Fake News Detection and Source Credibility Analysis
Using Machine Learning and Deep Learning

Lingala Anwesh [B200462]

Santhapuri Nishanth [B200628]

Pittala Mahesh [B200240]

Guide: Govardhani

Dept of CSE,RGUKT Basar

Problem Statement

- Rapid spread of fake news on social media & online platforms.

- Difficult to verify source credibility manually.

- Need for automated NLP + DL approaches for reliable detection.

Objectives

- Develop a machine learning and deep learning–based system for fake news detection.

- Perform source credibility analysis by integrating multiple datasets.

- Use NLP techniques to preprocess and analyze news text.

Datasets (Multi-source Input)

- FakeNewsNet → GossipCop + PolitiFact news.

- LIAR Dataset → Political statements labeled by fact-checkers.

- Kaggle Fake News Dataset → News articles from various domains.

- ✅ Combining datasets → Robust, large-scale dataset for training ML/DL models.

Methodology

- Data Preprocessing (NLP): Tokenization, Stopword Removal, Lemmatization.

- Feature Engineering: Word counts, TF-IDF, embeddings.

- Models (ML + DL): Logistic Regression, Random Forest, XGBoost.

- Deep Learning: LSTMs, CNNs for text.

- Transformer Models: BERT, RoBERTa for state-of-the-art performance.

- Source Credibility: Domain trust score + historical fact-checking.

Expected Outcomes

- Accurate fake vs. real news classification.

- Source credibility scores for news domains.

- Visualization: Classification probabilities, Attention heatmaps (Transformer explainability).

Future Scope

- Real-time fake news monitoring using live social media streams.

- Multi-lingual fake news detection.

- Mobile app + browser extension.

- Cloud-based scalable deployment.