

Detection and prevention of Fake news

Thesis report



June 4, 2020

Student Name: xxxx

Student ID: xxxxxxx

Table of Contents

**Short Abstract2**

**Introduction3**

**Background4**

**Related Works7**

**Motivation8**

**Problem Statement & Analysis9**

**Methodology10**

**Result, Evaluation & Discussion12**

**Conclusion13**

**References15**

**Appendices15**

**Short Abstract**

Due to a significant contribution of technology, lifestyle of current generation’s people is highly dependent on digital platforms like social media (e.g. Facebook, Instagram, Twitter, Blogs etc.). As people are using them in day-to-day life, these digital platforms are gathering information so randomly that it is certainly becoming a big challenge to keep track of them. And due to this random growth, the rate of misuse is increasing as well. Like every other digital innovation, these platforms can be misused to generate and deliver unauthentic or tampered information which is really harmful to any individual or society as a whole. A dangerous example of that, is spreading fake news and propaganda with eye catching captions or titles. It is now very frequent in practice due to open access of data. So, to identify and prevent misleading information and fake news in the digital platform has become a global issue. Digital platforms like social media are complex networks of users where any kind of news can be generated and propagated very easily in an unstructured and scattered nature. That’s why researchers are working very hard to find ways to automatically stop spreading misleading information from digital platforms. They have proposed several automated techniques in order to classify and detect fake news, prevent them from spreading, identify and trace the sources, cross-matching with original contents etc. Advanced technologies like machine learning and block chain are proposed and shown most promising results in recent research. In this thesis, a broad synopsis of a technique using combination of these two promising technologies to identify and prevent fake news spreading is proposed with critical comparisons to existing ones. The paper also discussed the motivation behind the work along with an effective methodology, evaluation and analysis criteria to justify that the proposed technique can be helpful.

**Introduction**

The world is moving towards digitalization with more advanced technologies and digital infrastructures. It is now a matter of a click to connect from one part to other and transmit information. This digital transmission of information can be so fast that it happens almost in a blink of eye. That’s why people are relying more on these technologies by sharing, storing and retrieving information through them. This digital revolution has changed people’s lives in every aspect. People are replacing their physical machines with virtual ones. Business models are changing and becoming more robust through digitalization. We can think of examples of digital transformation, how e-commerce and online retailer websites like Amazon, Alibaba etc. are dominating the current retailer business without focusing on to create physical business place. This has become possible because of the smooth and reliable data transmission.

Due to this fast and reliable digital transmission, access of data has become more frequent and data is now a powerful tool. The digital platforms are generating and storing huge amount of data in every minute. Availability and good use of data can ease and enhance our quality of life with better insights and analytics. It is now possible to understand and predict future outcomes by analyzing data. Due to the contribution of digital platforms, any kind of data is available to have a better look at it. But it has drawbacks. Because of ultra-fast and frequent data generation and transfer in digital platforms, it is now very getting very difficult to keep track of every bit of data, meaning, keeping track of where it is generating from and transferring to, also filtering out unwanted information. It is certainly an almost impossible task for humans. Due to this problem, data is losing authenticity or getting tampered more often. And misleading information or data can be harmful at a large scale to the society.

As the current generation is more relying on the internet and especially on digital platforms like social media, data is generating in a rapid and unstructured manner [(Kai Shu et al., 2017)](file:///C:\Users\Muiz%20Ahmed\Desktop\Xtra\Muiz\Block%20chain\Project%20Proposal.docx#r6). That’s why now data integrity became a big question. News is a kind of data which also transmitted through the digital networks very frequently. Spreading fake news is one of the aspects of current society which is highly affected by unauthentic and fake data where internet and social media is playing a vital role. And it is very easy to grab people’s attention through the social media as they are always connected to each other.

Spreading of fake news essentially indicates manipulation of original content, misleading or tampered information, completely false or unsubstantial news which can highly impact society or any individual’s reputation [(Chengcheng Shao et al., 2017)](file:///C:\Users\Muiz%20Ahmed\Desktop\Xtra\Muiz\Block%20chain\Project%20Proposal.docx#r8). Social media is a virtual gathering of people where they can share ideas, information with each other (Leticia Bode, 2016). Spreading news is very easy in social media because anyone can share any kind of information of any forms such as text, audio, video etc. And it’s also easy to attract people with exciting and eye-catching titles or captions. There are also different ways to promote fake news as well like click bates, misguidance etc. Interesting videos and images can attract attention. News can lose its authenticity in social media because it is now generating from uncountable sources. Literally, anyone can post a news and reach thousands of reader’s eye. And it is published without any verification most of the times.

Fake news can be identified as a social diseases. It is a really difficult job to distinguish between fake and real news. Even for humans. And that’s why it managed to attract global attention [(Antino Kim et al., 2019)](file:///C:\Users\Muiz%20Ahmed\Desktop\Xtra\Muiz\Block%20chain\Project%20Proposal.docx#r7) and mostly after US president election in 2016 (Hunt Allcott and Matthew Gentzkow, 2017). A study showed (Alexandre Bovet and Hernán A. Makse, 2019) that 25% of tweets in Twitter spread fake or extremely biased news during that time by Donald Trump supporters to reach more people and increase publicity. People are now convinced that preventing fake news requires digital solution like automated algorithms or robots etc. Because the growth of generating fake news is immensely increasing. And previously automated tools have proved that it is possible to share our responsibilities with robots. With contribution of advanced technologies and artificial intelligence, people can think of replacing tasks by machines or robots.

In recent years, many research methods using advanced technologies of automatically detecting fake news were established. Although mathematical models (Kristina Lerman, 2007) and traditional algorithms showed good results but machine learning and block chain are the most exciting technologies that showed great promise in recent times. It is possible to classify, cluster fake news and detect anomalies in order to identify and prevent fake news of any kind with various machine learning techniques. Block chain has solved critical problems recently like preventing data alteration, tampering in different domain of research. And it also showed promise to solve spreading fake news problem from digital platforms. Researchers also proposed hybrid solutions by combining these two cutting edge technologies to solve the problem. Because of wide range of features and applications of these two technologies, a hybrid framework can be a better option. And this paper mainly focuses on proposing a framework that can explore more opportunities and possibilities by critically analyzing and evaluating the current contributions in the research field.

In this thesis, a hybrid solution using machine learning and block chain is discussed with detailed information and implementation plan. The hybrid solution or frame work showed possible enhancement and improvement of the current situation with comparisons between existing techniques. The paper first presented backgrounds of these two technologies such as machine learning and block chain, their types and scope for implementation, also their possibility to solve critical problems by combining the key features. Then a few recent literatures on these topics are reviewed based on their proposals, ideas and results. The paper tries to incorporate their ideas into the proposed technique by critically analyzing and proposing an improved solution with possible methodology. Proposed methodology is also evaluated in order to justify that the model can be a feasible solution and implemented in real life context. Finally the research outcomes are identified to conclude and summarize the findings from the whole research idea.

**Background**

In this section, two major technologies that showed promises with their features and application scopes in order to solve the problem is discussed.

* **Machine Learning:** Machine learning itself is a broad domain of research with many techniques and methods. What a machine learning model does, is essentially learn from historical data, improve its predicting accuracy and then it can automatically predict on new data based on its knowledge. It requires historical data, also called as training data where the model learns from. Most of the machine learning algorithms try to implement mathematical theory in the core of their theory. Statistical methods, logistical analysis, numerical analysis are the common background of developing a machine learning model. And because of the strong mathematical background, these model can be perform better to solve different problem scenarios. Machine learning models can be tweaked and improved by basic concepts of math. Different machine learning models have shown good performance in classifying news articles, understanding their patterns and behavioral differences, tracking sources, context mining in form of linguistic cues, temporal analysis (Natali Ruchansky et al., 2017) etc. In machine learning approach, one of the few drawbacks, is the performance of the model is highly dependent on the availability and quality of dataset (Ray Oshikawa et al., 2018). And they usually require a large amount of well-organized and clean data to perform well. There are terms in order to evaluate a model such as overfitting when model performs very well in the training environment but performs poorly when it faces new or unseen data (Douglas M. Hawkin, 2004), and under fitting when model is not trained with sufficient data.

Supervised learning is one of the popular techniques to classify fake news.In supervised learning the model learns with knowing the targeted output. Generally, training data of a supervised learning model consists of features and corresponding labels (Rich Caruana and Alexandru Niculescu-Mizil, 2006). In this case, the features can be source of news data, body of the news article, the author etc. and as the labels are fake and real. From the occurrence in the training dataset, the model will learn and try to predict on new data. The model will try to determine, for what combination of features, the probability of a news to be fake or real is higher. Then it will set a criteria to classify them. There are two different techniques in supervised learning, namely, classification and regression.

Classification algorithms directly classify and predict a label whereas regression outputs a probabilistic score with a threshold to finally classify the labels. A few effective classifying techniques with supervised learning are Naïve Bayes, Logistic Regression and Support Vector Machine (SVM) etc. Naïve Bayes classifier tries to determine the label by the probabilities based on applying the Bayes’ theorem with strong independence assumptions between features. It essentially tries to find out the probabilistic relations among the features. For example, it is more likely that a news source and author would be same. Naïve Bayes classifier identifies this relationship and learns that this combination occurred to be fake or real in the training data (Irina Rish, 2001). Logistic regression is also a statistical model that uses basic form of logistic function to determine dependent variables (Raymond Wright, 1995). It is very useful in binary classification problems like fake news. SVMs are combination of classification and regression techniques (Xuchan Jua and Yingjie Tian, 2012). The problem of supervised learning is, the training data must be labeled with the targeted output. Dataset must be constructed with original and fake news and they should be labeled properly. The model is going to perform based on the labels of training data.

Other techniques such as anomaly detection, clustering, generative approaches are also proposed by many literature. They are mostly unsupervised, meaning the classified labels are not defined when the model is trained. Algorithms try predict the similarity or difference in their patterns and behavior, then separate into different categories. Clustering can make collections of similar data and that’s how unnatural or irregular data can be found. K-means clustering, support vector clustering are popular unsupervised techniques that can identify common, or unknown, unfamiliar sources of fake news from where it can be generated.

Natural language processing has many techniques. Recently Neural Network has showed brilliant results by trying to mimic how human brain works. In news data, it is really important to understand the temporal meaning or connection of each segment of the data because data is usually large and sequential. That’s why natural language is one of the most exciting techniques in machine learning domain. Recurrent Neural Networks, Long-Short-Term-Memory (LSTMs) are very popular algorithms with sequential data like text, audio and video (Ronan Collobert et. al., 2011). Convolutional Neural Networks can be used to identify social structure and propagation of the news which is important factors of spreading fake news (Federico Monti et al., 2019). But the drawbacks of these architectures, they are usually very complex models with a lot of learning parameters. And of course, a large dataset is needed. Which makes the model very expensive in terms of processing or learning. That’s why to perform these operations, high configured machines are required.

* **Block chain:**  Block chain is still an emerging technology and growing its dominance over various domains. Block chain technology is a distributed data storage, which uses features like distributed ledger, decentralized peer-to-peer networks, and cryptography. In simple words, block chain is a collection of blocks of information. A block consists of header which contains cryptographic hash of its previous block’s header, timestamp, cryptographic hash of its body content etc. and body containing the data. Since the block header contains the hash value of the previous block header, the information in each block changes, and the subsequent block changes accordingly. That’s how block chain ensures tamper-proof, forgery and traceability of information in the process of information collection, transfer and sharing (Wenqian Shang et al., 2018).

In decentralized networks all the peers are connected to each other. Data is stored and validated in every nodes (Zonyin Shae and Jeffrey J P Tsai, 2019). That’s why data has transparency and authenticity in a block chain. And as it stores data in blocks and they are cryptographically connected to each other, authors showed traditional data tracing technology to extract and evaluate source of the news. This feature is likely very much applicable to social media because it has uncountable nodes (Mohammad Torky et at., 2019). Creating a peer-to-peer secure platform for storing and exchanging information can combat against digital deception (Paula Fraga-Lamas and Tiago M. Fernández-Caramés, 2019)

There are different cryptographic techniques such as public-key cryptography, cryptographic functions, digital signature (Steve Huckle and Martin White, 2017). In public-key cryptography, data is hashed with a private key by which it can be decrypted. The private key belongs only to the owner of the data. It can protect original content of news from any alteration when propagating through internet. Cryptographic functions essentially means mathematically create a hash string based on the data. It is a unidirectional function (Adnan Qayyum et al., 2019). For creating the hash, a popular algorithm is used called SHA256 which produces 256-bit hash. By using cryptography, digital signatures can be generated. This provides data privacy, ownership and authenticity. Cryptographic mining of block chain requires extensive computational power. That’s why, to encrypt all news on the internet is a very challenging task. Also the rate of illegal mining by hackers (like bit-coin mining) can be increased.

Combining these two techniques to automatically classify and prevent fake news can be an effective and exciting solution to the problem. Because in order to prevent fake news, we first need to identify the fake news and machine learning has the capability to understand the key features of particular class and identify them. So machine learning can be used to automatically classify the fake news without human intelligence or intervention. A decentralized peer-to-peer cryptographic block chain can be used only to store authentic news so that it cannot be altered or changed by anyone. An immutable distributed database can solve the problem of spreading fake news. This proposed framework also can be implemented on large scale applications because it will be easier to scale as machines can be scalable.

But as mentioned, these techniques has some challenges, a few challenges are pointed out following below.

* Finding available datasets to train the machine learning models.
* The performance of the machine learning models are highly dependent of size and quality of dataset to train on.
* Overfitting and under fitting problems may occur after training the machine learning models.
* The datasets usually contain predefined data with specific content and timeframe which is not good for train machine learning models to be trained on because news data is universal and generating every day. The model should be flexible and robust in order to classify news data.
* Identify authentic source.
* Extensive computational power requires to train machine learning algorithms
* Extensive computational power requires to mine cryptographically in order to create block chain.
* A good amount of peers needed to make a distributed peer-to-peer network and to implement block chain.

**Related Work**:

Researchers have shown several techniques to solve the detection and prevention of spreading fake news problem. Machine learning and block chain has shown the most promising results among them. A very few research works are proposed by the researchers using machine learning and block chain individually and together as well in order to solve this problem of spreading fake news, also showed promising results.

Different machine learning models have shown good performance in classifying news articles, understanding their patterns and behavioral differences, tracking sources, context mining in form of linguistic cues, temporal analysis (Natali Ruchansky et al., 2017) etc. Supervised techniques were used in different literatures like Bayes Theorem (Mykhailo Granik and Volodymyr Mesyura, 2017), Logistic Regression, Support Vector Machine (SVM) (Kai Shu et al., 2019). LSTMs, RNNs (Natali Ruchansky et al., 2017) and CNNs (Federico Monti et al., 2019) were also used as well to classify sequential data like text. With proper dataset, supervised learning can classify news very well and the results are promising as it is binary classification problem. But there are some limitations as mentioned in the previous section. In order to perform well and overcome machine learning problems, researchers used several techniques. Benchmark datasets are used to experiment the models such as Twitter, wiebo (Natali Ruchansky et al., 2017), Buzzfeed, LIAR (Shuo Yang, 2019) etc. A data repository is introduced called FakeNewsNet (Kai Shu et al., 2019) where multidimensional data is captured from news and social media context. Also preparation of dataset is important and many techniques are well-practiced such as data partitioning (e.g. train-test-validation, k-fold validation, cross-validation), data pre-processing (e.g. cleaning, tokenization, stemming, generalization, sampling). Authors have used different techniques to reduce overfitting problems such as dropout, regularization etc. And to evaluate their models they have presented performance comparison in different metrics (precision, recall, F1 score, accuracy) and graphical figures.

Authors also proposed hybrid models with multiple machine learning algorithms. Three different modules were shown in order to extract temporal representation, capture the source characteristics and integrate these two to classify the news as fake or original article (Natali Ruchansky et al., 2017). Because of literacy and cognitive limitations of machine-based approach, a solution is proposed to combine human interventions to find the truthfulness of data and reliability of source (E. M. Okoro, 2018).

Authors showed great results to prevent news data from alteration and temperament by distributive ledgers and peer-to-peer block chain systems (Tee Wee Jing and Raja Kumar Murugesan, 2019). Results showed that because of the characteristics of connected nodes storing information, data authenticity can be maintained. The different uses of cryptographic techniques (Steve Huckle and Martin White, 2017) also proved that cryptography can provide special security to data.

Researchers also incorporates Artificial Intelligence (AI) and block chain in order verify and prevent fake news from social media. AI used for fact-checking for a deeper linguistic comparison of differing types of fake news e.g. satire, propaganda etc.

An AI and block chain based trusting news platform is proposed (Zonyin Shae and Jeffrey J P Tsai, 2019) with several components including a block chain based crowd sourcing platform to rank fake news from all the news posted in the platform; AI algorithms as well as ecosystem models are used to provide incentive mechanisms for the general public working as trusting news validators. Also a fake multimedia detection component which uses AI algorithms to detect the tampering of multimedia materials and another component for fake text detection component which uses AI algorithms to detect the fake news text from the social and humanistic perspective. And also mechanisms and system design research component to integrate AI and block chain technologies for supporting the scalability of AI's smart contracts, trusting news validator, etc. are presented in the literature with a goal to build an AI block chain platform which can leverage AI and block chain accountable and traceable crowd sourcing mechanism to rank and detect fake news.

**Motivation**:

The main motivation of this work is to identify and propose an effective framework using advanced techniques to solve a global problem which is becoming more and more impossible to solve by only intervention of humans. This problem needs help from machine intelligence and it has to be an automated solution. As discussed, the recent research works are proving that this is possible to automate detection and prevention of spreading fake news, our motivation for the work is to justify the techniques and improve with special architecture or framework. Also put our research contribution in order to analyze the problem based on current scenario with elaborated insights and details.

In our solution, we will try to incorporate a scalable machine learning application which is less complex and feasible to implement in real life context. Also a distribute peer-to-peer block chain system that ensure data authenticity and security.

**Problem statement and analysis**:

The first step to start solving a problem is to think about the problem from different perspectives. The problem needs to be clarified by making all the core assumptions associated with it. The best case and worst case scenario also needed to be considered. Then we should start designing a roadmap by analyzing the problem. To analyze a problem first we need to understand each and every criteria of the problem. Because analyzing a problem and designing an efficient roadmap are essential tasks in order to solve a problem, no matter how critical the problem is. But before doing the analysis, the problem needs to be identified and understood first. In this section, the core assumptions of the problem and the problem statement are revisited. Also it is analyzed critically to find the problem structure, domains and aspects which are used in the following section to compare the performance of the proposed framework.

The core idea of the problem is detecting fake news and preventing it from spreading on digital platforms. But it is a broad problem scenario. For our convenience, we can simplify the problem structure and narrow the whole problem scenario. We can start by dividing it into smaller problems and set criteria to solve them. Then we can take actions to solve the smaller problems one by one. So, firstly we can divide it into two sub problems, the present problem and the future problem.

The present problem refers to identify or detect the existing news as fake or real. As we have a huge amount of news already on the digital platforms, it should be the first concern to detect the fake news from existing data and how can we automate this process. To have a closer look into the present problem, we also should think of the data we are going to be dealing with. Digital platforms has uncountable sources of news. We need to narrow down the sources to collect news. We can select a specific social media platform to collect data and try solving our present problem. To narrow down the problem even more, we should consider the form of data. If we select Facebook as the selected social media, people can post any forms of news there like audio, video, image, textual etc. For simplicity we can collect text based data from Facebook. We can even select special categories of news such as sports, politics and travel etc. because news from these categories will contain similar source and information which will help us to relate.

The output from the solution of the present problem should provide a binary answer against a text based news article from Facebook. It should classify a news data which generated by any user as fake or authentic news. Then comes the future problem by referring to protect and prevent the news from temperament or alteration if it is authentic. And also mark the fake news in order to avoid them. For doing that, we need to provide a solution of an immutable database, where authentic news will be stored. And this system will be considered as the news source. People will not have the access to make any changes to data. That’s how the news data will remain authentic. The solution of the future problem will also provide a reliable and authentic news source in the social media platform. Where people will know that the data cannot be changed at any cost.

Digital platforms like social media has complex structure and network of users. If the data source is centralized, it will be easier to tamper the data. And the news source will not be as authentic and tamper free as we want. That’s why proper verification and validation process of news data is needed which can be possible by implementing a decentralize platform to store data and use it as a news source. In decentralized networks data can be replicated to multiple storages and if anyone wants to change a bit of data, a verification process will assure that he or she cannot change it. That’s why the ultimate source of authentic news should be implemented in a decentralized and distributed system where multiple sources will assure the authenticity of data.

**Methodology**:

The proposed framework is a combination of machine learning classifier and a distributed block chain. Appendix A shows the overall architecture of the framework. The classifier will be trained with text based examples of fake and real news. After training the model, we can use it to identify new text based news data. When the classifier will find an authentic news, the data will be added to a block chain which will be used as an immutable database of authentic news. People can avoid all other sources and follow this database to collect real news. In this section, a detail methodology of implementing these technologies along with some identified features are presented. We can divide the framework into two modules like a machine learning model and a block chain. They are discussed below.

* A Machine Learning Model:
  + Supervised Learning model: For binary classification like identifying fake news, supervised learning can provide better results with good accuracy. As discussed in the background section, various types of supervised learning can be used. For our proposal, we will be mostly focusing on support vector machine algorithm. Text based dataset can be classified using multi-word with SVMs (Wen Zhanga et. al., 2008). The main idea of an SVM classifier to identify a hyperplane that can divide the dataset into two classes by separating them linearly (Xuchan Jua and Yingjie Tian, 2012). Appendix B shows, the model annotates data points by numbers based on the frequency of the data and plot them in a graph, then try to fit a hyperplane in order to separate the data points. The hyperplane is set by maximizing the distance between hyperplane and nearest data points on either side during training. That’s how it is determined that two classes are well classified from the dataset. SVM was selected for the as the machine learning model to solve this problem by considering the research challenges, mentioned in background section. SVM is one of the better algorithms in terms of accuracy which needs only a small dataset. It is a combination of regression and classification. This supervised technique is less complex and also computationally less expensive than other natural language processing or deep learning techniques like RNNs, LSTMs etc.
  + Dataset collection: Dataset is the first prerequisite in order train a machine learning problem. We need to collect a proper dataset considering important factors like what type of algorithm we are using, what features are important, preparing and cleaning techniques of the dataset in order to prevent the model from learning any unwanted features. For supervised learning dataset must contain the targeted label. As we are considering text based datasets, we should consider the lengths of news articles. Because SVMs perform well in smaller datasets. For collecting the dataset, the problem statements needs be clearly identified which is stated in the previous section.
  + Key features and parameter analysis: We need to find out some of the key features that can be associated with fake news data. As we are interested only in detecting fake news by the machine learning model, we are going to focus on the content or body and title of a text based news article. We will try to train our model in order to classify the news by analyzing the title and content. So, in this case we are going to focus on only two key features and our data will contain only the titles and contents as input variables and their corresponding labels as output variable.
  + Data preprocessing: After collecting the dataset and finding the key features as input and output variables, we need to process the dataset according to our needs. There are several preprocessing techniques for textual data. First, we need to clean the dataset, remove unnecessary information and noises such as symbols, extra spaces etc. Cleaning datasets removes the unwanted information from the dataset. After that we need to tokenize the dataset. By tokenizing technique we can index every words of the dataset by numbers. Which is important because machines don’t understand texts. Tokenizing helps the model to identify texts. Other preprocessing techniques are also important such as stemming, stop word removal, lemmatization, parts of speech tagging, normalization (Eleanor Clark and Kenji Araki, 2011) etc. After the preprocessing is done, we need to split the dataset into train, test and validation sets. Training and validation sets are used in the training phase and test set is for evaluating the model after the training is completed. Test set is usually used to see how the model performs on the unseen data. There are different techniques to split the dataset as well. We can use general percentages to split the data like 80-10-10 where 80% of the dataset will be used as training data.
  + Training: First, we need create a one dimensional vector containing numbers. These numbers will represent frequency for each words. This technique is called word to vector (Word2Vec) which is a multi-word vector for text representation (Wen Zhanga et. al., 2008). By plotting these numbers on a XY plan like in appendix B, will show the frequency of the words for our two classes such as fake and real. During training with SVM, each input will contain a series of numbers representing the contents which will draw a linear line in the plot. After the successful completion of the training phase, a best fitted hyperplane will be identified by the SVM. When a new article content will feed to the model this hyperplane will determine the words that are more likely to be in the contents of a fake or real news article. And finally classify the label. To overcome limitations of machine learning models, a few techniques can be applied like Regularization etc. Regularization is technique that can be used to avoid overfitting during training by balancing the bias and variance of the data. This makes the model more flexible to unseen data. Also to improve accuracy, various hyper-parameters should be tweaked and changed during training like learning rate, number of training iterations etc.
* Block chain:
  + Structure: After getting an authentic news from the classifier, the news data will be stored in a database which will be basically a block chain system. The reason for choosing block chain based system as the data storage is because of its structure. A block chain contains series of blocks connected cryptographically. These blocks contain data and cryptographic hashes. We need to build these blocks by storing the news data. The basic structure of each block will consists of header which contains cryptographic hash of its previous block’s header, timestamp, cryptographic hash of its body content etc. and body containing the news data (M. Niranjanamurthy et. al., 2018). Since the block header contains the hash value of the previous block header, the information in each block changes, and the subsequent block changes accordingly. And it is easier to identify any changes on any of the blocks. It is more difficult to change the data of block that is part of a longer block chain. Suppose if the user wants to change the news data from his machine, he will first have to change the body of that block. Due to this change, the hash of that particular block will get changed. As the following block contains its hash function in its header, it should also be changed at once. And all the other blocks will have to change the hash as well.
  + Cryptographic Hashing: There are different cryptographic techniques such as public-key cryptography, cryptographic functions, digital signature. In public-key cryptography, data is hashed with a private key by which it can be decrypted. The private key belongs only to the owner of the data. It can protect original content of news from any alteration when propagating through internet. Cryptographic functions essentially means mathematically create a hash string based on the data. For creating the hash, a popular algorithm is used called SHA256 which produces 256-bit hash. By using cryptography, digital signatures can be generated. This provides data privacy, ownership and authenticity. Cryptographic mining of block chain requires extensive computational power. That’s why, to encrypt all news on the internet is a very challenging task. Also the rate of illegal mining by hackers (like bit-coin mining) can be increased. For hashing, we can use SHA256 algorithm which is widely used in block chain systems and a reliable one. It is a unidirectional function that means hash functions cannot be decrypted, so there is no possibility to know what the hash string contains. And that makes harder even if anyone manages to duplicate the hash. SHA256 algorithm can be applied to any kind of data such as texts, audio, video, binary, files, combination of these types etc. For our purpose, text data can be easily hashed by the algorithm.
  + Storing data: To store data and add a new block in the block chain, news data needs to be cryptographically mined. This “mining” terminology is used in block chain system to refer a validation process that is needed to ensure the cryptographic hashing was correct. The primary goal of this mining is, to find the best hash string with hashing algorithm that is unique, and not being used by any other blocks in the block chain. SHA256 algorithm generates 256 bytes of hash. That means it is harder to mine because of the length. But after completion of a successful mining, the data can be stored in the block chain which provides security to the data. It is almost impossible to mine and find the same hash by using different data. That’s why mining is important. If anyone wants to alter the data of a block, he will have to mine and find the exact same hash for every following blocks at once.
  + Peer-to-peer distributed network: One of the key features of block chain is to ensure a distributed peer-to-peer network where the whole block chain is stored on each nodes of the networks. That means block chain relies on no centralized storage system, rather it replicates its data on different nodes. Centralized storage systems can easily be hacked and data can be changed. Centralized systems has no records of previous versions of its data after it gets changed. So, this distributed storage feature provides data safety because it is difficult to alter data in every nodes at the same time. Block chains stored in every nodes are always validating and updating the blocks. If new block is added to the chain, it is instantly added to all the nodes. To make a peer-to-peer distributed network, we can store the news data block chain in every local machine from where the users are accessing the news. In this system, even if altering data and mining the whole block chain could be possible by extreme computational power, the distributive storage system will not allow if it sees any abnormal alteration of hashes in particular node. The peer-to-peer network will remain the same and data will be intact. As social media is a complex networks of users and apparently this large amount of users who are connected to each other very frequently can help the system to be more reliable and efficient as decentralized networks proposed in the framework requires more peers in the network. Due to huge amount of users in social media and they are connected to each other this structure will become stronger and the proposed framework has more promising outcomes in the future.

**Result, Evaluation & Discussion**:

Based on the problem statement, an analysis is done in this section with the context of proposed technique. The analysis tries to justify that the proposed framework is feasible and it is aligned with the problem statement. Its effectiveness, possibilities and evaluation criteria are also discussed for the justification. Also an implementation plan for the future is discussed later in the section along with a few research outcomes.

We can consider the proposed framework is providing feasible solution to solve both of the problems mentioned in the problem statement section. The present problem can be well identified and solved by the machine learning approach of the proposed framework. We have proposed to solve the problem by using SVM which can separate the classes by looking at the content and title of news articles. By narrowing down the sources, dataset can be collected from Facebook news sources and fed to the SVM classifier. Unlike logistic regression or Naïve bayes classifier where model learns to identify relationships and dependency among the features, SVM doesn’t require too many features other than the content and title of the news to classify news. Which makes the data collection process as well as the model structure even simpler. The reason we considered the combination of title and contents for the classification, is fake news are more likely to have same kind of contents and titles to catch reader’s attention. And SVM can identify the words which are more likely to occur in a particular class. Block chain can provide a secure network and immutable database which solves the future problem that is identified in the problem statement. The features like cryptographic mining and peer-to-peer network and validation of data makes very difficult to alter the data in a block. And its decentralized data storage feature also shows a promise that proper verification and validation of news data will be assured at any cost. Implementation plan according to methodology and problem statements is presented below.

*Prototype Designing:* The prototype of our framework can be implemented by first, training the classifier machine learning model with sample data which can be collected from different news sources in Facebook. We can collect information of news including their contents and titles from different Facebook pages. And then train the model on the dataset in order to teach the model to classify a news article by analyzing its content and title. After the model is trained, we can collect more random news from the Facebook again and feed to our model and it should output a probability that the news is some percentage fake. For the prototype we can cross check with the actual label repeatedly to set a threshold that justifies the news as fake or real. Let’s assume the threshold is 70%, meaning if the classifier outputs that the news has 70% chance to contain fake words in the content and title, we will mark it as a fake news. Other than that we will move forward to store the authentic news data in a block chain. We need to mine cryptographically to find the perfect hash of the contents and add them in a block chain. We can assume each user’s local machine who are reading the news as each node. We can store the block chain in their machine to form a distributed storage system. We need to verify and validate the block chain every single time in order to consider it as an immutable news data source. As social media like Facebook has millions of users, the block chain will be stored in millions of machine and it will be practically impossible to violate or tamper the data in the block chain in the distributed and decentralized system.

*Evaluation Criteria:* Every research should set some evaluation criteria in order to justify the work’s performance outside the development environment. If the research passes the mentioned criteria, it is likely to be a successful research project. To evaluate our framework we need set up our evaluation criteria. We need to evaluate that our model is trained well with a real life dataset, then only authentic news are added to the block chain. We also need to assure the block chain is formed with the correct structure and adding new blocks with a correct mining process. The evaluation process also needs to confirm that the block chain is stored and can be retrieved from each nodes to ensure the distributed storage system. First in order to evaluate the performance of our classifier, we can follow some evaluation metrics which are widely used to evaluate machine learning algorithms such as accuracy, precision, recall, F1 score, ROC and AUC curve etc. Accuracy is a rate or fraction of predictions that the model got correct. We can calculate accuracy by dividing the number of correct predictions by the total number of predictions. Also calculating the true positives (tp), true negatives (tn), false positives (fp) and false negatives (fn) can provide better metric evaluation. True positives and negatives refer to how many time the model predicted positive and negative class correctly. In our case the positive class is real news and true negative is fake news. False positives and negatives refer to the number of incorrect predictions to classify positive and negative class. Tp, tn, fp, fn are also used to calculate in other metrics like precision, recall, F1 score etc. Receiver Operator Characteristic (ROC) curves are also a commonly used metric to evaluate binary classification problems solved by machine learning. (Jesse Davis and Mark Goadrich, 2006) The area under to ROC curves (AUC) measures that how a parameter can classify between two classes.

The evaluation criteria for block chain would be evaluating its usability, performance, scalability, security, availability, quality of transactions etc. We should evaluate its usability in terms of it is serving our initial purpose or just making the life complicated. As block chain networks can be huge in a social media, it is very important to check on its performance and scalability every once in a while. We know that cryptographic mining, distributed ledger are very expensive features of block chain which can harm the performance. To improve the performance we need to think about the scalability as well.

*Research Outcomes:* The main purpose of the project was to find challenges in our day-to-day lives on digital platforms as we are spending more time, gathering and more information about ourselves. This research proposal answers several questions due that critical issues. The main focus of this research was to identify the problems, analyze and find out the problem statements. And finally find the best solution by evaluating and considering all the facts associated with it.

First of all, this research shows why spreading fake news is a problem by clearly outlining the issues regarding data integrity, transparency and transmission. The research also tried to convince that the solution of this problem should be automated as it is going to be more difficult for humans to do manual screening process and prevent spreading fake news. And also because digital platforms are generating news articles at a random growth in every minute or maybe in every second.

Paper also discusses most effective research areas and related works that are covered by the researchers in their literature. A brief discussion of their techniques and promising results that motived the work is also presented. The paper highlighted key features that can be used for further implementation and how combining these feature could improve the current results. After analyzing and studying the background and related works, the core problem statement was revisited in the light to identify the problem more closely.

We tried to divide the problem several sub tasks so that big problems can be solved by solving smaller problems. It also helps to clearly understand the requirements of the problem and what needs to be solved and improved by the proposed framework. And finally we proposed a framework based on the background study and problem context. We tried to design the roadmap by analyzing and implementing key features of two advanced technologies like Machine Learning and Block chain. A detailed methodology and evaluation techniques are proposed which ensures the feasibility of the proposed framework. Also suggests that it aligns with the actual problem statement and the main purpose of the research contribution has served properly.

**Conclusion**:

Due to digitization and digitalization, current world is accelerating towards more advanced technologies and digital. Digital data transmission is so easy due to the significant contribution of technology. This digital transmission of information can be so fast that it happens almost in a blink of eye. That’s why people are relying more on these technologies by sharing, storing and retrieving information through them. This digital revolution has changed people’s lives in every aspect. People are replacing their physical machines with virtual ones. Digital transformation has changed the way of living of the people in this era of modern science and technology. Data has become a powerful tool because of the open access of it. But also a few challenges are raised. Data integrity and privacy has become big challenges in the current generation due to this open access of data. As data is generating and propagating through the digital platforms so frequently, it is getting out of manual control by humans and due to that, we need to find ways to automatically get rid of the problem. As machine intelligence and data security systems are always showing new promises, we tried to provide a solution based on these concepts by proposing a hybrid framework. This solution can be implemented on a large scale of digital platforms where complex networks of users exist. Apparently large amount of users can help the system to be more reliable and efficient as decentralized networks proposed in the framework requires more peers in the network. Due to huge amount of users in social media and they are connected to each other this structure will become stronger and the proposed framework has more promising outcomes in the future.

**References**:

Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang and Huan Liu. 2017. Fake News Detection on Social

Media: A Data Mining Perspective. In *ACM SIGKDD Explorations Newsletter, Volume 19.*

Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Alessandro Flammini, and Filippo

Menczer. 2017. The spread of fake news by social bots. In *Nature Communications.*

Leticia Bode. 2016. Political News in the News Feed: Learning Politics from Social Media. In *Journal of*

*Mass Communication and Society, Vol 19, Issue 1.*

Antino Kim, Patricia L. Moravec and Alan R. Dennis. 2019. Combating Fake News on Social Media

with Source Ratings: The Effects of User and Expert Reputation Ratings. In *Journal of Management Information Systems, Volume 36*, Pages 931-968.

Hunt Allcott and Matthew Gentzkow. 2017. Social Media and Fake News in the 2016 Election. In *Journal*

*of Economic Perspectives, Vol, 31, No.2.*

Alexandre Bovet and Hernán A. Makse. 2019. Influence of fake news in Twitter during the 2016 US

presidential election. In *Nature Communications volume 10, Article number: 7.*

Kristina Lerman. 2007. Social Information Processing in News Aggregation. In *IEEE Internet Computing.*

Rich Caruana and Alexandru Niculescu-Mizil. 2006. An empirical comparison of supervised learning

algorithms. In *Proceedings of the 23rd international conference on Machine learning, June 2006 Pages 161–168https://doi.org/10.1145/1143844.1143865*

Irina Rish. 2001. An empirical study of the naive bayes classifier. In *IJCAI 2001 Work Empirical Methods*

*Artificial Intelligence.*

Raymond Wright. 1995. Logistic regression. In *L. G. Grimm & P. R. Yarnold (Eds.), Reading and*

*understanding multivariate statistics (p. 217–244). American Psychological Association.*

Ronan Collobert, Jason Weston, Leon Bottou, Michael Karlen, Koray Kavukcuoglu and Pavel Kuksa. 2011.

Natural Language Processing (Almost) from Scratch.In *Journal of Machine Learning Research, no. 76, vol. 12, pages 2493-2537.*

Natali Ruchansky, Sungyong Seo and Yan Liu. 2017. CSI: A Hybrid Deep Model for Fake News

Detection. In *ACM on Conference on Information and Knowledge Management*.

Ray Oshikawa, Jing Qian and William Yang Wang. 2018. A Survey on Natural Language Processing for

Fake News Detection. Available Online at: [*https://arxiv.org/abs/1811.00770*](https://arxiv.org/abs/1811.00770)

Shuo Yang, Kai Shu, Suhang Wang, Renjie Gu, Fan Wu and Huan Liu. 2019. Unsupervised Fake News

Detection on Social Media: A Generative Approach. In *Proceedings of the AAAI Conference on Artificial Intelligence.*

Federico Monti, Fabrizio Frasca, Davide Eynard, Damon Mannion and Michael M. Bronstein. 2019. Fake

News Detection on Social Media using Geometric Deep Learning. Available Online at: [*https://arxiv.org/abs/*](https://arxiv.org/abs/)*1902.06673*

Kai Shu, Deepak Mahudeswaran, Suhang Wang, Dongwon Lee and Huan Liu. 2019. FakeNewsNet: A Data

Repository with News Content, Social Context and Spatialtemporal Information for Studying Fake News on Social Media. Available Online at: [*https://arxiv.org/abs/1809.01286*](https://arxiv.org/abs/1811.00770)

Douglas M. Hawkins. 2004. The Problem of Overfitting, In *Journal of Chemical Information and*

*Computer Sciences 2004 44 (1), 1-12 DOI: 10.1021/ci0342472*

Mykhailo Granik and Volodymyr Mesyura. 2017. Fake news detection using naive Bayes classifier. In

*IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON).*

E. M. Okoro, B. A. Abara, A. O. Umagba, A. A. Ajonye and Z. S. Isa. 2018. A hybrid approach to fake

news detection on social media. In *Nigerian Journal of Technology, Vol 37, No 2.*

Wenqian Shang, Mengyu Liu, Weiguo Lin and Minzheng Jia. 2018. Tracing the Source of News

Based on Blockchain. In *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS 2018).*

Steve Huckle and Martin White. 2017. Fake News: A Technological Approach to Proving the Origins

of Content, Using Blockchains. In *Big Data, volume 5,* pages 356-371.

Zonyin Shae and Jeffrey J P Tsai. 2019. AI Blockchain Platform for Trusting News. In *IEEE 39th*

*International Conference on Distributed Computing Systems (ICDCS).*

Tee Wee Jing and Raja Kumar Murugesan. 2019. A Theoretical Framework to Build Trust and

Prevent Fake News in Social Media Using Blockchain. In *International Conference of Reliable Information and Communication Technology*, pages 955-962.

Mohamed Torky, Emad Nabil and Wael Said. 2019. Proof of Credibility: A Blockchain Approach for

Detecting and Blocking Fake News in Social Networks. In *International Journal of Advanced* *Computer Science and Applications,* pages 321-327.

Paula Fraga-Lamas and Tiago M. Fernández-Caramés. 2019. Fake News, Disinformation, and Deepfakes:

Leveraging Distributed Ledger Technologies and Blockchain to Combat Digital Deception and Counterfeit Reality. Available Online at: [*https://arxiv.org/abs/1811.00770*](https://arxiv.org/abs/1811.00770)

Adnan Qayyum, Muhammad Umar Janjua, Junaid Qadir and Falak Sher. 2019. Using Blockchain to

Rein in the New Post-Truth World and Check the Spread of Fake News. In *IT Professional volume 21,* pages 16-24.

Xuchan Jua and Yingjie Tian. 2012. Knowledge-based Support Vector Machine Classifiers via Nearest

Points. In *Procedia Computer Science, Volume 9, 2012, Pages 1240-1248.*

Eleanor Clark and Kenji Araki. 2011. Text Normalization in Social Media: Progress, Problems and

Applications for a Pre-Processing System of Casual English. In *Procedia - Social and Behavioral Sciences Volume 27, 2011, Pages 2-11*.

Wen Zhanga, Taketoshi Yoshidaa and Xijin Tangb. 2008. Text classification based on multi-word with

support vector machine. In *Knowledge-Based Systems Volume 21, Issue 8, December 2008, Pages 879-886*.

M. Niranjanamurthy, B. N. Nithya and S. Jagannatha. 2018. Analysis of Blockchain technology: pros, cons

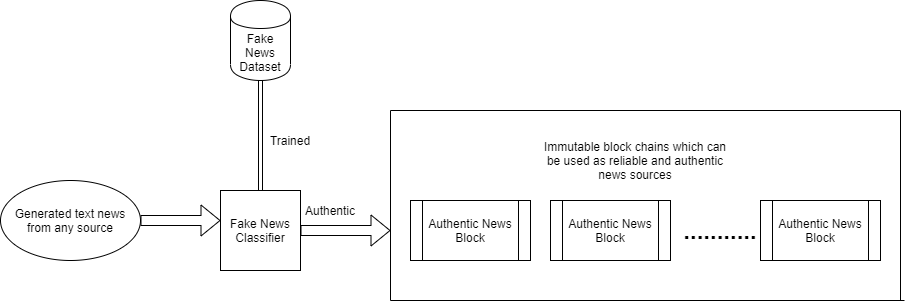
and SWOT. In *Cluster Computing volume 22, pages 14743–14757*.

Jesse Davis and Mark Goadrich. 2006. The relationship between Precision-Recall and ROC curves. In

*Proceedings of the 23rd international conference on Machine learning, Pages 233–240*.

**Appendices**:

Appendix A: Structure of the proposed framework



Appendix B: Drawing hyperplane to separate data points by SVM

