Fake News Detection Using Machine Learning

Project By:

Kavish Jain 23UCS614

Raghav Malani 23UCC590

Submitted To: Dr. Nirmal Sivaraman

Problem & Motivation

- Fake news spreads very quickly on social media platforms.
- Short posts make misinformation easy to create and share.
- Manual verification is slow and not scalable.
- Fake news affects public trust, elections, and social stability.
- Automated detection is needed for fast and reliable filtering.
- Machine Learning helps identify hidden patterns and detect misleading content effectively.

Project Objectives

- Implement the fake news detection methodology from the research paper on news articles.
- Apply the same machine learning models and features to one-line social media posts.
- Enhance the feature set to handle short, noisy, informal text.
- Compare performance across all three stages to understand domain effects.

Datasets Used

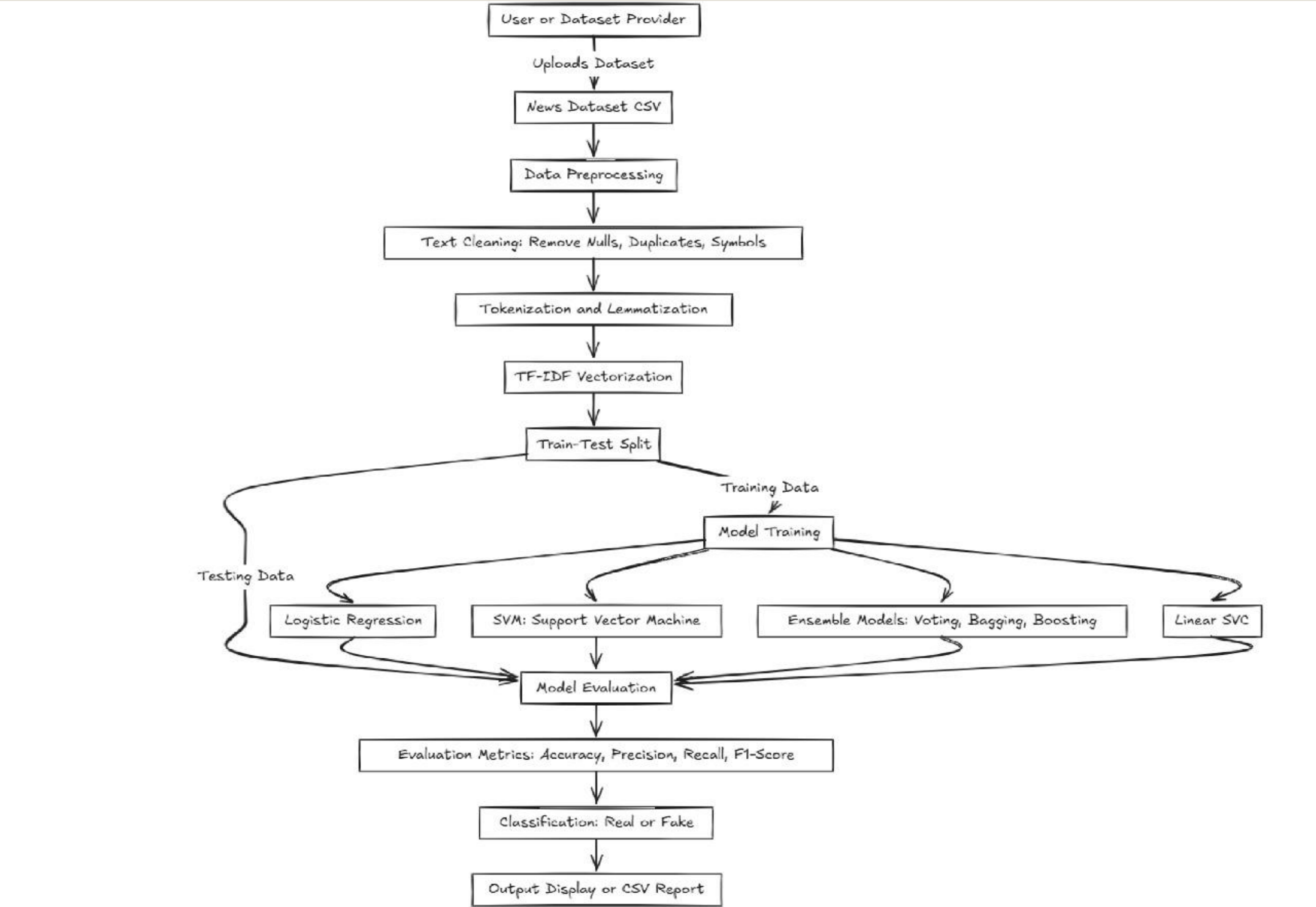
ISOT News Dataset (True.csv, Fake.csv)

- Long, structured news articles
- Clear writing patterns and proper grammar
- Good for baseline ML performance

Manually Collected Social Media Posts

- Very short one-liner posts
- Informal, noisy, low context
- Harder for TF–IDF and classical ML models
- Collected from multiple online sources (Instagram, satire pages, news handles)

Overall Workflow



Stage 1: Research Paper Implementation

- Implemented original methodology from Research Paper 1
- **Features used:**
  – TF–IDF (unigrams + bigrams)
  – Token count
  – Average word length
  – Punctuation ratio
- **Models applied:**
  – Logistic Regression
  – Support Vector Machine (Linear SVM)
  – Random Forest
  – AdaBoost
  – Gradient Boosting
  – Voting Classifier

Stage 1 Results:

- Achieved **very high accuracy (up to 99%+)** on news articles
- **Ensemble models** (Bagging, AdaBoost, Gradient Boosting) performed the best
- TF–IDF works extremely well because news articles are:
  – Long
  – Structured
  – Rich in vocabulary
- Models easily capture patterns due to clear sentence structure and context

Stage 2: Same Features on Social Media Posts

- Applied **Stage 1 features** (TF–IDF + basic linguistic features) on **one-line social media posts**

Problems Identified

- Posts are **extremely short** → almost no context
- **TF–IDF becomes sparse** and weak
- Writing style is **informal, noisy, inconsistent**
- Same feature set **does not generalize** from long articles

Performance

- Model accuracy dropped to **0.67 – 0.83**
- Shows that article-based features are **not suitable** for short posts

Stage 3: Enhanced Feature Engineering

- Introduced **advanced linguistic + stylistic features** to handle short, noisy posts
- Designed to capture **writing style**, **emotion**, and **behaviour patterns**
- Helps overcome lack of context in one-line posts

Enhanced Features Used

- **Uppercase character ratio**
- **Digit ratio** (numbers used in fake posts)
- **Punctuation patterns** (e.g., !, ??, …)
- **Unique word ratio**
- **Lexical diversity**
- **Sentiment Polarity Score**

  – Helps detect emotional exaggeration often present in fake posts

Stage 3 Results:

- Applied **enhanced linguistic, stylistic, and sentiment features**
- Major improvement compared to Stage 2

Key Outcomes

- **Accuracy increased up to 91%**
- **Large boost** across all models (LogReg, SVM, RF, Boosting)
- Models captured **writing style + sentiment** better than TF–IDF alone
- Significant improvement in **precision, recall, and F1-score**

Comparative Study (Stage 2 vs Stage 3)

Overall Comparison

- **Stage 2 Accuracy:** 0.67 – 0.83
- **Stage 3 Accuracy:** 0.85 – 0.91
- **Clear improvement** after adding enhanced linguistic, stylistic & sentiment features.

Key Improvements Across Metrics

- **Precision:** Increased significantly (models make fewer wrong predictions)
- **Recall:** Higher ability to detect fake news correctly
- **F1-Score:** Balanced improvement in both precision & recall
- Shows that **Stage 3 features are more effective for short social media posts**

Key Insights:

- **TF–IDF works extremely well for long news articles**, but performs poorly on short one-line posts due to lack of context.
- **Enhanced linguistic + stylistic + sentiment features** help compensate for missing context in social media text.
- **Sentiment polarity and writing style** (uppercase, punctuation, digit usage) provide strong signals for fake news detection in short posts.
- **Domain shift matters:**
  Models trained on long articles **do not generalize** to social media without redesigning features.
- **Feature engineering is more important than the model itself** when dealing with noisy short text.

Conclusion

- **Stage 1:** Models perform exceptionally well on long, structured news articles → accuracy ~99%.
- **Stage 2:** Accuracy drops sharply on one-line social media posts due to lack of context and sparse TF–IDF features.
- **Stage 3:** Performance improves significantly after adding enhanced linguistic, stylistic, and sentiment features.
- **Key Takeaway:**
  Feature engineering is the most important factor when adapting fake-news detection from formal articles to short, noisy social media text.

Thank you