

FENCE: Fake News Classifier

CSCE 590: Project Presentation

Problem Statement

With about **5 billion internet users** out of which about a **billion are active social media users**, it becomes an easy platform to inform the population with an information. At the same time, it makes it easier to influence them with wrong information or fake news.

Hence, detecting fake news becomes an important task to contain the veracity (truthfulness) of the online data.

Intended Users

Social Media Platforms. While various social media platforms are spending a high amount of time and research to **tackle the issue of fake news**. The proposed algorithm can help them tackle the problem easily.

Dataset Used

1. Tweets

- a. COVID-19 related tweets
- b. Syria War related tweets
- c. Political related tweets

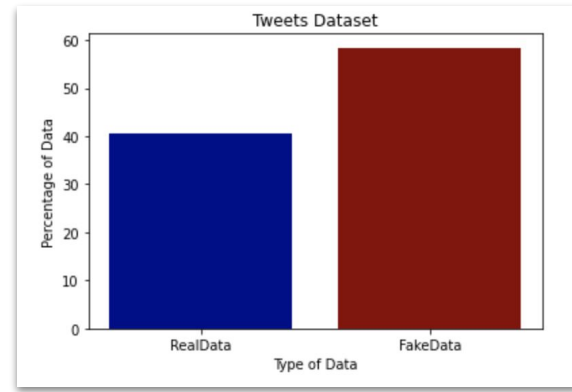
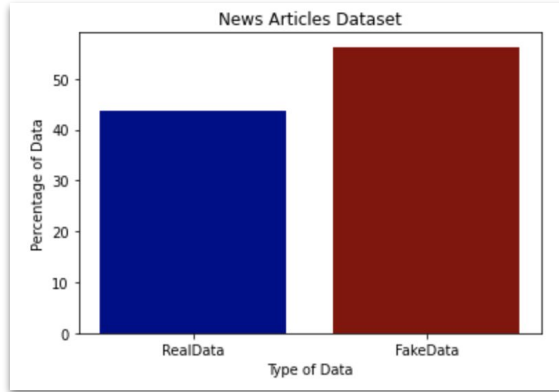
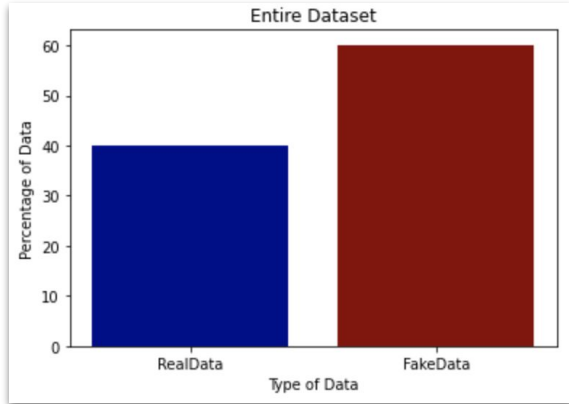
2. News Articles

- a. Politics related News Headlines
- b. World Events related News Headlines
- c. Government Policies related News Headlines

Fake News: “A post claims compulsory vaccination violates the principles of bioethics, that coronavirus doesn’t exist, that the PCR test returns many false positives, and that influenza vaccine is related to COVID-19”

Real News: “Hays County in Central Texas has seen an “incredible 845% increase in (COVID-19) cases since June 7.”

EDA: Exploratory Data Analysis: Label Distribution



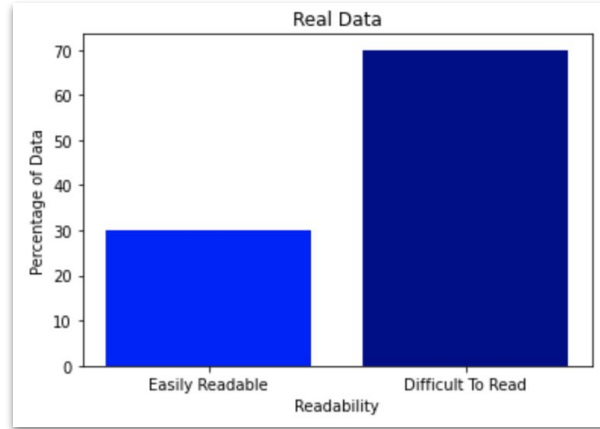
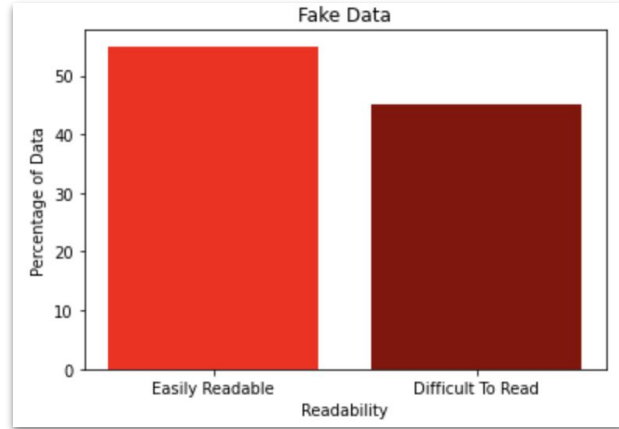
The label distribution is imbalanced.

Entire Dataset: 60% (fake), 40% (Real)

Tweets Dataset: 55% (fake), 45% (Real)

News Articles Dataset: 58% (fake), 42% (Real)

EDA: Exploratory Data Analysis: Readability



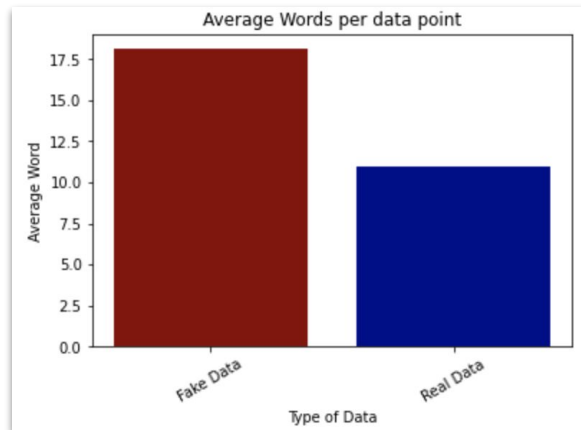
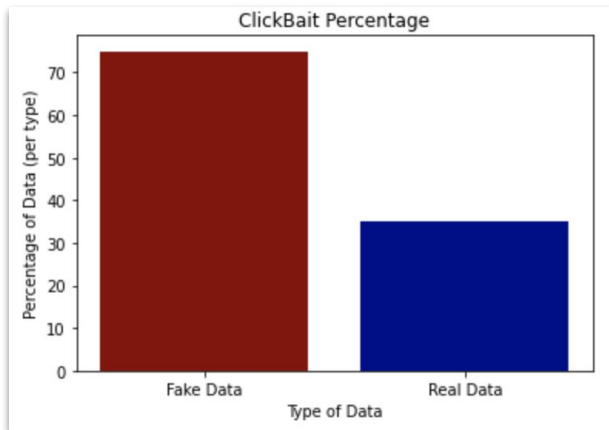
Readability Score:

Reading Ease score = $206.835 - (1.015 \times \text{AWL}) - (84.6 \times \text{ASW})$

AWL: Average word length in a sentence

ASW: Average word length in syllables (number of syllables divided by number of words) [One: Life, love; Two: About, Tuesday and so on]

EDA: Exploratory Data Analysis: ClickBait

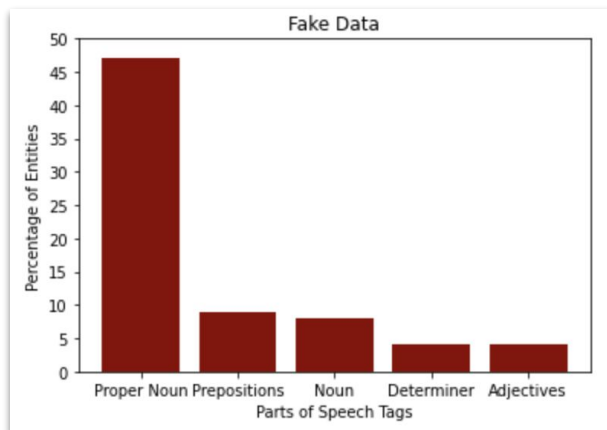


What is Clickbait?

Clickbait is a text or a thumbnail link that is designed to attract attention and to entice users to follow that link/ thread and read, view, or listen to the linked piece of online content, with a defining characteristic of being deceptive, typically sensationalized or misleading.

Example: You'll Never Believe This _____, How to Achieve Results Using This One Weird Trick?

EDA: Exploratory Data Analysis: POS Tags (Stanford NLP)

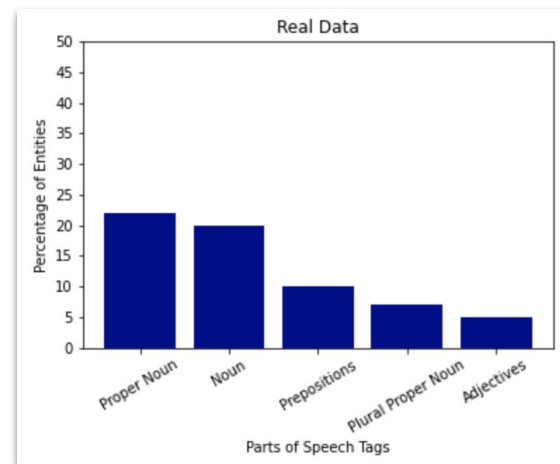


POS Tags Examples:

Proper Noun: Syria, USA, North, Helicopter;

Noun: nerve, agent, army, regime;

Determiner: a, an, the;



POS Tags Examples:

Prepositions: with, when, whenever, at;

Adjectives: big, heavy, human, dead

Plural Proper Noun: Rebels, Fractions, Students

Approach: Siamese Neural Network

TRIPLET LOSS



Anchor(A)



Positive(P)



Negative (N)

$d(A,P)$



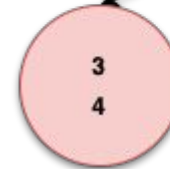
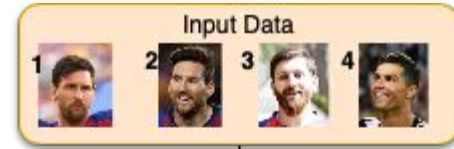
$$\begin{aligned} d(A,P) &\leq d(A,N) \\ d(A,P) - d(A,N) &\leq 0 \quad \text{X} \end{aligned}$$

Trivial Solution = $\text{Img}(P) = \text{Img}(A) = \text{Img}(N) = \text{NULL}$

Introducing margin (alpha)

$$\begin{aligned} d(A,P) - d(A,N) &\leq 0 - (\alpha) \\ d(A,P) - d(A,N) + (\alpha) &\leq 0 \end{aligned}$$

$d(A,N)$

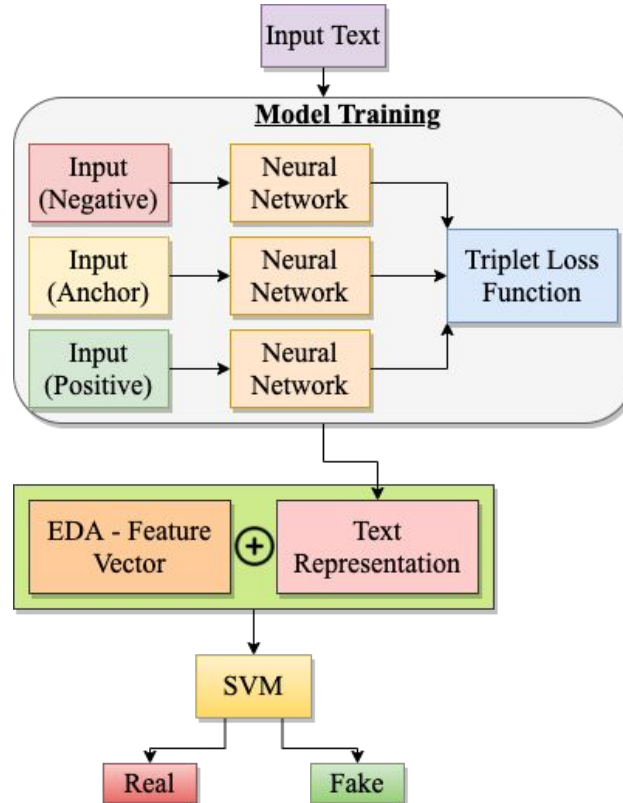


Negative Examples



Positive Examples

Approach: Siamese Neural Network



Results: Tweets trained model tested over News Headlines

Feature Set	Accuracy	Precision	Recall
Embeddings	73.87	73.87	75.22
Embeddings + EDA Features	77.32	77.54	79.33

Results: Trained Model Over Various Domains (Tweets + News Articles)

Feature Set	Accuracy	Precision	Recall
Embeddings	78.38	78.38	79.12
Embeddings + EDA Features	82.44	81.98	83.63

Conclusion and Future Work

- Trained a neural network model to detect fake news over different types of contexts (known and unknown) from different sources (news articles and tweets)
- Obtain more linguistic features which helps in modelling the language styling of each class (real and fake)
- Experiment other state of the art models.

Thank You!