BANGLA FAKE NEWS DETECTION USING MACHINE LEARNING

BY

AFSANA KHANOM
ID: 161–15–7396

AND

HUMAYRA KHANUM
ID: 161-15-7325

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering.

Supervised By

Mr. Sheikh Abujar
Lecturer
Department of CSE
Daffodil International University

Co – Supervised By

Mr. Shaon Bhatta Shuvo
Senior Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

DECEMBER 2019

APPROVAL

This thesis titled **"Bangla Fake News Detection Using Machine Learning"** submitted by Afsana Khanom, ID No: 161-15-7396, Humayra Khanum , ID No: 161-15-7325 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 6 December, 2019.

BOARD OF EXAMINERS

Dr. Syed Akhter Hossain — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Nazmun Nessa Moon — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Fizar Ahmed — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Mohammad Shorif Uddin — **External Examiner**
Professor
Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Mr. Sheikh Abujar, Lecturer, Department of CSE**Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma. [Font-12]

Supervised by:

Mr. Sheikh Abujar
Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Mr. ShaonBhattaShuvo
Senior Lecturer
Department of CSE
Daffodil International University

Submitted by:

AfsanaKhanom
ID: -161-15-7396
Department of CSE
DaffodilInternationalUniversity

HumayraKhanum
ID: -161-15-7396
Department of CSE
DaffodilInternational University

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We really grateful and wish our profound our indebtedness to **Mr. Sheikh Abujar**, **Lecturer**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Machine learning*" to carry out this project. His endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to Professor Dr. Syed Akhter Hossain, Head**,** Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

Nowadays, Social platform has become an exoteric means for audience to devour news. The proliferation of perplexing fact in regular access media outlets such as social media sites, news blogs, and online portals have created it difficult to detect actual news sources, thus increasing the need for computational tools able to provide clear-sightedness into the reliability of social media content. Extending fake news in social media is often higher than traditional news sources. The augmentation of Bangla fake news and its extension on social media has become a main anxiety due to its caliber to make demolishing dominance. Different machine learning intercourse have been endeavor to identify English fake news. But in our survey we build machine learning model called “Random Forest Machine Learning Model” through decision tree. In this research, we propulsion a benchmark study to detect fake news from online portal through comparing both true and false ne We also implemented some advanced deep learning models (CNN,CRNN,GRU,LSTM) that have shown promising results for detecting fake news. In our research based project we also build two more models called “Support Vector Machine” and “k-Nearest Neighbor”. We do comparison between three models for getting better accuracy. We have applied random forest, SVM and KNN on the same data test set and train set and found that random forest performs far better than the other two models.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURES **PAGE NO**

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Nowadays social media has become an elementary origin of fake news consumption. Online portal is less cost consumption, easy to access, and can fast proclaim posts. Therefore, the likelihood of extending fake news in social media is often higher than traditional news sources. It was also found that social media now outperforms television as the major news source [2].

Fake news is news, stories or hoaxes created to deliberately misinform or deceive readers [3]. Fake news can spread in several ways such as parody, sloppy journalism, misleading/puzzling headings, biased news and so on [3].Fake news detection is a critical matter that needs to be found out. Fake news can Intermission the authoritativeness balance of the news ecosystem. For example, it is evident that the most popular fake news was even more widely spread on Facebook than the most popular authentic mainstream news during the U.S. 2016 the president's election [6]. Some fake news was just created to trigger people's distrust and make them confused, impeding their abilities to differentiate what is true from what is not [5].

There are so many works have done to detect English fake news detection. But rarely can find any research which can detect Bangla fake news detection. However, we have decided to work on this issue by applying natural language processing and machine learning model. In this article, we immediate an overview of Bangla mesh news detection and prattle promising research monition. To control the extension of fake news we have done survey. In our research based project we are proposing a system where user can spontaneously copy paste the link of that news in the input section and as output it will show them the probability of that news being mesh which can surely reduce chaos about the news is being false or not.

We have collected some data of both true and false from social media. We have trained our dataset. Through comparing both types of news we can detect which is false or which is true

by applying machine learning model .To identify whether it is a false/ true, we built a binary classifier using deep learning methods. A review of the learned models uncloses the subsistence of major sentences that control the presence of fake news. Our model services at the sentence level as blockaded to paragraph level attention in .The overall workflow of our study has been demonstrated in figure 3.1.

1.2 Motivation

Nowadays, people post on Social media like Facebook, twitter and other online sites and traditional news portals often disclose Bangla mesh news for which these sites have obtained negative animus for flourishing articles that later substantiated to be false and these fake news mislead the readers and originate many issues. In our research based project we are proposing a system where user can spontaneously copy paste the link of that news in the input section and as output it will show them the probability of that news being mesh which can surely reduce chaos about the news is being false or not.

There are also some specific motives in our research based project

1. Bangla fake news on social media has been spreading rapidly. To better exhibitor the future directions of Bangla fake news detection research, eligible ramification are obligate.
2. Through a review study we find that there are some parting patterns that can be turned to advantage for Bangla fake news identification in social media.
3. Nowadays, people don't have that much time to read through the full news. Often they just get wrong information by reading only misleading/perplexing headlines. So we decided to do this survey to identify Bangla fake news.

In addition the illegal trade of spreading Bangla fake news detection is almost possible. It is possible to identify illegal trading information from social media using random forest model. Classification is a method of finding the fake news in a document and can be used as classifying text of bulk size where the decision tree methods are much harder and impossible to operate.

1.3 Rationale of the Study

In research area, there is so much research that were prosecuted to detect fake news. But for the Bangla fake news detection is so scarce. Very few work has been done in this topic. Also, Bangla news detection carriages much challenges that English fake news detection don't have. Fake news can be classified into various types. Similar some works have been done previously for Bangla fake news detection such as "A Benchmark Study on Machine Learning Methods for Fake News Detection" [6].

1.4 Outcome

This research work aims at

1. Classifying Bangla news in binary format to know whether the news is fake or real.
2. Identifying Bangla fake news from social platform.
3. Comparing both true and false news, show the possibility of that news being fake.
4. Increase accuracy on Bangla fake news detection.

The total research is about model building which aims how to give the accuracy of any news. Here a short description how does this model RF works. First of all root node splits into 2 parts. The two parts are decaled as next decision node and then it splits into more braches. After the part by part splitting we get our expected leaf node which actually is the main output.

1.5 Report Layout

In this chapter we have discussed about the introduction of the sufficiency to identify Bangla fake news, motivation, rational of the study and the outcome of the thesis. Later followed by the report layout.

In chapter 2, we will discuss about the background of our research topic.

In chapter 3, we will discuss about the methodologies employed in our study.

In chapter 4, we will discuss about the acquired results and discussion.

In chapter 5, we will discuss about the conclusion and future work.

CHAPTER 2

BACKGROUND

2.1 Introduction

Previously several researches have been conducted to find “English Fake News Detection from social media. In such cases, the work of Glowacki et al. [15] showed that the public concern about the misusages can bring havoc. To tackle those situations much research has been done, though are not sufficient.

There are many illustrations where knowingly schematic Bangla mesh, now the world is digital platform where everything is relying on the internet. We don’t use take efforts to buy newspaper or don’t spend so much time to click a link and read the full news. We just scroll down in social media and read the headlines. The swiftness with which misinformation create its possibilities online and find an audience is singular in the world of communication. Identifying Bangla fake news is in chief risky for two issues. First of all, the nature and the characteristics of the news in social media largely withstands any generalization of the non-verbal behavior of the sender of the fake news (see Zhou & Zhang, 2008).

Furthermore, now the world is digital platform where everything is relying on the internet. We don’t use take efforts to buy newspaper or don’t spend so much time to click a link and read the full news. We just scroll down in social media and read the headlines. By this we easily get misinformation and starting misjudge. So in our research based work we make a survey where the misinformation headline can be detected by applying machine learning model. We collect dataset in both cases true/ false. Comparing this two cases our model can give proper output about the news is true or fake through decision tree.

2.2 Related Works

There are so many works have done to detect English fake news detection. But rarely can find any research which can detect Bangla fake news detection. However, we have decided to work on this issue by applying natural language processing and machine learning model. To control the extension of fake news we have done survey. We have collected some data of both true and false from social media. We have trained our dataset. Through comparing both types of news we can detect which is false or which is true by applying machine learning model .To identify whether it is false/ true, we built a binary classifier using deep learning methods. A review of the learned models uncloses the subsistence of major sentences that control the presence of fake news. Our model services at the sentence level as blockaded to paragraph level attention in .The overall workflow of our study has been demonstrated in figure 1. There are many illustrations where knowingly schematic Bangla mesh, now the world is digital platform where everything is relying on the internet. We don't use take efforts to buy newspaper or don't spend so much time to click a link and read the full news. We just scroll down in social media and read the headlines. From our study we have seen some works has been done on English fake news detection but we have not seen any work on Bangla fake news detection except one recent research based work. From our research study we find 1 related work which is done but Buet student's group which help us a lot to know about related work. We consider the English fake news detection pattern as related work for our survey.

2.3 Comparative Studies

Even before this, a lot has been written and a lot of research has been done on the English Fake news detection. Everyone has an idea about fake news. The brain of the human being is remedied. But at the same time, they trust what they read at a glance. We have tried to build a system for fake news detection for Bangla news. So that people can't be misjudged by wrong information at any time.

We don't use take efforts to buy newspaper or don't spend so much time to click a link and read the full news. We just scroll down in social media and read the headlines. By this we easily get misinformation and starting misjudge. So in our research based work we make a survey where the misinformation headline can be detected by applying machine learning model. We collect dataset in both cases true/ false. Comparing this two cases our model can give proper output about the news is true or fake through decision tree.

Table 2.3: Comparison of accuracy from different models

MODEL	ACCURACY
K-Nearest Neighbor (KNN)	68.7%
Support Vector Machine (SVM)	70.84%
Random Forest (RF)	76.37%

Depend on the accuracy results in the table 2.3 (comparison of accuracy) and also the attached graph in figure is applied by putting various algorithms on input and the accuracy o the output. In the table we can see that there are three models such as KNN, SVM, RF. KN has 68.7% accuracy, SVM has 70.84% accuracy and RF has 76.37% accuracy. Here we ca identify that Random Forest has far better accuracy than KNN, SVM. That's why we buil random forest model for our research based project.

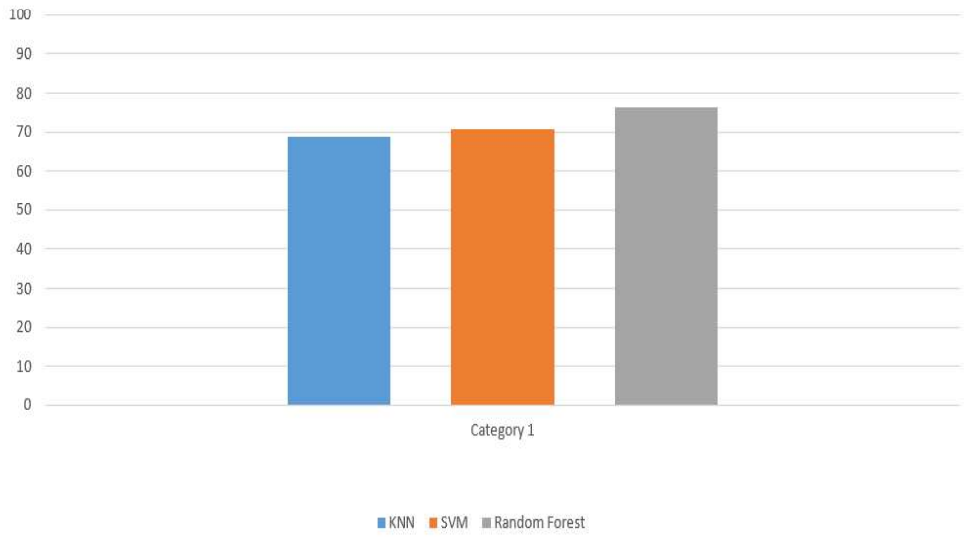

Figure : 2.3.1 Bar chart of algorithms' results comparison

CHAPTER 3

RESEARCH METHODOLOGY

3.1Generating keywords

Among the reputed portals Daily Prothom alo, Dhaka Channel, BD news Bangladesh Pratidin, Ittefaq, Daily KalerKantho were notable and Dhaka Channel Khbor24.com was used as a non-reputed news portal. The authenticity of the news were checked using Google scraping where we have to search by the news title then found the composition of the fakeness or realness of the news. Now the world is digital platform where everything is relying on the internet. We don't use take efforts to buy newspaper or don't spend so much time to click a link and read the full news. We just scroll down in social media and read the headlines. By this we easily get misinformation and starting misjudge. So in our research based work we make a survey where the misinformation headline can be detected by applying machine learning model. We collect dataset in both cases true/ false. Comparing this two cases our model can give proper output about the news is true or fake through decision tree.

At last we have also created an audience opinion where people can give their vote to identify whether a news is false or real.In this section, we discuss some important keywords which are related to our study. Here is some major keywords in our research. We use 1.5k dataset in our survey. Dataset are in excel file.we build the machine learning model python. Some specific keywords we identify such as Linguistic-base news, visual-base news,emotion base (anger, happiness, depression, frustration, hatred and so on emotions). We also identify the length of the news. Cause usually the length of mesh news is shorter than the length of original news. Mesh news descriptions is always differ from its misleading headlines.

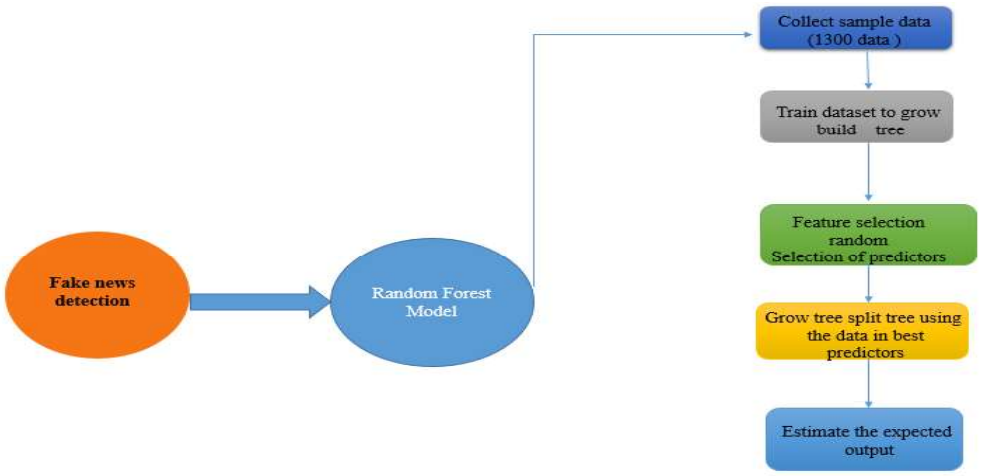

Figure 3.1 Overall working procedure

Random Forest model provides actuality reach out of the total dataset which actually gives th accuracy scores 72.12%. It bears less time compared to the other machine learning model. ensues the decision tree through splitting the trained dataset. In this research we used 110 Bangla dataset which consists of 550 fake news and 550 true news which we collected fror many online portals.

3.2 Data Collection

In this Research, we have deliberated dataset of both true and false to measure the performance of different methods. The characteristics of the dataset is mentioned here. We have collected our dataset from social news sites.Through Scrapping we collect news from from differ news portal like Prothom Alo, Prothom Alu, Dhaka Tribune,Dhaka kantha, Dhaka channel and so on. As our research based project is on bangla fake news detection, so we pluck all bangla news from news portal. It will remove missing data also which is assigned in NULL value.

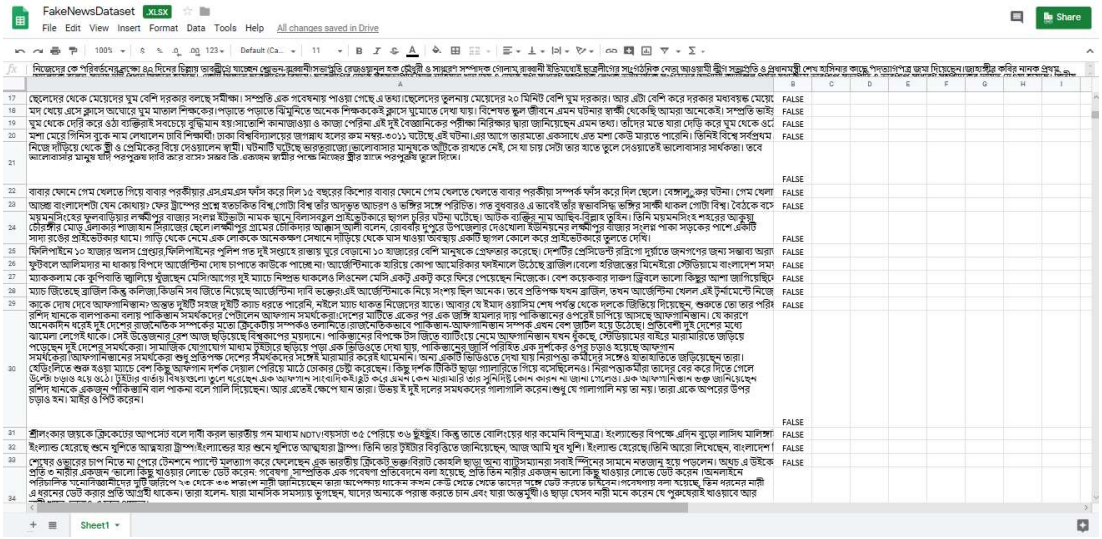

Figure 3.2.1 Data Collection

We collect data of portal news in this procedure. Now in the below we show a table where the demo of our data collection is shown. We use two columns for data such as one is for news description and other one is the specification of the news whereas it is true or false. As we are using classification model so we need to specify the input and output. From the input and output value the machine learning model can predict new input is false or true. In our dataset we have total fifteen hundred data.

Table 3.1: Data collection procedure

ভুল দিনে ঈদ পালন করে ১৬০ কোটি রিয়াল কাফফারা সৌদির!ভুল দিনে ঈদ পালন করে ১৬০ কোটি রিয়াল কাফফারা সৌদির!	FALSE
প্রধানমন্ত্রী শেখ হাসিনা চামড়াজাত পণ্য থেকে কাঙ্ক্ষিত রপ্তানি আয়ের লক্ষ্য অর্জনে আগামী পাঁচ বছর এ খাতে আর্থিক প্রণোদনা অব্যাহত রাখার ঘোষণা দিয়েছেন।	TRUE
বাংলাদেশ জিতে যাওয়ায় রাতে ম্যাককলামকে বাসায় ঢুকতে দিচ্ছে না তার আম্মু!	FALSE
বেসরকারী ভার্সিটি গুলোতে টিউশন ফি হিসেবে টাকার বিকল্প পেঁয়াজ নির্ধারণ	FALSE

Finally on this way, we collect both FALSE and TRUE news from web portal, facebook, online news portal for comparing between them and get the output by applying Random Forest Algorithm through decision tree in classification.

3.3 Dataset Preprocessing

As social sites are so unstructured in nature and contain a lot of unicode characters, we have to preprocess the raw form of the data.

Table 3.2: Preprocessing steps of a sample.

বেসরকারী ভার্সিটি গুলোতে টিউশন ফি হিসেবে টাকার বিকল্প পেঁয়াজ নির্ধারণ	A
“বেসরকারী’’ “ভার্সিটি” “ গুলোতে” ”টিউশন” “ফি” “হিসেবে” “টাকার” “ বিকল্প” “পেঁয়াজ” “নির্ধারণ”	B
“ভার্সিটি”:0 “ গুলোতে”:1 “পেঁয়াজ”:2 “বেসরকারী’’:3 ”টিউশন”:4 “ফি”:5 “টাকার”:6 “নির্ধারণ”:7 “ বিকল্প”:8 “হিসেবে”:9	C
3 0 1 4 5 9 6 8 2 7	D

Here,

A) Portion of texts from news

B) Tokenized version of the

C) Dictionary of the corpus

D) Representing the text in a single vector

Before going into the model, raw strings of news required some preprocessing. We should eliminated inessential IP and URL addresses from our texts. Next plod is to eliminate stop words. After that, we cleaned our Data frame by identifying the spelling of words. We split every text of Bangla news by white-space and remove some tokens from words.

3.4 Feature Extraction

The staging of machine learning models relies on a large deal on features design. Hence we have removed a vast range of features for behind those requires in other to the point work.

1. Lexical Features Extraction:
2. Sentiment Features Extraction
3. Pre-trained Word Embedding
4. n-gram Feature Extraction

Other features of news are two types

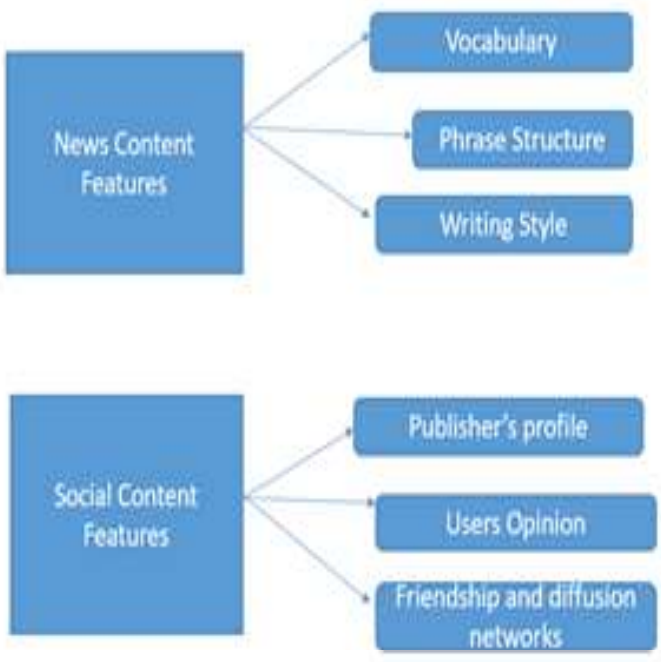

Figure 3 4. Feature extraction

3.5 Manual annotation

After preprocessing we got a total amount of data for manual annotation. We annotated the body in specific ways. One for the binary classification to determine the dataset. The total news story have been unsheathed from some Bangla reputed and non-reputed social online news portal. Among the reputed portals Daily Prothom alo, Bangladesh Pratidin, Ittefaq, Daily KalerKantho,Dhaka Tribbiun are mark-able and Dhaka Channel Khbor24.com was utilized as a non-reputed news portal. The authoritativeness of the news were restrained using Google scraping where we have to search by the news title then found the composition of the realness of the news. As we are using classification model so we need to specify the input and output. From the input and output value the machine learning model can predict new input is false or true. In our dataset we have total fifteen hundred data. At last we have also created an audience vote where people cast their vote to identify whether a news is fake or real.

CHAPTER 4

CLASSIFICATION AND MODEL STUDY

4.1 Classification

Between two types of supervised machine learning model (Classification and regression) , classification is one of them . Classification is a way more easy process of prediction for data analysis . Classification gives us the exact value which can be yes or no. More precisely it can give a result whether a prediction is true or false. It can be used for low range and high range of any dataset to predict the exact result. In our fake news detection we used both true and false classed from where these model can detect where the given data is true or false.

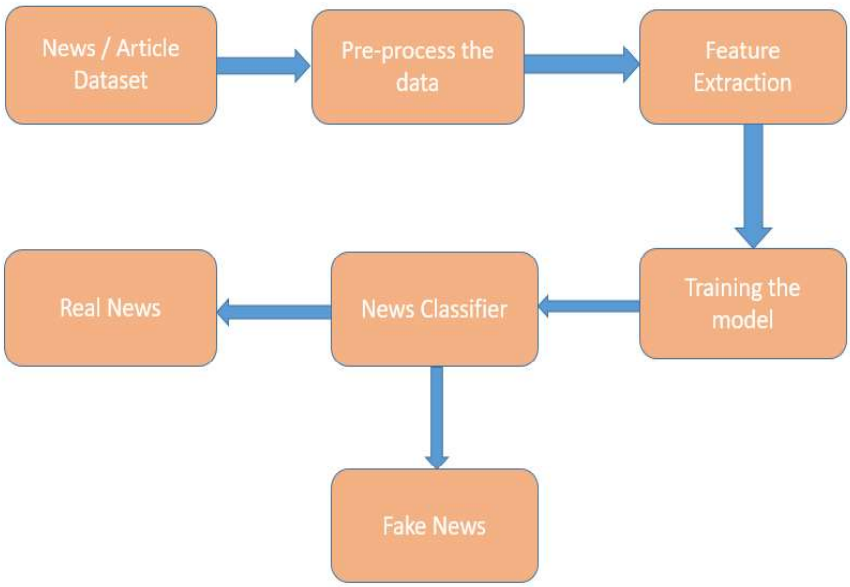

Figure : 4.1 Process of classification

4.2 Model Study

For fake news detection we used 3 models to compare the accuracy rate . Basically 3 model we used to find more accurate value which model can give us. We used Random forest model, KNN K- nearest neighbors) and SVM model (Support Vector Machine) . These 3 models give us 3 different accuracy whereas KNN (K- nearest neighbors) gives us 68.7% accuracy, SVM (Support Vector machine) gives us 70.84% accuracy and lastly Random Forest model gives us 76.37% accuracy. we discuss some important keywords which are related to our study. Here is some major keywords in our research. We use 1.5k dataset in our survey. Dataset are in excel file.we build the machine learning model python. Some specific keywords we identify such as Linguistic-base news, visual-base news,emotion base (anger, happiness, depression, frustration, hatred and so on emotions). We also identify the length of the news. Cause usually the length of mesh news is shorter than the length of original news. Here we can see that random forest model gives us more stable and accurate value so this model can be used to detect fake news. Depend on the accuracy results in the table 2.3 (comparison of accuracy) and in the table we can see that there are three models such as KNN, SVM, RF. KNN has 68.7% accuracy, SVM has 70.84% accuracy and RF has 76.37% accuracy. Here we can identify that Random Forest has far better accuracy than KNN, SVM. That’s why we build random forest model for our research based project.

4.2.1 KNN Model

KNN model is a part of supervised model where we use classification. It is a simple algorithm used in machine learning. Its working strategy depends on similarities between data which we inserted before. Mainly it estimates output from previous analysis. It evaluates the distance from new data to old data. Then it calculates the distance and go for the closest value and match the new data with the old data. In this way it gives the final decision value from comparing the distance between the old and new data using classification. We build this KNN algorithm for our research base work for detecting bangla fake news from online news portal. We have get good accuracy from this model but not more than random forest algorithm.

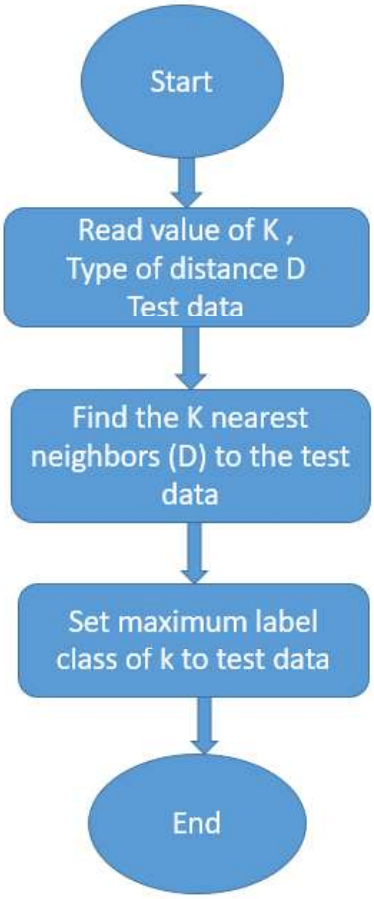

Figure : 4.2.1 KNN Algorithm's Flowchart

In this figure we can see the working steps of KNN algorithm. It reads the new data and the test it. For testing mainly it tests output from previous analysis. It evaluates the distance fror new data to old data. Then it calculates the distance and go for the closest value and match th new data with the old data. In this way it gives the final decision value. This is the full workin process of KNN algorithm.

4.2.2 SVM Model

Super vector model is a part of supervised model where we use classification. It is a simple algorithm used in machine learning. Its working strategy depends on similarities between data which we inserted before. It works on word vector which has hyper line. It adds a hyper line in the middle of dataset x axis and y axis in 45 degree. From calculate the distance of the line between the old and new data distance. It calculates the distance and go for the closest value and match the new data with the old data. In this way we get the final output of prediction.

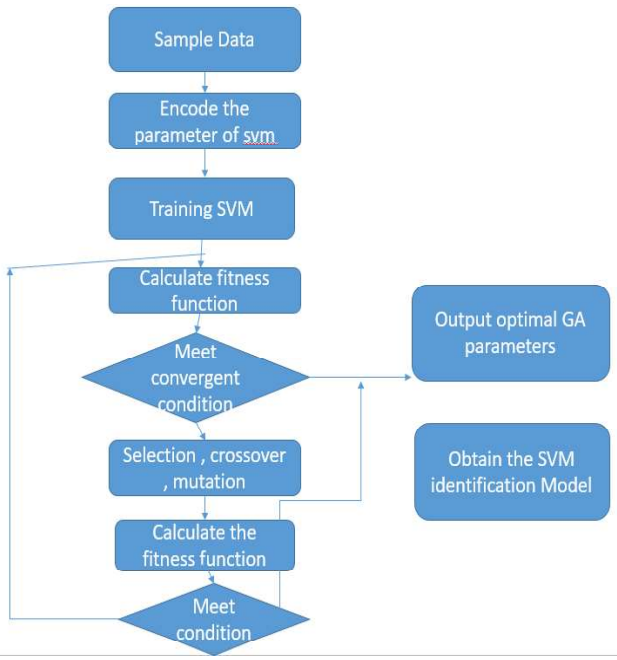

Figure : 4.2.2 SVM Model's Flowchart

4.2.3 Random Forest Model

Random forest algorithm is a very simple machine learning algorithm which is part of supervised model. It can be used for both in classification and regression. Random forest model is based on decision tree. From decision tree random forest indicates the best splits. It compares between branches and gives the best leaf tree from the parent trees. From calculate the best comparison of splits from the old and new data similarities. It goes for the closest value and match the new data with the old data. In this way we get the final output of prediction.

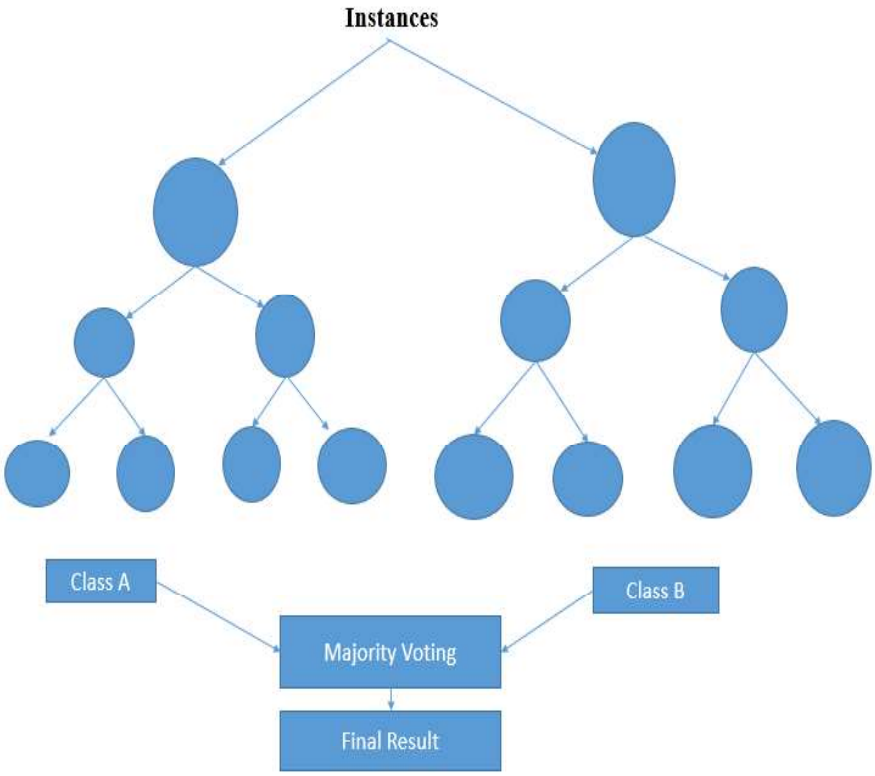

Figure : 4.2.3 Random Forest algorithm's flowchart

CHAPTER 5

EXPERIMENTAL RESULTS AND DISCUSSION

5.1 Results

The picture demonstrates us an insight of confusion matrix where we can see the results out of 1100 dataset we find out true positive 74, true negative 156 false negative 72 and false positive 26 which used train test split . We have evaluated our different methods using (recall, precision and accuracy) by splitting our dataset into 80:30 ratio where 80% data were used for training and 20% data for testing. The performance of the models.Random forest can gain high classification results through a classification ensemble with a set of decision trees that grow using randomly selected subspaces of data [8].

5.2 Descriptive Analysis

Here we can see the accuracy 76.37% Random forest model has shown the preeminent performance among all the models.

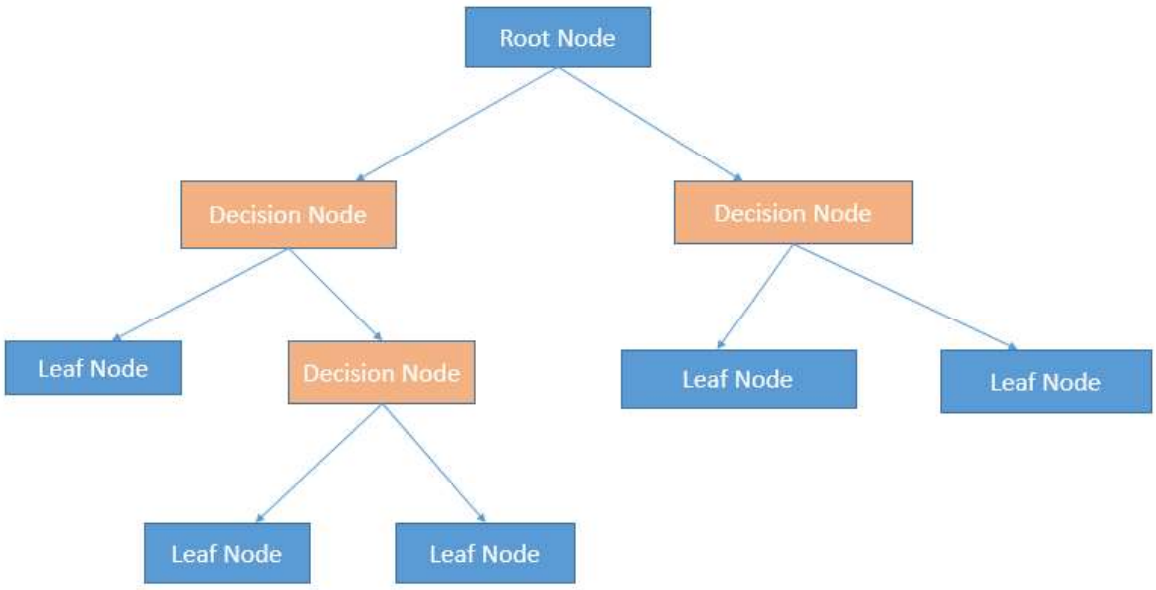

Figure 5.2 Working flow of decision tree

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

In our research we have combined such techniques which built a model for detecting whether the news is fake or true. In this digital era the news spreads faster than anything but due to deceptive news many unexpected situations occur. We basically used 3 models KNN, SVM and Random forest model which gives us 3 different accuracy. But the high accuracy we can gain by random forest model.The augmentation of Bangla fake news and its extension on social media has become a main anxiety due to its caliber to make demolishing dominance. Different machine learning intercourse have been endeavor to identify English fake news. But in our survey we build machine learning model called “Random Forest Machine Learning Model” through decision tree. We used binary classifier more specifically random forest model which gives us 76.37% accuracy. We also provided an output of confusion matrix to find out the insight of fake and true news which shows 60 , 165 , 86 and 17 respectively as true positives , true negatives , false positives and false negatives . Any user can easily find out the accuracy of the and also can be certain about a news whether the news is true or false. This model gives us best score of accuracy and also it is not time consuming. So Random forest model is the ideal model which can be used for fake news detection.

6.2 Future plan

Here we only implemented 3 models and from three different model we acquire 3 different value of accuracy and we can see that random forest model can give us best accuracy rate which is a machine learning model . We only implemented the model which shows the accuracy score but in future we want to build a website where any user can input the particular news from any social media, any online news portal or from any other sites to our website so that they can verify whether the news is true or false which will be very efficient to reduce the confusion through fake news. The website will have options for user to detect whether a news true or false. User can easily access the website and can copy paste the news and can find out if the news is true or false. This will be the first website because there is fake news detector but there is no Bangla Fake News detector. So this website can easily find a if a news is true or not and can reduce confusions.

References :

[1]Shodhgang ,available at , https://shodhganga.inflibnet.ac.in/simplesearch?query=acknowledgements.pdf last accessed on 06-10-2019 at 11:00 AM . .

[2] BBC , available at https://www.bbc.com/news/uk-36528256 , last accessed on 06-10-2019 at 11:25 AM .

[3] BBC , available , https://www.bbc.com/news/uk-36528256https:/www.webwise.ie/teachers/what-is-fake-news , accessed on 06-10-2019 at 12:00 PM .

[4] Buzzfeed , available at https://www.buzzfeed.com/craigsilverman/viralfake-election-news-outperformed-real-news-onfacebook?utm term=.nrg0WA1VP0#.gjJyKapW5y , accessed on 07-10-2019 at 12:00 PM .

[6] The New York Times , available at https://www.nytimes.com/2016/11/28/opinion/fakenews-and-the-internet-shell-game.html? r=0 , accessed on 07-10-2019 at 12:05 AM .

[6] Khan, Junaed & Khondaker, Md. Tawkat Islam & Iqbal, Anindya & Afroz, Sadia. (2019). A Benchmark Study on Machine Learning Methods for Fake News Detection.

[7] Data Flair, available at https://data-flair.training/blogs/advanced-python-project-detecting-fake-news/?fbclid=IwAR1CITEP6NfsSXkJxqZGS34WK4MVLzhcFay9UQ0U_hQVEReCkAS0_3RxzxQ , accessed on 07-10-2019 at 12:05 AM .

[8] S. Bharathidason, C. Jothi Venkataeswaran “Improving Classification Accuracy based on Random Forest Model with Uncorrelated High Performing Trees”, Journal, Volume, 06-09-2019.