Deliverable #1 - Fake News Detection

Eya Ibrahim (261113450)

David Medcalfe (260639056)

Maëlle Guislain (261070058)

MAIS 202 - Introduction to Accelerated Machine Learning

Assigned TPM: Emma Kondrup

McGill University

October 8, 2024

Word Count: 589 words

Deliverable Description

1. Choice of Dataset

We have selected the WELFake dataset from Kaggle, which contains 72,134 news articles, with 35,028 labeled as real and 37,106 labeled as fake. The dataset merges four popular news datasets (Kaggle, McIntire, Reuters, and BuzzFeed Political), providing a diverse and robust set of text data to train our machine learning model. This dataset is ideal for our project as it includes both titles and full-text articles, offering rich content for classification. Additionally, the large size of the dataset ensures that our model will be less prone to overfitting, allowing for better generalization.

You can find the dataset here:

https://www.kaggle.com/datasets/saurabhshahane/fake-news-classification

2. Methodology

First, we will preprocess the data by cleaning the text, removing stop words, punctuation, and converting the text to lowercase for consistency. We will then tokenize the text to break it into individual words and apply TF-IDF to convert it into a numerical format that can be processed by machine learning models.

For the machine learning model, we will start with logistic regression, a simple yet effective approach for classification problems. We will also test Naive Bayes to compare results and determine which model performs better for detecting fake news.

To evaluate the models, we will use metrics such as accuracy, precision, and recall to measure how well the model distinguishes between real and fake news. Our goal is to develop a straightforward yet effective solution to classify news articles.

a. Data Preprocessing

The WELFake dataset is feasible for our project, as it is well-structured and contains labeled real and fake news articles. The most useful information is the Title, Text, and Label columns, which provide the content and classification for each article.

We will preprocess the data by cleaning the text (removing stop words, punctuation, and converting to lowercase), tokenizing it into individual words, and applying TF-IDF to transform it into a numerical format for use in machine learning models.

b. Machine Learning Model

We want to predict if a news article is real or fake. We will start with a basic method called logistic regression, which is good for sorting things into two categories, like real or fake. We will also try other models like Naive Bayes and SVM to compare results.

Factors like the author, website, and content will be key in making these predictions. If time allows, we might explore more advanced methods like BERT for better accuracy. We will evaluate how well the models perform using precision, recall, and F1-score.

c. Evaluation Metric

We will evaluate our models using a confusion matrix to visualize the number of correct and incorrect predictions. Key metrics will include accuracy, which measures the overall correctness of the model, and precision and recall, which help assess how well the model handles fake news detection. We will also use F1-score to balance precision and recall, and logistic loss to evaluate how confident the model is in its predictions.

3. Application

We will integrate our machine learning model into a simple web-based landing page. Users will provide inputs such as a website link, headline, or social media post description via a text box.

The model will analyze the text and output a result, showing whether the content is "Real" or "Fake". This output will be displayed clearly on the page along with a confidence score (e.g., "85% likely fake").

If the content is real, we may offer links to credible sources. If the content is fake, we will suggest alternative, reliable sources to provide accurate information.