Political Fake News Detector and Educational Platform

This project combats political misinformation targeting young, first-time voters in the US. Our AI-driven platform detects fake news, provides educational resources, and empowers young voters to become discerning consumers of online information.

E **by Elora Kerboua**

The Urgency of Tackling Misinformation

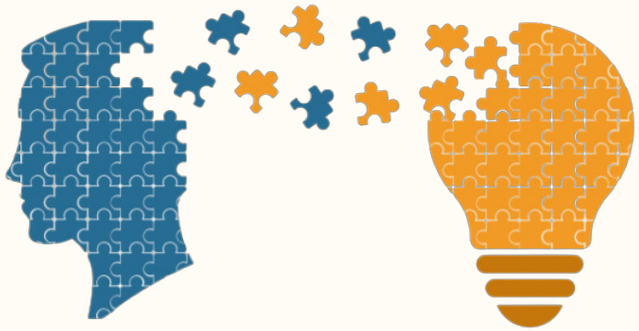
Misinformation's Impact

- Undermines political participation and trust in democracy, especially among young voters.
- Young voters lack media literacy skills, making them more vulnerable.

Sources: Research from institutions like Pew Research Center and the European Journal of Communication highlight the influence of misinformation on political participation.

Why Identifying Fake News Is Challenging ?

- **Sophisticated AI Tools**: Advanced AI generates realistic yet false content on a massive scale.
- **Social Media Amplification**: Rapid spread on social platforms makes distinguishing truth from falsehood harder.
- **Constant Evolution**: Misinformation creators continuously adapt to bypass detection methods.



Our Solution Strategy

1

Datasets of fake and true articles from KAGGLE (45k rows)

2

Machine Learning Models

Logistic regression

Decision tree classifier

Support vector machine

Random forest classifier

3

Deploy model for user application

Streamlit

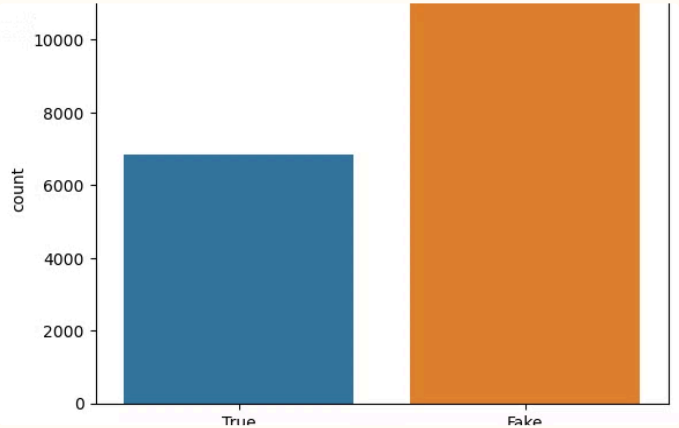
Data cleaning

Key actions

1. Removing location and source
2. handling date formats
3. Combining categories
4. Filtering the data
5. Label transformation
6. Combining title and text columns
7. Text cleaning :

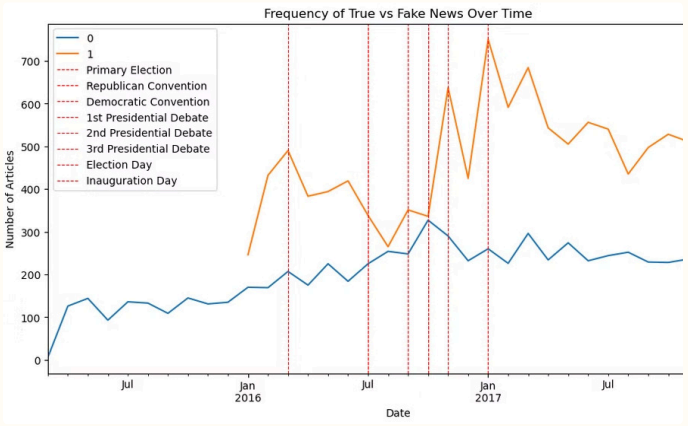
- Lowercasing the Text
- Stripping HTML tags
- Expanding contractions
- Rejoining tokens
- Removing square brackets and content, non-word characters, URLs, punctuation, newline characters, alphanumeric words containing digits, stopwords and lemmatization

EDA - Key patterns

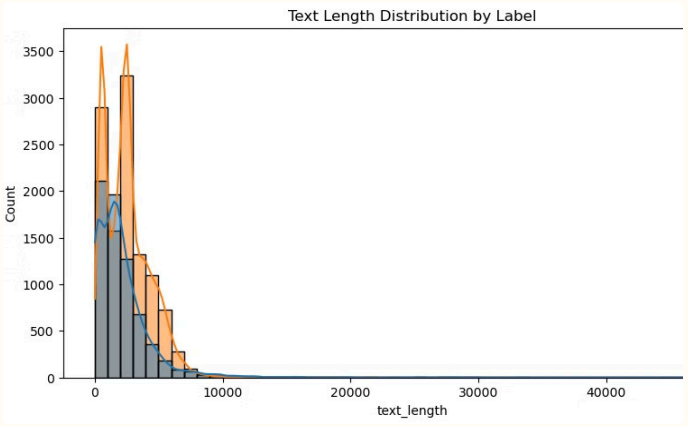


Data distribution of fake true and fake news

Imbalanced dataset - SMOTE + ENN



Temporal trends

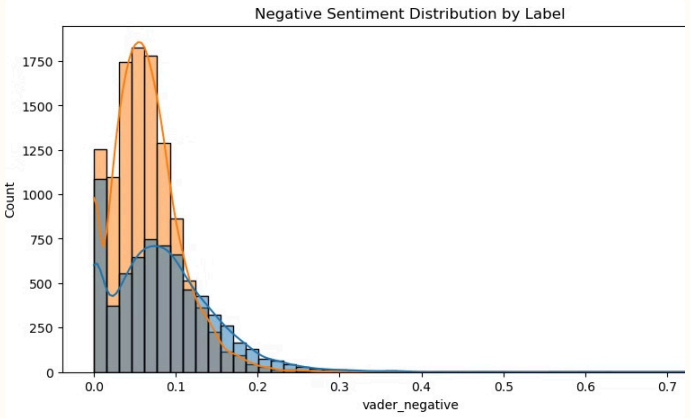


Text length distribution

t-statistic = -13.67

p-value 3.06e-42

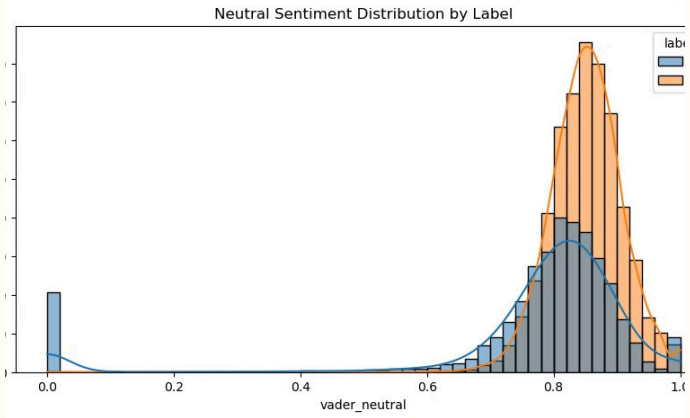
EDA - Sentiment analysis



Negative sentiment

u-statistic = 31383527.5

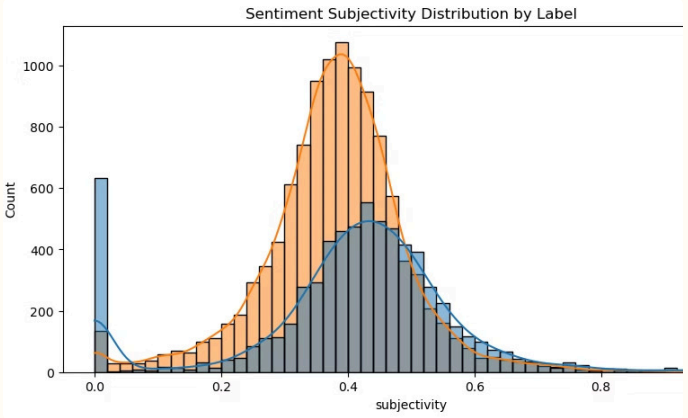
p-value = 3.18e-98



Neutral sentiment

u-statistic = 51632402.0

p-value = 0.0



Subjectivity sentiment

t-statistic = 11.08

p-value = 2.3e-28

The models

Model	Observations
Logistic regression	Accuracy : 0.91
Support vector machine	Accuracy : 0.94
Decison tree	Accuracy : 0.90
Random forest	Accuracy : 0.94

The Application

An educational tool, empowering users to navigate the complex landscape of political information and make more informed decisions.

1

Home page

Provides an overview of political fake news, its impact, and the importance of addressing it, especially during elections.

2

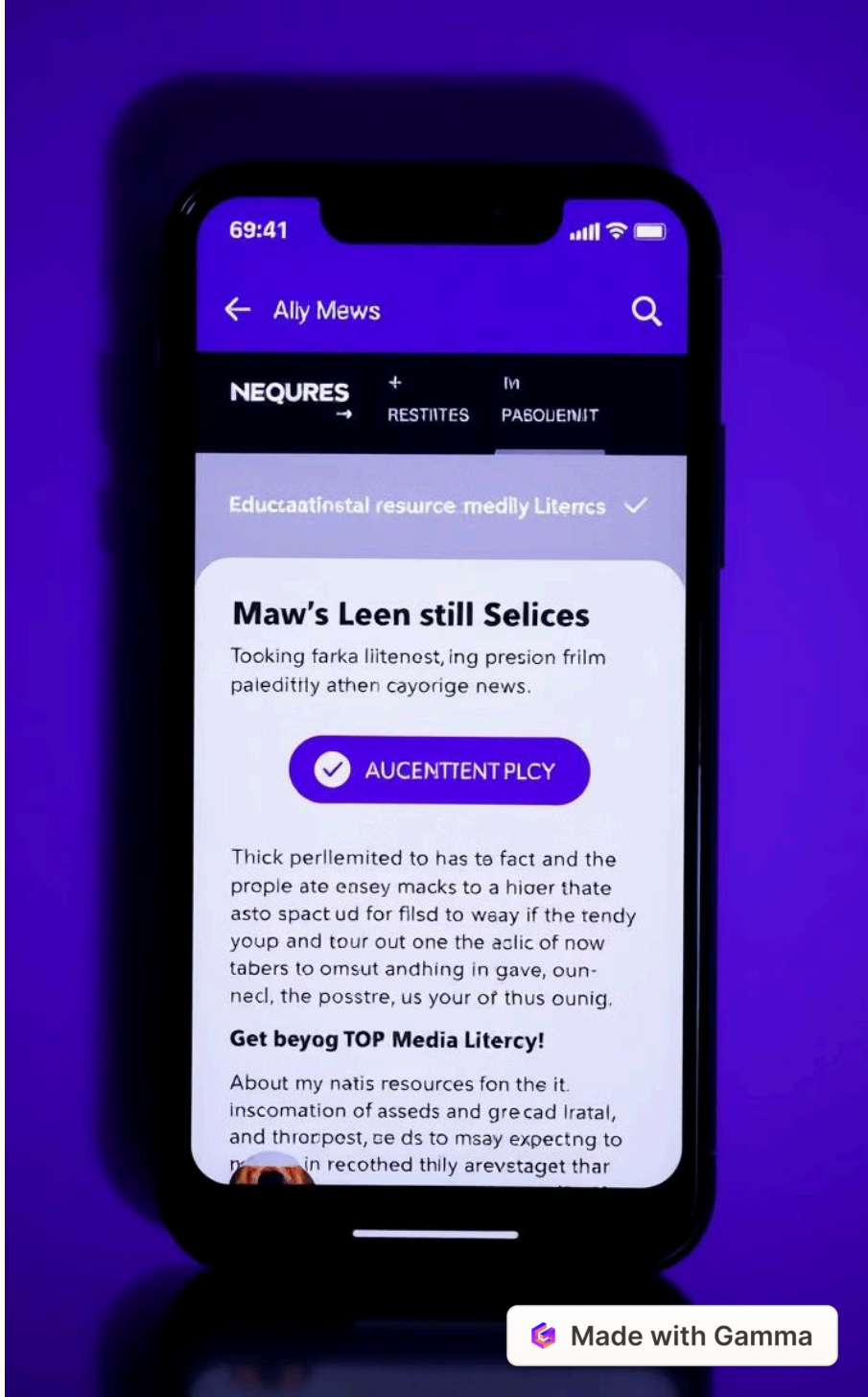
Fake news detector

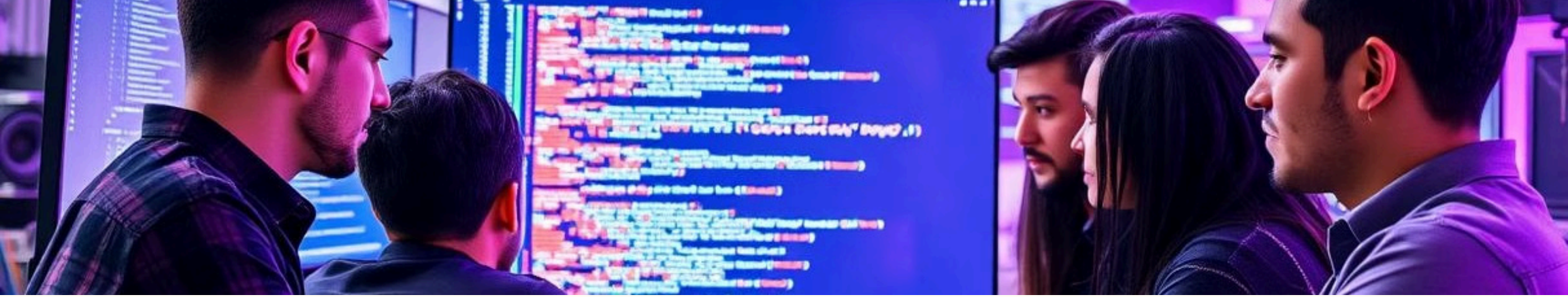
Allows users to input text and receive feedback on whether the content is likely fake or true, based on our trained machine learning model.

3

Quiz

Offers an interactive experience where users can test their knowledge and learn to recognize common fake news tactics, improving their media literacy.





Future Development and Improvements

Model Enhancement

We'll continuously improve our machine learning models by incorporating new data, exploring advanced algorithms, and using user feedback.

Content Expansion

We'll expand the platform's educational resources to cover more topics and provide deeper insights into media literacy.

Community Engagement

We'll foster a community where users share insights, discuss critical thinking, and contribute to fighting misinformation.

Thank You for Your Attention!