

## **INTRO :**

Assalamualaikum and good morning to everyone. Our research topic is comprehensive study on Bangla misinformation, disinformation, satire and fake news. We are working under the supervision of Mahruba Mam. I am Nasrullah Sami. I am ...

## **OUTLINE :**

Let's briefly talk about what is ahead. First we will discuss some basic terminologies. Then we will show some related research works and show the gaps that has not been solved. Then we will discuss our approach to solve those problems. Finally we will describe our data collection, proposed method and future planning.

## **Introduction :**

After the introduction of The internet in our day to day life it has significantly changed the way people receive news. Nowadays radio sets are on the verge of extinction. Young people hardly sit in front of television. Most of our news come from social media such as Facebook, Twitter, Youtube etc. But this raises a big question. In social media anyone can spread any type of information. How do we differentiate which news are authentic and which are fake? For this reason, information disorder has been an important issue in recent years.

But how many of us are aware of the types of fake news? For example, what is misinformation? Or what comes to our mind

when we say disinformation? How do we differentiate between these concepts?

Misinformation is misleading information that is spread unintentionally. For example, during covid-19, 'black cumin seeds' can cure covid, such information was circulating throughout social media which had no scientific background. People believed it to be true and spread it as misinformation. (book)

Disinformation is wrong information that is spread intentionally to deceive or achieve certain goals. For example, during the 2016 election of the United States, Russia launched a disinformation campaign in order to bias the outcome of the election. (book)

Satire is a way of criticizing people or ideas in a humorous way, especially in order to make a political point, or a piece of writing that uses this style. (Cambridge)

## **Comprehension with examples :**

Let's look at some more examples.

"Eat..." - this is wrong information which someone believed to be true and spreads as misinformation.

"A celebrity..." - this type of information is spread to defame public figures. And when it is fake, it falls under the category of disinformation.

"A town plans..." - this is a humorous way of criticizing the planning of the town which is called satire.

## **Objectives :**

The Objective of our research paper is to create a new benchmark multi-labelled dataset. In order to do that, we will collect existing data, and merge them with our newly collected data. We will implement our proposed model which is based on Siamese Neural Network.

Due to the lack of classified fake news, our dataset might turn out to be imbalanced. We aim to achieve formidable accuracy overcoming this problem.

## **Related Works :**

There has been a number of work done on fake news detection. We have mentioned some that are most relevant to our work. BanFakeNews dataset contains almost 50k Bangla real news and Bangla 1k fake news. There has been research on Bangla satire detection and multiple researches on Bangla fake news detection using different models. Majority of the research work is based on BanFakeNews dataset.

## **Gaps in Literature :**

As we have discussed before, fake news can be classified into further sub-categories. But all the existing works on Bangla fake news uses binary classification that detects fake or not-fake, satire or not-satire etc.

Existing Bangla fake news datasets have less number of fake news due to the scarcity of fake news collection. But the number of authentic news is a lot which results in imbalanced datasets.

Traditional model performances on imbalanced datasets are not always reliable.

## **Our Approach:**

News can be either fake or real. We will classify fake news into 3 classes: misinformation, disinformation and satire. And the 4th class will be Real.

We will implement a model based on Siamese Neural Network that gives more accurate results even on imbalanced datasets.

## **Why Classify:**

One question may arise why do we want to classify fake news.

1. Different types of fake news have different mitigation approaches.

For misinformation, education and fact-checking are effective.

For disinformation, strategies for countering malicious intent has to be developed.

For satire, it is not important to overreact.

2. Fake information classification will further help social media platforms to improve their content moderation algorithms and limit the spread of fake news.

3. Researchers and analysts can use classification of fake news to better understand its origins, motivations and effects.

## **Data Collection:**

We have collected data from multiple research works and merged them together. We are also collecting new fake news data and classifying them into the aforementioned 4 categories. So far we have collected 67428 real news, 5540 fake news and 2773 satire news. Each record of our dataset contains 3 attributes : title of the article, content of the article, and classification label.

## **Methodology:**

**Embedding** - Each word is transformed into a dense vector representation. This layer learns to map words to dense vectors where semantically similar words are close to each other.

**Encoding** - This layer captures the contexts of the sequence from the embeddings. This layer is designed to extract meaningful features that capture the textual content's semantics.

**Global max pooling:** It extracts the most important features from the entire sequence by selecting the maximum value from each feature dimension and captures the most relevant information.

**Fully Connected Layer:** It is a dense neural network with multiple hidden layers which represents the higher level representation that captures the relationship between texts.

**Hidden Representation:** In this layer the result is represented in lower dimension and more abstract form to capture important features. It is a compact form of the features the network has learned.

**Loss function using similarity matrices:** Similarity matrix is computed from the hidden representations of all pairs of classes. The similarity matrix is then used as the input for loss function. We will use contrastive loss function which minimizes the distances between similar pairs and maximizes the distances between dissimilar pairs.

**Multiclass softmax layer:** It produces class probabilities for each of 4 classes : misinformation, disinformation, satire and real. The class with the highest probability is predicted as the final class label for the article.

**Future Works:**