Leveraging AI in Combatting Fake News through Stance Detection

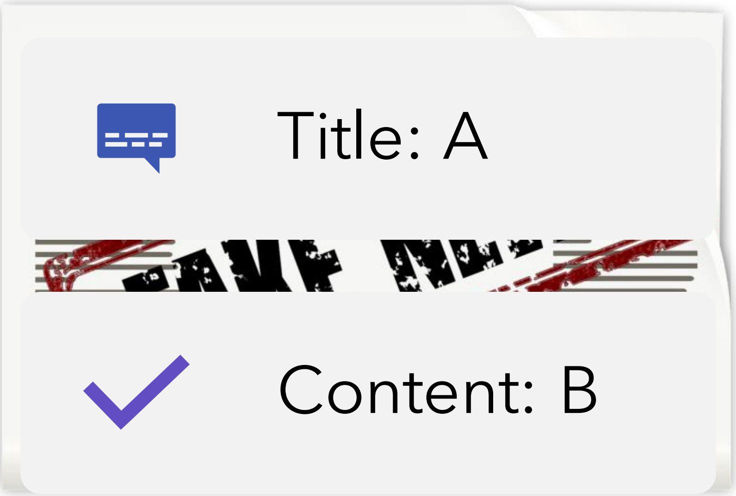
Zheng Zheng

Content

- Background
- Preprocessing and feature engineering
- Explanatory data analysis
- Implementing embedding layers
- Modeling
- Evaluation
- Business implementation advantages

Background

- **Identification of the Problem:**
  - Title mismatch the content



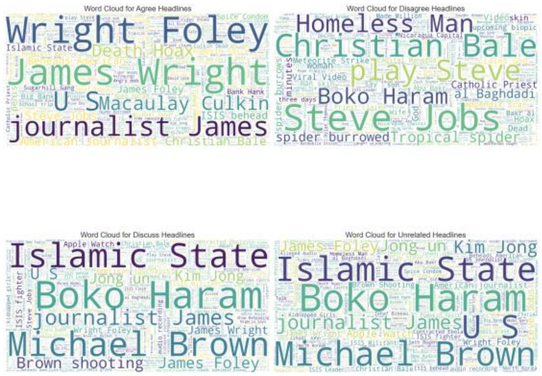
- **Purpose of the Project**:
  - developing an automated Stance Detection project to estimate the relative perspective of news article bodies in relation to their headlines

- **Data Source:**
  - FNC-1 Fake News competition GitHub repository (https://paperswithcode.com/dataset/fnc-1)
  - Stance Labels
    - Agrees: The body text agrees with the headline.
    - Disagrees: The body text disagrees with the headline.
    - Discusses: The body text discusses the same topic as the headline but does not take a position.
    - Unrelated: The body text discusses a different topic than the headline.

Data Preprocessing and feature engineering

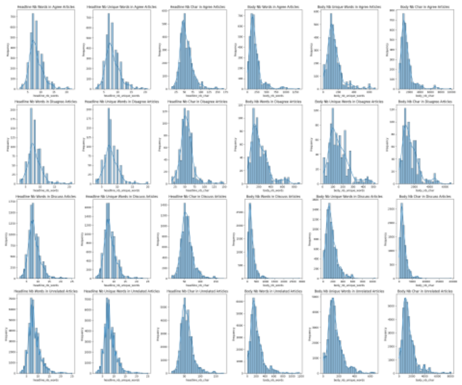
Subset of the dataset after basic processing(loading & merging)

	Body ID	articleBody	Headline	Stance
0	0	A small meteorite crashed into a wooded area i...	Soldier shot, Parliament locked down after gun...	unrelated
1	0	A small meteorite crashed into a wooded area i...	Tourist dubbed ‘Spider Man’ after spider burro...	unrelated
2	0	A small meteorite crashed into a wooded area i...	Luke Somers 'killed in failed rescue attempt i...	unrelated
3	0	A small meteorite crashed into a wooded area i...	BREAKING: Soldier shot at War Memorial in Ottawa	unrelated
4	0	A small meteorite crashed into a wooded area i...	Giant 8ft 9in catfish weighing 19 stone caught...	unrelated



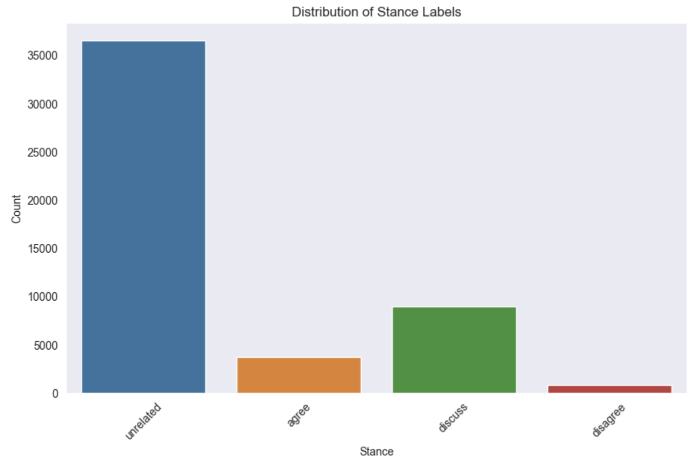
Feature engineering (six new numerical features)

1.Total Number of Words_headline: The total number of words in each headline.
2.Total Number of Unique Words_headline: The number of unique words used across the headlines, which indicates the diversity of vocabulary.
3.Total Number of Characters_headline: The character count in each headline, including spaces and punctuation.
4.Total Number of Words_body: The total number of words in each text body.
5.Total Number of Unique Words_body: The number of unique words used across the text bodies, which indicates the diversity of vocabulary.
6.Total Number of Characters_body: The character count in each text body, including spaces and punctuation.



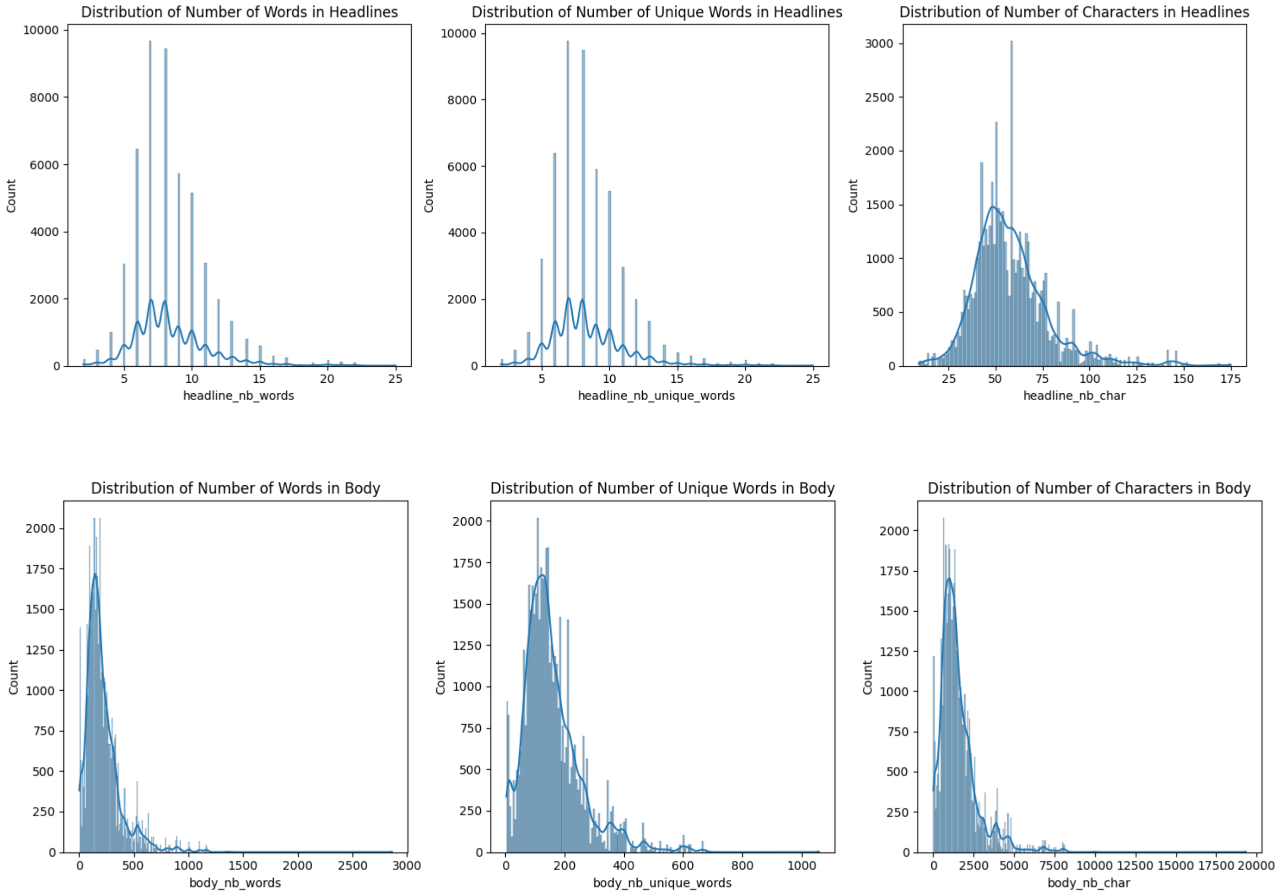
Explanatory data analysis

The distribution of stance labels



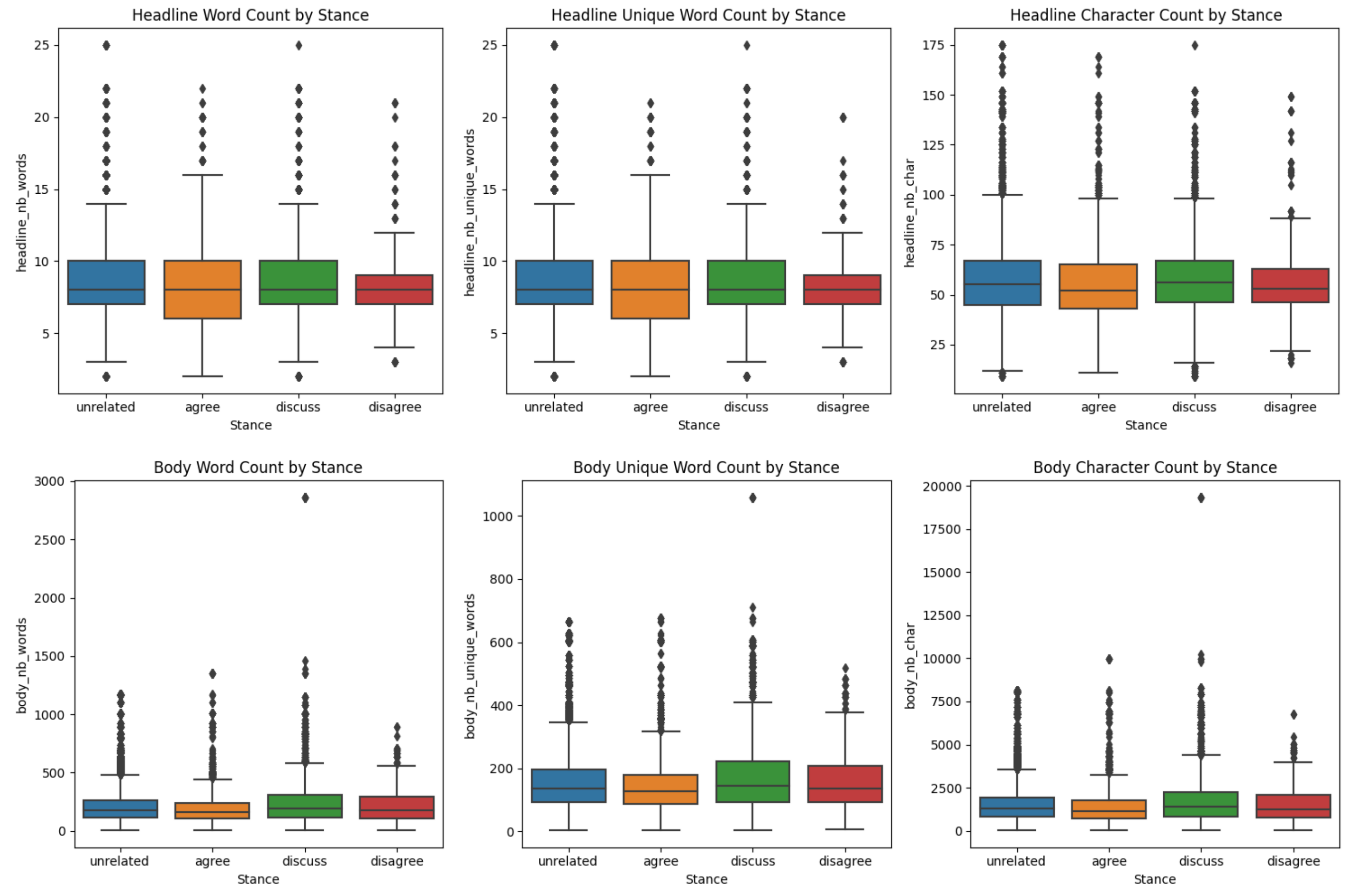
Imbalanced distribution database

The distribution of stance labels

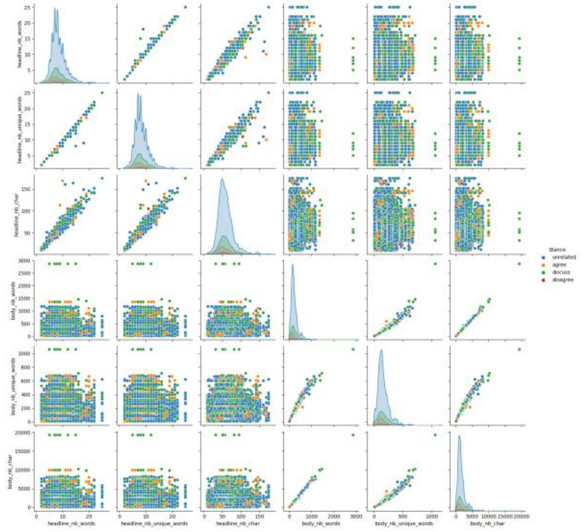


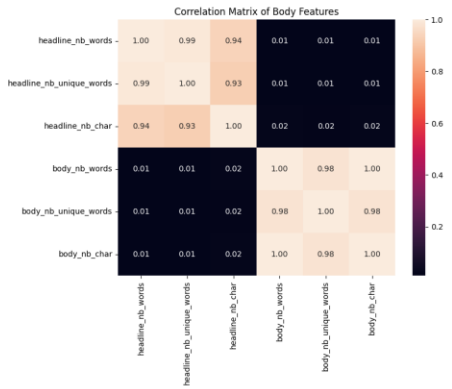
Explanatory data analysis

The distribution of these features across different stances



Correlation Between features





Implementing Embedding Layer

Embedding methods used

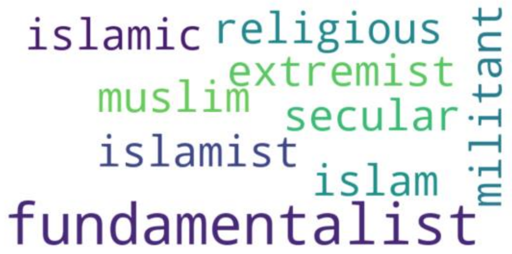
	Loaded words number	Running Time
GloVe Model	40000	20s
FastText Model	999994	3m
Word2Vec Model	3000000	1m

Evaluation for different methods(Quantitative Evaluation)

(Cosine similarity between 'islamic' and 'muslim')

- GloVe Model: 0.837352561
- FastText Model: 0.8157998
- Word2Vec Model: 0.7151465

Evaluation for different methods(qualitative evaluation- 'islamic')



Modeling

GloVe Embedding with Long Short-Term Memory(LSTM)

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding (Embedding)       (None, 100, 100)          2257800

 lstm (LSTM)                 (None, 50)                30200

 dropout (Dropout)           (None, 50)                0

 dense (Dense)               (None, 1)                 51

=================================================================
Total params: 2,288,051
Trainable params: 30,251
Non-trainable params: 2,257,800
_________________________________________________________________
```

Key Parameters:

MAX_SEQUENCE_LENGTH = 100
EMBEDDING_DIM = 100
MAX_VOCAB_SIZE = 20000
VALIDATION_SPLIT = 0.2
BATCH_SIZE = 128
EPOCHS = 10

GloVe Embedding with Convolutional Neural Network (CNN)

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding (Embedding)       (None, 100, 100)          2257800

 conv1d (Conv1D)             (None, 96, 128)           64128

 global_max_pooling1d (Glob  (None, 128)               0
 alMaxPooling1D)

 dense (Dense)               (None, 128)               16512

 dropout (Dropout)           (None, 128)               0

 dense_1 (Dense)             (None, 4)                 516

=================================================================
Total params: 2338956 (8.92 MB)
Trainable params: 81156 (317.02 KB)
Non-trainable params: 2257800 (8.61 MB)
_________________________________________________________________
```

Key Parameters:

MAX_SEQUENCE_LENGTH = 100
EMBEDDING_DIM = 100
MAX_VOCAB_SIZE = 20000
VALIDATION_SPLIT = 0.2

EPOCHS = 5

Modeling Evaluation

GloVe Embedding with LSTM

313/313 [==============================] - 9s 27ms/step

	precision	recall	f1-score	support
0	0.00	0.00	0.00	779
1	0.01	1.00	0.03	142
2	0.00	0.00	0.00	1816
3	0.00	0.00	0.00	7258
accuracy			0.01	9995
macro avg	0.00	0.25	0.01	9995
weighted avg	0.00	0.01	0.00	9995

- The accuracy is nearly 0.01, which is extremely low, suggesting the model is almost always incorrect in its predictions.
- The macro average F1-score is 0.01, and the weighted average F1-score is 0.00, both of which confirm the model's poor performance across all classes.
- The results indicate that the model is performing very poorly. It seems to be guessing nearly all instances as class 1, regardless of their true label, which is not a useful predictive behavior.

GloVe Embedding with Convolutional Neural Network (CNN)

157/157 [==============================] - 1s 6ms/step

	precision	recall	f1-score	support
0	0.60	0.29	0.39	372
1	0.25	0.04	0.07	80
2	0.67	0.55	0.60	911
3	0.85	0.94	0.89	3635
accuracy			0.81	4998
macro avg	0.59	0.46	0.49	4998
weighted avg	0.79	0.81	0.79	4998

- The accuracy (0.81) tells us that for the entire dataset, the model correctly predicts the class 81% of the time.
- The macro average F1-score (0.49) is much lower than the accuracy, which indicates a disparity among the class-specific performances.
- The weighted average F1-score (0.79) accounts for class imbalance by weighting the F1-score of each class by its support.

Business Implementation Advantages

- **Operational Efficiency**: By automating the process of Stance Detection, organizations can achieve a greater scale of operation in filtering and verifying news content. This reduces the reliance on manual fact-checking resources, thereby saving time and reducing operational costs.
- **Enhanced Accuracy**: The application of deep learning models to discern the stance of news articles relative to their headlines has demonstrated a high level of accuracy. This precision is vital in ensuring that only verified information reaches the public, maintaining the integrity of news dissemination.
- **Scalability**: The developed model is designed to handle the vast influx of data characteristic of the digital age. It is capable of processing large volumes of articles rapidly, ensuring that misinformation can be identified and addressed promptly.
- …

Thanks!