PANIMALAR ENGINEERING COLLEGE

An Autonomous Institution, Affiliated to Anna University, Chennai
A Christian Minority Institution
(JAISAKTHI EDUCATIONAL TRUST)
Approved by All India Council for Technical Education

DETECTION OF NEWS WHETHER ITS FAKE OR NOT USING DATA ANALYSTICS

Team Members: VIGNESH KUMAR R(211423104732),
VIMALRAJ M(211423104738)
Department: Computer Science and Engineering
Guide Name:Dr. SATHIYA V
Coordinator Name: Dr. Subedha.V ,Mrs. SHARMILA S
Domain: Data Science
SDG 4: Quality Education
SDG 16: Peace, Justice and Strong Institutions

Social relevance

- Safeguard of public health by combating medical misinformation
- Support informed decision-making in society.
- Linked SDGs:
- SDG 4: Quality Education 

- SDG 16: Peace, Justice and Strong Institutions 

- Prevent social unrest caused by rumors and false report.

Base Paper Details

- Title:Fake News Detection Using Deep Learning and Transformer-Based Model
- Authors:Paliwal Mohan Subhash; Deepa Gupta; Suja Palaniswamy; Manju Venugopalan
- Journal:IEEE Xplore
- Year: 2023
- Publisher: IEEE

Abstract

- Fake news detection is a critical data science application that uses machine learning and natural language processing to identify misleading content.
- With the rapid spread of misinformation on social media, manual fact-checking is insufficient.
- Data-driven methods provide scalable and accurate solutions to maintain information integrity and trust in digital communication.

- Fake news detection employs a combination of content-based, user-based, and propagation-based analysis.
- Content-based methods analyze linguistic and semantic features of text to identify deceptive writing patterns.
- User-based analysis evaluates the credibility and past behavior of information sources, while propagation-based methods study how news spreads across social networks.

Agenda

- Introduction
- Objective
- Literature Survey
- Problem Statement
- Architecture Diagram
- Use Case and Other Diagrams
- Modules
- Algorithms / Methodology
- Testing
- Test Cases and Validation
- Performance Analysis
- Screenshots
- Conclusion
- Future Work
- References

Introduction

- Fake news has become a global issue, affecting political, economic, and social systems.
- Fake news is false or misleading information disguised as news, aiming to manipulate opinions or gain attention. With social media, it spreads rapidly, influencing public opinion and causing confusion. Manual detection is slow and unreliable.
- We use data analystics and machine learning algorithm to detect the fake news

Objective

- Detect fake news using data analytics and ML
- Improve accuracy and reliability of information
- Build scalable and automated detection system
- Enhance decision-making with credible sources

Literature Survey

#	AUTHORS	YEAR	TITLE	KEY CONTRIBUTION
1	**Shu, Sliva, Wang, Tang, Liu** Kai Shu; Amy Sliva; Suhang Wang; Jiliang Tang; Huan Liu	2017	Fake News Detection on Social Media: A Data Mining Perspective	Early survey framing content, social context, and knowledge cues.
2	**Ruchansky, Seo, Liu** Natali Ruchansky; Sungyong Seo; Yan Liu	2017	CSI: A Hybrid Deep Model for Fake News Detection	Hybrid model combining content, user responses, and source behavior.
3	**Wang** William Yang Wang	2017	"LIAR": A New Benchmark Dataset for Fake News Detection	Introduces the LIAR dataset (12.8K statements) with fine-grained labels.
4	**Ahmed, Traoré, Saad** H. Ahmed; I. Traoré; S. Saad	2018	Detecting Hoaxes, Fake News, and Clickbait Using ML	Classical ML feature engineering for news credibility signals.
5	**Singh et al.** A. Singh; et al.	2019	Review on Fake News Detection Methods	Overview of content, context, and hybrid approaches.

6	**Thorne, Vlachos, Christodoulopoulos, Mittal** J. Thorne; A. Vlachos; C. Christodoulopoulos; A. Mittal	2018	FEVER: Fact Extraction and VERification	Large-scale fact-verification dataset and task.
7	**Shu, Wang, Liu** K. Shu; S. Wang; H. Liu	2019	Beyond News Contents: The Role of Social Context for FND	Models leveraging user/social signals with content.
8	**Pérez-Rosas, Kleinberg, Lefevre, Mihalcea** V. Pérez-Rosas; B. Kleinberg; A. Lefevre; R. Mihalcea	2018	Automatic Detection of Fake News	Linguistic features and new datasets for news veracity.
9	**Yang, et al.** Y. Yang; et al.	2019	Social Context-based Detection	Incorporates user-interaction and propagation signals.
10	**Zhou, Zafarani** X. Zhou; R. Zafarani	2020	Fake News Early Detection: A Survey	Early-stage detection before wide propagation.

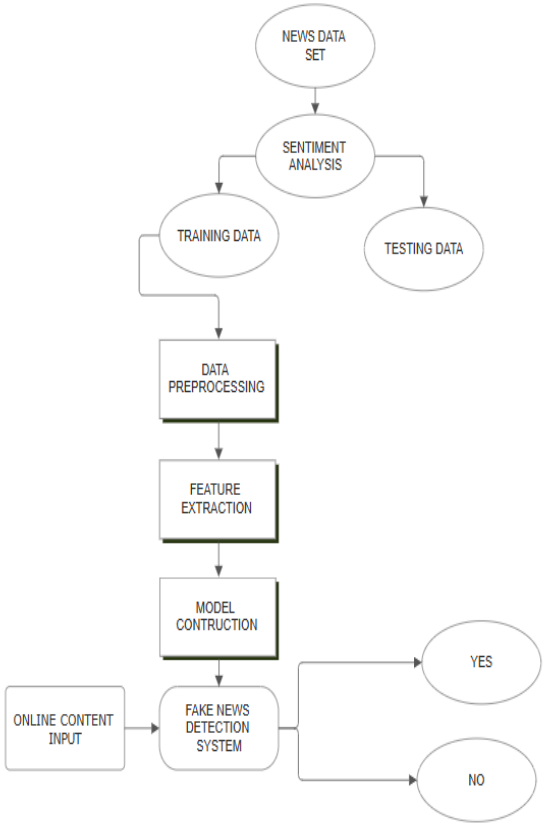
11	**Sharma, Qian, Jiang, Ruchansky, Zhang, Liu** K. Sharma; F. Qian; H. Jiang; N. Ruchansky; M. Zhang; Y. Liu	2019	Combating Fake News: Identification and Mitigation	Comprehensive review of detection & mitigation techniques.
12	**Oshikawa, Qian, Wang** R. Oshikawa; J. Qian; W. Y. Wang	2020	A Survey on NLP for Fake News Detection	Focus on text/NLP pipelines and datasets.
13	**Volkova, Shaffer, Jang, Hodas** S. Volkova; K. Shaffer; J. Jang; N. Hodas	2017	Separating Facts from Fiction: Linguistic Cues of Misinformation	Stylometry and affect for deception signals.
14	**Shu, Mahudeswaran, Wang, Lee, Liu** K. Shu; D. Mahudeswaran; S. Wang; D. Lee; H. Liu	2018	FakeNewsNet: A Data Repository with News Content and Social Context	Unified benchmark repository for FND research.
15	**Tacchini, Ballarin, Della Vedova, Moret, de Alfaro** E. Tacchini; F. Ballarin; M. L. Della Vedova; S. Moret; L. de Alfaro	2017	Some Like it Hoax: Automated Fake News Detection in Facebook	User-interaction features for social hoax detection.
16	**Vosoughi, Roy, Aral** S. Vosoughi; D. Roy; S. Aral	2018	The Spread of True and False News Online	Large-scale diffusion analysis; false news spreads faster.
17	**Zhou, Zafarani, Shu, Liu** X. Zhou; R. Zafarani; K. Shu; H. Liu	2019	Fake News: Fundamental Theories, Detection Strategies, and Challenges	Foundational taxonomy and open problems.

Problem Statement

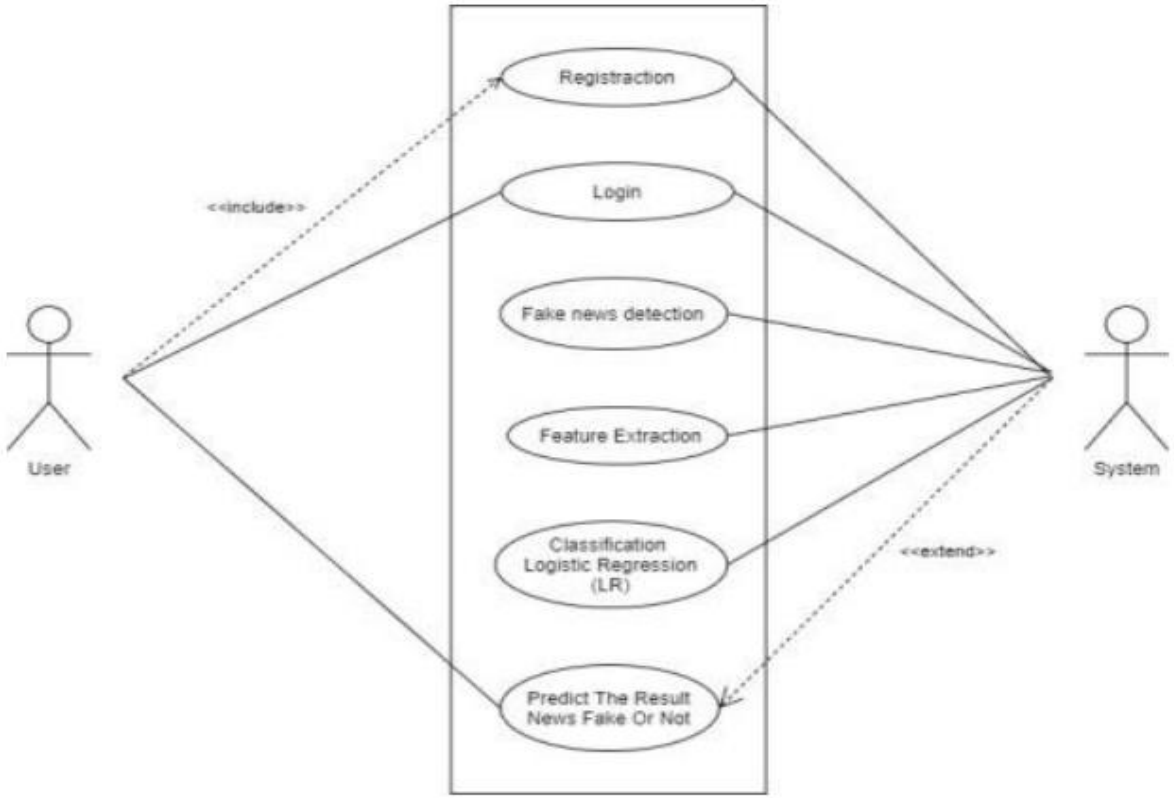
- Fake news misleads public opinion
- Difficult to manually identify misleading content
- Existing methods lack accuracy and scalability
- Need an automated, data-driven fake news detection approach

System Architecture

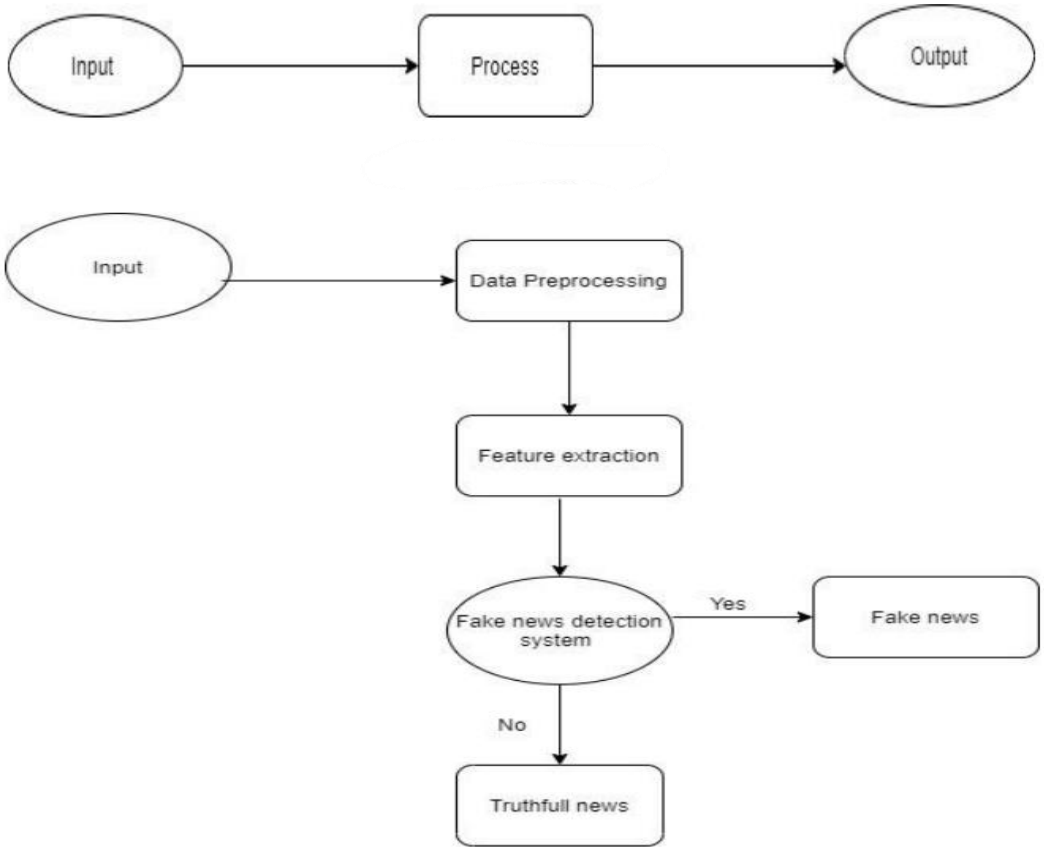
1. Data Input (News articles)
2. Text Preprocessing (cleaning, tokenizing)
3. Feature Extraction (TF-IDF/Word2Vec)
4. ML Model (SVM/Logistic Regression/LSTM)
5. Prediction Output (Fake or Real)



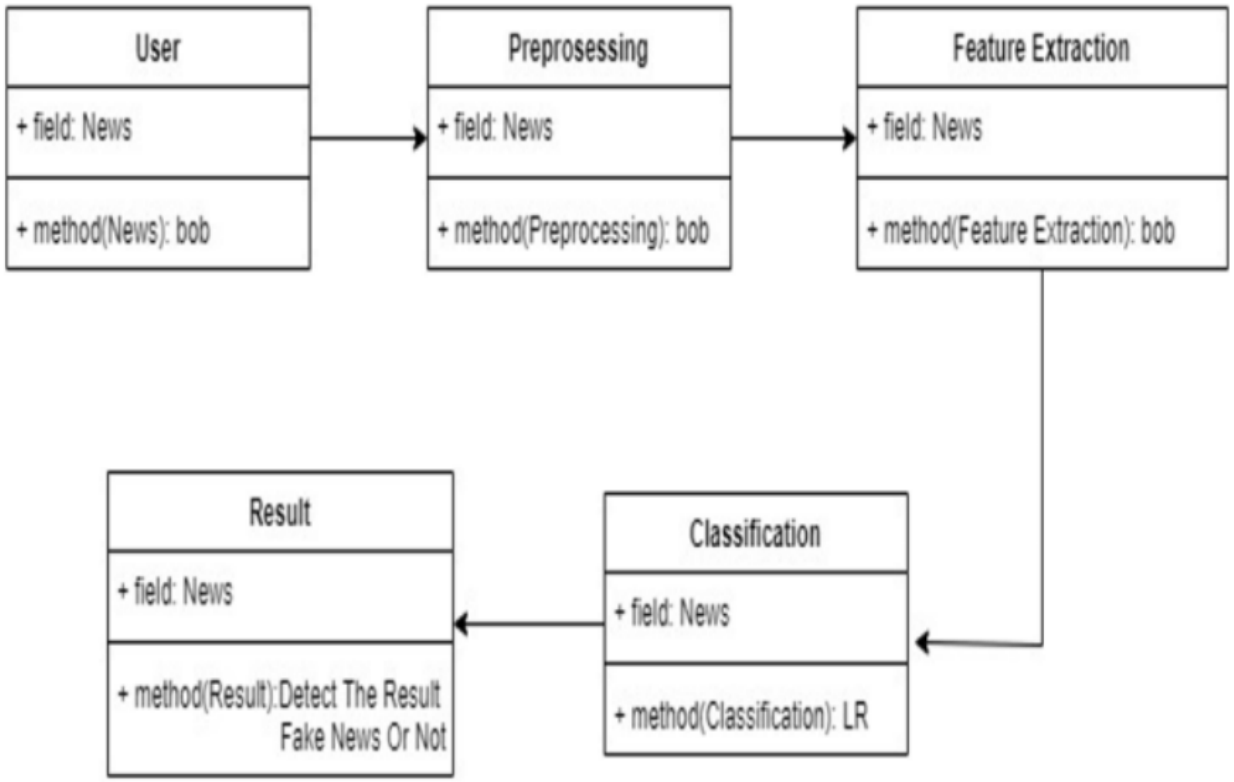
Use Case Diagram



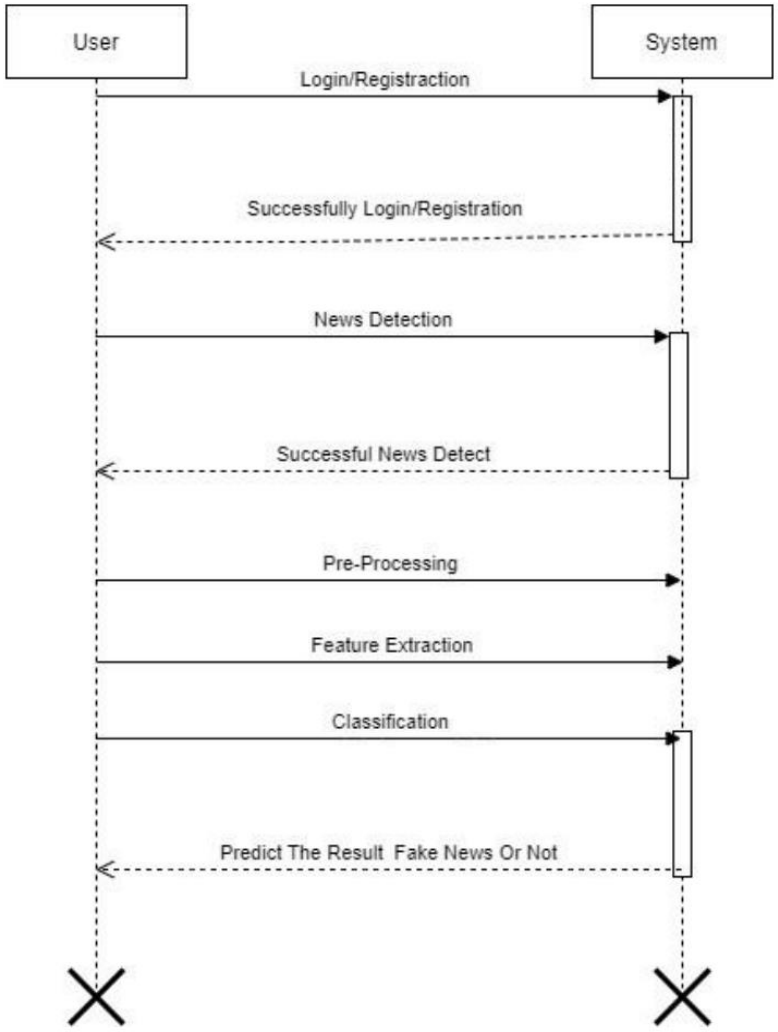
Data Flow Diagram (DFD)



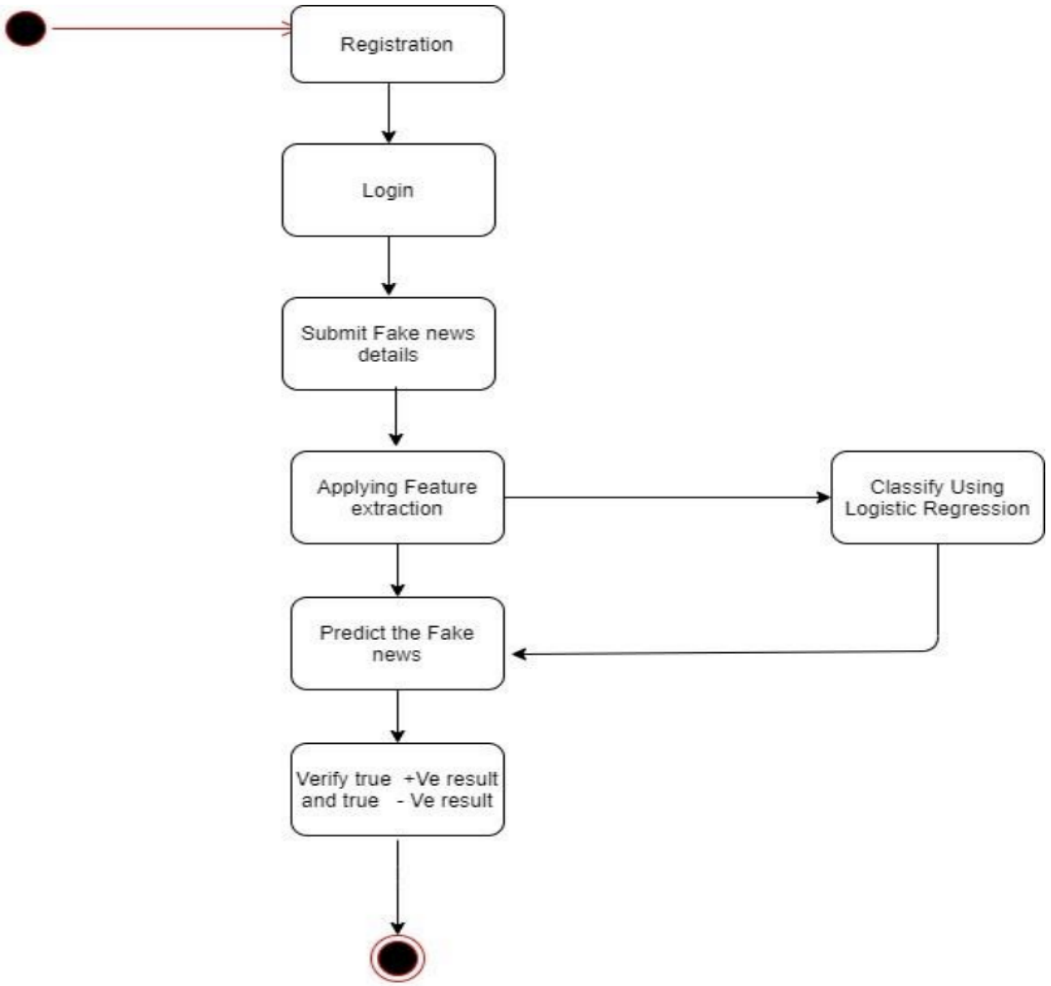
Class Diagram



Sequence Diagram



Activity Diagram



Modules

- Data Collection Module
- Preprocessing Module
- Feature Extraction Module
- Machine Learning Module
- Evaluation Module
- Interface Module

Algorithms / Methodology

- Step 1: Data Collection
- Step 2: Text Preprocessing (clean, tokenize, remove stop words)
- Step 3: Feature Extraction (TF-IDF, Word2Vec)
- Step 4: Model Training (SVM, Logistic Regression, LSTM)
- Step 5: Evaluation (Accuracy, Precision, Recall, F1-Score)

Testing

- **Unit Testing:** Module verification
- **Integration Testing:** Combined module analysis
- **System Testing:** Complete flow testing
- **Validation Testing:** Model accuracy and output validation

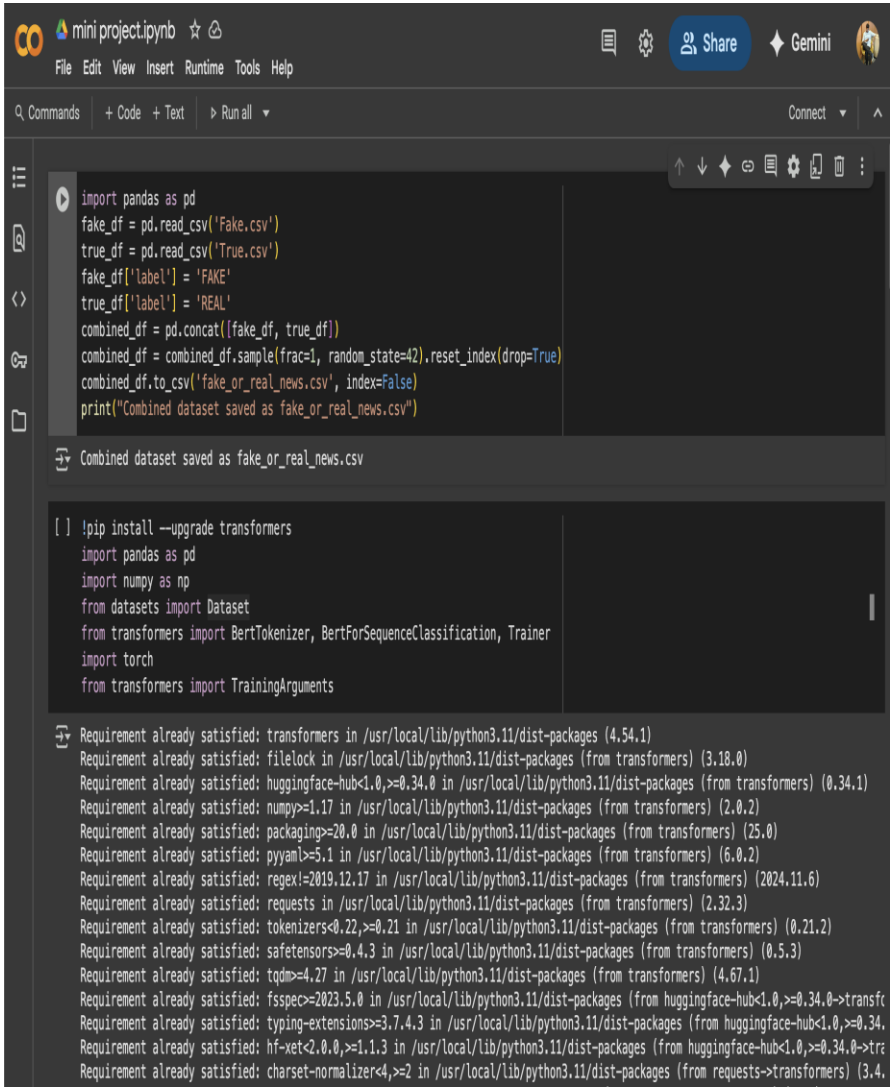
Test Cases / Validation

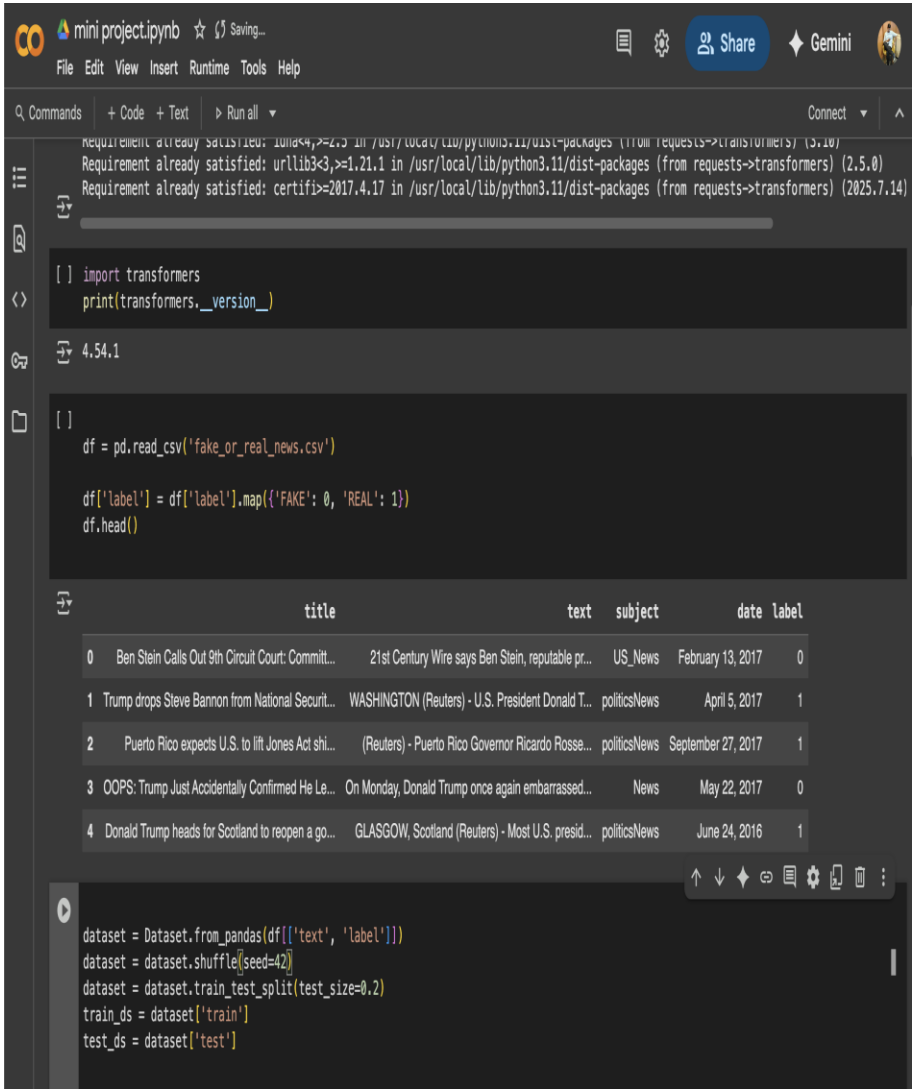
- **Input:** Sample news text
- **Expected Output:** Fake / Real
- **Actual Output:** Matches expected result
- **Metrics Used:** Accuracy, Precision, Recall

Performance Analysis

- Compared models: Logistic Regression, SVM, LSTM
- Evaluated metrics: Accuracy, Precision, Recall, F1-score
- LSTM achieved best accuracy
- Deep learning provides robust detection

Screenshots / Simulation





Conclusion

- Machine learning effectively detects fake news
- Automated detection ensures real-time verification
- Helps maintain trustworthy digital communication

Future Scope

- Real-time API-based fake news monitoring
- Multilingual dataset integration
- Transformer and hybrid model improvements
- Deployment as mobile/web application

References

- [1] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, “Fake news detection on social media: A data mining perspective,” ACM SIGKDD, 2017.
- [2] N. Ruchansky, S. Seo, and Y. Liu, “CSI: A hybrid deep model for fake news detection,” CIKM, 2017.
- [3] W. Y. Wang, “Liar, liar pants on fire: A new benchmark dataset for fake news detection,” ACL, 2017.

- [4] H. Ahmed, I. Traoré, and S. Saad, “Detecting opinion spams and fake news using text classification,” *Security and Privacy*, vol. 1, no. 1, 2018.
- [5] A. Singh et al., “A review on fake news detection using machine learning techniques,” in *Proc. ICCSP*, pp. 568–573, 2019.
- [6] J. Thorne, A. Vlachos, C. Christodoulopoulos, and A. Mittal, “FEVER: Fact extraction and verification,” in *Proc. NAACL-HLT*, pp. 809–819, 2018.
- [7] K. Shu, S. Wang, and H. Liu, “Beyond news contents: The role of social context for fake news detection,” in *Proc. WSDM*, pp. 312–320, 2019.

- [8] V. Pérez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea, “Automatic detection of fake news,” in *Proc. COLING*, pp. 3391–3401, 2018.
- [9] Y. Yang, L. Zheng, J. Zhang, Q. Cui, Z. Li, and P. S. Yu, “TI-CNN: Convolutional neural networks for fake news detection,” *arXiv preprint arXiv:1806.00749*, 2018.
- [10] X. Zhou and R. Zafarani, “Fake news: A survey of research, detection methods, and opportunities,” *arXiv preprint arXiv:1812.00315*, 2020.