Experimental Setup	63
4.2 Experiment and Validation Results	64
4.3 Adversarial scenarios	71
4.4 Error Analysis and Observations	72
4.5 Comparison of system with existing solutions.....	73
4.6 Summary	74
Chapter 5: Conclusion	76
5.1 Conclusion	76
5.2 Achievements and Impact.....	77
5.2.1 Achievements.....	77
5.2.2 Impact of study.....	78
5.3 Future scope.....	78
5.3.1 Advanced perceptual hash techniques	79
5.3.2 Larger and more diverse training datasets	80
5.3.3 Hybrid model architecture	80
5.3.4 Semantic analysis of text.....	81
5.3.5 Implementation on web browsers.....	81
5.3.6 Threat mitigation.....	82
References.....	83
Appendix	88

List of Tables

Table 1: Data structure table	46
Table 2: Probability spread example	55
Table 3: Comparison of TPS for existing systems	61
Table 4: Distribution of news articles	63
Table 5: Results from news classification of Content	64
Table 6: Total distribution of true/false values	66
Table 7: Performance metrics of all groups (Content)	67
Table 8: Results from news classification of Headlines	68
Table 9: Total distribution of True and False values	69
Table 10: Comparison of original and paraphrased headlines	71
Table 11: Comparison of performance metrics between solutions	74
Table 12: Comparison of regular and deep hashing techniques	79

List of Figures

Figure 1: Changes in rates of media consumption over the years 2011 to 2019 [7]	17
Figure 2: Number of internet users worldwide from 2005 to 2022 [9]	18
Figure 3: Process of adding a new node to the blockchain [34]	30
Figure 4: Proof of authority architecture created by Chen [36]	32
Figure 5: Schematic representation of workflow from the authority and organisation perspective [38]	34
Figure 6: Breakdown of data into text and image hashes	44
Figure 7: Example records of verified real news	47
Figure 8: Example records of verified fake news	48
Figure 9: Perceptual hash generation	48
Figure 10: Example hash output of text article	49
Figure 11: Example of hash output after alternation	50
Figure 12: Workflow of proposed system	51
Figure 13: Genesis block data	52
Figure 14: Block data of Real news	52
Figure 15: Block data of Fake news	52
Figure 16: Block data of Opinionated news	53
Figure 17: Block data of Partially True news	53
Figure 18: Flowchart of news classification	56
Figure 19: Classification of labelled news (Content)	65
Figure 20: Confusion matrix of true/false values	66
Figure 21: Graph of the Distribution of True/false values	66
Figure 22: Graph of the Classification of labelled news (Headlines)	69
Figure 23: Graph of the Distribution of Fake/Real values	70
Figure 24: Terms of usage for news articles from Reuters	88
Figure 25: Terms of usage for news articles from Reuters	89

Abstract

The rise of fake news poses a significant challenge to public trust in digital media and social media sites. This study presents a novel system for detecting fake news by combining perceptual hashing and blockchain technology to classify articles. The proposed solution stores hashes of news articles on a blockchain to ensure data integrity and immutability. To determine whether a user-submitted article is fake, the system compares its perceptual hash against the stored hashes using Hamming distance and employs a multinomial logistic regression model to classify the article as either real, fake, opinion or partially true. The system's performance is evaluated and compared to existing solutions using metrics such as accuracy, precision, recall, and F1 score, which highlights its efficiency in detecting fake news. Experimental results provided a high system accuracy of 70.25% in identifying fake news. Additionally, the paper addresses the limitations and potential threats that would affect the performance of the solution as well as potential future work and improvements that can be added to mitigate the specified issues, which means that there is potential for this concept to become a reliable tool for fact-checking in the digital age.

Keywords

Fake news, fake news detection, blockchain, perceptual hashing, hashing, logistic regression, hamming distance, multi-class classification, misinformation, disinformation, text classification, python.

Declaration

I, Hayden Gooi, can confirm that this work submitted for final assessment is my own and is stated in my own words. Any use of other literature, ideas, text, figures, or data are properly acknowledged and referenced within the paper within a list of references.

I confirm that, none of the work from this dissertation has been submitted in support of an application for another degree, qualification of this, any other university or other institute of learning.

Copyright

I confirm that the work submitted is my own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material. Any reuse must comply with the Copyright, Designs and Patents Act 1988 and any licence under which this copy is released.

© 2024 York St John University and Hayden Gooi

The right of Hayden Gooi to be identified as Author of this work has been asserted by them in accordance with the Copyright, Designs and Patents Act 1988

Acknowledgements

I would like to thank my supervisor, Somdip Dey, for his guidance and inspiration during the write up of this thesis because without his feedback and advice, I would not have been able to produce work of this quality. I also want to thank Mike O'Dea for his continued support from the beginning of this investigation to the end of it, even after switching university institute. I want to show appreciation for another supervisor, Malak Olamaie, who had also taken time to provide feedback and answer any questions that I had during the project.

I appreciate the time and effort my supervisors have given up in order to help aid my progression in the writing of this study.

Finally, I want to express my gratitude to my friends, Charlie and Ami, and my family who have always shown so much support for me throughout my academic years and provided the encouragement that I needed during any hard times encountered.

Chapter 1: Introduction of the research

1.1 Background of Fake news

With the rapid development of technology and the internet, the way we share information and data online has been revolutionised, this continues to improve and progress throughout each passing day. Due to the increased presence of media sites and advancements in search engines, users are able to have fast access to all kinds of information, which includes news articles and social media posts. This also means that any single user on the internet is able to post and share information very quickly, which, if proven to be false, but is advertised as being legitimate, can mislead people or potentially cause bias in opinion. The internet is made easily accessible and because of this, there is a wide variety of audiences who will browse content online, and these users will come from different backgrounds, ages and technological literacy. Today's children have a much higher exposure to technology compared to any previous generation and many believe that this is because technology and software companies have tailored their products towards preteens and have made a focus on making their product easy to use [1]. Previous research, performed in 2017, has found that 31% of children aged 10 to 18 have shared at least one news story online that later proven to be inaccurate or false [2]. Advancements in technology have also caused the arrival of bots, especially on social media sites, which are self-automated accounts that can generate and spread false information at an alarming rate which can eventually lead to human accounts sharing this information from their accounts through reposts [3].

The spread of fake news can have some serious consequences such as damaging the reputation of an individual or organisation, influencing public opinion which can alter political outcomes, and causing fear or panic by invoking certain emotions through the use of certain language. The presence of fake news in circulation around the internet has eroded users' trust in media and institutions because it presents a challenge to readers who have to determine whether any content presented to them on screen is trustworthy or not.

In the past, the issue of fake news was addressed by manually fact-checking the information but due to the volume and speed at which news is shared today, this proves to be ineffective and inadequate. Recent advances in technology have allowed researchers to develop an automated approach to detecting fake news which relies on artificial intelligence and natural

language processing to analyse the content of online articles. Unfortunately, these methods face limitations and can have difficulty when trying to identify subtle content changes and can struggle with ensuring data integrity.

To address these issues, this research proposes a unique approach which integrates perceptual hashing and regression into a blockchain system to create a secure and reliable method of detecting fake news content online.

1.2 Problem statement and the research objectives

Current solutions may often overlook the importance of securing the content against unauthorised changes to data, and they struggle to adapt to the rapid and continuous evolution of fake news.

This thesis addresses the following problem: How can perceptual hashing algorithms and blockchain technology be integrated to create a system that can detect fake news with high accuracy?

The hypothesis (H_1) of this study is that the combination of blockchain technology, multinomial logistic regression, and perceptual hashing algorithms will be effective and accurate at classifying articles when compared to existing methods.

The null hypothesis (H_0) is that the integration of blockchain technology and perceptual hashing algorithms cannot classify news articles as real or fake any better than existing methods and provide no significant improvement.

The research aims to focus on designing a novel solution that not only detects fake news but also ensures the integrity of news data through blockchain technology, thus making the verification process transparent, secure, and tamper-proof. The paper also aims to apply logistic regression to classify news articles as real or fake.

To achieve this aim, the following research objectives were defined:

- To use perceptual hashing algorithms to generate hash values for every news article record based on the content.
- To create a blockchain system to securely store the data of these articles as well as their hash value for classification and comparison.

- To apply multinomial logistic regression in the solution to use the hamming distance between news hashes to determine similarity and classify whether the content is legitimate or not.
- To evaluate and calculate the final accuracy of the system using metrics such as accuracy, precision, recall, and F1 score.

1.3 Research questions

The study will be directed using the following research questions:

- How can perceptual hashing be effectively used to detect small changes in different news articles?
- How effective is blockchain at detecting fake news articles?
- How can we predict fake news articles using multinomial logistic regression?
- How can we measure the effectiveness of a blockchain solution?

1.4 Significance of the study

The significance of this study lies in its novel combination of perceptual hashing, blockchain, and multinomial regression algorithms to address fake news detection so that it also ensures both accuracy and security of the data. The research that will be undertaken has the potential to impact the way we tackle the issue of fake news and its detection going forwards. Success from this study could mean that news sites and social media platforms can consider adopting a different approach when verifying content, and widespread implementation may restore public trust in online information.

1.5 Scope

Defining the scope of this project is important because it sets out the focus of the work within the limits of what is able to be achieved with the resources available and within the designated time frame.

The scope of this study is limited to detecting fake news in only text-based news articles using perceptual hashing for content comparison. While the system can identify similarities between articles and flag potential misinformation, it may not be suitable for detecting multimedia

forms of news, such as images and videos. The system may also not be suitable for articles that have been reworded slightly or paraphrased with the intent to bypass detection.

One key limitation that may affect our results is the sample size of the dataset, which will be stored on the system and used to train the regression classification model. Developing and expanding a blockchain can cause significant hardware demands, especially with larger data sets. As the number of blocks on the blockchain increases, the storage capacity and processing power needed increase exponentially because each node is required to maintain a copy of the blockchain ledger, and therefore, there is more data to process and validate [4]. These constraints have meant that the sample size of the experiment will be limited and would therefore lack representability of real-world scenarios. Future work could focus on optimising blockchain implementation for less demanding systems, which would allow future investigations to include larger datasets, which would prevent the generalizability of the results from being constrained by the sample size.

Unlike binary classification, which can simply distinguish between true and false information, this project aims to develop a multiclass classification model capable of using four distinct labels to classify each news article. Each record will either be assigned a 'real', 'opinion', 'partially true', or 'fake' label to highlight the content's truthfulness, but it is still important to acknowledge that this classification approach simplifies the complexity of real-world news, where articles may not fit into either of these labels or may have overlapping categories that it can be classified. Rectification would require the implementation of multi-labels or a different classification model that incorporates probabilistic features to capture ambiguous types of content.

1.6 Structure overview of dissertation

The paper is structured as follows:

- Chapter 1: Explains the background of the issue with fake news and the aims of the research to tackle the problem.
- Chapter 2: Literature Review explores previous research in fake news detection, perceptual hashing, blockchain technology, and regression applications. It explores the existing methods and potential other approaches, and details why blockchain was chosen as the most suitable solution.

- Chapter 3: Methodology outlines the technical approach to system development, including the design of the blockchain-based storage and perceptual hashing algorithms. The experimental setup describes the dataset, structure of the data, tools, and testing environment used to evaluate the system. It specifies the results that will be entered into the system and what outputs will be recorded for analysis in Chapter 4.
- Chapter 4: Results and Analysis presents the outcomes of the experiments, including the system's accuracy in detecting fake news. Throughout this chapter, the discussion interprets the results, discussing the system's performance compared to existing state-of-the-art fake news detection methods such as transformer-based models, and outlines reasons for the results obtained as a result of conducting the investigation. An error analysis is conducted to explain the reasoning for false positives and negatives. It also examines the effect of any attempts to manipulate the blockchain and bypass the hashing algorithm.
- Chapter 5: Conclusion and Future Work summarises the research findings and suggests areas for further research and improvement if further work were carried out. It discusses potential threats and attacks that may be conducted to control the blockchain and suggests some solutions to mitigate them.

Chapter 2: Literature Review

There are many methods to aid in the detection of fake news online. With the analysis of previous research in the area as well as the evaluation of existing methods, this chapter aims to highlight which forms of online fake news detection are more efficient or successful than others. Previous work on the topic will also help to point out which techniques of detection are suitable for certain scenarios or uses. The review of papers and studies will deliver insight into how this investigation can be carried out and will also bring a lot of information in this area of computer science, through the reflection of similar studies that have been performed beforehand and published.

2.1 Online News Media

Due to advancements in technology, news and information can be delivered to consumers very quickly and efficiently and any corrections can be made to articles as soon as possible with the aid of the internet as well as handheld devices. The availability of vast amounts of information, kept up to date and constantly increasing, makes the method of obtaining news using online media more desirable than other conventional methods such as from television broadcasts and physical newspapers.

A study on news consumption in the UK conducted by Jigsaw Research [5], found that different groups of people, separated by age groups, will tend to prefer different types of news media. It was discovered that adults, particularly the elderly (65+), aged 24 and above mostly got their news from popular television channels such as 'BBC One' and 'ITV' whereas on the other hand, younger viewers aged 16-24 years old chose online media apps and websites to obtain information about any news. These consisted of well-known sites such as 'Facebook', 'Instagram' and 'twitter' with a sharp increase in the rate of news being received by younger viewers on the app, 'TikTok' over the year 2020 (1%) compared to 2022 (7%). In the near future, the rates of news consumption in physical forms will decrease due to the proportion of younger consumers increasing over time. Statistics shows that internet consumption is ever

increasing over time while TV consumption shows a slow but steady decline within the past ten years [6].

The graph above demonstrates that the use of online platforms is increasing and had overtaken TV media in 2019, showing that in the future, online news media will possibly be the main method of obtaining news for most people.

Accessing the internet has become increasingly easy and convenient nowadays as most people

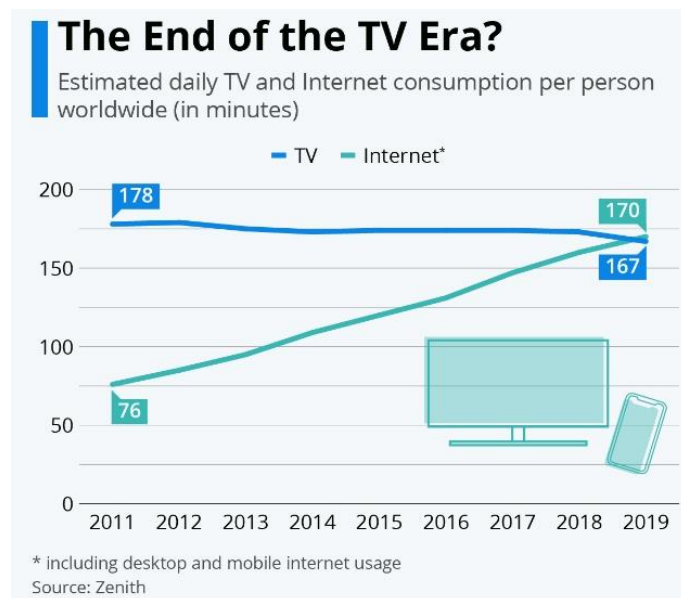


Figure 1: Changes in rates of media consumption over the years 2011 to 2019 [7]

will have easy access to smartphones and computers at any point in their day. In 2005, it was estimated that 1.02 billion people had access to internet and statistics show that this has increased to an overwhelming 5.3 billion active users worldwide [8]. This highlights how future generations will consume mass amounts of online media including news articles.

Number of internet users worldwide from 2005 to 2022 (in millions)

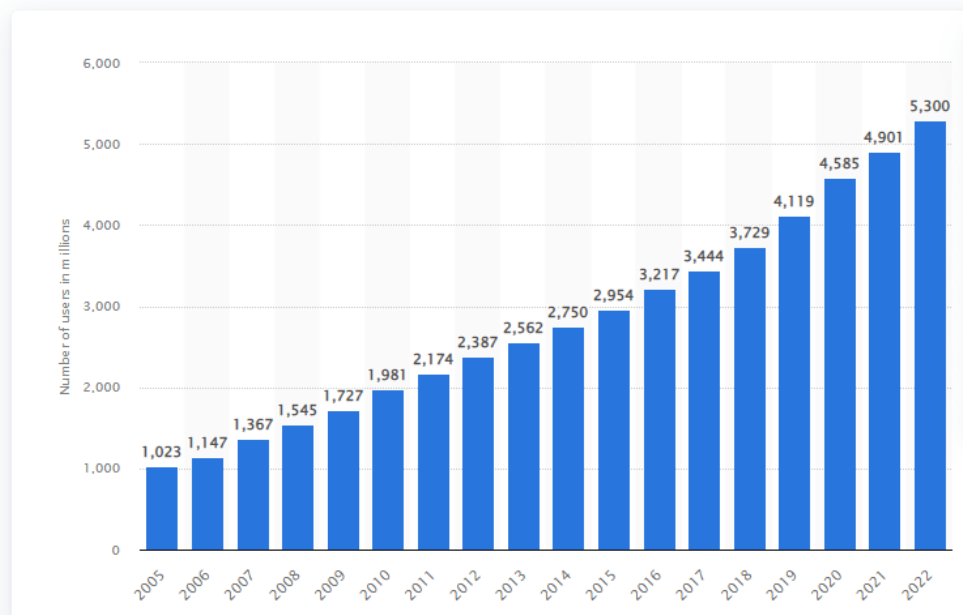


Figure 2: Number of internet users worldwide from 2005 to 2022 [9]

2.2 Artificial intelligence

AI or artificial intelligence can be described as the simulation of human behaviour using computer systems. The main aim of an artificial intelligence system is to mimic human intelligence processes, and an AI may be considered successful if it passes the Turing test by becoming indistinguishable from a normal human [10].

AI works by using a large collection of labelled data, which it can use to find patterns and correlations within the dataset to 'learn' and 'train' itself. This allows it to make predictions in the future on what to output when a similar prompt to the data with which it was trained is entered. For example, an AI can be trained to identify pictures with cars present if given millions of images of a car, so that it can look for similarities between the images and work out the defining characteristics of what a car should look like.

AI is considered to be important because it will be able to perform tasks at a significantly higher level of efficiency than human workers. AI is also cheaper to maintain, does not need to take breaks, and is less susceptible to human errors, which makes them desirable for companies to invest in [11].

Hype for AI has taken off over the past few years with the increase in popularity of sites such as 'ChatGPT', which utilises artificial intelligence to provide a smart chatbot that a user can interact with as if it were talking to another human online. It would be logical to believe that AI will only develop and improve in future years to the point where we can utilise these systems for most services that we use such as customer services or cashiers. Although AI may be smart, it won't be able to replace certain jobs such as graphic design or writing, as these jobs require creativity and coming up with new ideas/perspectives, but AI bases its output off previous data and ideas that already exist from its training data. An article by Garner [12] predicts that "1.8 million jobs will be taken by AI but will create 2 million jobs" by the year 2025.

2.3 Fake News

Baptista [13] defines fake news as being information that is "intentionally designed to mislead" by containing and presenting "false statements that may or may not be associated with real events". A piece of news, such as a web article, can be classified as fake news if it was factually correct at the time it was written/posted but later became outdated or disproven. The main aim of fake news is to influence readers to view a topic in a certain way or to damage the reputation of a person or company. Fake news may be in the form of 'disinformation' or 'misinformation'.

Disinformation refers to deliberately false or misleading information that is shared with the malicious intent to deceive, manipulate public opinion, or behaviour [14]. Unlike misinformation, which can be spread inadvertently or due to misunderstandings, disinformation is intentionally created and published to attain specific goals, such as political, ideological, or financial reasons. Disinformation often involves advanced methods, including the use of fake news websites, made-up evidence, false social media accounts, and incorrect information sources. Its primary aim is to introduce confusion and shape narratives in ways that benefit those who are responsible for the creation of the fake news. An example of this would be a news article created to influence a large number of people towards a certain political view, or it could be a post online that causes discrimination towards a certain group of people.

Misinformation refers to false or inaccurate information that is spread without a deliberate intent to cause harm. It can be unintentional, such as through misunderstandings or errors [14]. Misinformation can take various forms, including rumours, conspiracy theories, and

satirical news sites such as 'The Onion' [15], which aims to be comedic with its content and does not expect readers to take any of its content seriously, and is solely for the purpose of entertainment. It can spread quickly, particularly with the influence of social media and digital communication platforms, which have the potential to cause harm by influencing public opinion and behaviour.

Fake news can be found in many forms on the internet, ranging from social media posts to websites made to look like news webpages, with images that have been manipulated. This makes fake news very hard to spot, especially for those who lack experience with online web surfing, and increases the chances that a fake news article will be shared by readers, increasing the total audience. Hackers will also use a large number of bots on social media sites such as Twitter to spread the same article, making it go viral and giving viewers the sense of reliability due to the perceived number of 'real' people sharing the post.

Online scams are a type of fake news as they present disinformation to their targets in the hope that they can steal personal information. This may be done via fake ads that inform the viewer that they have won money or free gifts in the hopes that they will click on the ad. This may also be present in email inboxes as hackers may use imposter emails to pretend to be banking or delivery companies in order to trick victims into entering their personal information into an online form.

A research paper written by Kumar [16], focuses on the topic of hoaxes that are created and shared on Wikipedia as well as their characteristics. Wikipedia is a huge source of information for many users, so the impact of misinformation on the site is impactful. The paper found that hoaxes had a negative correlation between the number of views per day and the 'survival time' of the hoax online. Having a low number of views means that it is less likely to be detected as fake and reported by a user, but it also results in decreased viewership and distribution. The researchers investigated the articles that spent the longest time online before being taken down to research the key characteristics that made them hard to detect.

Successful hoaxes were found to have been referenced by other articles and pages, making them seem more genuine. Writers who accomplish this tend to plan ahead by publishing numerous articles that all reference each other, giving each other false credibility. Authors of successful hoaxes also used longstanding accounts that had many articles published which created the impression that they were a well-trusted journalist, as opposed to the failed

articles, which were created by accounts that were mainly created a few days before publishing. Future work could build upon this study by performing the same experiment on different media sites, such as social media trends, to see if similar findings are present, which would reinforce the outcome of this investigation.

Previous research [5] shows us that social media platforms score relatively poorly on trust ratings provided by the public, yet a third of all people who use social media actively trust these sites for news. There is no requirement for who can post information online, and with millions of people using the internet every day, it is nearly impossible to fact-check everything that gets posted online, whereas news delivered through television channels have to be well written and validated by professionals before being broadcast. Any information shown on TV has to be correct or else the reputation and integrity of news companies will be damaged, but people can remain anonymous online, so any fake news posted online by reporters doesn't pose many risks.

It is important to address the issue of fake news because exposure to fake news can reduce a person's trust in news sites and media organisations. This kind of issue will affect newer generations of users as they will get most of their news from online sources and their age may impact their ability to recognise real and fake news online due to their literacy skills. Fake news can have a large impact on the health of the public; during the 2019 COVID-19 pandemic, there was a large amount of fake news that circulated within all sorts of media, suggesting that 5G technology was able to spread the coronavirus [17]. Alongside this, anti-vaccination propaganda also caused fear among people who knew little about how vaccines worked, resulting in people refusing the immunisation against COVID-19 and encouraging the spread of the virus [18].

2.4 Characteristics of Fake News

The main characteristic of fake news consists of three main features [19]:

- Factually inaccurate – an article may be incorrect due to having unreliable sources or an older article may have had its information disproven in following years. There are many fact checking tools and sites, such as <https://www.factcheck.org/scicheck/>, which aims to perform evidence based and content analysis to output a reliability score.

- Optimised for sharing – Fake news websites may provide multiple ways to share its content on social media platforms or may even encourage the reader to spread the misinformation written on the site to gain more popularity. They may even include pop-ups or large banners on the side/bottom of the screen asking readers to share or subscribe to any newsletters and updates.
- Meant to distort emotions using prejudice or bias – posts and websites may use eye-catching vocabulary or manipulative writing for sway an audience to perceive a topic or person to be negatively viewed. For example, due to the misconception that vaccines cause autism, the fear of being vaccinated by some individuals may be used by fake news to promote other products than can be substituted for vaccines like essential oils which would cause more harm to health and safety.

Characteristics of Fake News vary between platforms, so it's vital to know what to look for by identifying which platform the news was sourced from. News generated by humans will look differently compared to news by bots and they may differ on the platforms or sites they are found on. News sites will be more likely to be made by humans because the idea of design and capturing attention of readers has to be considered which is difficult for bots whereas social media sites such as Instagram will have tons of bot due to the low difficulty of making a large number of bot accounts and how easy it is to post and share information online.

Disinformation shares several key features that distinguish it from other forms of false or misleading information:

- Deliberate Intent: Disinformation is created and spread intentionally with the purpose of deceiving, manipulating, or influencing viewers.
- Strategic Goal: It is typically published as part of a strategy to achieve specific goals, such as financial or political objectives [20].
- Targeted Audience: Disinformation often targets specific audiences or demographics, tailoring messages to exploit vulnerabilities and biases. For example, badly punctuated fake emails have intent to target the elderly or technologically illiterate population who won't easily pick up on these grammatical errors and figure out the scam. This is the audience which has the highest probability of falling victim to proceeding with the scam.

- **Coordinated Efforts:** Disinformation may involve coordinated efforts by individuals or organizations working together to strengthen false narratives and manipulate public opinion. Cisa [21] gave an example of this type of disinformation when stating how the “Chinese government was suspected of hiring as many as 2 million people” to flood the web in with pro-regime messages. This kind of disinformation manipulates public opinion by overwhelming the web with positive ideas about a certain organisation or individual while simultaneously drowning out an opposing statements or critics.
- **Exploitation of Emotions:** Disinformation campaigns frequently exploit emotions such as fear, anger, or outrage to prompt strong reactions and garner attention, often at the expense of accuracy or truthfulness. This is done through use of powerful and emotional language used to induce emotion from the reader.
- **Fabricated Evidence:** Disinformation may be accompanied by false evidence, such as modified images or videos, forged documents, or misleading statistics, to contribute credibility to false statements and claims. Cisa [21] claims that the development of AI in the present day has meant that deepfakes created by AI algorithms are almost “indistinguishable from real life”.
- **Uncredible sources of information:** articles which display disinformation may have a lack of references/source of information or they may source their information from other fake websites or papers. An article may be fake if the original information has been through many alterations by many different authors.

Misinformation exhibits several characteristics that distinguish it from accurate information:

- **Falsehood:** Misinformation contains inaccuracies or falsehoods that deviate from the truth. Previous studies have investigated the effects of propaganda during the Covid19 pandemic which spread a lot of disease misinformation, endangering the lives of many people who are factually misinformed about the virus [22].
- **Intent or Lack Thereof:** It can be spread intentionally to deceive or mislead, or it can be shared unknowingly due to misunderstandings, biases, or negligence.
- **Expansion:** Misinformation often spreads rapidly and widely, aided by modern communication technologies such as social media platforms, where it can be shared, liked, and reposted quickly via automated software as mentioned by Dallo [23].
- **Lack of Verification:** Misinformation typically lacks credible sources or evidence to support its claims. It may rely on gossip, rumours, or fabricated evidence.

- **Emotional Appeal:** Misinformation may exploit emotions such as fear, anger, or excitement to garner attention and engagement, often aiming to provoke strong reactions rather than provide information. This is done through use of powerful and emotional language used to induce emotion from the reader.
- **Persistence:** Even when corrected, misinformation can continue in public discourse due to reasons such as cognitive biases, ideological beliefs, and the echo chamber effect, where individuals are exposed to more misinformation that reinforces their existing views. They may persist due to lack of information that contradicts or disproves what a reader may believe in.
- **Manipulation:** Misinformation may be deliberately crafted or manipulated to serve agendas, such as political or commercial interests, by selectively presenting information or distorting facts.
- **Variability:** Misinformation can take various forms, including rumours, hoaxes, conspiracy theories, fake news articles, misleading images or videos, and deceptive advertisements.

Evaluation and analysis of the key features of fake news will allow us to find an appropriate solution that is able to effectively examine and spot these characteristics when defining whether a piece of text or an article is genuine or not. Having a clear definition of what makes a piece of text fake also grants us to pick out features of existing methods and combine or improve them to accurately group and categorise new articles.

2.5 Sources of Fake News

2.5.1 Fake News made by Machines

Computer programs, commonly referred to as ‘bots’, are often used on social media to post and spread fake news. These bots act as humans to seem more trustworthy but when a large number of bots are put together on the same platform, they can all be used to create a viral sensation around any topic they choose, such as the promotion of a certain product or exploitation of the stock market [24]. Bots will tend to look for trending topics on sites such as twitter which has a public ranking about which tags are popular at the current point in time. Bots can use these trending hashtags such as ‘#CoronaVirus’ or ‘#ClimateChange’ which are topics that are relevant and would be easy to incite fear or influence readers to believe false

information. Bots may also comment on real posts, sharing fake stories or false facts under legitimate news to make other readers believe in made up reports. BBC Bitesize [25] found that “Cybersecurity company, Radware, has suggested that a 27% increase” in the number of harmful bots online during the 2019 pandemic with the reasoning suspected to be bots exploiting the fears of Covid19.

Most fake news bots will be found on social media sites online. It is estimated that around 5% of all twitter accounts are bot/fake accounts [26]. For example, a bot account may be found on Twitter which spams a lot of posts within a short span of times. Investigating these bots’ profiles may highlight a lack of followers and bio information if the account is suspicious and not creditable. Fake accounts may also include spelling mistakes in their profile, and they tend to stick with default profile images. It can be easy to spot fake accounts sometimes if the account is new and has been created recently to promote a product and they tend to talk about the same topic or a certain viewpoint, using the same hashtags frequently as this is how they are programmed to operate.

With the development of artificial intelligence, AI can help find fake news, however, this technology can also be used to generate fake news at an alarming rate due to the efficiency of AI and this may include fake articles written by AI or deepfakes of famous people such as political leaders. These deepfakes mimic facial movements of people as well as vocal audio to make a made-up video of an individual seem real and creators of deepfakes can choose whatever they want the deepfake AI to say which can be harmful to a person’s reputation if the deepfake is realistic.

2.5.2 Fake News made by Humans

Fake news made by humans tends to consist less of mass social media posting because bots and algorithms can achieve this with more efficiently, but they involve more design-based forms of disinformation such as news sites which have been devised to look legitimate as well as ‘bad ads’ which needs to be eye-catching and convincing enough to trick unsuspecting users follow through with the scam. Humans may also use ‘sock puppet accounts’ which impersonate someone’s identity to mislead public opinion or to gain information that isn’t authorised to the hacker [27]. An example would be a hacker using a fake account imitating a celebrity to gain personal data from users. Humans are more likely to achieve this because it

involves talking and convincing other people online to give information and having a chatbot AI to replicate this, and act like a human on the other end would be difficult.

Satire content will be often written by humans because writers themselves will understand comedy and what they find funny/not funny whereas having an AI generate satire content may lead to confusing headlines or articles that don't make sense or may be offensive to certain groups of people. Due to human error, misinformation can be caused by humans who write news using unverified pieces of evidence or research. News can still be classified as fake even if it has some truth in the text content, despite being intentional or not.

A certain group of internet users, referred to as internet trolls, hold accounts on social media platforms for the purpose of generating negative comments under posts and arguing with other internet users who share opposing views. They will name-call public figures and challenge the credibility of ideas and views that clash with their own and aim to provoke other users to get a reaction for the sake of their own entertainment [28]. Anonymity has been made easy to achieve online which makes it nearly impossible to identify users online most of the time and it means that internet trolls can comment online with less repercussions. Trolls on the internet can damage reputations of individuals by arguing with people who support those individuals which may get those other users to also be nasty and behave in similar fashion.

2.6 Detection of Fake News

2.6.1 Detection using manual checking

One easy solution to detecting fake news is manual checking each post to decide whether it looks false or not. This could be an admin team on a website looking at images to see if they have been altered or checking the context of posts for consistency with similar posts. However, this method is extremely slow and inefficient when considering the amount of information and social media posts is uploaded online every minute so an alternate method is desired where the validation of text can be automated by machine.

There are many websites such as <https://www.politifact.com> or <https://www.washingtonpost.com/politics/fact-checker/> which are free to use websites that aim to fact check any political claims or statements made by elected officials. They give a rating on different statements made in the past, clearly displaying whether the claims made

are correct or if the content is likely to be consisting of lies/incorrect information. The webpage also uses references to support and back up any ratings that it assigns to statements. For non-text-based media, free web services such as <https://reverse.photos> or image search on google allows for the reverse image searching of a picture online. This will allow users to view any similar images on the internet which could allow them to spot any changes that may have been made to an image, if the original image is found by the search engine.

These methods of spotting fake news are mostly done manually and performs checks on one article at a time. This is most suitable for individual use such as a casual reader who is reading an article online and is sceptical about its content. Performing checks for a larger dataset of articles will require alternative methods.

2.6.2 Detection using NLP for content analysis

There are many uses of AI algorithms, and these algorithms can be tailored towards completing different tasks. Machine vision allows the computer to visualise data input through devices such as camera recordings or microphone inputs. This algorithm may be helpful for investigating the design or layout of online web pages which are known to contain fake news to train the AI instead of analysing and comparing the actual content and language used in the articles. This kind of algorithm may find that fake news websites use more vibrant colours and bold eye-catching headlines with the aim of achieving a higher number of foot traffic and to clickbait users into reading the article.

Another main form of detection algorithm is using 'Natural Language Processing' (NLP) which is defined as the processing of human language based on machine learning, using a computer program. NLP is used for services such as spam email detection to sort our inboxes by looking and analysing the heading of emails as well as the text of the email to look for keywords and language that is commonly used in other spam and harmful emails [29]. Using this algorithm, an AI would be able to look at a news article and categorise it into safe news or potentially harmful or incorrect news. Fake news would use specific language which aims to mislead readers into thinking of believing a certain way as well as clickbait consumers.

Aldwairi [30] designed an online web solution which included a tool that would detect and remove any web links that contained fake news webpages from user's search results. This tool was developed using AI NLP which analyses web links provided by the search engine for certain

words that may suggest a misleading tone as well as looking for slang vocabulary which implies that the article may be informally written and less likely to be reliable. The tool also studies the punctuation used in the article and the headlines of the webpage because webpages with a lot of repeated punctuation in the title such as “The Prime Minister does what????!!” will likely be seen as clickbait and be flagged by the AI. Clickbait websites will lure readers in with an interesting title, but the contents of the article may differ from what the headline suggests; this means that anyone who clicks on the misleading headline will see the mismatch of content and click off the site immediately. This metric is referred to as the “bounce rate” and the tool views website with high bounce rates as being fake and misleading. This paper used a formed a database of clickbait URLs to train the NLP AI before calculating attributes using a Python script. This file as then used in WEKA machine learning which aims to validate the solution. The study resulted in a “99.4% accuracy using logistic classifier” which shows that the research was a success, and the tool sorted fake news reliably. This study provided significant advancements in fake news detection but an issue with this solution is that it may misclassify satire web pages as being the same and fake news and remove them from search engines. Satire content may use the same punctuation as actual fake news pages to create humour or entertainment for readers. Future work could incorporate this tool into social media sites to remove web links from posts that contained fake news pages.

Another similar study by Raza [31] used a transformer based deep learning solution which aimed to learn representations from fake news data using an encoder but also used a “decoder to predict future behaviour based on past observations”. This future prediction has been highlighted to be important because a problem in the field of detecting fake news is ‘concept drift’. Many models that are built to look for fake news gets trained using current news data available at the current time, but this data may not be generalised for the future and many labelled samples of verified fake news articles gets outdated very quickly as soon as a new major even occurs. The study aimed to tackle this problem and stated an example of concept drift was a model trained on fake news data before the Covid19 pandemic may have some trouble verifying news after the pandemic occurs. The researchers used a ‘weak supervision module’ so that better predictions can be made when unforeseen news occurs and is inputted into the program. This model looks at many features such as ‘source of news’, ‘headline content’, ‘author’, and ‘publication time’ which when analysed together by the deep learning program, can give a fairly accurate classification of whether the news is fake or not.

When features of the news content are unavailable, it will look for social contexts for decision making such as 'user credibility of the post', the 'number of comments' and the overall 'opinion of the post' from investigating the comments made. The study was considered a success with an F1-score of 74.95% and the final accuracy of this model was calculated to be 74.89% with a precision of 72.40% showing the existence of false positives when the algorithm is run on a dataset. Limitations of the study are that the dataset used was quite small, having 260 news sources, for a training set which meant that the training data may not have covered more recent news or news catered towards many different readers or languages. Training the AI with different social media platforms or with a larger dataset that includes news sources of different countries or target audiences would help improve the accuracy of this solution.

Kumar [16] creates an automatic hoax detection using forest classification to flag Wikipedia articles as hoaxes. They performed the experiment on a sample set of around 12,901 articles consisting of a spread of real and fake news texts. The algorithm is used to flag articles if they are suspected to be a hoax, and the accuracy of these tests are recorded and compared to the accuracy of human validators. The findings of the research concluded that classification algorithms outperformed human validators with an accuracy of 86% to 66% respectively and this was suspected to be due to human bias affecting judgement. This research is impactful because any reader is able to contribute and amend data on any page on the Wikipedia site which presents a strong need to have methods of validating data and preventing fake news from being entered. Wikipedia pages have plenty of links which reference other Wikipedia articles and to improve this algorithm, links present to previously verified pages that have been processed by the system can be considered when evaluating the legitimacy of the page. Pages with links to previously verified pages may suggest that the data is more likely to be truthful and accurate.

2.6.3 Detection using Blockchain

Blockchain is described as 'a secure database that is shared across a network of participants' where the information stored in the blockchain is kept up to date and available to anyone who is within the blockchain to view. Any new transactions or changes can be easily added onto the end of the chain and is always recorded and kept up to date in an account booked referred to as a decentralised ledger [32]. Each block of this chain will contain the participant's ledger

which states a number of transactions that is updated when any new additions are made to the network as well as timestamps and a hash of the previous block.

Any new blocks that are requesting to be added must be sent and validated by every other node on the network. On a public blockchain, this authorisation of new nodes is performed by consensus where the majority of other nodes must agree that the transaction is valid. People who own computers on the network are encouraged to verify transactions through a process called 'proof of work' which rewards users who solve a complex mathematical problem to add blocks to the chain with cryptocurrency [33].

Stored transactions are encrypted using a SHA-256 hashing algorithm which provides it with a unique unchangeable hash. Any new data blocks are appended together alongside older blocks which makes it possible to monitor changes and allows for visibility of past versions of the blockchain [32]. These hashes are used to check integrity of the data and ensure that it has not been changed without authorization because using a hashing algorithm repeatedly on the same input data will always output the same hash so if any changes are made, it will affect the output hash, affecting every other hash all the way to the root of the blockchain network.

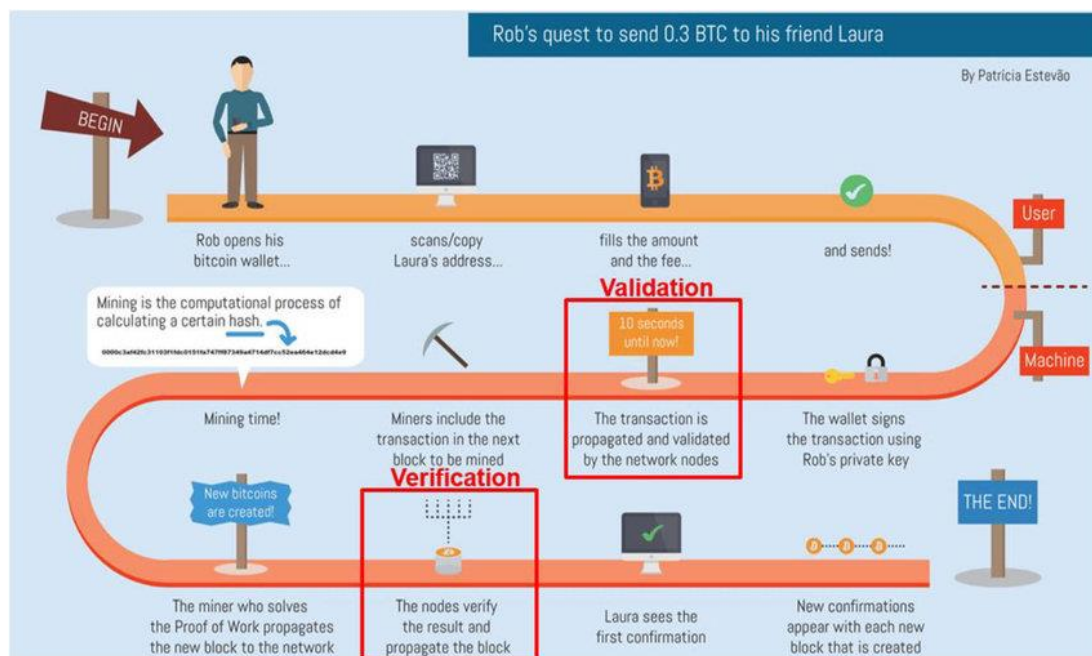


Figure 3: Process of adding a new node to the blockchain [34]

Huckle [35] created a blockchain system which was named "The Provenator" that would allow content creators the ability to store relevant metadata about their creations on the blockchain for authentication and verification later on when retrieved. This was implemented by using the

PREMIS (Preservation Metadata: Implementation Strategies) definitions to record the provenance of digital items onto the blockchain using the help of smart contracts. This is well designed experiment which strengthened the idea that the issue of fake news can be tackled by verifying the origins of stories and images that are shared on the web however, the solution in this paper was a very early prototype. This prototype uses cryptographic hashing on images stored onto the system which means that if two hashes match, the images' origins can be accurately reflected and traced but this means that if there is a slight alteration of an image, an entirely different hash will be produced. As this solution doesn't have any algorithms to filter out any noise, an improvement that could be implemented would include including object classification to help group similar images together to find fakes. Another limitation of this study is that it only tested the prototype on one scenario which doesn't accurately show its effectiveness in different scenarios which may occur in a non-experimental scenario; future work should include more diverse situations where fake news applications would be suitable.

A previous study written by Chen [36] develops a novel blockchain based system that aims to not only detect fake news but also prevent fake news by using a proof of authority consensus algorithm and weighted ranking system to determine the reliability of fake news online. Within this proof of authority protocol, nodes for journalists are set and verified using achievements and characteristics that indicate how the individual is reliable and qualified to practice in the journalism sector. On top of this, a credibility score is used which indicates how reputable an individual or news company stands. This considers various factors such as how many news articles a user submits which is verified to be accurate and vice versa with fake articles. When nodes have a high enough credibility score and have been verified, it can become validators which allows them to help validate blocks and transactions.

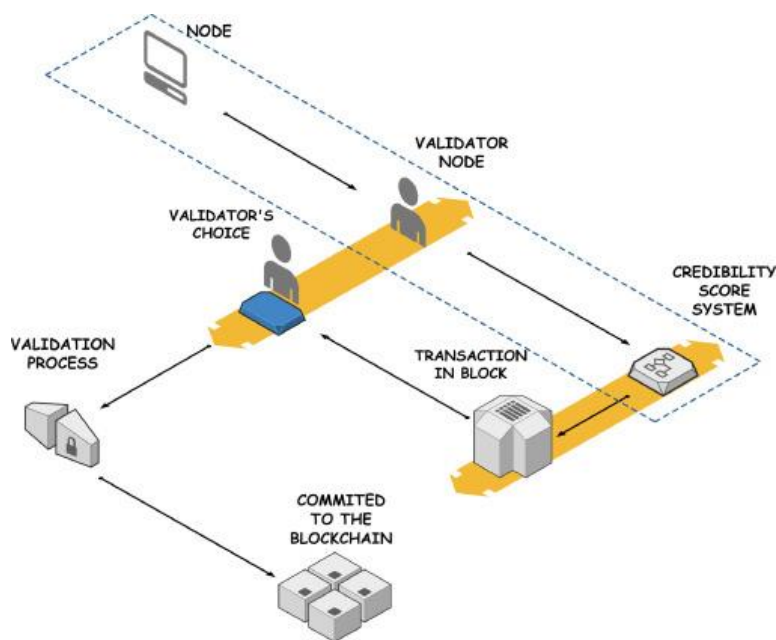


Figure 4: Proof of authority architecture created by Chen [36]

This solution is said to provide a “mechanism of trust” towards news and information posted online and targets multiple types of media that can be stored and used in recorded transactions in validation processes. This paper only included news organisations in the blockchain but to improve this study and make it more representative of how it can perform for various real-life media, this technology can be tested on different forms of media and be expanded to include other news sources and individual journalists.

Paul [37] proposes a similar approach where a blockchain is used and having anonymous validators within the system as to not induce any bias or external pressure. Unlike Chen [36], this study suggests integrating social media onto blockchain and having random users including journalists be requested to verify transactions once they are made onto the network. Multiple validators will judge and assign a correctness value (a number ranging from one to five where five is dependable and reliable news) onto the new transaction and the mean of those scores will be the final correctness value assigned to that transaction. This final score can be added to wherever the news article is shared to show readers the precision of the literature. Professional validators who rate and score many articles accurately can then be rewarded with incentives such as cryptocurrencies like ether. It is suggested that having an anonymous and decentralised system where a large number of validators rate news articles will result in more transparent and constant method of verifying fake news. It isn't stated how many validators will rate each transaction and this would probably be dependent on how

many users of the blockchain exist where having a larger amount would mean more reliability. This means that a few incorrect ratings may influence the mean rating of an article if there are a low number of validators who respond and give a rating, unless outlier removal precautions are put in place which would perhaps remove a few of the extreme scores if they deviated greatly from the mean. Having human validators means that they can understand the context and ethics of a block of text including any tones within the content as opposed to AI methods which may struggle or are more likely to make mistakes when gaining contextual background to an article. These human validators, although can detect bias in news articles, can also be subject to their own biases as their independent views and beliefs and influence the way that they perceive an article compared to one another. In similar solutions which use people as validators, depending on the amount of validators that are used, it can be quite costly as you would need to provide an incentive or payment for these validators for their time and work unless they were willing to volunteer to validate for no cost. Due to the vast amounts of data being shared online on a daily basis, in order for a blockchain solution to be effective at detecting fake news outside of a test scenario, it needs to be able to be easily expanded and this expansion needs to be cost effective. The massive increase in the number of nodes being added to the system as new articles are written every day, creates a need for an increasing quantity of human validators.

Contributions of a previous written paper involves the development of blockchain-based smart contracts using the deployment of the Ethereum blockchain system [38]. All the blocks in the proposed system are linked together in reverse order and can be used to track the source of a transaction which offers traceability. Authorities on the network deploys smart contracts on the Ethereum network and subsequently, an Authority admits access to Authorised organizations using the: ETH organization address, organization name and organization website. When an organization is successfully added on the smart contract a specific ID is assigned to it as well as a status. This status reflects how trustworthy a news source is, and the status can be altered by the Authority if the organization publishes fake news which results in the organization's smart contract access to be revoked as they are no longer deemed as a reliable news source.

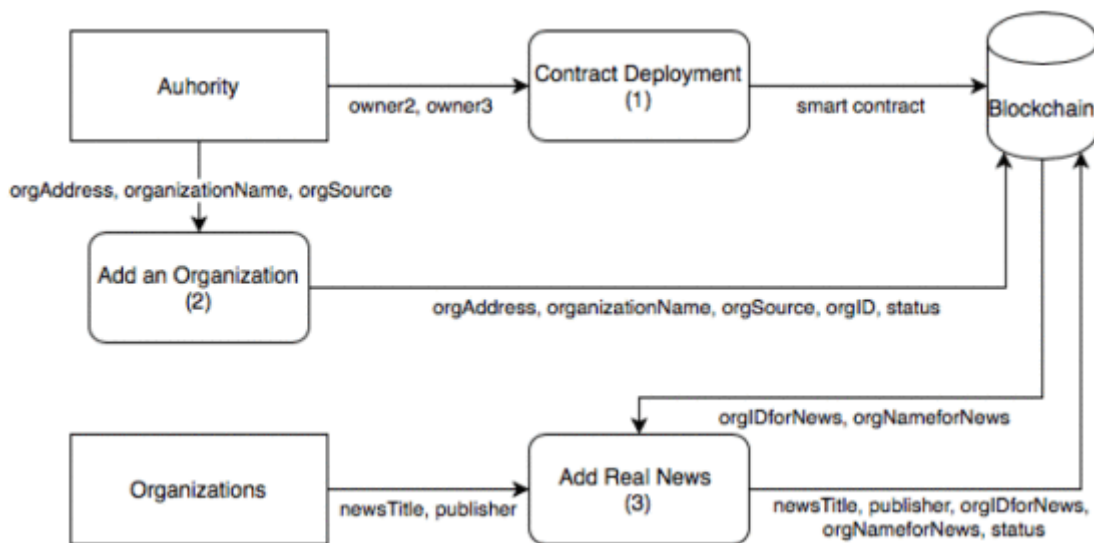


Figure 5: Schematic representation of workflow from the authority and organisation perspective [38]

This allows users to search for a specific news source using the organisation ID to verify if the source can be trusted.

Using previous research that has been conducted [37], we can create a similar blockchain system which uses hashing to verify the source of any new news article to give an idea of the integrity of the content. In order to verify these news sources, they would have to be checked by varying third parties such as web organisations who rate the reliability of websites or authors and fact-checking agencies before they are added onto the blockchain so that they can be labelled as verified or potentially harmful [41]. Any news article on the system will be given a perceptual hash and can be signed by the author which creates a trace of the original creator to each article which cannot be altered or deleted. Metadata can be paired with this to provide additional traits which creates an immutable record to timestamp and verify the source and integrity of the content. This study can then build upon this existing system by having the inclusion of verification badges or links that are assigned to new articles when they are added onto the system. This verification badge will confirm any article which has verified by the blockchain, and readers will be able to use the link to access the record of the article within the blockchain to allow a display of the content's hash, digital signature and any metadata which has been assigned to the article. Users can also compare the hash of the current version of an article and compare it to the hash on the blockchain system to verify if the content has been altered. The idea of perceptual hashing will be explored and added to the system to mitigate the issue of having a light change of the content create a completely different hash, making it

difficult to link the changed content to its origins [42]. The perceptual hash can be matched against its original using its hamming distance which would allow a link or trace to be formed if the content of text is slightly altered. The perceptual hash allows us to link articles to the original source of information and it can be an indicator to highlight whether the content of a record has been altered or not.

When considering using blockchain as the main solution for the detection of fake news, it's important that we analyse and evaluate different frameworks of blockchain due to the fact that different types of blockchain have their own advantages and disadvantages as well as their tailored use cases which make them more applicable to certain situations and scenarios.

Public: A public blockchain is a completely decentralised and open network. No permissions are needed to join the system which means that anyone can join the network, participate in the consensus process, and view the shared ledger. Transactions are visible to everyone, ensuring full transparency and these transactions cannot be altered once they are recorded. These systems are typically very secure as they implement cryptographic security methods as well as consensus algorithms like proof of work and proof of stake. These are most used to cryptocurrencies and decentralised applications as they are highly scalable and is open access to anyone.

Private: A private blockchain is a permissioned protected network controlled by a single organization or device. Access is restricted, and participants need permission to join. These blockchains can be tailored to specific needs of an organisation and allows for more privacy than other networks due to the need to be granted permission to access the blockchain. Private blockchains offer less decentralisation as it is controlled by a single entity and there is potential for abuse and misuse by the controlling entity.

2.7 Perceptual hashing

Perceptual hashing is an essential technique in the domain of content-based similarity detection. Unlike traditional cryptographic hashing, which generates drastically different outputs for minor input changes, perceptual hashing generates hash values that reflect the similarity between two inputs [47]. If two pieces of data are perceptually similar, their perceptual hashes will be closely aligned, making this algorithm particularly useful in identifying similarities in media, such as text, audio, or images.

Perceptual hashing has increased in popularity for plagiarism detection, image copy searching and copyright protection. For example, previous research has explored perceptual hashing for image-based multimedia and concluded that it is effective for identifying slight variations and proved useful for performing searches using an image input [54]. Most existing research in perceptual hashing focuses on visual or multimedia data, which presents an opportunity to incorporate the idea for text-based data.

McKeown [55] is a similar paper which evaluates the effectiveness of perceptual hashing algorithms when matching visual features of photographic images and videos to show similarity. The study mentions how hamming distance scores between similar as well as unrelated data can be used alongside perceptual hashing to discuss and point out similar content online. This investigation is useful because it shows that these ideas can be incorporated into a single solution to help detect related pieces of information.

In an age where misinformation is generated and spread at alarming rates, the novel idea of utilising perceptual hashing and blockchain technology to detect and mitigate fake news proves to be a promising approach as demonstrated in previous studies. This paper aims to build upon existing knowledge and provide new ideas in the field of research by focusing on utilising hashing algorithms to detect structurally similar content online and classify if new pieces of information is predicted to be real or fake by categorising them into one of four distinct labels.

2.8 Regression algorithms

Logistic regression is a widely used machine learning algorithm for binary and multiclass classification tasks. It allows the analysis of binary outcomes with two mutually exclusive labels [56] but can also classify non-binary outcomes where results may fall into more than two distinct categories [59] such as the example studied in this paper which will be articles classified as 'real', 'fake', 'partially true', or 'opinion'. Binary logistic regression models the probability that a given input belongs to a particular class by estimating the relationship between the input features and the binary outcome. The model processes a weighted sum of the input features and applies the sigmoid function to present a probability between 0 and 1, making it perfect for problems requiring classification into two categories.

Shah [57] is a paper that utilises logistic regression for the purpose of text classification. It mentions how similarly to fake news; huge amounts of textual documents are generated which presents a need for a system to sort and organise these files into categories. To address the issue, the author compares the use of logic regression algorithms with alternative methods such as random forest and k-nearest neighbour and uses similar metric such as accuracy to measure the success of each approach. The findings of this study concluded that logistic regression provided the highest accuracy of 97% when classifying data. Analysis of this study showed the same limitations as this paper's; the dataset that the experiment used was purely text-based and improvements would include alternations to the algorithm to include image recognition and classification.

In this study, we will be choosing to use multinomial logistic regression (MLR) instead of binary regression to classify our news articles. The reasoning for this is that binary logistic regression is limited to only being able to separate articles into real and fake news which is not an accurate depiction of how modern journalism is depicted in real life. The target variable in our study will have more than two non-ordinal values and requires the use of more than two labels to be assigned to blocks in the chain, making binary logistic regression unsuitable for the system. Multinomial regression extends the binary case by using a Softmax function [60] to calculate the probabilities of each class and assign the most likely label to each article based on previously learned feature weights from training. Multiclass classification allows the model to consider the relationship of classes at the same time rather than making one-on-one comparisons with only two classes.

Unlike ordinal regression models, which assume a natural order among categories, MLR treats each class as mutually exclusive and is suitable for situations where the labels represent qualitatively distinct types of content rather than numbers on a continuous linear scale. MLR models the log-odds of each class relative to a reference class using a set of linear equations corresponding to the number of classes in order to estimate probabilities of fitting into each class by using the Softmax function. This method allows evaluation of results and performance of the system through metrics such as cross-entropy loss or multiclass F1-score [61]. Previous studies that also utilise natural language processing alongside MLR have successfully applied this concept into text categorisation tasks such as document classification [63] and early fake news detection [64].

Rennie et al. (2003) [62] critically examines the multinomial Naïve Bayes classifier, which is a widely used baseline model for the purpose of text classification for tasks such as spam detection and document categorisation. This study highlighted that multinomial Naïve Bayes is a very fast and easy architecture to implement that performs very well; however, the performance of the solution can degrade significantly when zero-count features are included in the testing set. Inclusion of words in the test set that was not seen in the training set resulted in biased or unstable predications during the experiment. To address these weaknesses, the authors of the paper proposed multiple improvements that could be added such as feature selection with information gain to reduce dimensionality and noise, log-space computation to avoid numerical underflow which is common in Naïve Bayes models, and a technique called 'TF normalisation' which adjusted the term frequency values in text classification so that longer documents don't influence the model by containing more words and this ensured fair comparison. This paper doesn't attempt to replace Naïve Bayes with a new architecture but instead focuses on making refinements and corrections in order to make the model more applicable to real-world textual data. Their investigation proved that even models with theoretical limitations can compare to existing more advanced architectures when the appropriate adjustments and changes are made.

Miotto et al. (2016) [64] introduced a novel approach called 'Deep Patient', which uses unsupervised deep learning on large-scale electronic health records with the aim to predict a wide range of patient health outcomes based on historical clinical data. This paper used stacked denoising autoencoders to learn features in an unsupervised manner which are then used as input for supervised classifiers such as multinomial logistic regression. The authors benchmarked MLR and other conventional classifiers such as decision trees and k-nearest neighbours against their own framework. MLR was used to model multiclass outcomes related to different disease groups and clinical events and MLR served as a strong baseline, providing reasonable performance across disease prediction tasks. This study demonstrated that even in complex and high dimension real-life scenarios, MLR remains a valuable tool due to its ease of training and resistance to overfitting. In the context of this study, the relevance of this work lies in their application of MLR to structured, high stakes decision domains, confirming that while fake news and healthcare differ in subject matter, we can utilise this technique to benefit from explainable, multi-outcome predictions.

The investigation conducted by Ahmed et al. (2017) [65] used a machine learning framework for detecting fake news by using n-gram-based textual analysis and their study focused on extracting linguistic patterns from online news which were then used as features for machine learning classifiers. Models in this paper that were evaluated included Multinomial Naïve Bayes, Support Vector Machines and Multinomial Logistic Regression. In this study, the authors used a dataset consisting of articles which were labelled as 'real' and 'fake' and utilised MLR to classify news as being genuine or misinformation after configuration for only two outcomes in the algorithm. The results of this investigation showed that support vector machines performed with the highest accuracy, but multinomial regression was able to perform competitively. This experiment was limited and would benefit from a more representative dataset which includes a higher number of labels such as 'satire' to utilise the full potential of multinomial regression, and the author could improve results by incorporating semantic analysis to understand the semantic context in the articles. A similar study by Singh et al. (2021) [66] applies a similar logistic regression model and multinomial Naïve Bayes to classify news into the same two distinct categories of 'real' and 'fake' and evaluated the performance based on accuracy, precision, recall and F1-score. Although the problem settings have been binary for these two papers, it lays the foundation for our study with the aim to expand logistic regression to multiclass problems where news will need to be categorised into more specific categories.

Abramovich et al. (2021) [67] presents a theoretical milestone in extending logistic regression from binary classification to multiclass problems. The model aims to address classification problems with many features and multiple output classes which makes this paper highly relevant because they indicate how multinomial logistic regression models can remain interpretable and efficient in high dimensional text classification tasks, such as labelling articles and differentiating between fake, real, opinionated and partially true news. Our solution will aim to use the background of this research to help adapt statistical theory to practical NLP tasks.

Although deep learning models such as BERT [62] have recently outperformed MLR in terms of accuracy on complex semantic content, MLR remains more widely adopted due to its computational efficiency. It is particularly appropriate as an approach where resource constraints and interpretability are prioritised and serves as a strong benchmark for understanding feature significance in multiclass scenarios.

2.9 Summary of Literature

After reviewing the literature on this area of research, it has been concluded that this study will aim to research how fake news can be detected using perceptual hashing and blockchain methods as opposed to other methods such as AI. It will also utilise multinomial logistic regression to help classify news articles using multiple labels. This approach has been chosen for the study because blockchain is a more secure technique of detecting fake news as it can record any changes or alterations made to any piece of information stored in its network and it is very hard to manipulate the network or alter records once it has been recorded into the blockchain without deleting and rewriting a whole chain of data. Blockchain is good at detecting any changes that have been made to an article and it stores information about who made any changes which other detection methods fail to be able to perform. Compared to alternative methods available, blockchain can be a more suitable solution as it is not language restricted, unlike content analysis which can be limited to a certain language and focuses on the properties of the articles such as author. Perceptual hashing will be able to generate hash values for news articles on the blockchain and accounts for small minor changes that may be made to the article without completely changing the entire hash value. A regression algorithm will be able to use the blockchain data to train the model and help the system to classify news articles by comparing the similarity of the perceptual hash values with the labelled dataset of news records.

This paper has chosen blockchain as its news verification method because it can also be easily expanded to grow large by adding new nodes and records onto the network after verification, assuming that computational power and storage resources available are not an issue. This is appropriate for the topic of fake news because all over the world new events are occurring every day, and these events will be recorded and published by both trustworthy news sources and organisations who will manipulate the details of events to sway public opinion or to gain views. The idea of collaboration among users is important for this blockchain approach as everyone in the network is working towards a similar goal and will work together to detect and verify different pieces of news.

Existing literature and past studies have shown more potential in blockchain within this area of study compared to alternative approaches as the idea of blockchain is relatively new and has shown many uses outside of fake news such as crypto and recording transactions for shopping

sites. Developing the idea of blockchains for the intention of preventing the impact of fake news would create a positive impact for online users by helping them browse sites such as news pages more safely. Blockchain can have a huge positive impact by flagging up potentially damaging news in an effective manner. After analysis on the key characteristics which define fake news, it can be concluded that blockchain is the most suitable detection method to find these features because of the properties of blockchains themselves and the functions that they can provide for this type of problem.

MLR offers a scalable, interpretable framework for the assignment of distinct semantic labels to news articles based on textual features and similarity metrics (such as hamming distance from perceptual hashing in our algorithm). This approach is chosen because of its balance in performance, interpretability, and feasibility while being able to maintain methodological clarity so that this experiment can be easily understood and reproduced. Multiclass classification is used instead of traditional binary classification because of the need to categorise news articles into a non-linear scale with multiple different groups which binary classification is not able to accommodate. In our study, we aim to build upon the study proposed by Ahmed et al. (2017) [65] by configuring a MLR model to categorise fake news into multiple categories which includes 'opinion' and 'partially true' instead of primarily focusing on binary classification which would make it more applicable to what real-life news would be structured.

Using the information gained from analysing previous literature, it can be concluded that the research question of this study will be to answer: 'How can perceptual hashing algorithms and blockchain technology be integrated to create a system that can detect fake news with high accuracy?'. As mentioned in Chapter 1, to answer to this question, we need to consider:

- How can we identify a fake news article compared to a real one using regression?
- How can perceptual hashing be effectively used to detect changes in different news articles?
- How effective is blockchain at implementing verification of news articles?
- How can we measure the effectiveness of a blockchain solution?
- Are there any limitations or potential threats to our final system?

Chapter 3: Methodology

3.1 Introduction

As mentioned in Chapter 1, we are aiming to find if the combination of perceptual hashing and blockchain systems can provide a significant improvement when trying to detect fake news.

This research aims to enhance present knowledge and techniques on how we approach the issue of fake news to create a more efficient method of validating pieces of text that are present online.

3.2 Hypothesis

This study hypothesises that blockchain technology, through the use of perceptual hashing algorithms, is an effective tool for verifying the authenticity of news sources and detecting fake news. By comparing the hash of a newly encountered article to existing hashes, the system can confirm close matches, suggesting authenticity if the hash has been previously verified on the blockchain.

It is proposed that the blockchain-based verification system that will be created will achieve high accuracy in authenticating news sources while minimizing false positives and false negatives. The system will be successful if it can achieve an accuracy of 90% or higher. In fake news detection, achieving high accuracy is important because both false positives and false negatives can have serious consequences. 90% is a reasonable target to balance both performance expectations and the critical nature of fake news detection. Similar studies that use regression algorithms have adopted an accuracy target of ~90% to determine the success of the system [57].

3.3 Research Approach

3.3.1 Type of research

This dissertation will employ a quantitative research approach, to measure the effectiveness and accuracy of perceptual hashing in detecting fake news on a blockchain. The accuracy of the solution will be measured by quantifying the rate of false positives and false negatives in

the system. The use of regression analysis can be used to predict the likelihood of content being fake based on perceptual hashing.

3.3.2 Dataset preparation

The primary dataset contains a large variety of real news that has been verified to be true by external fact checking sources as well as a large variation of fake news articles which have been sourced from less trustworthy sites. The dataset will be obtained from Kaggle [44] which contains truthful articles from Reuters [45] which state that the terms of use for their articles cannot be commercial use but they allow the use of articles for personal and academic reasons as long as the user does not attempt to alter the articles or make them available on any third-party websites. Terms of use can be found at Reuters terms of use webpage [46] and from figures 1 and 2 from the appendix of this paper. This dataset contains a total of '21417' news articles that are real-news and have been verified to be truthful as well as a total of '23481' fake news articles which have been sourced from unreliable sources that were flagged as untrustworthy by PolitiFact. PolitiFact is a fact-checking organization that evaluates the accuracy of statements made by politicians, public figures, and organizations, primarily in the U.S. PolitiFact uses a Truth-O-Meter rating system, ranging from "True" to "Pants on Fire," to assess statements based on evidence and factual accuracy and it has become a widely respected resource for verifying the truthfulness of claims in politics and public news.

The dataset is labelled as either 'fake', 'real', 'opinion' or 'partially true' and the data is balanced to ensure representation across different content types, sources, and topics such as politics, middle east news and more general news topics. By having a varied dataset containing a large number of articles, the experiment can aim to establish a solid foundation for evaluating the effectiveness of the blockchain-based fake news detection solution in a real-world context. The aim of the final product is to have at least 500 labelled pieces of media, consisting of different types of news such as articles, on the network which will then be used for analysis. This number will be significant enough to provide a large network of nodes that can be used to provide good training and test results. Also, this amount will give us a large number of trustworthy records and a good sample of fake articles to add to the blockchain.

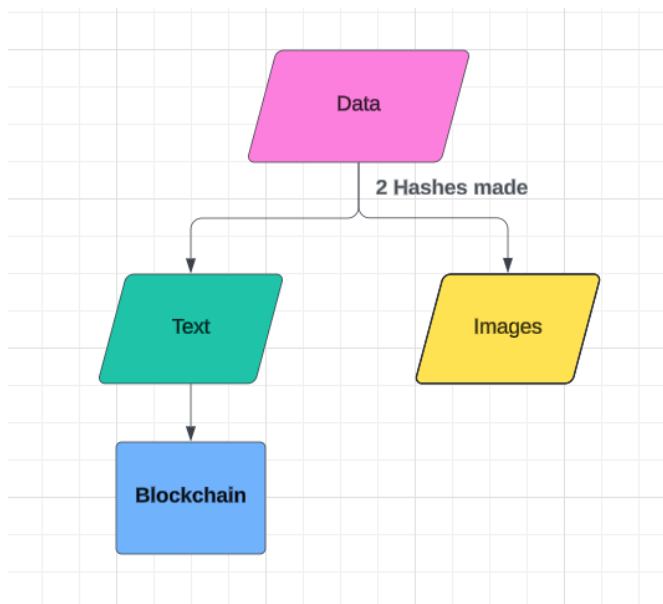


Figure 6: Breakdown of data into text and image hashes

Data in the diagram above depicts the news articles that will be used in the study, this includes real and fake news articles. Running the hashing algorithm using the data will provide two different hashes, one for the images included in the news article and the other will consist of the hash value for the text contained within the article. The hash value for the text data for all the articles will be used and stored onto the blockchain and the hash output for images will not be included and will be discarded.

3.3.3 Sampling method

A random sampling method is used to select news articles from each of the four distinct labels to ensure the dataset is representative of both types of news. From each label, we will also use random sampling to pick out news from each subcategory such as 'political' and 'world' news and the number of articles used that belong to each subcategory will be representative of the percentage quantity within the dataset. For example, in the fake news dataset, the number of 'left news' is half of the total quantity of 'general' news which means that we will use twice the amount of 'general' news in the training dataset.

3.3.4 Variables

There will be variables in the experiment that must be kept constant throughout the experiment to make sure that any changes in the dependant variables, are as a result of a change in the independent variables. Control variables will not be measured or changed as

doing so may affect the outcome of the experiment. The time period in which articles will be selected will be within the past 5 years of publication. This ensures that the age of the articles can be similar which reduces the occurrence of temporal biases which may affect the results. The format of the articles will be kept consistent within test groups as only text-based articles will be used for both groups. The sources of news will be consistent in terms of the credibility of where the articles are sourced from. We are making the assumption that all the labelled true news is coming from websites which are creditable and have been verified by external sources and tools and any labelled fake news are being sourced from websites which do not have any credibility and have been proven to spread misinformation in the past. This means that we are also assuming that real news cannot be published by fake news websites and vice versa.

The factor that will be changing in this study will be the label of the news article that is being fed into the system. The system should provide different output depending on whether the input is verified and true news, an opinion, partially true or misinformation. To measure the accuracy of the system. This variable will be changed by using a combination of articles of all labels.

The dependant variable is the outcome that is measured when different labelled news articles are inputted into the system. We will be measuring the classification that the regression algorithm predicts which allows for the calculation of the false positive rate, false negative rate, true positive rate and the true negative rate. Using these measurements, we will be then able to calculate the proportion of fake news that have been correctly identified.

3.3.5 Research stages

- Data Collection and Labelling: Collect real and fake news articles and label them before file storage.
- Hashing: Generate perceptual hashes from the pre-processed text of news content.
- Model Training: Train the multinomial regression model using the dataset of perceptual hashes and their labels.
- Model Evaluation: Test the model using unseen articles to measure its performance and then record the output.

3.4 Methodology Design

The methodology design is structured around four key components: data preprocessing, perceptual hash generation, blockchain implementation regression model classification. Blockchain is used for storing perceptual hashes to ensure data integrity and enable decentralised validation. The system's effectiveness is assessed through various performance metrics using a labelled dataset of real and fake news.

In this chapter, we will discuss the following steps in further detail:

1. Collection and definition of data: Collection of articles and the definition of the data structure.
2. Perceptual hash generation: Text preprocessing (removal of stop words, stemming, tokenization). Each pre-processed news article is transformed into a perceptual hash that represents the content of the news record using an algorithm.
3. Blockchain implementation: The generated perceptual hashes are stored on a private blockchain to ensure that data is tamper-proof and secure.
4. Regression model classification: Using the Hamming distance between a user-inputted article and the stored perceptual hashes, a regression model is trained and used to classify the article depending on similarities between hash values and use the label with the highest similarity.

3.4.1 Collection and definition of data

As mentioned earlier in the article, real news articles were collected from a trusted source [45] and have been verified by fact checking sites whereas fake articles are collected from less trustworthy sites that have been issued a low reliability rating from fact checking sites.

In the following table, the main data structures used in the system are listed, including those involved in perceptual hashing, blockchain storage, and model training.

Table 1: Data structure table

Data structure	Description	Type	Fields/Attributes
NewsArticle	Consists of the fields retrieved from the csv. file	class	Headline, content, date, subject, label

PreprocessedText	Result of the processing of text	List[str]	Tokens which represent a processed word from the text
PerceptualHash	A value that represents the perceptual hash of the content	str	A hexadecimal string
Hash	A value that represents the hash of the block itself	str	A hexadecimal string
Blockchain	A list of blocks representing the blockchain.	List[block]	Index, hash, previoushash, label and timestamp
Block	Represents a single block in the blockchain,	class	Index, perceptualhash, previoushash, timestamp, label, title, subject, date, text
HammingDistance	The number of differing bits between two perceptual hashes.	int	Integer
Dataset	A list of tuples representing the dataset for training and testing the regression model.	List[tuple]	Perceptualhash, label
TrainingData	A list of tuples, where each tuple consists of the Hamming distance and the actual label used to train the logistic regression model.	list[tuple]	Hammingdistance, label
TestingData	Training data used to evaluate the accuracy of the system	list[tuple]	Hammingdistance, label

The following 2 figures show what an example of both real and fake news article records look like. These records are stored in the same .csv file called 'test.csv':

title	text	subject	date
As U.S. budget fight looms, Republicans flip their fiscal script	WASHINGTON (Reuters) - The head of a conservative R politicsNews		December 31, 2017
U.S. military to accept transgender recruits on Monday: Pentagon	WASHINGTON (Reuters) - Transgender people will be a politicsNews		December 29, 2017
Senior U.S. Republican senator: 'Let Mr. Mueller do his job'	WASHINGTON (Reuters) - The special counsel investig politicsNews		December 31, 2017
FBI Russia probe helped by Australian diplomat tip-off: NYT	WASHINGTON (Reuters) - Trump campaign adviser Ge politicsNews		December 30, 2017
Trump wants Postal Service to charge 'much more' for Amazon shipments	SEATTLE/WASHINGTON (Reuters) - President Donald Tr politicsNews		December 29, 2017
White House, Congress prepare for talks on spending, immigration	WEST PALM BEACH, Fla./WASHINGTON (Reuters) - The ' politicsNews		December 29, 2017

Figure 7: Example records of verified real news

3.4.2 Perceptual hash generation

A perceptual hash for text data is a simplified representation that captures the essence of the content, rather than focusing on the exact sequence of characters [47]. Unlike SHA-256

title	text	subject	date
Donald Trump Sends Out Embarrassing New Year’s Eve Message; This is	Donald Trump just couldn t wish all Americans a Happy	News	December 31, 2017
Drunk Bragging Trump Staffer Started Russian Collusion Investigation	House Intelligence Committee Chairman Devin Nunes is g	News	December 31, 2017
Sheriff David Clarke Becomes An Internet Joke For Threatening To Poke People	On Friday, it was revealed that former Milwaukee Sheriff D	News	December 30, 2017
Trump Is So Obsessed He Even Has Obama’s Name Coded Into His Website	On Christmas day, Donald Trump announced that he wou	News	December 29, 2017
Pope Francis Just Called Out Donald Trump During His Christmas Speech	Pope Francis used his annual Christmas Day message to	News	December 25, 2017
Racist Alabama Cops Brutalize Black Boy While He Is In Handcuffs (GRAPHIC	The number of cases of cops brutalizing and killing peopl	News	December 25, 2017

Figure 8: Example records of verified fake news

hashing algorithms which produce a different output with slight modifications, perceptual hashing provides similar outputs when the input is slightly altered which allows them to detect similar pieces of data [48].

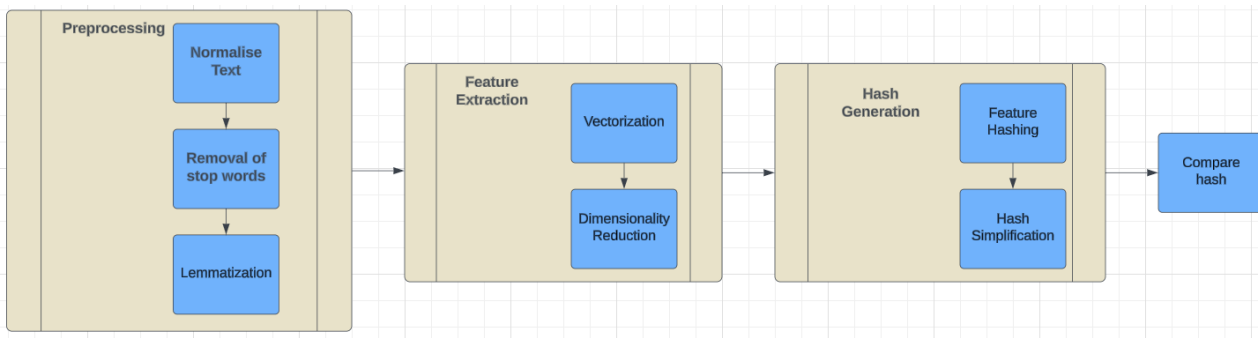


Figure 9: Perceptual hash generation

The figure above depicts the process that will take the text data from an inputted article and generate a perceptual hash value which can then be compared with the perceptual hash value of news already on the blockchain to detect any similarities between any two articles.

Using the text from article that needs a perceptual hash value to be generated, the text is normalised which means that all text is converted to lowercase, punctuation is removed, and words are converted to common variations such as “he’s” to “he is” so that the hash is not sensitive to small differences. Stop words are removed which consist of uninformative words such as “and” and “is” which don’t contribute much to the overall meaning of the text. During Lemmatization, words are reduced to their base form so words such as “making” gets reduced down to “make”.

After the text is pre-processed, the content is converted to numeral representation using TF-IDF (Term Frequency-Inverse Document Frequency), which is a measure of the importance of the word in relation to the whole document, which extracts the essential features of the text through the process of vectorization. Principal Component Analysis (PCA) or truncated SVD will be used to reduce the dimensionality of the vectorised text.

With the result of the feature extraction, a hash can be generated based on the features of the text. The sign of each feature value is taken which results in a binary hash meaning that a positive value gets converted to the binary value '1' and a negative value is converted to the binary value of '0'. This produces a compact binary string which represents the overall meaning of the text, and this string can then be simplified by applying bitwise operations or combining feature values to make a smaller hash. The final result can then be used for hash comparisons with the hashes on the blockchain.

When running a quick SHA-256 program from Xorbin [49] on an example article, the output of the program will consist of a 256 bit almost unique hash value which can be used to check the integrity of the article to ensure that no changes have been made to the article itself. For example, using the text from an article published by the BBC and written by Kotecha [50] as an input to the SHA program, it will output the string:

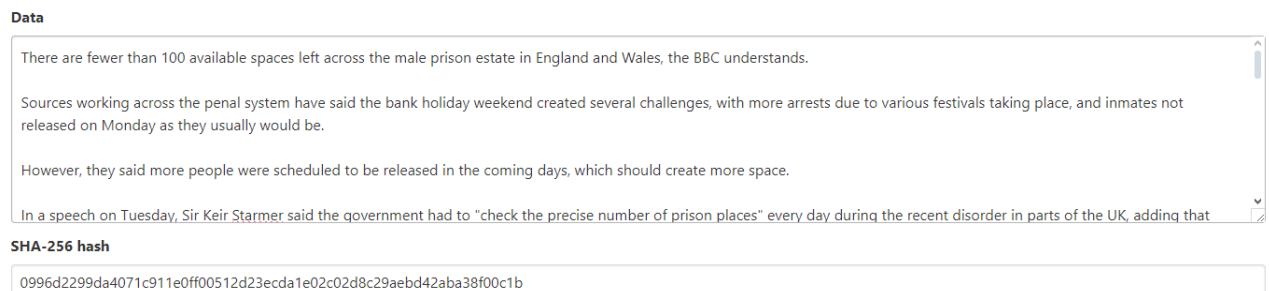


Figure 10: Example hash output of text article

The same hash value will always be calculated given that none of the content of the news article has been changed. If we alter the input of the program by adding a single character to the text article, the hash that will be calculated will look vastly different.

So, in the next given example, the number of available spaces stated has been changed from '100' to '1000' by adding an extra character of zero.

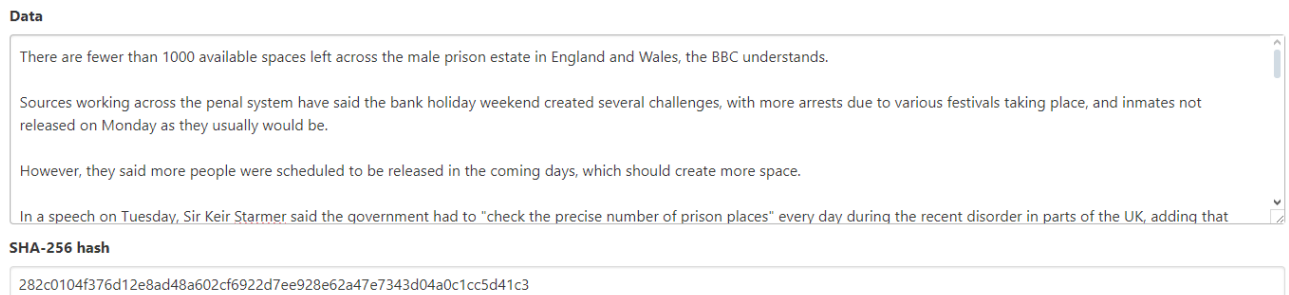


Figure 11: Example of hash output after alternation

The alteration of one character alone had a massive impact of the resulting hash value that was computed from the SHA program.

Sometimes an author may make alterations to a news article by correcting punctuation or fixing spelling errors which would still make the news article truthful but when running it through the hashing algorithm again, an entirely different output would be provided which suggests a need for a solution for this issue. Perceptual hashing addresses this by generating similar hashes for inputs which are similar to each other allowing articles that have been slightly altered to be traced back to its origins.

3.4.3 Blockchain implementation

Data stored in the blockchain system will include a dataset which contains a large variety of real news that has been verified to be true as well as a large variation of fake news articles which have been sourced from less trustworthy sites. Blocks added onto the chain will be timestamped with their appropriate information such as creation date and any relating nodes already existing on the chain. This will allow users to view whether a piece of text has been altered or deleted in any way and may suggest misinformation if an article has been changed. Any new data that is added to the system will have to previously be verified by external

methods before being added to the chain. Nodes that are verified will include a unique hash.

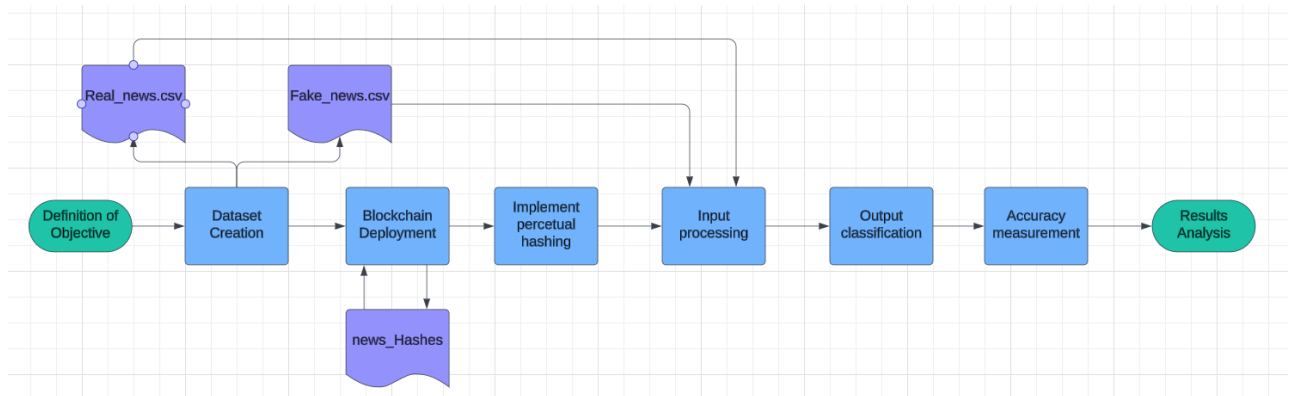


Figure 12: Workflow of proposed system

To create the blockchain system, a private blockchain will be implemented using Python-based libraries. Each block will contain the perceptual hash of the article, a timestamp, the label of the news, and relevant metadata such as date or subject of the news article.

Python is particularly suitable for implementing a blockchain system for detecting fake news using perceptual hashing for several reasons. Python's syntax is clean and easy to understand, which makes it an ideal choice for developing blockchain applications [58]. Blockchain concepts such as hashing, linking blocks, and maintaining immutability can be complex, but Python's simplicity and ease of usage reduces the occurrence of errors and makes the code easier to maintain. Python also offers a wide range of libraries and frameworks that make blockchain development and perceptual hashing easier such as scikit-learn which is a library used for logistic regression and is necessary for the classification of the news articles. Python has inbuilt JSON support which means that storing the blockchain data in a blockchain.json file for example, very convenient.

Labelled articles from each dataset will be inputted into the system one article at a time which allows the system to generate a hash value for each article to be outputted to the user. The hash can then be used to compare it against existing hashes on the blockchain and if there is a match to the value of the hash, it may indicate that the article falls in the same classification as the article which the hash belongs to. Any slight changes to an article should generate the same perceptual hash which will allow the user to trace the article back to its origins if stored on the blockchain system.

The first block of a blockchain is always the genesis block and the block will look like this on the system:

```
{
  "index": 0,
  "previous_hash": "0",
  "timestamp": 1727293704.0775204,
  "title": "Genesis Block",
  "text": "This is the genesis block",
  "subject": "N/A",
  "date": "N/A",
  "label": "N/A",
  "nonce": 0,
  "hash": "N/A",
  "titlehash": "N/A",
  "perceptual_hash": "N/A"
},
```

Figure 13: Genesis block data

Along the blockchain there will be a series of fake and real labelled news with its own unique hash as well as the hash of the previous block.

```
{
  "index": 2161,
  "previous_hash": "38a8e42a1f7041f441ecf3190ef44866a69a2585c7968933987ef49a99da8fc1",
  "timestamp": 1727293860.110792,
  "title": "Trump-backed candidate for Senate heads to Alabama run-off",
  "text": "(Reuters) - A Republican candidate backed by President Donald Trump for a U.S. Senate seat from Alabama",
  "subject": "politicsNews",
  "date": "August 15, 2017 ",
  "label": "Real",
  "nonce": 0,
  "hash": "f4acf0bba8c128505d0f8a9fb2e378e57e90b6f7a7933aa65acace1c906a9ef7",
  "titlehash": "f343fe123d3d51c51db2c7354ea0cf70",
  "perceptual_hash": "4b051e4d43bc8f6e7f08fa617549c6ad"
},
```

Figure 14: Block data of Real news

```
{
  "index": 6696,
  "previous_hash": "dfcafca5d367e569966cec59691348fe90781832b8a7d1db7f04cb43dca7c35d",
  "timestamp": 1727295003.8178093,
  "title": "DISGUSTING GOP Plan Would Make People Prove They Are Worthy Of Healthcare",
  "text": "Poverty is a moral failure, don t you know? Being born rich is the only thing that washes away all sins",
  "subject": "News",
  "date": "April 24, 2017",
  "label": "Fake",
  "nonce": 0,
  "hash": "9b8267b7299d5dac125da90d21bbb6635ff75f7747653d2729b23eb2b83a7bb1",
  "titlehash": "fd6a11bd4a3159937b717962eec2f361",
  "perceptual_hash": "c68be059a50961a15e768cbf42062d1e"
},
```

Figure 15: Block data of Fake news

```

"index": 484,
"previous_hash": "9fa6011fd7a028ed2b6045ec6ecb194e5e6e25749e78146721d90a6b54ff2448",
"timestamp": 1751864810.8131258,
"title": "MEDIA OBSESSES OVER Ted Cruz\u00e2\u20ac\u2122s Alleged Infidelities, While I",
"text": "Why is there so much secrecy when it comes to the mainstream media and the all",
"subject": "politics",
"date": "Mar 31, 2016",
"label": "Opinion",
"nonce": 0,
"hash": "65134a74e6269b1659185a1d2aae336378d2b5830a82e4c6373d93f11b8ed483",
"titlehash": "0f950f603295e9f45628faf7b0efd1d3",
"perceptual_hash": "6089f67f9cb421bf2eb3a2788537992d"

```

Figure 16: Block data of Opinionated news

```

"index": 681,
"previous_hash": "4defea0e7b073f248da9be969b5272865ad545a19c2fa79391534fb4ae6cdaca",
"timestamp": 1751864819.1702766,
"title": "Trump\u00e2\u20ac\u2122s \u00e2\u20ac\u2122 Storms Unarmed Legal Reside",
"text": "A Chicago man is in critical condition after the mother of all epic f*ckups by Trump s Deportation Force.",
"subject": "News",
"date": "March 27, 2017",
"label": "Partially True",
"nonce": 0,
"hash": "debca05119b3ac7b1dc6883aeb6ef665df82755429fd8d58c15d7aebdcd67dc1",
"titlehash": "d6e09529b84579d330928a9262120eff",
"perceptual_hash": "d242749f839299c243d3baac297992eb"

```

Figure 17: Block data of Partially True news

3.4.4 Regression model classification

Multinomial Logistic regression is a supervised machine learning algorithm used to predict non-binary outcomes based on input features. In this case, the system uses logistic regression to predict whether a given news article is real, opinionated, partially true or fake based on the Hamming distance between the perceptual hash of the user's input article and those stored in the blockchain. In our model, the logistic regression model outputs a probability value, which is then converted into one of four classes. The value of zero for a particular class will indicate that the news shared similar features and likely fits into the category whereas the value of one will suggest that the news is doesn't share much similarity with news from that category [51].

Logistic regression is suitable for this approach because it is simple, efficient, and interpretable. It calculates a combination of input features (Hamming distance and article labels) and uses a Softmax function to model the probability that the article belongs to one of our distinct classes. It is also used because logistic regression gives a probability output for each class, which is useful when dealing with uncertainty, such as determining whether an article fits into one category or another.

The primary feature used in the logistic regression model is the Hamming distance between the perceptual hash of the user-inputted news article and the hashes stored on the blockchain. A low value for the hamming distance suggests that two articles are similar in content whereas a higher value means that there is a greater distance between the values, and this establishes dissimilarity.

Model Training Process:

Input Feature: The Hamming distance between the perceptual hash of the input article and the stored hashes for each defined class are used as features. These are then combined with other features such as TF-IDF or metadata to form the input vector for the model

Training: The model is trained on the pre-labelled dataset where each article is assigned one of several categories. The MLR model learns the statistical relationship between the input features and each class label by optimising the Softmax-based cross entropy loss. It then outputs the probability distribution across all possible classes

The model computes the probability of each class using the following Softmax function:

$$P(y = k | x) = \frac{e^{w_k^T x}}{\sum_{j=1}^K e^{w_j^T x}}$$

Where $P(y=k|x)$ is the predicted probability that input x belongs to class k

x is the input feature vector representing a news article

w_k is the weight vector for class k , learned during training

$w_k^T x$ is the dot product between weights and features. It provides a raw score for class k

e is Euler's number for exponential function

K is the number of classes in total (4 in our study)

j is the index used in the denominator to sum over all classes

This result is that all class probabilities are non-negative and all sum to one

The model is training by minimising the cross-entropy loss:

$$L = - \sum_{i=1}^n \sum_{k=1}^K 1(y_i = k) \cdot \log P(y_i = k | x_i)$$

Output: After training, the model will produce a probability vector (P) for any new input article.

$$P = [P_{Real}, P_{Fake}, P_{Opinion}, P_{Partially True}]$$

These values represent the likelihood that the article belongs to each class and the total of these values will sum up to 1.

The article inputted will be assigned to the class with the highest probability:

$$\text{Predicted class} = \arg \max_k P_k$$

Where P_k is the predicted probability for class k , output by the regression model and $\arg \max_k$ is a mathematical operator that returns the value of k that maximises P_k .

An optional probability threshold can be used to flag low confidence predictions:

If $(P) \geq 0.5$, accept the prediction

If $(P) < 0.5$, don't accept the prediction unless its equal to the value of $\text{Max}(P)$

Example:

Table 2: Probability spread example

CLASS	PROBABILITY
REAL	0.20
FAKE	0.60
OPINION	0.15
PARTIALLY TRUE	0.05

Since $\text{max}(P)$ is 0.6 and the highest value corresponds to the fake category, the system will assign the appropriate to the article.

Python files for the blockchain solution can be found at:

<https://github.com/RickshawDriver/PerceptualHashBlockchain>

3.5 Evaluation metrics

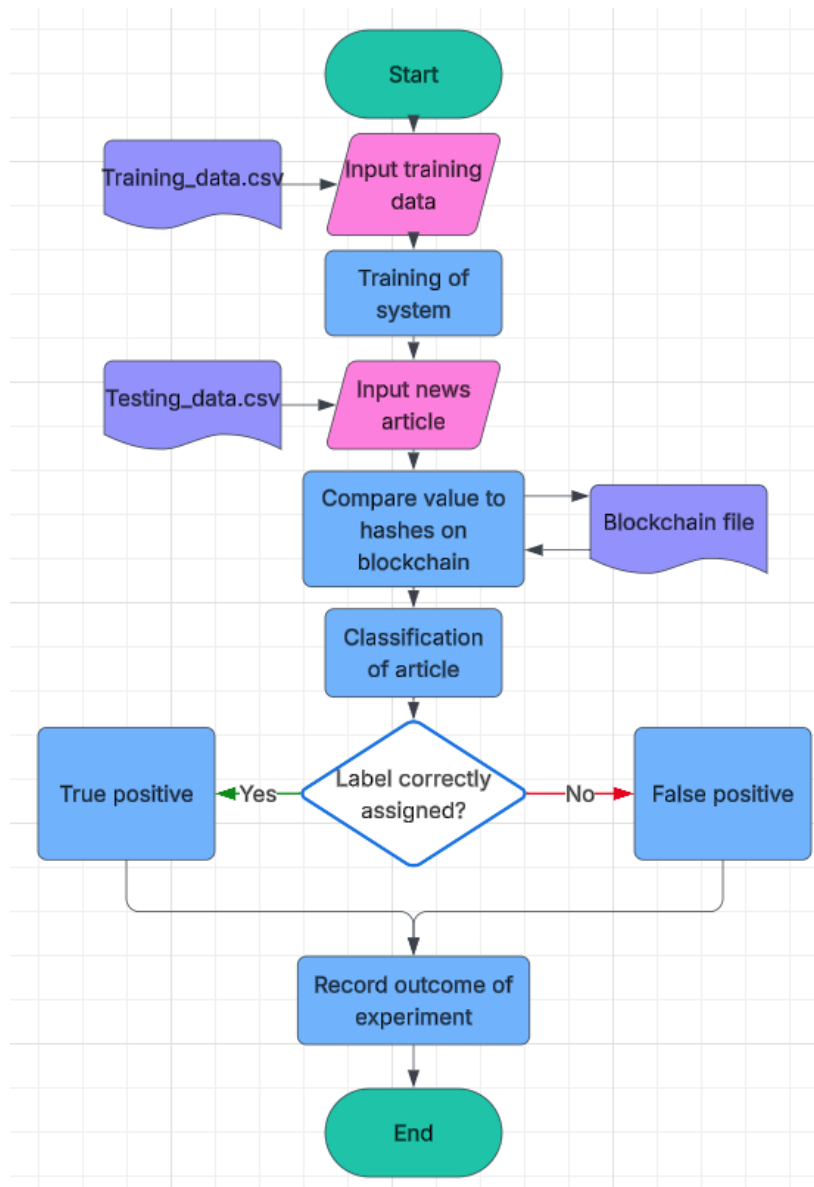


Figure 18:Flowchart of news classification

The study will evaluate the accuracy of the blockchain-based fake news detection system. Accuracy is the measurement of the proportion of news articles that are correctly classified by the solution, based on the comparison of perceptual hashes and blockchain verification. It will also measure the precision (correct identification of fake news), recall (minimising missed detections) and the f1-score (harmonic mean of precision and recall) to determine the success of the final system.

Accuracy of the proposed system can be calculated using the formula:

$$Accuracy = \frac{TP}{TP + FP + FN} * 100$$

Where TP (True Positives) represents the number of news that have been correctly identified by the perceptual hash and by using blockchain verification. FP (False Positives) are articles that have been predicted to class k but doesn't belong there and FN (False Negatives) are the number of articles that have belong to class k but has been predicted as another class by using the hashing algorithm and authentication from the blockchain.

The formula can be simplified to:

Equation 1: Calculation of Accuracy

$$Accuracy = \frac{\text{Number of Correct predictions}}{\text{Total number of predictions}} * 100$$

Precision measures the proportion of correctly predicted fake news out of all news that the system identified as fake. The calculation for this is as follows:

Equation 2: Calculation of Precision

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

The recall score measures the proportion of actual fake news that the system correctly identified. The calculation for this is as follows:

Equation 3: Calculation of Recall

$$Recall = \frac{TP}{TP + FN}$$

The F1 score is the harmonic mean of precision and recall. The calculation for this is as follows:

Equation 4: Calculation of F1 Score

$$F1 = \frac{Precision \times Recall}{Precision + Recall} * 2$$

If there was a total of 1000 news articles consisting of real and fake news in the experiment which concluded 929 true positives, 30 false positives and 41 false negatives, the calculation to find the metrics of the system would look like:

$$Accuracy = \frac{929}{478 + 451 + 30 + 41} * 100 = 92.9$$

The final accuracy of the solution would be calculated to be 92.9%

$$Precision = \frac{929}{929 + 30} = 0.969$$

The precision score of the system would be 96.9%

$$Recall = \frac{929}{929 + 41} = 0.958$$

The recall score of the system would be 95.8%

$$F1 = \frac{0.969 \times 0.958}{0.969 + 0.958} * 2 = 0.963$$

The F1 score of the system would be 96.3%

3.6 Ethical considerations

Ethical considerations will encompass privacy protection, data security, and adherence to copyright regulations regarding the use of news articles and social media content. Blockchain's transparency and immutability raises privacy concerns, particularly when dealing with sensitive information or personal data that can identify or be linked to an individual. Protecting user privacy while maintaining transparency in news verification poses a challenge and will require careful design considerations.

Owners of text may oppose of having their work being added to the blockchain and being verified by users of the network. Some pieces of text such as social media posts may be private and only available for certain viewers but anyone who has access will be able to take these private posts and add them onto the network for anybody with access to the blockchain to view the text without direct authorisation from the original owner, breaching their private data.

News companies may oppose the system because if text is incorrectly verified and a fake positive/negative is created, it may harm the reader into believing misinformation or it may damage the reputation of a legit news company if the article verified is incorrectly believed to be fake. Furthermore, news sites that are misleading but for the sole purpose of entertainment such as the Onion, will have their articles labelled as fake. While the articles of this site will be correctly labelled, readers will still want to read the text because of the

humorous aspect of these news articles. This problem can be addressed by adding a new label to the system which would label these types of news articles as 'satire' which would inform the user that the news is fake but with the intentions of entertaining the reader.

Once articles are classified, by assigning one of our classification labels (e.g. satire, opinion, fake, true or partially true), they are added onto the blockchain as a permanent record. The immutability of the records within the blockchain means that records cannot be changed or altered which raises concerns if the initial classification of an article is incorrect or even thought to be correct at the time but is then proven to be false and misleading [68]. For instance, an article that contains satire content may be incorrectly labelled as misinformation which may lead to reputational damage to a third party or organisation or perhaps a paper which focuses on the topic of medicine or healthcare may have information which gets proven to be untrue or not entirely factual with the introduction of further research after publication. Due to the decentralised nature of blockchain systems, it can be difficult to assign blame or have trust in the system if there is no central authority to oversee the classification process [69]. Incorrect labelling of news or malicious submissions to the blockchain can have serious consequences such as defamation of political figures and companies which can interfere with elections and campaigns. To resolve these issues, it would need to be clear to the end users how fairness is ensured during classification, if there are any protocols in place in the case of disputes/appeals, and the party which is responsible in ensuring that the system is consistently accurate with its labelling.

3.7 Discussions and Limitations

The solution struggles with determining the accuracy of the of the information itself within articles without external verification provided, if blockchain verification itself shows no signs of misinformation. This is due to the fact that blockchain itself as a technological tool is not equipped with the features to check the content of the articles itself as this requires human judgement or AI analysis because they can provide an understanding of context and interpretation [37].

Expanding the number of nodes would eventually slow the system down because as more blocks are added, the size of the blockchain grows exponentially and any new blocks must contain a set of transactions which results in an increased amount of data for the system to

traverse when searching and to process [39]. It will become a genuine limitation when applied to real world scenarios as hundreds of pieces of news would be written and needed to be added every day. Blockchain systems face inherent scalability limitations due to their decentralised nature and their reliance on consensus protocols. Using real world blockchain models of that utilise proof of work such as Bitcoin as an example, the typical number of transactions that can be done per second (TPS) averages at around 7TPS [72]. Georgiadis [71] calculated the exact upper bound for the maximal transaction throughput of the Bitcoin protocol and attained a TPS of 27 which is low compared to the TPS of PayPal of 100 and significantly inferior to the 1700 TPS that Visa performs at [70]. The low transaction throughputs of Blockchain technology hinders its widespread adoption when compared to existing solutions because in real-world situations where thousands of transactions are taking place at the same time, it can create a bottleneck in the system which results in delays and potential data loss.

The integration of the blockchain system into fake news detection will include considerable storage requirements and introduce computational complexity which must be considered before it's fully implemented to ensure system efficiency and viability. The reason for these intense requirements comes from blockchain's architecture as well as the additional data processing needed to generate a perceptual hash and classify each block before it is added to the blockchain [73]. Even after the processing is completed, all the blocks in the chain need to agree on the current state of the ledger through a consensus mechanism, which require the completion of resource-intensive operations such as complex cryptographic puzzles and smart contract validation [72]. These issues would affect the user experience of the solution because each time a news article is processed, the user would need to wait for perceptual hashing, AI classification and blockchain verification before receiving an output which confirms the news type of the text input. The proposed solution in this paper introduces a two-layer workload: The initial cost of resources of the use of AI in classification and hashing and the second layer that follows which consists of the computational demand of validating and adding new blocks to the blockchain.

Table 3: Comparison of TPS for existing systems

System	TPS	System Type	Scalability	Storage Requirements
Bitcoin (PoW)	7	Blockchain	Low	Low
Ethereum (PoW)	15	Blockchain	Low	High
Ethereum (PoS)	30	Blockchain	Moderate	High
PayPal	100	Centralised Financial	High	Low
Visa	1700	Centralised Financial	Very High	Very Low

The truth or accuracy of information can change over time as new facts emerge, or old data gets proven to be untrue with new research but may have been believed to be genuine at the time of publication. Blockchain is designed to be immutable, so once data is recorded, it cannot be changed. This is excellent for ensuring data integrity but raises an issue for content that might need to be updated or corrected in the future.

Introducing blockchain-based solutions requires widespread adoption from various users including companies, media organisations and regulatory bodies. It requires cooperation from a large number of news sources in order to be effective because the larger the blockchain database is, the more news articles that can be verified using our system because the existence of the original article is present on the solution.

Any discrepancy between the perceptual hash and the blockchain record could influence the accuracy. Thus, the system's success is tied to both components functioning correctly together.

3.8 Potential threats and attacks

The proposed solution in this paper may be vulnerable to various tactics that are designed to evade or manipulate the detection algorithm. One threat may arise when content is subtly altered – such as paraphrasing, rewording sentences or using synonyms – with the intent to trick the perceptual hashing algorithm because any small changes may significantly change the calculated hash and result in a lower similarity score if the system fails to detect the similarity between two articles [81]. This poses an issue because the system may incorrectly classify a reworded fake news article as new or original because the content hash would vastly different

due to minor changes, so the system would not flag the paraphrased article as being similar to the already labelled fake news article on the blockchain.

Beyond content manipulation, the infrastructure of the blockchain is also susceptible to attacks from malicious users, particularly on public networks. A major concern that needs to be considered is the possibility of a 51% attack where an attacker gains control of more than half of the system's computing power (hashrate). This allows the attacker to dominate the network's consensus algorithm and lets them manipulate the blockchain's ledger by changing the order of transactions, change parts of the blockchain or even prevent certain transactions from taking place [82]. This type of attack in our solution would allow an attacker to erase any records of disinformation, censor content from certain news sources, or ruin reputation of individuals/groups by mislabelling articles as being fact rather than opinion.

To mitigate these threats, we can enable the combination of multiple hashing methods such as perceptual, semantic [78] and deep hashing [77] which would improve its resistance to paraphrased or slightly modified content by allowing the system to understand the context of articles and it also permits the analysis of multiple dimensions of similarity. Instead of just relying on perceptual hashing, we can also use semantic embeddings from existing models like BERT to compare meanings of multiple texts. The cosine similarity of the embeddings can be used to detect similar or reworded articles which would relieve the issue of paraphrased content in the system [83]. Integrating multi-party consensus ensures that the validation of content is distributed between different trusted validators, making 51% attacks more difficult to execute because it increased the resource demand for an attacker to be able to successfully execute these attacks on the network [84]. To improve security even further, instead of a public blockchain, we can use a permissioned blockchain where only trusted validators, such as educational institutions, can write or validate data on the blockchain [79]. This does come with some drawbacks when applied to our system however because this would restrict the functions that would be available to the public and anyone who has not been granted permission by the blockchain. For example, a user would be able to check the label of an existing article by querying the blockchain, but they would not be able to submit any new articles onto the blockchain without having a trusted validator to review the submission beforehand which would require time and resources if many users are submitting articles to be added to the blockchain.

Chapter 4: Experimental and Validation Results

In this chapter, we discuss the experimental and validation results of the research methodology as proposed in chapter 3.

4.1 Experimental Setup

The experiment was conducted to evaluate the effectiveness of using perceptual hashing and a blockchain-based system to detect fake news and verify real news. The aim was to measure the system's accuracy in differentiating between real, fake, opinionated and partially true news articles, based on perceptual hash similarity. Additionally, the blockchain system was evaluated for its ability to store and verify news articles efficiently, ensuring tamper-proof data storage.

A total of 10,000 articles were stored on the blockchain and these articles were used to train our regression algorithm to provide more accurate predictions. A training dataset of 100 articles was first used to test the function of the system before being increased to 1000 when there was confidence that the system was able to process the data properly. Using larger quantities of training data such as 30,000 records resulted in slow processing times because the system had to calculate the perceptual hash of many records and therefore, it made the classification of a single user input too sluggish to be effective and would have an impact on the user experience. A final training dataset size of 1000 was finalised because it provided a sufficient number of records when training the system to provide a good accuracy that was above 70% using the formula mentioned in chapter 3. Increasing the number higher than 1000 in the training set caused exponentially long delays and wait times during training when a higher number of records were used. Throughout the testing process, the proportion of news within each distinct class in the dataset remained constant and as equal as possible to prevent bias when training the system because having an imbalance in the quantity of each news type can cause the classification algorithm to favour one label over the other.

Table 4: Distribution of news articles

Category	Number of articles
Fake news	250

Opinionated news	250
Partially true news	250
Real news	250

Using a testing sample of 400 news articles that were not in the training sample but from the same source file (100 labelled fake news, 100 labelled opinionated news, 100 labelled partially true news, and 100 labelled real news), a mixture of different types of text content was fed into the blockchain system to test the accuracy of the solution. A perceptual hash was calculated and compared with the stored perceptual hashes on the blockchain using regression algorithm and Hamming Distance to determine whether the user's headline or article content fit into a particular news class based on existing training data and learned features of news articles from each class. The generated perceptual hashes for both real and fake news articles were 64-character hexadecimal strings.

4.2 Experiment and Validation Results

Table 5: Results from news classification of Content

Category	Labelled as Fake on system	Labelled as Real on system	Labelled as Opinion on system	Labelled as Partially True on system
Fake news	78	12	7	3
Real news	4	82	6	8
Opinionated news	16	9	68	7
Partially True news	13	26	8	53

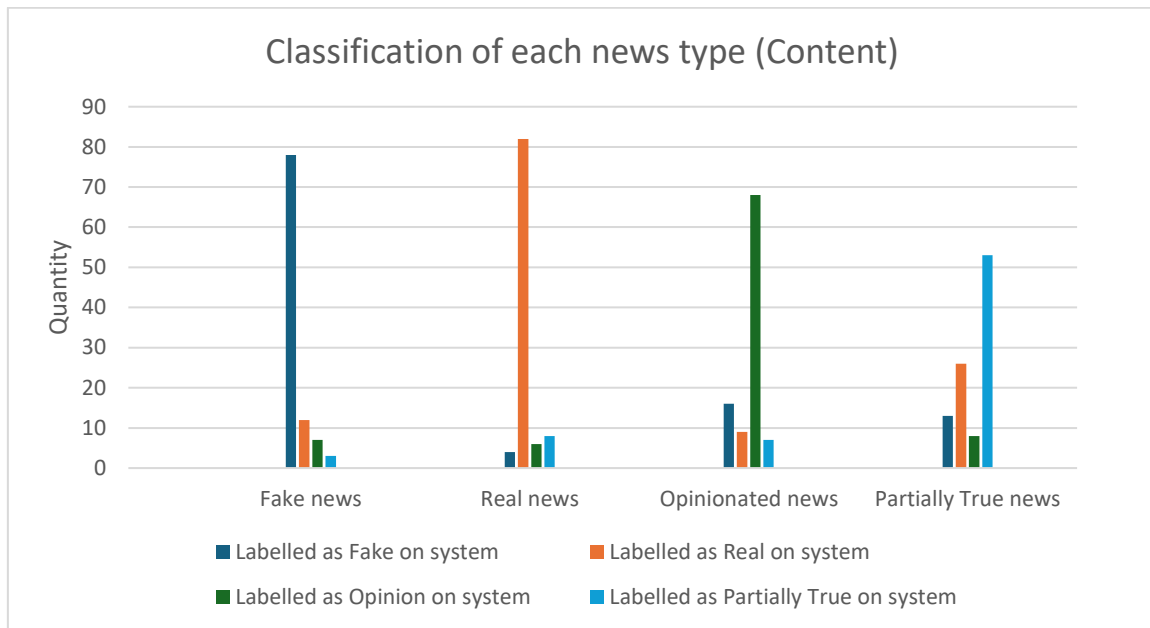


Figure 19: Classification of labelled news (Content)

Using the results gathered from the experiment, we can calculate an accuracy for the system which will help us determine whether the blockchain solution is successful or not in the detection of fake news. Using the following formula to calculate the percentage accuracy from the results that we recorded above.

$$Accuracy = \frac{TP}{TP + FP + FN} * 100$$

In the equation above, this is where:

TP (True positives): the number of news that have been correctly classified

FP (False positives): number of articles that have been incorrectly classified into the specified category

FN (False Negatives): the number of articles that belong in the specified category but have been incorrectly identified into another class

Category	Labelled as Fake on system	Labelled as Real on system	Labelled as Opinion on system	Labelled as Partially True on system
Fake news	TP(Fake)	FN->Real	FN->Opinion	FN->Partially True

Real news	FP->Fake	TP(Real)	FN->Opinion	FN->Partially True
Opinionated news	FP->Fake	FN->Real	TP(Opinion)	FN->Partially True
Partially True news	FP->Fake	FN->Real	FN->Opinion	TP(Partially True)

Figure 20: Confusion matrix of true/false values

Table 6: Total distribution of true/false values

Outcome	Total Count
True positives	281
False positives	33
False negatives	86

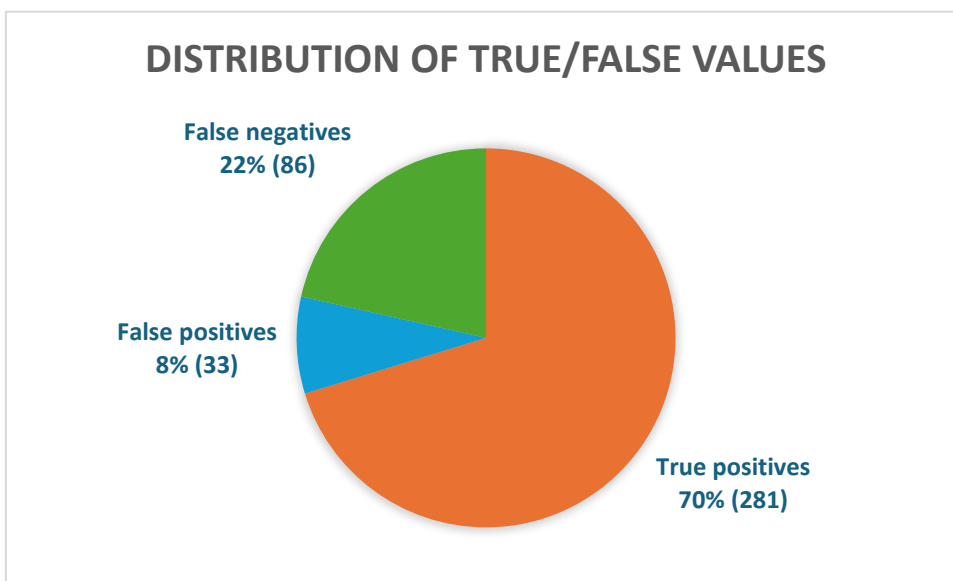


Figure 21: Graph of the Distribution of True/false values

$$Accuracy = \frac{281}{281 + 33 + 86} * 100 = 70.25\%$$

It has been concluded, using equation 1 from Chapter 3, that the final accuracy for the system is calculated to be 70.25% accurate when performing the experiment. Unfortunately, the

system has provided an accuracy below 90% which means that the solution is not consistent enough at predicting fake news to be able to show that the blockchain is effective.

Precision is the proportion of articles that the system as a certain class that was correct

$$Precision_{Fake} = \frac{78}{78 + 33} = 0.702$$

When the system predicts fake for an article, there is a 70.2% chance its correct

The recall is the proportion of articles from a true class that the system found. In the following calculation, we use equation 3 from Chapter 3:

$$Recall_{Fake} = \frac{78}{78 + 22} = 0.780$$

If there are 100 fake articles, the system will correctly identify 78 of them as fake

Finally, we calculate the F1 score, which is the harmonic mean of precision and recall, providing a balanced evaluation metric using equation 4.

$$F1 = \frac{0.702 * 0.78}{0.702 + 0.78} * 2 = 0.740$$

The result of this calculation resulted in an F1 score of 74.0%.

Now the same equations will be calculated for real, opinion and partially true classes:

Table 7: Performance metrics of all groups (Content)

Class	Precision	Recall	F1-Score
Fake	0.703	0.780	0.740
True	0.636	0.820	0.715
Opinion	0.764	0.680	0.719
Partially True	0.747	0.530	0.619

The reason that the model failed to achieve our target accuracy of 90% may have been as a result of multiple reasons:

- Our dataset which was used to train the regression model may have been too small which meant the model may not generalise as well, leading to poor performance and accuracy.
- Perceptual hashing is designed to catch small changes in content, but this can become a problem when it misidentifies minor edits, such as content that is reworded for clarity, as different content, affecting similarity detection.

- Fake news can often rephrase real news articles with little changes. The perceptual hash might generate similar hashes for news that are semantically the same, but one may be factually false, leading to false results.
- Using a basic regression model may not be advanced enough to capture the complex patterns in news data. More complicated models, such as neural networks or support vector machines, might perform better.
- The model might be overfitting, learning too many details from the training dataset, causing poor accuracy on the unseen data we used for testing. The model may also be underfitting, where the model is not capturing the patterns of the data, leading to poor predictions.
- When classifying opinionated or partially true news, there is a lot of overlapping language and structure which makes it difficult to separate uses hashes which is why the performance metrics for these 2 classes fell short compared to the 'fake' and 'real' tags.

Using a testing sample of a different set of 400 news articles (100 from each label), the headlines of the news articles were classified by the system instead of the main body of text. The blockchain also stores a perceptual hash of the title which is used to predict the headline's news category using the regression algorithm also.

Table 8: Results from news classification of Headlines

Category	Labelled as Fake on system	Labelled as Real on system	Labelled as Opinion on system	Labelled as Partially True on system
Fake news	71	16	9	4
Real news	16	74	8	2
Opinionated news	26	11	57	6
Partially True news	16	27	10	47

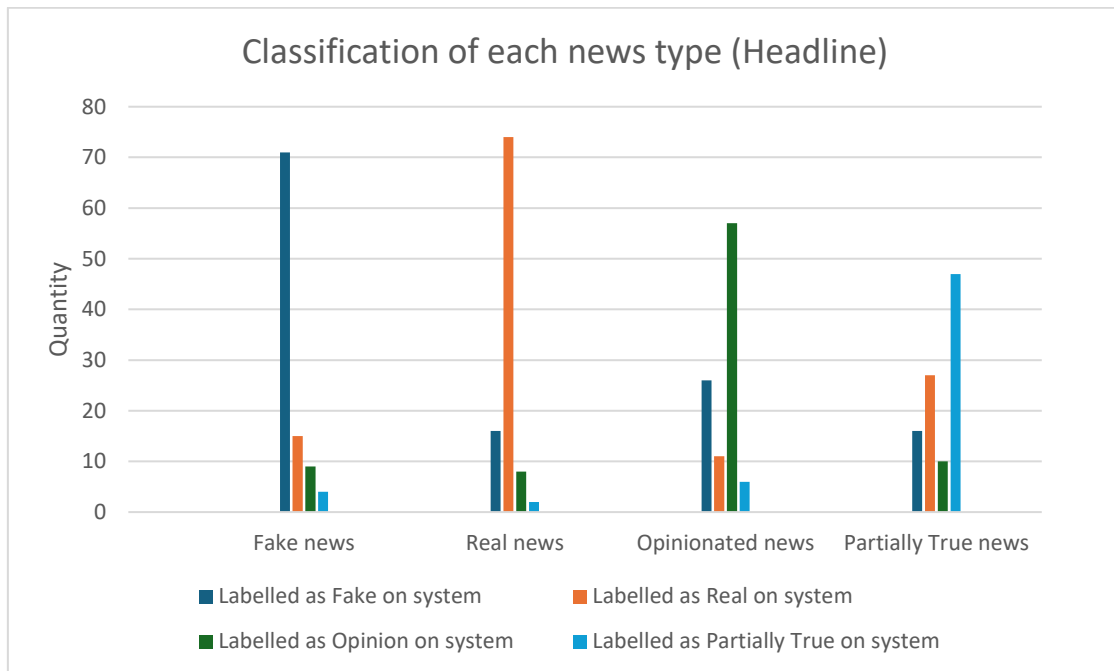


Figure 22: Graph of the Classification of labelled news (Headlines)

Table 9: Total distribution of True and False values

Outcome	Total Count
True positives	249
False positives	58
False negatives	93

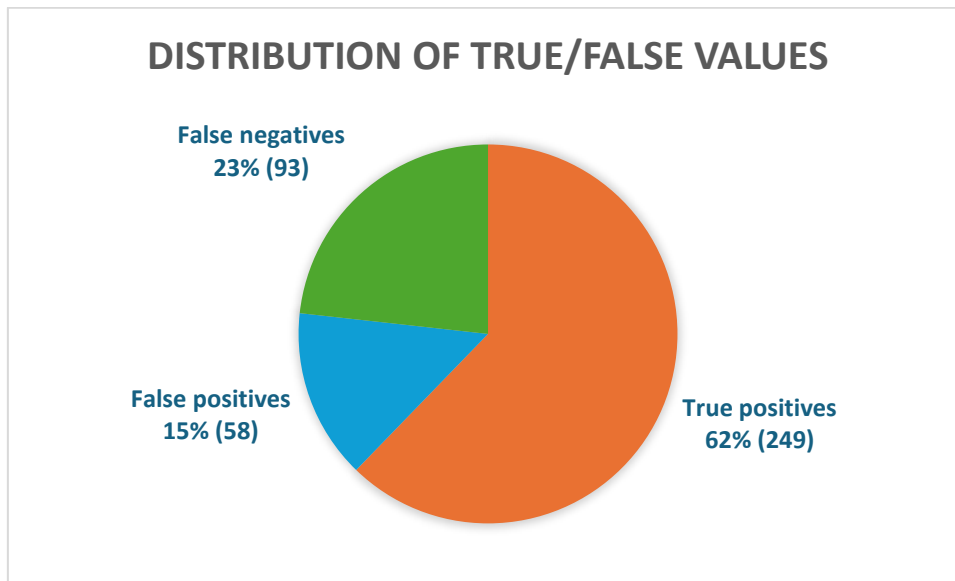


Figure 23: Graph of the Distribution of Fake/Real values

Using the equation defined previously, we can use these results to calculate the accuracy of the blockchain system when trying to detect fake news using the headline of the article

Using the equations defined in Chapter 3, we get the following results:

Class	Precision	Recall	F1-Score
Fake	0.550	0.710	0.621
Real	0.578	0.740	0.649
Opinion	0.679	0.570	0.618
Partially True	0.797	0.470	0.592

The final accuracy for the system when predicting the label of news articles when given its headline is 62.25% which also falls below the needed accuracy of 90% to be certain that the system is concise enough to be effective.

The system performs better when given the content of the news article and provides a correct classification majority of the time, but its accuracy is not high enough to say that the hypothesis can be proved to be true. One possible reason for the increased performance for content hashing is that there is more data for the model to train with. The content of the articles may contain paragraphs of sentences whereas within the headlines of the articles, they contain significantly less words. This means that for short headlines, a subtle change in one word may dramatically alter the hash value of the headline compared to changing one word of the article's content because the amount of change is proportionally larger in the headline scenario, even if it's just one word.

4.3 Adversarial scenarios

Multiple news articles were paraphrased and inputted into the system to see if the system would be able to detect the articles as being similar to existing records on the blockchain. A total of 5 examples of paraphrased news titles were used to test the perceptual hashing system under edge cases. The following table highlights the original title of each news article which was used in testing as well as the reworded counterpart.

Table 10: Comparison of original and paraphrased headlines

Input No.	Original News Headline	Paraphrased Headline
1	FATHER SPEAKS OUT After Son Was Attacked By Mob At Trump Rally	Father Breaks Silence After Son Assaulted by Crowd at Trump Event
2	BUSTED! OBAMA AND HILLARY LIED: Former Hostage Reveals Proof The \$400 Million Dollars Was Ransom Given To Iran	Ex-Hostage Claims Evidence Shows \$400 Million Was Ransom Paid to Iran, Accuses Obama and Clinton of Deception
3	The White House Has Gone Dark And CNN Is NOT Having It	CNN Outraged as White House Falls Silent on Key Issues
4	Trump Sends Out Email Begging Supporters To Send Him Money To Defend Against His Scandals	Trump Appeals to Backers via Email for Donations Amid Scandal Allegations
5	Trump to meet with long list of leaders in New York next week -White House	Trump plans to meet with big list of leaders in New York next week -White House

The system generated the hash of both the original text and the paraphrased versions and computed the hamming distance between them. A similarity of 0.8

Input no.	Hamming distance	Similarity (%)	Same class detection? (y/n)
1	73	42.9	n
2	65	49.2	n
3	73	42.9	n
4	54	57.8	n
5	55	57.0	n

The current similarity scores fall within the range of 40% to 60%, even when a few words are changed in the headline. This suggests that the current perceptual hashing method is moderately sensitive to structural changes in text. An average similarity score of approximately 50% indicates that half of the bits in the binary representation of the two hashes differ from each other which shows us that the paraphrased versions retain some core vocabulary and

elements as the original article but there is also enough variation to alter the hash significantly to the point where it bypasses our systems. To improve the results from the test, more context-aware techniques such as semantic embeddings could be used as the current solution lacks the semantic understanding needed to detect paraphrased content effectively and improve performance in adversarial scenarios.

4.4 Error Analysis and Observations

In order to understand the limitations and performance of the proposed system, we conducted a focused analysis of misclassifications – specifically, false positives false negatives. The aim of this analysis was to uncover any patterns and issues that need to be addressed when classifying complex news content using perceptual hashing and multinomial logistic regression.

A thorough review of the results revealed that ambiguity in news content was a recurring factor when it came to misclassifications. News articles written in a neutral or calm tone but covering disinformation often bypassed the classifier and was categorised with other real news, especially when the language used was similar to that of verified content. An example that came up in this study was “Christian cake baker Jack Phillips gained national infamy when he violated Colorado’s non-discrimination law by refusing to bake a cake for a gay couple’s wedding. He was subsequently sued, and now his case is headed to the Supreme Court.” which was classified as being true news due to the lack of impactful words or shocking punctuation. Additionally, articles that presented truthful information but quoted information or statements from articles which contained fake news often were misclassified as being fake themselves, depending on their phrasing. Also, articles which were created with the purpose of entertainment for readers and contained purposeful misinformation such as an article titled “STREET ARTIST Censored For Painting Of Hillary Clinton In Bikini” in the dataset, were also often marked as fake despite not having any malicious intent which highlights the difficulty in differentiating intent or context through hashing alone.

Some misclassifications were also present when dealing with opinionated articles or partially true content. These types of content normally contained factual information but presented in a way that creates a strong bias which makes the reader see a certain point of view and not the full perspective. In such cases where context is important, the perceptual hash algorithm

focused on word frequency patterns and failed to capture any deeper meaning behind the articles which lead to incorrect predictions.

To address these limitations, several improvements could be made:

- Features such as sentiment polarity, or named entity analysis can provide understandable context for the system
- Incorporating additional algorithms alongside perceptual hashing such as semantic embeddings could improve the classification accuracy
- Having human validators in the system for when the system has a low confidence score during classification could reduce the number of misclassifications
- Having a satire class for humorous articles or coding in hierarchical labels may help with distinguishing between exact content types

In summary, this error analysis confirmed that perceptual hashing offers a lightweight and efficient approach to tackling the issue of fake news but struggles when faced with the detection of intent or tone of an article. The implementation of additional methods and models alongside the system would significantly improve its accuracy and robustness.

4.5 Comparison of system with existing solutions

A comparative analysis was conducted against existing approaches to fake news detection on our completed system. The goal of this benchmarking was to determine how our solution performs in relation to more computationally intensive models, especially those which use deep learning and different architectures.

Two different solutions were used with the same dataset for consistency and fairness of the experiment to compare the accuracy that each system was able to achieve. The two systems used was a Support Vector Machine (SVM) [65] and a transfer-based model – BERT. BERT was fine tuned for our scenario using Hugging Face transformers library with a classification added for multiclass outputs as we need to classify the articles into one of four categories. The models were training using identical training data and the same exact classification labels but for BERT, tokenisation and embedding was handled using pre trained weights from Hugging Face and fine-tuned for five epochs [62]. We evaluated performance using the same four metrics: Accuracy, Precision, Recall and F1-Score.

Table 11: Comparison of performance metrics between solutions

Model	Accuracy	Precision	Recall	F1-Score
Perceptual Hashing/ Multinomial Logistic Regression	0.703	0.713	0.703	0.698
Support Vector Machine	0.741	0.722	0.718	0.720
BERT	0.887	0.875	0.862	0.868

All results were calculated as an average between all 4 classes.

As shown in the table above, BERT significantly outperformed both SVM and our perceptual hashing approach across all metrics and this is to be expected. As discussed previously in this paper, BERT is a transformer-based model and is pre-trained on a large dataset and is able to capture the semantic meaning of text rather than just analyse the structure or word count. Despite this, perceptual hashing does offer notable advantages when it comes to speed and resource efficiency, without sacrificing too much accuracy. It also has increased security because it uses hashed content for analysis and the figures from our calculations show that with further improvements, it shows potential in hopes of achieving a similar accuracy to BERT and other transformer systems.

4.6 Summary

The experimental results show that the blockchain-based perceptual hashing system does not consistently perform effectively in detecting fake news. Key findings include:

- The system has a total accuracy of 82% for content prediction and 77.6% for headline classification
- When the system has labelled an article as being fake, the system has correctly classified the article 79.62% of the time
- The system performed best when the Hamming distance was small, correctly classifying most fake and real news.
- As the size of the dataset increased from around 100 to 1000, the model's accuracy and generalization ability improved, highlighting the importance of a large and diverse training set in fake news detection systems.

In conclusion, the integration of perceptual hashing with a blockchain-backed system has the potential to provide a robust method for detecting fake news by comparing content similarity in a decentralised and transparent manner but due to limitations, the accuracy and precision have not been high enough to prove our hypothesis.

Chapter 5: Conclusion

In this chapter, the proposed work is concluded along while highlighting the achievements and the impact of this study. This chapter also highlights some potential improvements that can be undertaken for any future work that builds upon this study.

5.1 Conclusion

In summary of this thesis, we outlined the issue of fake news in Chapter 1 and the impact that it has when spread online. This chapter also presented the hypothesis of the study as well as the main aims of the experiment.

Chapter 2 consisted of an in-depth review of the literature which aimed to provide an overview of the current main methods of fake news detection which was machine learning and NLP. It also discussed blockchain as a potential solution to combat fake news because it allows the authentication of news sources. The review also covered multinomial logistic regression in cases where the outcome can be classified multiple categories. The literature highlighted some of the limitations with existing solutions for detecting fake news and research has suggested a potential for the combination of blockchain and perceptual hashing in one system to create a more robust answer to the issue of fake news.

Chapter 3 details the blockchain architecture and the methods that would be used to create the final system. It explains what kind of data will be stored in the blockchain and it also details how the data is acquired and prepared before being used in the experiment. It also explains the process of perceptual hashing on the system, which converts the content of the news articles that are stored on the system into hash values which are then stored securely on the blockchain. Multinomial logistic regression is also mentioned in Chapter 3 which explains how the model utilises the hamming distance between hashes to classify news articles and how the Softmax function is used to train the data. Towards the end of Chapter 3, it states the process of user input into the system, and we explained what metrics and formulas will be used to evaluate and analyse the success of the final solution.

Chapter 4 presented the final results of the experiments which were conducted to test the accuracy of the system. A confusion matrix is used to provide a visual aid for the reader to understand what is meant by true positives/negatives and false positives/negatives. A detailed analysis of the system is detailed and also backed up by quantitative results such as the final system accuracy of 70.25%. This chapter presented some reasoning and understanding as to why the final accuracy of the system is relatively high but not high enough to pass the 90% accuracy needed to prove the hypothesis. Chapter 4 also interpreted the conclusions in relation to the research objectives, and it reflects on the success of perceptual hashing when paired with blockchain technology for fake news detection. It also compares its results to existing solutions such as BERT when ran with the same dataset as a benchmark comparison. Paraphrased content was analysed and run through the system to see how the system would handle adversarial attempts to manipulate the system, and any wrong classifications are discussed through an error analysis of the false positives and false negatives.

Chapter 5 concludes the whole thesis and underlines the system's novel approach to combining perceptual hashing and blockchain for fake news detection. It summarises the findings and contributions of the research, but it also discusses any limitations and possible expansions that can be undertaken to improve the system if future work was to be conducted. Any possible threats to the solution are also mentioned as well as some possible solutions which would help prevent and mitigate these attacks.

5.2 Achievements and Impact

5.2.1 Achievements

After successfully calculating a final accuracy of the solution from the results of the experiment, this research has successfully developed a hybrid solution that combines the use of perceptual hashing to identify content similarity and blockchain, which ensures the security and immutability of stored data. This system also integrates logistic regression which uses the hamming distance between perceptual hashes to classify whether the article is more likely to belong to a particular class or not, depending on its similarity to existing articles which are stored on the blockchain. This is a novel approach in the context of fake news detection, where previous methods have primarily relied on text-based analysis and artificial intelligence without ensuring data integrity.

Although the system did not achieve the required accuracy of 90%, it did achieve a relatively high accuracy of 70.25% which demonstrates that the system is effective the majority of the time, but with some improvements and future work towards the consistency of predictions, reaching the accuracy threshold is very achievable.

The system has been designed to be scalable which means that additional news articles can be added to the blockchain system which would increase the accuracy of the model, assuming there is a large available quantity of computational resources and storage.

The experiment was conducted using real-world news articles with known labels sourced from existing news sites, highlighting the fact that the system has the potential to deliver real-life applications.

5.2.2 Impact of study

The integration of perceptual hashing and blockchain in one system creates a verification system for news content, helping to combat the growing problem of misinformation in media online. It will enable media platforms and sites to verify articles against existing records on the blockchain system and check its authenticity before releasing any information to the public, preventing the circulation of fake news. The use of this blockchain ensures that the system maintains a permanent, transparent, and unalterable record of all news articles and their classifications. Users can be confident that the records on the system are secure and have not been altered in any way.

This system acts as an educational tool for the public, raising awareness of the techniques that can be used to detect fake news. Showing how automated systems can detect fake or misleading content, encourages the development of media literacy and critical evaluation of news sources that are unverified.

5.3 Future scope

Despite the fact that the system was not able to achieve an accuracy that was sufficient enough to prove the hypothesis mentioned in Chapter 1, there are several avenues for future research that could improve the performance and scope of this system which would help increase the overall accuracy and precision.

5.3.1 Advanced perceptual hash techniques

To improve the scalability and robustness of the proposed fake news detection solution, several improvements can be considered and implemented. Firstly, the use of more advanced hashing methods, such as deep perceptual hashing [77] or semantic hashing [78], can be added to increase the system's ability to detect paraphrased or semantically similar articles. Implementation of these techniques allow the solution to generate feature-rich representations of content which results in a more robust similarity detection between articles when compared to simple surface-level hashing techniques. To implement this, we would need to replace our current hashing algorithm with a neural network-based model that generates similarity-preserving hash codes for text. This would be a system that allows the user to input an article which would firstly be tokenised and cleaned to remove common words that have little to no meaning before the system encodes the result into dense semantic embedding using tools such as BERT or Distil BERT which are pre-trained transformer models. The BERT embeddings would be passed through the hashing layer to calculate a binary hash code. The hash generation can follow the HashNet architecture presented the paper from [77], where deep neural networks use a training method called 'supervised continuation' to learn similarity-preserving hash codes. We also use supervised (pairwise or triplet training) or semi supervised methods (contrastive loss or Quantisation loss) to train the algorithm. The hash codes and metadata of articles inputted can then be added to the blockchain so that we can use hamming distance to compare articles are flag content which is similar to known fake news [80].

Table 12: Comparison of regular and deep hashing techniques

Capability	Regular Perceptual Hashing	Deep Perceptual Hashing
Duplicate content detection	Yes	Yes
Semantic understanding	No	Yes
Ability to deal with paraphrased content	No	Yes
Accuracy	Limited	High
Efficiency	Fast	Depends on the resources available

Future work could explore the utilisation of Multimodal Hashing to allow the combination of visual, textual, and audio hashing for articles that contain many forms of media [52]. This would result in a significant improvement to the system because fake news often includes

manipulated images and even video media. The system is limited to text-based data at the moment which restricts its versatility when detecting fake news in different media forms and therefore, would benefit from perceptual hashing algorithm improvements or even combined use of multiple algorithms to allow multimedia detection.

5.3.2 Larger and more diverse training datasets

The dataset that was stored on the blockchain and used to train and test the system was relatively small compared to the amount which would be representative of real-world scenarios. As mentioned in Chapter 3, the training dataset for the system consisted of 10,000 unique articles where within the dataset, there was an equal number of fake news and real news. However, with more news articles being written every day, the number of fake and legitimate news articles increases dramatically which suggests a need for a system which would be able to store a large amount of data. In a scenario where computational power was less of a limitation, having a larger and more diverse dataset to store and train the system with would improve the model's ability to generalise across different types of news content. It would also reduce overfitting to a specific category of news articles and enhance the robustness of the solution in real-world applications.

5.3.3 Hybrid model architecture

Newer iterations of our solution would benefit from separating the storage of data from the verification of data through the use of a hybrid model architecture. News articles would be stored off-chain using a decentralised storage method and only the essential metadata such as hashes of content and verification results are stored on the blockchain itself to be used to verify the integrity of the data. Each data object that is stored off-chain would have a content-addressable hash which links it to the on-chain counterpart [75]. This promising approach allows efficient and quick verification as well as increased scalability [76], without overloading the blockchain network. By reducing the on-chain data volume, we avoid encountering any bottlenecks and it allows us to achieve a higher TPS which is essential when classifying hundreds of articles per day. Having data stored off the chain means that it can independently be updated or corrected easily, and it allows the system to support ethical obligations under GDPR such as the right to data rectification.

5.3.4 Semantic analysis of text

Fake news often involves small linguistic manipulations, where the structure of an article might remain the same, but the meaning could be altered through changes in word choices, or tone. The integration of semantic analysis using natural language processing could also be added to capture deeper content differences that the hash algorithm implemented in the study may not have detected. By implementing one of many semantic analysis techniques available such as Latent Semantic Analysis (LSA) [53] or word embeddings, we can create a focus on the analysis of subtle manipulations and biases presented in the article. The result of this addition, alongside the perceptual hashing framework, would improve the quality of fake news detection of the blockchain and would represent a significant step forward in preventing the spread of fake news online.

The current solution currently offers traditional perceptual hashing methods which excels at identifying articles which are similar based on the structure/lexicity of the content; however, this technique is not as effective when analysing content which has been reworded differently, paraphrased or has been altered contextually. Semantic analysis mitigates this limitation by enabling the system to understand the intent of text as well as its overall meaning rather than just examining its structure and appearance. Techniques such as Named Entity Recognition (NER) can be used to provide to provide insight on the meaning of individual words and sentences within articles [74]. NER is a natural language processing technique that is often built using deep learning models or transformer based models to categorise what the potential meaning of each word may be in both forwards and backwards directions For example, NER would help the system distinguish whether the word 'Amazon' refers to the company or the location, depending on how the word is used within the sentence. It helps with clarifying the true context of any claims made by recognising who the core subject it and any action, which reduces ambiguity and helps prevent false positives where real news is identified as misleading due to vague statements [62].

5.3.5 Implementation on web browsers

Coding the system into a browser extension could allow end-users to verify the authenticity of news articles in real-time when they encounter a news article online as they surf through webpages. The web extension would calculate and check the perceptual hash of the content

on the user's current webpage with the information stored in the blockchain to verify the text's authenticity. The web extension would then be able to confirm the news article on the blockchain system and display a label on the webpage to inform the user about whether the information on the screen is fake or not after it has been verified by the system. This could help prevent the spread of fake news directly at the user level.

This research establishes that the addition of some future work and improvements, the combination of perceptual hashing and blockchain technology offers a viable solution for addressing the challenges of fake news detection while ensuring data integrity through blockchain's decentralised architecture. While there are limitations in terms of dataset size when training the model and classification complexity, this research may alter the way we detect misleading content online and will provide a strong foundation for further research that aims to help tackle the ongoing issue of fake news in the future.

5.3.6 Threat mitigation

As mentioned in chapter 3 in this study, while the system demonstrates the potential of combining perceptual hashing and machine learning for fake news classification and detection, it remains susceptible to several adversarial threats. These included deliberate paraphrasing of content to bypass perceptual hashing with the aim of flooding the blockchain with inaccurate data by manipulating the system into thinking that a slightly reworded fake article, is an entirely new article which may result in it being labelled differently. Addressing this issue could include integrating multiple hash methods at the same time such as perceptual and semantic hashing [78] to improve the system's ability to detect rephrased content. There is also the risk of a 51% attack on the system, especially in public blockchain architectures. Mitigation of this type of attack would include a permission-based blockchain framework with multi-party consensus algorithms to improve the trust in content validation, making it more difficult for an attacker to gain control of the blockchain [84].

References

- [1] Hatch, K.E., 2011. Determining the effects of technology on children.
- [2] Figueira, Á. and Oliveira, L., 2017. The current state of fake news: challenges and opportunities. *Procedia computer science*, 121, pp.817-825.
- [3] Shao, C., Ciampaglia, G.L., Varol, O., Flammini, A. and Menczer, F., 2017. The spread of fake news by social bots. *arXiv preprint arXiv:1707.07592*, 96(104), p.14.
- [4] Hughes, L., Dwivedi, Y.K., Misra, S.K., Rana, N.P., Raghavan, V. and Akella, V., 2019. Blockchain research, practice and policy: Applications, benefits, limitations, emerging research themes and research agenda. *International journal of information management*, 49, pp.114-129.
- [5] Jigsaw Research (2022). News Consumption in the UK: 2022. [online] https://www.ofcom.org.uk/_data/assets/pdf_file/0027/241947/News-Consumption-in-the-UK-2022-report.pdf
- [6] Richter, F. (2020). MEDIA CONSUMPTION. The End of the TV Era? Available at: <https://www.statista.com/chart/9761/daily-tv-and-internet-consumption-worldwide/>.
- [7] Richter, F. (2018). Infographic: Is TV's Reign Nearing Its End? [online] Statista Infographics. Available at: <https://www.statista.com/chart/9761/daily-tv-and-internet-consumption-worldwide/>.
- [8] Statista (2021). Number of internet users worldwide 2005-2017 | Statista. [online] Statista. Available at: <https://www.statista.com/statistics/273018/number-of-internet-users-worldwide/>.
- [9] Petrosyan, A. (2024). *Number of internet users worldwide 2005-2017 | Statista*. [online] Statista. Available at: <https://www.statista.com/statistics/273018/number-of-internet-users-worldwide/>.
- [10] Dobrev, D., 2012. A definition of artificial intelligence. *arXiv preprint arXiv:1210.1568*.
- [11] Ergen, M., 2019. What is artificial intelligence? Technical considerations and future perception. *Anatolian J. Cardiol*, 22(2), pp.5-7.
- [12] Gartner. (2017). Gartner Says By 2020, Artificial Intelligence Will Create More Jobs Than It Eliminates. [online] Available at: <https://www.gartner.com/en/newsroom/press-releases/2017-12-13-gartner-says-by-2020-artificial-intelligence-will-create-more-jobs-than-it-eliminates>.
- [13] Baptista, J.P. and Gradim, A., 2022. A working definition of fake news. *Encyclopedia*, 2(1).
- [14] Baines, D. and Elliott, R.J., 2020. Defining misinformation, disinformation and malinformation: An urgent need for clarity during the COVID-19 infodemic. *Discussion papers*, 20(06), pp.20-06.
- [15] The Onion. (2019). *The Onion | America's Finest News Source*. [online] Available at: <https://www.theonion.com>.

- [16] Kumar, S., West, R. and Leskovec, J., 2016, April. Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes. In Proceedings of the 25th international conference on World Wide Web (pp. 591-602).
- [17] www.kaspersky.co.uk. (2023). How to identify fake news. [online] Available at: <https://www.kaspersky.co.uk/resource-center/preemptive-safety/how-to-identify-fake-news>
- [18] Rocha, Y.M., de Moura, G.A., Desidério, G.A., de Oliveira, C.H., Lourenço, F.D. and de Figueiredo Nicolete, L.D., 2021. The impact of fake news on social media and its influence on health during the COVID-19 pandemic: A systematic review. *Journal of Public Health*, pp.1-10.
- [19] Libguides.tru.ca. (2017). Research Guides: Fake News: Characteristics of Fake News & Media Bias. [online] Available at: <https://libguides.tru.ca/fakenews/characteristics>.
- [20] Hetler, A. (2022). 10 ways to spot disinformation on social media. [online] WhatIs.com. Available at: <https://www.techtarget.com/whatis/feature/10-ways-to-spot-disinformation-on-social-media>.
- [21] www.cisa.gov. (2022). Tactics of Disinformation | CISA. [online] Available at: <https://www.cisa.gov/resources-tools/resources/tactics-disinformation>.
- [22] Chen, K., Luo, Y., Hu, A., Zhao, J. and Zhang, L., 2021. Characteristics of misinformation spreading on social media during the COVID-19 outbreak in China: a descriptive analysis. *Risk Management and Healthcare Policy*, pp.1869-1879.
- [23] Dallo, I., Elroy, O., Fallou, L., Komendantova, N. and Yosipof, A., 2023. Dynamics and characteristics of misinformation related to earthquake predictions on Twitter. *Scientific reports*, 13(1), p.13391.
- [24] Hajli, N., Saeed, U., Tajvidi, M. and Shirazi, F., 2022. Social bots and the spread of disinformation in social media: the challenges of artificial intelligence. *British Journal of Management*, 33(3), pp.1238-1253.
- [25] BBC Bitesize. (n.d.). What are 'bots' and how can they spread fake news? [online] Available at: <https://www.bbc.co.uk/bitesize/articles/zjhg47h>.
- [26] Young Scot. (2022). How Online Bots Spread Fake News. [online] Available at: <https://young.scot/get-informed/ysdigiknow-fake-news-bots/>.
- [27] Liu, D., Wu, Q., Han, W. and Zhou, B., 2016. Sockpuppet gang detection on social media sites. *Frontiers of Computer Science*, 10, pp.124-135.
- [28] Zezulka, L.A. and Seigfried-Spellar, K., 2016. Differentiating cyberbullies and internet trolls by personality characteristics and self-esteem.
- [29] BURNS, E. (2022). What is artificial intelligence (AI)? [online] TechTarget. Available at: <https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence>.
- [30] Aldwairi, M. and Alwahedi, A., 2018. Detecting fake news in social media networks. *Procedia Computer Science*, 141, pp.215-222.
- [31] Raza, S. and Ding, C., 2022. Fake news detection based on news content and social contexts: a transformer-based approach. *International Journal of Data Science and Analytics*, 13(4), pp.335-362.

[32] McKinsey (2022). What is blockchain? | McKinsey. [online] [www.mckinsey.com](https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-blockchain). Available at: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-blockchain>.

[33] Investopedia. (2023). Proof of Work (PoW). [online] Available at: [https://www.investopedia.com/terms/p/proof-work.asp#:~:text=Proof%20of%20work%20\(PoW\)%20is](https://www.investopedia.com/terms/p/proof-work.asp#:~:text=Proof%20of%20work%20(PoW)%20is).

[34] Saiedi, E., Broström, A. and Ruiz, F., 2021. Global drivers of cryptocurrency infrastructure adoption. *Small Business Economics*, 57(1), pp.353-406.

[35] Huckle, S. and White, M., 2017. Fake news: A technological approach to proving the origins of content, using blockchains. *Big data*, 5(4), pp.356-371.

[36] Chen, Q., Srivastava, G., Parizi, R.M., Aloqaily, M. and Al Ridhawi, I., 2020. An incentive-aware blockchain-based solution for internet of fake media things. *Information Processing & Management*, 57(6), p.102370.

[37] Paul, S., Joy, J.I., Sarker, S., Ahmed, S. and Das, A.K., 2019, June. Fake news detection in social media using blockchain. In 2019 7th International Conference on Smart Computing & Communications (ICSCC) (pp. 1-5). IEEE.

[38] Christodoulou, P. and Christodoulou, K., 2020, November. Developing more reliable news sources by utilizing the blockchain technology to combat fake news. In 2020 second international conference on Blockchain computing and applications (BCCA) (pp. 135-139). IEEE.

[39] Webisoft. (2024). 16 Disadvantages of Blockchain: Limitations and Challenges - Webisoft Blog. [online] Available at: <https://webisoft.com/articles/disadvantages-of-blockchain/>.

[40] Quasim, M.T., Khan, M.A., Algarni, F., Alharthy, A. and Alshmrani, G.M.M., 2020. Blockchain frameworks. *Decentralised Internet of Things: A Blockchain Perspective*, pp.75-89.

[41] Humprecht, E. (2019) 'How Do They Debunk "Fake News"? A Cross-National Comparison of Transparency in Fact Checks', *Digital Journalism*, 8(3), pp. 310–327. doi: 10.1080/21670811.2019.1691031.

[42] Du, L., Ho, A.T. and Cong, R., 2020. Perceptual hashing for image authentication: A survey. *Signal Processing: Image Communication*, 81, p.115713.

[43] Bertagnoli, L. (2022). Advantages of Blockchain: 8 Worth Considering | Built In. [online] [builtin.com](https://builtin.com/blockchain/advantages-of-blockchain). Available at: <https://builtin.com/blockchain/advantages-of-blockchain>.

[44] www.kaggle.com. (n.d.). Fake News Detection Datasets. [online] Available at: <https://www.kaggle.com/datasets/emineyetm/fake-news-detection-datasets>.

[45] Reuters Editorial (2021). *Business & Financial News, U.S & International Breaking News* / Reuters. [online] U.S. Available at: <https://www.reuters.com>.

[46] <https://www.reuters.com/info-pages/terms-of-use/>


[47] Farid, Hany. "An overview of perceptual hashing." *Journal of Online Trust and Safety* 1.1 (2021).

- [48] Zhao, H. and He, S., 2016, August. A retrieval algorithm for encrypted speech based on perceptual hashing. In *2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)* (pp. 1840-1845). IEEE.
- [49] xorbin.com. (n.d.). SHA-256 hash calculator | Xorbin. [online] Available at: <https://xorbin.com/tools/sha256-hash-calculator>.
- [50] Kotecha, S. (2024). *Under 100 spaces in men's prisons in England and Wales*. [online] BBC News. Available at: <https://www.bbc.co.uk/news/articles/c0rw48nj282o> [Accessed 3 Sep. 2024].
- [51] Du, L., Ho, A.T. and Cong, R., 2020. Perceptual hashing for image authentication: A survey. *Signal Processing: Image Communication*, 81, p.115713.
- [52] Cao, W., Feng, W., Lin, Q., Cao, G. and He, Z., 2020. A review of hashing methods for multimodal retrieval. *IEEE Access*, 8, pp.15377-15391.
- [53] Choi, F.Y., Wiemer-Hastings, P. and Moore, J.D., 2001. Latent semantic analysis for text segmentation. In *Proceedings of the 2001 conference on empirical methods in natural language processing*.
- [54] Huang, Z. and Liu, S., 2020. Perceptual image hashing with texture and invariant vector distance for copy detection. *IEEE Transactions on Multimedia*, 23, pp.1516-1529.
- [55] McKeown, S. and Buchanan, W.J., 2023. Hamming distributions of popular perceptual hashing techniques. *Forensic Science International: Digital Investigation*, 44, p.301509.
- [56] LaValley, M.P., 2008. Logistic regression. *Circulation*, 117(18), pp.2395-2399.
- [57] Shah, K., Patel, H., Sanghvi, D. and Shah, M., 2020. A comparative analysis of logistic regression, random forest and KNN models for the text classification. *Augmented Human Research*, 5(1), p.12.
- [58] Dutta, S. and Saini, K., 2022. Blockchain implementation using python. In *Advancing Smarter and More Secure Industrial Applications Using AI, IoT, and Blockchain Technology* (pp. 123-136). IGI Global.
- [59] Kwak, C. and Clayton-Matthews, A., 2002. Multinomial logistic regression. *Nursing research*, 51(6), pp.404-410.
- [60] Franke, M. and Degen, J., 2023. The softmax function: Properties, motivation, and interpretation.
- [61] Tan, K. and Bellec, P.C., 2023. Multinomial logistic regression: Asymptotic normality on null covariates in high dimensions. *Advances in Neural Information Processing Systems*, 36, pp.70892-70925.
- [62] Devlin, J., Chang, M.W., Lee, K. and Toutanova, K., 2019, June. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)* (pp. 4171-4186).
- [63] Rennie, J.D., Shih, L., Teevan, J. and Karger, D.R., 2003. Tackling the poor assumptions of naive bayes text classifiers. In *Proceedings of the 20th international conference on machine learning (ICML-03)* (pp. 616-623).

- [64] Miotto, R., Li, L., Kidd, B.A. and Dudley, J.T., 2016. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. *Scientific reports*, 6(1), p.26094.
- [65] Ahmed, H., Traore, I. and Saad, S., 2017. Detection of online fake news using n-gram analysis and machine learning techniques. In *Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments: First International Conference, ISDDC 2017, Vancouver, BC, Canada, October 26-28, 2017, Proceedings 1* (pp. 127-138). Springer International Publishing.
- [66] Singh, A., Ugale, A., Shah, N. and Sankhe, A., 2021. Fake News detection using logistic regression & multinomial naive Bayes.
- [67] Abramovich, F., Grinshtein, V. and Levy, T., 2021. Multiclass classification by sparse multinomial logistic regression. *IEEE Transactions on Information Theory*, 67(7), pp.4637-4646.
- [68] Qayyum, A., Qadir, J., Janjua, M.U. and Sher, F., 2019. Using blockchain to rein in the new post-truth world and check the spread of fake news. *IT Professional*, 21(4), pp.16-24.
- [69] Agerskov, S., Pedersen, A.B. and Beck, R., 2023, April. Ethical Guidelines for Blockchain Systems. In ECIS.
- [70] Zhang, L., Wang, T. and Liew, S.C., 2022. Speeding up block propagation in Bitcoin network: Uncoded and coded designs. *Computer Networks*, 206, p.108791.
- [71] Georgiadis, E., 2019. How many transactions per second can bitcoin really handle? Theoretically. *Cryptology ePrint Archive*.
- [72] Baliga, A., 2017. Understanding blockchain consensus models. *Persistent*, 4(1), p.14.
- [73] Lashkari, B. and Musilek, P., 2021. A comprehensive review of blockchain consensus mechanisms. *IEEE access*, 9, pp.43620-43652.
- [74] Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K. and Dyer, C., 2016. Neural architectures for named entity recognition. *arXiv preprint arXiv:1603.01360*.
- [75] Storage, B.U.C.A., Design And Implementation Of Msha256 On Blockchain Using Content Addressable Storage Patterns.
- [76] Buterin, V., 2014. A next-generation smart contract and decentralized application platform. *white paper*, 3(37), pp.2-1.
- [77] Cao, Z., Long, M., Wang, J. and Yu, P.S., 2017. Hashnet: Deep learning to hash by continuation. In *Proceedings of the IEEE international conference on computer vision* (pp. 5608-5617).
- [78] He, L., Huang, Z., Chen, E., Liu, Q., Tong, S., Wang, H., Lian, D. and Wang, S., 2023. An efficient and robust semantic hashing framework for similar text search. *ACM Transactions on Information Systems*, 41(4), pp.1-31.
- [79] Androulaki, E., Barger, A., Bortnikov, V., Cachin, C., Christidis, K., De Caro, A., Enyeart, D., Ferris, C., Laventman, G., Manevich, Y. and Muralidharan, S., 2018, April. Hyperledger fabric: a distributed operating system for permissioned blockchains. In *Proceedings of the thirteenth EuroSys conference* (pp. 1-15).

- [80] Qiu, Z., Pan, Y., Yao, T. and Mei, T., 2017, August. Deep semantic hashing with generative adversarial networks. In Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval (pp. 225-234).
- [81] Krishna, K., Song, Y., Karpinska, M., Wieting, J. and Iyyer, M., 2023. Paraphrasing evades detectors of ai-generated text, but retrieval is an effective defense. Advances in Neural Information Processing Systems, 36, pp.27469-27500.
- [82] Aponte-Novoa, F.A., Orozco, A.L.S., Villanueva-Polanco, R. and Wightman, P., 2021. The 51% attack on blockchains: A mining behavior study. IEEE access, 9, pp.140549-140564.
- [83] Giabelli, A., Malandri, L., Mercorio, F., Mezzanzanica, M. and Nobani, N., 2022. Embeddings evaluation using a novel measure of semantic similarity. Cognitive Computation, 14(2), pp.749-763.
- [84] Zhou, J., Feng, Y., Wang, Z. and Guo, D., 2021. Using secure multi-party computation to protect privacy on a permissioned blockchain. Sensors, 21(4), p.1540.

Appendix


World ▾ Business ▾ Markets ▾ Sustainability ▾ Legal ▾ Breakingviews ▾ Technology ▾ Investigations ▾ More ▾
My News 🔍
Sign In
Register

8.Restrictions on Use

a.THE SERVICE AND THE CONTENT IS PROVIDED BY REUTERS AND ITS LICENSORS TO YOU FOR YOUR PERSONAL USE AND INFORMATION ONLY. YOU MAY NOT USE THE SERVICE OR THE CONTENT FOR ANY COMMERCIAL PURPOSE.

b.You promise that you are accessing, using, and/or registering to the Service in your personal, individual capacity (except, if applicable, where your access falls under a group subscription agreement in accordance with Section 1(f)). You agree not to use, transfer, distribute, and/or dispose of the Service or Content in any manner that could compete with the business of Reuters or any of its partners. You may not use the Content or Service, including without limitation, any Content made available through an RSS feed, in any commercial product or service, without our prior written consent.

c.Without our prior written consent, you may not:

- remove, alter, forward, scrape, frame, in-line link, copy, sell, distribute, retransmit, create derivative works or otherwise make available (to third parties and/or on another website, app, blog, product, or service) the Content, except as occasionally permitted by certain sharing features in the Service that explicitly allow you to share Content or links to Content with a few other individuals.
- use any robots, spiders, scripts, service, software or manual or automatic device, tool, or process designed to data mine or scrape the Content, data or information from the Service, or otherwise access or collect the Content, data or information from the Service using automated means.
- use the Service or Content for any purpose relating to the development and training of any machine learning ("ML") and artificial intelligence ("AI") activities and technologies, including, but not limited to using the Service or Content to build, train, enhance or tune any AI or ML technologies. You may not, at any time, directly or indirectly, use the Service or Content for any purpose relating to the development and training of any machine learning ("ML") and artificial intelligence ("AI") activities and technologies, including, but not limited to using the Service or Content to build, create, train, retrain, enhance or tune any AI or ML technologies (whether belonging to you or a third party).
- reverse engineer, circumvent, decrypt, decompile, disassemble, modify, or otherwise attempt (using services, software, or any manual or automatic device, tool, or process) to alter or interfere with the Service, any feature or component of the Service, any restriction, condition, or technological measure that controls access to the Service in any way, or any content in the Service; or make any unauthorized use thereof.

Figure 24: Terms of usage for news articles from Reuters

Figure 25: Terms of usage for news articles from Reuters