



# Fast Tracking Media Intelligence

Natural Language Processing Model for Text  
Summarization of CNN Daily News articles

By Andrew Levinton



# Business Problem

- Summarize CNN DailyMail articles
- Streamline and automate the understanding of complex news topics.
- Help track news developments to enhance market research.



# Data

- The CNN / DailyMail Dataset written by journalists at CNN and the Daily Mail.
- Data contains Articles and Highlights(reference summaries created by journalists)
- Train set: 287,113
- Test set: 11,490



# Article and Summary Length

The average CNN Daily Mail article is 683 words in length

- 5-7 minutes of reading, and potentially 10+ minutes to paraphrase

The average reference summary is 55 words

- Less than a minute to read, possibly 1-3 minutes to paraphrase



# Data Preprocessing

Contraction Mapping - expanding shortened words into readable phrases.

**{"ain't": "is not",  
"aren't": "are not",  
"can't": "cannot",  
"cause": "because"}**



# Data Preprocessing

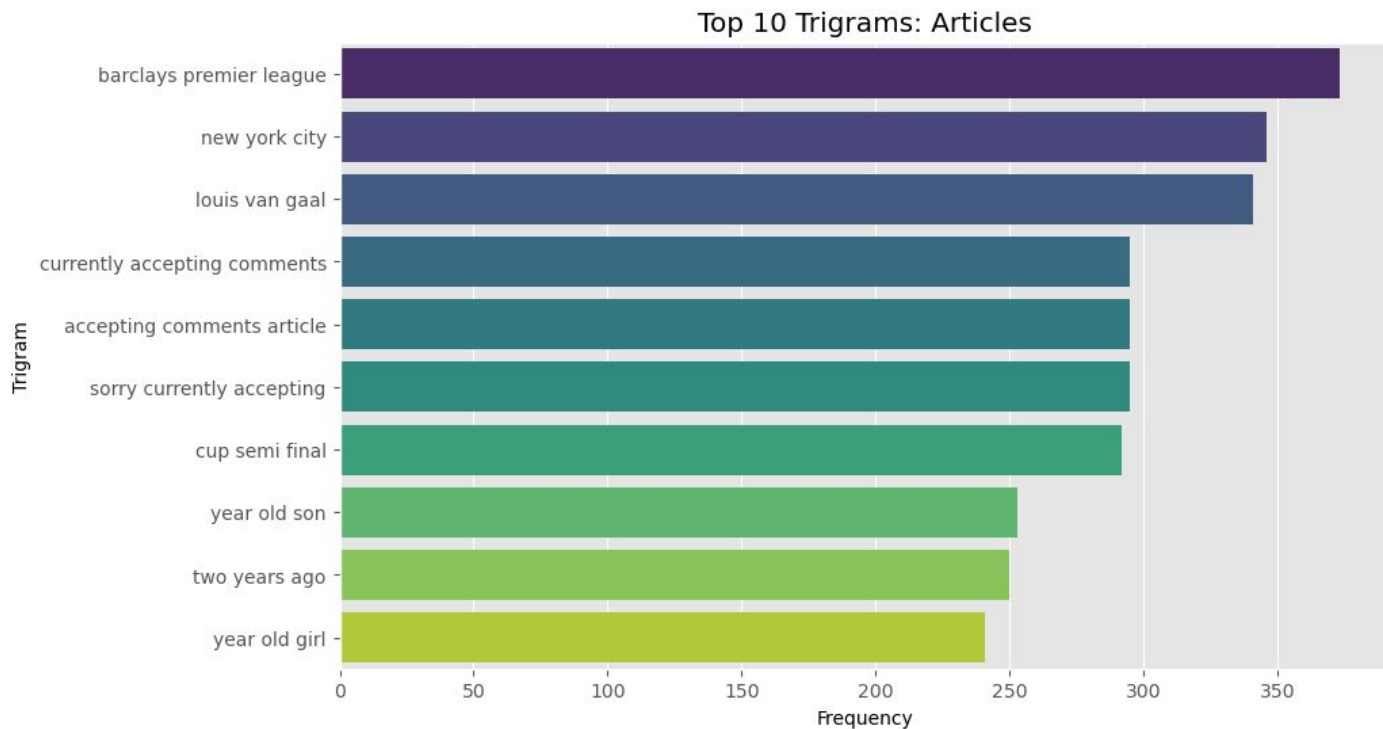
## Before Cleaning

Ever noticed how plane seats appear to be getting smaller and smaller? With increasing numbers of people taking to the skies, some experts are questioning if having such packed out planes is putting passengers at risk. They say that the shrinking space on aeroplanes is not only uncomfortable – it's putting our health and safety in danger. More than squabbling over the arm rest, shrinking space on planes putting our health and safety in danger? This week, a U.S consumer advisory group set up by the

## After Cleaning

ever noticed plane seats appear getting smaller smaller increasing numbers people taking skies experts questioning packed planes putting passengers risk say shrinking space aeroplanes uncomfortable putting health safety danger squabbling arm rest shrinking space planes putting health safety danger week consumer advisory group

# Findings





# Modeling

EDA

Preprocessing

Modeling

Model Evaluation

Recommendations





# Model Demo

## Original Article

A drunk teenage boy had to be rescued by security after jumping into a lions' enclosure at a zoo in western India. Rahul Kumar, 17, clambered over the enclosure fence at the Kamla Nehru Zoological Park in Ahmedabad, and began running towards the animals, shouting he would 'kill them'. Mr Kumar explained afterwards that he was drunk and 'thought I'd stand a good chance' against the predators. Next level drunk: Intoxicated Rahul Kumar, 17, climbed into the lions' enclosure at a zoo in Ahmedabad and began running towards the animals shouting 'Today I kill a lion!' Mr Kumar had been sitting near the enclosure when he suddenly made a dash for the lions, surprising zoo security. The intoxicated teenager ran towards the lions, shouting: 'Today I kill a lion or a lion kills me!' A zoo spokesman said: 'Guards had earlier spotted him close to the enclosure but had no idea he was planning to enter it. 'Fortunately, there are eight moats to cross before getting to where the lions usually are and he fell into the second one, allowing guards to catch up with him and take him out. 'We then handed him over to the police.' Brave fool: Fortunately, Mr Kumar fell into a moat as he ran towards the lions and could be rescued by zoo security staff before reaching the animals (stock image) Kumar later explained: 'I don't really know why I did it. 'I was drunk and thought I'd stand a good chance.' A police spokesman said: 'He has been cautioned and will be sent for psychiatric evaluation. 'Fortunately for him, the lions were asleep and the zoo guards acted quickly enough to prevent a tragedy similar to that in Delhi.' Last year a 20-year-old man was mauled to death by a tiger in the Indian capital after climbing into its enclosure at the city zoo.

### Model Summary

"Rahul Kumar, 17, clambered over enclosure fence at Kamla Nehru Zoological Park. He ran towards the animals shouting 'Today I kill a lion or a lion kills me!' Fortunately, Mr Kumar fell into a moat and was rescued by zoo security. He has been cautioned and will be sent for psychiatric evaluation."

### Human Summary

Drunk teenage boy climbed into lion enclosure at zoo in west India .  
Rahul Kumar, 17, ran towards animals shouting 'Today I kill a lion!' Fortunately he fell into a moat before reaching lions and was rescued .



# Model Performance

## Article 1

### Predicted Summary:

U.S consumer advisory group set up by the Department of Transportation said that while the government is happy to set standards for animals flying on planes, it doesn't stipulate a minimum amount of space for humans. Tests conducted by the FAA use planes with a 31 inch pitch, a standard which on some airlines has decreased. Many economy seats on United Airlines have 30 inches of room, while some airlines offer as little as 28 inches.

### Reference Summary:

Experts question if packed out planes are putting passengers at risk.

U.S consumer advisory group says minimum space must be stipulated.

Safety tests conducted on planes with more leg room than airlines offer.

**Rouge-L Score:**  
**0.39**

The ROUGE-L score is relatively low, indicating poor recall and overlap with the reference summary.



# Model Performance

## Article 2

### Predicted Summary:

Rahul Kumar, 17, clambered over enclosure fence at Kamla Nehru Zoological Park. He ran towards the animals shouting 'Today I kill a lion or a lion kills me!' Fortunately, Mr Kumar fell into a moat and was rescued by zoo security. He has been cautioned and will be sent for psychiatric evaluation. Last year a 20-year-old man was mauled to death by a tiger in Delhi.

### Reference Summary:

Drunk teenage boy climbed into lion enclosure at zoo in west India.

Rahul Kumar, 17, ran towards animals shouting 'Today I kill a lion!'

Fortunately he fell into a moat before reaching lions and was rescued.

**ROUGE-L Score: 0.6471**

The ROUGE-L score is relatively high, indicating good recall and overlap with the reference summary.

# Model Demonstration





# Business Benefits to this model

## Competitive Intelligence:

- Track news and developments in their industry
- Stay ahead of competitors and adapt to market changes more effectively.

## Market Research:

- Analyze emerging trends, and competitor strategies



# Future Work

- Fine tuning BART model to further improve performance
- Further Developing a front end app to be used at scale by stakeholders



# Thank you!

Email: [alevinton151@gmail.com](mailto:alevinton151@gmail.com)

Github: <https://github.com/andrewkoji>

LinkedIn: <https://www.linkedin.com/in/andrew-levinton-23603b22a/>