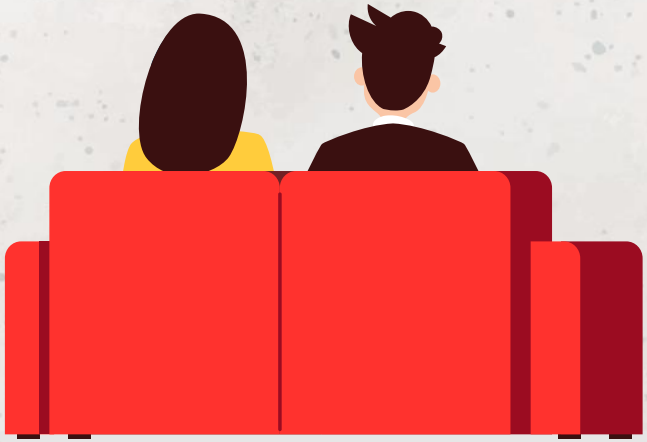
Fake News Classifier

Tan Kai Yong Alvin
DSI-32

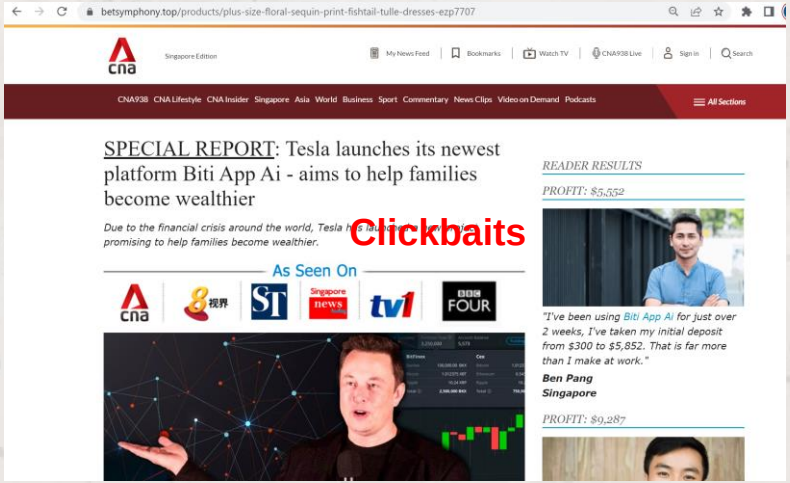
What is Fake News?

- **Fake news** is false or misleading information presented as news.



Problem Statement

- Fake news is a global epidemic posing threat to our society.
- Social media is fuelling the spread of fake news.
- Fake news are Increasingly challenging to detect.



4 in 5 Singaporeans confident in spotting fake news but 90 per cent wrong when put to the test: Survey

Source: The Straits Times (https://www.straitstimes.com/singapore/4-in-5-singaporeans-confident-in-spotting-fake-news-but-90-per-cent-wrong-when-put-to-the)

BEIJING (Reuters) - The failure of Catalonia's effort to declare independence shows that Taiwan independence is also doomed to fail as national unity is paramount for all countries, a Chinese official said on Wednesday. China takes a dim view of independence or secessionist movements around the world, while saying…

A

…it adheres to a policy of non-interference in other nations internal affairs. Beijing has said it opposes any actions to split Spain. Last month Madrid assumed direct rule over Catalonia after the Spanish prime minister sacked the region s secessionist government which had declared independence from Spain following an attempted referendum.

B

… Taiwan is too small and fragile to become a rival to Asia. "Taiwanese are facing a terrible, sad situation. Many Taiwan students and teachers are going into the streets every day to protest, demanding independence . Some even take to the streets, just to demand independence," said a Hong Kong-based official. Hong Kong's main newspaper said on a front page of its online edition last week that it would publish pictures of student and teacher protests on August 9.
.

BEIJING (Reuters) - The failure of Catalonia s effort to declare independence shows that Taiwan independence is also doomed to fail as national unity is paramount for all countries, a Chinese official said on Wednesday. China takes a dim view of independence or secessionist movements around the world, while saying…

A

…it adheres to a policy of non-interference in other nations internal affairs. Beijing has said it opposes any actions to split Spain. Last month Madrid assumed direct rule over Catalonia after the Spanish prime minister sacked the region s secessionist government which had declared independence from Spain following an attempted referendum.

B

… Taiwan is too small and fragile to become a rival to Asia. "Taiwanese are facing a terrible, sad situation. Many Taiwan students and teachers are going into the streets every day to protest, demanding independence. Some even take to the streets, just to demand independence," said a Hong Kong-based official. Hong Kong's main newspaper said on a front page of its online edition last week that it would publish pictures of student and teacher protests on August 9.
.

FAKE

Demonstration

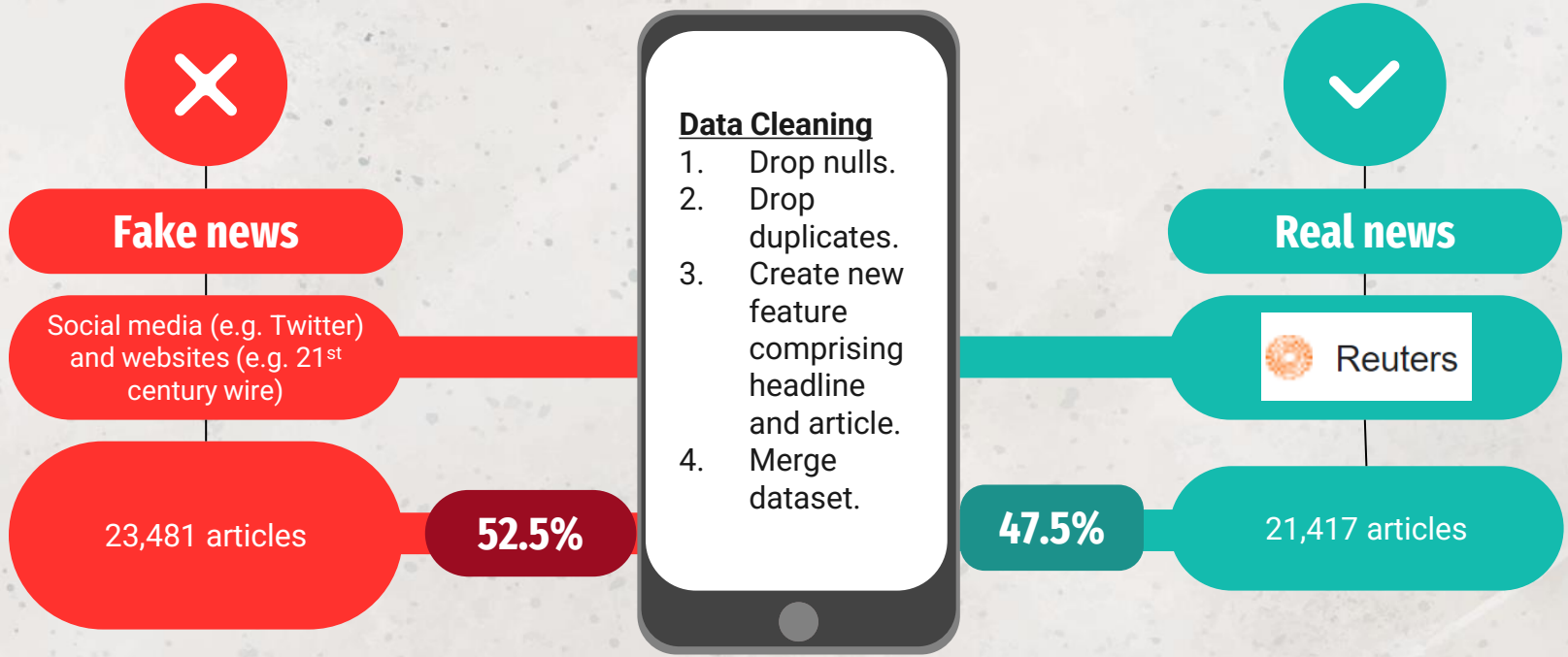
AGENDA

1. Data Pre-processing and Exploratory Data Analysis
2. Modeling
3. Sentiment Analysis
4. Conclusion and Recommendations

01

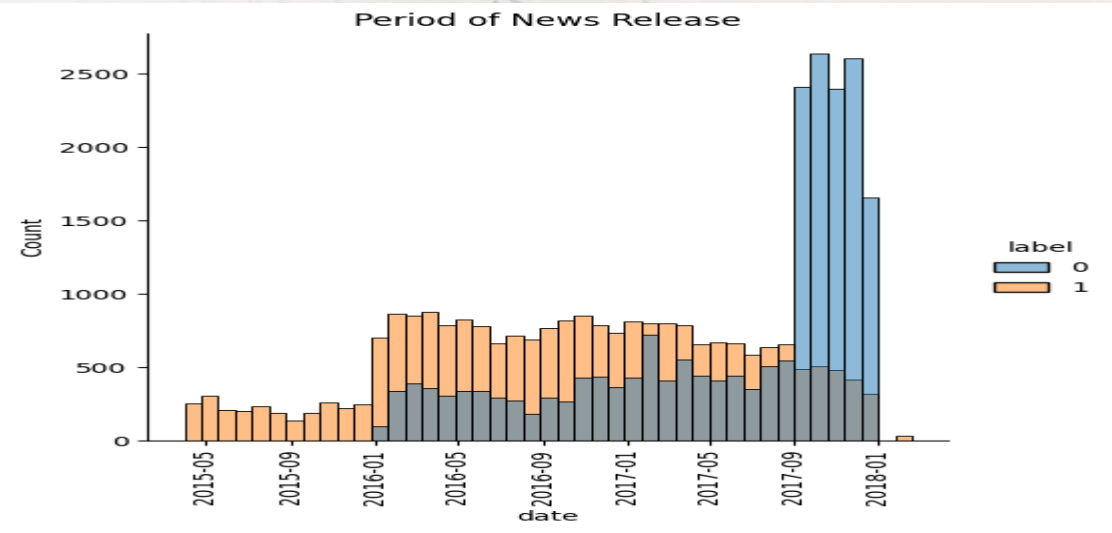
Data Collection / Cleaning and EDA

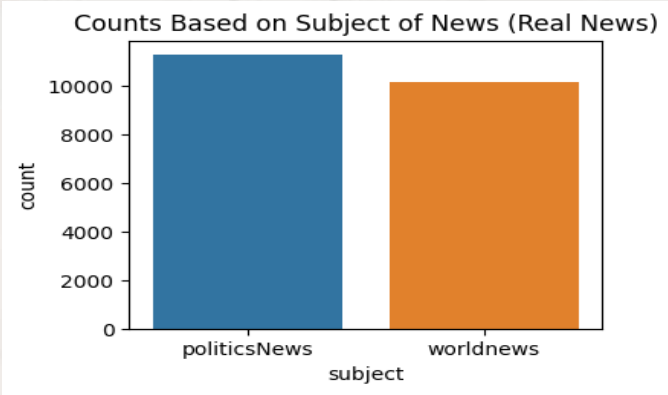
About the Dataset

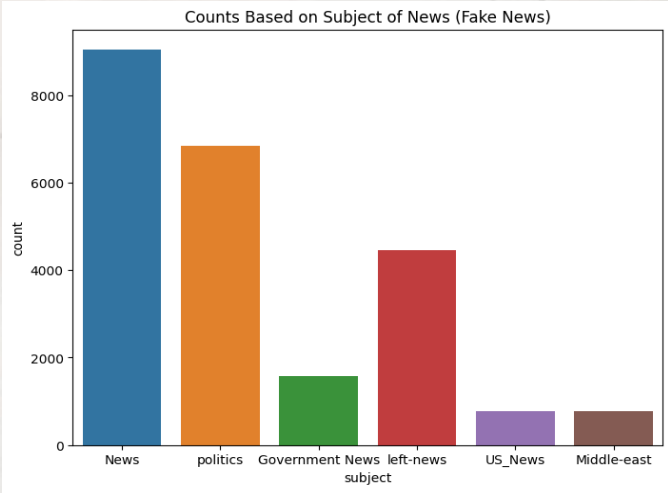


EDA

- Similar subjects reported, dominated by political news.
- Similar period of publication.

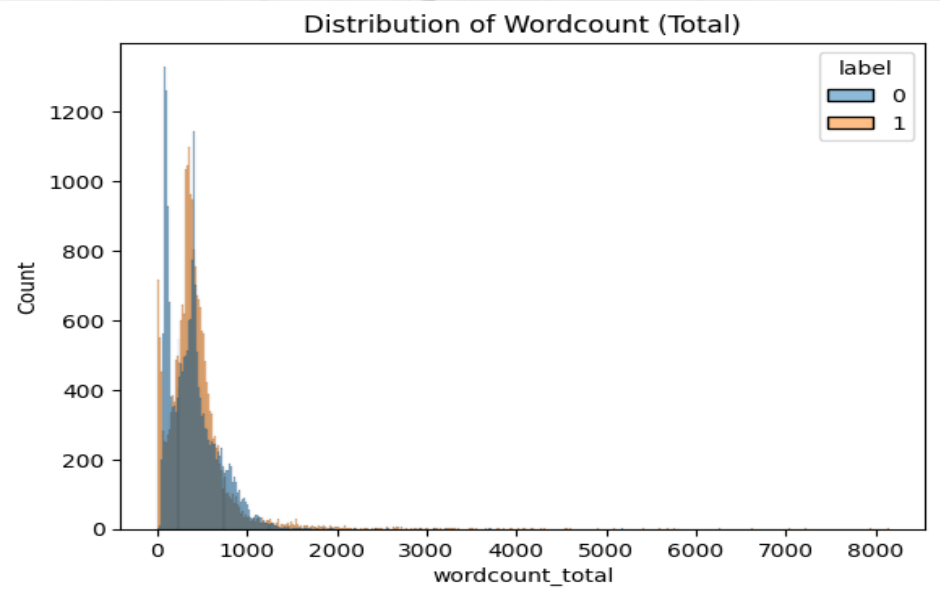


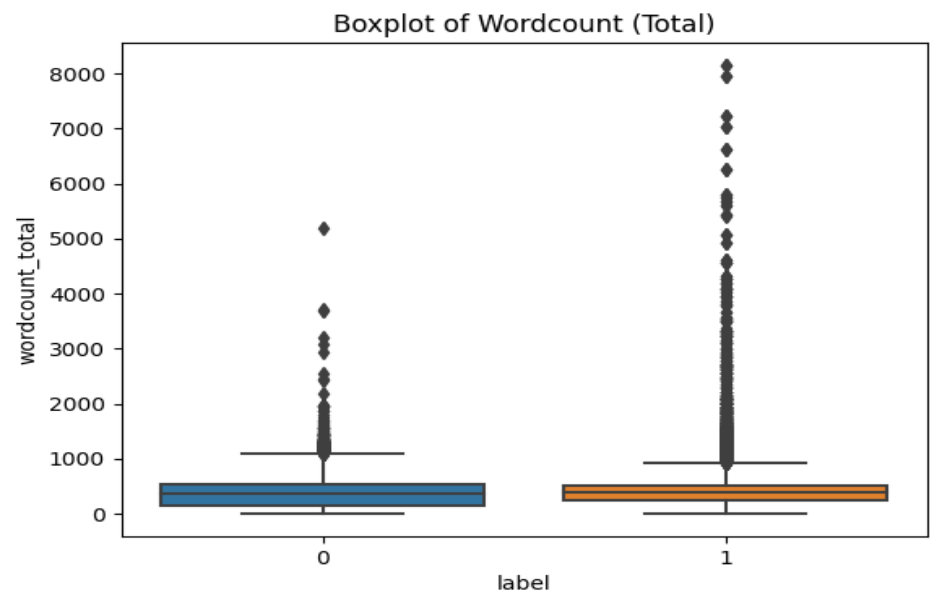




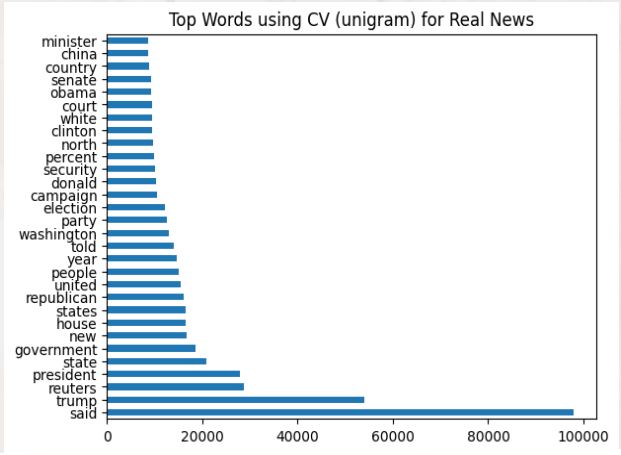
EDA

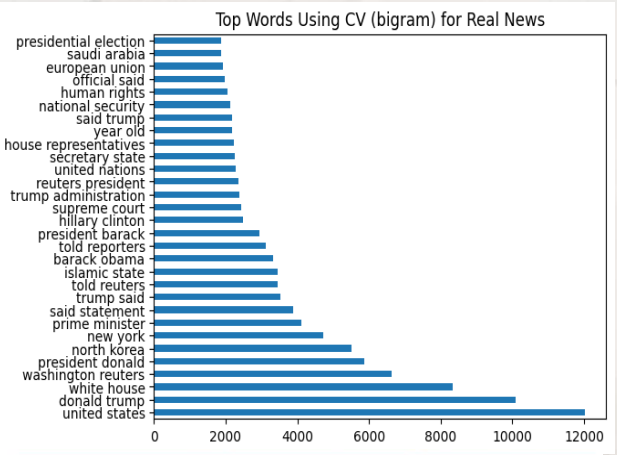
- Word count (mean) is similar with ~400 words per article.
- Large number of outliers in fake news text.

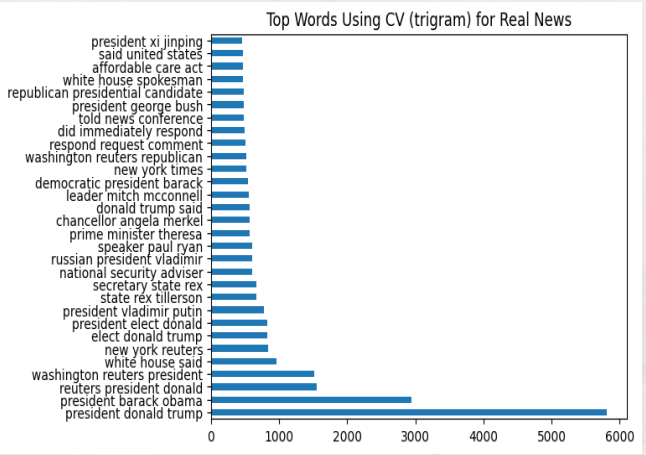


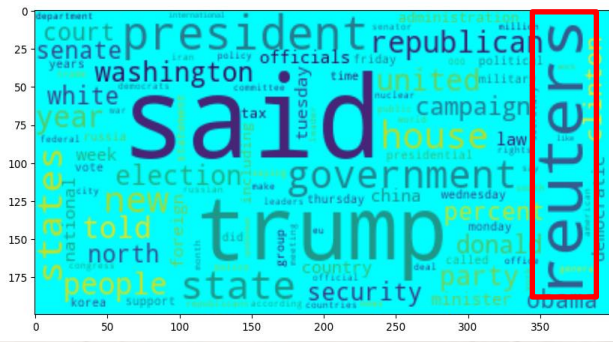


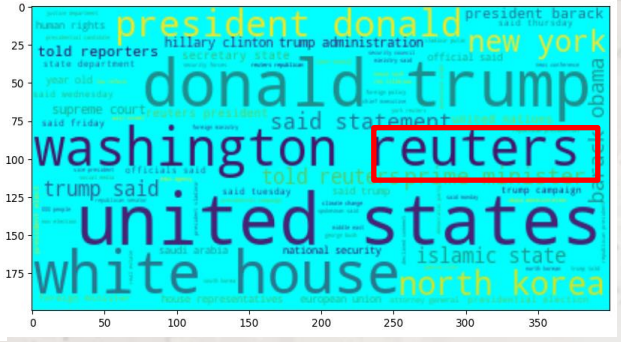
EDA (Real News)

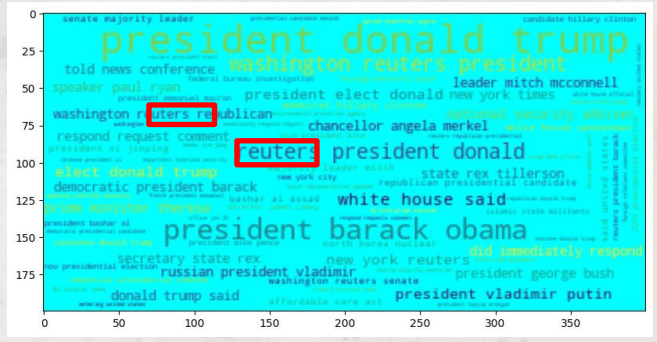




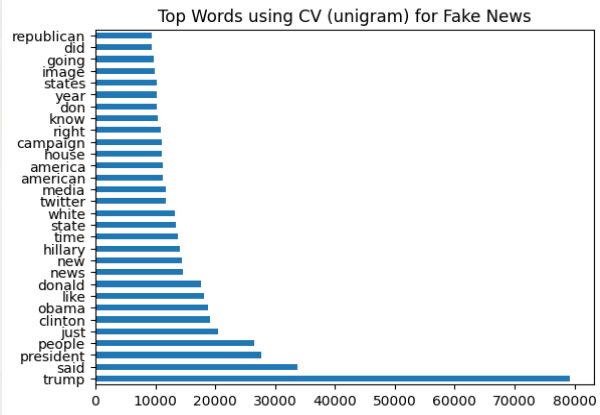


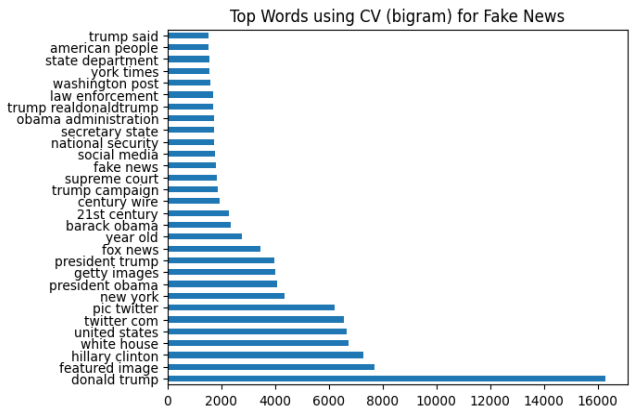


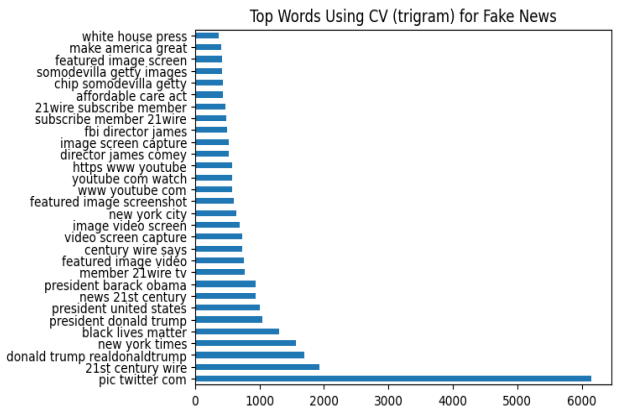


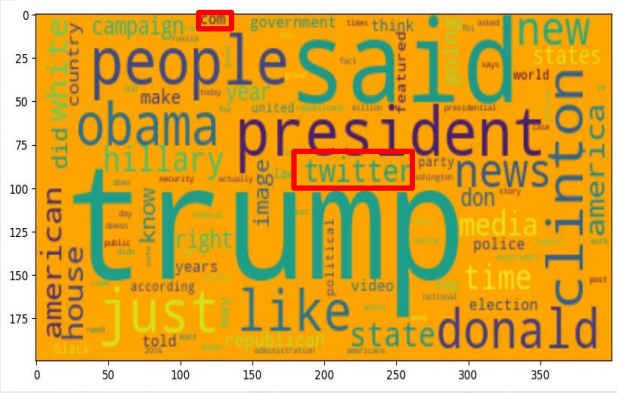


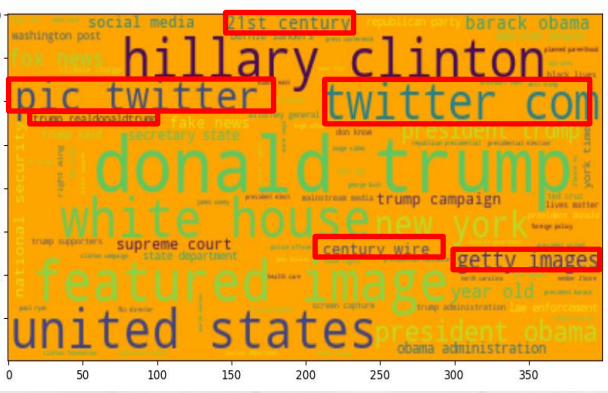
EDA (Fake News)

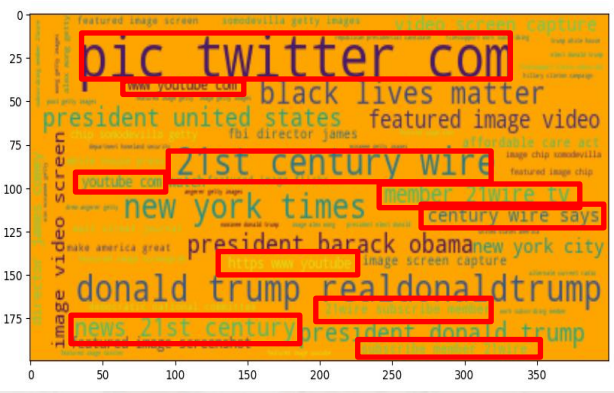












02

Modeling

Modeling

Data Import

Import ‘combined.csv’

Data Pre-processing

Apply regex
1. Lower-case
2. Drop url
3. Drop punctuations

1. Tokenization
2. Drop stop-words

Lemmitization
~~Stemming~~

Vectorization
1. Count Vectorizer
~~2. TF-IDF Vectorizer~~

Modeling

PYCARET

Baseline Model

Modeling

- Models with the best accuracy and recall scores namely Light Gradient Boosting Machine, Logistic Regression and SVM (Linear Kernel).

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
lightgbm	Light Gradient Boosting Machine	0.9865	0.9989	0.9850	0.9893	0.9871	0.9730	0.9730	1.6400
lr	Logistic Regression	0.9832	0.9966	0.9844	0.9837	0.9840	0.9664	0.9664	11.7240
svm	SVM - Linear Kernel	0.9824	0.0000	0.9840	0.9826	0.9833	0.9648	0.9648	1.0820
rf	Random Forest Classifier	0.9807	0.9981	0.9746	0.9885	0.9815	0.9614	0.9615	5.7480
et	Extra Trees Classifier	0.9805	0.9982	0.9723	0.9904	0.9813	0.9610	0.9612	11.3450
ridge	Ridge Classifier	0.9705	0.0000	0.9725	0.9712	0.9719	0.9408	0.9408	0.8990
lda	Linear Discriminant Analysis	0.9703	0.9939	0.9723	0.9712	0.9717	0.9405	0.9405	28.7240
gbc	Gradient Boosting Classifier	0.9605	0.9940	0.9526	0.9716	0.9620	0.9208	0.9211	17.2550
ada	Ada Boost Classifier	0.9591	0.9923	0.9577	0.9641	0.9609	0.9180	0.9180	4.0540
dt	Decision Tree Classifier	0.9362	0.9359	0.9423	0.9366	0.9394	0.8721	0.8721	4.7160
nb	Naive Bayes	0.8985	0.9559	0.8561	0.9454	0.8985	0.7975	0.8015	0.5820

Modeling

Model	Transformer	Train Score (Accuracy)	Test Score (Accuracy)	Recall
Multinomial Naïve Bayes (Baseline Model)	Count Vectorizer	0.9505	0.9543	0.9553
Multinomial Naïve Bayes	TF-IDF Vectorizer	0.9344	0.9359	0.9489
Light Gradient Boosting Machine	Count Vectorizer	0.9903	0.9931	0.9935
Logistic Regression	Count Vectorizer	0.9847	0.9873	0.9889
SVM (Linear Kernel)	Count Vectorizer	0.9838	0.9851	0.9852

Modeling

Data Import

Import ‘combined.csv’

Data Pre-processing

Apply regex
1. Lower-case
2. Drop url
3. Drop punctuations

1. Tokenization
2. Drop stop-words

Lemmitization
~~Stemming~~

Vectorization
1. Count Vectorizer
~~2. TF-IDF Vectorizer~~

Modeling

Neural Ntwork: LSTM (with GloVe embedding)

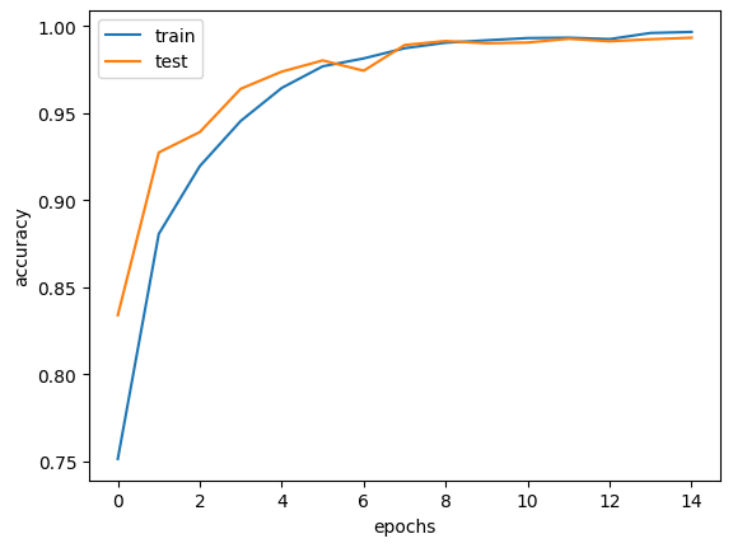
PYCARET

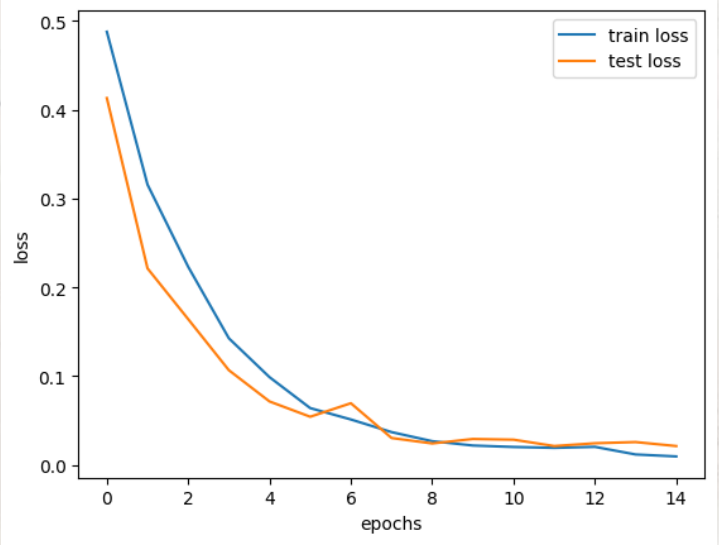
Baseline Model

Modeling

- Count based word embedding does not consider sequence order of texts and the context / meaning of the words.

- Long Short Term Memory (LSTM), a neural network model with word embedding (Global Vector for Word Representation) is explored.





Train Accuracy:
0.9967

Test Accuracy:
0.9934

Are the models that accurate?

Modeling

Data Import

Import ‘combined.csv’

Data Pre-processing

Apply regex
1. Lower-case
2. Drop url
3. Drop punctuations

1. Tokenization
2. Drop stop-words

Lemmitization
~~Stemming~~

Vectorization
1. Count Vectorizer
~~2. TF-IDF Vectorizer~~

Modeling

Neural Ntwork: LSTM (with GloVe embedding)

PYCARET

Baseline Model

Validation on Unseen Dataset

Generation of Unseen Data

- Dataset of 1000 news text comprising of 500 (each) of fake and real news.
- Generated by the GPT-2 word transformer with an initial input of 50 words.

Modeling

Validation on Unseen Dataset

Model(s)

Model	Accuracy
Light Gradient Boosting Machine Model	0.8460 (Best Model)
Logistic Regression	0.8150
LSTM (with GloVe Embedding)	0.5700

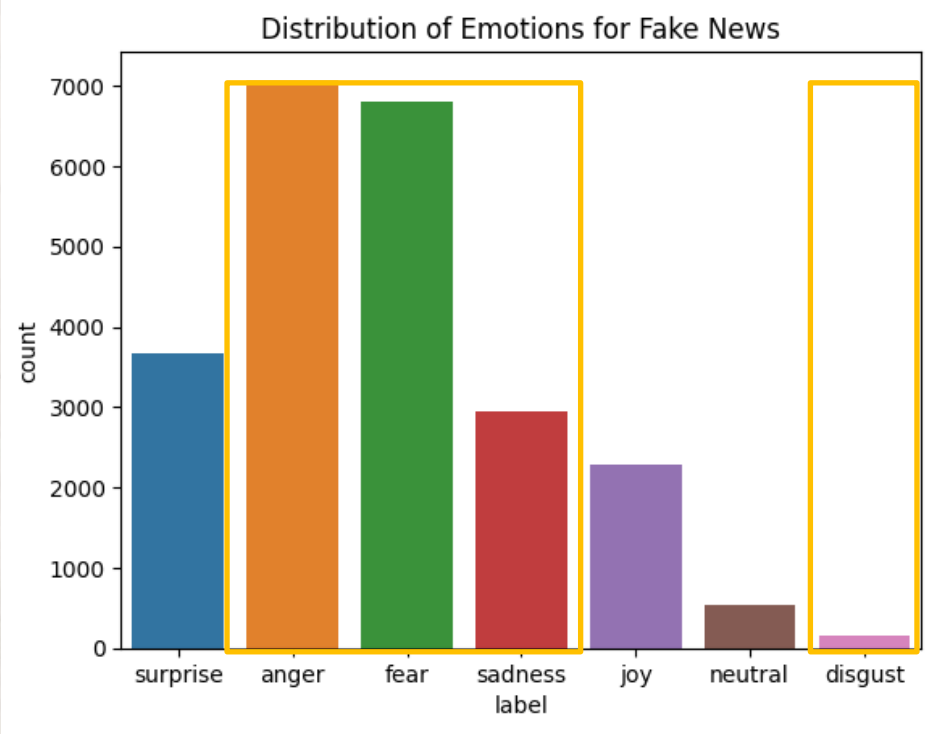
03

Sentiment Analysis

Sentiment Analysis

- Applied emotion sentiment analysis to understand the emotions behind the fake news.
- Pre-trained **emotion-english-distil roberta-base** (hugging face) model. classifies text into 7 emotions ‘surprise’, ‘anger’, ‘fear’, ‘sadness’, ‘joy’, ‘neutral’ and ‘disgust’..
- ~16,000 fake news portrays negative emotions..
- Key words from fake news with negative sentiments (e.g. ‘police officers’, ‘law enforcement’, ‘black lives matters’, ‘affordable care act’).

Hugging Face



04

Conclusion and Recommendations

Key Takeaways

- **Light Gradient Boosting Machine (LGBM) model** with Count Vectorizer is the best model with an accuracy score of 0.9903 (train) and 0.9935 (test).

- Traditional ML approaches which adopts bag of n-grams performed well on unseen data generated by GPT-2 transformer.

- LSTM neural network model with GloVe word embedding performed well with train and test accuracy (0.99) but was unable to perform on unseen dataset. This is likely due to inadequate sequence and context learned from the training data.

Challenges

- Hardware constraint.
- Difficulty in sourcing and label fake news dataset.
- Requires constant fine-tuning.
- Too many context for model to learn. Large dataset is required.

Future Works (Recommendations)

- Training and modelling with a larger datasets with varying contexts, focusing on social media news post and Artificial Intelligence (AI) generated fake news.

- Explore multi-classification problem, predicting the varying levels of truth in the fake news article.

Thank
You

FIRSTDRAFT

7 TYPES OF MIS- AND DISINFORMATION

SATIRE OR PARODY

No intention to cause harm but has potential to fool

MISLEADING CONTENT

Misleading use of information to frame an issue or individual

IMPOSTER CONTENT

When genuine sources are impersonated

FABRICATED CONTENT

New content is 100% false, designed to deceive and do harm

FALSE CONNECTION

When headlines, visuals or captions don't support the content

FALSE CONTEXT

When genuine content is shared with false contextual information

MANIPULATED CONTENT

When genuine information or imagery is manipulated to deceive

The failure of Catalonia’s effort to declare independence shows that Taiwan independence is also doomed to fail as national unity is paramount for all countries, a Chinese official said on Wednesday. China takes a dim view of independence or secessionist movements around the world, while saying it adheres to a policy of non-interference in other nations internal affairs. Beijing has said it opposes any actions to split Spain. Last month Madrid assumed direct rule over Catalonia after the Spanish prime minister sacked the region s secessionist government which had declared independence from Spain following an attempted referendum.

The failure of Catalonia’s effort to declare independence shows that Taiwan independence is also doomed to fail as national unity is paramount for all countries, a Chinese

The failure of Catalonia’s effort to declare independence shows that Taiwan independence is also doomed to fail as national unity is paramount for all countries, a Chinese on Wednesday. Chin takes a dim view of independence or secessionist movements around the world while it adheres to a policy of non-interference in other internal affairs. Beijing has it opposes any actions to split Spain. Last month Madrid assumed direct rule over Catalonia after the Spanish sacked the region s secessionis which had declared independence from Sp following an attempted referendum.