Fake News Detection in Bengali Language using Transfer Learning Approach

Authors

S M Abu Rafi Sadat
Student ID: 170042082

Al Muhtasim Srijon
Student ID: 180042107

Hasnaine Ahmed Dihan
Student ID: 180042109

Supervisor

Md. Hamjajul Ashmafee
Assistant Professor,
Department of Computer Science and Engineering (CSE)
Islamic University of Technology (IUT)

A thesis submitted to the Department of CSE

in partial fulfillment of the requirements for the degree of

Bachelor of Science in Software Engineering

Academic Year: 2021-2022

Department of Computer Science and Engineering (CSE)

Islamic University of Technology (IUT)

A subsidiary organ of Organization of Islamic Cooperation (OIC)

Board Bazar, Gazipur-1704, Dhaka, Bangladesh

May, 2022

Declaration of Authorship

This is hereby declared that the work presented in this literature is the result of scrutinized experiments carried out by the authors under the supervision of Md. Hamjajul Ashmafee in the Department of Computer Science and Engineering (CSE), Islamic University of Technology (IUT), Gazipur, Dhaka, Bangladesh. In addition, neither this thesis nor any part of this has been submitted in any degree, diploma, or other certifications to this or any other institution. The guidelines of conduct have been acknowledged and respected by the authors, along with existing literature from their respective authors, which are mentioned at the closing chapter.

Authors:

S M Abu Rafi Sadat	Al Muhtasim Srijon	Hasnaine Ahmed Dihan
Student ID: 170042082	Student ID: 180042107	Student ID: 180042109

Supervisor:

Md. Hamjajul Ashmafee
Assistant Professor
Department of Computer Science and Engineering (CSE)
Islamic University of Technology (IUT)

Dedicated to all the victims of violence instigated by Fake News

Acknowledgement

First and foremost, We would like to express our heartfelt gratitude to our supervisor, MD. Hamjajul Ashmafee, for his unwavering support and encouragement throughout our research. He was always willing to assist us anytime we needed it. Then we would like thank the Network and Data Analysis research group (NDAG) for their unconditional support and suggestions that only lead us to a better path. Also, thanks to all our well wishers, without their support this feat would have been unachievalbe. Finally, we would like to acknowledge and appreciate the minds behind all the researches we've utilized to accomplish this.

List of Abbreviations

BERT	Bidirectional Encoder Representations from Transformers
RoBERTa	Robustly optimized BERT approach
mBERT	Multilingual BERT
XLM	Cross-lingual Language Model
TLM	Translation Language Modeling
MLM	Masked Language Modeling
DNN	Deep Neural Network
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short-term Memory
LR	Logistic Regression
SVM	Support Vector Machine
KNN	K-Nearest Neighbor
MNB	Multinomial Naive Bayes
ReLU	Rectified Linear Unit

Abstract

The rapid proliferation of fake news poses grave consequences for civil discourse, political environments, and social cohesion. From public elections to mob violence, fake news has been leveraged to achieve personal and political gain. The influence of fake news in this era of information is undeniable. Misinformation can cause mass disruption, and we need a way to stop that from happening. This study experiments with the widely used multilingual pre-trained transformers *XLM-Roberta* and *Multilingual-BERT*, along with *Bangla-BERT*. It also explores the impact of stemming in Bengali and demonstrates the effectiveness of combining deep neural network *(DNN)* layers with pre-trained transformers. A major setback faced in this research was the lack of a well-balanced dataset, which led to inconsistent performance from the models. We undersampled two datasets from the original one[1], one with the ratio *fake:authentic = 1:1*, another with *fake:authentic = 1:3*. We were able to achieve 0.95 precision and a 0.92 F1 score in a heavily undersampled but well-balanced dataset derived from the original one. *XLM-Roberta* and *Bangla-BERT* based models achieved recall scores of 0.94 and 0.93 respectively on the dataset where ratio of *fake:authentic* is *1:1*. Overall, the models trained on the *1:1* dataset delivered consistent scores across all the metrics, which emphasizes the importance of collecting more fake news data for future research. The best model, based on *Bangla-BERT*, achieved an accuracy of 96.2% which sets a new benchmark accuracy for transformer based models in fake news detection in Bengali.

Keywords: *Transformers, BERT, Roberta, DNN, Multilingual, Bangla, Fake News*

Contents

Figures

Tables

Chapter 1

Introduction

1.1 Motivation

The dissemination of fake news has become a major social issue with potentially grave consequences for civil discourse, political environments, and social cohesion. False and erroneous material deliberately manufactured to lead others to believe it is true is referred to as fake news. The phrase ”fake news” became well-known during the 2016 U.S. election campaign when politicians, most notably Donald Trump, began using it as a catchphrase to deflect criticism [4]. Fake news has been used for an array of reasons, including political opinion-shaping or financial gain. Fake news can persuade people to support particular ideas or goals by tampering with their perceptions and thoughts. It is crucial to understand that fake news frequently combines truths and lies, making it difficult to distinguish between the two. As information circulates rapidly through social media and other technology, fake news spreads swiftly and reaches a broad demographic. Online news outlets frequently serve as hubs for propagating erroneous and misleading information, deluging social media with made-up stories. Individuals or groups fabricating distorted information are primarily motivated by profit rather than the issues they construct. They can draw visitors to their websites using tools like Facebook, which enables them to make revenue through advertising. These news pieces have been designed to be used as clickbait to lure in readers and haul in profits for the authors. It might be difficult to tell fraudulent news sources apart from reliable ones.

The detrimental effects of fake news go beyond the internet sphere; they frequently have grave repercussions in the offline sphere as well. Unsettlingly violent incidents occurred in several nations for transmitting misinformation. In Mexico, a mob burned two guys alive without checking the veracity of the information after a bogus news article about a kidnapping went viral and circulated on WhatsApp [5]. Similar occurrences happened in Sri Lanka, India, and Myanmar [6] when false information spread on Facebook and WhatsApp sparked deadly violence.

Fake news has a profound psychological effect that fosters hatred and widens social gaps. In a nationwide survey, the Management and Resources Development Initiative (MRDI) in Bangladesh found that rural areas (66%), urban areas (62.3%), and metropolitan areas (52.5%) all experienced high rates of fake news [7]. Tragically, half of the population cannot tell fact from opinion when looking for information online. The lack of adequate digital literacy among a significant portion of internet users in Bangladesh further exacerbates the issue. During the COVID-19 pandemic, Bangladeshi netizens sought health-related information on social media, leading to the spread of misinformation. One such example involved a religious claim suggesting that protection against COVID-19 could be achieved by regularly consuming Thankuni leaves (Indian pennywort) while saying "Bismillah" (in the name of Allah) [8]. Additionally, a falsified report that claimed human sacrifices were essential to the construction of the Padma Bridge circulated on Facebook and WhatsApp, leading to mob violence and the murder of alleged kidnappers [9]. A mob of over 25,000 people destroyed temples and homes in Ramu in 2012 due to a misleading report of the mutilation of a Quran that was falsely attributed to a Buddhist person via a Facebook account [10]. The prevalence of Facebook profiles and groups soliciting support for fictitious people or organizations has also grown. This thesis suggests the creation of a model leveraging transformers to increase the precision of fake news identification in resource-constrained settings to stop the spread of false information.

1.2 Problem Statement

The proliferation of fake news has resulted in detrimental consequences globally, including instances of lynching and exacerbating the impact of the COVID-19 pandemic. According to a countrywide survey carried out by the Management and Resources Development Initiative (MRDI) in partnership with Unicef, the problem of false news in Bangladesh has grown to frightening proportions. The survey revealed that 63.6 percent of individuals in Bangladesh encountered news on social media or online platforms, initially perceiving it as authentic only to discover it was falsified [7]. Additionally, nearly two-thirds of the population either rarely or sporadically verify the credibility of news sources [7]. This situation necessitates urgent measures to mitigate the spread of fake news and its associated consequences in Bangladesh. While fact-checking websites such as BD FactCheck [11], Jachai [12], and Rumor Scanner [13] exist in Bangladesh, they heavily rely on user-based reporting, which may introduce biases and skew the contextual integrity of reported incidents. Therefore, it is crucial to create an automated system that guarantees correctness and integrity. Furthermore, the Bengali language cannot be processed by the existing fake news detection systems because they were created primarily to evaluate English letters and words. As a result, a comprehensive solution that can accurately analyze Bengali letters and words is essential.

Currently, individuals who receive news through messaging applications such as WhatsApp, Imo, or Viber lack the means to discern the authenticity of the information they encounter. Consequently, a system that tackles these drawbacks and offers a reliable method to identify and stop fake news in Bengali is crucial. This comparative study aims to address the aforementioned challenges by investigating and evaluating various approaches for fake news detection in the under-resourced Bengali language. By examining existing techniques and adapting them to accommodate the linguistic and socio-cultural characteristics specific to Bengali, this research seeks to identify the most effective methods for detecting fake news in this language. The findings of this study will not only help Bengali fake news detection develop, but they will also offer important insights for tackling fake news issues in other under-resourced languages. With the findings at hand, decision-makers, social media platforms, and users may promote media literacy while halting the spread of false information inside under-resourced linguistic communities.

1.3 Research Challenges

This research endeavor entails several inherent challenges that need to be addressed in order to effectively detect fake news in the under-resourced Bengali language. The following challenges are anticipated in the course of this study:

- One of the primary challenges in developing a robust fake news detection system for Bengali is the scarcity of labeled and well balanced datasets. Fake news datasets specifically curated for the Bengali language are scarce, hindering the training and evaluation of machine learning models. Constructing a comprehensive, well balanced, and representative dataset that encompasses various types of fake news scenarios in Bengali poses a significant challenge. We had to address the hugely imbalanced proportion of our collected dataset forcing us to undersample to an extent that made the utilization of the dataset nearly futile.

- Detecting fake news in Bengali is challenging due to the language's unique linguistic characteristics, including complex grammar, rich morphology, and contextual nuances. Existing techniques designed for simpler languages may not be directly applicable. Developing language-specific approaches to capture the subtleties and context-specific features of fake news in Bengali is a major challenge.

1.4 Thesis Objective

The main objectives of this thesis can be described below.

- To analyze and evaluate transformer based fake news detection approaches and methodologies in the context of under-resourced Bengali language.

- To study the effectiveness of the developed fake news detection models and techniques through rigorous experimentation and comparison with state-of-the-art methods.

1.5 Key Contribution

Here are the key contributions of our work.

- This thesis conducts a study of fake news detection techniques and methodologies that could be potentially superior to the existing solutions in the context of under-resourced Bengali language. By analyzing and evaluating these approaches, it provides insights into their strengths, limitations, and applicability to Bengali.
- This thesis validates and compare transformer based approaches and strategies for detecting and mitigating the spread of fake news in under-resourced Bengali language communities.
- This empirical evaluation provides insights into the strengths and limitations of the proposed approaches and benchmarks their performance against existing methods.

1.6 Thesis Structure

This thesis is divided into several chapters and each chapters to several sections and subsections for a comprehensive view. This chapter introduces the thesis through problem statement, key contribution, research challenges, and the thesis objectives. The following chapters dive deep into the research activities and document them. Chapter 2 includes literature review to provide an idea of what has been going on in the domain of fake news detection and their potential impact on further research. Chapter 3 describes the dataset used in this thesis. Chapter 4 documents the experimentation and methodology. Chapter 5 analyses the result. Finally, chapter 6 draws a conclusion and provides potential future directions.

Chapter 2

Literature Review

2.1 Machine learning, deep learning and linguistic analysis

Fake news poses a significant challenge to the credibility of online information. To combat this issue, researchers have increasingly turned to machine learning algorithms for automated fake news detection. However, the majority of existing studies have focused on English-language content, leaving a significant gap in research concerning other languages. An overview of the current research on fake news detection and associated methods in the context of the Bengali language is provided in this literature review.

For automatic fake news detection, many machine learning and deep learning algorithms have been used. Techniques like feature engineering, text classification, and neural network models have all been investigated [14]. To identify between fake and real news stories, these techniques frequently use linguistic, semantic, and contextual data. Linguistic analysis plays a crucial role in identifying fake news. Researchers have leveraged linguistic cues, such as sentiment analysis, syntactic patterns, and stylistic features, to detect the manipulation and deception in news articles [15]. Understanding the linguistic characteristics specific to the Bengali language is essential for effective detection. Linguistic cues and features play a crucial role in identifying fake news. Researchers have utilized sentiment analysis, syntactic patterns, and lexical features specific to the Bengali language to detect deceptive news articles [16]. These linguistic features capture the manipulation of information and stylistic variations in fake news content.

2.2 Cross-Lingual Transfer Learning and Machine Translation Techniques

Cross-lingual transfer learning has been investigated as a method to address the scarcity of labeled data in languages other than English. Techniques like pre-training on large English datasets and fine-tuning on smaller Bengali datasets have shown promising results in cross-lingual fake news detection [17]. Another approach to multilingual fake news identification is to translate Bengali news pieces into a language containing labeled data, such as English, and then apply existing false news detection models to the translated content. To bridge the language divide and enable cross-lingual detection, machine translation techniques have been investigated [18].

2.3 Comparative Analysis of Machine Learning Algorithms

Anika Anjum et al. have found that, among the traditional classifier algorithms such as *Support Vector Machine, K-Nearest Neighbors, Naive Bayes, Random Forest, Decision Tree, Random Forest* gave the best accuracy of 82% in detecting fake news in Bangla [19]. Another research paper has implemented and compared several popular machine learning algorithms, including *Naive Bayes, Support Vector Machines, Random Forests, and Neural Networks.*to detect fake news in Bangla. Based on their experiment they proposed a model based on *Gaussian Naive Bayes* algorithm which got 87% accuracy [20]. The attributes were selected by using an *Extra Tree Classifier* and a text feature depending on *TF-IDF*. Although *Extra Tree Classifier* is used as a classifier generally, it was used as a feature selection technique to achieve better accuracy in detecting false headlines in the Bangla language.

2.4 Bangla Newspaper Scraper and Random Forest Classifier Model

Farzana Islam et al. have also made a Bangla newspaper scraper and a web URL checker which can be used to classify fake news in Bangla [21]. They also proposed a model with *Random Forest Classifier* which showed a maximum of 85% accuracy. It works best with the combination of news body and news headline. At least 2-3% of improvement has be shown by using both the body and headline of the news.

2.5 Deep Learning and BERT-based Models

With the success of deep learning in various *NLP* tasks, researchers have explored the application of neural network architectures for fake news detection. Models such as *Convolutional Neural Networks (CNN)* and *Recurrent Neural Networks (RNN)* have shown promising results in capturing semantic and contextual information for distinguishing between fake and genuine news [22]. Kaliyar et al. have proposed a novel model that can classify a English news as authentic or fake which obtained an accuracy of 98.90% [23]. Their model was based on a *BERT*-based deep learning approach by combining different parallel blocks of the single-layer *CNN*s with the *Bidirectional Encoder Representations from Transformers (BERT)*. The suggested method improves outcomes by 4% when compared to baseline methods, and it shows promise for the identification of fake news.

2.6 Performance of ML and Transformer Models

Incorporating different deep neural networks can achieve a higher accuracy in the detection of Bengali fake news. Multiple machine Learning (*LR, SVM, KNN, MNB, Adaboost,* and *DT*), deep Neural Networks (*LSTM, BiLSTM, CNN, LSTM-CNN, BiLSTM-CNN*), and transformer (*Bangla-BERT, m-BERT*) models were experimented on 4678 distinct news. The most accurate models, with accuracy rates of 95.9%, 95.5%, and 95.3%, respectively, are *CNN, CNN-LSTM,* and *BiLSTM* [24]. These models achieved higher accuracy than *BanglaBERT* and *multilingual BERT*. Another study shows that, deep learning model such as *RNN* performs better than *Bangla BERT* in detecting fake news in Bengali. *RNN* demonstrated the greatest performance among deep learning models, with 96.55% accuracy with a f1 score of 0.96. The pre-trained Bangla BERT model has an accuracy of 93.35% and an F1-Score of 0.96 [25]. The accuracy of *Bangla BERT*'s headline-based classification of crime news was compared to that of 8 other machine learning and language classifier models. Over 6293 training samples and 1574 testing samples, the *Bangla BERT* model achieved the accuracy of 90.15% [26].

Traditional machine learning methods have been used in earlier research to categorize false news in the Bengali language. Additionally, efforts have been made to use cross-lingual transfer learning to identify false information in Bengali. These methods, meanwhile, have not produced accuracy levels that are sufficient. Transformer model adoption has recently come to light as a more efficient method, producing greater accuracy rates. The accuracy of our suggested methodology, which combines Transformer and DNN layers, is substantially greater than that of earlier methods, coming in at 96.2%.

Chapter 3

Dataset

The primary challenge in detecting fake news in Bengali has been collecting the input data. There has not been sufficient research conducted on the Bengali language in this domain. In contrast, a sizable amount of data is available for widely used languages like English, French, and Spanish, as false news detection algorithms are currently the topic of greater research. Finding a useful dataset for a language with few resources like Bengali was fairly difficult. The only dataset we have found on Bengali is from a research work published in 2020 titled- *'BanFakeNews: A Dataset for Detecting Fake News in Bangla'* [1]. It had about 50k collection of labeled content along with the headline, source, date, category, domain, relation, and article id.

3.1 Description

- **Article ID:** ID of the news
- **Domain:** Publisher's name
- **Date:** Published date
- **Category:** Type of News
- **Source:** Origin source of the news.
- **Relation:** Related or Unrelated
- **Headline:** Headline of the news

- **Content:** Body of the news
- **Label:** 0 → fake, 1 → Authentic

Here's how the dataset looks-

articleID	domain	date	category	source	relation	headline	content	label
3759	bangla.bdnews24.com	2018-09-21 01:01:40	Finance	প্রতিষ্ঠানটির সচিব	Related	বিদ্যুৎ বিল সংগ্রহে ওজিপাডিকোর সঙ্গে চুক্তি বি...	সম্প্রতি ওজিপাডিকোর খুলনার প্রধান কার্যালয়ে এক...	1.0
470	earki.com	২০:৪৫, জানুয়ারি ২৪, ২০১৯	Miscellaneous	Reporter	Unrelated	ফ্ল্যাট পাওয়ার জন্য যেভাবে মাত্র ৫টি ধাপে সাংব...	আবাসন সঙ্কট নিরসনে সাংবাদিকদের ফ্ল্যাট দেওয়ার ...	0.0
7180	bangla.thereport24.com	2018-09-29 15:17:42	Crime	খিলক্ষেত থানার ভারপ্রাপ্ত কর্মকর্তা	Related	রাজধানীতে ফ্লাইওভারে উল্টে গেলো কাভার্ড ভ্যান	রাজধানীর কুড়িল ফ্লাইওভারে একটি কাভার্ড ভ্যান উ...	1.0
494	motikontho.wordpress.com	2012-10-30T11:33:15+00:00	Miscellaneous	Reporter	Unrelated	সমঝোতায় এলেন শাজাহান ও তৈমুর | দৈনিক মতিকণ্ঠ	নিজস্ব মতিবেদকপরস্পর সমঝোতায় এসেছেন আওয়ামী ল...	0.0
3785	kalerkantho.com	2018-09-22 11:49:41	National	ঢাকা মেডিকেল কলেজ হাসপাতাল পুলিশ ক্যাম্পের এসআই	Related	শ্যামপুরে দুর্ঘটনায় দুই শ্রমিক নিহত	রাজধানীর শ্যামপুরে বড়ইতলা এলাকায় দুর্ঘটনায় দুই...	1.0

Figure 3.1: Snapshot of the Dataset

3.2 Undersampling

Against nearly 48k authentic news, there was just about 1k fake news in the dataset, skewing the it toward authentic news. Although people come across authentic news more than fake news, this ratio of distribution in the dataset is beyond that rationality. A UCLA report shows that only 62% of the data we come across on internet are unreliable [27]. So to utilize the dataset we had to undersample it. We experimented with two ratio of distribution-

- 1:3 where, Fake : Real = 1300 : 4000
- 1:1 where, Fake : Real = 1300 : 1300

3.3 Pre-Processing

3.3.1 Data Cleaning

To maximize the use of our undersampled data, we got rid of the elements that don't contribute much to the understanding the meaning of a sentence and its context. These would be-

- Extra white spaces.

- Hyperlinks and URLs.
- Punctuation.
- Stopwords.

We have used the '*bnlp_toolkit*' [28] for stopwords' list.

Figure 3.2: Stopwords in Bengali

We have selected the text information from the '*content*' column as our data for training and testing.

Text from '*content*' before pre-processing

Table 3.1: Unprocessed Text

content	label
...চলতি বছরের এপ্রিলে উত্তর কোরীয় নেতা কিম জং উন ও দক্ষিণ কোরীয় প্রেসিডেন্ট মুন জায়ে ইন পানমুনজামে এক ঐতিহাসিক বৈঠকে মিলিত হন। ওই বৈঠকেই পরবর্তী একটি বৈঠকে মিলিত হওয়ার ব্যাপারে সম্মত হন দুই নেতা।...	1

Text from '*content*' after pre-processing

Table 3.2: Processed Text

content	label
...চলতি বছর এপ্রিল কোরী নেতা কিম জং উন দক্ষিণ কোরী প্রেসিডেন্ট মুন জা ইন পানমুনজাম এক ঐতিহাসিক বৈঠ মিলিত বৈঠকে পরবর্তী বৈঠ মিলিত হওয় সম্মত নেতা...	1

You can see that words like 'হন' and 'ও' are removed from the text along with all the punctuation. Words are also stemmed to smaller tokens using Bengali stemmer from the '*bnltk*' library [29]. This would help the vocabulary of the dataset from exploding.

3.4 Final Selection and Split:

Our selected dataset [1] for the experiment has the processed content and the label column. As we will analyze and compare to find out the best-performing model, we need to ensure that the dataset for training and testing is consistent across the models. So we initially split the dataset into the ratio of approximately 75 : 25 for training and testing. So the two types of the ratio of real and fake news we have selected previously in (section 3.2) will be split in the following way-

- For the ratio *fake : authentic = 1 : 1 -> 1300 : 1300*, we first separated 1000 fake and 1000 authentic data and combined them to make a training set. Then we combined the rest of the 600 (300 in each class) to create the test set. So **Train : Test = 2000 : 600**
- And for the ratio *fake : authentic = 1 : 3 -> 4000 : 1300*, we first separated 1000 fake and 3000 authentic data to combine them for training. After that, like in the previous case, we combined 300 fake and 1000 authentic data that were left for testing. So **Train : Test = 4000 : 1300**

This pre-split dataset is going to be used in across all the experiments to ensure minimum bias and maximum comparabilty.

Chapter 4

Methodology and Experiment

4.1 Transformers

Transformers, a highly effective class of neural network models, have revolutionized tasks involving natural language processing. In contrast to conventional recurrent neural networks (RNNs), transformers use self-attention mechanisms to identify contextual associations between words in a phrase. This allows transformers to effectively handle long-range dependencies and capture global information during language understanding. With their parallelizable and scalable architecture, transformers have delivered superior performance in various NLP tasks, including machine translation, sentiment analysis, and question-answering. In the context of fake news detection, transformers offer the potential to capture complex linguistic patterns and semantic nuances, making them essential for the accurate and robust identification of misinformation in textual data [30].

We have chosen to experiment with variations of *mBERT* [31] [32] and *XLM-Roberta* [2], which were trained on huge corpora of cross-lingual data. Another one is *Bangla-BERT* [33], which was specifically trained on Bengali corpora to use as pre-trained transformers. All of these models are bidirectional in the sense that they capture the context of a sentence by processing it from both left to right and right to left. We'll discuss their pre-training and characteristics briefly.

4.1.1 Pre-Training and Characteristics of the Transformers

As mentioned earlier, pre-trained transformer models are trained on large corpora of lingual data. Which allows them to effectively understand the linguistic intricacies and complexities. The techniques involved in the pre-training phase are:

Masked Language Modeling (MLM): Both *mBERT* and *XLM-Roberta* variations undergoes a pretraining phase in which it draws up knowledge from an enormous amount of multilingual text data. Whereas *Bangla-BERT* was trained with only Bengali corpora. A piece of the input text is randomly masked during MLM, and the model is trained to anticipate the original mask words depending on the context. Through this procedure, the model acquires an enhanced understanding of the associations between words and strengthens its capacity to fill in the blanks. While *BERT* masks approximately 15% of the input tokens, *RoBERTa* adopts a different masking strategy. *RoBERTa* employs dynamic masking, where different training instances receive different mask patterns. This approach avoids the inconsistencies caused by static masking used in *BERT*.

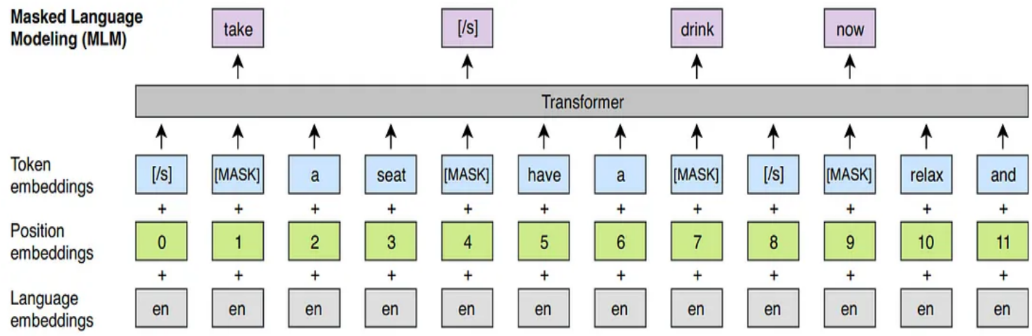


Figure 4.1: MLM from *XLM-RoBERTa* paper [2]

Next Sentence Prediction: In addition to MLM, *BERT* models also utilize the next sentence prediction task. This task involves training the model to determine whether two input sentences are consecutive in the original text or not. By learning to predict the continuity of sentences, the model gains the ability to capture contextual dependencies between different parts of a document or article.

Translation Language Modeling (TLM): *XLM-RoBERTa* uses TLM in addition to MLM, where it is trained to anticipate the original version of a sentence given its translated counterpart in another language. This challenge aids in capturing semantic and syntactic similarities between languages and encourages the model to learn cross-lingual representations.

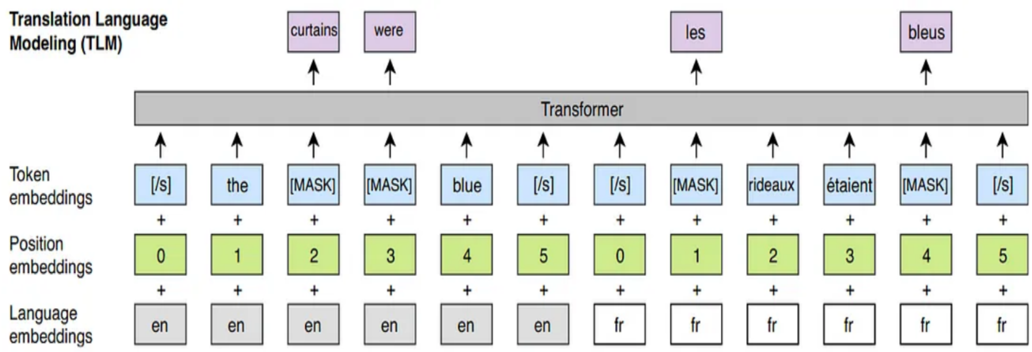


Figure 4.2: TLM from *XLM-RoBERTa* Paper [2]

Transformer Layers and Contextualized Word Representations: Both *BERT* [34] and *RoBERTa* [35] variations consist of multiple stacked transformer layers. Each layer utilizes a self-attention mechanism that allows the model to attend to different parts of the input sequence and capture contextual dependencies between words. This attention mechanism helps the model to extract informative representations from the input text, resulting in contextualized word representations that encode both the meaning of individual words and their relationships with the surrounding context.

4.1.2 Transfer Learning and Fine Tuning

Utilizing transformers depends on two strategies-

Transfer Learning: The pretraining phase of a transformer equips the model with a strong understanding of language and, in the case of multilingual ones, the ability to capture cross-lingual patterns. This knowledge could be further valuable for any other specific task if subjected to further training, as it enables the model to leverage contextual information and linguistic nuances.

Fine-tuning: After pre-training, transformers are fine-tuned on task-specific labeled data for any classification task. Fine-tuning involves updating the model's parameters using a smaller, domain-specific dataset. By training on this specific dataset, transformers learn to understand the differences between the classes. During fine-tuning, the classification layer is typically added on top of a transformer's pre-trained layers, allowing the model to make predictions based on the contextualized representations generated by the transformer.

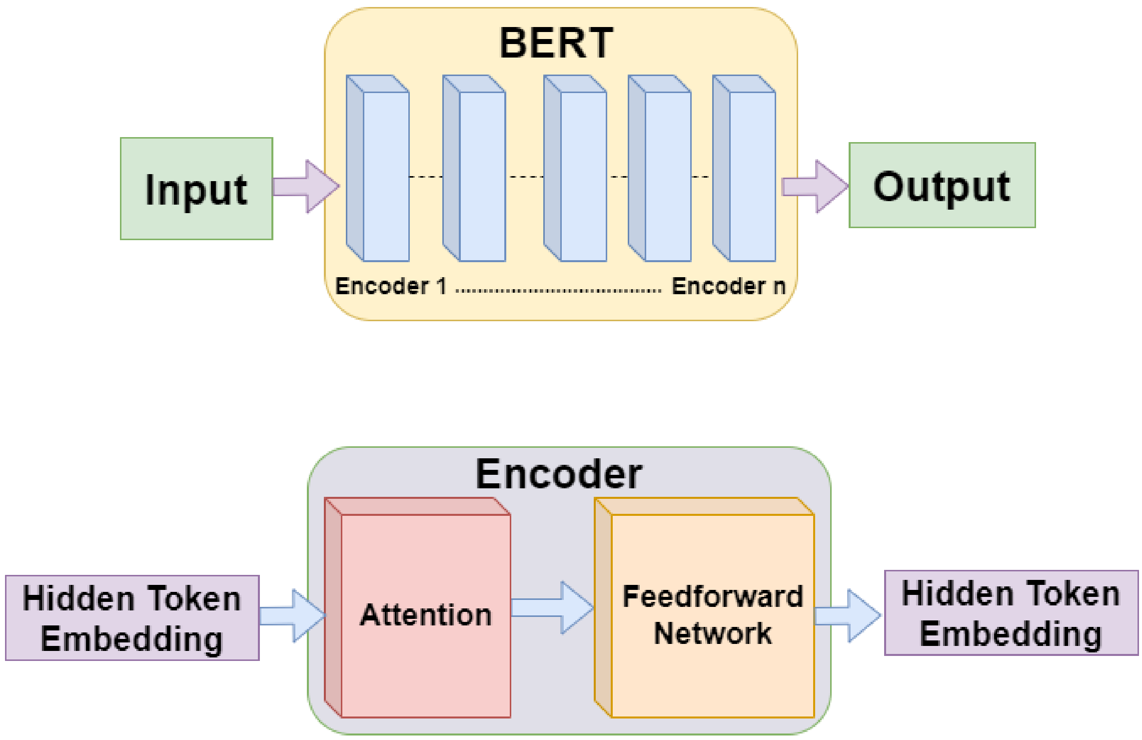


Figure 4.3: General architecture of the BERT

The general architecture of the *BERT* [34] or *Roberta* [35] model is illustrated in the figure 4.3

4.2 Training Steps

4.2.1 Setup

Selected Transformers:

List of transformer models chosen for the experiment and represented as-
Name of the model: corresponding pre-trained model in the *transformer* library [30]

- Bangla-BERT: *sagorsarker/bangla-bert-base* [33]
- mBERT: *bert-multilingual-base* (both cased [31] and uncased [32])
- XLM-RoBERTa: *xlm-roberta-base* [2]

Hyper Parameters:

After training for a certain range of epochs the models tend to hit plateau. Also batch size of more than 8 starts to clog the memory. Use of more than one *DNN* layer makes it easier for the model to memorize seen data resulting into overfitting of the model. After performing number of experiments we have found that the models train well on these set parameters. Also computing resource constraint factored into the choice of a few of the parameters.

- Size of added the DNN layer on top of pre-trained encoders → 1024
- ADAM optimizer with Learning Rate → 10^{-5}
- Loss function → Binary cross-entropy
- Train : Validation split (*of the 75% from Train:Test split*) → 80:20
- Batch Size → 8
- Tokens Per Sequence → 512
- Number of Epoch → within 10 to 15 (inclusive)

4.2.2 Feature Extraction

We have used pre-trained tokenizers of the corresponding pre-trained trasformers to generate meaningful embeddings that captures the context and also is optimally understandable to the corresponding transformer. Both BERT and Roberta uses *wordpiece* tokenizer [36] which helps to capture out of vocabulary words. The tokenizers take pre-processed text as input and return two fixed and equal sized vectors, *input_ids* and *attention_mask*, to feed the network for training. In our case the fixed size is 512 as we have selected that to be the maximum length of a sequence. Anything below that will be padded with special '*[PAD]*' tokens upto 512 and anything more than that will be truncated. The reason behind is that, texts in our dataset either contain less than 512 tokens, or tokens after that are highly unlikely to add new information and varying context to the input.

input_ids is a vector that represent tokens with their unique ids that also captures their relationship in the context of the language as well. Whereas *attention_mask* indicates which tokens to focus on and which ones to ignore for the transformer model. Tokens with mask '**0**' means those tokens have nothing to contribute to semantic understanding of the text.

So for the example given in table 3.2 the generated embedding would look like the following if maximum length of the sequence is set to 50 for that length is good enough as an example to show how this generation of embedding works-

...চলতি বছর এপ্রিল কোরী নেতা কিম জং উন দক্ষিণ কোরী প্রেসিডেন্ট মুন জা ইন পানমুনজাম এক ঐতিহাসিক বৈঠ মিলিত বৈঠকে পরবর্তী বৈঠ মিলিত হওয় সম্মত নেতা...

Figure 4.4: Portion of a processed input text

```
[[   101  3317  2091 44836 41585  2285 20039  2208  6402  7911 17035 10706
  77260  2454  2772  2285 20039  2193 39459  7052  4804  3244 11833  3710
   2490 10951 20708  2066  3249 14877 19519  6273  3810 37921  8833  2765
  14877 19519  6273 13669  9294  7505 12262  2208   102     0     0     0
      0     0]], shape=(1, 50), dtype=int32)
```

Figure 4.5: Generated input_ids for the text in fig 4.4

```
[[1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
  1 1 1 1 1 1 1 1 1 0 0 0 0 0]], shape=(1, 50), dtype=int32)
```

Figure 4.6: Generated attention_mask for the text in fig 4.4

You can see that the last few tokens have '0' as their id because the sequence length is less than the maximum length 50. So the extra tokens get padded with '*[PAD]*' tokens and the id for that is '0'. Similarly the *attention_mask* has '0' in those places to indicate that those tokens are to be ignored by the model during training.

4.2.3 Transformer Architecture

In this step the embeddings generated from the previous step will be fed to the designed transformer based network. The pattern we followed is to add a DNN layer of size 1024 on top of the pre-trained transformer and make the whole network trainable. As this is a binary classification task, we added a classification DNN layer of size 2.

Embedding generated by the selected pre-trained trasformers is a vector of size 768 [34] [35]. Then that is fed to the DNN layer to generate a vector of size 1024 as mentioned in section 4.2.1. This layer uses the ReLU activation function to generate the output as ReLU helps speeding up the training process. Finally that vector is fed to the classification layer and produced a binary distribution which is followed by a Softmax activation function to generate the final output. The figure below perfectly illustrates the architecture we used in the experiment.

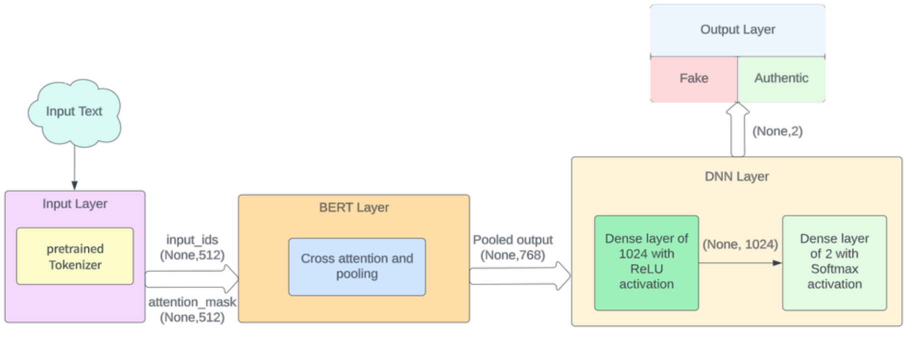


Figure 4.7: Architecture Used

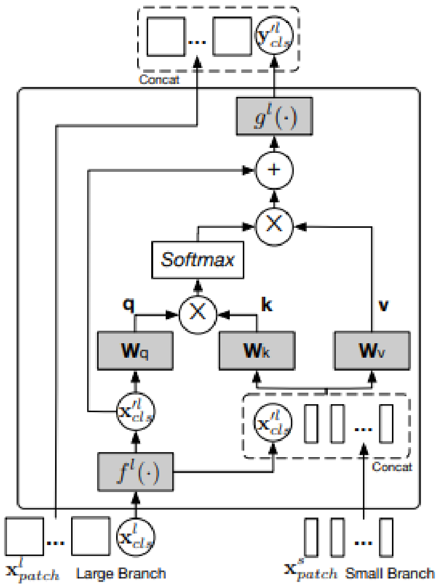


Figure 4.8: Cross Attention (source: CrossViT Paper [3])

4.3 Evaluation Metrics

The metrics we used to evaluate our models are- *Precision, Recall*, *F1-score*, and *accuracy*. They all use certain direct measurements to derive them. The measurements can be attained from the generated confusion matrices.

4.3.1 Confusion Matrix

A confusion matrix, also known as an error matrix, is a table that is often used to evaluate the performance of a classification model. It lists and contrasts a model's classification task predictions with the data's actual labels. Usually, it is laid up as a square matrix, with the rows denoting the actual labels and the columns denoting the predicted labels. The matrix's diagonal members indicate accurate predictions, while its off-diagonal elements account for incorrect classifications.
The matrix denotes TP (true positives) and TN (true negatives) as correctly predicted data. Whereas FP (false positives) and FN (false negatives) as incorrectly predicted ones.
The confusion matrix enables you to calculate numerous evaluation metrics, including accuracy, precision, recall, and F1 score, and it gives you significant information about how well a classification model is performing. These metrics can assist you in identifying the model's advantages and disadvantages concerning accurately classifying cases. So the labels that will be use to calculate the derived metrics are:

- $TP = TruePositives$
- $FP = FalsePositives$
- $TN = TrueNegatives$
- $FN = FalseNegatives$

Table 4.1: Confusion Matrix

	Predicted Positive	**Predicted Negative**
Actual Positive	TP	FN
Actual Negative	FP	TN

4.3.2 Precision

This measure tells us how many of the predictions were accurate by the model. We can calculate the rate of accurate prediction, i.e. precision using the following equation:

$$Precision = \frac{TP}{TP + FP} \tag{4.1}$$

4.3.3 Recall

This measures the ability of the model to predict accurate class of an input from the original inputs of that class. This is the ratio of the number of accurate predictions to the original number of inputs in a particular class.

$$Recall = \frac{TP}{TP + FN} \tag{4.2}$$

4.3.4 F1-Score

F1 score is the metric we used to find the balance between precision and recall. The equation is as follows:

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4.3}$$

Chapter 5

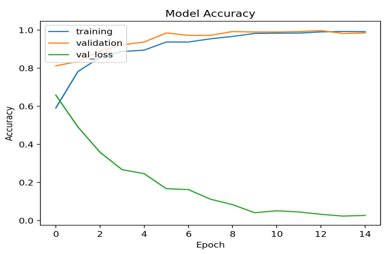
Result and Discussion

5.1 Training Evaluation

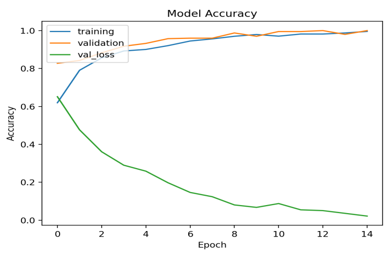
We adopted two types of strategy to understand the effect of stemming in Bengali language. We conducted all the experiments both with stemming and without stemming. Also we intended to observe the effect of ratio between fake and authentic on the training. So we our experiments also include two ratios of *fake:authentic* as mentioned in section 3.4.

5.1.1 mBERT-uncased

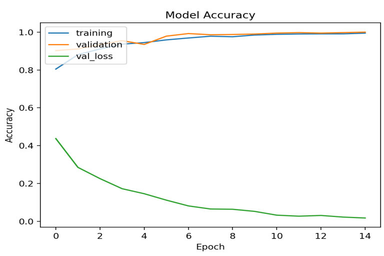
This transformer was trained on over a hundred languages to learn the semantic understanding of those languages. *Uncased* here means that the model is not case sensitive.



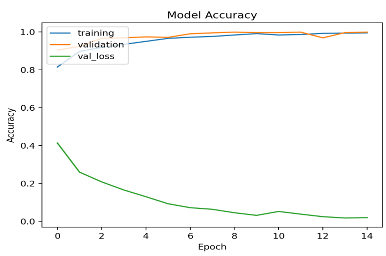
(a) Without Stemming for 1:1



(b) With Stemming for 1:1

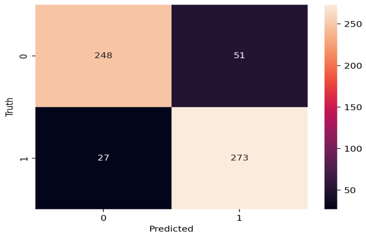


(c) Without Stemming for 1:3

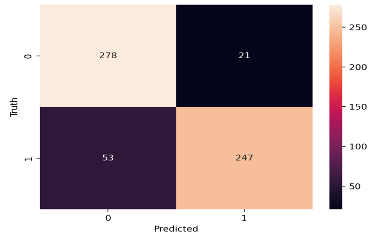


(d) With Stemming for 1:3

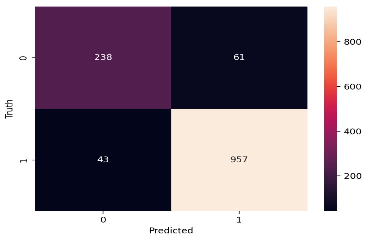
Figure 5.1: Learning Curves of mBERT-uncased



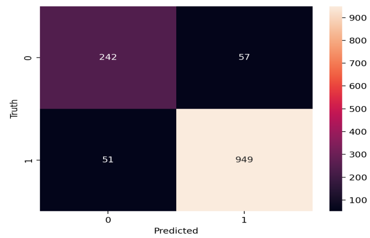
(a) Without Stemming for 1:1



(b) With Stemming for 1:1



(c) Without Stemming for 1:3

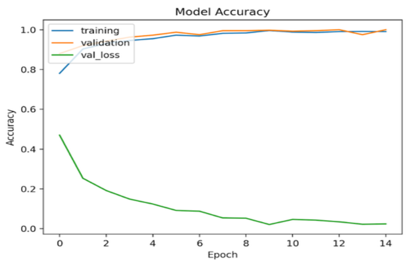


(d) With Stemming for 1:3

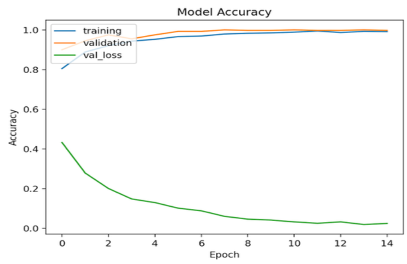
Figure 5.2: Confusion Matrices of mBERT-uncased

5.1.2 mBERT-cased

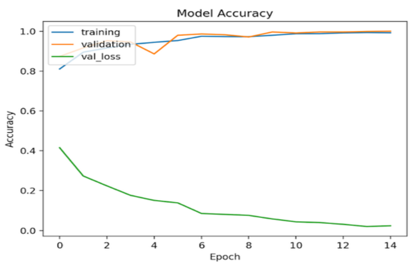
Like the other variation of *mBERT*, this transformer was trained on over a hundred languages to learn the semantic understanding of those languages. Evidently, *cased* here means that the model is case sensitive.



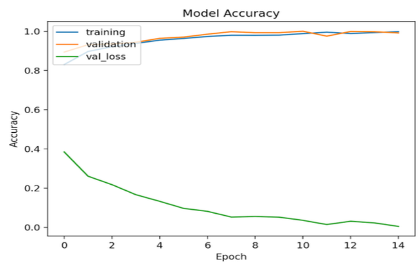
(a) Without Stemming for 1:1



(b) With Stemming for 1:1

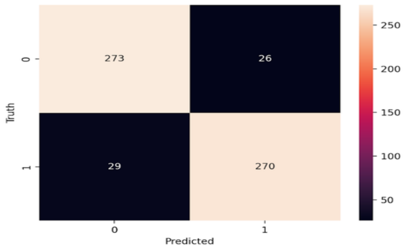


(c) Without Stemming for 1:3

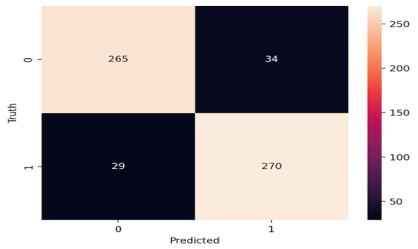


(d) With Stemming for 1:3

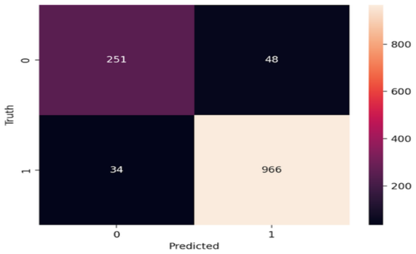
Figure 5.3: Learning Curves of mBERT-cased



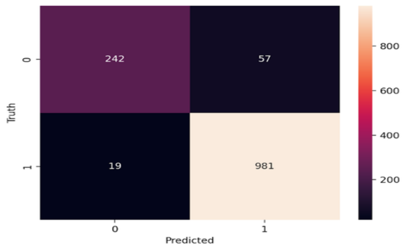
(a) Without Stemming for 1:1



(b) With Stemming for 1:1



(c) Without Stemming for 1:3

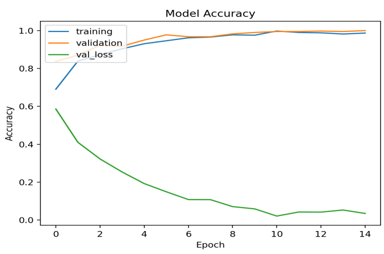


(d) With Stemming for 1:3

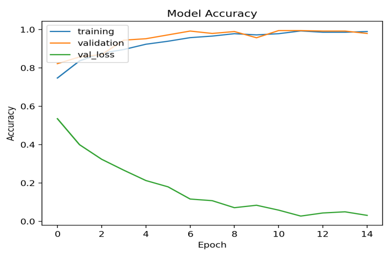
Figure 5.4: Confusion Matrices of mBERT-cased

5.1.3 XLM-RoBERTa-base

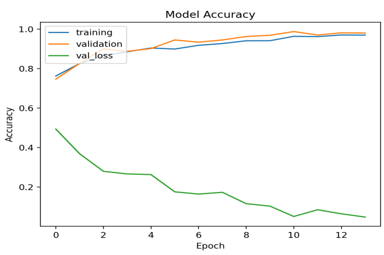
This transformer was trained on exactly a hundred languages to learn the semantic understanding of those languages. This used the TLM technique, as described in section 4.1.1, while training so that the model gains inter-language understanding as well.



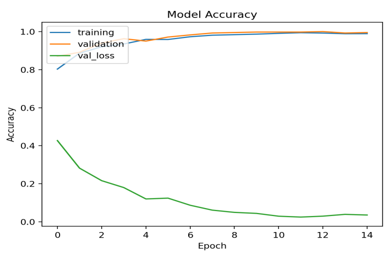
(a) Without Stemming for 1:1



(b) With Stemming for 1:1

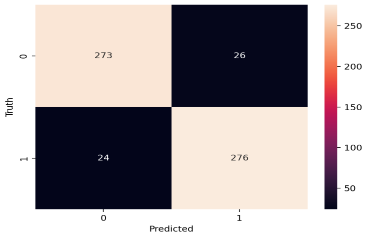


(c) Without Stemming for 1:3

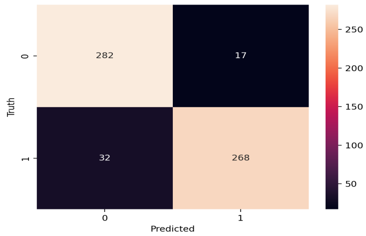


(d) With Stemming for 1:3

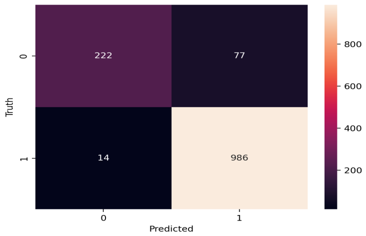
Figure 5.5: Learning Curves of XLM-RoBERTa



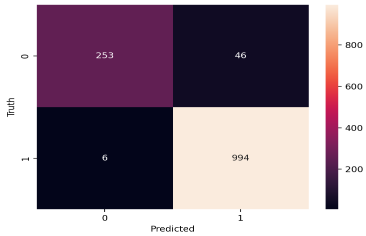
(a) Without Stemming for 1:1



(b) With Stemming for 1:1



(c) Without Stemming for 1:3

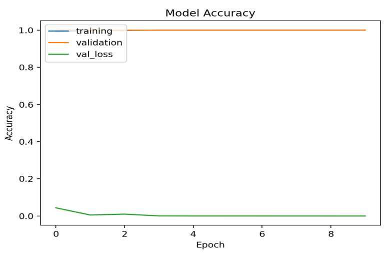


(d) With Stemming for 1:3

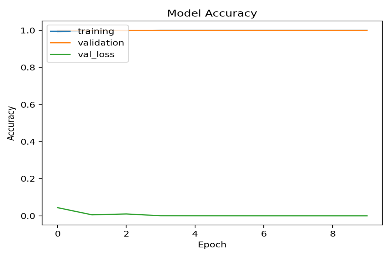
Figure 5.6: Confusion Matrices of XLM-RoBERTa

5.1.4 Bangla-BERT

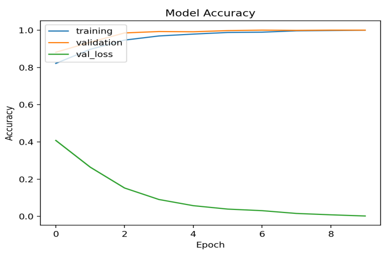
This was trained on large Bengali corpora to gain semantic understanding of the language exclusively.



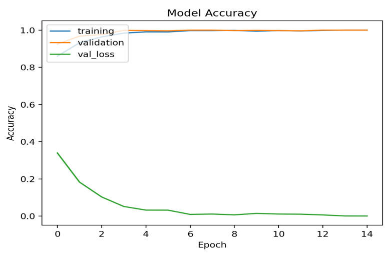
(a) Without Stemming for 1:1



(b) With Stemming for 1:1

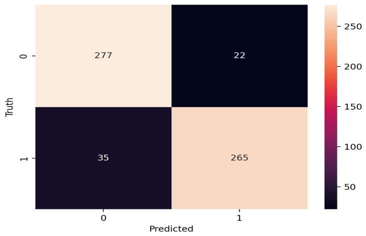


(c) Without Stemming for 1:3

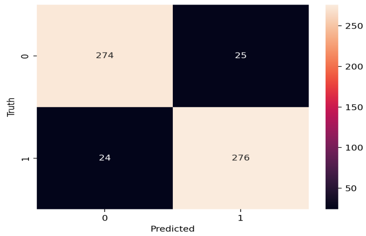


(d) With Stemming for 1:3

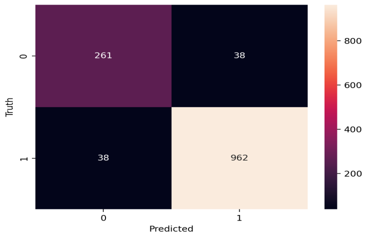
Figure 5.7: Learning Curves of Bangla-BERT



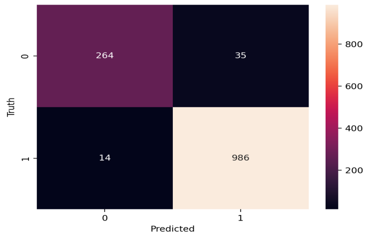
(a) Without Stemming for 1:1



(b) With Stemming for 1:1



(c) Without Stemming for 1:3



(d) With Stemming for 1:3

Figure 5.8: Confusion Matrices of Bangla-BERT

5.1.5 Result Analysis

The evaluation metrics calculated from those confusion matrices will equip us with significant insights into the performance of our experimented models. The scores are presented in the following tables.

Fake:Authentic = 1:1

Table 5.2 shows that *XLM-RoBERTa* performs the best when the models are trained with texts not stemmed. But table 5.1 shows that *Bangla-BERT* matches the performance of *XLM-RoBERTa* if texts are stemmed during the pre-processing stage. Because of the use of TLM as mentioned in section 4.1.1, XLM-RoBERTa performs better than the other two multilingual models, mBERT-cased and mBERT-uncased, though they all were pre-trained on large multilingual corpora. mBERT-cased performed better than mBERT-uncased.

Table 5.1: 1 to 1 without Stemming

Transformer	Authentic			Fake			Accuracy(%)
	Precision	Recall	F1	Precision	Recall	F1	
mBERT-uncased	0.84	0.91	0.88	0.90	0.83	0.86	87
mBERT-cased	0.91	0.90	0.90	0.90	0.91	0.90	90.7
XLM-Roberta	0.91	**0.92**	**0.92**	**0.92**	0.91	**0.92**	**91.7**
Bangla-BERT	**0.92**	0.88	0.90	0.89	**0.93**	0.91	90.4

Table 5.2: 1 to 1 with Stemming

Transformer	Authentic			Fake			Accuracy(%)
	Precision	Recall	F1	Precision	Recall	F1	
mBERT-uncased	0.92	0.82	0.87	0.84	0.93	0.88	87.6
mBERT-cased	0.89	0.90	0.89	0.90	0.89	0.89	89.3
XLM-Roberta	**0.94**	0.89	**0.92**	0.90	**0.94**	**0.92**	**91.8**
Bangla-BERT	0.92	**0.92**	**0.92**	**0.92**	0.92	**0.92**	**91.8**

Ratio of Fake:Authentic = 1:3

Unlike the Bangla-BERT based model trained on the dataset of ratio 1:1, the one that is trained on the dataset of ratio 1:3 performs better than all the multilingual models. The reason could be the use of more data to train the models. The training graph was almost flat for Bangla-BERT trained on the smaller dataset, but with a bit of larger dataset our model seemed to learn better than all the multilingual models. This could be because Bangla-BERT was already trained on a large Benglai corpora shortening it's learning curve with a very small Bengali dataset. XLM-RoBERTa based model again came out to be a better performing model because of the use of TLM in the pre-training phase. mBERT-cased performed better than mBERT-uncased again.

Table 5.3: 1 to 3 without Stemming

Transformer	Authentic			Fake			Accuracy(%)
	Precision	Recall	F1	Precision	Recall	F1	
mBERT-uncased	0.94	0.96	0.95	0.85	0.80	0.82	92
mBERT-cased	**0.95**	0.97	0.96	0.88	**0.84**	0.86	93.7
XLM-Roberta	0.93	**0.99**	0.96	0.94	0.74	0.83	93
Bangla-BERT	**0.95**	**0.99**	**0.97**	**0.95**	0.82	**0.88**	**94.1**

Table 5.4: 1 to 3 with Stemming

Transformer	**Authentic**			**Fake**			Accuracy(%)
	Precision	Recall	F1	Precision	Recall	F1	
mBERT-uncased	0.94	0.95	0.95	0.83	0.81	0.82	91.7
mBERT-cased	0.94	0.98	0.96	0.93	0.81	0.86	94.1
XLM-Roberta	0.96	**0.99**	0.97	**0.98**	0.85	0.91	96
Bangla-BERT	**0.97**	**0.99**	**0.98**	0.95	**0.88**	**0.92**	**96.2**

5.2 Discussion

- From the tables 5.3, 5.1, 5.2, and 5.4 we can see that the best performing model in each type of experiment is the one which used the *Bangla-BERT* pre-trained transformer. The reason could be it is specific pre-training on large Bengali corpora and specific knowledge of semantics of Bengali language. Although the scores are very high for authentic news, recall scores of fake news classification is comparatively less. We can see that the disparity between the scores is almost absent in case of *Bangla-BERT+DNN* model trained on the dataset with ratio *fake:authentic = 1:1*. Recall scores drop in case of the other dataset, indicating lack of performance in accurately classifying fake news from the actual ones. The reason could be the lack of fake news data and the imbalance in the dataset leading to learning curves that results inconsistent scores for fake news detection. It could be inferred from this that if the dataset had more fake news data we could used larger dataset with the ratio 1:1, eventually getting better and consistent performance from the models.

- We wanted to examine the effect of stemming as stemming in Bengali language can be tricky. Like চা, পা are words in Bengali, but they can be the stemmed version of চাওয়া, and পাওয়া respectively. Which potentially could have a negative effect in capturing contexts. Results from the experiments show that stemming has overall a positive effect on the models.

- After the *Bangla-BERT* transformer, close second is *XLM-Roberta*. It performed quite well on the *1:1 dataset*. The reason could be the use of *Translation Language Modeling(TLM)*4.1.1 in the pre-training of the transformer. Which resulted in a better understaning of the language.

- *mBERT* models could not perform better than any of the previous two. The reason could be that the *Bangla-BERT* was trained on Bengali corpora exclusively while *mBERT* trained on more than a hundred languages [32] [31] allowing *Bangla-BERT* a better and faster learning curve and eventually a better result.

Chapter 6

Conclusion

6.1 Summary

In this research, we experimented with some of the widely used pre-trained transformers, which were trained on large multilingual corpora. Also, we experimented with *Bangla-BERT*, which was specifically trained on Bengali corpora. Moreover, throughout our study, we made an effort to observe the effect of stemming in the Bengali language. This array of experiments led us to results that have profound significance in the domain of natural language processing in Bengali, especially in Bengali fake news detection. The use of the *DNN* layer on top of pre-trained transformers brought us good results. This promotes the importance of experimenting with transformers with other neural networks to achieve better results, which otherwise are not possible with traditional networks or just transformers alone. The use of pre-trained tokenizers produced meaningful token embeddings that resulted in better training performance. Also, it saved us a lot of time and effort and allowed us to invest that elsewhere, thus broadening our experimental reach. We compared all the models of our experiments with *precision*, *recall*, *F1-score*, and *accuracy*. Finally, we conclude with the suggestion of a *Bangla-BERT* transformer-based model, which can detect Bengali fake news most efficiently. Overall, by leveraging transformers, we created an efficient fake news detection model in Bengali.

6.2 Future Scope

Lack of fake news data and utilizing other neural network with transformers indicate toward a future full of possibilities and opportunities.

- Collection of more fake news data is the most evident step that one can take to improve fake news detection. As we've seen from the metrics that the ratio of 1:1 produced consistent training performance and evaluation scores. Lack of fake news in the dataset led to inconsistency in the performance.

- Neural layers like *CNN*, *RNN* and *LSTM* variations can be used instead of simple fully connected *DNN* layer to improve performance. As LSTM and RNN can extract and utilize contextual information and gain a better semantic understanding of sequential information.

- Other transformer models like T5 and its variations can be explored to in this classification task.

Bibliography

[1] M. Z. Hossain, M. A. Rahman, M. S. Islam, and S. Kar, "BanFakeNews: A dataset for detecting fake news in Bangla," in *Proceedings of the Twelfth Language Resources and Evaluation Conference*. Marseille, France: European Language Resources Association, May 2020, pp. 2862–2871. [Online]. Available: https://aclanthology.org/2020.lrec-1.349

[2] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, and V. Stoyanov, "Unsupervised cross-lingual representation learning at scale," *CoRR*, vol. abs/1911.02116, 2019. [Online]. Available: http://arxiv.org/abs/1911.02116

[3] C.-F. Chen, Q. Fan, and R. Panda, "Crossvit: Cross-attention multi-scale vision transformer for image classification," 2021.

[4] N. R. Bethan Staton, "Donald trump catchphrase 'fake news' wins word of the year," Nov. 2, 2017, [Feb. 21, 2023][Online]. [Online]. Available: https://news.sky.com/story/donald-trump-catchphrase-fake-news-wins-word-of-the-year-11109333

[5] M. Martínez, "Burned to death because of a rumour on whatsapp," (Nov. 12, 2018), [Accessed Nov. 12, 2022][Online]. [Online]. Available: https://www.bbc.com/news/world-latin-america-46145986

[6] P. Mozur, "A genocide incited on facebook, with posts from myanmar's military," (Oct. 15, 2018), [Oct. 15, 2022][Online]. [Online]. Available: https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html

[7] Management and R. D. Initiative, "News literacy in bangladesh, national survey," (2020), [Jan. 11, 2023][Online]. [Online]. Available: https://mrdibd.org/wp-content/uploads/2021/07/News-Literacy-in-Bangladesh-National-Survey.pdf

[8] R. Team, "Rumor: Thankuni will prevent coronavirus," (18 March, 2020), [Jan. 11, 2023][Online]. [Online]. Available: https://rumorscanner.com/fact-check/archives/750

[9] R. Rafe, "Bangladesh: Fake news on facebook fuels violence," (11 Jan. 2019), [Jan. 11, 2023][Online]. [Online]. Available: https://www.dw.com/en/bangladesh-fake-news-on-facebook-fuels-communal-violence/a-51083787

[10] B. R. COX'S BAZAR, "Muslim protesters torch buddhist temples, homes in bangladesh," (30 Sep. 2012), [Jan. 11, 2023][Online]. [Online]. Available: https://www.reuters.com/article/uk-bangladesh-temples-idUKBRE88T03Q20120930

[11] B. F. Check, "Bd fact check," [Feb. 21, 2023][Online]. [Online]. Available: https://bdfactcheck.com/

[12] Jachai, "Jachai," [Feb. 21, 2023][Online]. [Online]. Available: https://www.jachai.org/

[13] rumor scanner, "Rumor scanner bd," [Feb. 21, 2023][Online]. [Online]. Available: https://rumorscanner.com/

[14] K. Shu, A. L. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ArXiv*, vol. abs/1708.01967, 2017.

[15] H. Rashkin, E. Choi, J. Y. Jang, S. Volkova, and Y. Choi, "Truth of varying shades: Analyzing language in fake news and political fact-checking," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Copenhagen, Denmark: Association for Computational Linguistics, Sep. 2017, pp. 2931–2937. [Online]. Available: https://aclanthology.org/D17-1317

[16] O. Sen, M. Fuad, M. N. Islam, J. Rabbi, M. Masud, M. K. Hasan, M. Awal, A. Fime, M. T. Fuad, D. Sikder, and A. R. Iftee, "Bangla natural language processing: A comprehensive analysis of classical, machine learning, and deep learning based methods," *IEEE Access*, vol. 10, pp. 1–1, 01 2022.

[17] S. Sazzed, "Cross-lingual sentiment classification in low-resource bengali language," 01 2020, pp. 50–60.

[18] S. Ruíz, E. Providel, and M. Mendoza, "Fake news detection via english-to-spanish translation: Is it really useful?" in *Social Computing and Social Media: Experience Design and Social Network Analysis*, G. Meiselwitz, Ed. Cham: Springer International Publishing, 2021, pp. 136–148.

[19] A. Anjum, M. Keya, A. K. Mohammad Masum, and S. R. Haider Noori, "Fake and authentic news detection using social data strivings," in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2021, pp. 1–5.

[20] S. B. S. Mugdha, S. M. Ferdous, and A. Fahmin, "Evaluating machine learning algorithms for bengali fake news detection," in *2020 23rd International Conference on Computer and Information Technology (ICCIT)*, 2020, pp. 1–6.

[21] F. Islam, M. M. Alam, S. M. Shahadat Hossain, A. Motaleb, S. Yeasmin, M. Hasan, and R. M. Rahman, "Bengali fake news detection," in *2020 IEEE 10th International Conference on Intelligent Systems (IS)*, 2020, pp. 281–287.

[22] "Detection of fake news using deep learning cnn–rnn based methods," *ICT Express*, vol. 8, no. 3, pp. 396–408, 2022.

[23] R. K. Kaliyar, A. Goswami, and P. Narang, "Fakebert: Fake news detection in social media with a bert-based deep learning approach," *Multimedia Tools Appl.*, vol. 80, no. 8, p. 11765–11788, mar 2021. [Online]. Available: https://doi.org/10.1007/s11042-020-10183-2

[24] R. I. Rasel, A. H. Zihad, N. Sultana, and M. M. Hoque, "Bangla fake news detection using machine learning, deep learning and transformer models," in *2022 25th International Conference on Computer and Information Technology (ICCIT)*, 2022, pp. 959–964.

[25] N. M. Jakilim, S. Mahamudul Hasan, and E. Hassan, "A benchmark of machine learning and deep learning algorithms for detecting fake news in bangla language," in *2022 4th International Conference on Sustainable Technologies for Industry 4.0 (STI)*, 2022, pp. 1–6.

[26] N. Khan, M. S. Islam, F. Chowdhury, A. S. Siham, and N. Sakib, "Bengali crime news classification based on newspaper headlines using nlp," in *2022 25th International Conference on Computer and Information Technology (ICCIT)*, 2022, pp. 194–199.

[27] C. S. Office, "Household internet security and information integrity 2021," (2021), [Dec. 25, 2022][Online]. [Online]. Available: https://www.cso.ie/en/releasesandpublications/ep/p-isshisi/householdinternetsecurityandinformationintegrity2021/informationintegrity/

[28] S. Sarker, "Bnlp toolkit," (Latest Release: Apr 29, 2023). [Online]. Available: https://pypi.org/project/bnlp-toolkit/

[29] A. Hossain, "Bnltk," (Latest Release: Jun 29, 2019). [Online]. Available: https://pypi.org/project/bnltk/

[30] H. F. Team, "Hugging face," [Feb. 21, 2023][Online]. [Online]. Available: https://huggingface.co/docs/transformers/index

[31] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: pre-training of deep bidirectional transformers for language understanding," *CoRR*, vol. abs/1810.04805, 2018. [Online]. Available: http://arxiv.org/abs/1810.04805

[32] ——, "BERT: pre-training of deep bidirectional transformers for language understanding," *CoRR*, vol. abs/1810.04805, 2018. [Online]. Available: http://arxiv.org/abs/1810.04805

[33] S. Sarker, "Banglabert: Bengali mask language model for bengali language understading," 2020. [Online]. Available: https://github.com/sagorbrur/bangla-bert

[34] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," 2019.

[35] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "Roberta: A robustly optimized bert pre-training approach," 2019.

[36] X. Song, A. Salcianu, Y. Song, D. Dopson, and D. Zhou, "Fast wordpiece tokenization," 2021.