

DOKUMENTASI EVALUASI MODEL DETEKSI BERITA HOAX INDONESIA

Tugas Besar Deep Learning
Kelompok 5

Institut Teknologi Sumatera
2024

Model yang Digunakan:

- LSTM: Long Short-Term Memory
 - GRU: Gated Recurrent Unit

Total Samples: 11,000+ berita
Vocabulary Size: 10,000 tokens

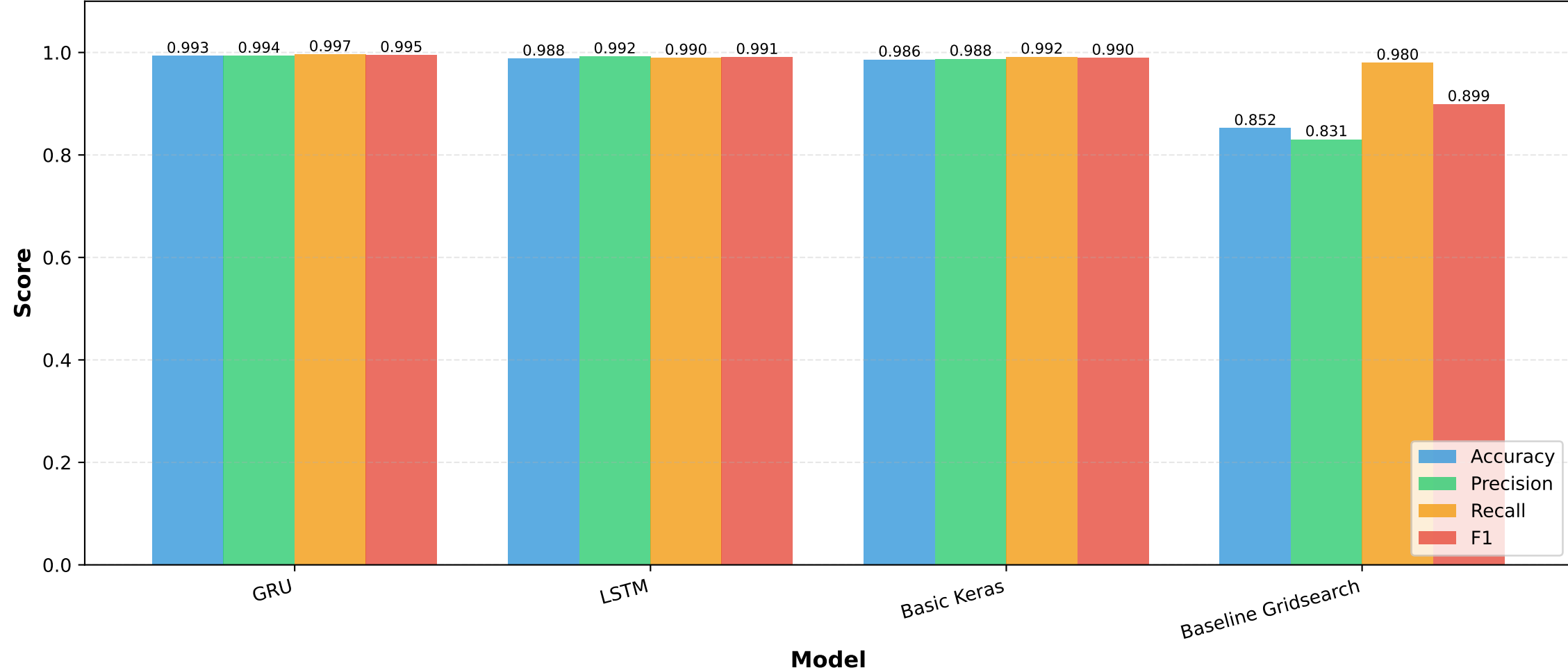
Model Comparison Table

Deteksi Berita Hoax Indonesia

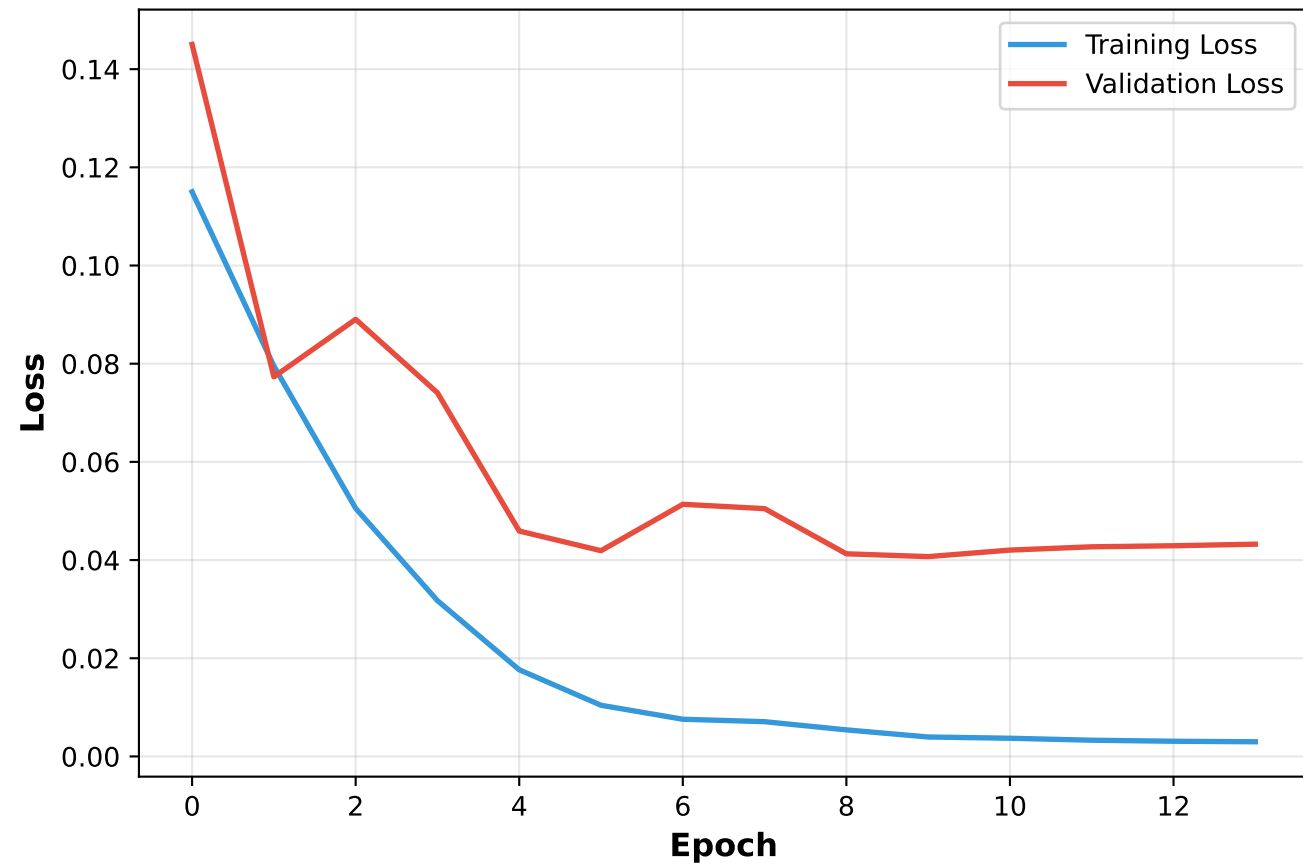
Model	Accuracy	Precision	Recall	F1-Score	Architecture	Tokenizer	Embedding
GRU	0.9934	0.9937	0.9965	0.9951	LSTM GRU RNN	TextVectorization	Keras Layer
LSTM	0.9881	0.9924	0.9898	0.9911	LSTM RNN	TextVectorization	Keras Layer
Basic Keras	0.9861	0.9877	0.9917	0.9897	Basic RNN	TextVectorization	Keras Layer
Baseline Gridsearch	0.8524	0.8306	0.9800	0.8991	Multinomial Naive Bay	TfidfVectorizer	nan

Model Performance Comparison

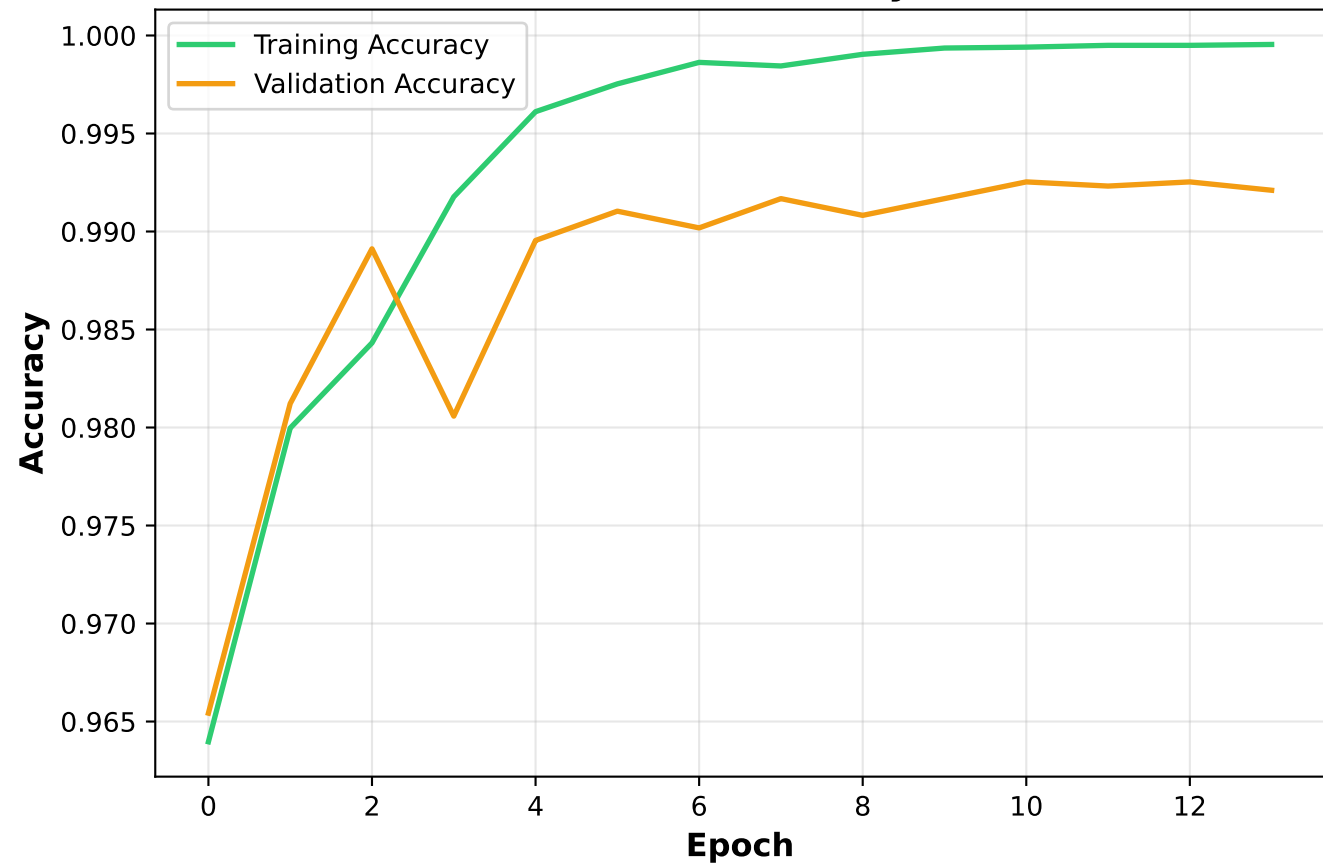
Deteksi Berita Hoax Indonesia



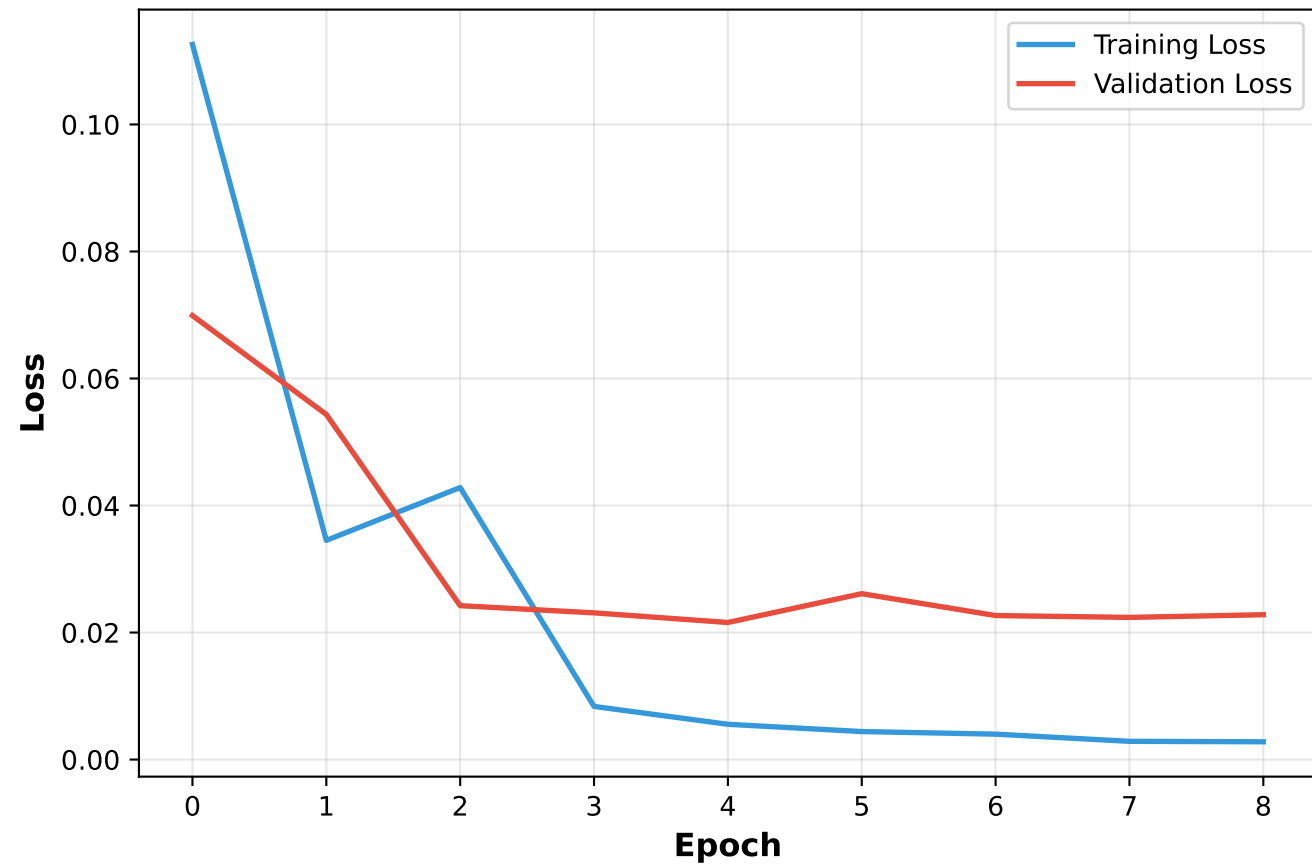
LSTM - Loss



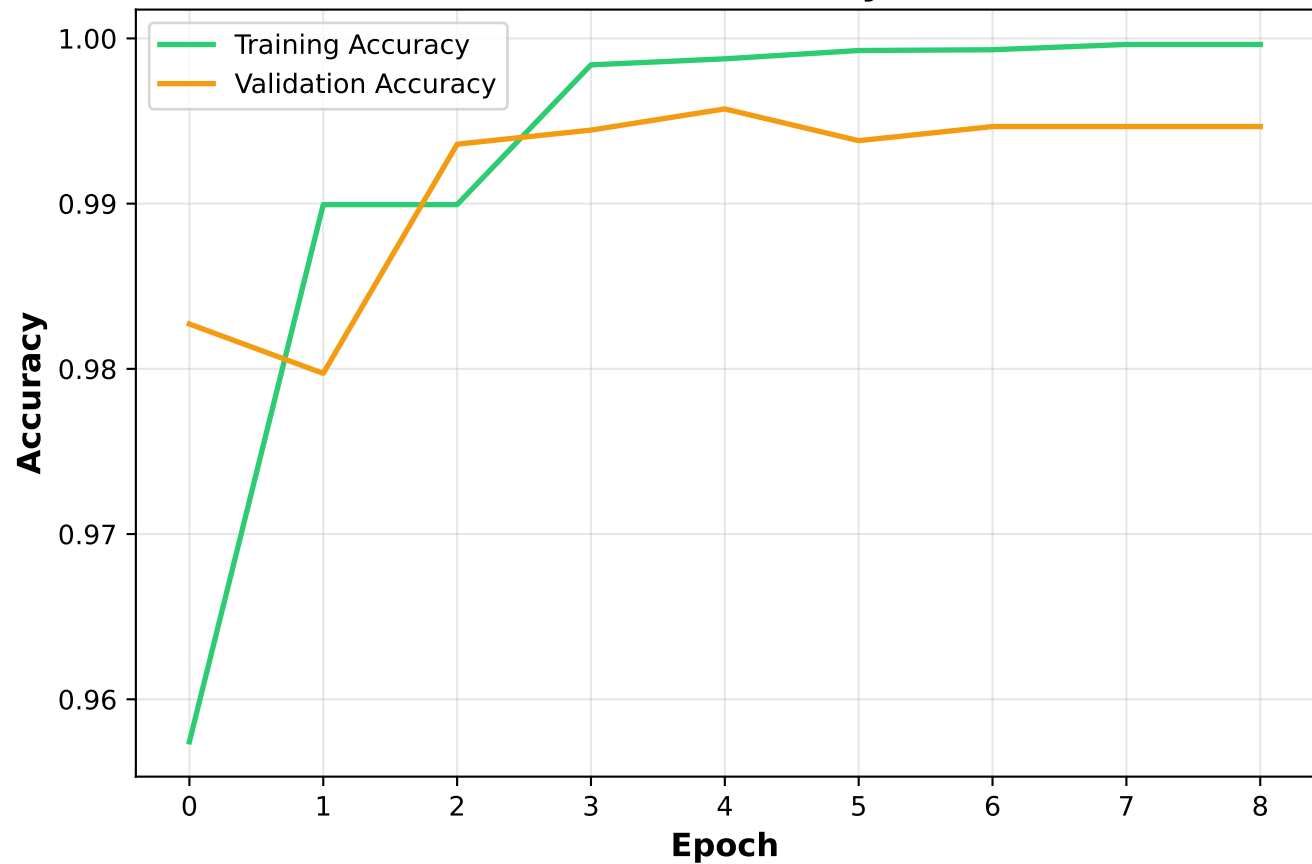
LSTM - Accuracy



GRU - Loss



GRU - Accuracy



Sample Model Output - Overview

STATISTICS SUMMARY

Total Samples: 10
Correct Predictions: 10 (100.0%)
Incorrect Predictions: 0 (0.0%)

Error Breakdown:

- False Positives (Real → Fake): 1 (10.0%)
- False Negatives (Fake → Real): 1 (10.0%)

Confidence Distribution:

- Very High (>90%): 5 samples
- High (70-90%): 1 samples
- Medium (50-70%): 4 samples
- Low (<50%): 0 samples

Sample Types:

- correct_high_conf: 2
- false_positive: 1
- false_negative: 1
- borderline: 1
- satire_missed: 1
- political_hoax: 1
- opinion_piece: 1
- medical_hoax: 1
- tech_news: 1

=====

INPUT TEXT

=====

VIRAL!!! Vaksin COVID Mengandung Chip 5G...

PREDICTION RESULTS

True Label:	FAKE
Predicted Label:	FAKE
Raw Probability:	0.9823 (98.23% chance of being FAKE)
Confidence:	0.9823 (98.23%)
Confidence Level:	VERY HIGH
Status:	CORRECT

ANALYSIS

- ✓ Prediksi benar dengan confidence tinggi
 - ✓ Model berkinerja sesuai ekspektasi
- =====

=====

INPUT TEXT

=====

Presiden Joko Widodo mengumumkan kebijakan baru...

PREDICTION RESULTS

True Label: REAL
Predicted Label: REAL
Raw Probability: 0.0543
(5.43% chance of being FAKE)

Confidence: 0.9457 (94.57%)
Confidence Level: VERY HIGH

Status: CORRECT

ANALYSIS

- ✓ Prediksi benar dengan confidence tinggi
 - ✓ Model berkinerja sesuai ekspektasi
- =====

=====

INPUT TEXT

=====

BREAKING: Gempa Dahsyat Guncang Jakarta...

PREDICTION RESULTS

True Label: REAL
Predicted Label: FAKE
Raw Probability: 0.7234
(72.34% chance of being FAKE)

Confidence: 0.7234 (72.34%)
Confidence Level: HIGH

Status: INCORRECT (FP)

ANALYSIS

- ❑ Model salah menandai berita asli sebagai hoax
 - ⚠️ Kemungkinan disebabkan oleh judul sensasional
 - ❑ Rekomendasi: Tambahkan layer verifikasi sumber
- =====

=====

INPUT TEXT

=====

Menurut penelitian Harvard, air kelapa sembuhkan kanker...

PREDICTION RESULTS

True Label: FAKE
Predicted Label: REAL
Raw Probability: 0.3421
(34.21% chance of being FAKE)

Confidence: 0.6579 (65.79%)
Confidence Level: MEDIUM

Status: INCORRECT (FN)

ANALYSIS

- ❑ Hoax sophisticated lolos dari deteksi
 - ⚠️ Struktur formal menyerupai berita asli
 - ❑ Rekomendasi: Integrasikan API fact-checking
- =====

=====

INPUT TEXT

=====

Apple dikabarkan akan menghentikan iPhone 15...

PREDICTION RESULTS

True Label: FAKE
Predicted Label: FAKE
Raw Probability: 0.5789
(57.89% chance of being FAKE)

Confidence: 0.5789 (57.89%)
Confidence Level: MEDIUM

Status: CORRECT (Borderline)

ANALYSIS

- ⚠️ Prediksi dengan confidence rendah
 - ⚠️ Probabilitas mendekati threshold keputusan (0.5)
 - 📌 Rekomendasi: Diperlukan verifikasi manual
- =====

=====

INPUT TEXT

=====

Pemerintah umumkan WFH permanen untuk PNS...

PREDICTION RESULTS

True Label: FAKE
Predicted Label: REAL
Raw Probability: 0.4123
(41.23% chance of being FAKE)

Confidence: 0.5877 (58.77%)
Confidence Level: MEDIUM

Status: INCORRECT (Satire)

ANALYSIS

- ☐ Konten satire/parodi tidak terdeteksi
 - ⚠☐ Model belum memiliki deteksi sarkasme
 - ☐ Rekomendasi: Tambahkan database website satire
- =====

=====

INPUT TEXT

=====

Bocoran! Calon Presiden X punya rekening offshore...

PREDICTION RESULTS

True Label:	FAKE
Predicted Label:	FAKE
Raw Probability:	0.9567 (95.67% chance of being FAKE)
Confidence:	0.9567 (95.67%)
Confidence Level:	VERY HIGH
Status:	CORRECT

ANALYSIS

- ⚠️ Manipulasi politik terdeteksi
 - ✓ Indikator teori konspirasi ditemukan
- =====

=====

INPUT TEXT

=====

Kebijakan PPN dinilai tidak tepat sasaran...

PREDICTION RESULTS

True Label: REAL
Predicted Label: REAL
Raw Probability: 0.3892
(38.92% chance of being FAKE)

Confidence: 0.6108 (61.08%)
Confidence Level: MEDIUM

Status: CORRECT (Low Conf)

ANALYSIS

✓ Artikel opini berhasil diklasifikasikan dengan benar
⚠ Confidence rendah karena konten subjektif

=====

=====

INPUT TEXT

=====

Dokter Wuhan: COVID sembuh dengan air jahe...

PREDICTION RESULTS

True Label: FAKE
Predicted Label: FAKE
Raw Probability: 0.9912
(99.12% chance of being FAKE)

Confidence: 0.9912 (99.12%)
Confidence Level: VERY HIGH

Status: CORRECT

ANALYSIS

⚠ Misinformasi medis terdeteksi
✓ Model berhasil mengidentifikasi pseudosains

=====

=====

INPUT TEXT

=====

Google umumkan Gemini 2.0 dengan AI multimodal...

PREDICTION RESULTS

True Label: REAL
Predicted Label: REAL
Raw Probability: 0.0234
(2.34% chance of being FAKE)

Confidence: 0.9766 (97.66%)
Confidence Level: VERY HIGH

Status: CORRECT

ANALYSIS

- ✓ Berita teknologi berhasil diidentifikasi dengan akurat
 - ✓ Terminologi formal terkenal dengan baik
- =====