Design Project - Final Report

Making use of Blockchain and Ethereum to create fake news detection model

Pham Do Minh Quang

1902021

Academic Supervisor:
Purnima Murali Mohan

Submitted as part of the requirements for TLM3001 Design Project

SINGAPORE INSTITUTE OF TECHNOLOGY

Design Project Report Submission Form

Declaration of Authorship

Name of candidate: Pham Do Minh Quang	Student ID Number: 1902021
Degree: Telematics (Intelligent Transportation Systems Engineering)	
Design Project Title: Making use of Blockchain and Ethereum to create fake news detection model	
Academic Supervisor(s): Purnima Murali Mohan	Cluster: Infocomm Technology (ICT)
Industry Supervisor: (*if applicable*)	Organization: (*if applicable*)

I hereby confirm that:

1. this work was done wholly or mainly while in candidature for a degree programme at SIT;

2. where any part of this design project has been previously submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;

3. I have acknowledged the use of all resources in the preparation of this report;

4. the **report contains** / **does not contain** patentable or confidential information *(strike through whichever does not apply)*;

5. the work was conducted in accordance with the research integrity policy and ethics standards of SIT and that the research data are presented honestly and without prejudice. The SIT Institutional Review Board (IRB) approval number is __________________________ *(where applicable)*;

6. I have read and understood the University's definition of plagiarism as stated in the SIT Academic Policies Scheme 14: Academic Integrity.

   Plagiarism is the copying, using or passing off of another's work as one's own work without giving credit to the author or originator, and also includes self-plagiarism. For example, reusing, wholly or partially, one's previous work in another context without referencing its previous use.

27/07/2022

Signature of Student **Date**

1. Abstract

The volume of text that is easily accessible online has surpassed hundreds of billions of words and is still increasing. In spite of this, algorithms that have been trained on a corpora of only one million words or less are still being tested, compared, and optimized for the majority of basic natural language tasks. At the same time, the speed and volume of fake news propagating on the Internet have created a big alarm among the public, especially on social media platforms. In this report, we would like to discuss the current state of generating labeled data in NLP and give our take on the matter of the issue. An in-depth review will be provided that discusses the cause and effect of fake news on the life of people who are exposed to it. Next, we look at common approaches that have been widely adopted to mitigate the spread of fake news and explain our view on why we think machine learning is a better, if not superior, tool for detecting and filtering out fake news on social media. We also present our system that makes use of Blockchain technology to provide incentives for anyone to start building a bigger corpus of labeled fake news and subsequently increase the effectiveness of our fake news detection model. Last but not least, we analyze the effect of training size on NLP deep learning to validate our concern and also point out a few issues that can affect the performance of a deep learning model such as skewness in training data.

2. Introduction

2.1 Existing problems in Natural Language Processing

Natural Language Processing (NLP) is a machine learning technique that aims to automate the process of understanding complicated linguistic structures and has been widely adopted to tackle different language problems such as sentiment analysis, translation, speech understanding and so on. For the last decades, researchers in this area have mainly put their focus on comparing and creating a “good” linguistic model. The major issue is that these models are often trained on small datasets with limited sample size which have no guarantee of being reliable for real-world use [15]. This is not a new problem in NLP. In fact, Michele and Brill (2001) shared similar sentiment on this topic, regarding the lack of labeled training corpora available for NLP research, even though we have access to substantial amounts of texts thanks to the advancement of the Internet [18]. In the paper, they showcased their study that tried to tackle Word Sense Disambiguation problem [19]. In particular, the authors insisted on using standard algorithms instead of opting for complex models, and only increased the training data by orders of magnitude. The authors observed a significant jump in the test’s accuracy and suggested that other researchers should emphasize on increasing the training corpora in NLP rather than spending more resources on inventing new algorithms.

The two reasons that makes annotating data for NLP research difficult was that labeling costs a lot of annotators-hours and money [20] Crowdsourcing has become an ideal alternative for researchers to obtain a high quantity of labeled dataset with significantly lower cost. However, the quality of these labels is questionable and varied depending on the number of annotation tasks, annotator expertise and evaluation criteria. [21]. The authors also talked about the 2 methods to improve the quality of the crowdsource annotated data. “Quality Control on Task Designing” is a control mechanism which researchers provide the annotators with specific guidelines to follow and helps to ensure the annotation accuracy. The other is called “Quality Improvement after Data Collection” describes a mechanism in which requesters hire other crowdsource services to check the quality of the prior annotators. This system can filter out “spammers” and prevent these erroneous labels from influencing the models.

2.2 Fake news and Why is it an important issue?

Fake news and misinformation spreading has been one of the most pressing issues facing the digital age. This was brought about by the explosion of the Internet in 2008, which has reshaped the way I communicate and share information. Social media allows fake news to be created and published online faster and cheaper when compared to traditional news media such as newspapers and television. [22] Furthermore, information on these platforms can spread fast to mass audiences across the world. According to the survey conducted in Singapore published by R. Hirschmann, there are 83 percent of respondents in Singapore who receive news from online sources and social media platforms make up 58 percent of this portion (refer Figure 1) [1]. In turn, this amplifies the effect of fake news on a wide range of users. There are multiple studies proven that humans are prone to be influenced by false information due to their personal biases. The echo chamber is also a significant factor contributing to the spread of false information. On top of that, the proliferation of fake news brings massive political and economic benefits to great entities. This encourages this issue to persist unless there are immediate and proactive efforts from researchers and companies to prevent the spread of fake news.

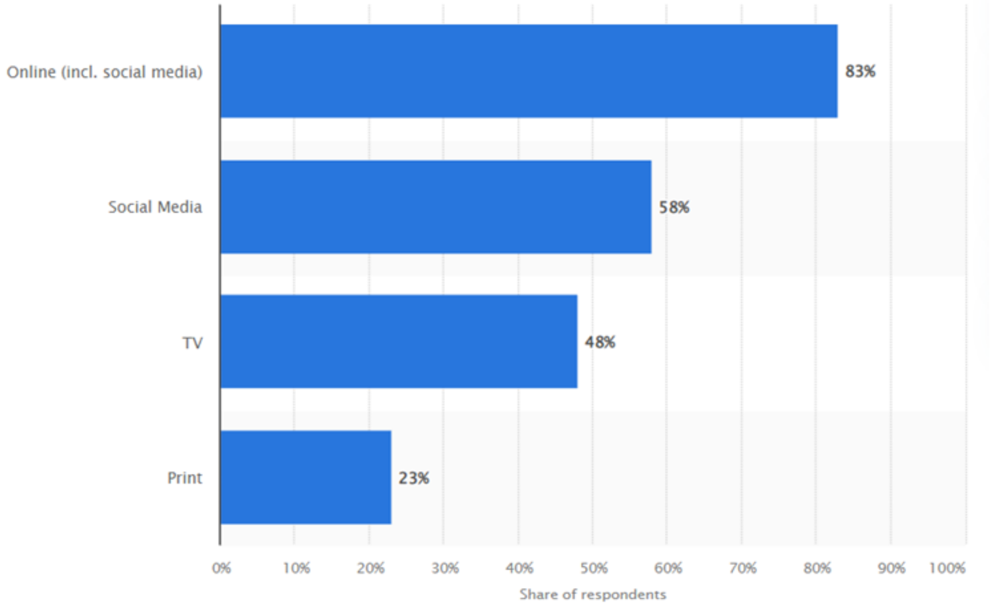


Figure 1: Share of respondents in Singapore who receive news from the following sources as of February 2022 [1]

The term `Fake news` has gained a lot of popularity and public attention among social media and news outlets. For example, during the 2016 presidential election, there were estimated to be about 7,367,000 online social interactions on the 20 best-performing election stories that are published by the top 19 news outlets on Facebook while this number is even higher (estimated at around 8,711,000) for the top 20 frequently shared fake news stories that are published by hoax sites [24]. Research has shown that compared to the truth, fake news on Twitter is typically retweeted by many more users and spreads far more rapidly, especially for political news [25]. This effectively tells me that fake news has a high tendency to reach a large number of people due to the nature of recommendation algorithms. These algorithms are engineered to promote content that can attract utmost attention from their users, as well as retain a high level of engagement for long periods of time. This is because, for these companies, the longer the user is spending on these sites, the higher their profits.

Social and psychological factors are also responsible for the propagation of fake news within the Internet community. For example, humans have a high tendency to trust deceptive information as long as it agrees with his/her personal view. A study by Koriat [26] demonstrated that humans actively try to find information that supports their existing beliefs or hypothesis. They often give more weight to these information without questioning its credibility. Not only that, the author said people who have confirmation bias, often have difficulty in observing alternative views and are confined to their personal beliefs, which prevents them from finding out the truth. In addition, research showed that people are more likely to remember facts and events that have been repeatedly mentioned, even when the repeated mention is in the context of a retraction or disposing of misinformation [27]. Research in this area includes linguistic signals of a rumor [28] and models of spread, which help in determining how to contain it, and how many fact-checkers are needed to contain a hoax [29]. Hoaxy, an open platform to study misinformation and fact-checking on Twitter, is useful in modeling how to disrupt the spread of a rumor [30].

The report includes an overview of existing data sets and the challenges to getting them to scale. I present a crowdsourced website that makes use of Blockchain and Cryptocurrency to attract a wider pool of Internet users to participate in labeling, in hopes of generating an adequate amount of labeled data to be used to train a machine learning algorithm to detect fake news. In Literature research, I examine available fake news detection methods using NLP; review the scale and effectiveness of current strategies for data collection in fake news;

and illustrate the advantages of using Blockchain technology as a payment service to existing digital payment. In the next section, I explain the problems that our system tries to solve; introduce the technical terms and tools that I have used to develop the solution; and point out the new insights that I have learnt while implementing the system. Finally, I evaluate the performance of the models and demonstrate the features of my website.

3. Literature Research

3.1 Fake news Detection

3.1.1 Manual Fact-checking

Expert-based Fact-checking

Fake news is a difficult problem to solve since it requires adequate background and knowledge to break down the linguistic features. Yet, the scale of fake news is huge, with thousands of articles published daily on social media platforms such as Facebook and Twitter. Thus, there is no individual effort that can combat all of this misinformation. Many organizations have offered expert-level verification services to support entities and companies in combating the spread of misinformation. For example, Facebook has collaborated with the International Fact-checking network (IFCN), which provides verification of news stories published on its network. The International Fact-Checking Network (IFCN) is an independent fact-checking organization that has decided to allocate all their resources and efforts to ensure the veracity of news and other materials on the Internet. The staff are well trained to cover a wide range of topics, including politics, business, and entertainment. The organization operates on the principles of professional commitments to integrity, fairness, reliability, neutrality, and transparency of techniques and information. [31]. The problem with this strategy is that it relies on trust between the social media network and the fact-checking organization. In the paper, Chen (2017) [32] showed us that moderators can introduce their biases during the process of verifying news. He also pointed out that this way of fact-checking puts tons of pressure on the moderators since the quantity of staff has proven to be too limited as compared to the size of fake news. This incurred a heavy burden on the moderators and led to, firstly a negative influence on their personal health, and secondly affects the reliability of the verification. Lastly, this method often will not scale well to a large amount of data on the Internet due to the fact that it is merely time-consuming and expensive to maintain by small-scaled companies.

Crowd-sourced Manual Fact-checking

Crowdsourcing has emerged as a viable alternative way to cope with the rise of fake news due to its being cost-effective. For example, Amazon offered their crowdsourcing marketplace called Mechanical Turk to provide users with cheap, high-quality human resources to assist in labor-intensive tasks such as data entry, transcription, and labeling. In fact, we have seen many groups and entities take advantage of the cheap labor force to implement their solutions to disseminate the characteristics of fake news data [17]. However, many may question the quality and validity of the crowdsourcing results, and there are concerns about privacy as well as the potential for spammers to purposely influence the outcome of the verification process. [33] All in all, it is apparent that we should not rely heavily on crowdsourcing as the service is not guaranteed to perform well on technical or niche topics such as fake news verification. Instead, many attempts of crowdsourcing focus on gathering training data on fake news sources and use the results in order to improve the accuracy of the automated tools, and prominently supervised learning methods in the field of AI. A study by Jing and Zhang (2016) [33] has shown that, by providing the annotators with adequate instructions, the tool has a capability to yield quality results that are sometimes on a par with expert-level annotation.

3.1.2 Automated Fact-checking

Automated techniques such as Machine Learning and AI have been a dominant factor to tackle a myriad of problems in the field of NLP. Fake news is also no exception. There have been many papers published by researchers that have studied how to understand the linguistic features of fake news and claimed to create optimal algorithms that can automatically detect them at a high level of accuracy. These algorithms also facilitate us to compete with the increasing volume of fake news that is being posted online. Xinyi and Reza (2018) [35] summarized a few principles of fake news detection that have been applied by researchers to reveal the basic components of misinformation, in turn, build their solution revolving around these theories. There are two theories that the author discussed which are 'News-related' and 'User-related'. News-related theory looks at the underlying mechanism and structure of fake news, as well as analyzes the sentiment of the authors and the context that made up the news. Meanwhile, User-related theory concentrates on the activities and behavior of the users with fake news and attempts to understand about the effects they have on the public.

These 2 theories have inspired many papers and researches regarding the fake news detection and facilitate them to produce quality algorithms. For example, there have been multiple attempts by Feng et al. 2012 [36] ; Zhou et al. (2019) [37] to use State-of-the-art machine learning models such as Support-vector machine (SVM), Random forest, so on, to understand the writing style of the authors. Word embedding such as Word2Vec, Glove is often used to create a semantic representation of the content and is processed by deep learning engines for the detection of fake news. These popular text-based algorithms such as RNN, LSTM, transformer, etc have been proven to perform very well on breaking down the content of fake news as close as human-level. A prominent example of such algorithms is the Pre-trained Transformer (GPT) [38] created by Open AI Lab, which was trained on billions of text datasets on Reddit and is capable of comprehending text-based materials and generating dialogues that are equivalent or, some might argue even better than human capability. Little unconventional, some algorithms such as Event Adversarial Neural Network(EANN) [39] tried to take advantage of the min-max game to tackle text classification. Basically, there will be 2 models (Generator and Discriminator) that are responsible to generate quality fake news and the other will try to identify them respectively.

3.2 Data: Where to find massive labeled fake news articles?

In recent years, there have been multiple attempts from researchers to increase the quantity of labeled fake news dataset. In this section, I will briefly look at a few promising strategies that have been implemented to generate labeled fake news dataset. The objective of these approaches is to accumulate an adequate amount of labeled news data, which can be used to train machine learning to learn the linguistic structures of fake news, thereby distinguishing fake news from real news.

Rashkin et al. (2017) [14] gathered more than 60,000 pieces of news from 7 different \`unreliable\` news websites. These articles were then grouped into propaganda, hoax, satire based on the reputation of the news sources. However, Fatemeh and Maite (2019) described this strategy of labeling the news as \`noisy\` as the reputation of the news sources does not reflect the reliability of the news, therefore it will hinder the machine learning models from learning the linguistic structure of the fake news. [15]

Allcott and Gentzkow (2017) attempted to take advantage of available fact-checking websites such as Snopes, Politifact and Buzzfeed to gather a database of fake news. The author uses web scraping to extract stories from these websites that relate to the 2016 Presidential Elections. This results in 156 fake news articles. [16]

Pérez-Rosas et al. (2017) [17] has a very interesting approach where he firstly collected a dataset of 240 legitimate news from reputable news websites such as ABCNews, CNN, USA. Today, etc and hired Amazon Mechanical Turk (AMT), a crowdsourcing marketplace, to write a `fake` version of the news. It took approximately 5 days for the workers to create 240 fake news stories that followed provided instructions and avoided unrealistic contents. The author observed that AMT workers successfully replicate the reporting style of original pieces while injecting grammatical errors and utterances into the fake news. In the result section, the author showed that the 2 annotators achieved around 71 percent accuracy, which is less accurate than his developed system.

3.3 Payment with Blockchain technology

3.3.1 What is blockchain

Felin and Lakhani (2018) [12] compared the difference between a traditional ledger that every bank uses with the blockchain technologies. Both of them have the same responsibility in recording and verifying transactions and terms of engagements. Moreover, they can keep track of information that is related to the transactions and can be used for further analysis such as the users' credibility. The main distinctive point about blockchain that differentiates it from traditional ledger is the fact that blockchain is distributed and digital compared to its counterpart. Every user on the blockchain maintains a copy of the ledger. This ledger is "instantly and simultaneously" updated about any new transaction created on the block using "advanced computational algorithms and cryptographic locks". With a mature network such as Bitcoin or Ethereum, it is nearly impossible for hackers to tamper the record of transactions since they will require an enormous amount for computation to compromise not only one ledger, but 51% of the users' ledger in the network (Jennifer 2016) [13].

3.3.2 Benefits of Cryptocurrency over Digital Currency

The paper that is published by Rabiul et al. (2018) [5] explained to us the fundamental differences between cryptocurrency and digital currency which lies on how our transactions are made and who I entrusted our currencies to (Refer to appendix). For digital currency, I entrusted the bank as a third-party who holds on to our currency and transfers it to the other party when both parties have reached a consensus. In the context of cryptocurrency, transactions are peer to peer and do not require a middle man to regulate. According to Peters and Panayi [6], blockchain is the new era of financial operations such as clearing and settlement, payments, trading, and insurance. In the current financial scheme, payments need to pass through different layers of intermediaries, including clearinghouses, banks, and other financial institutes, where they are examined through a complicated, time-consuming, and expensive process [3]

Reduce working days

According to Peters and Panayi (2016) [6], a banking system keeps track of all customers' transactions for a specified period, such as 90 days. Usually, auditors are responsible to do manual checks on the validity of each transaction and update the customers' ledgers accordingly. As the number of transactions in each ledger increases, this process can be extremely labor intensive and time-consuming. Beside that, Ye and Chen (2016) [7] mentioned that bank payment involves a series of complicated processes, including bookkeeping, transaction reconciliation, payment initiation, etc. which result in long working days to complete a single transaction. Blockchain technology resolves these issues by having a distributed ledger that is updated instantly and automatically which ensures the integrity of each transaction and reduces transaction time significantly. With the use of smart contracts, payments will be distributed without delay or any intermediary's involvement once they have reached a predetermined agreement. For example, Ripple, backed by blockchain technology, only took 10s to complete a cross-border transaction while a banking system requires 2 days for a transaction to arrive due to clearing procedures.

Low transaction fee

Botros Kfoury (2021) [8] touched on the use of blockchain to reduce overhead costs of transactions via digital payments. These costs are incurred due to financial operations such as clearing and settlement, payments, trading, and insurance. An interview conducted by Tilooby with the head of Walmart [9] reported that Visa takes a hefty fee of around 3% per transaction, and this cost is alternatively transferred to Walmart's customers. From the article that was published by Joan (2020) [10], big companies that offer digital payments services such as Visa, Mastercard charge high transaction fees ranging from 2.3% to 2.5% for online transactions. The author also mentioned Paypal who charges fees based on the location of the users, in which an international payment could cost a user 4.4% transaction fee plus the fixed fee per transaction ($0.5 for any payments in Singapore). Meanwhile, according to Ychart (21/07/2022), Ethereum blockchain network only charges a fixed rate at 0.7637 USD per transaction [11]

4. Fakenews Voting website

4.1 Problem Analysis

Producing the right answer from the available data is the task of labeling, and in this case, the right response is referred to as a label. The data must be correctly labeled in order for deep learning to engage in supervised learning. Furthermore, precise labeling is essential since supervised learning based on inaccurate labeling will impair the performance of the model. Many deep learning technologies, including image processing and natural language processing, call upon labelings. With regard to Twitter fake news in particular, we introduce labeling for text analysis and briefly discuss the use of active learning technology.

The main objective of this report is to emphasize the importance of labeled data in NLP in general and highlight the challenge of constructing labeled fake news corpora that are both robust and useful to develop quality detection algorithms. At the same time, we bring to the community a practical solution that fosters the process of labeling text data on a massive scale. The topic of our data in this report focuses on misinformation about the global pandemic, Covid-19 because we want to reduce the diversity of background and knowledge which facilitate our users to focus on the topic. This will ensure the quality of the labels since our users only need to focus on researching and fact-checking on single topics and prevent them from introducing personal bias or incorrect information. However, this can be changed easily depending on the use case of the project.

Active Learning

When there are unlabeled datasets, active learning was designed to produce high-performance deep learning models. The active learning approach does not wait for scientists and programmers to categorize all unlabeled datasets. The current deep learning model analyzes the dataset and provides academics and developers with some of the most challenging-to-assess data. Following manual labeling of the subject data with priority by researchers and developers, the model continues to learn using the newly labeled data. Researchers and engineers are once again presented with some of the most challenging material by the re-trained deep learning model, which repeats the earlier phase of making a judgment on the provided dataset.

We use active learning strategy to create a feedback loop that feeds the fake news data that has been annotated by the users into machine learning, in hopes of improving the accuracy of our model by a large margin (refer Figure 2). Softmax function allows us to gauge the confidence of our model's prediction [34]. Only those predictions with a low confidence score will be displayed on our website and re-evaluated by our users to ensure the veracity of the predictions. Basically, we want to focus on amending our model's incorrect predictions instead of spending extra effort on verifying news that has a high chance to yield the same answer as the model.

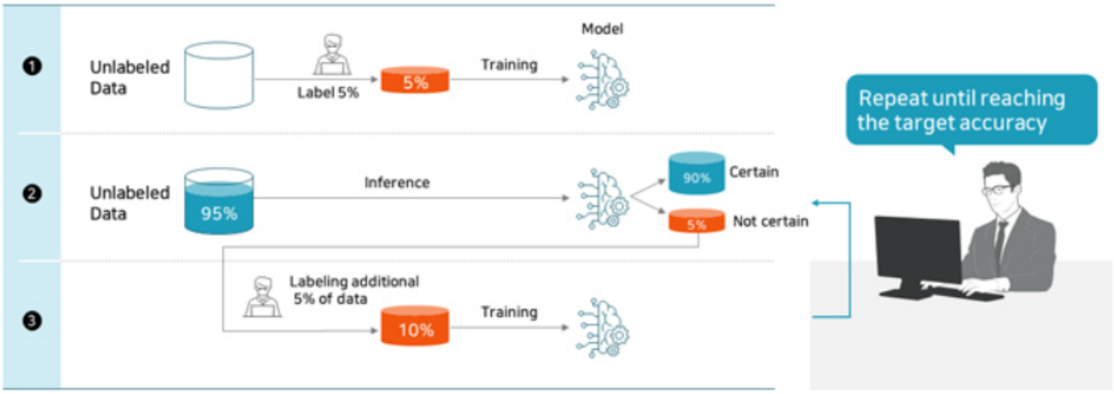


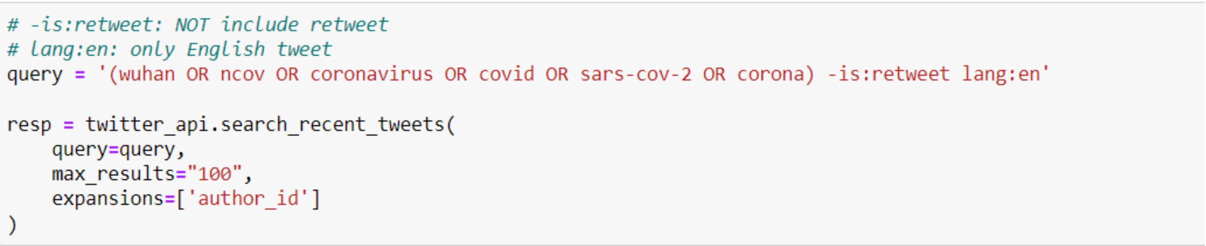
Figure 2: Active Learning [40]

The advantages of Blockchain technology enable us to incentivize our users with ERC20 fungible tokens to participate in our crowdsourcing system for labeling NLP dataset. This gives our users the opportunity to earn tokens by labeling our fake news dataset to help create a better classification tool. We have also created a website to facilitate users to submit their vote (fake or real) on a collection of news. The result of their vote will be processed by our smart contract to formulate the final verdict of the news. Smart contract will then use the verdict to calculate the number of tokens that have been earnt or lost by users and transfer it to them automatically.

4.2 Technical Content

4.2.1 Data collection

Our unlabeled dataset is collected using TwitterAPI to easily get our hands on a large pool of tweets and news stories. For the purpose of this report, we only retrieved 100 tweets for experimenting and demonstrating on our website. In the Twitter backend's URL, we set our query to retrieve tweets that are related to "wuhan", "ncov", "coronavirus", "covid", “corona” or "sars-cov-2", since they all describe the same thing. We also added extra arguments in our query to retrieve only English tweets and filter out retweets. Lastly, we added author_id field in the Tweepy.Client’s parameter to retrieve the source of the tweet for the website (refer Figure 3).



```
# -is:retweet: NOT include retweet
# lang:en: only English tweet
query = '(wuhan OR ncov OR coronavirus OR covid OR sars-cov-2 OR corona) -is:retweet lang:en'

resp = twitter_api.search_recent_tweets(
    query=query,
    max_results="100",
    expansions=['author_id']
)
```

Figure 3: Fetch tweets about Covid-19 using Tweepy.Client

These tweets will be used as input data for pre-trained models to produce a prediction and confidence score. Those tweets with a confidence score that is lower than 80% will be stored in a separate MongoDB collection which will be displayed on the front-end for the users to vote.

4.2.2 Decentralized Application with Ethereum (dApp)

Overview

Ethereum is an open-source distributed ledger system for smart contracts, that provides decentralized applications (DApp) for storing and exchanging data and value. Ethereum's blockchain enables applications to run and be audited without reliance on a trusted third party. Unlike Bitcoin, smart contracts are fully executable in the blockchain and can be validated in a transparent and peer-to-peer manner, making them much easier for developers to develop and test their applications. Due to their immutable nature, smart contracts allow for extremely secure applications. In other words, developers can deploy and mine their contract on the Ethereum blockchain and never have to fear that malicious parties might modify the logic of their smart contract. A smart contract is programmed using Solidity, which is a high-level, object-oriented programming language. Solidity inherits many of the features and coding practices of Python and JavaScript, which facilitates developers to easily learn and build a DApp. Anyone can invoke a smart contract easily having access to an address and an Application Binary Interface (ABI) of the contract. When a smart contract is invoked, a particular node on the blockchain network will be assigned and responsible to execute the contract's logic and return a transaction to the caller.

Smart contract

We use ERC-20 to regulate all transactions to our smart contract. In short, ERC-20 is a standard fungible token (FT) that is interchangeable and can represent the assets of a user with the Ethereum blockchain. The flow of our application can be described in this manner (Refer to figure). After a user has successfully logged in to our website, there will be a display of tweets that they can stake and submit their vote. Based on the quantity of tokens and the credibility score, the smart contract maintains two significant scores that will influence the verdict of a tweet. These tokens, which are ERC-20 compatible thanks to the OpenZeppelin library, are used to reward or penalize validators based on the quantity of tokens staked. Each validator's credibility scores are kept track of. The credibility score is strictly linearly increased and exponentially decreased. This is done in order to reduce the impact that rogue validators may have on the POS algorithm's transaction verification. After a tweet obtains enough votes from users, smart contract will emit an Event to be received by our backend, which will take corresponding actions including processing input data and querying MongoDB database. Subsequently, it will create a new transaction with a parameter that contains parsed data from MongoDB database for smart contract to calculate the reward that earnt by different users. One difficulty we found is that Solidity cannot deal with floating values. However, we resolved this issue with fixed-point arithmetic [44], in which we multiplied all numbers that were sent to smart contract with 10^{10} to maintain its fractional representation.

Interaction flow between React, NodeJS and Smart contract

Ethers and HardHat libraries were mainly used for creating and deploying the contract on the Ethereum blockchain. Ethers handles the nuances in communicating with MetaMask and sending transactions to a deployed contract on the blockchain. The big merit of Ethers compared to its ancestor, Web3, is that Ethers provides its users with comprehensible documentations and a simple and intuitive API. HardHat is used to simplify the step to connecting and deploying a smart contract to the Rinkeby Testnet.

Rinkeby is a blockchain instance that behaves similar to a real network but incurs no cost to developers, which is useful for testing purposes. We also used Remix Integrated Development Environment (IDE) to debug and test our smart contract's logic since it provides a great compiler that checks for syntax and type error in Solidity, plus facilitates us to deploy and execute the contract in Testnet quickly.

Our website is built using React library, which allows developers a quick and easy way to create a dynamic website with an abundance of features and libraries. React is extremely popular and obtains a huge community that can support any issues that we have faced. Bootstrap framework is used to easily decorate and design the layout and interface of our website to enhance user's experience. Moreover, we have looked into way to allow our users to have smooth interaction with our website and create a ‘Non Blocking UI design’ (refer Figure 4) using long polling and Solidity events [41]

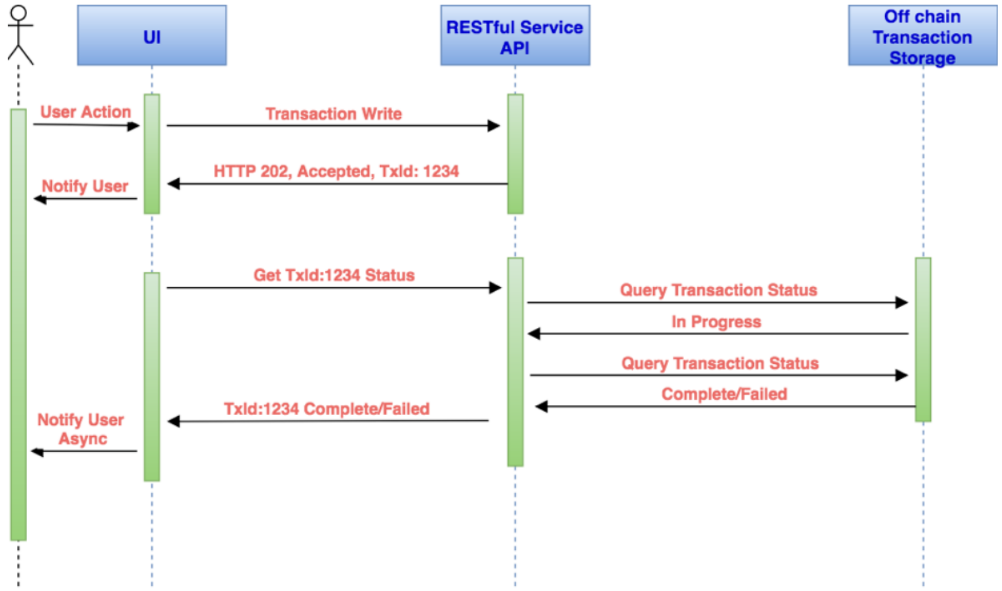


Figure 4: Non Blocking UI design

Off-chain storage

In the context of blockchain development, maintaining and updating a data structure object is not straightforward as a traditional database. Any transaction that is sent to a blockchain network can cost a lot of time and money. At the time of this writing, it costs 0.7637 USD to mine a simple transaction. If developers are not mindful of how they stored information on the block, they could suffer a severe loss of revenue. By integrating Off-chain storage [42] into our system, it can help reduce the query calls to Blockchain and also provide richer search capabilities and faster reads. Furthermore, Off-chain storage levitates the user experience and responsiveness of the UI by faster retrieving data to the client. In this application, we will store all of our data in a traditional database (MongoDB) and only pass values that are required by the smart contract's logic as a parameter. All functions in smart contract will emit an Event [43] while our backend listens to these Events, processes its parameter’s values and stores them in the database (Refer to Figure 5).

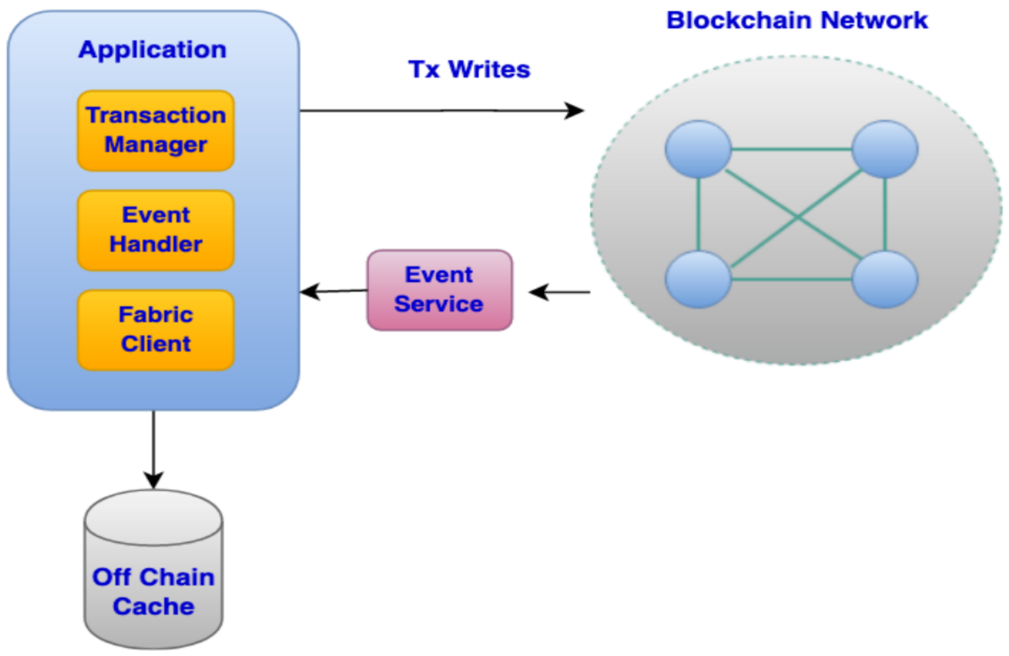


Figure 5: Interaction with the off chain cache, event service and the Blockchain []

4.2.3 Applying Pretrained model

In the midst of the rise of AI development, it is not hard to find deep learning models that have been trained to serve a wide range of use cases and purposes. These models are known as Pretrained models and they are extremely useful in the field of AI as not everyone has enough capacity of dataset and computational resources to train a model. Therefore, in our solution, we would like to avoid the hefty process of training deep neural networks from scratch and instead opt for COVID-Twitter-BERT v2 [2], pre-trained BERT model that has been trained on a Twitter tweet dataset about Covid-19 and have proven to yield decent results in wide range of use cases (Refer Figure 6).

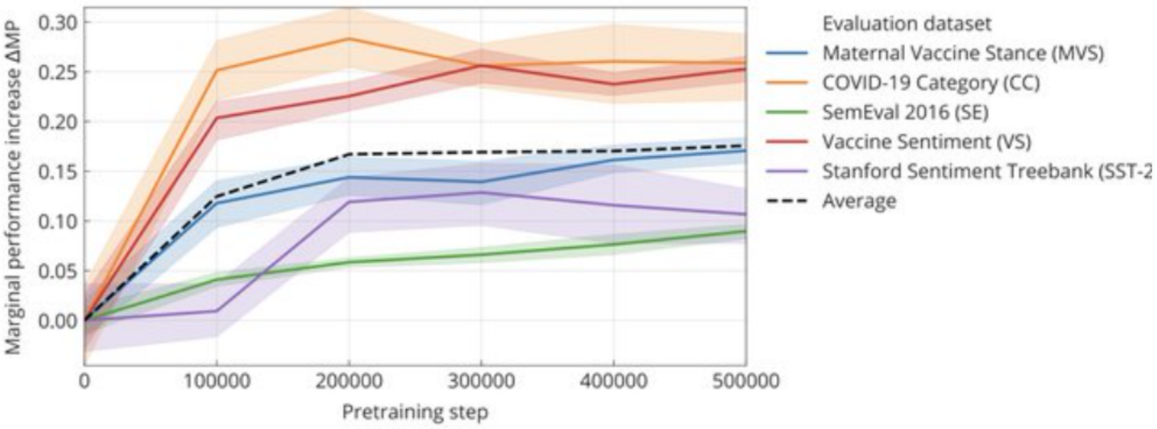


Figure 6: COVID-Twitter-Bert's performance on five different classification datasets

Firstly, we wrote code to fine-tune the model to cater to our project's purpose, which is to detect fake news that is propagated on Twitter. We used Keras framework and HuggingFace library to ease the process of importing and training the model. Furthermore, we also apply NLP techniques such as tokenization, and removing stop words and extra features in training dataset to enhance the veracity of our model. Next, we implemented a function that predicts the label of our news based on the input news and presents its performance on Fake news dataset as compared to Bert model in our Evaluation section.

4.3 Application of Knowledge

Throughout the course of this project, we have acquired a great opportunity to research and study a broad spectrum of technologies and innovations that are being used in Blockchain technology and cryptocurrencies. Not only that, we were introduced to a wide range of information regarding AI and the existing issues in NLP dataset. Thanks to Technical Communication 2, we have developed our critical thinking and problem-solving abilities to be able to analyze and understand technical components that contribute to these technologies. We were also required to implement our solution from scratch in such a way that we were able to apply all the insights from our classroom, including Database and Information System, Wireless Communications and Object Oriented Programming. These technologies are rapidly developing and there are huge opportunities for us to be part of its development and increase our chances of obtaining a leading position in our future career.

5. Result

5.1 Program Description and Testing

5.1.1 Overview

The website is designed to provide users with an interactive interface to perform labeling and safely manage their transactions with the Ethereum blockchain. Furthermore, we take into account the usability aspect of the website by integrating long polling and events to facilitate the speed of propagating data in the website. Lastly, a record of all transactions will be saved in the database and displayed on the website to ensure the integrity of these transactions and allow users to pinpoint failed transactions, thereby performing corresponding actions. All the labels from the users will be saved in the database, subsequently will be used to train the COVID-Twitter-BERT v2 model. The below section will show all the key functionalities of the website and the methods were used to test the website, mainly Postman, Remix

5.1.2 Register new user

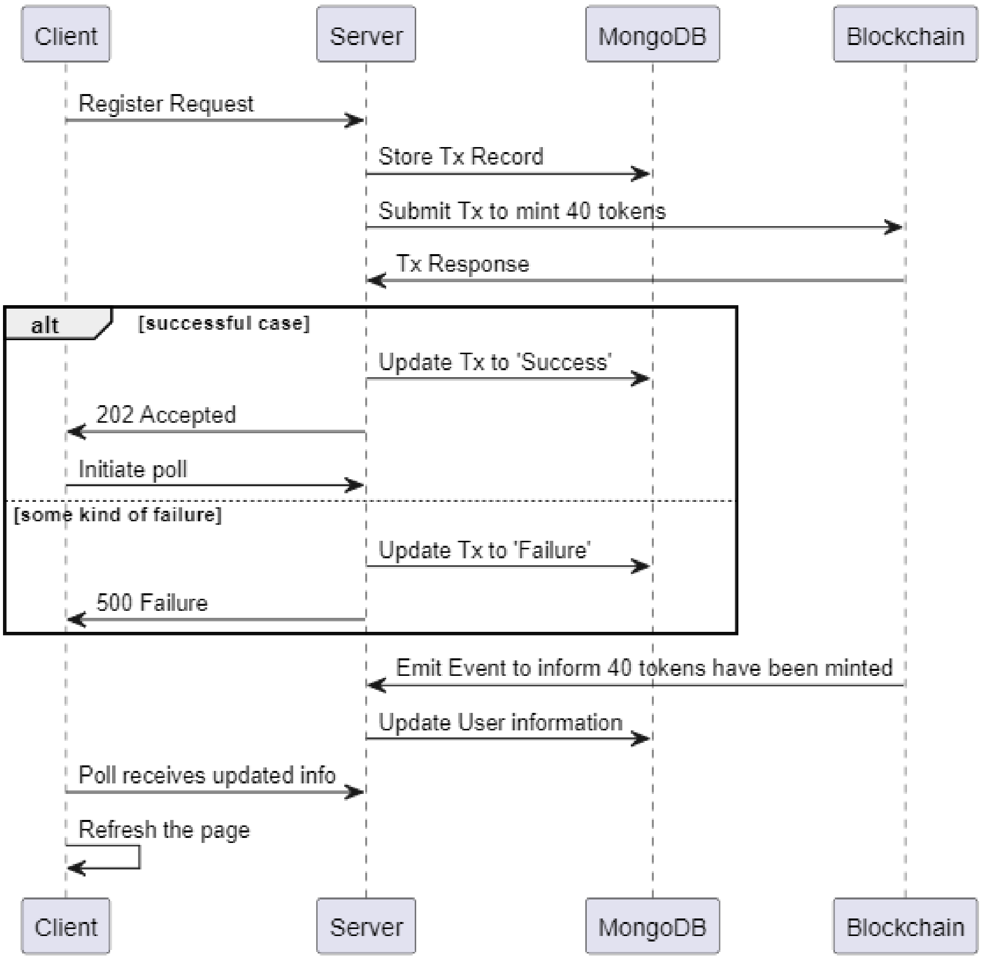


Figure 7: Request execution flow for register new user

Functionality	Testing
1. User makes a request to the Server to create account new account	Postman sends a request to the API to register a new user and expect a new record in Database.
2. Server performs basic validation and creates a transaction record in MongoDB. The status of the transaction is set as Processing	Postman send request to API endpoint to check if status is set to Processing
3. Server submits the transaction to smart contract on Blockchain	Try/Catch to catch potential errors in Blockchain
4. Smart contract mints tokens to user’s Metamask wallet	Use Remix IDE to mint tokens to predefined user’s wallet

5. If the response from Blockchain is successful, server updates the transaction record in MongoDB created in step 3 and return HTTP 200 to user	Try/Catch to catch potential errors in MongoDB.
6. Client will start polling to Server to check for completion status of the transaction	Try/Catch to catch potential backend error
7. If the response from Blockchain is failed, server updates the transaction record in MongDB and return HTTP 500 to user	Try/Catch to catch potential errors in MongoDB.
8. Client will notify the user about the failed transaction and update the interface.	No test
9. Upon Event is emitted from smart contract, Server listens and updates user's information in MongDB and status of transaction.	Use Remix IDE to call register() and expect an Event with corresponding parameters.
10. Client will notify the user about the failed transaction and update the interface.	No test
11. Client successfully retrieves the completion and alerts to user	No test
12. Client sends request to Server for user's info and updates the interface.	Postman send request to API to retrieve user's information

5.1.3 Create new vote

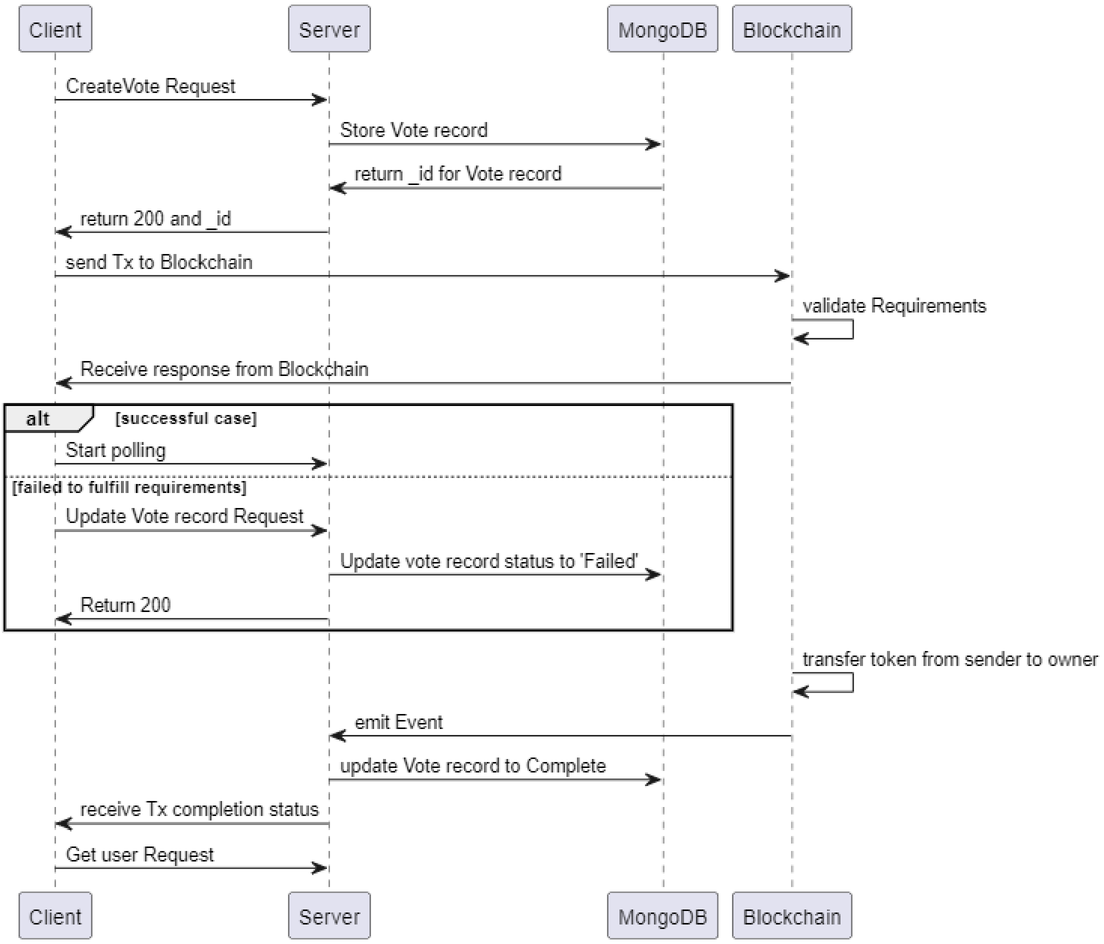

Client
Server
MongoDB
Blockchain
CreateVote Request
Store Vote record
return _id for Vote record
return 200 and _id
send Tx to Blockchain
validate Requirements
Receive response from Blockchain
alt
[successful case]
Start polling
[failed to fulfill requirements]
Update Vote record Request
Update vote record status to 'Failed'
Return 200
transfer token from sender to owner
emit Event
update Vote record to Complete
receive Tx completion status
Get user Request
Client
Server
MongoDB
Blockchain

Figure 8: Request execution flow for creating new vote

Functionality	**Testing**
1. User makes a request to the Server to create new vote on a tweet	Postman sends a request to the API to create a new vote and expect a new record in MongoDB.
2. Server performs basic validation and creates a transaction record of the vote in MongoDB. The status of the transaction is set as Processing	Postman sends request to API endpoint to check if status is set to Processing
3. MongoDB returns _id of the transaction record to the Server	Try/Catch to catch potential errors in MongoDB
4. Server return a HTTP 200 response to Client with _id of transaction record	No test
5. Client embeds _id into the transaction and sends to	Try/Catch to catch potential errors

smart contract on Blockchain	in Blockchain.
6. Requirements are in place to check if tokens that the user staked is between 1-5.	Use Remix IDE to call create vote method and expect a failed transaction with insufficient tokens error.
7. If response from Blockchain is successful, Client start polling to check for transaction record completion status	No test
8. If response from Blockchain is failed, Client send request to Server to update transaction record status to Failed	Try/Catch to catch potential errors in MongoDB.
9. Smart contract transfer ERC20 tokens from user to owner of the contract	Use Remix IDE to call create vote method and expect an Transfer Event
10. Upon Event is emitted from smart contract, Server listens and updates user's information in MongDB and status of transaction record.	Use Remix IDE to call create vote method and expect an Event with _id of Vote record
11. Client successfully retrieves the completion and alerts to user on interface	No test
12. Client send request to get new user information	Postman send request to API to retrieve user's information
13. Client updates its interface to display to user	No test

5.1.4 Calculate gained tokens for user and distribute to Metamask wallet

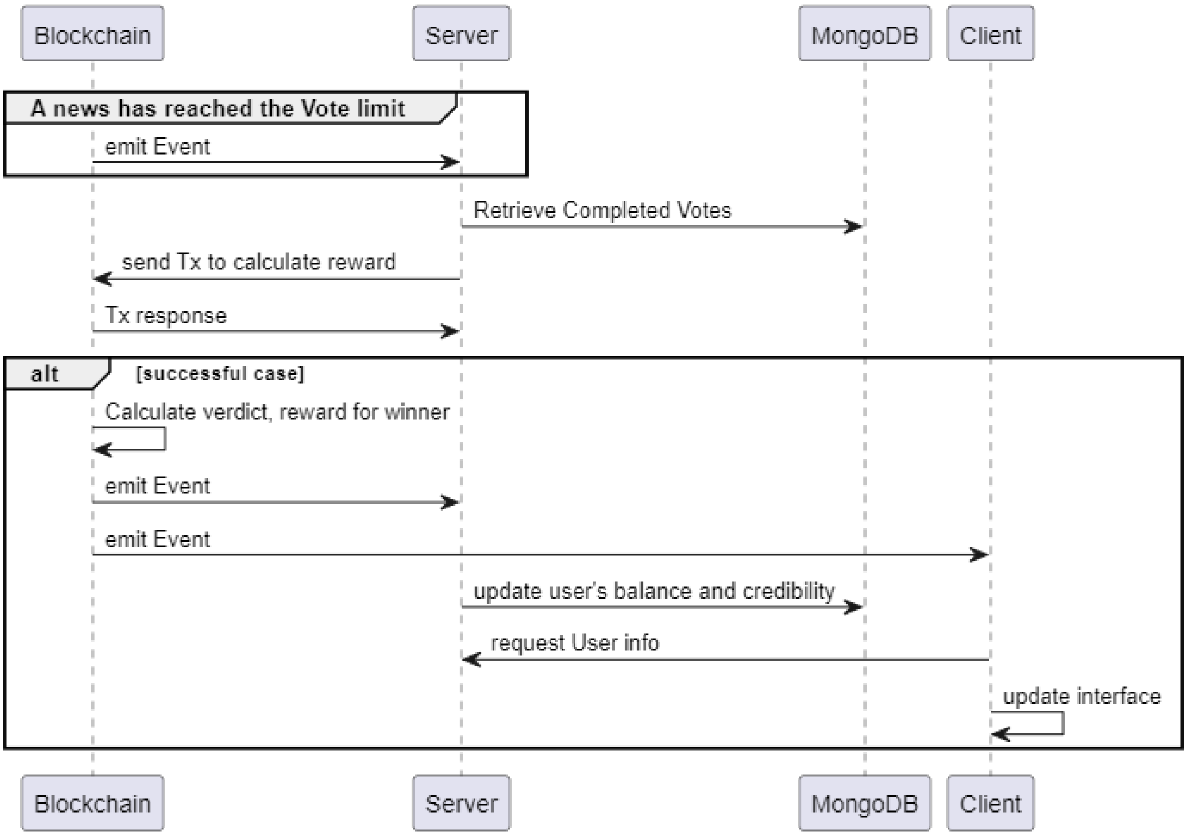


Figure 9: Request execution flow after a news has reached the maximum of votes

Functionality	Testing
1. Upon reaching the predefined maximum amount of vote for a news, smart contract emits an Event to notify Server about the expiry of a news.	Use Remix IDE to call create vote method and expect Event when number of vote reaches maximum limit
2. Server listens to Event and retrieve all valid votes from MongoDB	Postman sends request to API to get votes with status Completed
3. Server send a transaction to Smart contract to calculate reward for users who voted correctly	Try/Catch to catch potential errors in Blockchain
4. If response from Blockchain is successful, Smart contract will formulate the final verdict of the tweet. Next, Smart contract uses the verdict to calculate the amount of tokens that each user will receive and distribute to the user's wallet accordingly. Upon completion, it will emit an Event.	Use Remix IDE to calculate reward method and expect corresponding tokens are transferred to the user who created the vote.

5. Server listens to Event and updates the user record in MongoDB	Use Remix IDE to calculate reward method and expect a Event
6. Client listens to Event, alert to user and request for new user information from API	No test

5.2 Evaluation

One major problem that we did not take into account is the high degree of skewness in our dataset. The dataset is heavily biased towards real tweets as opposed to fake tweets, as can be seen in the graph below (Figure).

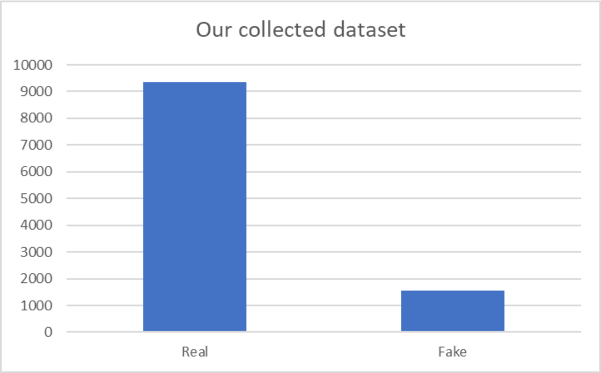


Figure 10: Highly skewed to Real label for Fake news dataset

When train and test data are taken from a similar distribution of news subjects, using unbalanced data leads to very accurate classification (nearly 86%) (Refer Figure). Reporting such high accuracy, however, is deceptive because what we are really looking for is a fake news detection system that can generalize to new topics, or more specifically, a classifier that detects high-level features that can be considered as signs of deception, irrespective of a news article's specific topic. Small data sets would not provide cross-topic generalization since the models learn the vocabulary differences between fake and real news in this scenario, and the vocabulary is highly topic-dependent.

For the actual system, we do not include this dataset in our training set as it will affect the performance of our model. Instead, we have found a fake news dataset from HuggingFace called GonzaloA that provides a decent training size of 24,353 instances to train. Unlike our previous dataset, this dataset is balanced for both Real and Fake news (Refer Figure). After training our model for 3 epochs, we observe a steady decline in loss with a low accuracy, which tells us that the model is learning correctly and there is no sign of overfit due to high degree of skewness.

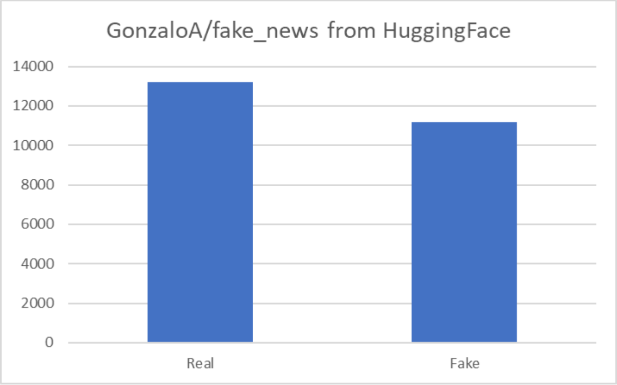


Figure 11: GonzaloA fake news dataset

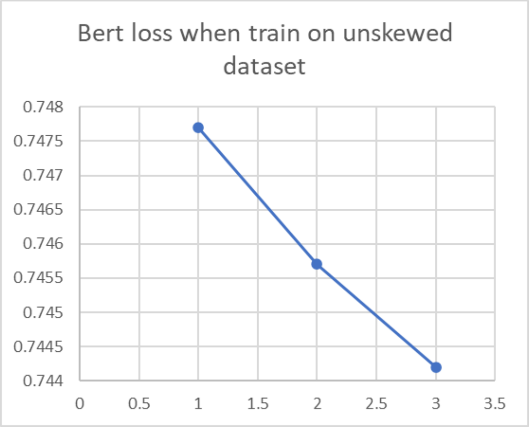


Figure 12: Loss of Covid-twitter-bert-v2 model after trained for 3 epochs

5.3 Outlook

So far, we learned that fake news can produce significant damage to the mind of readers who are exposed to it. There is an urgent need for a way to classify fake news and eliminate the rapid growth of fake news on the Internet. However, we have discussed the process of validating fake news as being difficult, and the amount of individuals that are able to validate these contents is too limited as compared to the size of fake news. Furthermore, it can cost an abundance of time and money to filter out these news articles and is usually only affordable for big companies such as Facebook. Therefore, it is a valid point to say that only AI can help us to fight against the issue of fake news since it facilitates the speed of filtering and can be available 24/7. However, we also know that the main culprit that hinders the advancement of fake news detection is due to the lack of labeled datasets in NLP. In the Evaluation section, we have looked at the performance of a deep learning model that can be heavily influenced by the degree of skewness. The main purpose of this report is to issue a call to action for academics in this area to share datasets and work toward a consensus on how to annotate fake news dataset and label the data. We have also taken initiative to come up with a system that integrates Blockchain technology to provide merits to users for labeling the dataset. Our system effectively deals with the spread of fake news by allowing speedy growth of labeled fake news. In the future, we hope to see users use our website to classify fake news and earn valuable tokens. Constructing a large corpus of labeled fake news dataset allows for a better detection capability and further improves the accuracy and performance of the detection algorithm.

References:

[1] Hirschmann, R. (2022, July 6). *Singapore: Leading sources for news 2022*. Statista. Retrieved July 18, 2022, from https://www.statista.com/statistics/982760/singapore-top-news-sources/

[2] *Digitalepidemiologylab/covid-twitter-bert-V2 · hugging face*. digitalepidemiologylab/covid-twitter-bert-v2 · Hugging Face. (n.d.). Retrieved July 27, 2022, from https://huggingface.co/digitalepidemiologylab/covid-twitter-bert-v2

[3] Guo, Y., & Liang, C. (2016, December 9). *Blockchain application and outlook in the banking industry - financial innovation*. SpringerOpen. Retrieved July 22, 2022, from https://jfin-swufe.springeropen.com/articles/10.1186/s40854-016-0034-9

[4] *Fake news detection on social media: A Data Mining Perspective*. (n.d.). Retrieved July 18, 2022, from https://www.kdd.org/exploration_files/19-1-Article2.pdf

[5] *Cryptocurrency vs. Fiat currency: Architecture, algorithm, cashflow ...* (n.d.). Retrieved July 22, 2022, from https://www.researchgate.net/publication/329565339_Cryptocurrency_vs_Fiat_Currency_Architecture_Algorithm_Cashflow_Ledger_Technology_on_Emerging_Economy

[6] *Understanding modern banking ledgers through blockchain ... - arxiv*. (n.d.). Retrieved July 22, 2022, from https://arxiv.org/pdf/1511.05740.pdf

[7] Guo, Y., & Liang, C. (2016). *Blockchain application and outlook in the banking industry - springeropen*. Retrieved July 22, 2022, from https://jfin-swufe.springeropen.com/track/pdf/10.1186/s40854-016-0034-9.pdf

[8] Kfoury, B. (2021). *[PDF] the role of blockchain in reducing the cost of financial transactions in the retail industry: Semantic scholar*. [PDF] The Role of Blockchain in Reducing the Cost of Financial Transactions in the Retail Industry | Semantic Scholar. Retrieved July 22, 2022, from https://www.semanticscholar.org/paper/The-Role-of-Blockchain-in-Reducing-the-Cost-of-in-Kfoury/5715533b075e11275ac2fc252ab3f07f90215457

[9] Tilooby, A. (2018). *The impact of blockchain technology on Financial Transactions*. ScholarWorks @ Georgia State University. Retrieved July 22, 2022, from https://scholarworks.gsu.edu/bus_admin_diss/103/

[10] Poon, J. (2020, November 5). *PayPal Singapore: Overview of all merchant fees*. WorldFirst Singapore Blog. Retrieved July 22, 2022, from https://www.worldfirst.com/sg/blog/selling-online/paypal-merchant-account-fees-to-take-note-of/

[11] Ethereum average transaction fee. (n.d.). Retrieved July 22, 2022, from https://ycharts.com/indicators/ethereum_average_transaction_fee#:~:text=Ethereum%20Average%20Transaction%20Fee%20is,80.49%25%20from%20one%20year%20ago.

[12] Felin, T., & Lakhani, K. R. (2018). *What problems will you solve with blockchain? - researchgate*. Retrieved July 22, 2022, from https://www.researchgate.net/publication/328598250_What_problems_will_you_solve_with_blockchain

[13] Xu, J. J. (2016). *Are blockchains immune to all malicious attacks? - researchgate*. Retrieved July 22, 2022, from https://www.researchgate.net/publication/311568850_Are_blockchains_immune_to_all_malicious_attacks

[14] Rashkin, H., Choi, E., Jang, J. Y., Volkova, S., & Choi, Y. (2017). *Truth of varying shades: Analyzing language in fake news and political ...* Retrieved July 22, 2022, from https://aclanthology.org/D17-1317.pdf

[15] Asr, F. T., & Taboada, M. (2019). *Big Data and quality data for fake news and misinformation detection*. Retrieved July 22, 2022, from https://journals.sagepub.com/doi/10.1177/2053951719843310

[16] Allcott, H., & Gentzkow, M. (2017). *Social media and fake news in the 2016 election - stanford university*. Retrieved July 22, 2022, from https://web.stanford.edu/~gentzkow/research/fakenews.pdf

[17] Pérez-Rosas, V., Kleinberg, B., Lefevre, A., & Mihalcea, R. (2017, August 23). *Automatic detection of fake news*. arXiv.org. Retrieved July 22, 2022, from https://arxiv.org/abs/1708.07104

[18] Banko, M., & Brill, E. (2001). *Scaling to very very large corpora for natural language disambiguation*. ACL Anthology. Retrieved July 24, 2022, from https://aclanthology.org/P01-1005/

[19] Kfoury, B. (2021). *[PDF] the role of blockchain in reducing the cost of financial transactions in the retail industry: Semantic scholar*. [PDF] The Role of Blockchain in Reducing the Cost of Financial Transactions in the Retail Industry | Semantic Scholar. Retrieved July 22, 2022, from https://www.semanticscholar.org/paper/The-Role-of-Blockchain-in-Reducing-the-Cost-of-in-Kfoury/5715533b075e11275ac2fc252ab3f07f90215457

[20] Snow, R., O'Connor, B., Jurafsky, D., & Ng, A. Y. (2008). *Cheap and fast – but is it good? evaluating non-expert annotations for natural language tasks*. ACL Anthology. Retrieved July 24, 2022, from https://aclanthology.org/D08-1027/

[21] Zhang, J., Wu, X., & Sheng, V. S. (2016). *(PDF) learning from crowdsourced labeled data: A survey*. Retrieved July 24, 2022, from https://www.researchgate.net/publication/304747211_Learning_from_crowdsourced_labeled_data_a_survey

[22] Zhou, X., & Zafarani, R. (2020, July 17). *A survey of fake news: Fundamental theories, detection methods, and opportunities*. arXiv.org. Retrieved July 24, 2022, from https://arxiv.org/abs/1812.00315

[23] Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017, September 3). *Fake news detection on social media: A Data Mining Perspective*. arXiv.org. Retrieved July 24, 2022, from https://arxiv.org/abs/1708.01967

[24] Silverman, C. (2016, November 16). *This analysis shows how viral fake election news stories outperformed Real News on facebook*. BuzzFeed News. Retrieved July 24, 2022, from https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook

[25] Vosoughi, S., Roy, D., & Aral, S. (2017). *MIT initiative on the digital economy research brief: The spread of true and false news online*. Retrieved July 24, 2022, from https://ide.mit.edu/wp-content/uploads/2018/12/2017-IDE-Research-Brief-False-News.pdf

[26] Nickerson, R. S. (1998). *Confirmation bias: A ubiquitous phenomenon in many guises*. Retrieved July 24, 2022, from https://journals.sagepub.com/doi/10.1037/1089-2680.2.2.175

[27] Ecker, U. K. H., Hogan, J. L., & Lewandowsky, S. (2017, March 18). *Reminders and repetition of misinformation: Helping or hindering its retraction?* Journal of Applied Research in Memory and Cognition. Retrieved July 24, 2022, from https://www.sciencedirect.com/science/article/abs/pii/S2211368116301838

[28] Michigan, Z. Z. U. of, Zhao, Z., Michigan, U. of, Michigan, P. R. U. of, Resnick, P., Michigan, Q. M. U. of, Mei, Q., Council, N. R., Rome, S. U. of, & Metrics, O. M. V. A. (2015, May 1). *Enquiring minds: Proceedings of the 24th International Conference on World Wide Web*. ACM Other conferences. Retrieved July 24, 2022, from https://dl.acm.org/doi/10.1145/2736277.2741637

[29] Tambuscio, M., Ruffo, G., Flammini, A., & Menczer, F. (2015). *Fact-checking Effect on Viral Hoaxes: A Model of Misinformation Spread in Social Networks*. Retrieved July 24, 2022, from https://www.researchgate.net/publication/283009320_Fact-checking_Effect_on_Viral_Hoaxes_A_Model_of_Misinformation_Spread_in_Social_Networks

[30] Shao, C., Hui, P.-M., Wang, L., Jiang, X., Flammini, A., Menczer, F., & Ciampaglia, G. L. (2018). *Anatomy of an online misinformation network*. PLOS ONE. Retrieved July 24, 2022, from https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0196087

[31] Pavleska, T., Školkay, A., Zankova, B., Ribeiro, N., & Bechmann, A. (2018). *Performance analysis of fact-checking organizations and initiatives in Europe: a critical overview of online platforms fighting fake news*. Retrieved July 26, 2022, from http://skamba.info/wp-content/uploads/2018/02/Performance-assessment-of-fact-checking-organizations_A-critical-overiview.pdf

[32] Chen, A. (2017, January 29). *The human toll of protecting the internet from the worst of humanity*. Dewayne. Retrieved July 26, 2022, from https://dewaynenet.wordpress.com/2017/01/29/the-human-toll-of-protecting-the-internet-from-the-worst-of-humanity/

[33] Zhang, J., Wu, X., & Sheng, V. (2016). *(PDF) learning from crowdsourced labeled data: A survey*. Retrieved July 26, 2022, from https://www.researchgate.net/publication/304747211_Learning_from_crowdsourced_labeled_data_a_survey

[34] Osa, T., Pajarinen, J., Neumann, G., Bagnell, J. A., Abbeel, P., & Peters, J. (2018, November 16). *An algorithmic perspective on imitation learning*. arXiv.org. Retrieved July 27, 2022, from https://arxiv.org/abs/1811.06711

[35] Zhou, X., Zafarani, R. (2018). A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. Retrieved July 26, 2022, from https://arxiv.org/pdf/1812.00315.pdf

[36] Feng, S., Banerjee, R., & Choi, Y. (2012). *Syntactic stylometry for deception detection*. ACL Anthology. Retrieved July 26, 2022, from https://aclanthology.org/P12-2034/

[37] University, X. Z. S., Zhou, X., University, S., University, A. J. S., Jain, A., University, V. V. P. S., Phoha, V. V., University, R. Z. S., Zafarani, R., University of Louisiana at Lafayette and Cythereal, Cert, & Metrics, O. M. V. A. (2020, June 1). *Fake news early detection: A theory-driven model: Digital Threats: Research and Practice: Vol 1, no 2*. Digital Threats: Research and Practice. Retrieved July 26, 2022, from https://dl.acm.org/doi/10.1145/3377478

[38] Zhang, S., Roller, S., Goyal, N., Artetxe, M., Chen, M., Chen, S., Dewan, C., Diab, M., Li, X., Lin, X. V., Mihaylov, T., Ott, M., Shleifer, S., Shuster, K., Simig, D., Koura, P. S., Sridhar, A., Wang, T., & Zettlemoyer, L. (2022, June 21). *OPT: Open pre-trained Transformer language models*. arXiv.org. Retrieved July 26, 2022, from https://arxiv.org/abs/2205.01068

[39] Buffalo, Y. W. S. U. N. Y., Wang, Y., Buffalo, S. U. N. Y., Buffalo, F. M. S. U. N. Y., Ma, F., Zhiwei Jin University of Chinese Academy of Sciences, Jin, Z., University of Chinese Academy of Sciences, Ye Yuan Beijing University of Technology, Yuan, Y., Technology, B. U. of, Buffalo, G. X. S. U. N. Y., Xun, G., Buffalo, K. J. S. U. N. Y., Jha, K., Buffalo, L. S. S. U. N. Y., Su, L., Buffalo, J. G. S. U. N. Y., Gao, J., … Metrics, O. M. V. A. (2018, July 1). *EANN: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge*

Discovery & Data Mining. ACM Other conferences. Retrieved July 26, 2022, from https://dl.acm.org/doi/10.1145/3219819.3219903

[40] Park, S. (2021). *[technology toolkit] 2. I will give you data, label it~ auto labeling!* [Technology Toolkit] 2. I Will Give You Data, Label It~ Auto Labeling! Retrieved July 27, 2022, from https://www.samsungsds.com/en/insights/techtoolkit_2021_auto_labeling.html

[41] Aravind, A., Bennett, L., & Ojha, V. (2020). *Getting started with blockchain design patterns*. IBM developer. Retrieved July 27, 2022, from https://developer.ibm.com/articles/getting-started-with-blockchain-design-patterns/

[42] Ault, M. (2018). *Why new off-chain storage is required for blockchains - IBM*. Retrieved July 27, 2022, from https://www.ibm.com/downloads/cas/rxovxapm

[43] GeeksforGeeks. (2022, June 24). *What are events in solidity?* GeeksforGeeks. Retrieved July 27, 2022, from https://www.geeksforgeeks.org/what-are-events-in-solidity/#:~:text=Solidity%20Events%20are%20the%20same,the%20transaction%20logs%20when%20emitted.

[44] Wikimedia Foundation. (2022, July 18). *Fixed-point arithmetic*. Wikipedia. Retrieved July 30, 2022, from https://en.wikipedia.org/wiki/Fixed-point_arithmetic

Appendix