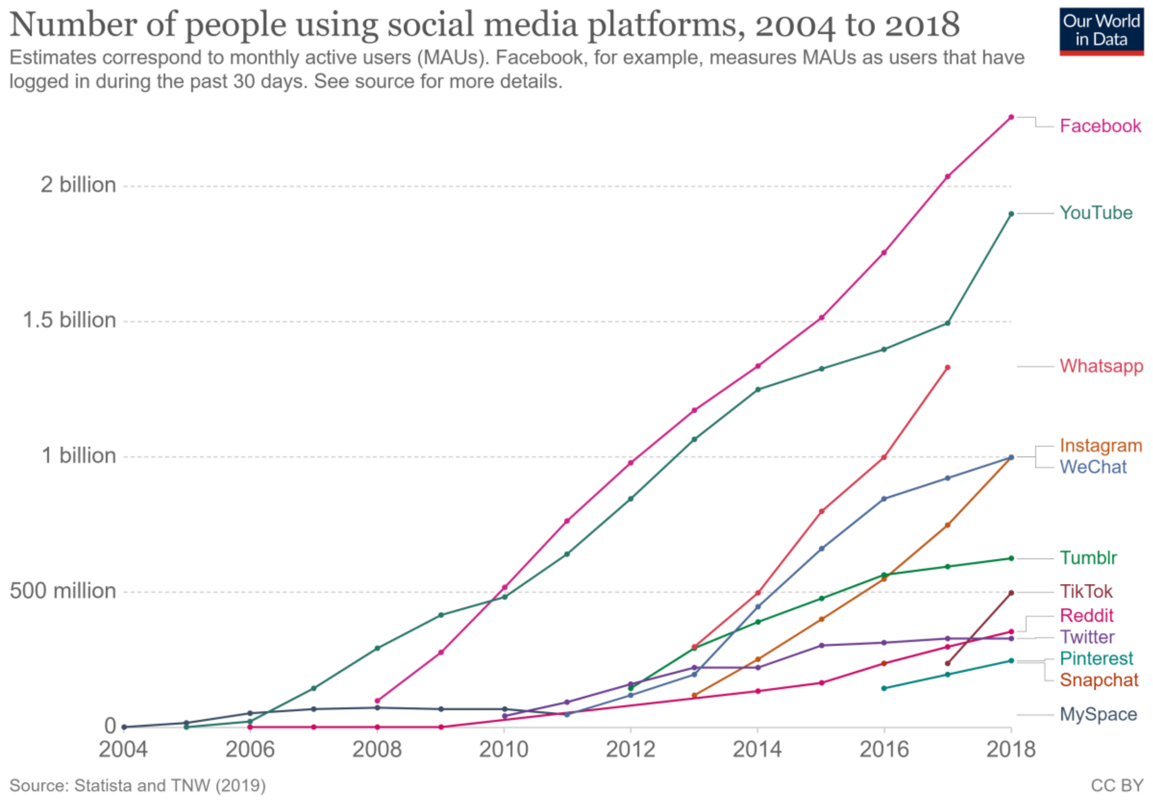
Real or Fake News? Using NLP with Disasters Tweets

Capstone project by **Lim Zheng Wei**

Background

- In 1990, 0.5% of world population were online
- 26 years later, over 3 billion people are using the net (Roser, 2018)
- Young adults are more likely to get news via social media((Perrin and Kumar, 2020)



(Ortiz-Ospina 2020)

Context in Singapore

- 4.45 million Facebook users, 2.11 million Instagram users, 2.83 million LinkedIn users
- Averages 6 hours 48 mins online, 2 hours 8 minutes on social media
- More than 4.6 million social media users as of Jan 2020
- Cyber crimes rise by nine-fold in the last 3 years (Shahari, 2020)

Problem

- People to announce an emergency they're observing in real-time
- People can post easily, cannot verify the source
- Not always clear whether a person's words are announcing a disaster.
- 4 in 5 Singaporeans confident in spotting fake news but 90 per cent wrong when put to the test(Hui Wen, 2020)

Objective, Scope

Objective: Build a machine model to differentiate between real and fake disaster news

Scope: Use dataset from Kaggle, use different classifier to separate the data.

Use accuracy, precision, recall, roc score to assess the classifier.

Fake News

- Fake news is an invention, false information that appears to be real news with the purpose of tricking people
- Fake news used to be common in print but has increased with the rise of social media(Lee, 2019)

Implications

- Information shape people's opinion
- People make important decisions based on information
- People form an impression about people or a situation by obtaining information
- Invented, false, exaggerated, or distorted information cause people to make bad decisions("Impacts of Fake News", 2019)

Impacts of Fake News

- Financial
- Fear
- Health
- Racist ideas
- Bully and Violence
- Democratic

Steps

- Pre-processing
- Text cleaning
- 20% of data used as test set
- 80% for modelling

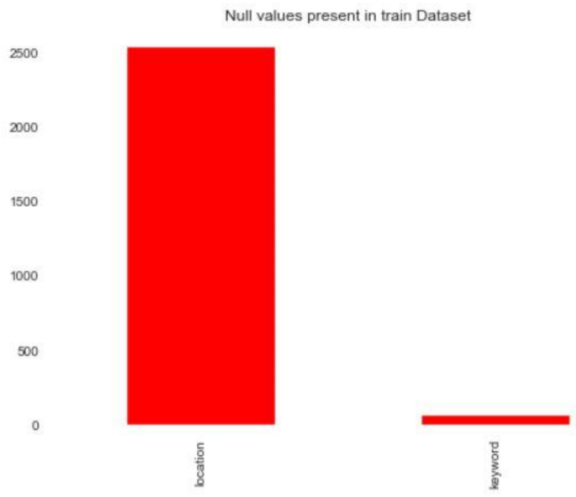
Trust the Process

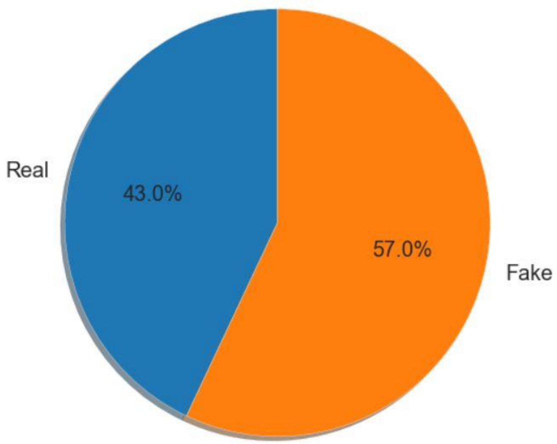
EDA → Basic Feature Extraction → Text Cleaning → Feature Engineering → Text Classification

EDA

	id	keyword	location	text	target
0	1	NaN	NaN	Our Deeds are the Reason of this #earthquake M...	1
1	4	NaN	NaN	Forest fire near La Ronge Sask. Canada	1
2	5	NaN	NaN	All residents asked to 'shelter in place' are ...	1
3	6	NaN	NaN	13,000 people receive #wildfires evacuation or...	1
4	7	NaN	NaN	Just got sent this photo from Ruby #Alaska as ...	1

- 7613 rows, 5 columns

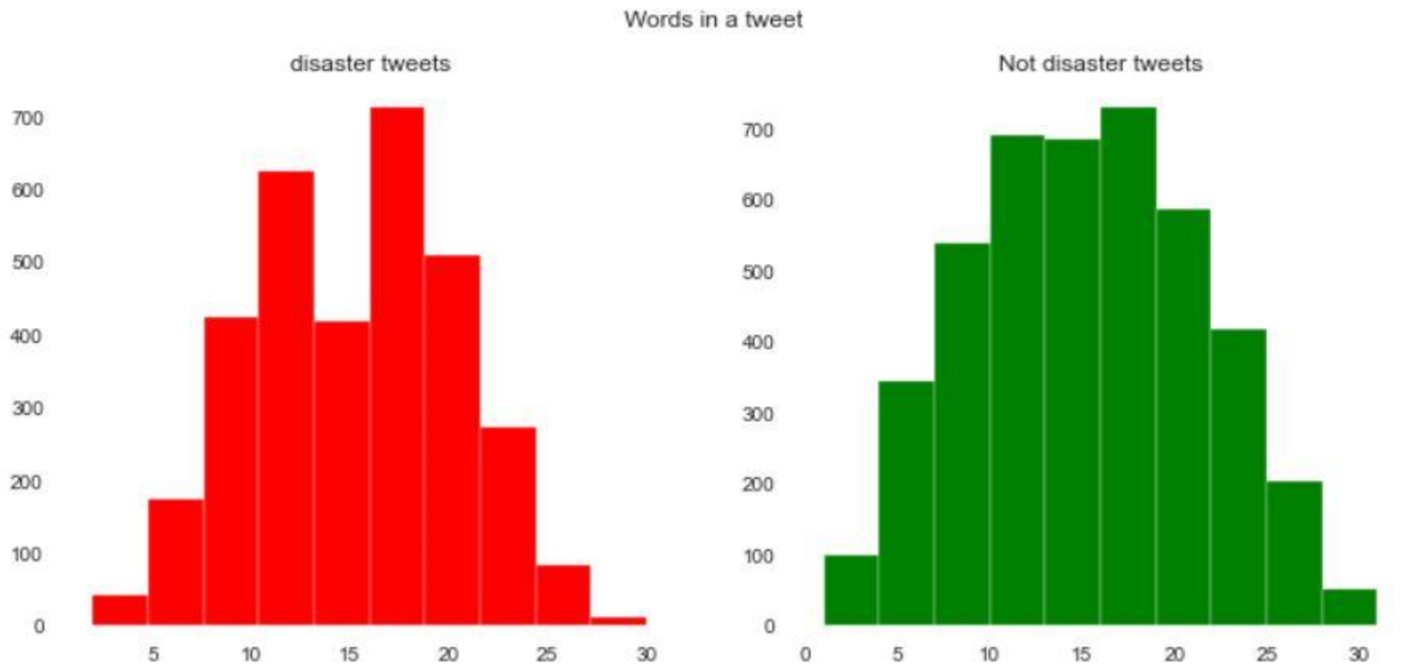




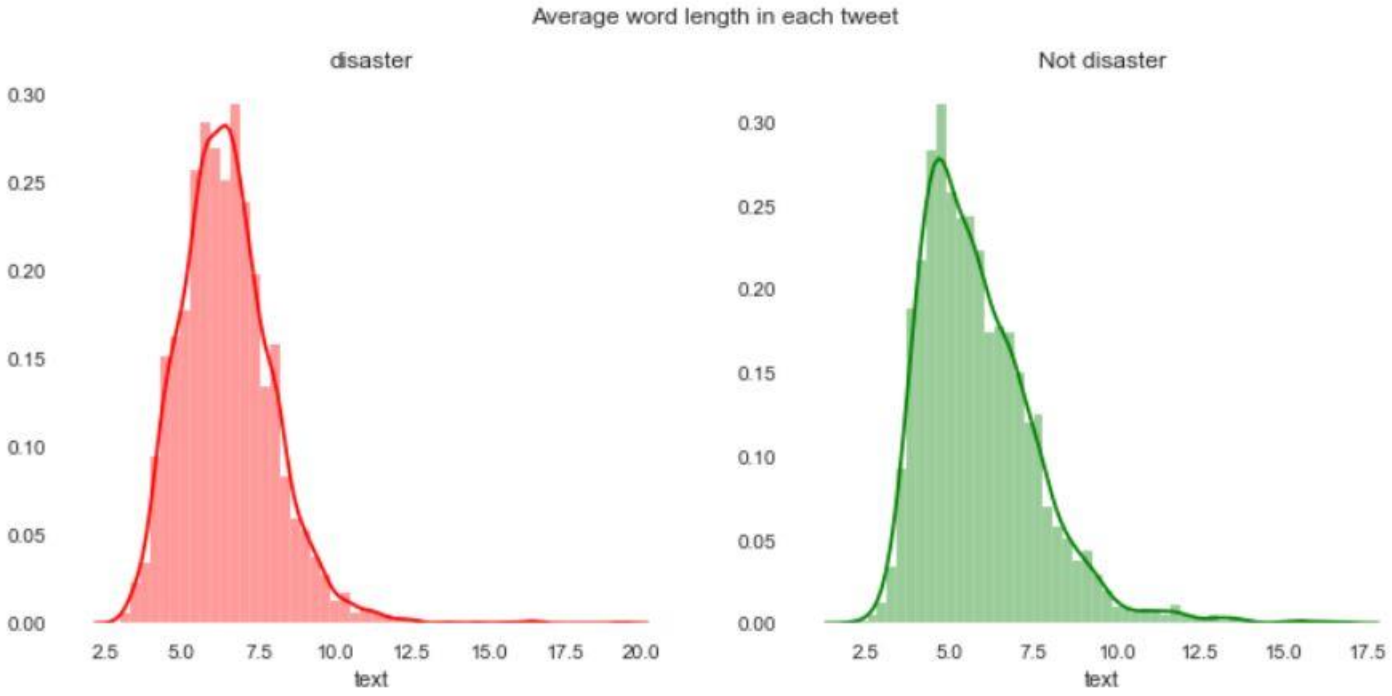
Basic Feature Extraction

- The following methods were being used to see if there were any distinct features to differentiate between a fake and a real disaster
- The outcome of the results would likely affect how the cleaning of the texts in the subsequent stages.

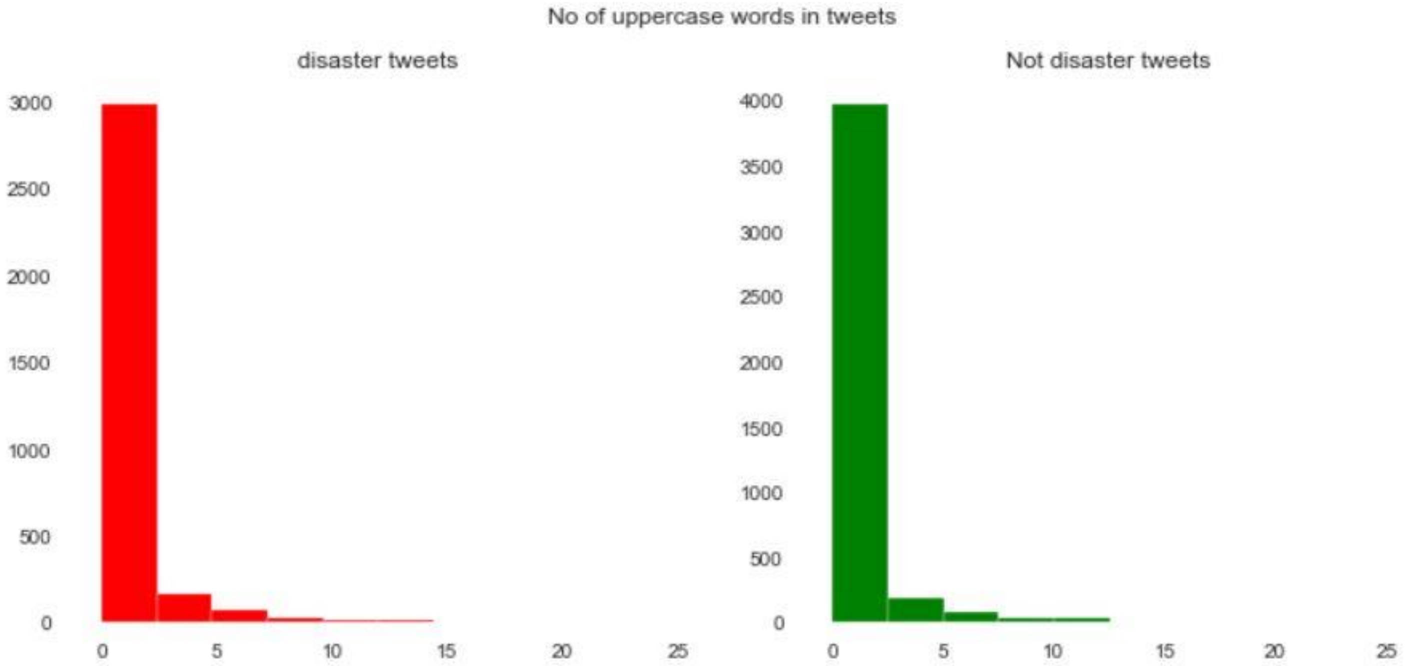
Number of Words



Average Word Length



Number of Uppercase Words



Feature Engineering

CountVectoriser – Transform text into word count

Data = ['The', 'quick', 'brown', 'fox', 'jumps', 'over', ' the', 'lazy', 'dog']

	The	quick	brown	fox	jumps	over	lazy	dog
Data	2	1	1	1	1	1	1	1

("CountVectorizer in Python", 2020)

Term Frequency Inverse Document Frequency (TF-IDF) – Determine the emphasise of a word/phrase

TF-IDF N-Gram – Takes a sequence of words into account

Text Classification

- Naïve Bayes
- Linear
- Support Vector
- Bagging (Random Forest)
- Boosting (Gradient Descent)
- Deep Learning (Keras)

Selected Results – Naïve Bayes

	CountVectoriser	TF-IDF Word	TF-IDF N-Gram
Accuracy	0.810	0.821	0.750
Precision	0.806	0.856	0.885
Recall	0.734	0.702	0.480
ROC_AUC	0.863	0.865	0.747

Selected Results – Linear

	CountVectoriser	TF-IDF Word	TF-IDF N-Gram
Accuracy	0.807	0.814	0.752
Precision	0.805	0.845	0.892
Recall	0.734	0.694	0.482
ROC_AUC	0.857	0.866	0.748

Types of Keras Models

- **Baseline** – Original neurons and layers chosen
- **Reduced** – Fewer neurons and/or layers
- **Regularised** – Adding penalty for overfitting a function
- **Dropout** – Random neurons being ignored

Selected Results – Keras

	Baseline	Reduced	Regularised	Dropout
Accuracy	0.897	0.893	0.920	0.891
Precision	0.915	0.916	0.947	0.911
Recall	0.844	0.834	0.868	0.834
ROC_AUC	0.891	0.887	0.915	0.885

Discussions

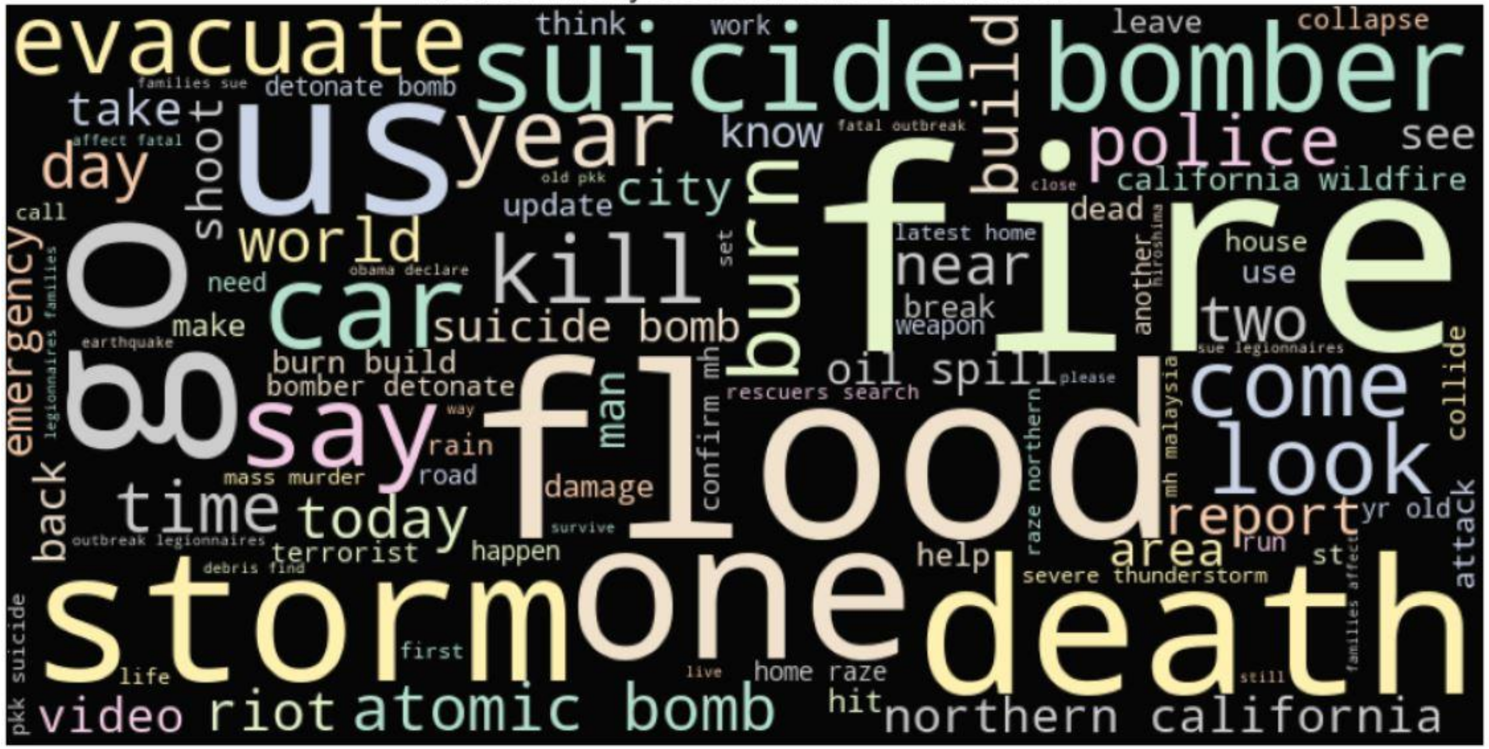
- CountVectoriser and TF-IDF did well
- TF-IDF did better between the 2
- TF-IDF N-Grams have best precision but worst recall
- All 4 Keras models performed better than the rest

Fake - ID 10258: Camping in a war zone with roving raccoons toughens city slicker http://t.co/oJuS08yZrq

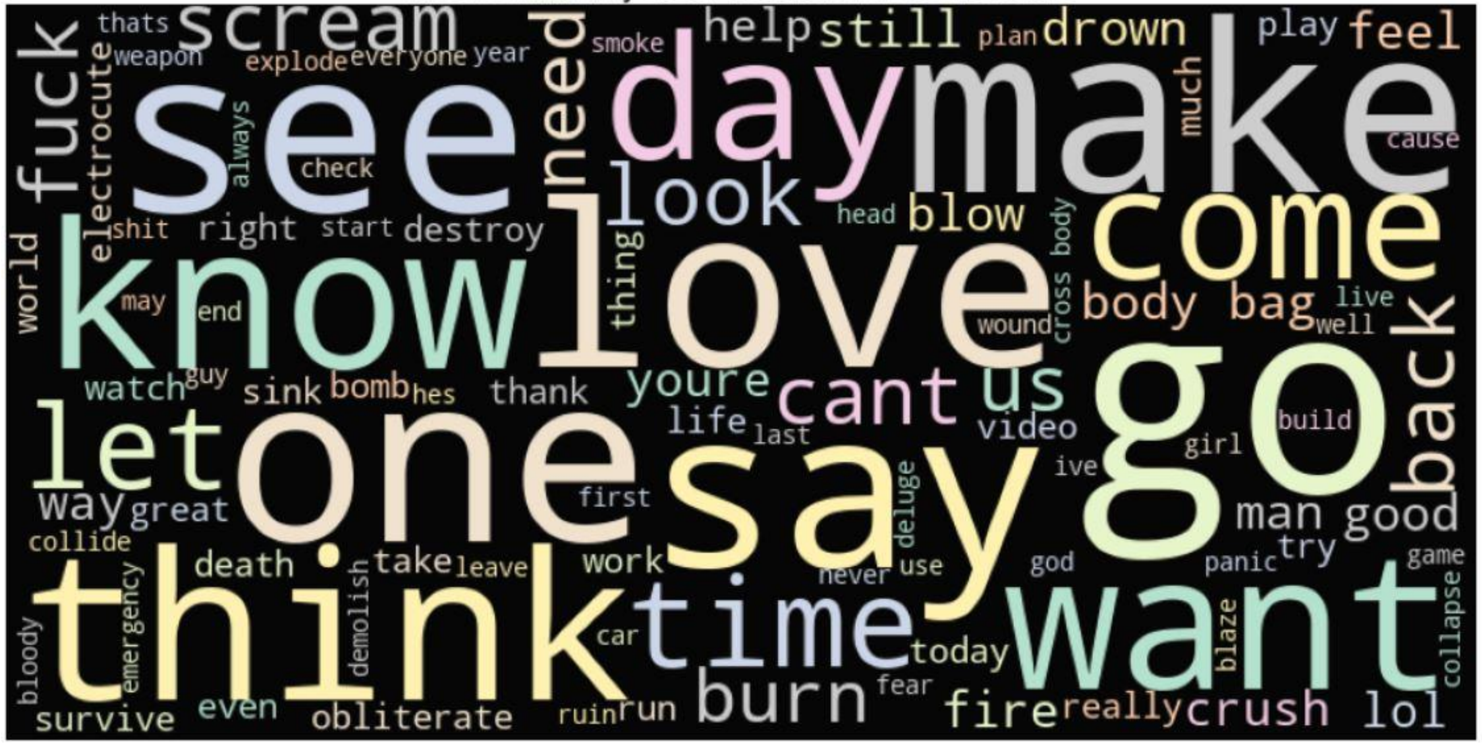
Real - ID 4641: Services are returning to normal #SouthLine after a medical emergency at Yennora and urgent track equipment repairs at Cabramatta earlier.

WordCloud

Most commonly used words for real disasters



Most commonly used words for fake disasters



Word Importance

	Words
Fake	Think, One, Say, Go, Love, See, Day, Make, Want, Time, Come, Know
Real	Evacuate, Suicide, Bomber, Flood, Storm, Riot, Kill, Emergency, Northern California, Wildfire, Shoot, Damage, Severe, Thunderstorm

Conclusion

- Fake news is real
- Regularised Keras is the best model overall
- Considering there are many many more texts out there, deep learning is the way forward
- Future Work: Try other DL methods

References

- Roser, M. (2018). The internet's history has just begun. Our World in Data. Retrieved 15 September 2020, from https://ourworldindata.org/internet-history-just-begun.
- Ortiz-Ospina, E., 2020. *The Rise Of Social Media*. [online] Our World in Data. Available at: <https://ourworldindata.org/rise-of-social-media> [Accessed 12 August 2020].
- Perrin, A. and Kumar, M., 2020. *About Three-In-Ten U.S. Adults Say They Are 'Almost Constantly' Online*. [online] Pew Research Center. Available at: <https://www.pewresearch.org/fact-tank/2019/07/25/americans-going-online-almost-constantly/> [Accessed 12 August 2020].
- Shahari, F., 2020. *10+ Interesting Social Media Statistics In Singapore (Updated 2019)*. [online] CloudRock® Digital Marketing Agency | Singapore. Available at: <https://cloudrock.asia/sg/blog/social-media-statistics-singapore/#:~:text=Singapore%20Social%20Media%20Statistics,in%20Singapore%20had%20touched%2079%25.> [Accessed 12 August 2020].
- Hui Wen, N., 2020. *4 In 5 Singaporeans Confident In Spotting Fake News But 90 Per Cent Wrong When Put To The Test: Survey*. [online] The Straits Times. Available at: <https://www.straitstimes.com/singapore/4-in-5-singaporeans-confident-in-spotting-fake-news-but-90-per-cent-wrong-when-put-to-the> [Accessed 12 August 2020].
- Lee, T. (2019). The global rise of "fake news" and the threat to democratic elections in the USA | Emerald Insight. Emerald.com. Retrieved 21 September 2020, from https://www.emerald.com/insight/content/doi/10.1108/PAP-04-2019-0008/full/html.
- Impacts of Fake News. 30sec to check it out. (2019). Retrieved 21 September 2020, from https://30secondes.org/en/module/impacts-of-fake-news/.
- CountVectorizer in Python. (2020). Retrieved 8 October 2020, from https://www.educative.io/edpresso/countvectorizer-in-python