

# News: Fake or Real?

Natural Language Processing Project  
Debora Bimbi

# Real or Fake?

lol! watch antifa thug cry like a baby  
after berkeley police unmask and  
arrest her

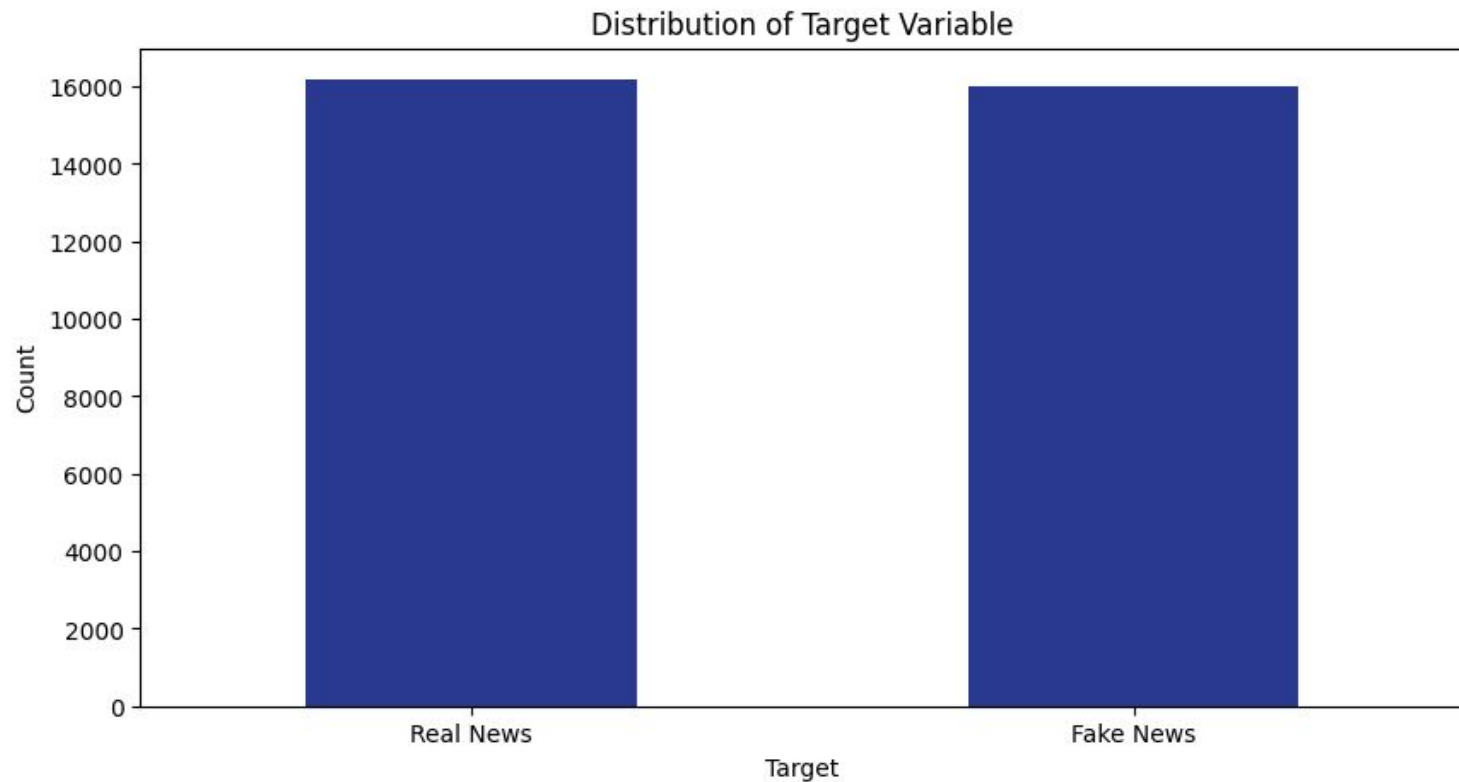
trump wants postal service to  
charge 'much more' for amazon  
shipments

---

# Unknown dataset

**9981 news quotes**

# Training dataset



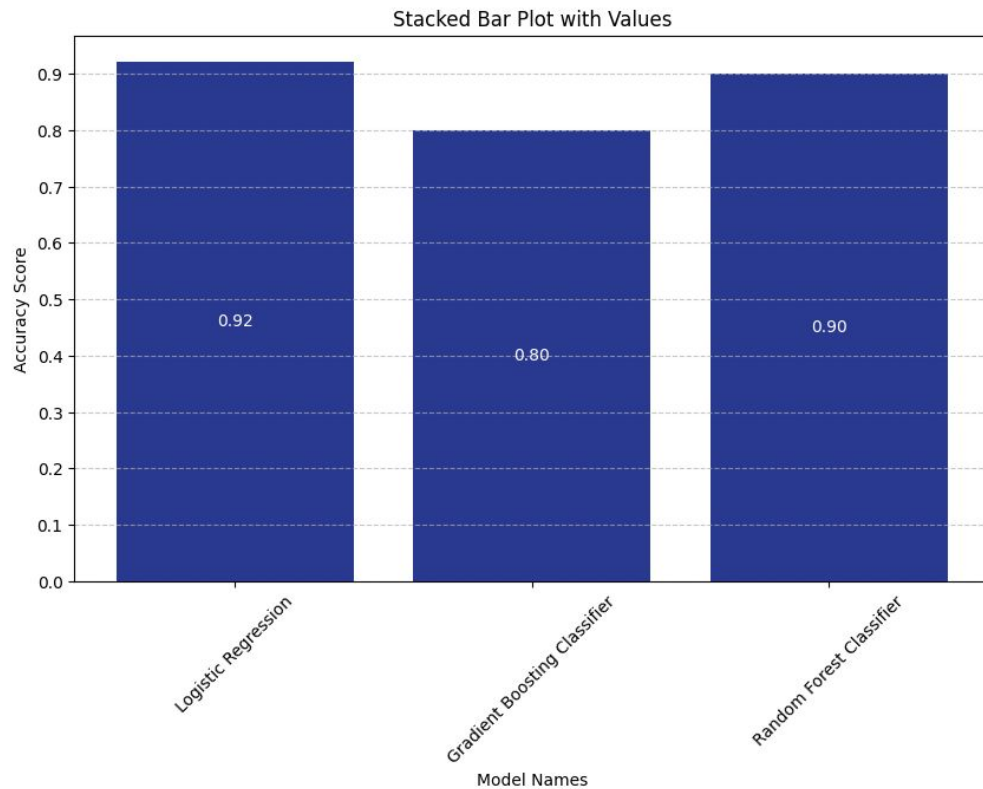
# TfidfVectorizer

Tf = Term Frequency

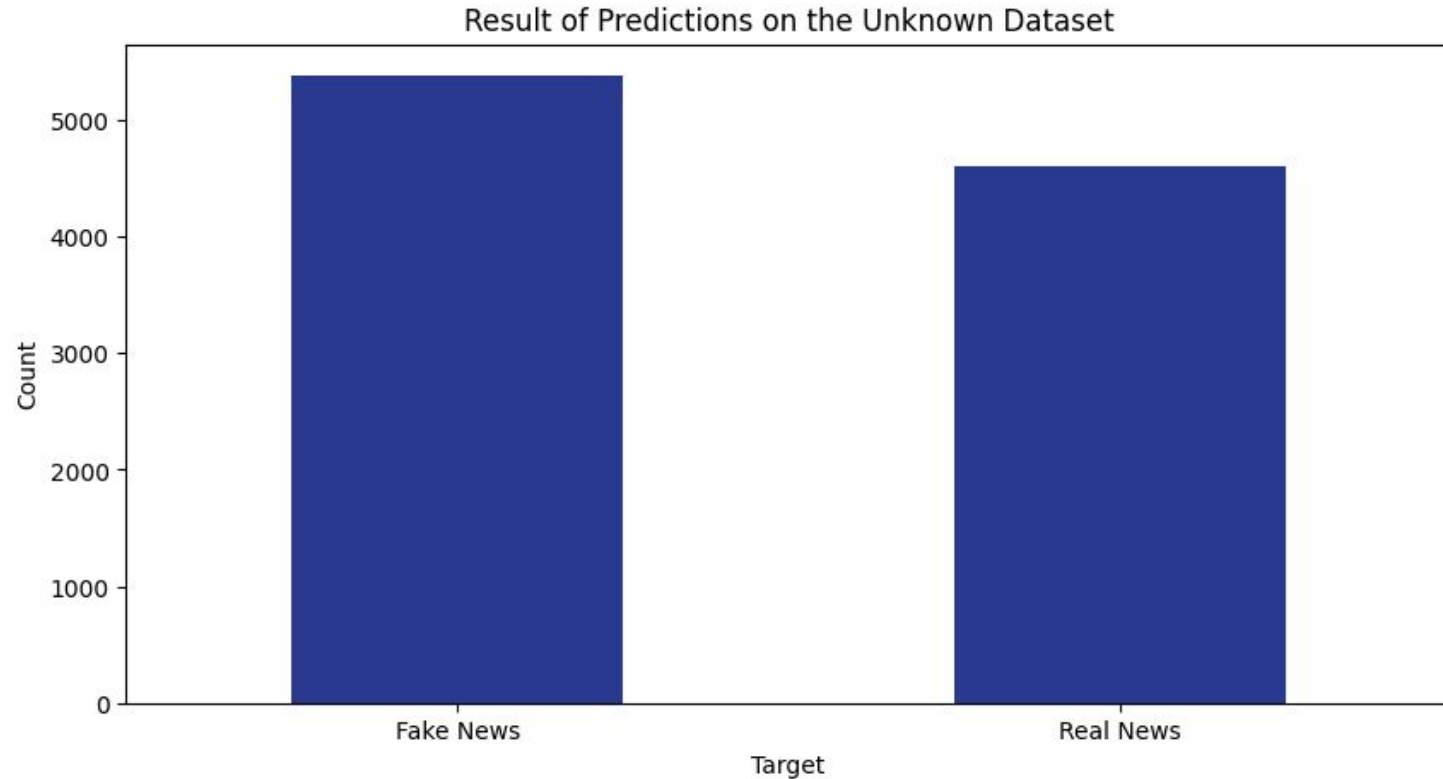
Idf = Inverse Document Frequency

Vectorizer

# Model's accuracy



# Final prediction result on unknown dataset



# What next ?

## **Semantic encoding**

Bag of Words → TFIDF → Word Embeddings → Deep Learning

## **Non semantic encoding**

One-hot Encoding, Label Encoding





Thank you for your attention!