Fake News Detection Using Machine Learning

By

Abidur Rahman (011 171 275)
Ranabir Devnath (011 171 288)
Shubhradev Chakrabarty (011 171 301)
Asif Ahammed (011 171 183)
Susmita Debnath (011 171 105)

June 23, 2021

Department of Computer Science and Engineering
United International University

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Social media for news consumption can be compared with a double-edged sword. It is comparatively low cost, easy to use and access, and rapid dissemination of information leads people to consume news from social media. But this however comes at the cost of questionable trustworthiness and significant risk of exposure to 'fake news', i.e., low quality news with intentionally false information.

1.1 Project Overview

We spent a significant amount of time in our lives interacting with people through social media platforms. We tend to seek out and consume news from social media rather than traditional news organizations. The reason behind this change in consumption behaviors can be derived as:

(i) it is less expensive to consume news on social media compared with traditional news media, such as newspapers or television.

(ii) it is easier to further share, comment on, and discuss the news with friends or other readers on social media.

Despite having these advantages, the authenticity of news on social media is significantly lower than traditional news organizations. Because of its easiness and cheapness, large volumes of fake news, i.e., those news articles with intentionally false information, are produced online for a variety of purposes, such as financial and political gain. One individual or society can be badly impacted by this immense spread of fake news. Nowadays, it is very difficult to identify a news whether it is fake or real. So, we will build a machine learning model that will automatically detect if a news is true or not.

1.2 Motivation

Fake news is misleading or untrue information which has the aim of damaging the reputation or entity or making money. It can be extremely dangerous and misleading in a

society. According to stats in the last ten years, there are several brutal incidents reported due to the spread of falsified news in social network platforms which are considered as the easiest media for spreading rumors. Mostly clumsy and innocent people are used to lead these violences because they are prone to believe these rumours.

In this digital era, it is probably the greatest challenge to control the spreading of bogus or deceiving news because of the free progression of data through social networking sites like Facebook, Twitter, YouTube and others. Another report says residents of US, Spain, Germany, UK, Argentina and South Korea guarantee they have seen bogus or deluding data via online media related to CoVID-19. Because of the abundant spreading of false and unreliable information, some commentators are now referring to the latest wave of misinformation that's accompanied by the CoVID-19 pandemic as a disinfodemic.

Research shows that the rumors spreading in web-based media leave an especially enduring impact on less smart individuals and get them far from settling on the right choices. False news is utilized to build individuals' fears, raise racist thoughts and lead to harassing and violence against innocent people.

Indeed, fake news has extraordinary democratic effects. For example, American presidential election shows how it disturbs people's feelings. Over the most recent couple of years, there have been numerous tragic occurrences in Bangladesh due to rumors. In July 2019, five people were beaten to death and ten injured by mobs as a result of widespread rumors about the expected human sacrifice in the construction of the Padma Bridge.[Ref] In September 2012, a series of attacks occurred on Buddhists, shrines, and places of Buddhist inhabitants in Ramu Upazila in Cox's Bazar by local mobs at midnight. [Ref]

Therefore, we can see that fake news is a very dangerous problem. In societies like ours, this evil can do considerably more harm than anywhere else. Because people here are less critical and more prone to believe any kind of campaign. To help mitigate the negative effects caused by fake news-both to benefit the public and the news ecosystem-It's critical to develop methods to automatically detect fake news on social media.

The diagram [1] below is showing the number of papers on fake news and rumors in several years.

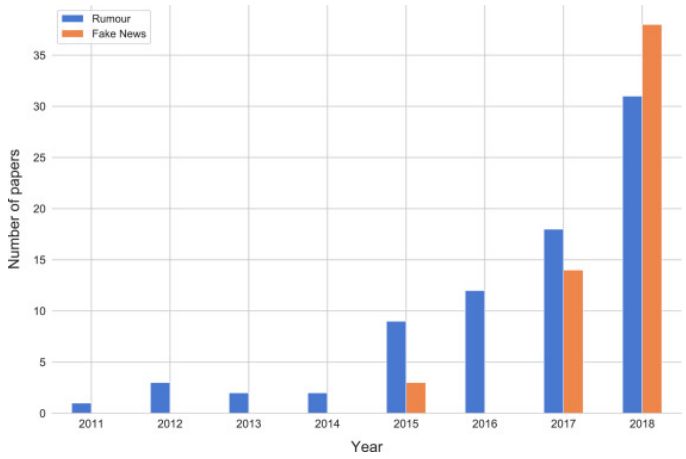

Figure 1.1: Survey report analysis on fake news and rumors detection techniques

1.3 Objectives

The objective of this project is to discover the viability and impediments of language-based systems for detecting any type of fake news which is detected using machine learning algorithms, AI calculations.

- The aim of this project is to provide a systematic evaluation of the machine learning algorithm for fake news prediction
- This can save time for people. Online news readers would be more benefited by using this project.

1.4 Methodology

- Pre-processed dataset on this field especially in Bangla is very rare. So, for training this system, curating data from different social media is needed.
- The dataset, which has been curated manually, is real-life data. First of all, they have to be labeled as “Real” or “Fake”, which is very difficult to insure. Also, the data-set is incomplete, unstructured, and noisy. So, the dataset has to be manually preprocessed.
- Feed the dataset to the NLP based machine learning algorithms to make a logical model.
- Then test any news post collected from social media to find the output. And measure the accuracy using various performance measuring tools.

After that a machine learning model can learn without being explicitly programmed. it can improve its performance by gaining more data. Then we can use our trained machine learning model to classify fake or true news.

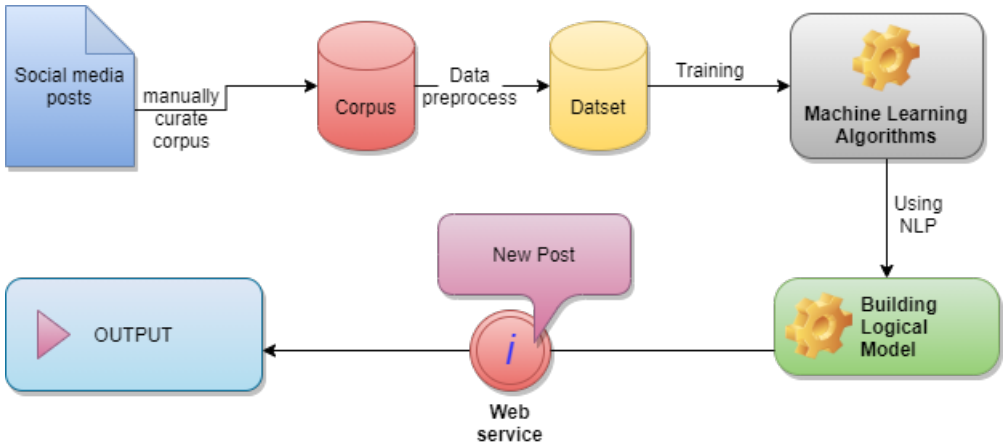

Figure 1.2: Working Process of this project

1.5 Project Outcome

Our goal is to identify fake news from social media. For reaching our goal, we are expecting to make a web service with the properties of identifying fake news. Which will work as an extension or plugin with social media websites. The service will take keywords from a specific post or news from a particular media and run different algorithms over it. Then it will finally predict whether the **social media post** or news is possibly fake or not.

Chapter 2

Background

2.1 Preliminaries

In this part, we will use some terms and definitions that which will be used throughout the paper.

2.1.1 Model Building

Machine Learning: Machine learning is a part of Artificial Intelligence that enables a machine to be functional from its input. Some of the basic key factors like Self-evaluation, Future information prediction, making a critical decision that defines machine learning in a modern way. Generally, machine learning algorithms can be divided into 4 subcategories: Supervised learning, Unsupervised learning, Semi-supervised learning, and Reinforcement learning. Judging the nature of this paper Supervised learning technique suitable here.

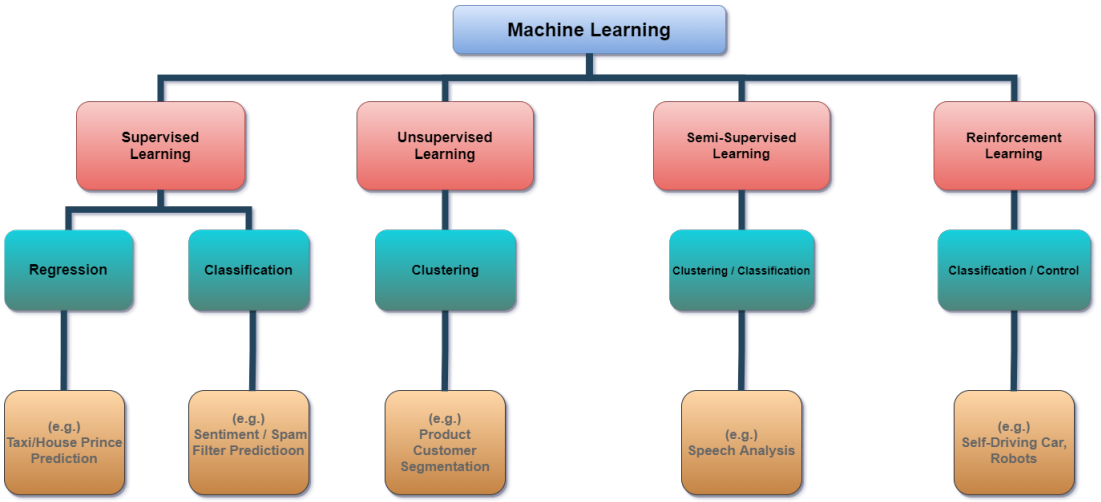

Figure 2.1: Classification of Machine Learning

Supervised learning: Supervised learning is one kind of machine learning algorithm that learns from a set of labeled examples to generalize possible outputs. In this process, the model is trained using labeled data. It means that those training instances are attached with the correct answer. After that some new unseen and new instances will be given to the model, then the model will analyze and generate an output for the unseen unlabeled data.

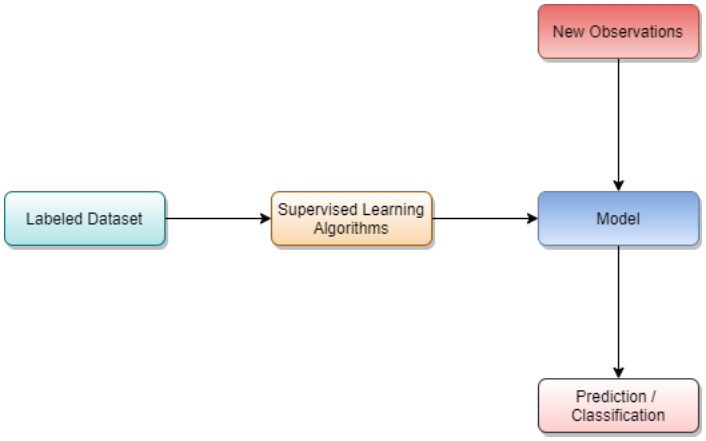

Figure 2.2: Supervised Learning

This project is text analysis based so Natural Language Processing(NLP) generally appears here.
NLP: Natural Language Processing is a field of machine learning with the ability to, analyze, understand, and potentially generate human language by a computer. A process of teaching machines about human natural sentiment or language. Here is a high-level overview of the NLP workflow.

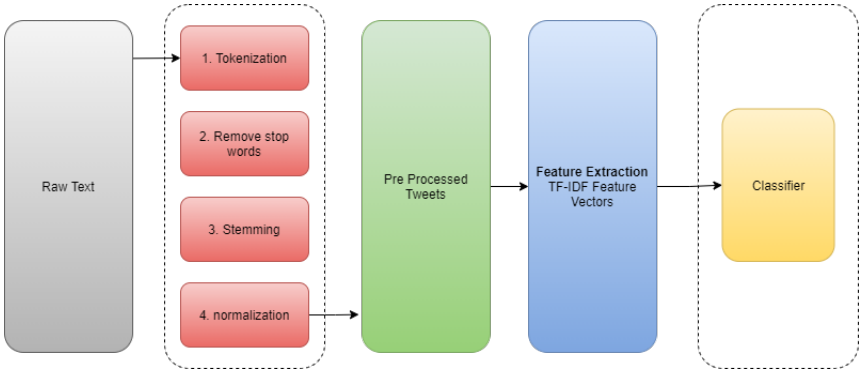

Figure 2.3: Natural Language Processing.

Since text data can not be fed to the machine, that is why it needs to be converted into some kind of numeric vector. After removing all the stopwords and Applying Tokenizer, a feature extraction technique is applied to convert them into vectors and these will go through a classifier to generate the desired output.

Tokenizer: Tokenizer divides a sentence into a single piece of a word. It is the process of tokenizing or splitting a string, text into a list of words or tokens. Visualizing this definition, "Natural language processing" will be converted into ["Natural", "Language", "Processing"].

Stopwords: They are the set of commonly used words in a language. Stop words are heavily used in Text Mining and Natural Language Processing. It obliterates words that are so frequently used. It because they carry very little useful information.

TF-IDF: TF-IDF vectorizer is one of the most popular feature extraction techniques used in NLP and Text Mining. A process that is used to find the meaning of sentences that contain words. It is a statistical measurement that evaluates how significant or relevant a word is in a document. By applying TF-IDF, information about the important words can be stored.

Neural network: Neural networks [2] is a deep learning technique, matched closely after the human brain. It is designed to recognize patterns from data. Neural networks refer to a system consisting of neurons, that is artificial in nature. It helps to cluster and classify or many tasks relevant to supervised learning. Neural networks can also extract information and can also create features from the data that are fed to it for clustering and classification. Generally, Neural Network consists of three-layer. An input layer, hidden layer, and output layer. The layers of functions between the input and the output are what make up the neural network.

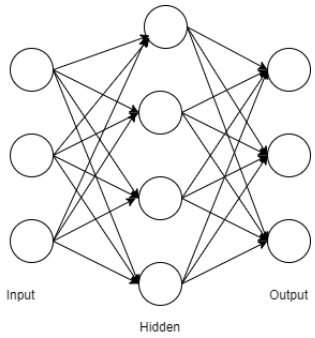

Figure 2.4: Simple Neural Network Model

A node layer is a row of those neurons. It is like a switch, when input is fed through the network it turns on or off. Here each layer's output is its subsequent layer's input. Starting from an input layer that receives the data and an output layer that predicts the result.
The layers are a place where all calculations take place. They are made of nodes. A node usually integrates input from the data with a set of weights or coefficients that either intensify or dampen that input. After that these input-weight products are summed together

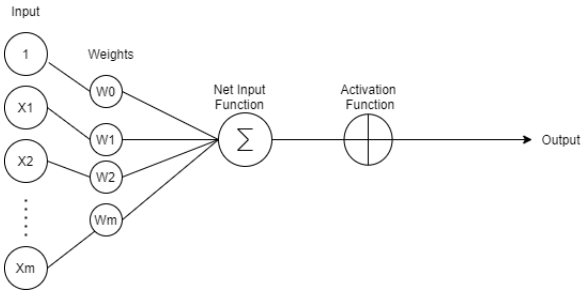

Figure 2.5: Diagram of Nodes

and then the result is passed through a node's activation function. It also determines whether that signal should progress further through the network to affect the ultimate outcome.

Support Vector Machines (SVM): SVM [3] is a type of classifier that deals with non-linear feature space by using kernel methods. It uses a multidimensional hyper-plane to find the maximum margin between the data points. It also separates them into different classes. It transforms the input space into higher dimension to reduce the classification error.

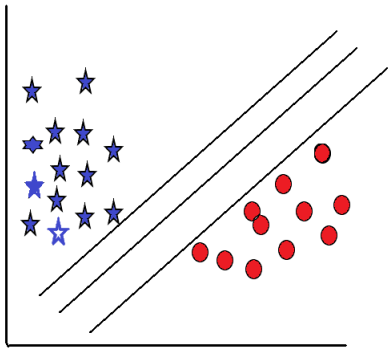

Figure 2.6: SVM-Classifier

2.1.2 Model Evaluation

Accuracy: The number of data points from all samples that is correctly predicted is called the accuracy. Accuracy gives the result how good the model is working.

$$Accuracy = \frac{TruePositive + FalsePositive}{AllSamples} \tag{2.1}$$

Precision: Precision is defined as the fraction of admissible instances amongst all redeemed instances, which means among all the positive results how much true positives are found.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \tag{2.2}$$

Recall: Recall, also known as sensitivity, is the fraction of redeemed instances amongst all admissible instances, means among actual positives how much true positives are found.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \tag{2.3}$$

F1 Score: The F1 score is defined as the weighted harmonic mean of the test's precision and recall. This score is calculated according to the formula :

$$F1Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{2.4}$$

Accuracy: The number of data points from all samples that is correctly predicted is called the accuracy. Accuracy gives the result how good the model is working.

$$Accuracy = \frac{TruePositive + FalsePositive}{AllSamples} \tag{2.5}$$

Precision: Precision is defined as the fraction of admissible instances amongst all redeemed instances, which means among all the positive results how much true positives are found.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \tag{2.6}$$

Recall: Recall, also known as sensitivity, is the fraction of redeemed instances amongst all admissible instances, means among actual positives how much true positives are found.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \tag{2.7}$$

F1 Score: The F1 score is defined as the weighted harmonic mean of the test's precision and recall. This score is calculated according to the formula :

$$F1Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{2.8}$$

2.2 Literature Review

In 2018, Wang et al [4] proposed an end-to-end framework which is called Event Adversarial Neural Network (EANN). It can derive event-invariant features so it has the benefit of detecting fake news on newly arrived events. Their framework has three components: the multi-modal feature extractor, the fake news detector, and the event discriminator. The multi-modal feature extractor extracts the visual and textual features from the posts. For detecting fake news it also collaborates with the fake news detector to learn the distinct

representation. Event discriminator removes features of any specific event. It also keeps shared features among the events. Their result showed 71% accuracy on Twitter dataset and 83% accuracy on Weibo dataset.

In 2017,Shu et al [5] worked on The characterization of fake news. In this survey, they presented a comprehensive review of detecting fake news on social media. It also includes fake news characterization on social theories and psychology, existing algorithms from a data mining perspective. They used some widely popular datasets such as "BuzzFeedNews", "LIAR", "BS Detector", "CREDBANK". They used a Knowledge-based approach and style based approach for solving this kind of fake news detection problem.

In 2019, Monti et al [6] said that frequently the evaluation of any fake news requires the knowledge of social or political context or 'common sense'. Real news and fake news spread differently on social media as they claimed. It forms a propagation pattern that could be used for detecting fake news. They proposed an automatic fake news detection model based on geometric deep learning. They observed that social network structure and propagation are the most important features that allow a model higher accuracy and also the model can detect fake news at an early stage.

In 2019, Zellers et al [7] worked on neural fake news and defending against them. They build a controllable text generation model named Grover. It has the capability to generate the rest of the article from a short given article. They found that the currently best discriminators can classify neural fake news from real, human-written, news with 73% accuracy. According to their claim the best defense against Grover turns out to be Grover itself, with 92% accuracy.

In 2017,Ruchansky et al [8] mentioned that there are generally three agreed-upon characteristics of fake news. the text of an article, the user response it receives, and the source users promoting it. So, their model includes all of the three characteristics for automated and accurate prediction. Their model consists of three modules: Capture, Score, and Integrate. The first module based on Recurrent Neural Network to capture the temporal pattern of user activity on a given article. The second module learns the behavior of users. The third module detects if an article is real or not.

In 2020, Hussain et al [9] offered a solution for fake news detection in the Bengali language. They utilized two supervised machine learning algorithms. They are Multinomial

Naive Bayes (MNB) and Support Vector Machine (SVM) classifiers. They are used for identifying Bangla fake news. They achieved 96.64% accuracy on SVM with linear kernel and 93.32% accuracy on MNB.

In 2017, Veronica Perez et al [10] used a linear SVM classifier and five-fold cross-validation, with accuracy, precision, recall, and F1 measures averaged over the five iterations.They have covered seven different domains of news and introduced two different datasets.One is the combination of manual and crowd-sourced news data set, and the other is directly collected from the web. For the dataset build up, they have extracted five different features,they are: Ngrams, Punctuation, Psycho-linguistic feature, Readability and Syntax. They conducted several exploratory analysis to identify linguistic properties that are predominantly present in fake content,and acquired 78% accuracy which is comparable to human ability to spot fake content.

In 2015,Conroy et al [11] defined "Fake news detection" as the task of categorizing news along a continuum of veracity, with an associated measure of certainty. They proposed a typology of several varieties of veracity assessment methods emerging from two major categories – linguistic cue approaches (with machine learning), and network analysis approaches. They achieved accuracy of 86% using Support Vector Machine.

A comparative analysis of reviewed papers:

Ref. No.	Algorithms	Research method	Data Set	Accuracy	Lang	Limitation
[4]	Machine Learning(RNN, CNN)	The multi-model feature extractor,and the event discriminator	Real-life dataset Source: Twitter, Weibo Size:10805 (real), 12647(fake)	Twitter-71.5% Weibo-82.7p%	English	Accuracy can be improved by implementing some hybrid-classifiers.
[5]	Data Minning	Fake news Characterization and Detection			English	Requires a pre-annotated fake news ground-truth dataset to train a model

[6]	Geometric Deep Learning	Feature Extraction, Classification	Type: real-life Source: Twitter Size:1080	92.7%	English	Network manipulations that are difficult to implement in practice
[7]	NLP(BERT, Transformer)	Feature Extraction, Classification	Source: common crawl Size: 5000 news articles	92%	English	Primarily leverage distributional features rather than evidence
[8]	RNN	Capture(Text analysis) Score(source behavior analysis) Integrate(integration of both module).	Type: real-life Source: Twitter, Weibo Size:2845(real), 2811(fake)	Twitter-89.2% Weibo-95.4%	English	Incorporate concepts from reinfor cement learning and crowdsourcing
[9]	MNB and SVM Classifiers	Data Collection, Data Preprocessing, Feature Extraction, Classifier	Type: real-life + manually created fake news Source: online news portal Size:1548(real), 993(fake)	SVM-96.64% MNB-93.32%	Bengali	Larger dataset to expand the number of features and sufficient lexicons.
[10]	Linear SVM classifier	Linguistic Features Extraction, classification	Type: Manual and Crowdsourced Source: From web Size: 1627(Fake and real) Size: 240(Fake)	78%	English	Worked only on seven specific domains
[11]	NB and SVM Classifiers	linguistic cue approaches (with machine learning), network analysis approaches	Type: real-life + manually created fake news Source: from web		English	The ability to determine alignment between attributes and descriptors

Table 2.1: Critical review.

2.2.1 Similar Applications

NewsCop: It basically gives a Verified Fake news from various social media sites like Facebook,WhatsApp,twitter and many more.In this application if a user posts news for verification, they listed the post in polling section and then NewsCop users can give their vote (as per their knowledge of the content) to verify the post. There is no use of Machine learning or Artificial intelligence here.

Oigetit fake news filter: This application uses Artificial Intelligence and different types of algorithms to compute the facts,validity of any news that one reads on the internet.It delivers valid news with AI filtering technology.

FakeFinder: It is a mobile app that can detect fake news from the live Twitter stream and alert users in real time.

BdFactCheck: They are Bangladesh's first fact-checking professional organization that aims to reduce the level of deception and confusion in Bangladesh. They monitor news from traditional media, social networking sites, and public places said political party leaders, public figures, intellectuals and their authenticity.

2.2.2 Related Research

2.3 Gap Analysis

After studying several papers and projects we found that most of them are in the English language. Because it is an international language of technology and communication. It is spoken and understood in most of the countries in the world, that is why it is very important. But there are few full research papers in detecting Bengali news.

As the grammar of the Bengali language is much different from English grammar, to identify Bengali fake news one has to build a new model from scratch that can detect Bengali news easily. It will not be an easy process.

Besides, the collection of data for this project is going to be difficult because the fake and real news has no other exceptional identity to detect them. So, one has to scrap more to collect all types of data.

Chapter 3

Project Design

3.1 Requirement Analysis

3.1.1 Functional Requirements

User Interface

As there will be several information needed from user side for finding fake news,the system must need an user interface to interact with the users.

Hardware Networking Interface

User must have the internet access to use the system.So they should have network interface on their device that they are using for connecting with the internet.

Communication Protocol

As the interaction between user and system will happen through internet, the network connection must maintain the TCP/IP protocol.

3.1.2 Non-functional Requirements

Accuracy

If the generated result is not accurate, the system means nothing. So accuracy is one of the most important requirements of this system.

Ease of Use

The system should be user friendly in use. Because if it is complicated,the user will loose attraction over it.

Information Security

As the system will be a subscription based system, so the user will share many personal or transaction information through this system. So it is very important to ensure the security of these information.

3.1.3 Use Cases

Use case diagram of our project has shown below

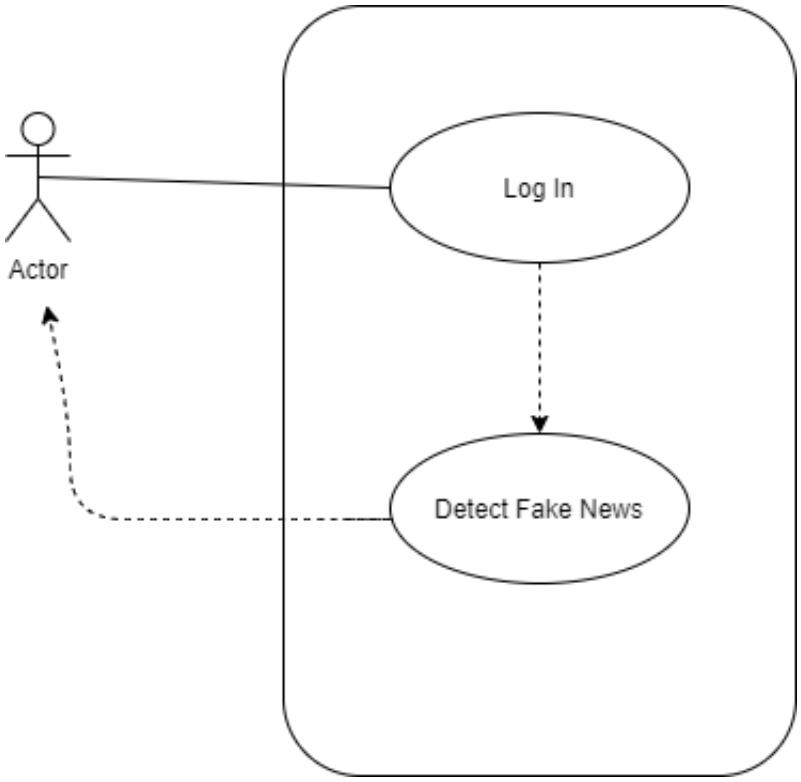

Figure 3.1: Use Case Diagram

Use case description: UC-01

	Name	UC-01: Log in	
	Primary Actor	User	
	Pre-condition	User must have internet access and signed up before.	
	Success Scenario	If the username and password has been entered correctly the system will show featured page to the user.	

	Alternate Scenario	If the user do not have any account, the system will take him/her to sign up page. If any of the field did not match with the database,the system will ask the user to put correct values.	
	Post-condition	After successfully logged in, user can put news for checking wheather the news is true or fake	

Table 3.1: UC-01

Use case description: UC-02

	Name	UC-02: Detect fake news	
	Primary Actor	System	
	Pre-condition	User must have logged in to the system and put the news for checking.	
	Success Scenario	The user will immediately get the result wheather the given nes is true or fake.	
	Alternate Scenario	If the user insert the wrong format of news then the system will stop detecting and will show a alert for giving the right format.	
	Post-condition	Detected result will displayed on the user interface.	

Table 3.2: UC-02

3.1.4 Data Flow Diagram level 1

Data flow diagram of our project has shown below

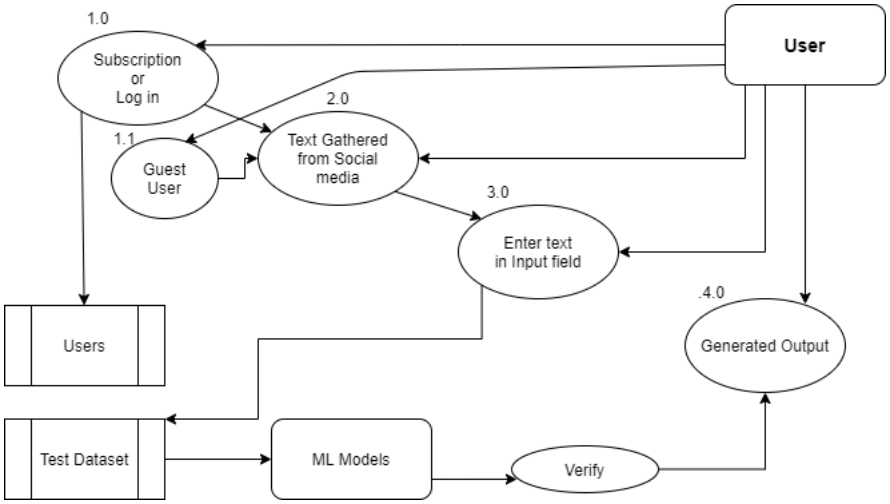

Figure 3.2: Data Flow Diagram

3.1.5 Task Allocation

Allocation of future tasks has shown below.

	Corpus Collection	Public Survey	Meet Stakeholders	Model Buildup	System develop	Report Writing
Abidur Rahman			√	√	√	√
Ranabir Devnath	√			√	√	√
Shubhradev Chakrabarty			√	√	√	√
Asif Ahmed	√	√	√	√		
Susmita Debnath	√	√	√		√	

Table 3.3: Task Allocation

Chapter 4

Implementation and Results

.

4.1 Environment Setup

4.2 Testing and Evaluation

4.3 Results and Discussion

Chapter 5

Standards and Design Constraints

5.1 Compliance with the Standards

5.1.1 Software Standards

Standards	Available Options	Chosen	Reason
Programming languages	Python, R, Julia	Python	Has extensive support libraries. simple and portable. High readability
Version Controller	BitBucket, GitHub	GitHub	Open source, allows branching and provides cloud-based repository, free and reliable
Design Standard	UML, OAA	UML	High readability, allows to design, thoroughly, visual representation, abundance of available tools
Project Management	Trello, JIRA	JIRA	Allows teams to easily plan, release, track any associated issues. supports several agile methodologies.
Markup Languages	Microsoft Word,Google Doc,Latex	Latex	Better for tables and illustrations. Consistent handling of references and bibliography.
Data Formats	Microsoft Word,Google Doc, CSV	CSV	Faster to load and import, consumes less memory. Parsing data from it is trivial and Python library makes it easier to handle CSV.

IDE	Jupyter Notebook, Weka,Google Colaboratory, PyCharm	Google Colaboratory, PyCharm	Colab for Automatic history and versioning and sharing. TPU support apart from its existing GPU and CPU instances. Pycharm for its Coding friendly environment on local machines

Table 5.1: Software Standards

5.1.2 Communication Standards

5.2 Design Constraints

5.2.1 Economic Constraint

This is an online-based machine learning process. So a subscription-based earning system could be applied.It will save time for social media users to verify the authenticity of any news and it is also cost-effective.It doesn't create or replace any jobs because there is no specific job for verifying the authenticity of social media news.

5.2.2 Ethical Constraint

It will help disadvantaged people to be aware of fake news of niche marketing over them.In terms of diplomacy, it's questionable that if the truth behind it is really favorable or unfavorable.

5.2.3 Social Constraint

Social entropy could be reduced by using this system.People will get instant awareness between true and fake news. Wide uses of this system will greatly reduce the tendency of spreading misinformation or fake news.Sometimes false identification of news, devastating entropy could occur in Society also.But we all know fake news poses a substantial threat to every society, with serious negative consequences. As we are trying to get rid of this problem by the system, so it's definitely not against any social norms.Everyone of the society should accept it.

5.2.4 Manufacturability

As it's a web based project, if we can store it in a big server, we can manufacture it in a large number. It will be profitable too as it will be a subscription base service system.

5.2.5 Sustainability

It is questionable until the validation stage. The machine needs to be updated from time to time. After the development phase, we can continuously update our product based on the customer requirements.

5.3 Cost Analysis

Analysis of required cost for developing the system is given below:

5.3.1 Management

Cost analysis for management of our project.

Component	**Cost(BDT)**
Project Management Tool	7000/Year
High End Desktop	120000

Table 5.2: Management Cost Analysis

5.3.2 Deployment

Cost analysis for deployment of our project.

Component	**Cost(BDT)**
Domain	10000/Year
Hosting	5000/Year
Virtual GPU	7000/Year

Table 5.3: Deployment Cost Analysis

5.4 Complex Engineering Problem

5.4.1 Complex Problem Solving

P's that are addressed through this project:

P1 Dept of Knowledge	P2 Range of Conflicting Requirements	P3 Depth of Analysis	P4 Familiarity of Issues	P5 Extent of Applicable Codes	P6 Extent of Stakeholder Involvement	P7 Interdependence
✓	✓	✓	✓			✓

Table 5.4: Mapping with complex problem solving.

P1: This project requires the study of existing models with similar goals(K8), corpus collection from social media(K4), statistical knowledge of data analysis(K3), knowledge of designing of the machine learning based model(K3, K4), integration of different components(K5, K6).

P2: Conflicting technical requirements: Collecting data from mainstream social media like Facebook is very difficult. To achieve higher accuracy from our model, we need a huge amount of data. We have to collect data that will be very tough to handle because they will be unstructured, incomplete, and noisy. So balancing good accuracy within limited data as well is a tough job.

P3: Due to the quantity and the quality of data of social media, no obvious formula can be applied to pre-process or classify as a Machine learning problem. Depth of analysis is needed to find a way to pre-process and select a specific algorithm.

P4: Since there is no agreed definition of the term "Fake News", it is very tough to characterize Fake news. It has a complex pattern. Many psychological and social theories are related to fake news. So, As a Computer Science and Engineering student, "Fake News Characterization" is very burdensome.

P7: The dependency of sub-systems in this project is common. Like data collection, labeling, and pre-processing, model training, building a user interface, etc.

5.4.2 Engineering Activities

A's that are addressed through this project:

A1 Range of resources	A2 Level of Interaction	A3 Innovation	A4 Consequences for society and environment	A5 Familiarity
√	√		√	√

Table 5.5: Mapping with complex engineering activities.

A1: In this project, our stakeholders are various kinds of social media users. We have to scrape data from mainstream social media like Facebook or Twitter. Our motive is to build an extension on social media.

A2: We have to interact with news experts and traditional newspapers to identify the authenticity of the news.

A4: The main consequence of classification social media news, is the false classification of news. As this is a machine, it's possible to get the wrong output. And to mitigate that, we think we have to feed our machine more and more data for increasing the accuracy of results.

A5: This project has been done previously in English and another language. But in Bengali, very few projects have been done. Also these are not widely popular.

References

[1] Alessandro Bondielli and Francesco Marcelloni. A survey on fake news and rumour detection techniques. *Information Sciences*, 497:38–55, 2019.

[2] Mohamad H Hassoun et al. *Fundamentals of artificial neural networks*. MIT press, 1995.

[3] Johan AK Suykens and Joos Vandewalle. Least squares support vector machine classifiers. *Neural processing letters*, 9(3):293–300, 1999.

[4] Yaqing Wang, Fenglong Ma, Zhiwei Jin, Ye Yuan, Guangxu Xun, Kishlay Jha, Lu Su, and Jing Gao. Eann: Event adversarial neural networks for multi-modal fake news detection. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery Data Mining*, KDD '18, page 849–857, New York, NY, USA, 2018. Association for Computing Machinery.

[5] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. Fake news detection on social media: A data mining perspective. *SIGKDD Explor. Newsl.*, 19(1):22–36, September 2017.

[6] Federico Monti, Fabrizio Frasca, Davide Eynard, Damon Mannion, and Michael M. Bronstein. Fake news detection on social media using geometric deep learning, 2019.

[7] Rowan Zellers, Ari Holtzman, Hannah Rashkin, Yonatan Bisk, Ali Farhadi, Franziska Roesner, and Yejin Choi. Defending against neural fake news, 2020.

[8] Natali Ruchansky, Sungyong Seo, and Yan Liu. Csi: A hybrid deep model for fake news detection. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, CIKM '17, page 797–806, New York, NY, USA, 2017. Association for Computing Machinery.

[9] M. G. Hussain, M. Rashidul Hasan, M. Rahman, J. Protim, and S. Al Hasan. Detection of bangla fake news using mnb and svm classifier. In *2020 International Conference on Computing, Electronics Communications Engineering (iCCECE)*, pages 81–85, 2020.

[10] Verónica Pérez-Rosas, Bennett Kleinberg, Alexandra Lefevre, and Rada Mihalcea. Automatic detection of fake news. In *Proceedings of the 27th International Conference*

on Computational Linguistics, pages 3391–3401, Santa Fe, New Mexico, USA, August 2018. Association for Computational Linguistics.

[11] Nadia K. Conroy, Victoria L. Rubin, and Yimin Chen. Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*, 52(1):1–4, 2015.