

# REAL VS. FAKE NEWS

Kevin Gross





# Agenda



# Business Problem

**Provide CNN with a mechanism that confidently identifies real and fake news stories**

- Has becoming increasingly difficult to make sense of what content is based on fact, half-truths or lies
- CNN needs to report on real news in an era of information overload



# Overview



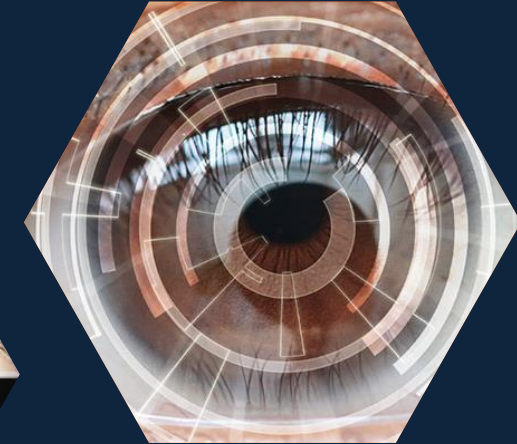
**Multinational 24-  
Hour Cable News  
Channel**



**Real vs. Fake News**



**"To Inform, Engage,  
and Empower the  
World"**



**"We are truth-  
seekers and  
storytellers"**



# Business Understanding

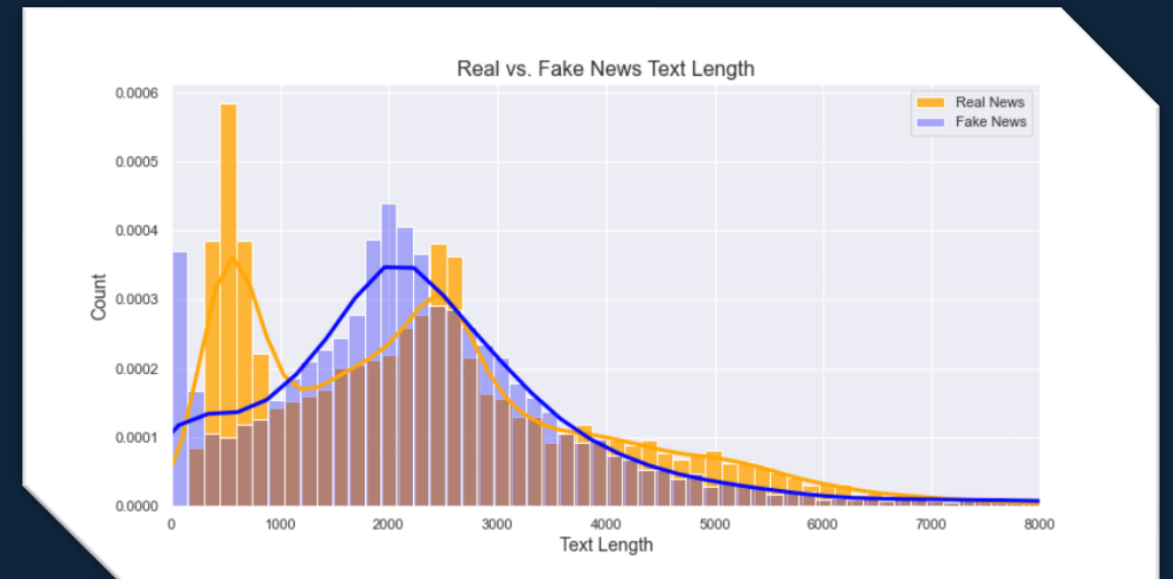
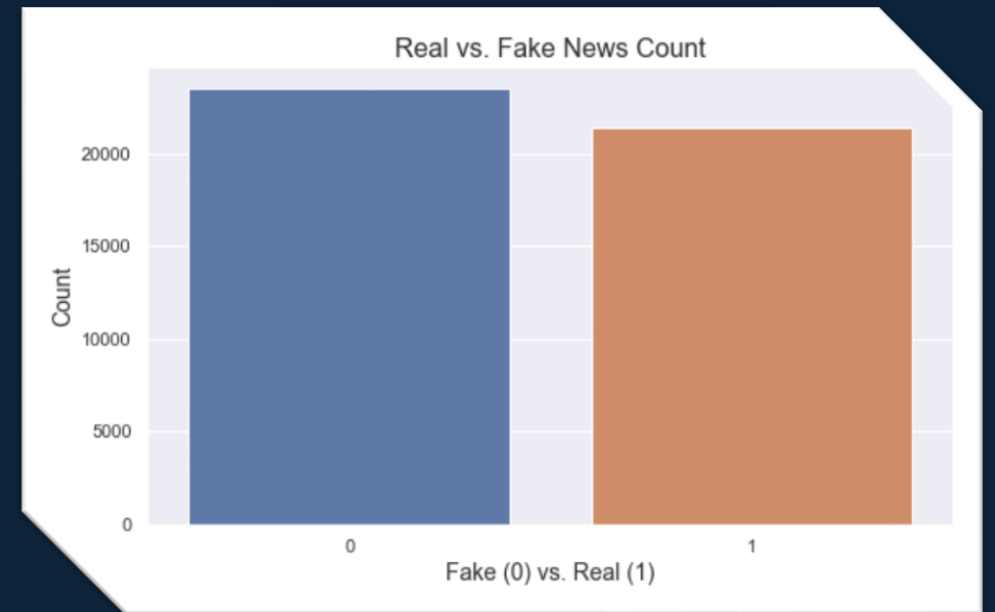
Information overload

Fact, half-truth, or lies?

Digital platforms ripple effect



# DATA UNDERSTANDING



# Modeling



Standardizing  
Case



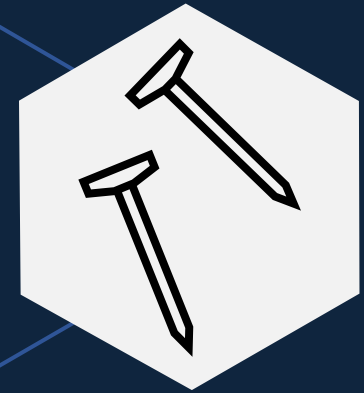
Removing  
Punctuation



Removing  
Stopwords

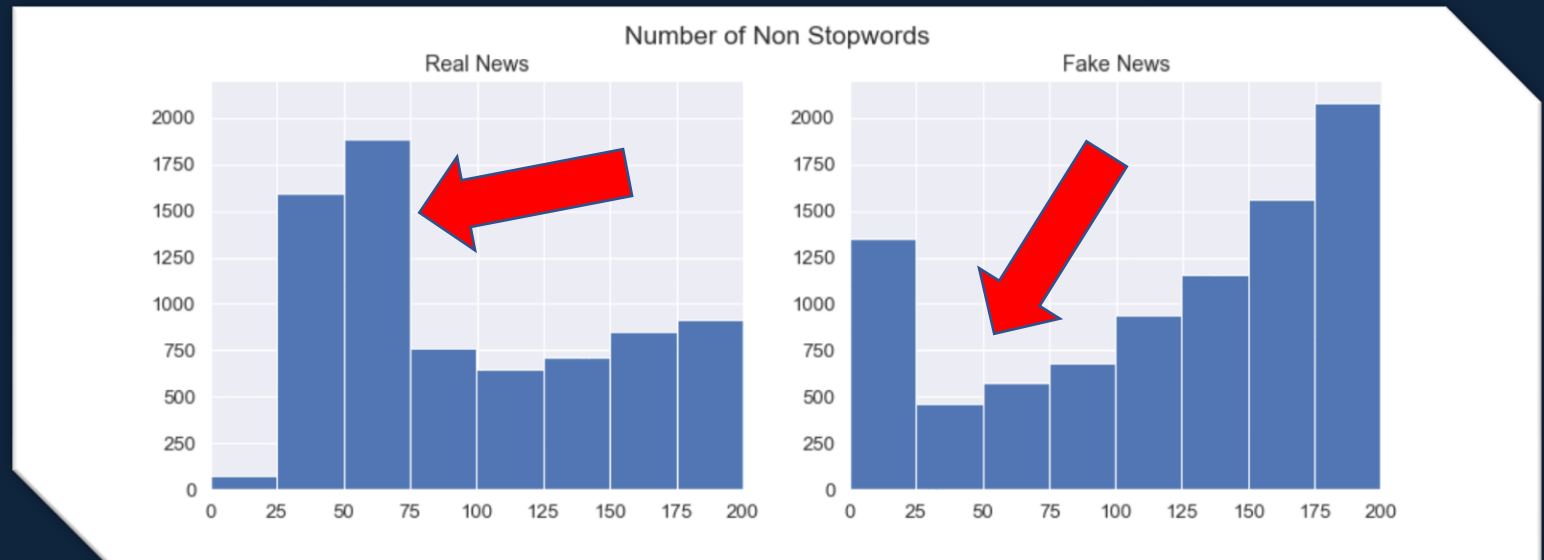
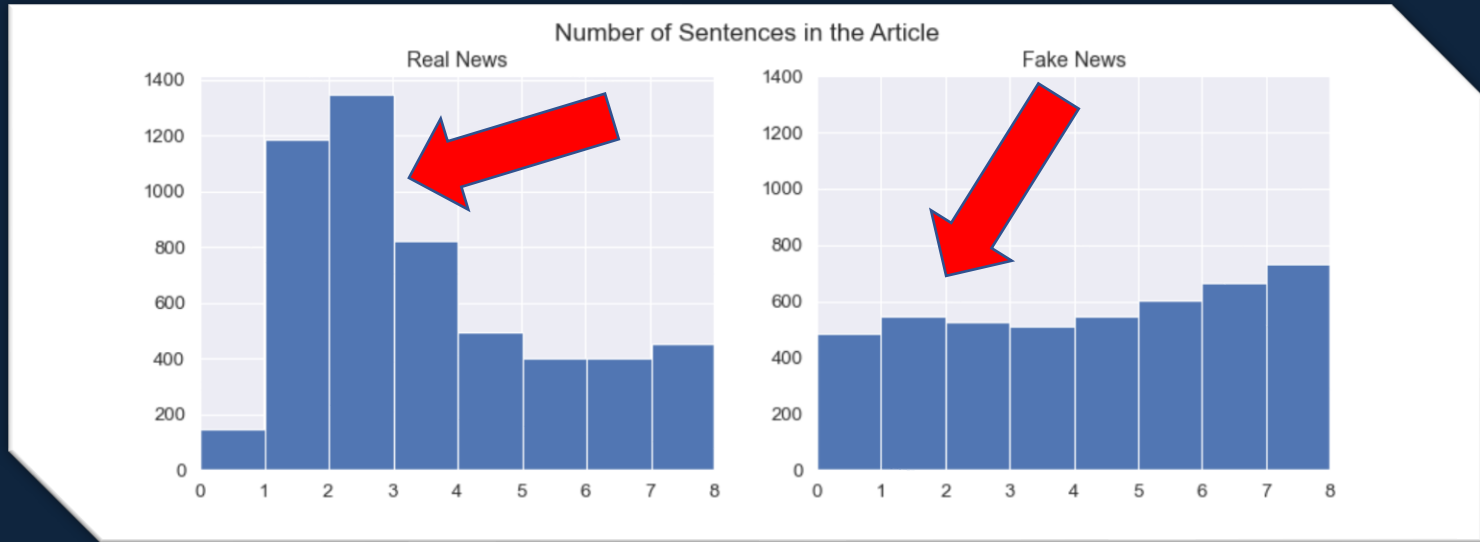


Stemming and  
Lemmatizing



Engineering  
Features

# Engineered Features







RECOMMENDATION

CNN should implement this model as it is **98.26% accurate** in identifying real vs. fake news articles

Automated process can help save time, money, and resources

# Next Steps



Analyze additional newer data



Engineer additional features



Run time analysis model



# Thank You!

Kevin Gross

[kevgross89@gmail.com](mailto:kevgross89@gmail.com)

[github.com/kevgross89](https://github.com/kevgross89)