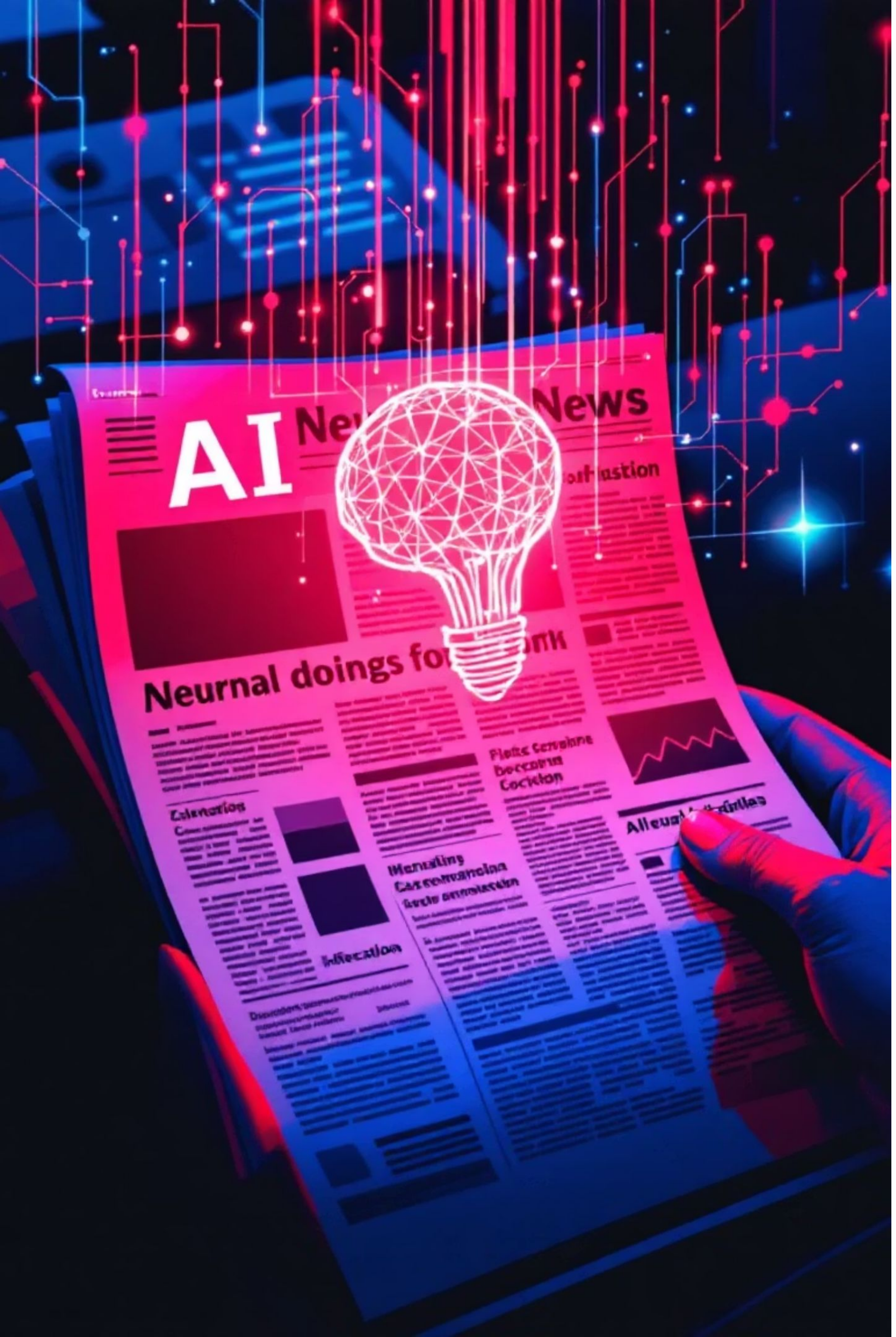


Fake vs Real News Detection

Using Natural Language Processing

Girish Kumar
20 November 2025

Can Machines Detect Clickbait & Fake News from Titles Alone?

34,152 Labeled Titles

Training dataset with fake and real news classifications

9,984 Test Titles

Unlabeled data to predict with highest accuracy

The Challenge

Titles are short, sensational, and full of deceptive tricks

Data Snapshot

Training Data

- **34,152 labeled news titles**

- Fake titles use "wow!", "shocking", "boom!", all caps

- Real titles are neutral and factual

  "germany's fdp look to fill schaeuble's big shoes"

Test Data

- 9,984 unlabeled titles
- Unknown labels (marked as 2) need prediction

Data Preprocessing & Model Selection

Preprocessing Techniques

- Lowercasing
- Punctuation
- Whitespace
- Stopword removal
- TF-IDF/BOW/Word2Vec Vectorization

ML Classifiers

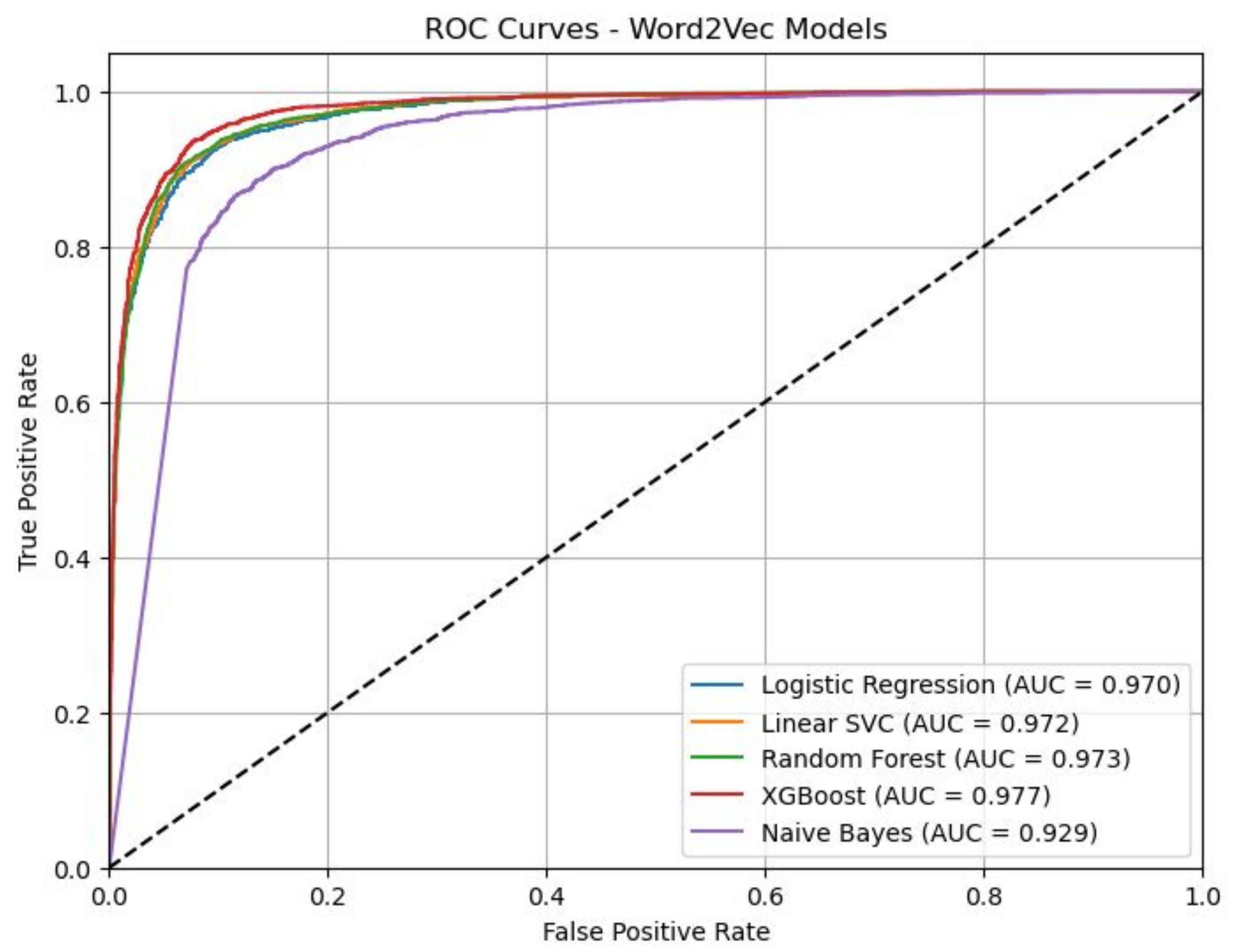
- LogisticRegression
- MultinomialNB
- LinearSVC
- RandomForest
- XGBoost

Word2Vec – Learning Word Meaning from Context

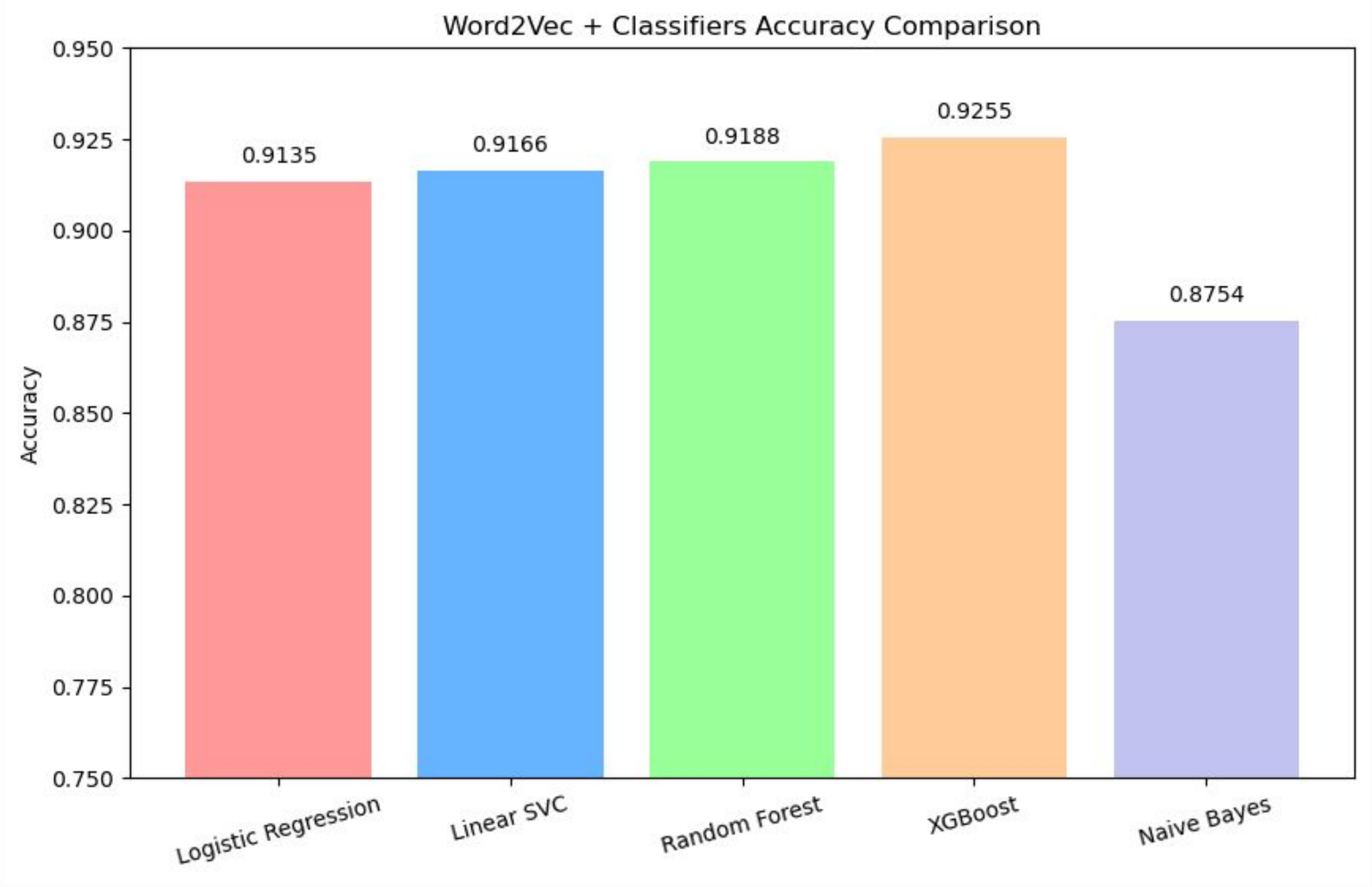
- Learns 300-dim, similar words = vector

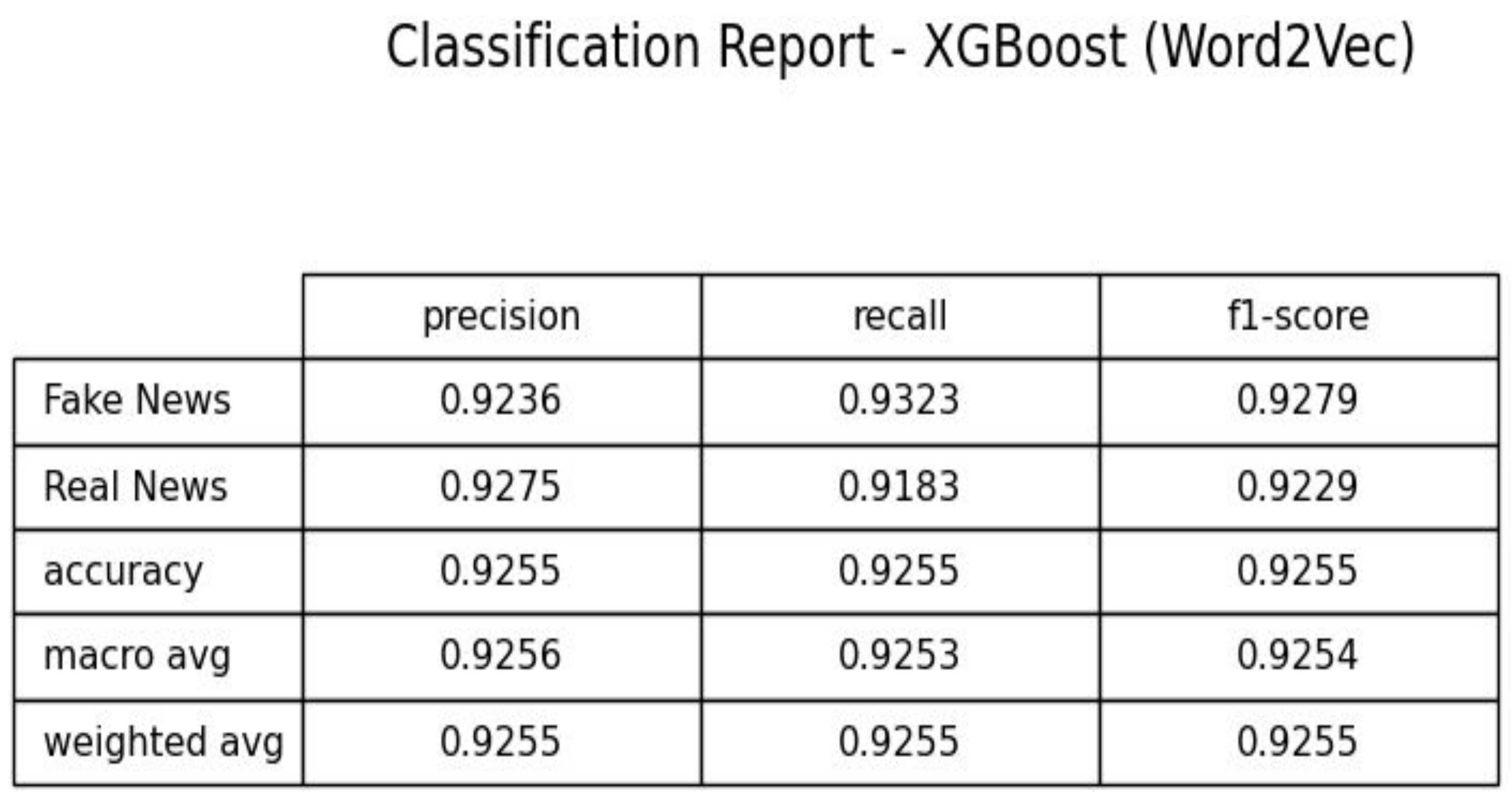
- Trained on 34k titles to capture word relationship

- Tested across multiple models



Word2Vec – Learning Word Meaning from Context



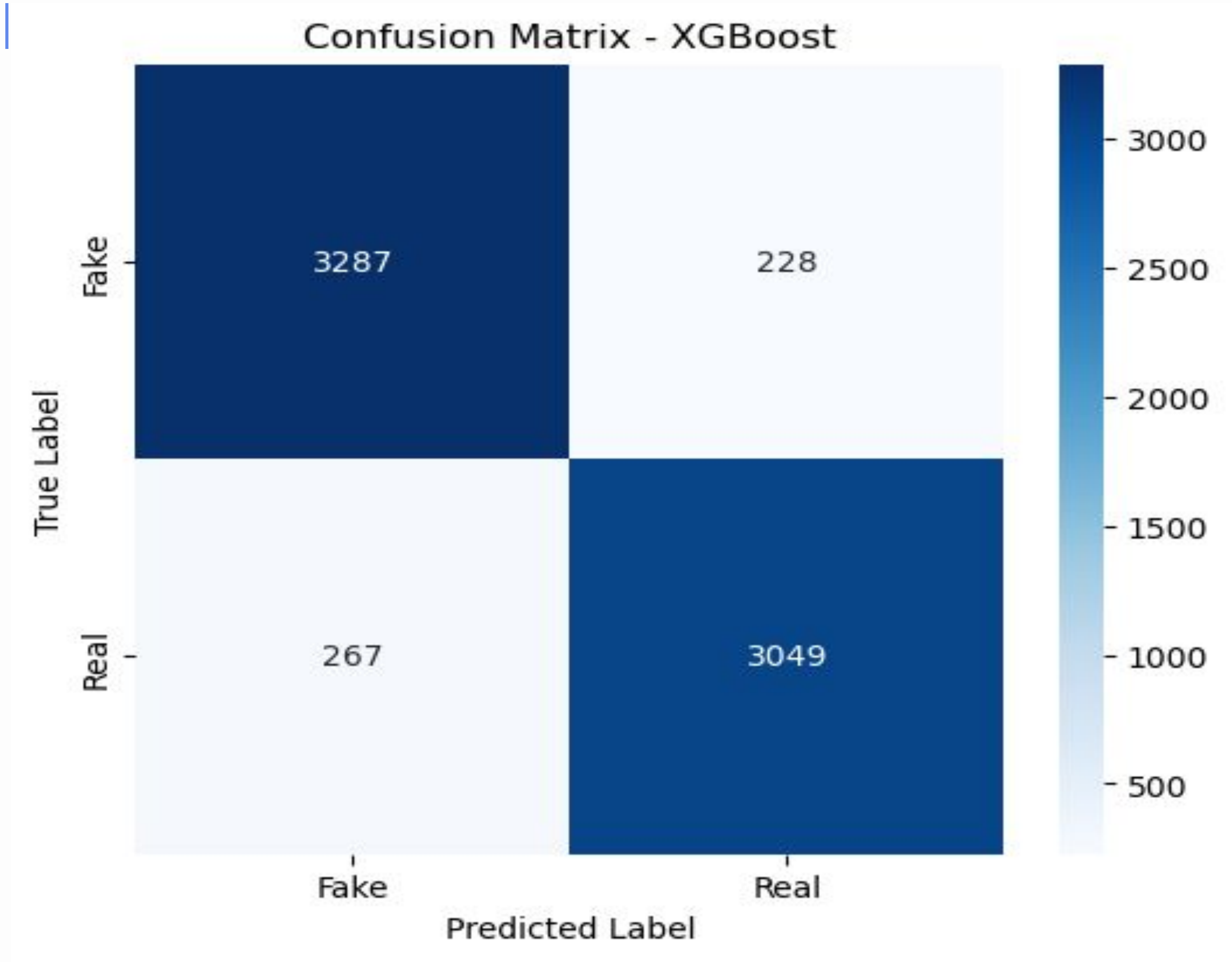


Classification Report - XGBoost (Word2Vec)

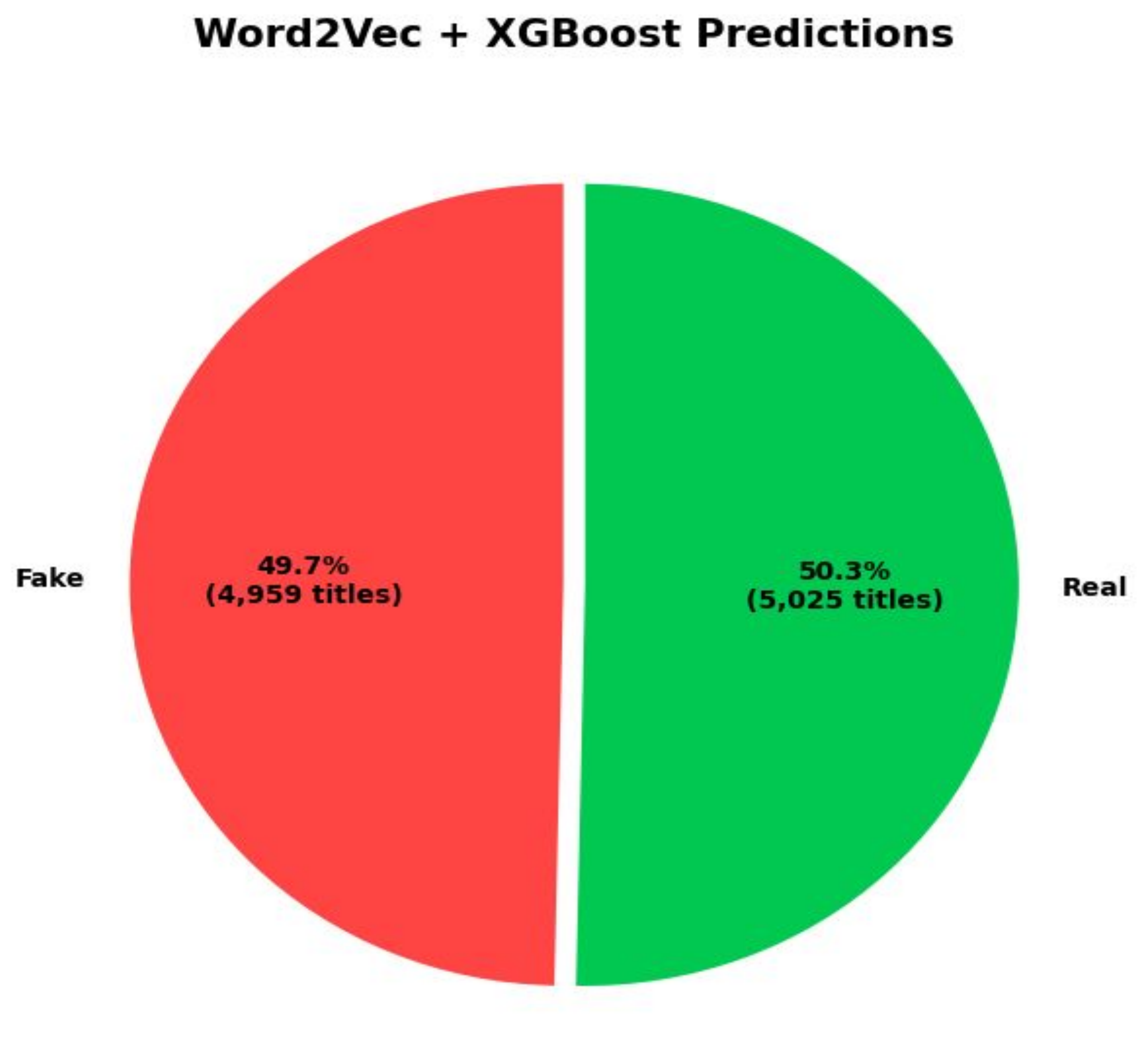
	precision	recall	f1-score
Fake News	0.9236	0.9323	0.9279
Real News	0.9275	0.9183	0.9229
accuracy	0.9255	0.9255	0.9255
macro avg	0.9256	0.9253	0.9254
weighted avg	0.9255	0.9255	0.9255

Word2Vec - Validation & Testing

Validation on split data



Testing on Unknown data



Bag of Words (BOW)

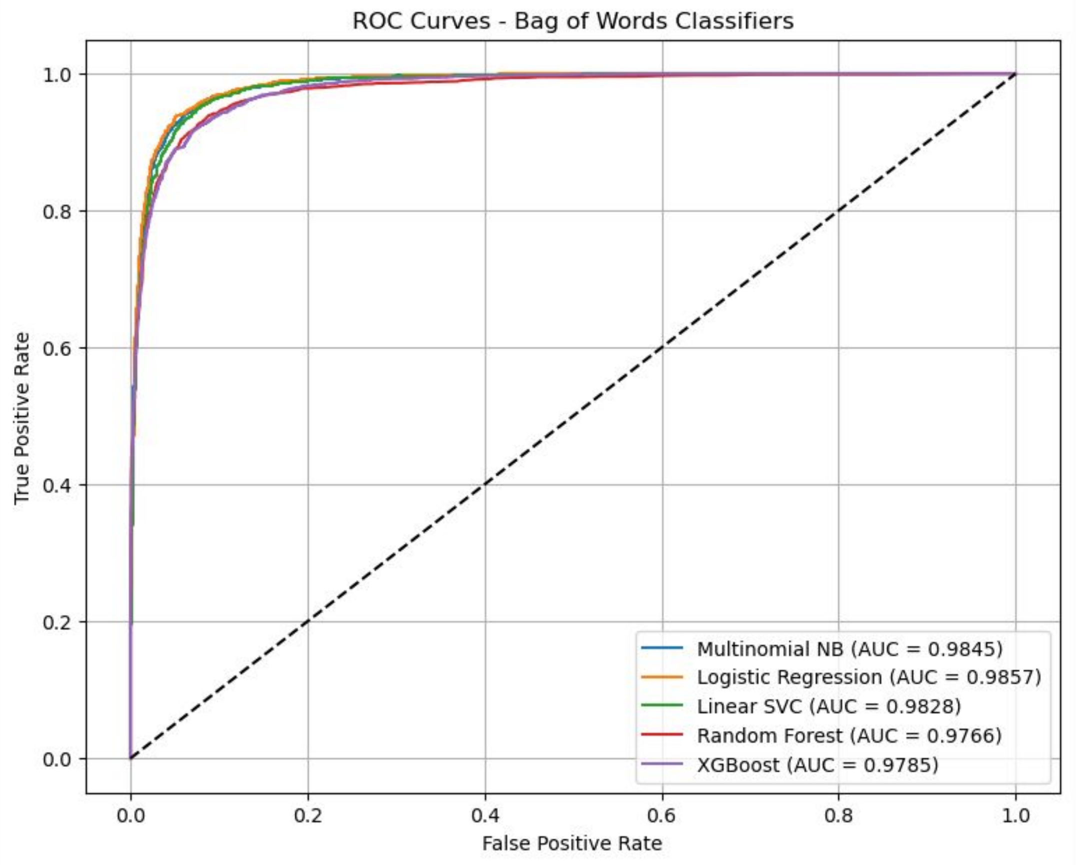
- Counts occurrences of words & n-grams parameters hypertuning

- No semantics — just exact matches

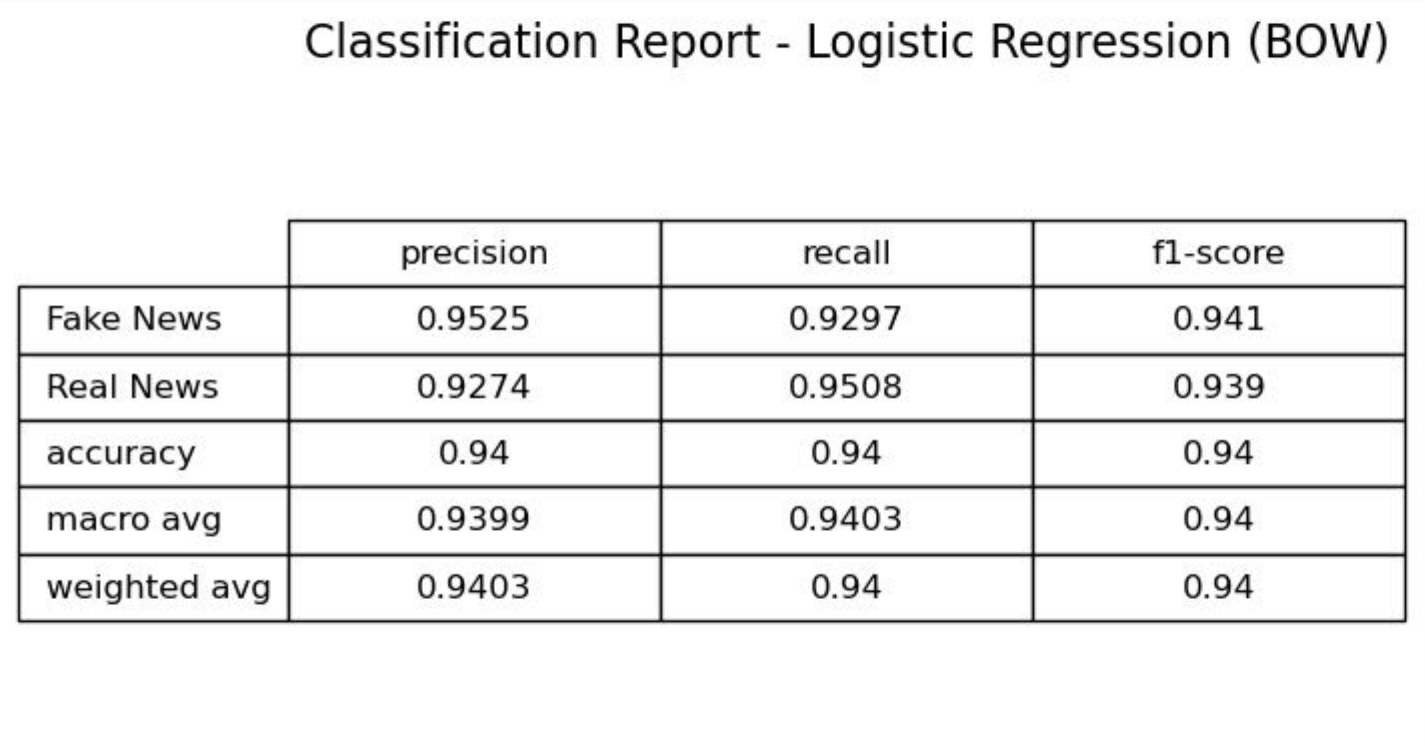
- Captures full phrases perfectly (“you won’t believe”, “breaking news”)

- Preserves sensational wording exactly

- Trained and tested across multiple models

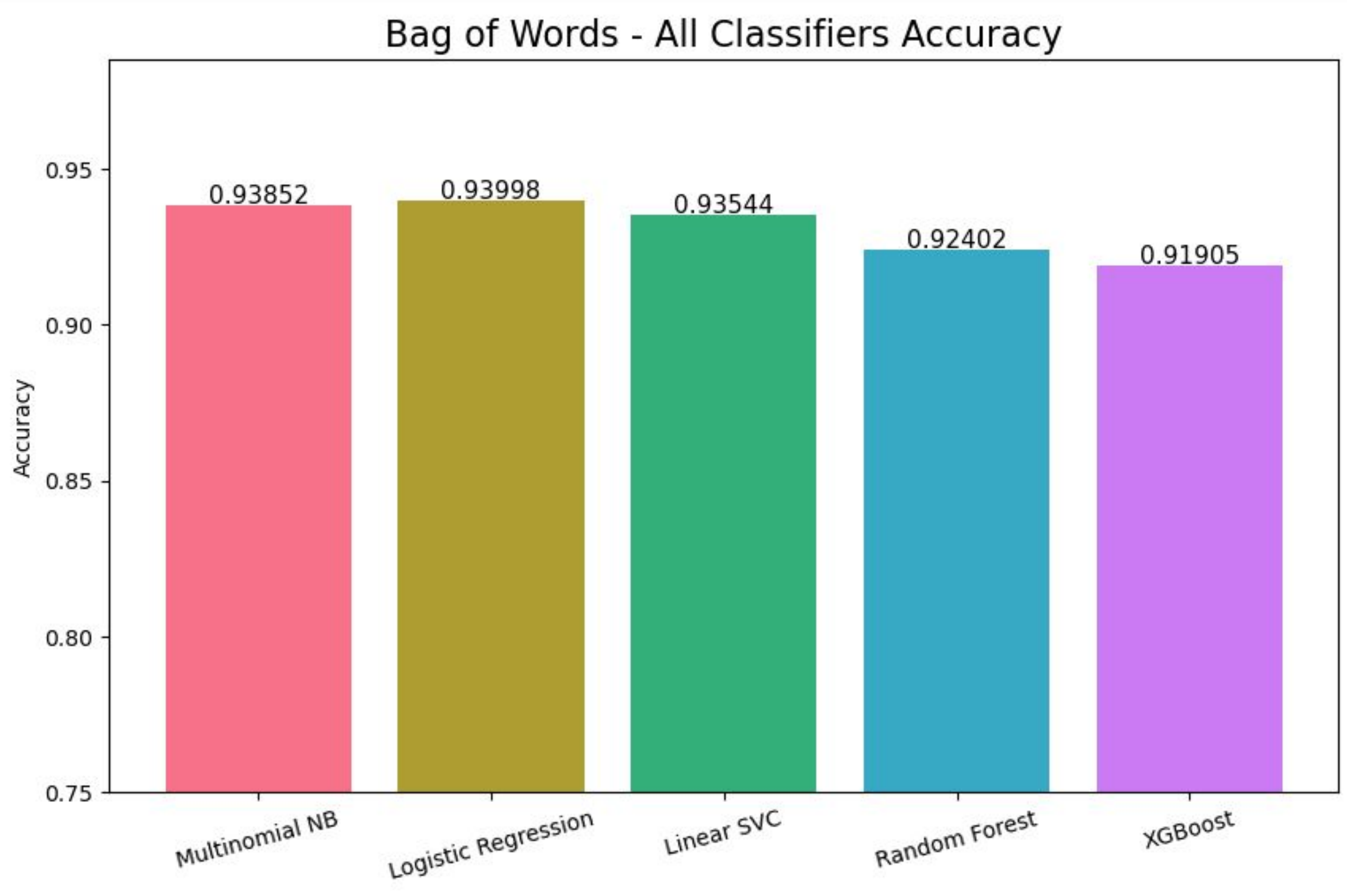


Bag of Words (BOW)



Classification Report - Logistic Regression (BOW)

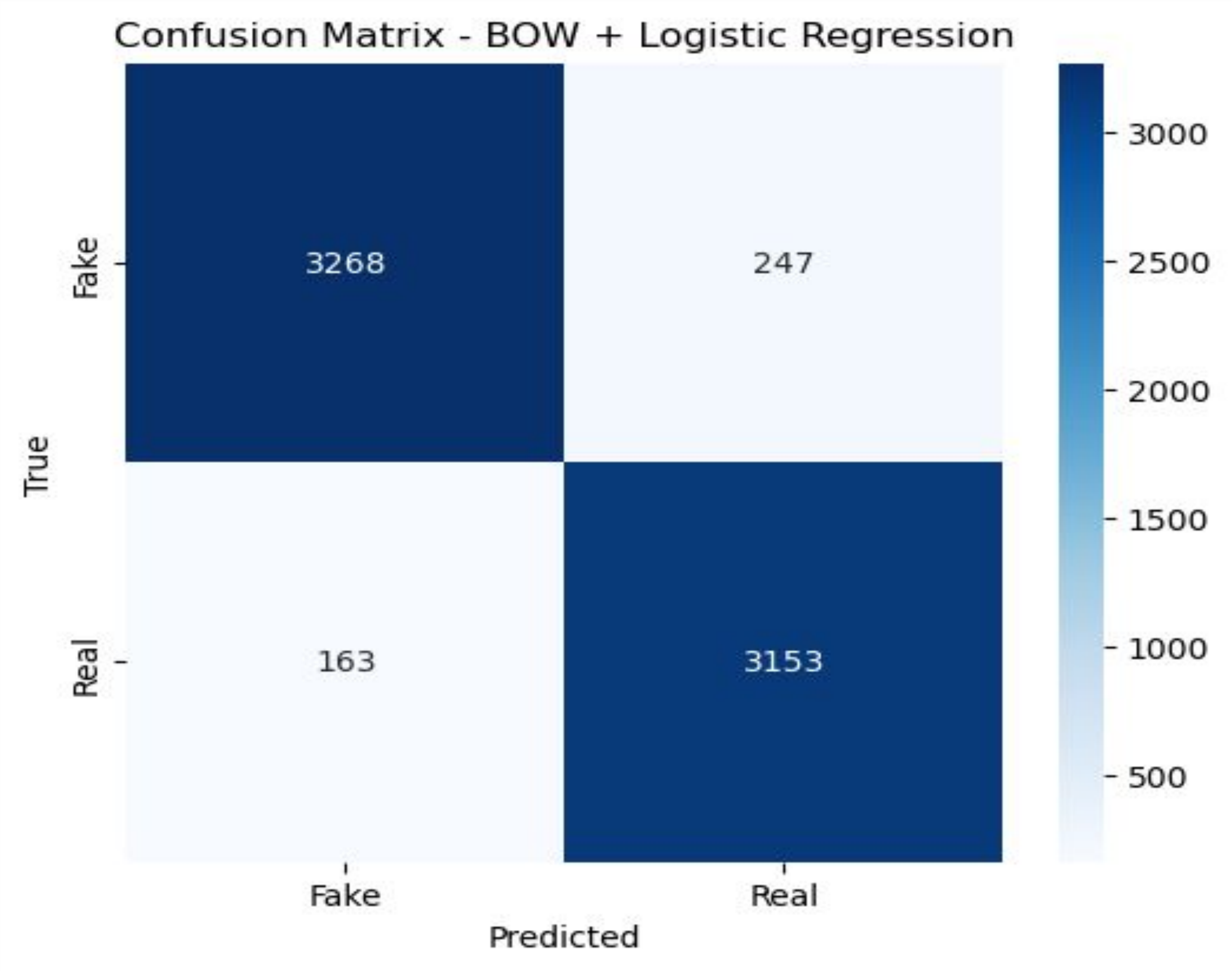
	precision	recall	f1-score
Fake News	0.9525	0.9297	0.941
Real News	0.9274	0.9508	0.939
accuracy	0.94	0.94	0.94
macro avg	0.9399	0.9403	0.94
weighted avg	0.9403	0.94	0.94



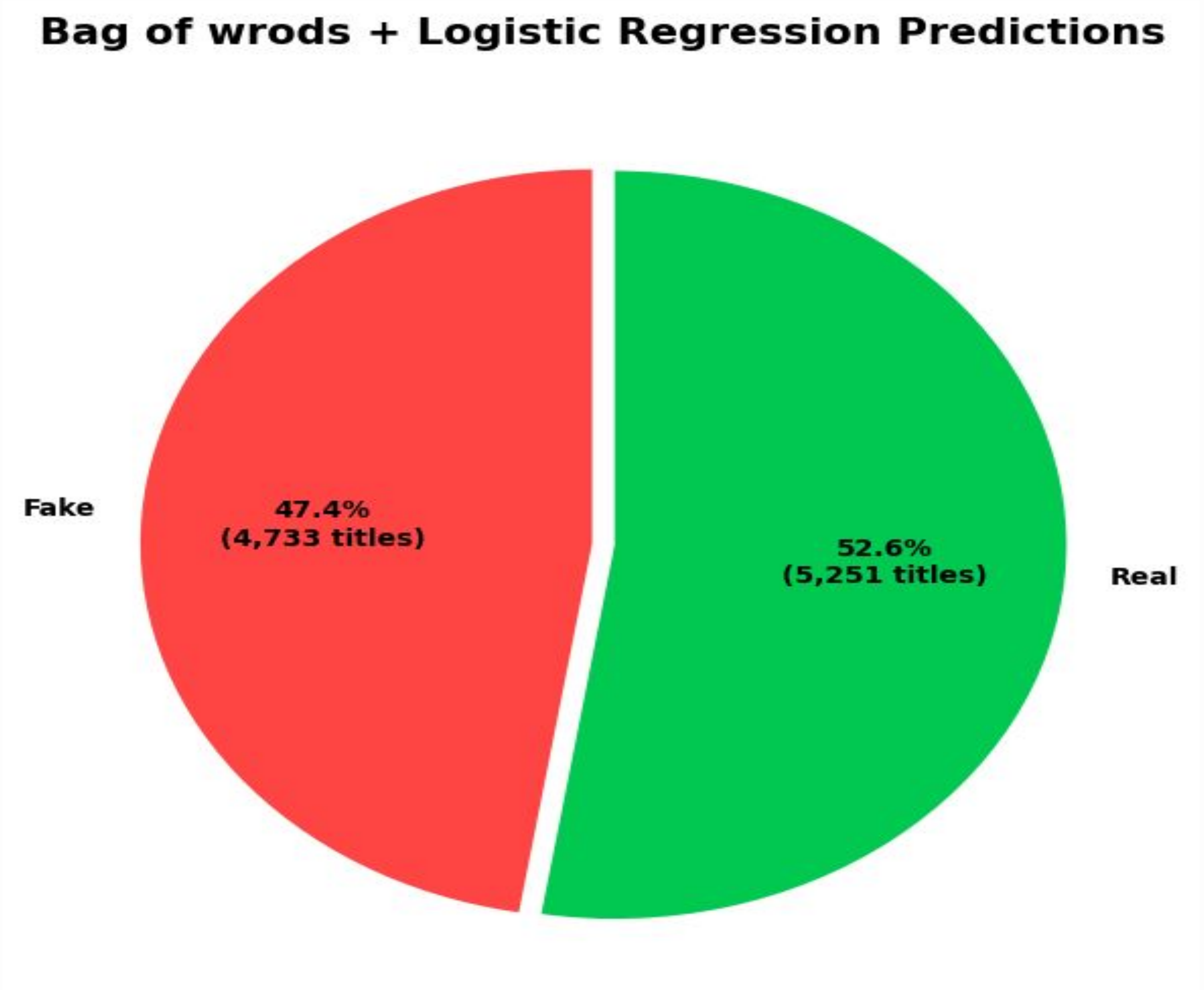
Bag of Words - All Classifiers Accuracy

Bag of Words (BOW) - Validation & Testing

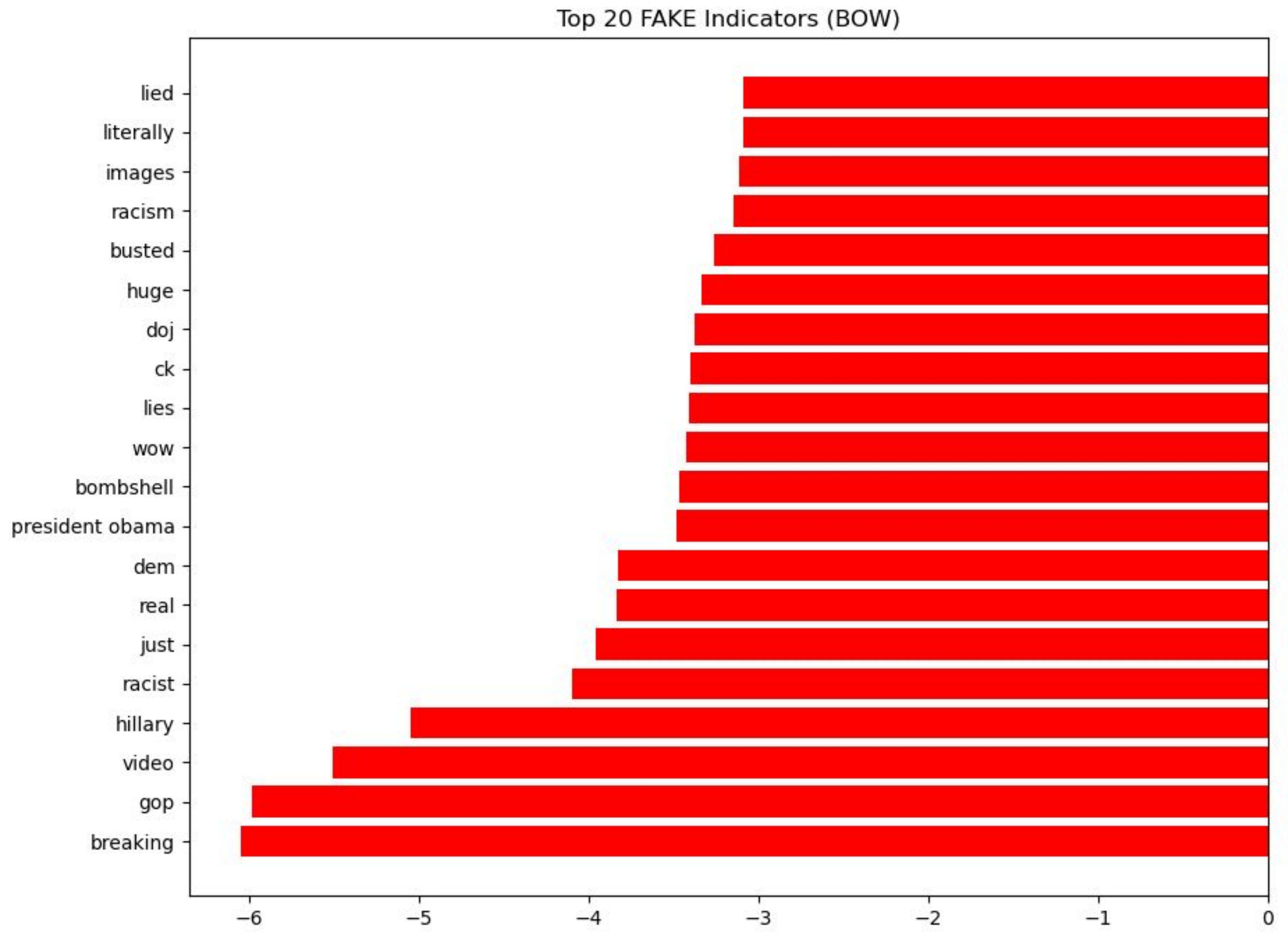
Validation on split data



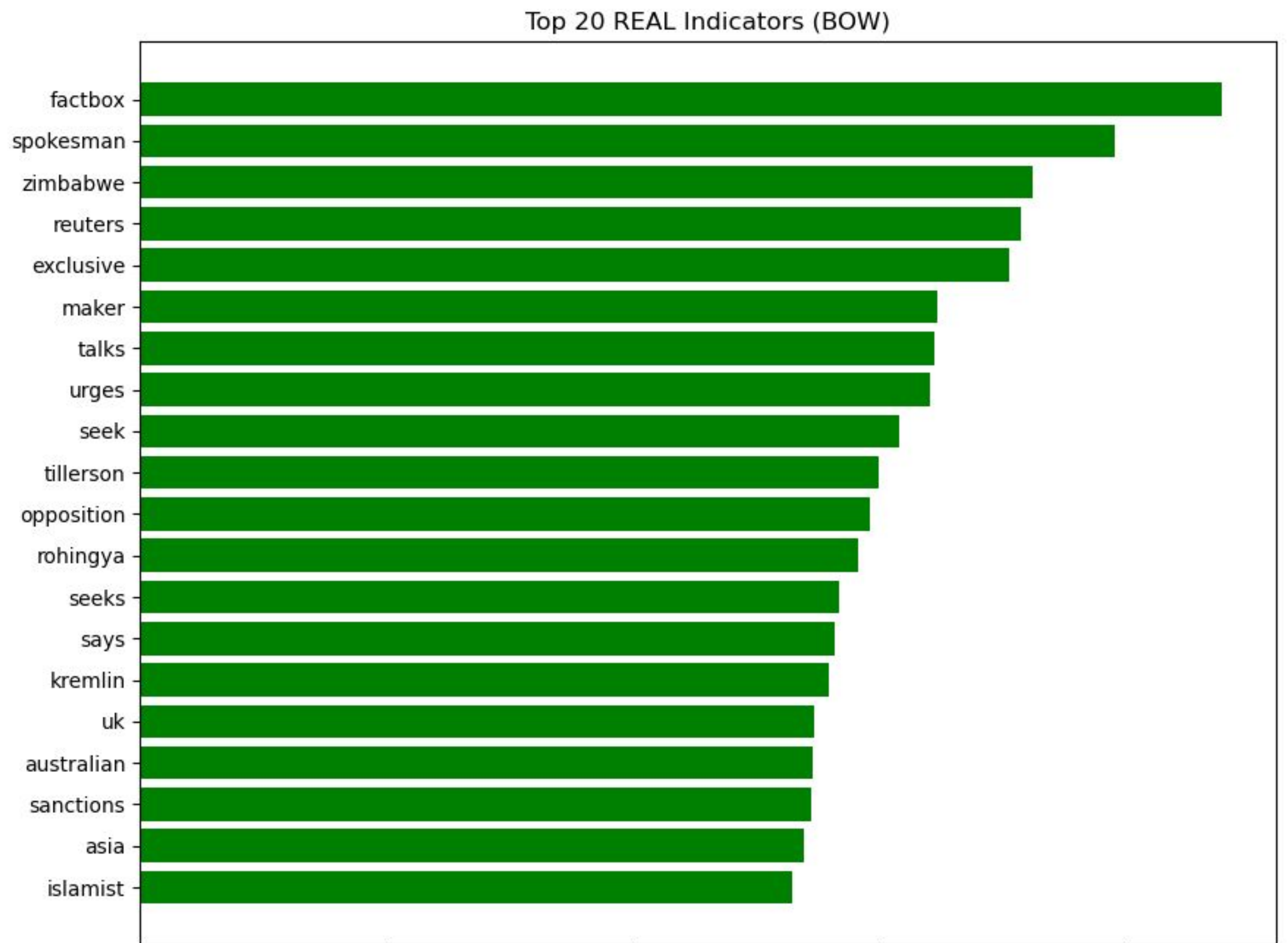
Testing on Unknown data



Bag of Words (BOW)



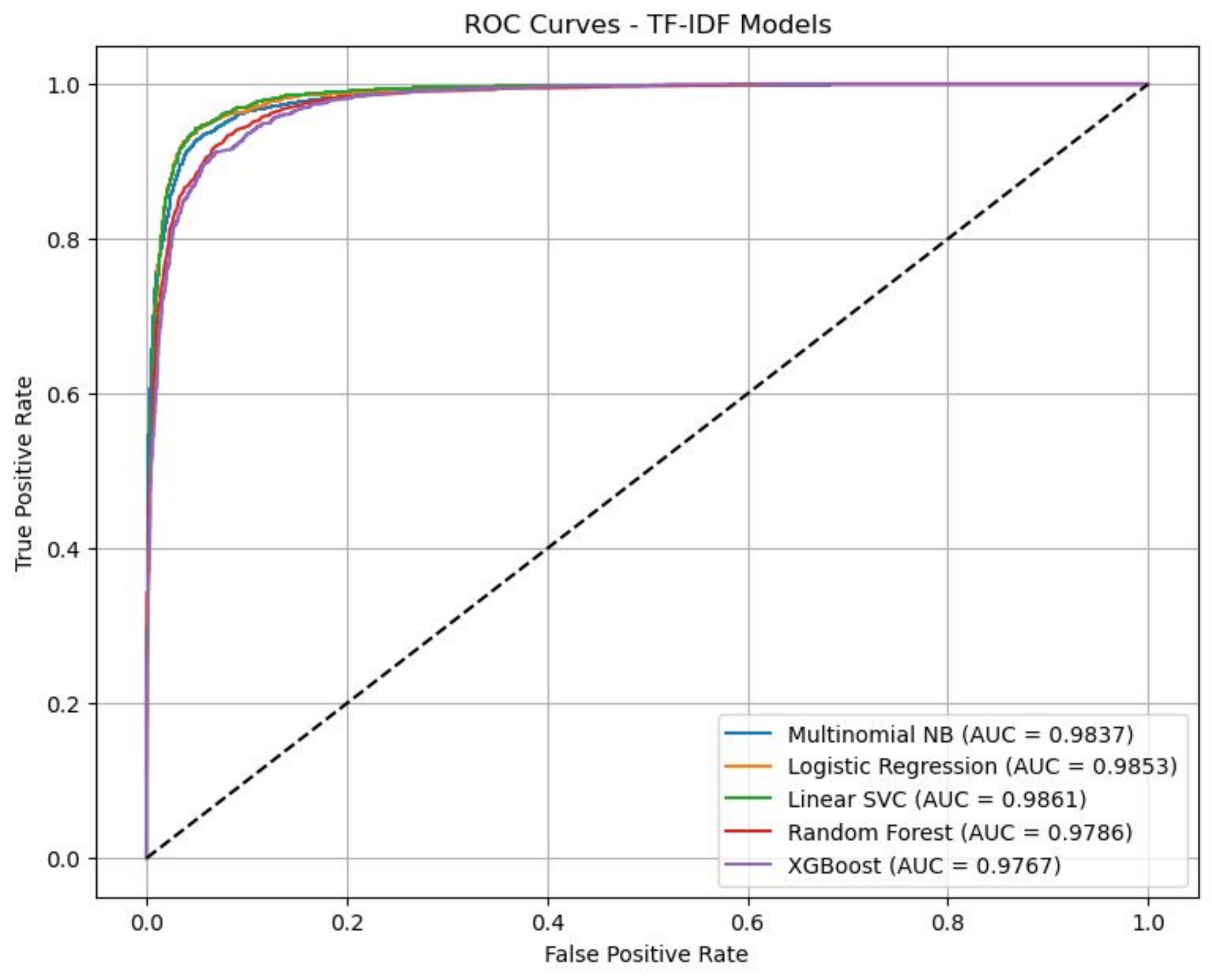
Top 20 FAKE Indicators (BOW)



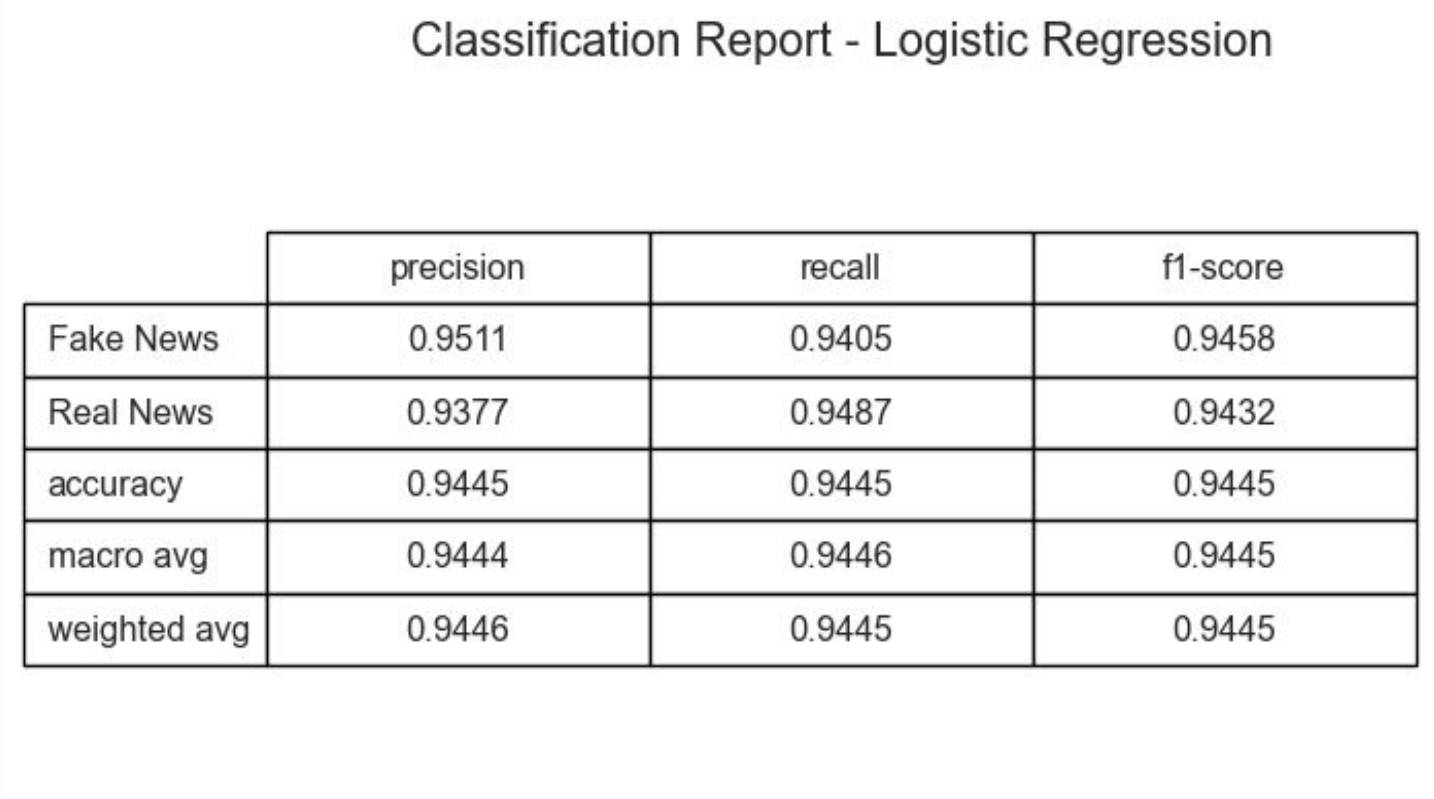
Top 20 REAL Indicators (BOW)

TF-IDF

- Counts n-grams (1-3) like BOW
- Weights rare/discriminative terms higher (“boom!”, “shocking” → high score in fake) (“says”, “minister” → high score in real)
- sublinear_tf + stop_words = clean signal

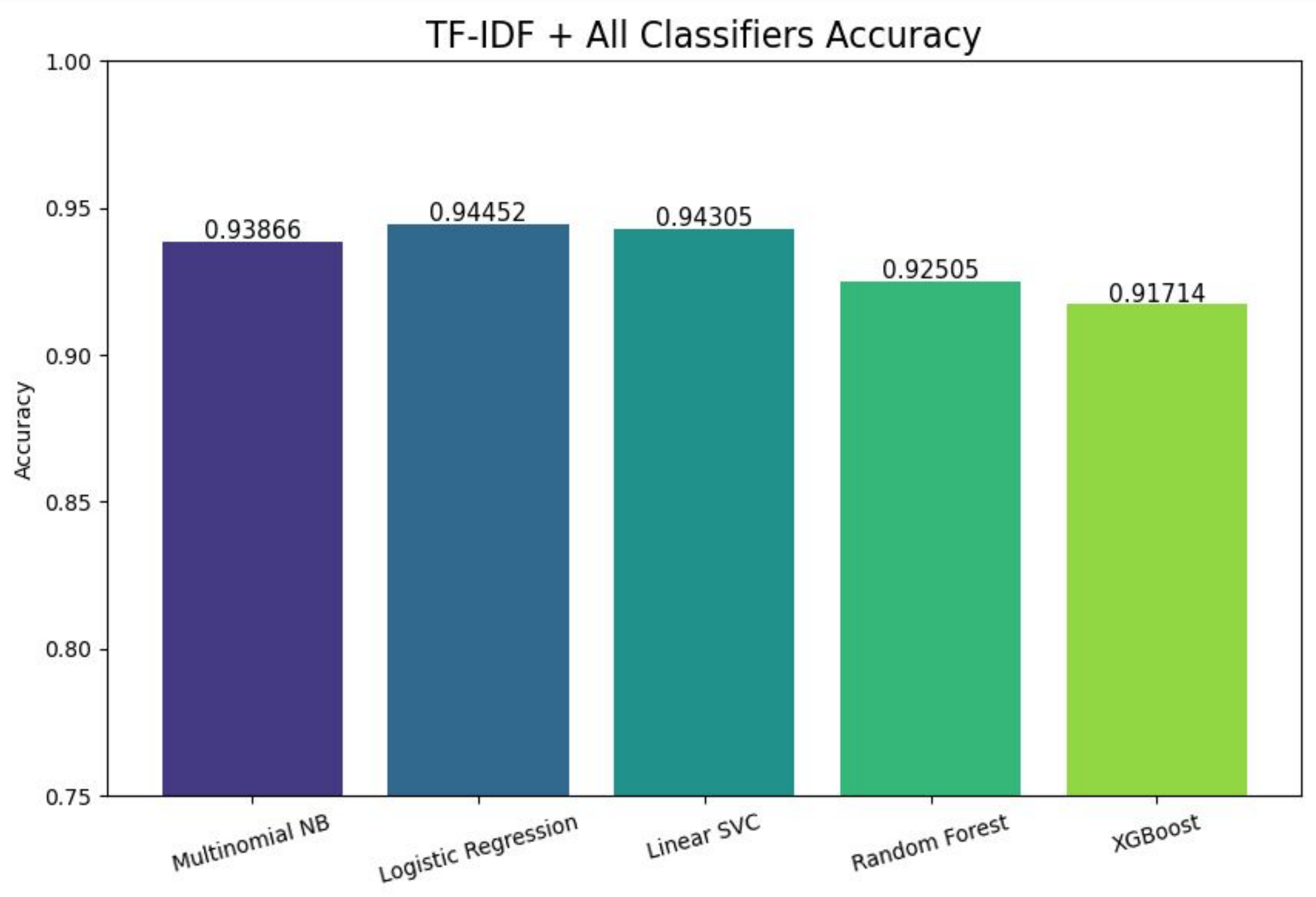


TF-IDF + All classifiers



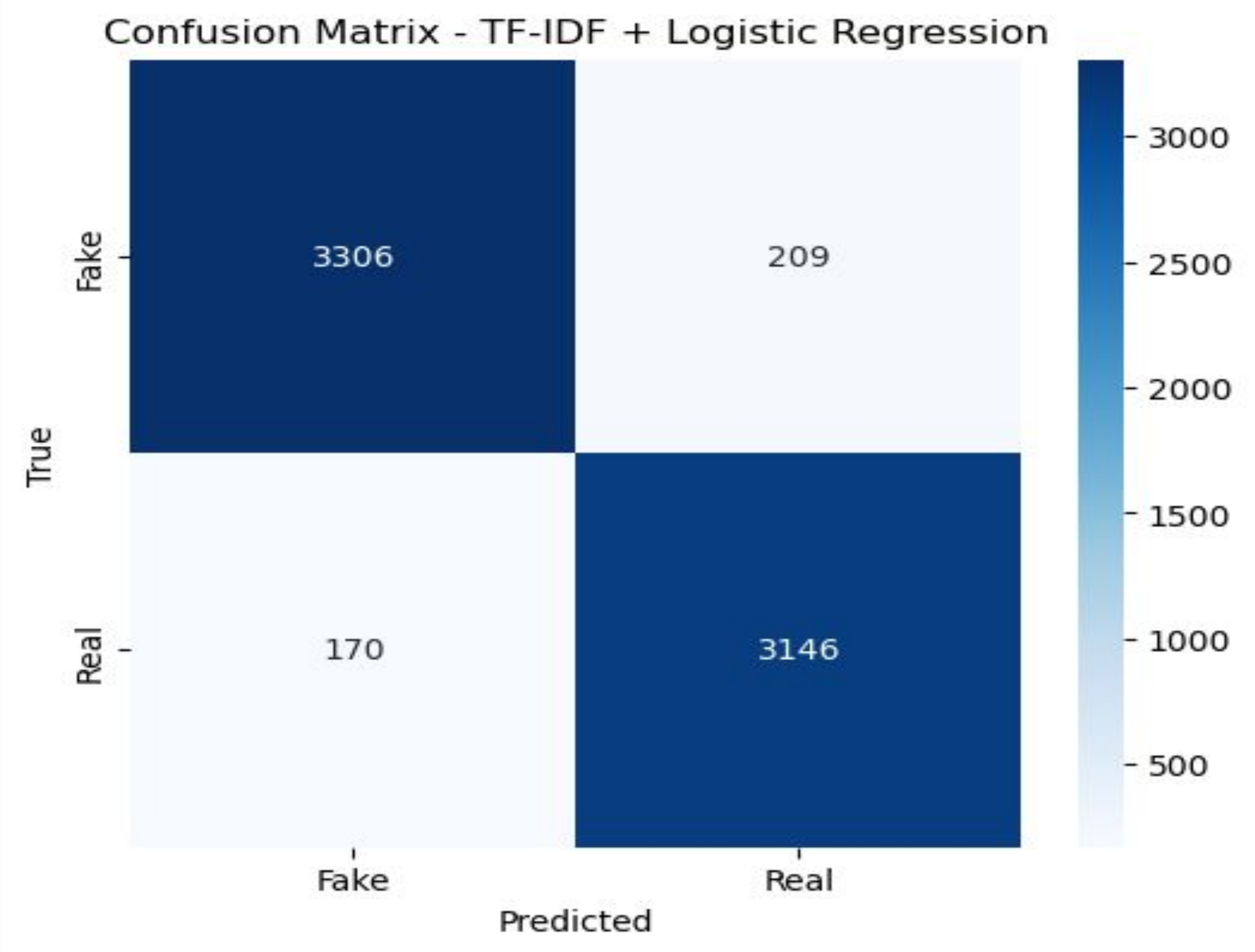
Classification Report - Logistic Regression

	precision	recall	f1-score
Fake News	0.9511	0.9405	0.9458
Real News	0.9377	0.9487	0.9432
accuracy	0.9445	0.9445	0.9445
macro avg	0.9444	0.9446	0.9445
weighted avg	0.9446	0.9445	0.9445

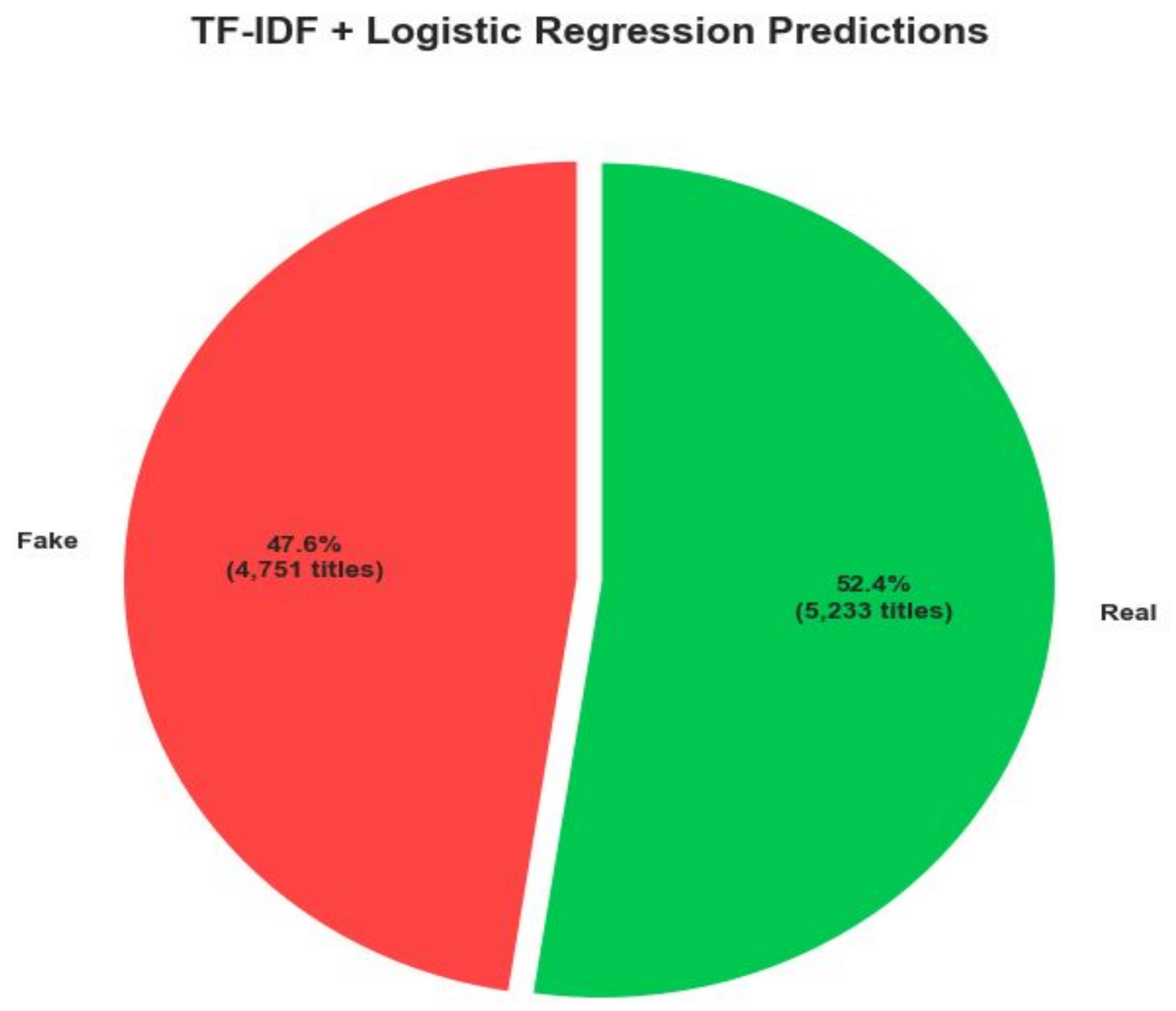


TF-IDF - Validation & Testing

Validation on split data



Testing on Unknown data



BERT: Deep Learning Model

Initial Run – No Hyperparameter Tuning

- Model: Fake-News-Bert-Detect

  (Pre-trained BERT specialized for fake news)
- Default settings → Poor generalization

Performance (Before Tuning)			
Precision	Recall	F1-Score	Accuracy
93.86%	28.55%	44.25%	63.86%

- Observation:

  Extremely high precision but catastrophically low recall

  → Model predicted nearly everything as “Real News”

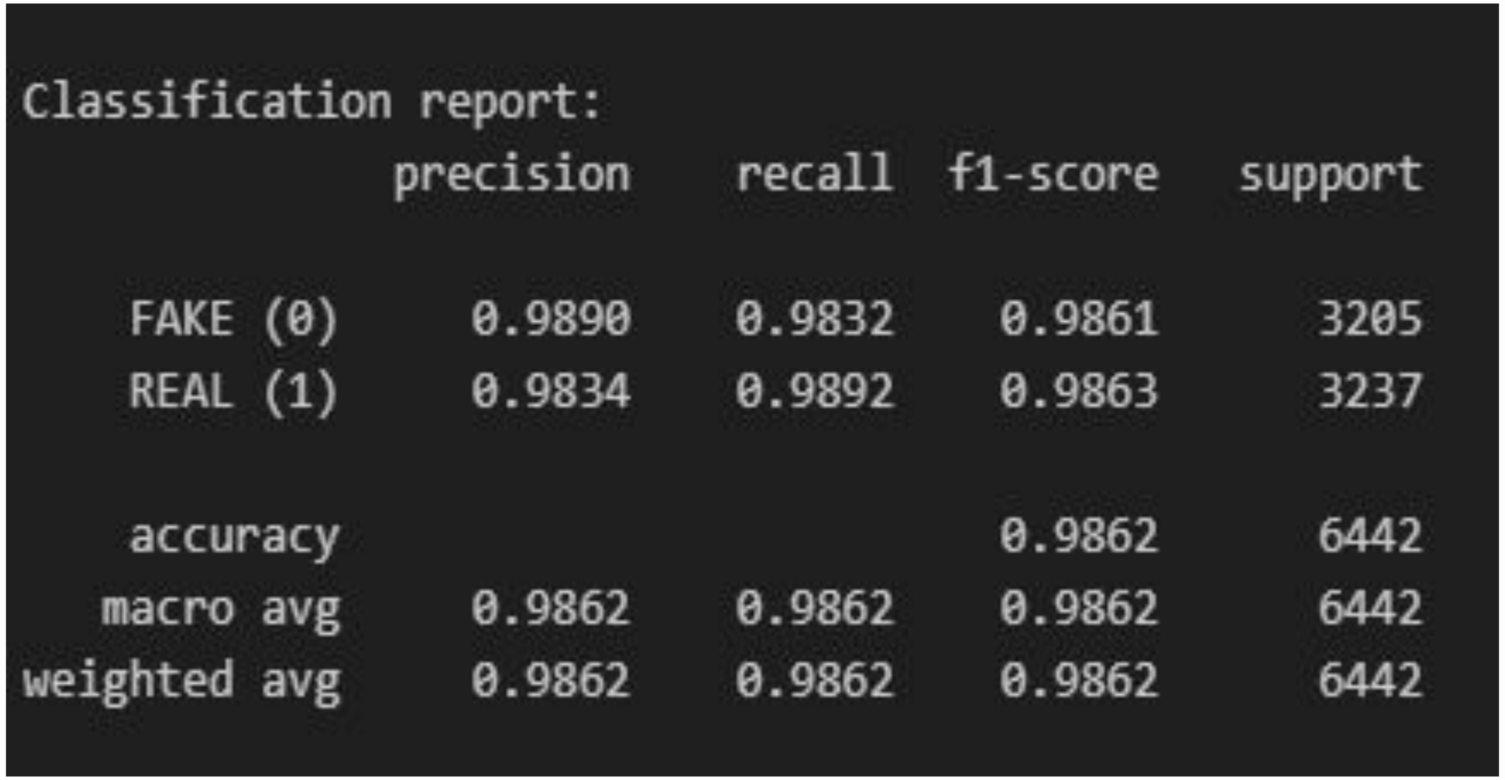
  → Missed ~71% of all fake/clickbait headlines!

  Conclusion: Default hyperparameters are insufficient

BERT: Deep Learning Model

After Proper Fine-Tuning → State-of-the-Art Performance

- Best Hyperparameters
  - Learning Rate : 2e-5
  - Batch Size : 16
  - Epochs : 3
  - Max Length : 512 tokens
  - Hardware : RTX 3060 GPU
  - Total Training Time: ~55 minutes



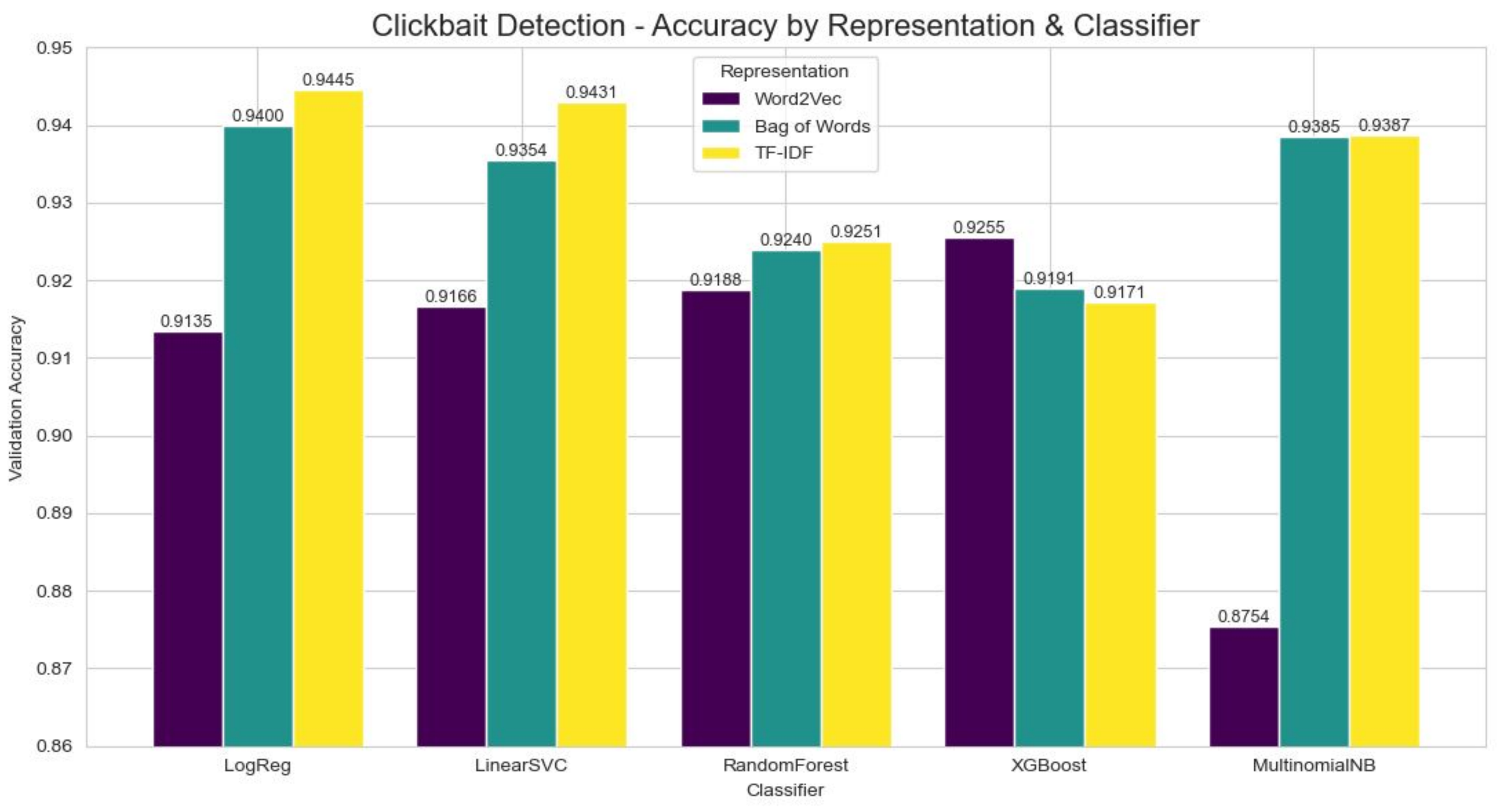
```
Classification report:
              precision    recall  f1-score   support

    FAKE (0)     0.9890    0.9832    0.9861      3205
    REAL (1)     0.9834    0.9892    0.9863      3237

    accuracy                         0.9862      6442
   macro avg     0.9862    0.9862    0.9862      6442
weighted avg     0.9862    0.9862    0.9862      6442
```

Accuracy Comparison: Classical ML vs Fine-Tuned BERT

Key Takeaways:

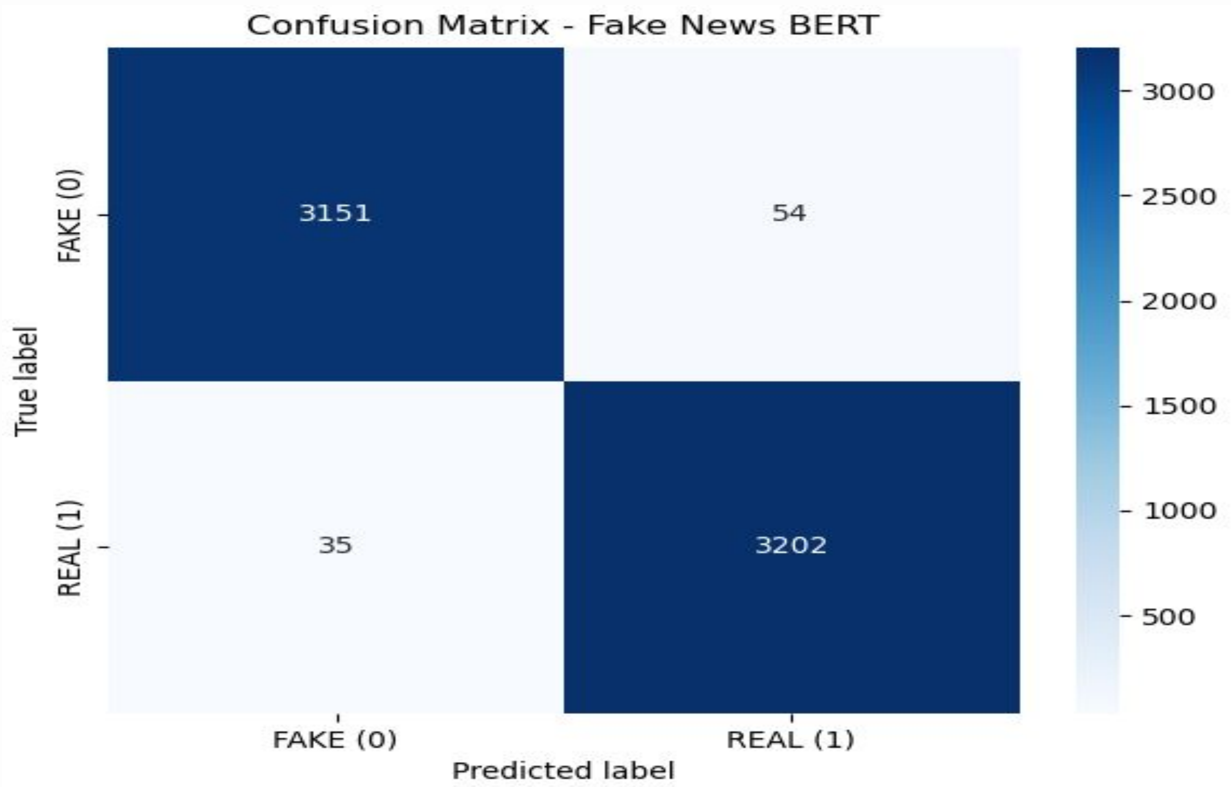


1. **Classical ML (Best): TF-IDF + LinearSVC → 94.45%**

   → Trains in <10 seconds on CPU • Model size <50 MB

2. **BERT (Initial – No tuning): 63.86%**

   → Extremely low recall, useless in practice



3. **BERT (After proper fine-tuning): 98.62%**

   → Learning Rate: 2e-5 | Epochs: 3 | Batch: 16

   → Training time: ~55 minutes on RTX 3060

   → Near-perfect confusion matrix (only 89 errors total!)

Top Fake vs Real Indicators

Top FAKE Signals

- boom!
- says
- wow!
- shocking
- you won't believe
- holy moly
- hilarious
- breaking:

Top REAL Signals

- minister
- government
- according to
- president
- reuters
- official
- statement
- announced

thank you