USC University of Southern California

# Sentiment and Semantic Driven Detection: Social Media Filters for Accurate Fake News Identification

Pathik Viramgama, Amartya Chaudhuri, Aditya Dutta, Mishal Patel, Karan Shah

**Our goal is to harness cutting-edge machine learning models to accurately classify and analyze news articles, distinguishing between real and fake news to combat misinformation effectively.**

USC University of Southern California

# Why It Matters

**Motivation:** In an era where misinformation can spread rapidly, distinguishing real news from fake has become crucial for informed decision-making.

**Problem Definition:** We tackle the challenge of developing a machine learning tool capable of accurately classifying news articles by their veracity, addressing data biases and the complexities of language.

USC University of Southern California

# Datasets In-Depth

## Dataset 1: Twitter News Dataset

- Number of Articles: 44,898
- Article Length Range: 1 to 51,794 characters (Average: 2,469)
- Balance: Almost equal number of real and fake news articles, around a 48.5 / 51.5 split.
- Features: Includes 'title', 'text', 'subject', and 'date' (primary features: 'title' , 'text')
- Consideration: Needed careful setting of max_length for sequence padding.

## Dataset 2: WELFake News Dataset

- Number of Articles: 72,134 (35,028 real and 37,106 fake)
- Columns: Serial number, Title, Text, Label (0 = fake, 1 = real).
- Source: A merger of four popular news datasets (Kaggle, McIntire, Reuters, BuzzFeed Political) to enhance diversity and prevent over-fitting.
- Utilization: Provides extensive text data for effective machine learning model training.

USC University of Southern California

# Models and Methodology

- We applied a multifaceted approach, employing both traditional machine learning and advanced deep learning models: BERT, RoBERTa, GPT-2 for nuanced text analysis, PCA with pycaret for dimensionality reduction, and LSTM for sequence modeling.

- This diversified methodology allows us to not only classify content accurately but also understand the semantic underpinnings of fake vs. real news.

# Preliminary Results

- Our LSTM model excelled in categorizing news types within both Dataset 1 & 2, achieving an accuracy rate of ~95%
- For fake news detection, BERT and RoBERTa models performed impressively on Dataset 2, with accuracies of 89.78% and 95.6%, respectively.
- GPT-2 achieved an impressive F1-score of 87% on Dataset 1.
- Machine Learning classifiers with 3-component PCA achieved 90.1% and 91.4% respectively for Dataset1 and Dataset2.

USC University of Southern California

# Machine Learning

| Model | Accuracy | AUC | Recall | Prec. | F1 |
|---|---|---|---|---|---|
| Extra Trees Classifier | 0.8818 | 0.9549 | 0.8714 | 0.8796 | 0.8754 |
| Random Forest Classifier | 0.8808 | 0.9502 | 0.8686 | 0.8798 | 0.8741 |
| Light Gradient Boosting Machine | 0.8701 | 0.9462 | 0.8687 | 0.8602 | 0.8644 |
| Gradient Boosting Classifier | 0.8675 | 0.9447 | 0.8731 | 0.8525 | 0.8626 |
| Ada Boost Classifier | 0.8604 | 0.9394 | 0.8523 | 0.8544 | 0.8533 |
| Quadratic Discriminant Analysis | 0.8603 | 0.9379 | 0.8706 | 0.8417 | 0.8559 |
| Ridge Classifier | 0.8598 | 0.0000 | 0.8668 | 0.8434 | 0.8549 |
| Linear Discriminant Analysis | 0.8598 | 0.9383 | 0.8670 | 0.8434 | 0.8550 |
| Naive Bayes | 0.8597 | 0.9387 | 0.8714 | 0.8402 | 0.8555 |
| Logistic Regression | 0.8590 | 0.9384 | 0.8558 | 0.8496 | 0.8526 |
| SVM - Linear Kernel | 0.8585 | 0.0000 | 0.8762 | 0.8355 | 0.8550 |
| K Neighbors Classifier | 0.8543 | 0.9218 | 0.8473 | 0.8471 | 0.8471 |
| Decision Tree Classifier | 0.8424 | 0.8416 | 0.8232 | 0.8426 | 0.8327 |
| Dummy Classifier | 0.5235 | 0.5000 | 0.0000 | 0.0000 | 0.0000 |

*Fig 1: On Dataset 1*

USC University of Southern California

# GPT-2

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Fake | 0.89 | 0.90 | 0.89 | 4524 |
| True | 0.89 | 0.88 | 0.89 | 4289 |
| accuracy | | | 0.88 | 8813 |
| macro avg | 0.90 | 0.91 | 0.88 | 8813 |
| weighted avg | 0.90 | 0.89 | | 8813 |

*Fig 2: On Dataset 1*

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Fake | 0.99 | 0.98 | 0.98 | 224 |
| True | 0.98 | 0.99 | 0.98 | 209 |
| accuracy | | | 0.98 | 433 |
| macro avg | 0.98 | 0.98 | 0.98 | 433 |
| weighted avg | 0.98 | 0.98 | 0.98 | 433 |

*Fig 3: On Dataset 2*

USC University of Southern California

# LSTM

```
276/276 ━━━━━━━━━━━━━━━━━━━━ 25s 89ms/step
```

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Fake | 0.99 | 0.99 | 0.99 | 4524 |
| True | 0.99 | 0.99 | 0.99 | 4289 |
| accuracy | | | 0.99 | 8813 |
| macro avg | 0.99 | 0.99 | 0.99 | 8813 |
| weighted avg | 0.99 | 0.99 | 0.99 | 8813 |

*Fig 4: On Dataset 1*

```
451/451 ━━━━━━━━━━━━━━━━━━━━ 53s 116ms/step
```

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Fake | 0.93 | 0.97 | 0.95 | 7089 |
| True | 0.97 | 0.93 | 0.95 | 7338 |
| accuracy | | | 0.95 | 14427 |
| macro avg | 0.95 | 0.95 | 0.95 | 14427 |
| weighted avg | 0.95 | 0.95 | 0.95 | 14427 |

*Fig 5: On Dataset 2*

# BERT

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.86 | 0.91 | 0.89 | 3213 |
| 1 | 0.92 | 0.87 | 0.89 | 3522 |
| accuracy | | | 0.89 | 6735 |
| macro avg | 0.89 | 0.89 | 0.89 | 6735 |
| weighted avg | 0.89 | 0.89 | 0.89 | 6735 |

*Fig 6: On Dataset 1*

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.99 | 0.99 | 0.99 | 7006 |
| 1 | 0.99 | 0.99 | 0.99 | 7302 |
| accuracy | | | 0.99 | 14308 |
| macro avg | 0.99 | 0.99 | 0.99 | 14308 |
| weighted avg | 0.99 | 0.99 | 0.99 | 14308 |

*Fig 7: On Dataset 2*

# RoBERTa

```
Accuracy: 0.9992204899777283
Precision: 0.9991550485847064
Recall: 0.999366152545954
F1 Score: 0.9992605894158656
```

*Fig 8: On Dataset 1*

```
Epoch 1/3 completed.
Epoch 2/3 completed.
Epoch 3/3 completed.
Accuracy: 0.9556192339949678, Precision: 0.9592991169977925, Recall: 0.9533799533799534, F1 Score: 0.9563303761777044
```

*Fig 9: On Dataset 2*

USC University of Southern California

# Challenges Encountered and Our Solutions

- We faced challenges related to data bias and computational resources. To mitigate these, we implemented rigorous data preprocessing and augmentation techniques and leveraged cloud computing platforms for efficient model training.

- Overfitting, especially in complex models like BERT, was addressed through regularization strategies and cross-validation, enhancing model robustness.

# Future Directions

- Going forward, we plan to expand our dataset collection to include multilingual news sources, improving our model's global applicability.
- We're also exploring the integration of more nuanced semantic analysis techniques and increasing the computational efficiency of our model training processes.
- Our ultimate aim is to develop a real-time fake news detection tool that can be deployed across various media platforms.

USC University of Southern California

# Welcoming Questions!