Identifying Deceptive Content: A Study on Clickbait and Fake News Detection

PRESENTED BY:

GROUP 53

Members:

Presenter: Shashank Rangarajan
Rishabh Ghosh
Michael Guastalla
Chia-Yu Tung
Edwin Wang

Problem Statement & Goals

Problem Statement: To accurately predict Fake News & Clickbait in Text data and study the relationship between them.

Goals:

- Implement existing SOTA models on Fake News and Clickbait detection
- Evaluate the models across various publicly available datasets, and study the results.
- Explore correlations between FakeNews and Clickbait predictions from these models, and draw conclusions

Datasets & Model Architecture

Datasets:

- **Fake News Detection:**
  - Fake News Corpus (Szpakowski, 2018) (fake v/s real)
  - Fakeddit (Nakamura et. al., 2020)
  - LIAR (Wang, 2017)
- **Clickbait Detection:**
  - Fake News Corpus (Szpakowski, 2018) (clickbait v/s real)
  - Webis (Potthast et. al., 2017)
  - Clickbait-detector (Mathur, 2020)

Model Architecture:

- **Models implemented:**
  - **Fake News:** RoBERTa -> FC, Siamese-BERT->FC
  - **Clickbait:** Glove-GRU->FC, Glove- LSTM->FC, BERT-GRU-FC, RoBERTa->FC
- **Final Model:**
  - **RoBERTa -> FC** for both Clickbait and FakeNews since our main goal is to study the relationship between the predictions of these models.

Experiments & Results

- **Fake News Detection Experiments:**

Dataset/Model	Fakeddit Model	FNC Model	Combined Model
Fakeddit Dataset	Accuracy:0.86 F1-score:0.82 AUC:0.85	Accuracy:0.54 F1-score:0.36 AUC:0.50	Accuracy:0.84 F1-score:0.79 AUC:0.83
FNC Dataset	Accuracy:0.52 F1-score:0.52 AUC:0.52	Accuracy:0.86 F1-score:0.86 AUC:0.86	Accuracy:0.85 F1-score:0.84 AUC:0.85
Combined Dataset	Accuracy:0.68 F1-score:0.65 AUC:0.68	Accuracy:0.71 F1-score:0.66 AUC:0.70	Accuracy:0.83 F1-score:0.82 AUC:0.83

- **An observation on the LIAR Dataset:**

Model/Dataset	AUC	F1
LIAR on LIAR	0.63	0.59
Fakeddit Model on LIAR	0.63	0.71

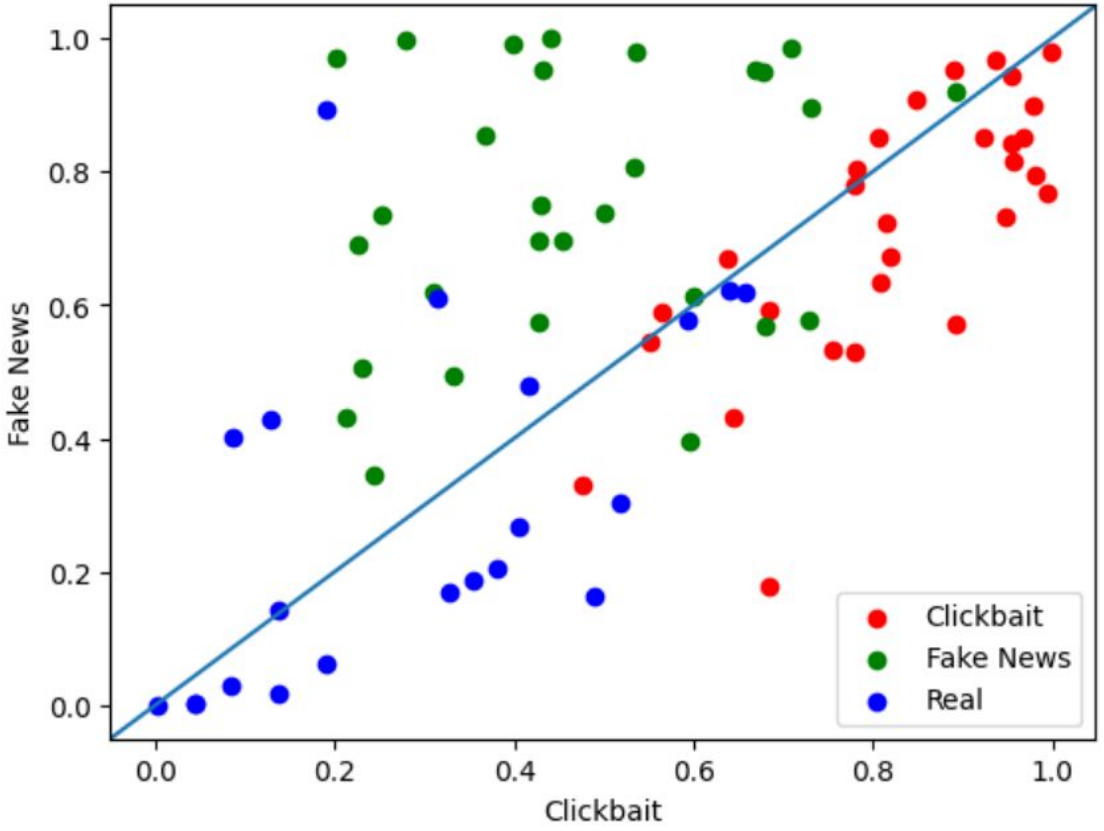
Experiments & Results

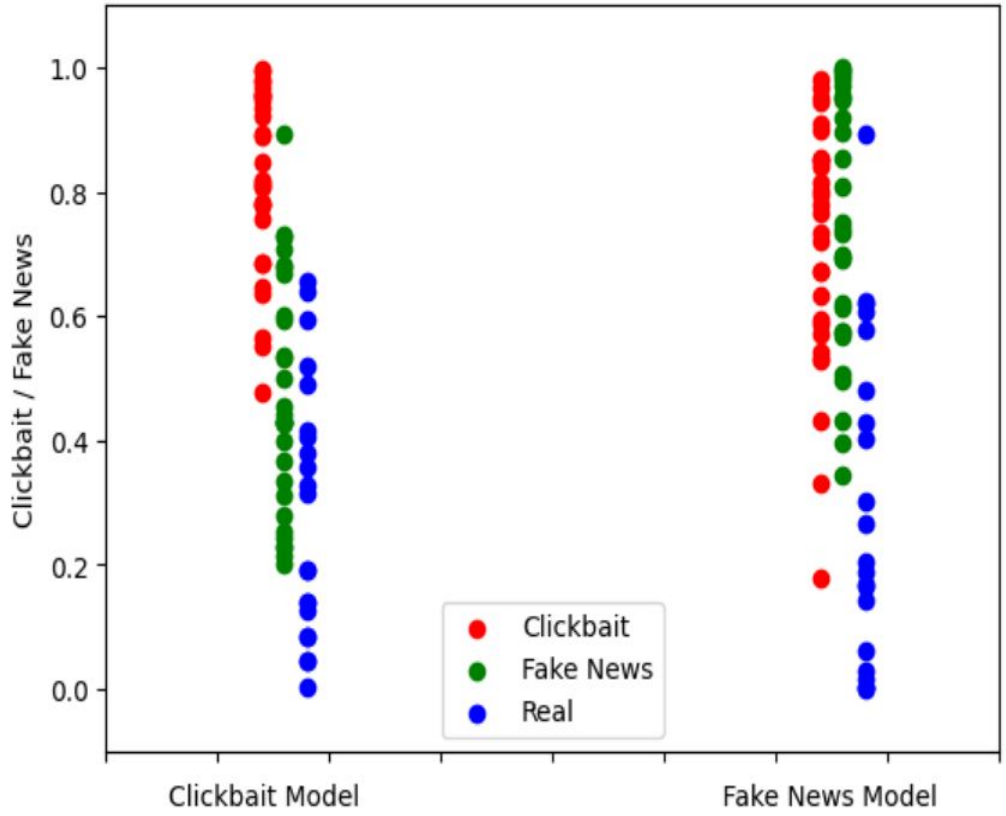
- **Clickbait Experiments:**

Dataset/Model	Mathur's Model	Webis Model	FNC Model	Combined Model
Mathur's Dataset	Accuracy:0.92 F1: 0.91 AUC: 0.92	Accuracy:0.75 F1: 0.68 AUC: 0.74	Accuracy:0.72 F1: 0.69 AUC: 0.72	Accuracy:0.82 F1: 0.81 AUC: 0.82
Webis 2017 Clickbait Dataset	Accuracy:0.53 F1: 0.45 AUC: 0.61	Accuracy:0.86 F1: 0.67 AUC: 0.77	Accuracy:0.54 F1: 0.35 AUC: 0.53	Accuracy:0.69 F1: 0.52 AUC: 0.68
FNC Dataset	Accuracy:0.56 F1: 0.46 AUC: 0.56	Accuracy:0.48 F1: 0.27 AUC: 0.48	Accuracy:0.89 F1: 0.89 AUC: 0.89	Accuracy:0.88 F1: 0.88 AUC: 0.88
Combined Dataset	Accuracy:0.59 F1: 0.52 AUC: 0.58	Accuracy:0.55 F1: 0.36 AUC: 0.54	Accuracy:0.84 F1: 0.83 AUC: 0.84	Accuracy:0.86 F1: 0.85 AUC: 0.86

Experiments & Results

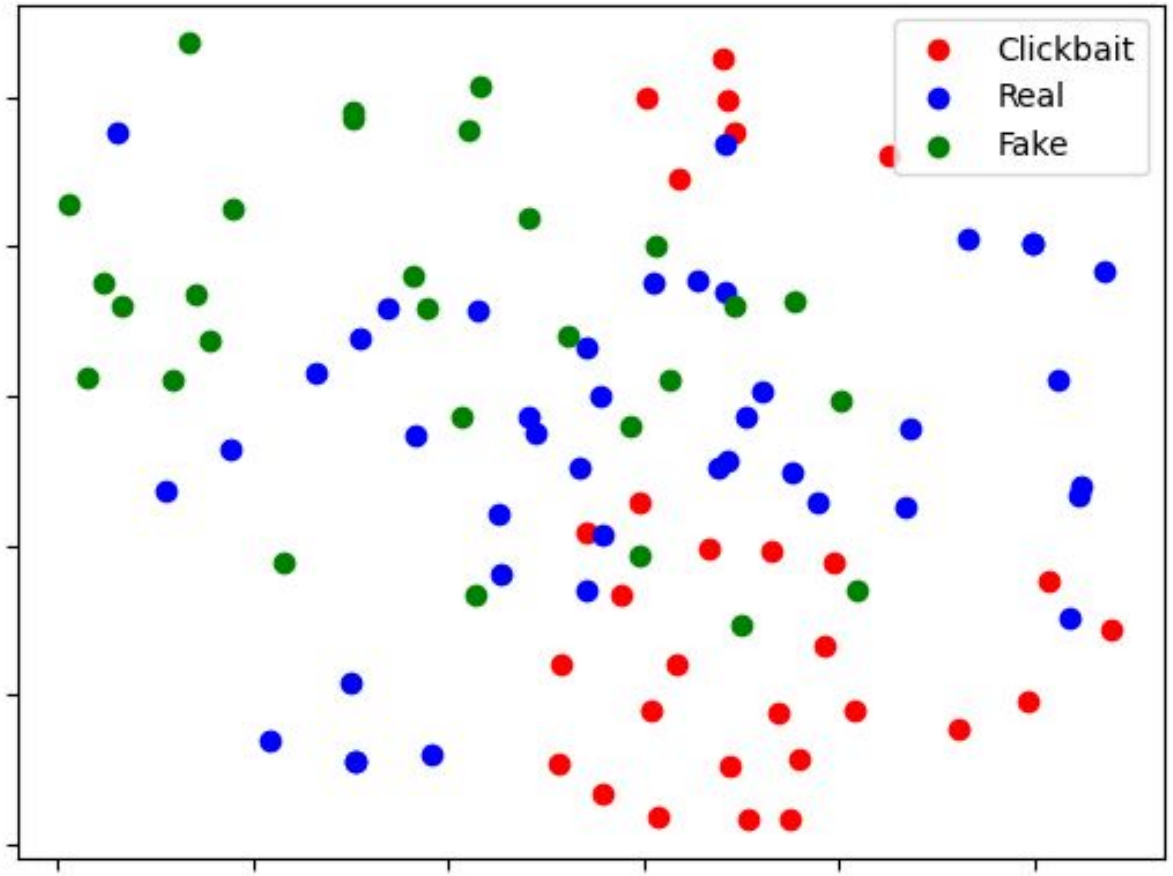
- Relationship between Clickbait and Fake News:

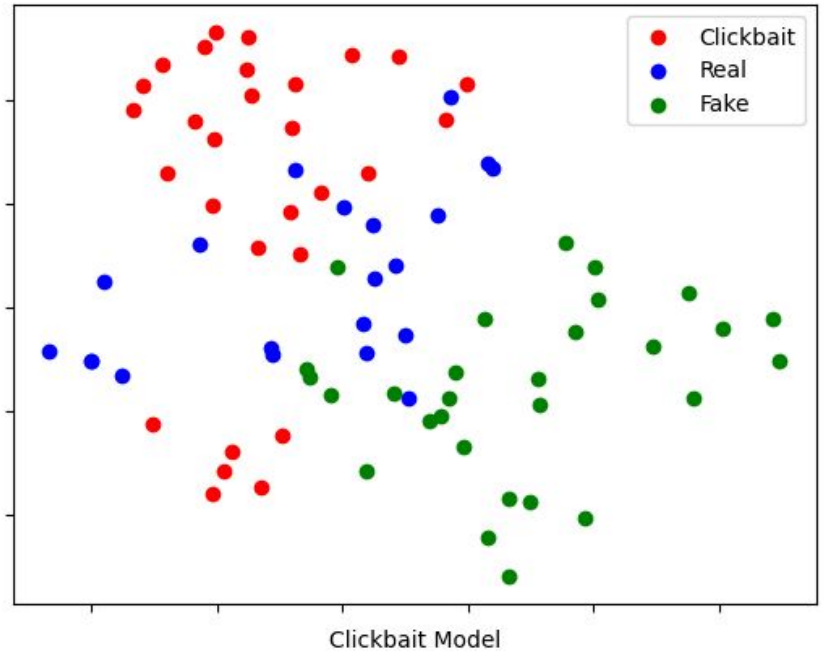


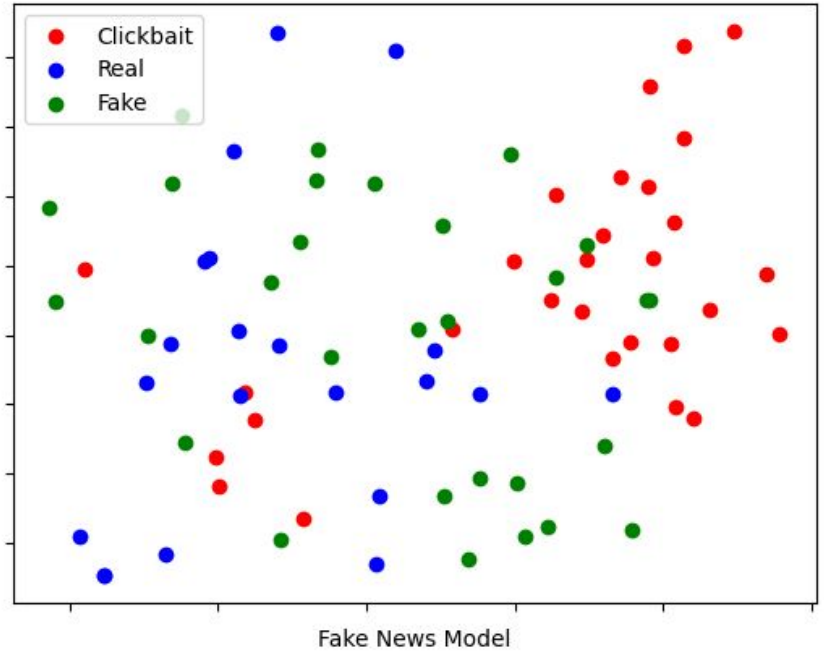


Experiments & Results

- Relationship between Clickbait and Fake News:







Conclusion & Future Work

- **Conclusions:**
  - Combining multiple datasets improves the robustness of fake news detection and clickbait detection, highlighting the importance of diversity in the training data.
  - There is a positive correlation between models trained for fake news detection and clickbait detection.
  - Models trained for fake news detection tend to be indispensable for clickbait detection, however the vice-versa does not hold for the experiments we performed.

- **Future Works:**
  - Future work could explore using graph neural networks for similar experiments and investigate the relationship between fake news and clickbait
  - Examine if fake news and clickbait share similar relationships in multimodality
  - The relationship between authenticity and clickbait in machine-generated text, specifically from AI models such as GPT, is an area for further exploration.

References

- Fake News Corpus (Szpakowski, 2018) - https://github.com/several27/FakeNewsCorpus
- Fakeddit (Nakamura et. al., 2020) - Kai Nakamura, Sharon Levy, and William Yang Wang. 2019. r/fakeddit: A new multimodal benchmark dataset for fine-grained fake news detection.
- LIAR - William Yang Wang. 2017. Liar. https://www.cs. ucsb.edu/~william/data/liar_dataset.zip
- Webis - Webis Clickbait Challenge Dataset. 2017. clickbaitchallenge dataset. https://zenodo.org/record/6362726#.YsbdSTVBzrk
- Clickbait Detector - Saurabh Mathur. 2017. clickbait-detector. https://github.com/saurabhmathur96/clickbait-detector/tree/master/data

Thank You !