# 

| **Name** | **Rollno** |
| --- | --- |
| Mohammad Ahmed Ansari | 03 |
| Om Ghat | 13 |
| Sagar Ghat | 14 |
| Samiksha Jagne | 17 |

# **Topic: Fake News Detection using NLP and Deep Learning.**

# **1. Introduction**

In the digital era, the spread of misinformation through social media and news outlets has become a significant societal threat. Fake news can manipulate public opinion, influence elections, and cause widespread panic. Detecting such misinformation using automated techniques is crucial. This project focuses on using Natural Language Processing (NLP) and deep learning to accurately classify news as real or fake. The solution will combine text preprocessing, feature extraction, and transformer-based models to improve detection performance.

### **2. Proposed System**

We propose a multi-stage pipeline for fake news detection:

1. Data Collection: Use publicly available fake news datasets (LIAR, FakeNewsNet).
2. Text Preprocessing: Tokenization, stop word removal, stemming/lemmatization, etc.
3. Feature Extraction:  
   * Traditional: TF-IDF, CountVectorizer
   * Semantic: BERT embeddings (CLS token)
4. Model Training:  
   * Baselines: Logistic Regression, SVM
   * Deep Learning: LSTM, BiLSTM
   * Transformers: BERT, RoBERTa
5. Evaluation:
   * Metrics: Accuracy, Precision, Recall, F1-score, ROC-AUC
6. Deployment: Streamlit or Flask app to classify user-inputted headlines/articles

### **3. Technologies & Tools Required**

| **Category** | **Tools/Technologies** |
| --- | --- |
| Languages | Python 3.9+ |
| Libraries | NLTK, spaCy, Scikit-learn, TensorFlow / PyTorch, Transformers (HuggingFace) |
| Models | BERT, RoBERTa, LSTM |
| Visualization | Matplotlib, Seaborn |
| Data | LIAR, FakeNewsNet, ISOT Fake News Dataset |
| IDE | Jupyter Notebook |
| Deployment | Streamlit or Flask for web deployment |

### **4. Research Papers**

| **No** | **Title** | **Authors** | **Year** | **Source** | **Link** |
| --- | --- | --- | --- | --- | --- |
| 1 | A Transformer-based approach for Fake News detection using Time Series Analysis. | [Aniq](https://ieeexplore.ieee.org/author/37089911253) et al. | 2023 | IEEE | [Link](https://ieeexplore.ieee.org/document/10178457) |
| 2 | Fake News Detection Using Deep Learning and Natural Language Processing. | Anand et al. | 2022 | IEEE | [Link](https://ieeexplore.ieee.org/document/10068440) |
| 3 | Fake News Detection using NLP. | Zubiaga et al. | 2022 | IEEE | [Link](https://ieeexplore.ieee.org/document/10100305) |
| 4 | A Survey on Natural Language Processing for Fake News Detection. | Ray et al. | 2020 | LREC | [Link](https://aclanthology.org/2020.lrec-1.747.pdf) |
| 5 | Ensemble Techniques for Robust Fake News Detection: Integrating Transformers, Natural Language Processing, and Machine Learning. | Roy et al. | 2024 | MDPI | [Link](https://www.mdpi.com/1424-8220/24/18/6062) |