# Tackling Fake News in Bengali: Unraveling the Impact of Summarization vs. Augmentation on Pre-trained Language Models

Arman Sakif Chowdhury†, G. M. Shahariar†, Ahammed Tarik Aziz, Syed Mohibul Alam, Md. Azad Sheikh and Tanveer Ahmed Belal

Ahsanullah University of Science and Technology, Dhaka, Bangladesh.

*Corresponding author(s). E-mail(s): armansakif90@gmail.com; Contributing authors: sshibli745@gmail.com; tarikaziz821@gmail.com; sadeemohib@gmail.com; mazad6932@gmail.com; belal.cse@aust.edu;
†These authors contributed equally to this work.

**Abstract**

With the rise of social media and online news sources, fake news has become a significant issue globally. However, the detection of fake news in low resource languages like Bengali has received limited attention in research. In this paper, we propose a methodology consisting of four distinct approaches to classify fake news articles in Bengali using summarization and augmentation techniques with five pre-trained language models. Our approach includes translating English news articles and using augmentation techniques to curb the deficit of fake news articles. Our research also focused on summarizing the news to tackle the token length limitation of BERT based models. Through extensive experimentation and rigorous evaluation, we show the effectiveness of summarization and augmentation in the case of Bengali fake news detection. We evaluated our models using three separate test datasets. The BanglaBERT Base model, when combined with augmentation techniques, achieved an impressive accuracy of 96% on the first test dataset. On the second test dataset, the BanglaBERT model, trained with summarized augmented news articles achieved 97% accuracy. Lastly, the mBERT Base model achieved an accuracy of 86% on the third test dataset which was reserved for generalization performance evaluation. The datasets and implementations are available at https://github.com/arman-sakif/Bengali-Fake-News-Detection.

**Keywords:** fake news, classification, summarization, Bengali, augmentation, transformers, BERT

# 1 Introduction

Fake news refers to news stories or articles that are deliberately misleading or fabricated and spread through various media channels, such as social media, news websites, or even traditional printed newspapers [1]. It can also be a misguided viewpoint of the author. "Fake news" is an old term, dating back to the time when newspapers had just gained popularity, long before the web was invented [2]. These types of news articles are intentionally created to deceive readers or viewers or to manipulate public opinion for political, financial, or other reasons. Fake news is created and spread by either those who have ideological interests, malicious intentions, or media individuals looking for some views, which are commonly termed as clickbait. Fake news can be presented in a variety of formats, such as articles, images, videos, and even memes [3].

The dissemination of fake news has led to significant societal consequences, including polarization, misinformation, and confusion. A study [4] shows that fake news spreads faster than the truth and humans are primarily responsible for the spread of misleading information. Due to the fact that most of these stories are not verified and are so much more interesting than regular news, it makes it difficult to resist spreading them. According to a 2018 study by Pew Research Center, 64% of U.S. adults believe that fake news has caused a great deal of confusion about the basic facts of current events, and 23% of people say they have shared a made-up news story [5].

On February 3, 2022, a prominent Bangladeshi TV station had a news story rated as false information by the fact-checking unit of the international news agency AFP [6]. The story, which was credited to the state-owned Saudi Press Agency (SPA), claimed that the Saudi government had approved a draft amendment to redesign its national flag by removing Kalima Tayyiba (Islamic declaration of faith) from it. However, this claim turned out to be untrue as neither the SPA nor any other Saudi news media had reported anything of that sort. The actual revision of the draft was related to proposing new regulations regarding the use of the flag, and several Saudi newspapers confirmed that no changes were proposed to the contents of the flag [7]. In a separate report by Boom Bangladesh [8], a third-party fact-checking partner of Facebook focusing on Bangladesh, it was revealed that they had detected 23 instances of fake news in mainstream media outlets between March and December 2020. Notably, the most popular TV channel in the country was caught spreading fake news 10 times during that 10-month period [7].

Fake news can also have serious consequences for public health. During the early days of the Coronavirus period in 2020, a rumor spread out that drinking disinfectants could cure Covid-19. Some people even ingested hand sanitizer with methanol which led 4 people to death and 26 hospitalized [9]. In another instance, people believed that 5G spread coronavirus, and more than 70 phone masts have been vandalized because of this false rumor in the UK [10]. Some people even claimed the whole coronavirus is a hoax and refused to take quarantine seriously or taking cure when it came out. Fake news can also impact businesses and their reputations. For example, false rumors about a company's financial situation can lead to a decline in stock prices and cause significant financial losses. As we can see, addressing the challenge of identifying and classifying fake news has become increasingly important in the context of the digital age.

Research supports that fake news are harder to detect [11]. There are various reasons for fake news being difficult to detect. With the proliferation of social media platforms and the internet, there are now millions of sources of news and information. Many of these sources are not reputable, and it can be difficult to distinguish between legitimate sources and fake news. As already discussed, fake news spreads even faster than real news. By the time a news article has been identified as fake, it may have already been shared by thousands or even millions of people. With the sheer number of people sharing the same news, people tend to believe the news to be real. Another reason can be confirmation bias [12]. People often seek out information that already confirms their existing beliefs and opinions. This means that fake news stories that confirm people's biases are more likely to be shared and believed.

In recent years, machine learning has evolved as an appropriate solution for fake news detection for several reasons. Firstly, for a machine learning model to work, it needs to learn with a significant amount of data. And with the proliferation of social media and online news sources, there is now a significant amount of data available to train machine learning models to detect fake news. Machine learning algorithms process large amounts of data automatically and find underlying patterns or characteristics which can help detect fake news. Pattern recognition is particularly useful as many fake news share an underlying pattern. Research in fake news has been conducted for a long time, especially in English. There has been extensive work on classifying fake news in the English language [13–19]. Recently other than the English language, research on fake news is being conducted in Spanish [20], French [21–24], Arabic [25], Chinese [26–32], and many other languages [33–42]. There are many different methods used in these researches but deep learning based methods seemed to be the most popular.

Bengali is one of the largest speaking languages, and although there are English newspapers in our country, most people read news in our native language- Bengali. But there has not been much research to detect fake news in Bengali. The probable reason that there is so little research in Bengali to detect fake news is because there is not enough labeled fake news data available. As there are authentic news sites and most big news organizations have online news portals, it is not difficult to collect authentic news data. But it is difficult to find a source with large amounts of data for fake news. In recent years, the few pieces of research that have been conducted in Bengali mostly used deep learning as a base method to detect fake news, especially using BERT models is prevalent. One of the most notable research is BanFakeNews [43]. The limitation of this research is that their built model has trouble detecting the minority class which is fake news. Sharma et al. [44] conducted another research to detect satirical news in Bengali using Convoluted Neural Network. In a recent research published in 2022- AugfakeBERT [45], the authors used the BanFakeNews dataset and tried to tackle the problem of insufficient fake news articles through text augmentation. The limitation of this research is that they did not address BERT's ability to only take input sequences up to 512 tokens. This shortcoming can be significant as many news articles contain large volumes of text.

In this study, by fine-tuning a few existing pre-trained transformers and employing summarization and augmentation techniques, we have proposed a novel methodology consisting of four distinct approaches for Bengali fake news identification. To tackle the class imbalance problem we increased fake news articles by generating synthetic Bengali fake news through English to Bengali translation and using different types of augmentation such as - token replacements and paraphrasing. To tackle BERT's 512 token limitation, large texts were summarized using pre-trained summarization models. Additionally, to test the generalization performance of the proposed approaches, we manually collected 102 fake news articles from various sources and selected equal numbers of random authentic news from *BanFakeNews* corpus. We utilized standard evaluation metrics to measure the performance of the proposed methodology. Our experimental results show the ability of the proposed approaches to detect Bengali fake news with high accuracy even with completely unseen test dataset. In summary, the followings are the notable contributions we made in this paper:

- We propose a novel methodology consisting of four distinct approaches using summarization and augmentation to detect Bengali fake news.
- Through extensive experimentation, we demonstrate the effectiveness of applying summarization and augmentation methods to enhance the generalization performance. To overcome the token length limitation of BERT-based models, we develop a summarization pipeline. Furthermore, we employ translation and augmentation strategies to address the class imbalance issue resulting from the scarcity of fake news articles in the training dataset.
- To evaluate the generalization performance, we curate a test dataset comprising 102 manually collected fake news articles that are entirely unseen

by the existing models. This dataset serves as a benchmark for comparing the performance of different approaches.
- We have made the datasets and implementations of the proposed approaches publicly available with the aim of garnering collaboration and encouraging future researchers to contribute to this particular field.

The rest of the paper is organized as follows: section 2 reviews the relevant literature to classify fake news. Section 3 illustrates all the corpus we used in our paper. Section 4 describes the research methodology. Our experimental results and findings are presented in section 5. The Final section 6 contains a description of the limitations of this work and future possibilities and then draws a conclusion.

## 2 Related Works

In this section, we discuss some of the past studies which are relevant to our research. Extensive research has been conducted to detect fake news in English, however, Bengali literary works are scarce. We studied a number of works in Bengali, English, and some other languages and we classify them into four main categories - traditional machine learning based approaches, neural network based approaches, hybrid & ensemble approaches and pre-trained transformer model based approaches. In our following narratives, we discuss each of the three categories.

### 2.1 Traditional Approaches

Using traditional machine learning based approach, there are several techniques to detect fake news. This approach typically involves feature engineering and the use of various classifiers such as Support Vector Machines (SVM), Logistic Regression (LR), Naive Bayes (NB), Random Forests (RF), K-nearest Neighbor algorithm (KNN), Stochastic Gradient Descent (SGD), Passive Aggressive Classifier (PAC) and Decision Trees (DT).

Khan et al. [46] conducted a benchmark study to evaluate the effectiveness of several practical approaches on three distinct English datasets, the largest and most diverse of which was created by them. They experiment with both traditional machine learning based and neural network based approaches. Within traditional models they used SVM, LR, DT, MNB and KNN - and among them Naive Bayes, with n-gram features, performed the best.

Baarir et al. [47] used two existing English datasets and divided news data into three categories - Textual, categorical and numerical. Category wise, they used different preprocessing techniques. They only used Support Vector Machine (SVM) classifier but experimented with different hyper-parameters to improve accuracy. Another unique aspect of their research is that, they experimented with different features and through experimentations denote that the best features are in this order - text, author, source, date and sentiment. Using all features together they achieved 100% accuracy in their research.

Hossain et al. [43] created *BanFakeNews*, which is an annotated dataset of around 48 thousand authentic and around 1300 fake news in Bengali. To our knowledge, this is the only publicly available dataset in on Bengali fake news and it paved research on Bengali fake news detection. They collected authentic news articles from 22 reputed sources and three types of fake news - misleading, clickbait and satire are collected from some popular websites. They initially experiment with traditional machine learning models - SVM, LR and RF and empirically find that SVM incorporated with all linguistic features outperforms the other two models, achieving 0.91 F1-score on fake class. They also observe that lexical features perform better than other linguistic features. Additionally, they claim that the use of punctuation in fake news is more frequent, and most of the time fake news is found on the least popular sites.

In the research of Hussain et al. [48] Multinomial Naive Bayes (MNB) and Support Vector Machine (SVM) are used to build Bengali fake news detection model using count vectorizer and TF-IDF as feature extraction methods. They created their own dataset of around 2500 news articles among which 1548 were authentic and 993 fake. In their experimentation, SVM with an accuracy of 0.9664 outperforms MNB with an accuracy of 0.9332.

The research of Sraboni et al. [49] utilized the *BanFakeNews* dataset [43] and another dataset of 2500 news [48] and trained several traditional machine learning algorithms - RF, PAC, MNB, SVM, LR and DT using TF-IDF as a feature extractor. Rather than using all of 51.8k available data, they used 3.5k authentic data and 2.3k fake data to remove bias. They also experimented with different train-test split ratios with 50-50, 60-40, 80-20 and 70-30 and empirically found that 70-30 split ratio gives the best result. Among the models, Passive Aggressive Classifier (PAC) and Support Vector Machine (SVM) achieve 0.938 and 0.935 accuracy respectively which are higher than the other models they trained.

Mugdha et al. [50] created their own balanced dataset of total 112 instances with an equal number of fake and real news. They tested multiple traditional algorithms with tenfold cross validation - Support Vector Machine (SVM), Logistic Regression (LR), Multi-layer Perceptron (MLP), Random Forest Classifier (RF), Voting Ensemble Classifier (VEC) and Gaussian Naive Bayes (GNB). Initially, none of the algorithms were performing well. Then they applied 3 feature selection techniques - PCA, Kernel PCA and Extra Tree Classifier. Through empirical research, they discovered that Extra Tree Classifier outperformed other techniques and that GNB with Extra Tree Classifier outperformed other models, with an accuracy of 0.8552.

The paper of Piya et al. [51] researched developing a model which can simultaneously identify Bengali and English fake news at the same time. Their approach for detecting fake news from a bilingual perspective uses feature extraction techniques such as TF-IDF and N-Gram analysis. They trained six supervised traditional machine learning models - Logistic Regression, Linear SVC, Decision Tree Classifier, Random Forest Classifier, Multinomial Naive Bayes, and Passive Aggressive Classifier. Among these six models, the linear

SVC performed best and score 0.9329 in accuracy and 0.93 F1-score. They adopted BanFakeNews for the Bengali portion of their Dataset. Additionally, they used ISOT, a dataset released in [52], for the English dataset. In their final DS, they had 42,324 bilingual samples. The majority of them were in English.

In the paper of Mugha et al. [53] they focused only on the headlines to detect fake news in Bengali. They developed their own Bengali dataset with a total of 538 cases, of which 269 were authentic and 269 were false. They have used 9 different traditional machine learning algorithms for their research - SVM, Logistic Regression(LR), multilayer perceptron (MLP), Random Forest Classifier (RF), Voting Ensemble Classifier (VEC), Gaussian Naive Bayes (GNB), Multinomial Naive Bayes (MNB), AdaBoost (AB) and Gradient Boosting (GB). With a score of 0.87, the Gaussian Naive Bayes classifier outperformed all other models in terms of accuracy. This algorithm chose the attribute using an Extra Tree Classifier and a text feature dependent on TF-IDF. Performance was assessed using stratified 10-fold cross-validation on the dataset. They contribute to this field by creating a novel dataset and also a new Bengali stemmer.

## 2.2 Neural Network based Approaches

A neural network is a type of computational model that draws inspiration from the structure and operation of the human brain. It is a hierarchy of neurons, or interconnected nodes, arranged in layers. When a neural network contains several layers, it becomes a deep neural network. Deep learning is the usage of neural networks that have numerous hidden layers, which enables them to learn hierarchical data representations. Deep learning approaches, as opposed to traditional machine learning, are able to recognize significant features and grasp the semantic context of textual input. In this section, we discuss the relevant literature which was based on neural network models.

Sharma et al. [44] adopted Convoluted Neural Network (CNN) based approach built upon a hybrid feature extraction model. They created their own dataset by crawling satire news portal - Motikontho for satire news and real news from Prothom ALo and Ittefaq, two popular news media. They gather 1480 satire news and to balance the dataset, they picked the same amount from the real news. With this balanced training, they were able to achieve 0.96 in accuracy.

Bahad et al. [54] proposed using Bi-directional LSTM recurrent neural network to detect fake news. They used two publicly available news articles datasets separately. They experimented with Bi-directional LSTM, CNN, vanilla RNN, unidirectional LSTM and found Bi-directional LSTM to be superior, achieving .9108 in accuracy.

In the paper of Khan et al. [46], along with traditional methods, they also experimented with deep learning models such as - CNN, LSTM, Bi-LSTM, C-LSTM, HAN, Conv-HAN, char level C-LSTM. Their interpretation of the result is that no deep learning model is uniquely superior to others as no model performs consistently well on all three of their datasets. In addition, they note that while neural network-based models exhibit good accuracy and f1-score on a fairly large dataset (Combined Corpus), they exhibit overfitting on a small dataset (LIAR). Keya et al. in their paper - AugFakeBERT [45] experimented with all three - traditional, deep learning and transformer based approach. For deep learning models they used CNN, LSTM, Bi-LSTM and CNN-LSTM.

## 2.3 Hybrid and Ensemble Approaches

Some research tried to merge multiple methods together to achieve better results. In this section, we discuss them.

Ruchansky et al. [55] propose a hybrid model named CSI consisting of three modules - Capture, Score and Integrate. The first module employs a recurrent neural network (RNN) to capture the temporal structure of user activity on a particular article and is based on the response and text. The second module picks up on the source characteristics based on user activity, and then the two are combined with the third module to determine whether or not an article is authentic. They used two datasets - **TWITTER** and **WEIBO**. Their hybrid CSI model was able to achieve 0.95 in accuracy and F1-score.

The research team of Nasir et al. [56] proposed a novel hybrid deep learning model for classifying fake news that integrates recurrent and convolutional neural networks. They used two publicly available datasets **FA-KES** and **ISOT**. Both datasets have almost equal amounts of fake and real news. They were able to achieve an accuracy score of 0.60 in DS1 and 0.99 in DS2.

The research of Ahmad et al. [52] proposed ensembling multiple traditional machine learning models. They picked three publicly available English datasets which contain news from multiple domains. They explored various ensemble techniques such as bagging, boosting and voting classifier. Their first voting classifier is ensemble of LR,RF and KNN and second one consists of LR, linear SVM, and CART. The bagging ensemble consists 100 decision trees and two boosting algorithms are XGBoost and AdaBoost. Comparing ensemble and individual learners, the ensemble learners have consistently outperformed the individual learners in their research.

## 2.4 Transformer based Approaches

The field of natural language processing has been transformed by pre-trained transformer models like BERT (Bidirectional Encoder Representations from Transformers), which have displayed astounding performance on a variety of natural language processing tasks. Intricate associations and meanings are captured by these models as they learn contextualized word representations from huge datasets. Researchers have fine-tuned pre-trained transformer models successfully for the task of fake news classification.

Till now the only sufficiently large publicly available Bengali news dataset is BanFakeNews [43] but the dataset is heavily skewed towards real news. Keya et al. [45] suggested using text augmentation techniques to somewhat reduce the imbalance. Their research experiments with traditional, deep learning and transformer - all three approaches and they propose pre-trained transformer based approach. They fine tuned BERT base uncased architecture with balanced dataset and their build model - AugFakeBERT outperforms other methods by achieving an accuracy score of 92.45%.

Saha et al. [57] used two Bengali dataset and experimented with Pre-trained BERT transformer, LSTM with regularization, CNN, SVM and NB. They used two Bengali datasets separately for training and testing. In their research, BERT achieves 94% F1-score which is superior to other methods explored.

**Table 1**: Notable studies on fake news detection using traditional methods, neural networks, hybrid, ensemble Approaches and transformers.

| Type | Reference | Model used | Performance |
|---|---|---|---|
| Traditional Approach | [46] | Traditional (SVM, LR, DT, MNB, KNN) | 95% accuracy (NB) |
| | [47] | SVM | 100% accuracy using 5 features |
| | [43] | SVM,LR, RF | 91% F1 score(SVM) |
| | [48] | MNB, SVM | 96.64% accuracy(MNV) |
| | [49] | RF, PAC, MNB, SVM, LR, DT | 93.8% accuracy (PAC)<br>93.5% accuracy(SVM) |
| | [50] | SVM,LR,MLP,RF,VEC,GNB | 85.2% accuracy(GNB) |
| | [51] | LR, Linear SVC, DT, RF, MNB, PAC | 93.23% accuracy<br>93% F1-score |
| | [53] | SVM, LR, MLP, RF, VEC,<br>GNB, MNB, AdaBoost (AB) , GB | 87% accuracy |
| Neural Network | [44] | CNN | 96% accuracy |
| | [45] | Traditional (LR, MNB, DT, RF, SGD, SVM, KNN)<br><br>DL (CNN, LSTM, Bi-LSTM, CNN-LSTM, M-BERT)<br><br>Transformer(BERT, M-BERT | 91% accuracy(LSTM) |
| | [46] | DL (CNN, LSTM, Bi-LSTM, C-LSTM, HAN, Conv-HAN, char level C-LSTM) | DS1 - Conv HAN 59% acc<br>DS2 - C-LSTM 95% acc<br>DS3 - Bi-LSTM and C-LSTM both 95% |
| | [54] | LSTM, CNN,<br>vanilla RNN, unidirectional LSTM | 91.08% accuracy (Bi-LSTM) |
| Hybrid and Ensemble | [56] | CNN + RNN | 99% accuracy |
| | [52] | Traditional ( {LR,RF,KNN}, {LR,SVM,CART}, DT, Adaboost, XGBoost) | 96.3% accuracy (XGBoost) |
| | [55], | hybrid RNN | 95% accuracy |
| Transformer | [45] | Traditional (LR, MNB, DT, RF, SGD, SVM, KNN)<br><br>DL (CNN, LSTM, Bi-LSTM, CNN-LSTM, M-BERT)<br><br>Transformer(BERT, M-BERT) | 92.45% accuracy |
| | [57] | Transformer(BERT) | 94% accuracy |

## 3 Corpus Description

In this section, we provide a brief description of the corpora that we used for training and test purpose. Throughout our research, we used a total of three corpora: two publicly available corpora and a manually collected corpus which we refer to as - 'CustomFake'.

**(a) BanFakeNews Corpus:** BanFakeNews [43] dataset contains approximately 50,000 Bengali news articles among which about 48 thousand are of *authentic* category and 1299 are of *fake* category. The dataset contains seven attributes - *articleID, domain, date, category, headline, content* and *label.* The label here is binary where 0 denotes *fake* class and 1 denotes *authentic* class. For our research, we merged the headlines and their corresponding news content.

**Table 2**: BanFakeNews dataset statistics

| | **Authentic** | **Fake** |
|---|---|---|
| Number of instances | 48678 | 1299 |
| Average length of articles | 1858.95 | 1817.73 |
| Average word count | 283.6 | 289.54 |
| Longest article word count | 4791 | 2713 |
| BERT Max token length | 12539 | 8107 |

**(b) Fake News Detection Corpus**: The second corpus that we utilized is an English fake news article corpus named Fake News Detection[1]. This corpus has around 17000 articles related to *news, politics, government, middle-east. News* category articles are roughly 39%. We only took the *fake* articles of *News* category and translated them to Bengali using the English-Bengali machine translation model provided by Hasan et al. [58]. We refer to the translated fake news dataset as 'TransFND' dataset. We removed empty cells and finally, we had 4309 *fake* news articles which were necessary to curb the data imbalance of BanFakeNews corpus.

**(c) CustomFake Corpus:** As discussed before, the BanFakeNews corpus lacks a sufficient number of fake news samples. To somewhat solve this problem, we collected fake news articles manually. We visited the fake news sources provided by Hossain et al. [43] along with some new sources and collected a total of 102 new Bengali fake news articles and constructed a corpus we refer to as 'CustomFake' corpus. We only use this corpus to evaluate the generalization performance of the proposed approaches. The fake news collection sources for the corpus are provided in Table 3.

[1] https://www.kaggle.com/datasets/jainpooja/fake-news-detection

**Table 3**: Sources of the fake news articles collected for 'CustomFake' dataset.

| **Sources** | **Num of articles** | **Sources** | **Num of articles** |
|---|---|---|---|
| Earki | 22 | bengali.news19 | 2 |
| jachai.org | 12 | Banladesh football ultra | 1 |
| hindustantimes.com | 6 | Qatar Airways | 1 |
| bddailynews69.bd | 6 | The Daily star | 1 |
| nationalistview.com | 5 | Bangla Tribune | 1 |
| daily-star | 4 | songrami71 | 1 |
| The Daily star Bangla | 3 | sports protidin | 1 |
| zeenews | 3 | jamuna.tv | 1 |
| newschecker.in | 3 | 71News24 | 1 |
| DailyNews96.com | 3 | ntv | 1 |
| bangla.hindustantimes.com | 3 | Jugantor | 1 |
| Cinegolpo | 2 | roarmedia | 1 |
| banglainsider.com | 2 | Anandabazar | 1 |
| kalerkontho | 2 | awamiweb | 1 |
| Bangladesh24Online | 2 | Aviation News | 1 |
| jagonews | 2 | Kolar kontho | 1 |
| priyobangla24.com | 2 | bengali.news18 | 1 |
| bengali.news18.com | 2 | | |

## 3.1 Training Corpus

For the training purpose, we have utilized both the *BanFakeNews* and *TransFND* corpora. Table 4a shows the number of instances in both of the training datasets. We describe the two training datasets below.

**(a) <u>Dataset 1</u>:** We acquired 4309 translated fake news from *TransFND* corpus and we had 1299 fake news articles available in *BanFakeNews* corpus. We merged these articles which resulted in a total number of 5608 Bengali *fake* news articles. We randomly selected an equal number of *authentic* news from BanFakeNews. We separated 1200 news articles for testing which has an equal number of fake and real news. Thus 'Dataset 1' was left with 10,016 news articles, where 5008 is fake and 5008 is authentic.

**(b) <u>Dataset 2</u>:** Augmentation techniques are used for reducing the imbalance of the minority class brought on by insufficient samples. To create the second training dataset, in this work, we applied two augmentation techniques - *token replacement* and *paraphrasing*. We augmented 1299 *fake* news articles of *BanFakeNews* two times and we got a total of 3897 (1299$\times$2 and original 1299) *fake* news articles. We performed experimentation with 3507 *fake* instances and an equal number of real news. We also experimented by taking four augmentations (using *back translation* along with other two to augment) for a single fake instance and observed that most augmentations only showed subtle changes which resulted in over fitting problem during training. Therefore, we created 'Dataset 2' by augmenting a single fake instance using *token replacement* by leveraging BanglaBERT[2] and *paraphrasing* by leveraging BanglaT5[3].

[2]https://huggingface.co/sagorsarker/bangla-bert-base
[3]https://huggingface.co/csebuetnlp/banglat5_banglaparaphrase

| Corpus | Auth | Fake |
|---|---|---|
| Dataset 1 | 5008 | 5008 |
| Dataset 2 | 3507 | 3507 |

(a) Training

| Corpus | Auth | Fake |
|---|---|---|
| Test DS1 | 600 | 600 |
| Test DS2 | 2000 | 2000 |
| Test DS3 | 102 | 102 |

(b) Test

**Table 4**: Number of instances in training and test datasets.

## 3.2 Test Corpus

To evaluate the performance of our developed models, we created three different test corpora. Table 4b shows the data distributions in each test corpus. We describe the three test datasets below.
**(a) Test DS1:** In this test dataset, there are a total of 1200 instances where 600 instances are *fake* and the rest 600 are *authentic* news articles. The instances of test dataset are separated from the training 'Dataset 1' where the instances were from both the 'BanFakeNews' and 'TransFND' corpora.
**(b) Test DS2:** To create the second test dataset, we randomly pick 2000 *fake* articles from 4309 'TransFND' dataset. Then we randomly pick 2000 *authentic* articles from 'BanFakeNews' dataset. We concatenate *fake* and *authentic* data instances and shuffle them. While creating this test dataset, we made sure that the 2000 authentic instances were not used in the training dataset. Moreover, we only use this test dataset to evaluate the performance of our developed models while training with 'Dataset 2' because there are no data instances of 'TransFND' in the second training dataset.
**(c) Test DS3:** We use the 'CustomFake' corpus as the third test dataset where we manually collected 102 new *fake* news articles from various news websites. We pair them with 102 randomly selected *authentic* articles from BanFakeNews corpus. While creating this test dataset, we made sure that the 102 *authentic* news articles were not used in the training datasets.

# 4 Methodology

In this section, we present a detailed overview of our proposed methodology.

## 4.1 Proposed Approaches

We have explored four distinct approaches. The first approach involves fine-tuning pre-trained transformers [59] that serve as the baseline models. In the second approach, we employ summarized news articles to fine-tune these baseline models. The third approach involves fine-tuning the baseline models using both actual and augmented news articles. The fourth approach involves utilizing both summarized actual and augmented news articles for fine-tuning the baseline models. Figure 1 illustrates all four approaches. We provide detailed descriptions of each approach below.

**(a) Approach 1: Fine-tuning Transformers.**
In this approach, we have performed fine-tuning on five pre-trained language models using 'Dataset 1'. We followed the standard fine-tuning procedure for sequence classification task. Each text from the dataset is tokenized, [CLS] token is added to represent the beginning of a sequence. Each text is padded or truncated to the maximum length of 512. During fine-tuning, through back propagation, the weights of the classification layer and some of the upper layers of the language model are updated. In this approach, we did not introduce summarization or augmentation. Figure 1a represents the fine-tuning approach.

**(b) Approach 2: Fine-tuning Transformers via Summarization.**
Figure 1b represents this approach. In this approach, we began by summarizing the news articles from 'Dataset 1'. Then, we fine-tuned the pre-trained language models using these summarized news articles. The reason for implementing summarization is that pre-trained transformers can only handle up to 512 tokens, and news articles are usually longer than that limit. As a result, it is possible that the classification model could potentially miss crucial words in a news article, which are vital for distinguishing fake news from authentic ones. Therefore, we felt the need to condense the knowledge within the entire news articles as much as possible by summarizing, allowing the classifier model to focus on the most important information. We discuss the details of the summarization procedure in the following section.

**(c) Approach 3: Fine-tuning Transformers via Augmentation.**
Figure 1c provides a visual representation of this approach. Initially, we constructed 'Dataset 2' where we only augmented the fake news articles of the *BanFakeNews* dataset leaving the authentic news articles as they are. We then fine-tuned the pre-trained language models using Dataset 2. The purpose behind employing augmentation techniques was to introduce variations and noise into the fake news data. This was achieved by incorporating random perturbations such as token replacements. Text augmentation methods such as paraphrasing was utilized to preserve the semantics of the original text while presenting it in a different manner. This aspect holds particular significance in fake news detection as subtle alterations in how information is presented can significantly impact the ability of the model to identify fake news. Pre-trained language models when fine-tuned on augmented data, can learn more diverse representations which enhances their capability to generalize across different contexts and sources of fake news.

**(d) Approach 4: Fine-tuning via Summarization and Augmentation.**
The primary concept behind this approach is to examine how summarizing the augmented data influences the process of detecting fake news as illustrated in Figure 1d. Initially, we condensed the content of 'Dataset 2' through summarization and then we performed the fine-tuning process using this summarized dataset. By distilling the augmented information into concise summaries, the model becomes capable of learning from these condensed representations, enabling it to identify shared characteristics and distinctive

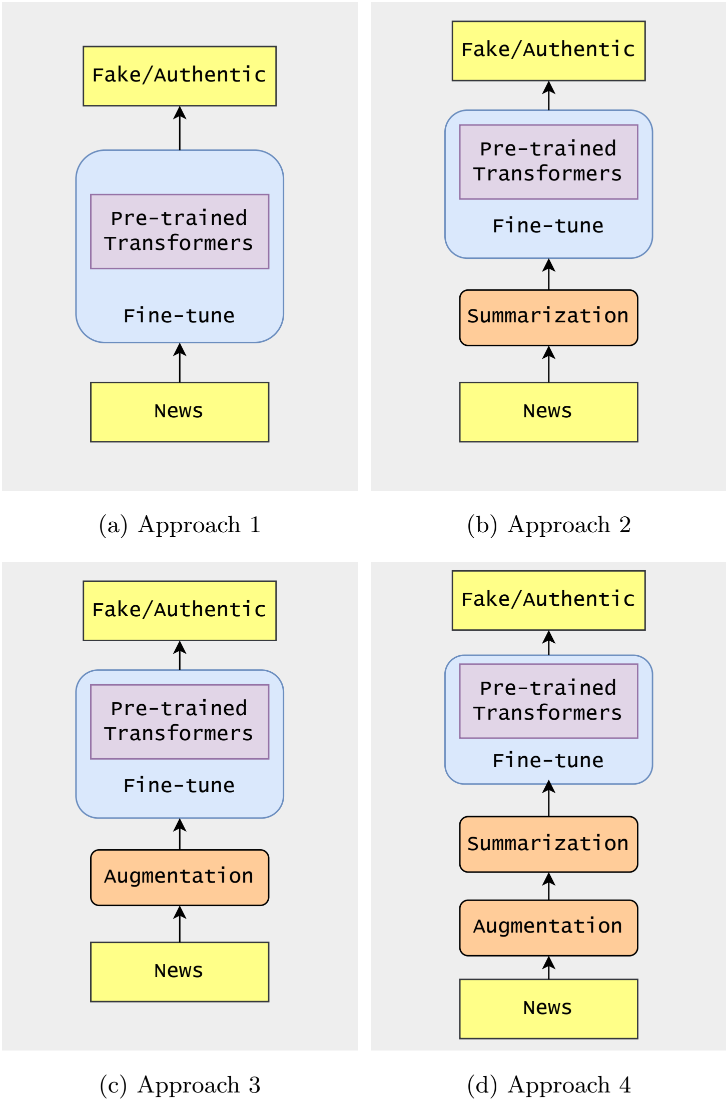



**Fig. 1**: Schematic diagram of the proposed approaches.

factors of fake news. This can enhance the capability of the model to detect fake news from previously unseen sources or variations of fake news.

## 4.2 Language Models

In this study, we fine-tuned few Bidirectional Encoder Representations from Transformers (BERT) [60] variant language models pre-trained on Bengali texts in order to detect fake news. BERT works bidirectionally and employs a Transformer [59] encoder architecture that includes multiple layers of self-attention mechanisms and feed-forward neural networks. By utilizing these layers, BERT can effectively understand the contextual connections between words within a sentence and produce high quality representations (embeddings) for each word. These embeddings enable BERT to make precise predictions by combining them with a task-specific architecture and further training. We have performed fine-tuning on five different variants of BERT models among which three are multi-lingual BERT models. We show the architectures of all the fine-tuned models in Table 5.

**(a) Multilingual BERT based.** Multilingual BERT (mBERT) language model was pre-trained on 104 languages using masked language modeling (MLM) objective. We fine-tuned the pre-trained base mBERT checkpoint[4] available in Hugging Face Transformers Library [61]. We also fine-tuned two other available mBERT checkpoints pre-trained on Bengali fake news detection task. We refer to those models as TM-mBERT[5] and DB-mBERT[6] in this study. TM-mBERT and DB-mBERT were developed by fine tuning the original base mBERT model using the *BanFakeNews* dataset.

**(b) BERT based.** To detect fake news, we fine-tuned two BERT variant Bengali language models: BanglaBERT and BanglaBERT Base. BanglaBERT [62] is pre-trained on a vast collection of Bengali texts using ELECTRA pre-training objective. On the other hand, BanglaBERT Base is pre-trained over two corpora: Bengali common crawl corpus and Bengali Wikipedia Dump Dataset. We fine-tuned the checkpoints[7,8] available in Hugging Face Library.

**Table 5**: Architectures of the pre-trained transformer models.

| | **TM mBERT** | **Bangla BERT** | **mBERT Base** | **DB mBERT** | **BanglaBERT Base** |
|---|---|---|---|---|---|
| Embedding size | 768 | 768 | 768 | 768 | 768 |
| No. of attention heads | 12 | 12 | 12 | 12 | 12 |
| No. of hidden layers | 12 | 12 | 12 | 12 | 12 |
| Max positional embedding | 512 | 512 | 512 | 512 | 512 |
| Vocabulary size | 119547 | 32000 | 119547 | 119547 | 102025 |

## 4.3 Augmentation

A popular technique in natural language processing (NLP) to increase the quantity of training data samples is to use *text augmentation*. Among several available techniques, we chose three different types of augmentation techniques - *token replacement, back translation* and *paraphrase generation.*

[4] https://huggingface.co/bert-base-multilingual-cased
[5] https://huggingface.co/Tahsin-Mayeesha/bangla-fake-news-mbert
[6] https://huggingface.co/DeadBeast/mbert-base-cased-finetuned-bengali-fakenews
[7] https://huggingface.co/csebuetnlp/banglabert
[8] https://huggingface.co/sagorsarker/bangla-bert-base

**Table 6**: Different augmentation techniques applied to a sample text.

| Text Process | Bengali Text | English Translated Text |
|---|---|---|
| Original | মুরগির হামলায় শেয়াল নিহত | Fox killed by chicken attack |
| Token replaced | মুরগির অভিযানে শেয়াল নিহত | Fox killed in chicken raid |
| Back translated | মুরগির আক্রমণে শেয়াল নিহত | Fox killed by chicken onslaught |
| Paraphrased | মুরগির আক্রমণে শিয়াল নিহত হয় | The fox was killed by the attack of the chicken |

Figure 2 illustrates the augmentation pipeline used in this study. We show the application of these augmentation techniques by taking an example sentence in Table 6.

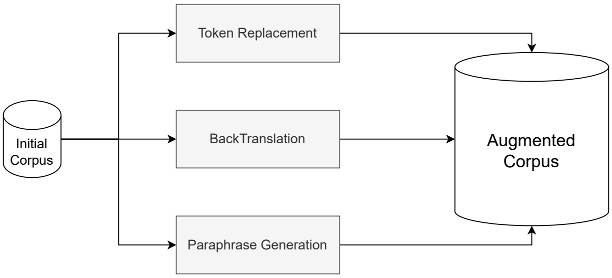



**Fig. 2**: Pipeline of the augmentation process.

**(a) Back translation:** Back translation method works as follows: we take some sentences (e.g. in Bengali) and translate to another language (e.g English). Then we translate the English sentences back to Bengali sentences. The final Bengali sentences will have subtle changes than the original Bengali sentences. We used the publicly available `bnaug`[9] library to perform back translation.
**(b) Paraphrase generation:** Paraphrase generation preserves the meaning of the sentence while changing the grammar and word choices. There are two ways to achieve this - Rule Based and machine learning based. In rule based method rules are created manually for example randomly selecting one or more words in sentences and changing them with words. This may also include changing active voice into passive, changing parts of speech etc. On the other hand, in machine learning based approach, paraphrase is automatically generated from data. Pre-trained seq2seq (sequence-to-sequence) language models are capable of generating paraphrases automatically. The `bnaug` library we used uses a BanglaT5 checkpoint[10] provided by Akil et al. [63] specifically pre-trained on Bengali paraphrasing task.

[9] https://github.com/sagorbrur/bnaug
[10] https://huggingface.co/csebuetnlp/banglat5_banglaparaphrase

**(c) Token replacement:** The last technique we used was token replacement using random masking. BERT models are pre-trained on a huge volume of texts using 'Masked Language Modelling' (MLM) objective where the model has to predict masked words based on a context. This can be used for augmentation where some tokens are randomly masked in a sentence and a language model has to predict the token for that mask. For token replacement, we used `nlpaug` library[11] with two BERT variations: BanglaBERT [62] and ShahajBERT[12].

### 4.4 Summarization

To tackle the issue of BERT models not being able to take token sequence length greater than 512, we summarized our news articles which exceeded the limit. Our hypothesis was that rather than taking top truncated part of the news, if summarized news was used for training, the models should perform better. As we know, sometimes large news articles start with irrelevant information. To summarize articles in our corpus, we used a mT5 checkpoint[13] pre-trained on multilingual summarization task provided by Hasan et al. [64] which generates the abstractive summarization. A base mT5 model was fine-tuned using one million annotated article-summary pairs covering 44 languages including Bengali. To make sure that summary contained texts from all parts of the article, we divided large articles into chunks, summarized those chunks and then merged the chunks to create a summarized text. The summarization pipeline is depicted in Figure 3. This was necessary as some of the articles were very large, even as large as 19,000 tokens. An example of summarization is provided in Table 7.

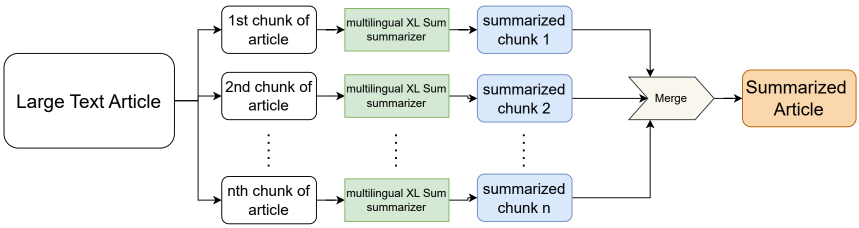



**Fig. 3**: Pipeline of the summarization process.

## 5 Experiments

In this section, we conduct a comprehensive evaluation to report the performance and effectiveness of the proposed approaches.

[11] https://github.com/makcedward/nlpaug
[12] https://huggingface.co/neuropark/sahajBERT
[13] https://huggingface.co/csebuetnlp/mT5_multilingual_XLSum

**Table 7**: Example of summarization of a sample news article instance. *B* represents Bengali and *E* represents the corresponding English translation.

| **Article** | **Summary** |
|---|---|
| কোকাকোলার কোমল পানীয়তে মিশ্রিত হতে যাচ্ছে গাঁজা! কোকা-কোলা এবার পানীয়তে গাজাঁর নির্যাস যোগ করতে যাচ্ছে। কোকা-কোলা বলছে, গাঁজা সংশ্লিষ্ট পানীয় বাজার পর্যবেক্ষণ করছে তারা। কানাডাভিত্তিক বিএনএন ব্লুমবার্গ টিভি এই তথ্য জানিয়েছে। এছাড়াও স্কাইনিউজ, ম্যাশেবল, এবিসি নিউজ সহ বিশ্বের প্রথম সারির গণমাধ্যম এ খবর প্রকাশ করেছে। বলা হচ্ছে, স্থানীয় উৎপাদক 'অরোরা ক্যানাবিস' এর সঙ্গে গাঁজার স্বাদযুক্ত কোমল পানীয় উৎপাদনের বিষয়ে অলোচনা করছে কোকা-কোলা। তবে গ্রাহকদের মাদকাসক্ত করতে নয়, তাদের শারীরিক যন্ত্রণা লাঘব করাই পানীয় তৈরিকারীদের লক্ষ্য। কোকা-কোলা এক বিবৃতিতে জানিয়েছে, অনেক উৎপাদকের মতো আমরাও পর্যবেক্ষণ করছি যে, কোমল পানীয় তৈরির ক্ষেত্রে নন-সাইকোঅ্যাক্টিভ ক্যানাবিডিওল বা চিত্ত উত্তেজিত করে না এমন গাঁজা জাতীয় দ্রব্যের ব্যবহার কতটা জনপ্রিয়তা পাচ্ছে। ক্যানাবিডিওল ক্যানাবিস বা গাঁজার একটি উপাদান, যা প্রদাহ, ব্যথা বা খিচুনির চিকিৎসার ক্ষেত্রে আরামদায়ক হতে পারে এবং এর কোনো চিত্ত উত্তেজক প্রভাব নেই। গাঁজার বিনোদনমূলক ব্যবহার আইনত বৈধ করতে যাচ্ছে কানাডা। চিকিৎসা কাজে অবশ্য অনেক আগে থেকেই গাঁজা বৈধ কানাডায়। কোকা-কোলার এ সিদ্ধান্তের ফলে কানাডায় গড়ে উঠেছে বিশাল আকারের গাঁজা শিল্প অন্যদিকে, উৎপাদনকারী সংস্থা মোলসন কুরস ব্রুয়িং বলেছে, তারা হাইড্রোপোথেক্যারি সংযোজন করে গাঁজা নিষিক্ত পানীয় তৈরি করবে। ইতোমধ্যে বিশ্বখ্যাত করোনা বিয়ার তৈরিকারী সংস্থা কনস্টেলেশান ব্র্যান্ডস গাঁজা উৎপাদনকারী প্রতিষ্ঠান ক্যানোপি গ্রোথের ওপর চার বিলিয়ন ডলার বিনিয়গ করেছে। কোকাকোলা আর অরোরা'র অংশীদারিত্বের ফলে গাঁজার পানীয়ের বাজারে প্রথম নন-অ্যালকোহলিক পানীয় হিসেবে যাত্রা শুরু হবে কোক'এর। জানা গেছে, অরোরা'র সঙ্গে কোকা-কোলা'র আলোচনা অনেকদুর অগ্রসর হলেও চুড়ান্ত কোনো চুক্তি হয়নি। সূত্রটি জানিয়েছে, এই পানীয়টি শুধু অবসাদই দূর করবে না , সতেজতা লাভেও সহয়তা করবে। আলাদা এক বিবৃতিতে অরোরা জানিয়েছে চুক্তি চূড়ান্ত না হওয়া পর্যন্ত তারা এ বিষয়ে বিস্তারিত কিছু জানাবে না। তবে তারা বলেছে, গাঁজা নিষিক্ত পানীয়ের বাজারে প্রবেশ করার বিষয়ে অরোরা যথেষ্ট আগ্রহী রয়েছে। | [ৰ] কোকাকোলার কোমল পানীয়তে মিশ্রিত হতে যাচ্ছে গাঁজা! এমন খবর প্রকাশ করেছে বিশ্বের প্রথম সারির গণমাধ্যম এবং সামাজিক যোগাযোগ মাধ্যমের বেশ কয়েকজন বিএনএন নিউজ স্ট্রিমিং এর সহযোগী সংগঠন 'কোকা-কোলা'। এখবর পাওয়া গেছে, যুক্তরাষ্ট্রের একটি উৎপাদকের সঙ্গে আলোচনা করার পর তারা বলছে যে, এই করোনাভাইরাসের সংক্রমণ ঠেকাতে যুক্তরাষ্ট্রের গাঁজা উৎপাদনকারী সংস্থা কোকা-কোলা আর ক্যানাডায় নতুন একটি নন-অ্যালকোহলিক পানীয় তৈরি করার বিষয়ে অরোরা'র অংশীদারিত্ব পাওয়ার সিদ্ধান্ত নিয়েছে দেশটির স্বাস্থ্য মন্ত্রণালয়ের সূত্র। কিন্তু এই চুক্তিতে কোক'এর সঙ্গে অলোচনা অগ্রসর হওয়ার পর তারা বলছে, এখন থেকে।<br><br>[E] Cannabis is going to be mixed with Coca-Cola's soft drinks! Such news has been revealed by the world's leading mass media and several social media, affiliated organization of BNN News Streaming - 'Coca-Cola'. It has been reported that, after discussing with a producer in the United States, they say that "to prevent the spread of this coronavirus, the American cannabis production company Coca-Cola and Aurora have decided to partner up in Canada to develop a new non-alcoholic drink, according to the country's Ministry of Health. But after the progress of negotiations with Coke on this deal, they say, from now on. |

## 5.1 Evaluation Metrics

For the purpose of evaluation, we have employed various standard metrics including accuracy, precision, recall, MCC (Matthews Correlation Coefficient), and ROC-AUC score. We provide a brief explanation of these metrics below. In the following equations, TP represents True Positives, TN represents True Negatives, FP represents False Positives, and FN represents False Negatives.

**Accuracy:** The ratio of correctly classified data and total number of data.

$$\text{accuracy} \;=\frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

**Precision:** The proportion of positive cases that were correctly classified.

$$\text{precision} \;=\frac{TP}{TP+FP} \tag{2}$$

**Recall:** The ratio of correctly classified positives and actual positives.

$$\text{recall} \;=\frac{TP}{TP+FN} \tag{3}$$

**MCC:** The Matthews correlation coefficient (MCC) is least influenced by imbalanced data. It is a correlation coefficient between the observed and predicted classifications. The value ranges from -1 to +1 with a value of +1 representing a perfect prediction, 0 as no better than random prediction and -1 the worst possible prediction.

$$MCC=\frac{TP\times TN-FP\times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{4}$$

**ROC-AUC Score:** The ROC or the Receiver Operating Characteristic curve calculates the sensitivity and specificity across a continuum of cutoffs. An appropriate balance can be determined between sensitivity and specificity using the curve. AUC is simply the area under the ROC curve. An area of 1 represents a perfect model and an area of 0.5 represents a worthless model.

## 5.2 Experimental Setup

To carry out all the experiments, we used Google Colaboratory platform with 12.7GB system RAM, 15GB google Tesla T4 GPU RAM and 78.2GB disk space. To process and prepare data we used `pandas` and `numpy` library. For implementation purpose, we used `python` language and transformer models were implemented using `Pytorch` framework. During the training and validation stages, we maintained an 85 : 15 ratio on the training datasets. For testing purposes, we created three separate test datasets.

## 5.3 Hyper-parameters Tuning

Hyper-parameters are parameters that are set before the training process begins such as the learning rate, batch size, loss function, activation function, the number of epochs etc. The hyper-parameters are adjusted until the best performance is achieved on the validation set. We report the hyper-parameter settings for all our experiments in Table 8. As our train and test datasets are balanced, we consider accuracy and f1-score as the primary metrics for assessing the superiority of the models. Since all classes have an equal number of samples, we specifically focus on the macro average. Moreover, we take note of other scores such as precision, recall, MCC, and ROC-AUC to gain a more comprehensive understanding of the performance of each model.

**Table 8**: Hyper-parameter settings

| **Hyper-parameter** | **Value** |
|---|---|
| Maximum Sequence Length | 512 |
| Epochs | 4 |
| Batch Size | 16 |
| Activation Function | Gelu, Softmax |
| Learning Rate | 2e-05 |
| Optimizer | AdamW |
| Loss Function | Binary Cross Entropy |

## 5.4 Evaluation

To assess the performance of the proposed approaches, we address the following research questions:

(a) **RQ1**: How does fine-tuning without summarization and augmentation perform when translated news (English-Bengali) are combined with Bengali news articles for the task of fake news classification?
(b) **RQ2**: What impact does introducing summarization before fine-tuning have on fake news detection?
(c) **RQ3**: How does the utilization of augmented fake data during fine-tuning affect fake news detection?
(d) **RQ4**: What is the influence of summarizing augmented news articles on fake news classification?

**RQ1: How does fine-tuning without summarization and augmentation perform when translated news (English-Bengali) are combined with Bengali news articles for the task of fake news classification?**

In *approach 1*, we focused on fine-tuning all the transformer models without the use of summarization or augmentation. During training, we incorporated translated fake news (English-Bengali) alongside Bengali news articles. The performance of these fine-tuned models was evaluated using 'Test DS1' and

'Test DS3'. We show the performance of the fine-tuned models on 'Test DS1' and 'Test DS3' in Table 9 and 11 respectively. The fine-tuned models in *approach 1* exhibited impressive performance on 'Test DS1', with the BanglaBERT model achieving the lowest accuracy of 88%. The mBERTBase model achieved the highest accuracy of 92%, while the other three models also scored above 90%. On the other hand, the BanglaBERT Base model attained the highest accuracy score of 80% on 'Test DS3'. This custom test dataset, designed to assess generalization performance, was completely unfamiliar to the fine-tuned models. The results obtained from *approach 1* indicate that it is feasible to develop a robust Bengali fake news detection model without relying on summarization or augmentation techniques, particularly when a substantial amount of fake data is available. In our case, we achieved this by translating fake English news into Bengali, thereby increasing the number of fake news instances. Overall, the BanglaBERT Base model demonstrated better performance on both test datasets.

**Table 9**: Performance comparison across all the approaches on the first test dataset 'Test DS1'. *Inference* represents performing classification without any training or fine-tuning.

| **Method** | **Model** | **A** | **P** | **R** | **F1** | **MCC** | **ROC** |
|---|---|---|---|---|---|---|---|
| Inference | TM-mBERT | 0.11 | 0.09 | 0.11 | 0.10 | -0.80 | 0.61 |
| | BanglaBERT | 0.50 | 0.25 | 0.5 | 0.33 | 0.00 | 0.50 |
| | mBERT-Base | 0.53 | 0.54 | 0.53 | 0.47 | 0.07 | 0.41 |
| | DB-mBERT | 0.11 | 0.09 | 0.11 | 0.10 | -0.79 | 0.57 |
| | BanglaBERT-Base | 0.52 | 0.66 | 0.52 | 0.39 | 0.12 | 0.50 |
| Approach 1 | TM-mBERT | 0.91 | 0.91 | 0.91 | 0.91 | 0.81 | 0.59 |
| | BanglaBERT | 0.88 | 0.89 | 0.88 | 0.88 | 0.77 | 0.59 |
| | mBERT-Base | 0.92 | 0.92 | 0.92 | 0.92 | 0.84 | 0.48 |
| | DB-mBERT | 0.91 | 0.91 | 0.91 | 0.91 | 0.82 | 0.55 |
| | BanglaBERT-Base | 0.90 | 0.90 | 0.90 | 0.90 | 0.80 | 0.49 |
| Approach 2 | TM-mBERT | 0.84 | 0.87 | 0.84 | 0.84 | 0.71 | 0.55 |
| | BanglaBERT | 0.76 | 0.80 | 0.76 | 0.75 | 0.56 | 0.56 |
| | mBERT-Base | 0.87 | 0.87 | 0.87 | 0.87 | 0.74 | 0.51 |
| | DB-mBERT | 0.88 | 0.89 | 0.88 | 0.88 | 0.77 | 0.57 |
| | BanglaBERT-Base | 0.88 | 0.88 | 0.88 | 0.88 | 0.76 | 0.51 |
| Approach 3 | TM-mBERT | 0.95 | 0.95 | 0.95 | 0.95 | 0.90 | 0.59 |
| | BanglaBERT | 0.85 | 0.87 | 0.86 | 0.84 | 0.71 | 0.43 |
| | mBERT-Base | 0.92 | 0.93 | 0.92 | 0.92 | 0.85 | 0.48 |
| | DB-mBERT | 0.94 | 0.94 | 0.94 | 0.94 | 0.89 | 0.54 |
| | BanglaBERT-Base | **0.96** | 0.96 | 0.96 | **0.96** | 0.71 | 0.43 |
| Approach 4 | TM-mBERT | 0.90 | 0.90 | 0.90 | 0.90 | 0.81 | 0.57 |
| | BanglaBERT | 0.91 | 0.91 | 0.91 | 0.91 | 0.83 | 0.49 |
| | mBERT-Base | 0.90 | 0.90 | 0.90 | 0.90 | 0.80 | 0.50 |
| | DB-mBERT | 0.89 | 0.89 | 0.89 | 0.89 | 0.79 | 0.56 |
| | BanglaBERT-Base | 0.93 | 0.93 | 0.93 | 0.93 | 0.86 | 0.48 |

**RQ2: What impact does introducing summarization before fine-tuning have on fake news detection?**

We introduced the utilization of summarization technique in *approach 2*. Analyzing the results presented in Table 9 and 11, we can observe that the BanglaBERT Base model achieved the highest accuracy scores of 88% and 78% on 'Test DS1' and 'Test DS3' respectively. The DB-mBERT model exhibited good performance on 'Test DS1' with an accuracy of 88%, but its performance on 'Test DS3' was comparatively weaker, with an accuracy of only 70%. When compared to the results of *approach 1*, we observed a slight decrease in performance after implementing the summarization technique in *approach 2*. However, this decrease is negligible, with only a marginal 1-2% reduction in accuracy.

**RQ3: How does the utilization of augmented fake data during fine-tuning affect fake news detection?**

In the third approach, we incorporated augmented fake news samples during the fine-tuning process and evaluated all the fine-tuned models using our three test datasets. It is important to note that 'Test DS2' was exclusively created for testing approaches involving augmentation, so it was not used to evaluate *approach 1* and *approach 2*. We only tested models in *approach 3* and *approach 4* where for fine-tuning we used the training dataset 'Dataset 2'. As we have not used any translated news articles during the training in these approaches, this test corpus is completely unknown for the models *approach 3* and *approach 4*. We show the performance comparison among the approaches on 'Test DS2' in Table 10. Considering the results on all three test datasets, the mBERT Base model emerged as the top performer in this approach. It achieved an accuracy score of 92% on both 'Test DS1' as well as 'Test DS2' and 86% on 'Test DS3'. The BanglaBERT Base model also demonstrated strong performance, achieving accuracy scores of 96%, 93% and 80% on 'Test DS1', 'Test DS2' and 'Test DS3' respectively. We observed that the utilization of augmented fake data significantly improved the performance of all the fine-tuned models compared to approaches 1 and 2. While all models exhibited good performance, the mBERT Base and BanglaBERT Base models outperformed the others in terms of accuracy.

**RQ4: What is the influence of summarizing augmented news articles on fake news classification?**

In the fourth approach, we introduced the technique of summarizing the augmented news articles. In this case, the BanglaBERT model exhibited exceptional performance across all three test datasets. It achieved accuracy scores of 91%, 97% and 82% on 'Test DS1', 'Test DS2', and 'Test DS3' respectively. This model demonstrated superior learning capabilities when trained on summarized data compared to the other four models. The BanglaBERT-Base model also performed well, achieving an accuracy of 93% on both 'Test DS1' as well

**Table 10**: Performance comparison between Approach 3 and 4 on the second test dataset 'Test DS2'. *Inference* represents performing classification without any training or fine-tuning.

| Method | Model | A | P | R | F1 | MCC | ROC |
|---|---|---|---|---|---|---|---|
| Inference | TM-mBERT | 0.37 | 0.22 | 0.37 | 0.28 | -0.37 | 0.61 |
| | BanglaBERT | 0.50 | 0.25 | 0.50 | 0.33 | 0.00 | 0.50 |
| | mBERT-Base | 0.51 | 0.51 | 0.51 | 0.45 | 0.02 | 0.41 |
| | DB-mBERT | 0.34 | 0.21 | 0.34 | 0.26 | -0.43 | 0.56 |
| | BanglaBERT-Base | 0.48 | 0.41 | 0.48 | 0.35 | -0.09 | 0.50 |
| Approach 3 | TM-mBERT | 0.89 | 0.90 | 0.89 | 0.89 | 0.80 | 0.63 |
| | BanglaBERT | 0.86 | 0.88 | 0.86 | 0.86 | 0.75 | 0.41 |
| | mBERT-Base | 0.92 | 0.93 | 0.92 | 0.92 | 0.85 | 0.48 |
| | DB-mBERT | 0.89 | 0.90 | 0.89 | 0.89 | 0.79 | 0.56 |
| | BanglaBERT-Base | 0.93 | 0.93 | 0.93 | 0.93 | 0.87 | 0.49 |
| Approach 4 | TM-mBERT | 0.91 | 0.91 | 0.91 | 0.91 | 0.82 | 0.51 |
| | BanglaBERT | **0.97** | 0.98 | 0.97 | **0.97** | 0.95 | 0.53 |
| | mBERT-Base | 0.95 | 0.95 | 0.95 | 0.95 | 0.50 | 0.90 |
| | DB-mBERT | 0.91 | 0.91 | 0.91 | 0.91 | 0.82 | 0.52 |
| | BanglaBERT-Base | 0.93 | 0.94 | 0.93 | 0.93 | 0.87 | 0.47 |

as 'Test DS2' and 80% on 'Test DS3'. Through empirical observation, we discovered that not all models adapt equally well to summarized training data. However, the BanglaBERT model showcased the highest performance when trained on summarized augmented data. When comparing these results with the previous approaches, it is evident that the summarization of augmented data positively influences the task of fake news classification.

## 5.5 Result Analysis

We analyze the results in two ways - quantitatively and qualitatively.

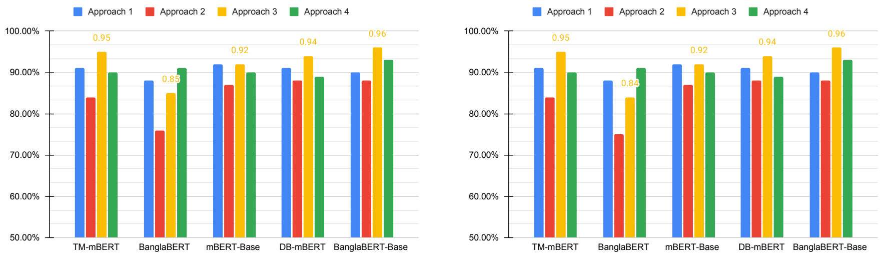
**(a) Quantitative Analysis**

This section seeks to experimentally support the models' performance. Among the three test datasets, we use the first two test datasets - 'Test DS1' and 'Test DS2' for quantitative analysis. Table 9 and 10 present the experimental findings that followed the evaluation of each model on these two test datasets. Figure 4 is the visual representation of the performance comparison among all the fine-tuned models across all the approaches on 'Test DS1'. As *approach 3* showed promising results on 'Test DS1', we decided to conduct further testing on this approach. In order to achieve a fair assessment, we excluded any samples from translated news articles during the training of models in *approach 3*. For this purpose, we created Test DS2', ensuring that the translated data remained unknown to the models in approach 3. This step allowed us to obtain a more comprehensive understanding of their true capability in detecting fake news. The performance comparison on 'Test DS2' is

**Table 11**: Performance comparison across all the approaches on the third dataset 'Test DS3'. *Inference* represents performing classification without any training or fine-tuning.

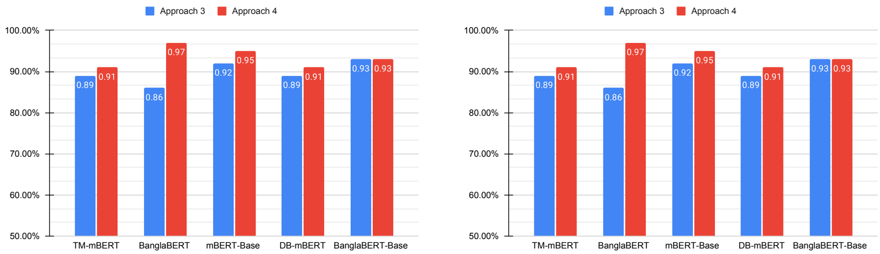
| Method | Model | A | P | R | F1 | MCC | ROC |
|---|---|---|---|---|---|---|---|
| Inference | TM-mBERT | 0.37 | 0.21 | 0.37 | 0.27 | -0.39 | 0.89 |
| | BanglaBERT | 0.50 | 0.25 | 0.50 | 0.33 | 0.00 | 0.50 |
| | mBERT-Base | 0.48 | 0.46 | 0.48 | 0.40 | -0.06 | 0.46 |
| | DB-mBERT | 0.38 | 0.22 | 0.38 | 0.27 | -0.37 | 0.88 |
| | BanglaBERT-Base | 0.47 | 0.37 | 0.47 | 0.34 | -0.13 | 0.54 |
| Approach 1 | TM-mBERT | 0.73 | 0.82 | 0.73 | 0.70 | 0.54 | 0.73 |
| | BanglaBERT | 0.76 | 0.83 | 0.76 | 0.75 | 0.59 | 0.66 |
| | mBERT-Base | 0.71 | 0.82 | 0.71 | 0.68 | 0.52 | 0.77 |
| | DB-mBERT | 0.70 | 0.79 | 0.70 | 0.68 | 0.48 | 0.79 |
| | BanglaBERT-Base | 0.80 | 0.84 | 0.80 | 0.79 | 0.64 | 0.63 |
| Approach 2 | TM-mBERT | 0.68 | 0.79 | 0.68 | 0.64 | 0.45 | 0.75 |
| | BanglaBERT | 0.73 | 0.82 | 0.73 | 0.71 | 0.54 | 0.83 |
| | mBERT-Base | 0.74 | 0.81 | 0.74 | 0.72 | 0.54 | 0.75 |
| | DB-mBERT | 0.70 | 0.79 | 0.70 | 0.67 | 0.48 | 0.80 |
| | BanglaBERT-Base | 0.78 | 0.83 | 0.78 | 0.77 | 0.60 | 0.84 |
| Approach 3 | TM-mBERT | 0.79 | 0.85 | 0.79 | 0.78 | 0.64 | 0.73 |
| | BanglaBERT | 0.71 | 0.71 | 0.71 | 0.71 | 0.41 | 0.52 |
| | mBERT-Base | **0.86** | 0.88 | 0.86 | **0.86** | 0.74 | 0.83 |
| | DB-mBERT | 0.78 | 0.85 | 0.79 | 0.78 | 0.64 | 0.73 |
| | BanglaBERT-Base | 0.80 | 0.83 | 0.80 | 0.79 | 0.63 | 0.81 |
| Approach 4 | TM-mBERT | 0.74 | 0.83 | 0.74 | 0.72 | 0.56 | 0.83 |
| | BanglaBERT | 0.82 | 0.86 | 0.82 | 0.81 | 0.68 | 0.75 |
| | mBERT-Base | 0.75 | 0.80 | 0.75 | 0.75 | 0.55 | 0.78 |
| | DB-mBERT | 0.74 | 0.80 | 0.74 | 0.74 | 0.54 | 0.73 |
| | BanglaBERT-Base | 0.80 | 0.85 | 0.80 | 0.80 | 0.65 | 0.70 |

illustrated in Figure 5. The results of our experiments demonstrate that both *approach 3* and *approach 4* perform exceptionally well in detecting fake news. *Approach 4* achieved an impressive accuracy score of 97%. It is worth noting that the incorporation of summarization improves the results for all models in *approach 4*, except for the BanglaBERT Base model which maintains the same level of performance. In *approach 1*, where 'Dataset-1' was used for fine-tuning, we observed that summarization does not significantly enhance the performance in most cases. *Approach 1* outperforms *approach 2* on 'Test DS1' and on 'Test DS3'. TM-mBERT, BanglaBERT, and BanglaBERT Base models performed worse in *approach 2*. However, DB-mBERT and mBERT Base models either perform equally or better. Comparing *approach 1* and *approach 3*, we find that *approach 3* generally performs better, except for the BanglaBERT model, which performs better in *approach 1* on both 'Test DS1' and 'Test DS3'. On the other hand, *approach 4* performed almost equally or better than *approach 1* on both 'Test DS1' and 'Test DS3'. Notably,




(a) Comparison in terms of accuracy. (b) Comparison in terms of F1-score.

**Fig. 4**: Visual representation of the performance comparison on 'Test DS1'.



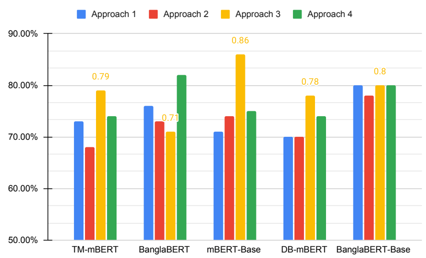

(a) Comparison in terms of accuracy. (b) Comparison in terms of F1-score.

**Fig. 5**: Visual representation of the performance comparison on 'Test DS2'.

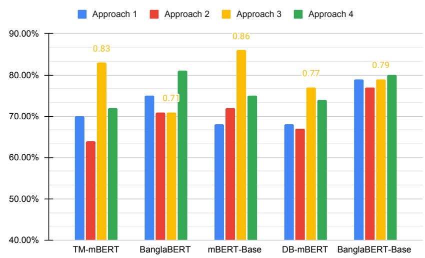
*approach 3* exhibits significantly better performance than *approach 2*. When considering the impact of summarization, we observe that it decreases the performance in *approach 4* compared to *approach 3* on 'Test DS1' and 'Test DS3'. However, the BanglaBERT model is an exception to this trend, as it performs better after summarization on both of these test datasets. We conducted further testing with 'Test DS2', where summarization consistently improves performance. In *approach 4*, all models perform either equally or better than *approach 3* on 'Test DS2'. Based on the discussions, we conclude that all approaches exhibit good accuracy in detecting fake news. *Approach 3* which incorporates augmentation, appears to perform better than the other approaches, but *approach 4* with the combination of augmentation and summarization, shows promising results.

**(b) Qualitative Analysis**

In this section, we evaluate the generalization ability of the model towards real-world test samples using 'Test DS3'. The results of the experiments on Test DS3 are presented in Table 11. Even on this completely unseen test dataset, *approach 3* outperforms the other three approaches and achieves the highest accuracy and F1 score of 0.86. The equality of the accuracy and F1 score indicates that the model is capable of correctly classifying both positive and negative classes. Figure 6 provides a visual representation of the performance comparison of all models on 'Test DS3'. From the figure, we observe




(a) Comparison in terms of accuracy.




(b) Comparison in terms of F1-score.

**Fig. 6**: Visual representation of the performance comparison on 'Test DS3'.

that the five transformer models we utilized exhibit varying performance characteristics. The TM-mBERT transformer model shows better performance without the use of summarization. It performs better in *approach 3* compared to *approach 1* indicating the effectiveness of the augmented training dataset. The BanglaBERT model performs exceptionally well in *approach 4*, where we employed summarization on the augmented dataset. The mBERT Base model excels in *approach 3* where the utilization of summarization is applied. However, we observe that summarization leads to a decrease in performance for this transformer model. Similarly, summarization adversely affects the performance of the DB-mBERT model. This model performs best in *approach 3*, but its performance decreases when summarization is implemented. On the other hand, BanglaBERT Base model displays a well-balanced performance across different strategies. It performs best in *approach 3* but it also achieves satisfactory scores in the other approaches.

## 5.6 Comparison with Existing Models

To evaluate and compare the performance of our proposed approaches, we conducted direct inference on two existing pre-trained models: TM-mBERT and DB-mBERT which are specifically designed for detecting Bengali fake news. Table 12 shows result comparison. We use 'Test DS3' for performance comparison between existing models and our approaches, because this test dataset is unknown to all the models. All our approaches performed better than the existing pre-trained models.

| **Models** | **Accuracy** | **F1-score** |
|---|---|---|
| TM-mBERT (inference) | 0.37 | 0.27 |
| DB-mBERT (inference) | 0.38 | 0.27 |
| BanglaBERT Base (Approach 1) | 0.80 | 0.79 |
| BanglaBERT Base (Approach 2) | 0.78 | 0.77 |
| mBERT Base (Approach 3) | **0.86** | **0.86** |
| BanglaBERT (Approach 4) | 0.82 | 0.81 |

**Table 12**: Performance comparison of our approaches with existing pre-trained models on the third test dataset 'Test DS3'.

# 6 Conclusion and Future Work

In this paper, we presented our approach to classify Bengali fake news using a combination of summarization and augmentation techniques with pre-trained language models. To ensure a thorough evaluation of our trained models, we conducted tests on three distinct test datasets. The results revealed that our models achieved remarkably high levels of accuracy and f1-score on the first two test datasets. For the third test dataset, which was kept separate to assess generalization, the best model demonstrated an accuracy and f1-score of 86%. These findings demonstrate the effectiveness of the model in accurately distinguishing between fake and authentic news articles. There are several promising directions for future work that can expand upon our research and broaden the scope of this paper. While our primary focus was on the classification of Bengali fake news, our approach holds potential for application in other languages with limited resources. Furthermore, our research primarily centered around the binary classification of authentic and fake news articles. To further advance the field, future studies could explore the realm of multi-class classification, involving the categorization of news articles into different types of fake news, including satire, propaganda, and clickbait.

# Statements and Declarations

## 6.1 Ethical Approval and Consent to participate

Not applicable.

## 6.2 Consent for publication

Not applicable.

## 6.3 Human and Animal Ethics

Not applicable.

## 6.4 Availability of data

The datasets generated during and/or analysed during the current study are available at - https://github.com/arman-sakif/Bengali-Fake-News-Detection

## 6.5 Code availability

The implementations can be found at - https://github.com/arman-sakif/Bengali-Fake-News-Detection

## 6.6 Competing interests

The authors have no competing interests to declare that are relevant to the content of this article.

## 6.7 Funding

No funding was received for conducting this study.

# References

[1] McGarrigle, J.: Explained: What is Fake news? | Social Media and Filter Bubbles. https://www.webwise.ie/teachers/what-is-fake-news/. [Online; accessed 2023-05-29] (2018)

[2] SOLL, J., White, J.B., Sutton, S., Sitrin, C., Mahoney, B., Gerstein, J., , P.: The Long and Brutal History of Fake News. http://politi.co/2FaV5W9 (2016)

[3] Sundar, S.S., Molina, M.D., Cho, E.: Seeing Is Believing: Is Video Modality More Powerful in Spreading Fake News via Online Messaging Apps? Journal of Computer-Mediated Communication **26**(6), 301–319 (2021) https://academic.oup.com/jcmc/article-pdf/26/6/301/41139661/zmab010.pdf. https://doi.org/10.1093/jcmc/zmab010

[4] Vosoughi, S., Roy, D., Aral, S.: The spread of true and false news online. science **359**(6380), 1146–1151 (2018)

[5] Barthel, M.: Many Americans Believe Fake News Is Sowing Confusion. https://www.pewresearch.org/journalism/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/. [Online; accessed 2023-05-29] (2016)

[6] Shishir, Q.: Reports misleadingly claim Saudi Arabia plans to 'remove Islamic oath from national flag'. https://factcheck.afp.com/doc.afp.com.9YR4EL. [Online; accessed 2023-05-29] (2022)

[7] Shishir, Q.: Bangladeshi media's troubling tango with fake news. https://www.thedailystar.net/views/opinion/news/bangladeshi-medias-troubling-tango-fake-news-2985186. [Online; accessed 2023-05-29] (2022)

[8] Check Team, B.F.: ২০২০: কোন সংবাদমাধ্যমে কতটি ভুয়া খবর প্রকাশিত হয়েছে? https://www.boombd.com/fact-file/fake-and-misleading-news-in-bangladeshi-news-outlets-in-2020-11830. [Online; accessed 2023-05-29] (2021)

[9] Jensen, A.: 4 dead, 26 hospitalized in Arizona after ingesting hand sanitizer with methanol. https://www.azcentral.com/story/news/local/arizona-health/2020/07/17/people-dead-hospitalized-drinking-hand-sanitizer-methanol-arizona/5448890002/. [Online; accessed 2023-05-29]

[10] Spring, M.: Coronavirus: The human cost of virus misinformation. https://www.bbc.com/news/stories-52731624. [Online; accessed 2023-05-29]

[11] Stecula, D.: Analysis | Fake news might be harder to spot than most people believe. https://www.washingtonpost.com/news/monkey-cage/wp/2017/07/10/fake-news-might-be-harder-to-spot-than-most-people-believe. [Online; accessed 2023-05-29] (2017)

[12] Perkins, D.: Confirmation bias - Wikipedia. https://en.wikipedia.org/wiki/Confirmation_bias. [Online; accessed 2023-04-11] (2010)

[13] Hansen, C., Hansen, C., Chaves Lima, L.: Automatic fake news detection: Are models learning to reason? In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers), pp. 80–86. Association for Computational Linguistics, Online (2021). https://doi.org/10.18653/v1/2021.acl-short.12

[14] Shu, K., Sliva, A., Wang, S., Tang, J., Liu, H.: Fake news detection on social media: A data mining perspective. ACM SIGKDD explorations newsletter **19**(1), 22–36 (2017)

[15] Zhang, X., Ghorbani, A.A.: An overview of online fake news: Characterization, detection, and discussion. Information Processing & Management **57**(2), 102025 (2020)

[16] Shu, K., Wang, S., Liu, H.: Beyond news content: The role of social context for fake news detection. In: Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining. WSDM '19, pp. 312–320. Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3289600.3290994

[17] Pérez-Rosas, V., Kleinberg, B., Lefevre, A., Mihalcea, R.: Automatic

detection of fake news. In: Proceedings of the 27th International Conference on Computational Linguistics, pp. 3391–3401. Association for Computational Linguistics, Santa Fe, New Mexico, USA (2018)

[18] Wang, W.Y.: "liar, liar pants on fire": A new benchmark dataset for fake news detection. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pp. 422–426. Association for Computational Linguistics, Vancouver, Canada (2017). https://doi.org/10.18653/v1/P17-2067

[19] Ruchansky, N., Seo, S., Liu, Y.: Csi: A hybrid deep model for fake news detection. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, pp. 797–806 (2017)

[20] Martínez-Gallego, K., Álvarez-Ortiz, A.M., Arias-Londoño, J.D.: Fake news detection in spanish using deep learning techniques. arXiv preprint arXiv:2110.06461 (2021)

[21] Meddeb, P., Ruseti, S., Dascalu, M., Terian, S.-M., Travadel, S.: Counteracting french fake news on climate change using language models. Sustainability **14**(18), 11724 (2022)

[22] Beauvais, C.: Fake news: Why do we believe it? Joint Bone Spine **89**(4), 105371 (2022)

[23] Liu, Z., Shabani, S., Balet, N.G., Sokhn, M.: Detection of satiric news on social media: Analysis of the phenomenon with a french dataset. In: 2019 28th International Conference on Computer Communication and Networks (ICCCN), pp. 1–6 (2019). IEEE

[24] Andriani, M., Kalsum, A.F., Elloianza, G.N.: Social discourse of fake news in french and its digital social media literacy. Lingua Cultura **16**(1), 105–115 (2022)

[25] Mahlous, A.R., Al-Laith, A.: Fake news detection in arabic tweets during the covid-19 pandemic. International Journal of Advanced Computer Science and Applications **12**(6), 778–788 (2021)

[26] Yang, K.-C., Niven, T., Kao, H.-Y.: Fake news detection as natural language inference. In: Proceedings of the 12th ACM International Conference on Web Search and Data Mining. WSDM '19. ACM, New York, NY, USA (2019). https://doi.org/10.1145/1122445.1122456

[27] Zhang, X., Cao, J., Li, X., Sheng, Q., Zhong, L., Shu, K.: Mining dual emotion for fake news detection. In: Proceedings of the Web Conference 2021, pp. 3465–3476 (2021)

[28] Zhang, Q., Guo, Z., Zhu, Y., Vijayakumar, P., Castiglione, A., Gupta, B.B.: A deep learning-based fast fake news detection model for cyber-physical social services. Pattern Recognition Letters **168**, 31–38 (2023)

[29] Chu, S.K.W., Xie, R., Wang, Y.: Cross-language fake news detection. Data and Information Management **5**(1), 100–109 (2021)

[30] Chen, M.-Y., Lai, Y.-W., Lian, J.-W.: Using deep learning models to detect fake news about covid-19. ACM Transactions on Internet Technology (2022)

[31] Du, J., Dou, Y., Xia, C., Cui, L., Ma, J., Philip, S.Y.: Cross-lingual covid-19 fake news detection. In: 2021 International Conference on Data Mining Workshops (ICDMW), pp. 859–862 (2021). IEEE

[32] Zhu, Y., Sheng, Q., Cao, J., Li, S., Wang, D., Zhuang, F.: Generalizing to the future: Mitigating entity bias in fake news detection. In: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 2120–2125 (2022)

[33] HaCohen-Kerner, Y., Manor, N., Bashan, N., Dimant, E.: Detecting fake news in urdu using classical supervised machine learning methods and word/char n-grams. In: CEUR Workshop Proceedings (2021)

[34] Kumar, S., Singh, T.D.: Fake news detection on hindi news dataset. Global Transitions Proceedings **3**(1), 289–297 (2022)

[35] Abd Rahim, N., Basri, M.: Malcov: Covid-19 fake news dataset in the malay language. In: 2022 International Visualization, Informatics and Technology Conference (IVIT), pp. 239–244 (2022). IEEE

[36] Vogel, I., Jiang, P.: Fake news detection with the new german dataset "germanfakenc". In: Digital Libraries for Open Knowledge: 23rd International Conference on Theory and Practice of Digital Libraries, TPDL 2019, Oslo, Norway, September 9-12, 2019, Proceedings 23, pp. 288–295 (2019). Springer

[37] Kuzmin, G., Larionov, D., Pisarevskaya, D., Smirnov, I.: Fake news detection for the russian language. In: Proceedings of the 3rd International Workshop on Rumours and Deception in Social Media (RDSM), pp. 45–57 (2020)

[38] Silva, R.M., Santos, R.L., Almeida, T.A., Pardo, T.A.: Towards automatically filtering fake news in portuguese. Expert Systems with Applications **146**, 113199 (2020)

[39] Isa, S.M., Nico, G., Permana, M.: Indobert for indonesian fake news

detection. ICIC Express Letters **16**(3), 289–297 (2022)

[40] Murayama, T., Hisada, S., Uehara, M., Wakamiya, S., Aramaki, E.: Annotation-scheme reconstruction for "fake news" and Japanese fake news dataset. In: Proceedings of the Thirteenth Language Resources and Evaluation Conference, pp. 7226–7234. European Language Resources Association, Marseille, France (2022)

[41] Taskin, S.G., Kucuksille, E.U., Topal, K.: Detection of turkish fake news in twitter with machine learning algorithms. Arabian Journal for Science and Engineering **47**(2), 2359–2379 (2022)

[42] Yun, T.-U., Ahn, H.: Fake news detection for korean news using text mining and machine learning techniques. Journal of Information Technology Applications and Management **25**(1), 19–32 (2018)

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

[44] Sharma, A.S., Mridul, M.A., Islam, M.S.: Automatic detection of satire in bangla documents: A cnn approach based on hybrid feature extraction model. In: 2019 International Conference on Bangla Speech and Language Processing (ICBSLP), pp. 1–5 (2019). IEEE

[45] Keya, A.J., Wadud, M.A.H., Mridha, M., Alatiyyah, M., Hamid, M.A.: Augfake-bert: Handling imbalance through augmentation of fake news using bert to enhance the performance of fake news classification. Applied Sciences **12**(17), 8398 (2022)

[46] Khan, J.Y., Khondaker, M.T.I., Afroz, S., Uddin, G., Iqbal, A.: A benchmark study of machine learning models for online fake news detection. Machine Learning with Applications **4**, 100032 (2021). https://doi.org/10.1016/j.mlwa.2021.100032

[47] Baarir, N.F., Djeffal, A.: Fake news detection using machine learning. In: 2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH), pp. 125–130 (2021). IEEE

[48] Hussain, M.G., Hasan, M.R., Rahman, M., Protim, J., Al Hasan, S.: Detection of bangla fake news using mnb and svm classifier. In: 2020 International Conference on Computing, Electronics & Communications Engineering (iCCECE), pp. 81–85 (2020). IEEE

[49] Sraboni, T., Uddin, M., Shahriar, F., Rizon, R.A., Polock, S.I.S., et al.:

Fakedetect: Bangla fake news detection model based on different machine learning classifiers. PhD thesis, Brac University (2021)

[50] Mugdha, S.B.S., Kuddus, M.B.M.M., Salsabil, L., Anika, A., Marma, P.P., Hossain, Z., Shatabda, S.: A gaussian naive bayesian classifier for fake news detection in bengali, 283–291 (2021). Springer

[51] Piya, F.L., Karim, R., Arefin, M.S.: Bdfn: A bilingual model to detect online fake news using machine learning technique. In: Soft Computing for Security Applications: Proceedings of ICSCS 2021, pp. 799–816 (2022). Springer

[52] Ahmad, I., Yousaf, M., Yousaf, S., Ahmad, M.O.: Fake news detection using machine learning ensemble methods. Complexity **2020**, 1–11 (2020)

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

[54] Bahad, P., Saxena, P., Kamal, R.: Fake news detection using bi-directional lstm-recurrent neural network. Procedia Computer Science **165**, 74–82 (2019)

[55] Ruchansky, N., Seo, S., Liu, Y.: Csi: A hybrid deep model for fake news detection. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. CIKM '17, pp. 797–806. Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3132847.3132877

[56] Nasir, J.A., Khan, O.S., Varlamis, I.: Fake news detection: A hybrid cnn-rnn based deep learning approach. International Journal of Information Management Data Insights **1**(1), 100007 (2021)

[57] Saha, S., Sarker, A., Chakraborty, P., Yousuf, M.A.: Bengali fake news detection: Transfer learning based technique with masked lm process by bert, 83–96 (2022). Springer

[58] Hasan, T., Bhattacharjee, A., Samin, K., Hasan, M., Basak, M., Rahman, M.S., Shahriyar, R.: Not low-resource anymore: Aligner ensembling, batch filtering, and new datasets for Bengali-English machine translation. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 2612–2623. Association for Computational Linguistics, Online (2020). https://doi.org/10.18653/v1/2020.emnlp-main.207

[59] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez,

A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems **30** (2017)

[60] Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: BERT: Pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pp. 4171–4186. Association for Computational Linguistics, Minneapolis, Minnesota (2019). https://doi.org/10.18653/v1/N19-1423

[61] Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., *et al.*: Transformers: State-of-the-art natural language processing. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, pp. 38–45 (2020)

[62] Bhattacharjee, A., Hasan, T., Ahmad, W., Mubasshir, K.S., Islam, M.S., Iqbal, A., Rahman, M.S., Shahriyar, R.: BanglaBERT: Language model pretraining and benchmarks for low-resource language understanding evaluation in Bangla. In: Findings of the Association for Computational Linguistics: NAACL 2022, pp. 1318–1327. Association for Computational Linguistics, Seattle, United States (2022). https://doi.org/10.18653/v1/2022.findings-naacl.98

[63] Akil, A., Sultana, N., Bhattacharjee, A., Shahriyar, R.: BanglaParaphrase: A high-quality Bangla paraphrase dataset. In: Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 2: Short Papers), pp. 261–272. Association for Computational Linguistics, Online only (2022)

[64] Hasan, T., Bhattacharjee, A., Islam, M.S., Mubasshir, K., Li, Y.-F., Kang, Y.-B., Rahman, M.S., Shahriyar, R.: XL-sum: Large-scale multilingual abstractive summarization for 44 languages. In: Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, pp. 4693–4703. Association for Computational Linguistics, Online (2021). https://doi.org/10.18653/v1/2021.findings-acl.413