

**ANALISIS PERFORMA BILSTM DAN BIGRU DALAM
KLASIFIKASI BERITA HOAKS BAHASA INDONESIA**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

ADIB RAMADHAN

22.11.4873

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

ANALISIS PERFORMA BILSTM DAN BIGRU DALAM KLASIFIKASI BERITA HOAKS BAHASA INDONESIA

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

ADIB RAMADHAN

22.11.4873

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS PERFORMA BILSTM DAN BIGRU DALAM
KLASIFIKASI BERITA HOAKS BAHASA INDONESIA**

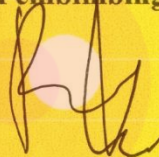
yang disusun dan diajukan oleh

ADIB RAMADHAN

22.11.4873

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 27 Januari 2026

Dosen Pembimbing,



Theopilus Bayu Sasongko, S.Kom. M.Eng.
NIK. 190302375

HALAMAN PENGESAHAN
SKRIPSI
ANALISIS PERFORMA BILSTM DAN BIGRU DALAM KLASIFIKASI
BERITA HOAKS BAHASA INDONESIA

yang disusun dan diajukan oleh

ADIB RAMADHAN

22.11.4873

Telah dipertahankan di depan Dewan Penguji
pada tanggal 27 Januari 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Sharazita Dyah Anggita, S.Kom., M.Kom.
NIK. 190302285

Firman Asharudin, S.Kom., M.Kom.
NIK. 190302315

Theopilus Bayu Sasongko, S.Kom., M.Eng.
NIK. 190302375



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 27 Januari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : ADIB RAMADHAN
NIM : 22.11.4873

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS PERFORMA BILSTM DAN BIGRU DALAM KLASIFIKASI BERITA HOAKS BAHASA INDONESIA

Dosen Pembimbing : Theopilus Bayu Sasongko, S.Kom. M.Eng.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 27 Januari 2026

Yang Menyatakan,



ADIB RAMADHAN

HALAMAN PERSEMBAHAN

Puji dan syukur peneliti panjatkan ke hadirat Allah SWT atas rahmat dan karunia-Nya sehingga skripsi ini dapat diselesaikan. Dengan penuh rasa hormat, skripsi ini peneliti persembahkan kepada:

1. Kedua orang tua dan keluarga peneliti yang senantiasa memberikan doa, dukungan, serta motivasi selama peneliti menempuh pendidikan dan menyelesaikan skripsi ini.
2. Bapak Theopilus Bayu Sasongko, S.Kom., M.Eng. selaku dosen pembimbing, yang telah memberikan bimbingan, arahan, dan masukan selama proses penelitian dan penyusunan skripsi ini.
3. Universitas Amikom Yogyakarta sebagai institusi pendidikan yang telah memberikan lingkungan akademik dan fasilitas pendukung selama proses studi.
4. Rekan-rekan seperjuangan yang telah memberikan dukungan dan motivasi selama masa perkuliahan.

KATA PENGANTAR

Puji dan syukur peneliti panjatkan ke hadirat Allah SWT atas rahmat dan karunia-Nya sehingga peneliti dapat menyelesaikan skripsi ini. Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta.

Dalam proses penyusunan skripsi ini, peneliti menyadari bahwa penyelesaiannya ini tidak terlepas dari bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini peneliti ingin menyampaikan rasa hormat dan ucapan terima kasih kepada:

1. Kedua orang tua dan keluarga peneliti, yang senantiasa memberikan doa, dukungan, serta motivasi sehingga peneliti dapat menyelesaikan skripsi ini.
2. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta, atas kesempatan dan fasilitas yang diberikan selama peneliti menempuh pendidikan.
3. Bapak Theopilus Bayu Sasongko, S.Kom., M.Eng., selaku dosen pembimbing, yang dengan penuh kesabaran telah meluangkan waktu, memberikan bimbingan, arahan, serta masukan yang sangat berharga selama proses penyusunan skripsi ini.
4. Bapak Firman Asharudin, S.Kom., M.Kom. dan Ibu Sharazita Dyah Anggita, S.Kom., M.Kom., selaku Dosen Penguji, yang telah memberikan saran, masukan, serta perbaikan guna menyempurnakan skripsi ini.
5. Rekan-rekan seperjuangan, yang telah memberikan dukungan, kerja sama, dan semangat selama masa perkuliahan hingga penyusunan skripsi ini.

Yogyakarta, 27 Januari 2026

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xv
INTISARI	xvii
<i>ABSTRACT</i>	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penelitian	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori.....	27

2.2.1	Berita Hoaks.....	27
2.2.2	Data Mining	28
2.2.3	Sentiment Analysis	30
2.2.4	Word Embedding	32
2.2.4.1	FastText.....	32
2.2.4.2	GloVe (Global Vectors for Word Representation)	34
2.2.5	Natural Language Processing (NLP)	37
2.2.6	Recurrent Neural Network (RNN).....	37
2.2.7	Long Short-Term Memory (LSTM)	40
2.2.7.1	Bidirectional Long Short-Term Memory (BiLSTM).....	44
2.2.8	Gated Recurrent Unit (GRU).....	45
2.2.8.1	Bidirectional Gated Recurrent Unit (BiGRU)	48
2.2.9	Hyperparameter Tuning	50
2.2.10	K-Fold Cross Validation	51
2.2.11	Evaluasi Model	52
2.2.11.1	Akurasi	52
2.2.11.2	Presisi	52
2.2.11.3	Recall	53
2.2.11.4	F1-Score	53
2.2.11.5	Confusion Matrix	53
2.2.12	Flask.....	54
BAB III METODE PENELITIAN		55
3.1	Objek Penelitian.....	55
3.2	Alur Penelitian	55
3.2.1	Identifikasi Masalah	56

3.2.2	Studi Literatur	57
3.2.3	Pengumpulan Data	58
3.2.4	Preprocessing Data.....	58
3.2.4.1	Case Folding	58
3.2.4.2	Cleaning	59
3.2.4.3	Normalisasi Singkatan	59
3.2.4.4	Stopword Removal.....	59
3.2.4.5	Stemming	60
3.2.4.6	Tokenizing	60
3.2.5	Exploratory Data Analysis.....	61
3.2.5.1	Distribusi Label.....	61
3.2.5.2	Visualisasi Word Cloud	61
3.2.6	Word Embedding	62
3.2.7	Pelatihan Model	63
3.2.7.1	Hyperparameter Tuning.....	63
3.2.7.2	K-Fold Cross Validation	63
3.2.8	Evaluasi Model	64
3.2.9	Analisis Hasil & Pembahasan.....	65
3.2.10	Deployment Model	66
3.2.11	Kesimpulan & Saran	67
3.3	Alat dan Bahan.....	67
3.3.1	Data Penelitian	68
3.3.2	Alat/Instrumen Penelitian	68
BAB IV HASIL DAN PEMBAHASAN		71
4.1	Pengumpulan Data	71

4.2	Preprocessing Data.....	72
4.2.1	Case Folding	73
4.2.2	Cleaning	74
4.2.3	Normalisasi Singkatan	75
4.2.4	Stopword Removal.....	76
4.2.5	Stemming	77
4.2.6	Tokenizing	78
4.3	Exploratory Data Analysis	80
4.3.1	Distribusi Label.....	80
4.3.2	Visualisasi Word Cloud	81
4.4	Word Embedding	83
4.5	Pelatihan Model	86
4.5.1	Hyperparameter Tuning	87
4.5.2	K-Fold Cross Validation	88
4.6	Evaluasi Model	95
4.6.1	Classification Report.....	95
4.6.2	Confusion Matrix	98
4.7	Analisis Hasil Perbandingan Model BiLSTM dan Model BiGRU.....	101
4.8	Deployment Model	103
4.8.1	Hasil Berita Non-Hoaks	104
4.8.2	Hasil Berita Hoaks	105
BAB V PENUTUP		106
5.1	Kesimpulan	106
5.2	Saran	107
REFERENSI		109

DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian	9
Tabel 2. 2 Tabel Confusion Matrix.....	53
Tabel 3. 1 Data Penelitian	68
Tabel 3. 2 Perangkat Keras (Hardware).....	68
Tabel 3. 3 Perangkat Lunak (Software).....	69
Tabel 4. 1 Distribusi Data	71
Tabel 4. 2 Dataset	71
Tabel 4. 3 Hasil Penggabungan Fitur Judul dan Narasi Berita	72
Tabel 4. 4 Hasil Case Folding.....	73
Tabel 4. 5 Hasil Cleaning.....	74
Tabel 4. 6 Hasil Normalisasi Singkatan.....	75
Tabel 4. 7 Hasil Stopword Removal	76
Tabel 4. 8 Hasil Stemming.....	77
Tabel 4. 9 Hasil Tokenizing.....	79
Tabel 4. 10 Hasil Word Embedding FastText dan GloVe	84
Tabel 4. 11 Hasil Konfigurasi Parameter Terbaik	87
Tabel 4. 12 Hasil K-Fold Cross Validation Model BiLSTM.....	89
Tabel 4. 13 Hasil K-Fold Cross Validation Model BiGRU	90
Tabel 4. 14 Hasil Classification Report Model BiLSTM (FastText).....	95
Tabel 4. 15 Hasil Classification Report Model BiLSTM (GloVe)	96
Tabel 4. 16 Hasil Classification Report Model BiGRU (FastText).....	97
Tabel 4. 17 Hasil Classification Report Model BiGRU (GloVe)	97
Tabel 4. 18 Hasil Perbandingan Model BiLSTM dan Model BiGRU	102

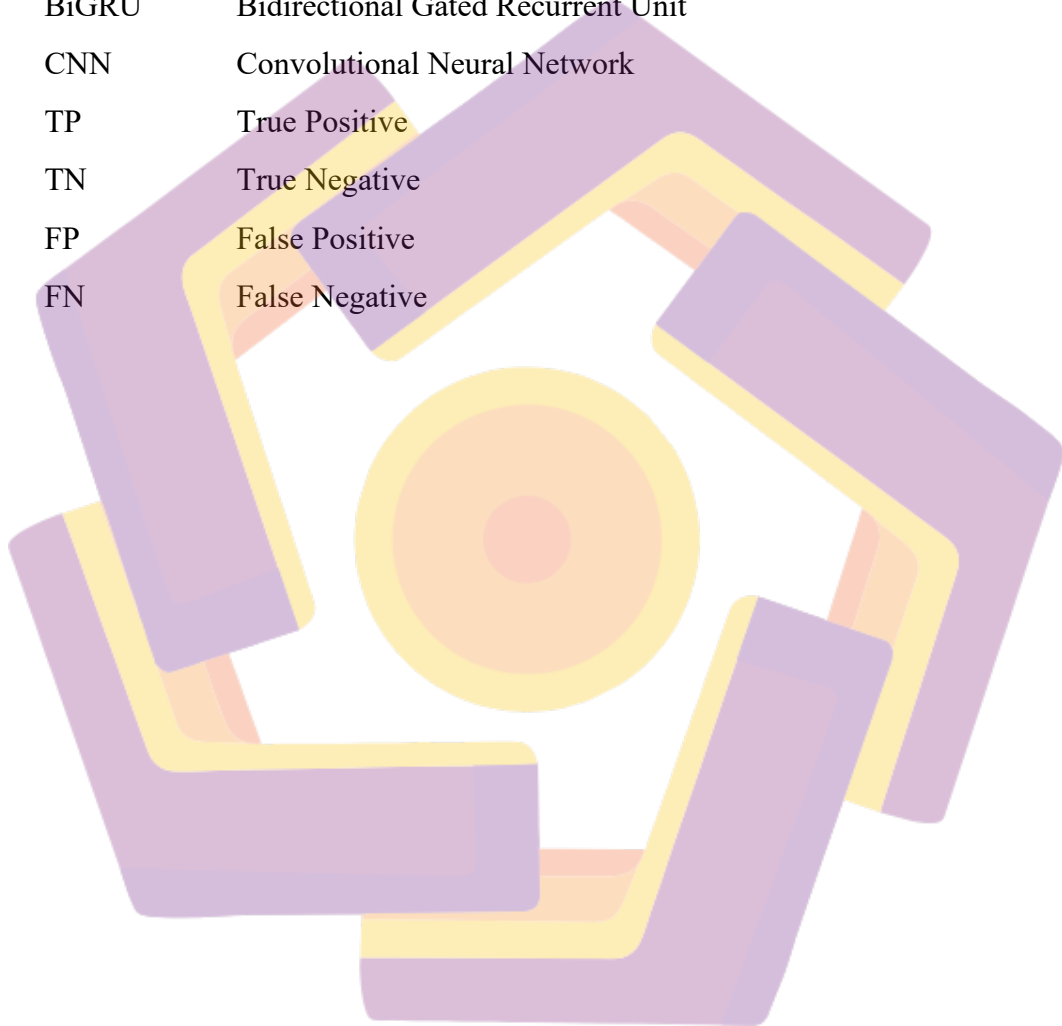
DAFTAR GAMBAR

Gambar 2. 1 Arsitektur RNN	38
Gambar 2. 2 Arsitektur LSTM.....	41
Gambar 2. 3 Arsitektur BiLSTM	44
Gambar 2. 4 Arsitektur GRU	46
Gambar 2. 5 Arsitektur BiGRU	49
Gambar 3. 1 Alur Penelitian	56
Gambar 4. 1 Distribusi Label.....	80
Gambar 4. 2 Visualisasi Word Cloud Berita Non-Hoaks	81
Gambar 4. 3 Visualisasi Word Cloud Berita Hoaks	82
Gambar 4. 4 Grafik Learning Curve Model BiLSTM dan BiGRU	92
Gambar 4. 5 Perbandingan Learning Curve BiLSTM (FastText) Sebelum dan Sesudah Menggunakan L2 Normalization.....	93
Gambar 4. 6 Confusion Matrix Model BiLSTM dan BiGRU	99
Gambar 4. 7 Antarmuka Sistem Klasifikasi Berita Hoaks	103
Gambar 4. 8 Hasil Berita Non-Hoaks	104
Gambar 4. 9 Hasil Berita Hoaks	105

DAFTAR LAMBANG DAN SINGKATAN

v_w	Vektor representasi kata w
G_w	Himpunan character n-grams dari kata w
z_g	Vektor embedding dari n-gram g
X_{ij}	Frekuensi kemunculan bersama kata i dan j
V	Ukuran kosakata
J	Fungsi objektif pada model GloVe
x_t	Vektor input pada waktu ke- t
h_t	Hidden state pada waktu ke- t
h_{t-1}	Hidden state pada waktu sebelumnya
C_t	Cell state LSTM pada waktu ke- t
C_{t-1}	Cell state LSTM pada waktu sebelumnya
f_t	Nilai forget gate pada LSTM
i_t	Nilai input gate pada LSTM
o_t	Nilai output gate pada LSTM
$\hat{C}_t$	Candidate cell state pada LSTM
R_t	Reset gate pada GRU
Z_t	Update gate pada GRU
$\hat{h}_t$	Kandidat hidden state pada GRU
W, U, V	Matriks bobot (weight) jaringan
b	Bias
σ	Fungsi aktivasi sigmoid
$\tanh$	Fungsi aktivasi hyperbolic tangent
AI	Artificial Intelligence
NLP	Natural Language Processing
GloVe	Global Vectors for Word Representation
OOV	Out-of-Vocabulary
TF	Term Frequency

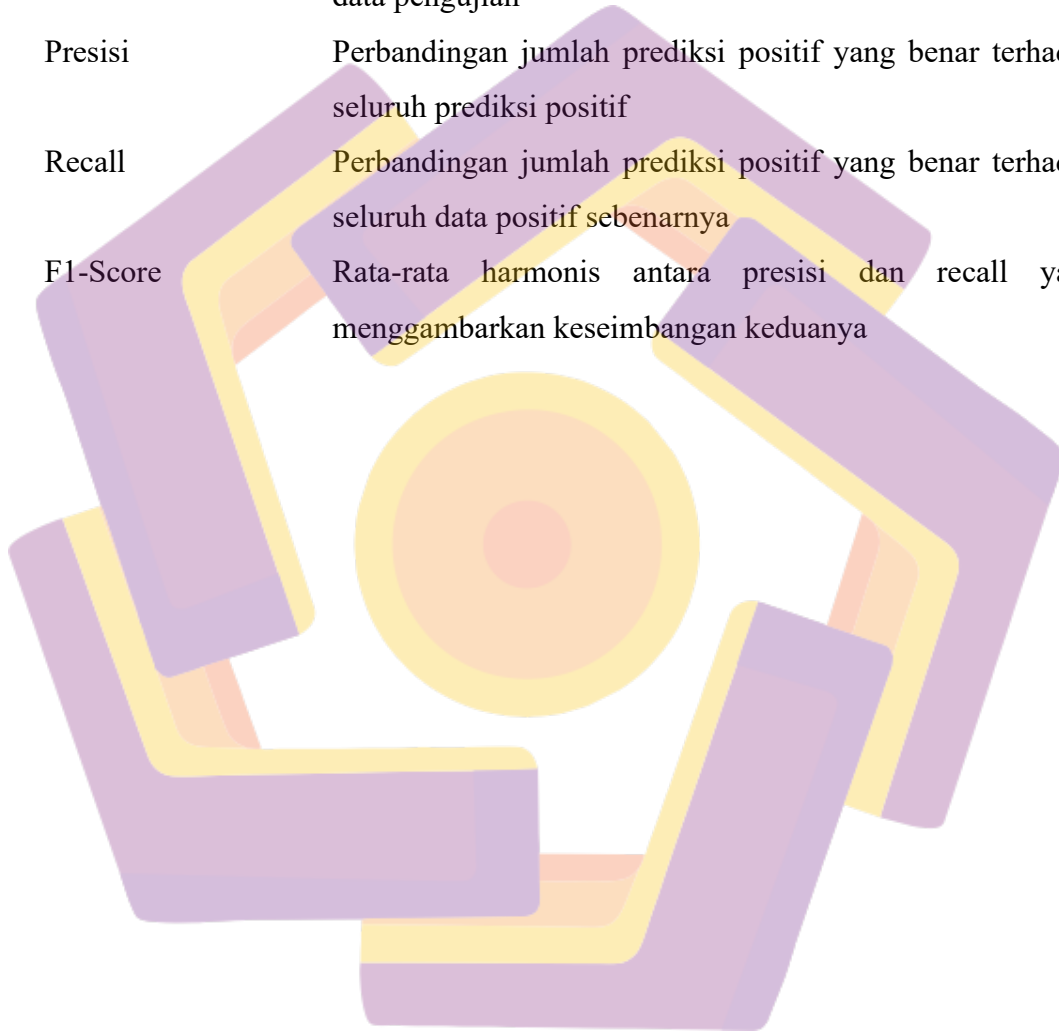
IDF	Inverse Document Frequency
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
BiLSTM	Bidirectional Long Short-Term Memory
GRU	Gated Recurrent Unit
BiGRU	Bidirectional Gated Recurrent Unit
CNN	Convolutional Neural Network
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative



DAFTAR ISTILAH

Klasifikasi	Proses pengelompokan data ke dalam kelas atau kategori berdasarkan pola yang dipelajari model
Dataset	Kumpulan data yang digunakan dalam proses pelatihan, validasi, dan pengujian model
Deep Learning	Cabang machine learning yang menggunakan jaringan saraf berlapis untuk mempelajari pola data kompleks
Bidirectional	Pendekatan pemrosesan data berurutan dari dua arah, yaitu dari awal ke akhir dan dari akhir ke awal
Preprocessing Teks	Tahap pembersihan dan normalisasi teks sebelum pemodelan
Word Cloud	Visualisasi data teks yang menampilkan kata berdasarkan frekuensi kemunculannya
Korpus	Kumpulan data teks yang digunakan sebagai sumber pembelajaran model word embedding
Out-of-Vocabulary	Kata yang tidak terdapat dalam kosakata model saat pelatihan
Word Embedding	Teknik representasi kata dalam bentuk vektor
Generalisasi	Kemampuan model untuk memberikan prediksi yang baik pada data yang belum pernah dilihat
Overfitting	Kondisi ketika model terlalu menyesuaikan data latih sehingga performanya menurun pada data baru
Hyperparameter	Proses penyesuaian parameter model untuk memperoleh performa terbaik
Batch Size	Jumlah data yang diproses model dalam satu iterasi pelatihan
Dropout	Teknik regularisasi dengan menonaktifkan sebagian neuron secara acak saat pelatihan untuk mencegah overfitting
Learning Rate	Parameter yang mengatur besar langkah pembaruan bobot model selama proses pelatihan

Epoch	Satu siklus penuh proses pelatihan model terhadap seluruh dataset
Confusion Matrix	Matriks evaluasi yang menunjukkan perbandingan antara label prediksi dan label sebenarnya
Akurasi	Persentase prediksi yang benar dibandingkan dengan seluruh data pengujian
Presisi	Perbandingan jumlah prediksi positif yang benar terhadap seluruh prediksi positif
Recall	Perbandingan jumlah prediksi positif yang benar terhadap seluruh data positif sebenarnya
F1-Score	Rata-rata harmonis antara presisi dan recall yang menggambarkan keseimbangan keduanya



INTISARI

Penyebaran berita hoaks di media digital semakin meningkat dan memberikan dampak negatif terhadap masyarakat, seperti menurunnya kepercayaan publik, munculnya kesalahpahaman informasi, serta potensi gangguan stabilitas sosial. Oleh karena itu, penelitian ini bertujuan untuk menganalisis dan membandingkan performa model klasifikasi berita hoaks berbahasa Indonesia menggunakan metode *Bidirectional Long Short-Term Memory* (BiLSTM) dan *Bidirectional Gated Recurrent Unit* (BiGRU) dengan fitur ekstraksi *word embedding* FastText dan GloVe. Dataset berita hoaks dan non-hoaks diproses melalui tahap preprocessing yang meliputi *case folding*, *cleaning*, normalisasi singkatan, *stopword removal*, *stemming* dan *tokenizing* untuk persiapan pelatihan model. Selanjutnya model dilatih menggunakan *hyperparameter tuning* dan *k-fold cross validation*, kemudian di evaluasi menggunakan metrik evaluasi akurasi, presisi, recall, F1-score dan confusion matrix. Hasil penelitian menunjukkan bahwa seluruh model menghasilkan performa yang sangat baik. Model BiLSTM dengan FastText memperoleh akurasi sebesar 98,90%, sedangkan BiLSTM dengan GloVe mencapai akurasi 98,72%. Model BiGRU menunjukkan performa yang lebih unggul, dengan akurasi 99,02% pada FastText dan 99,27% pada GloVe, disertai nilai Presisi, *recall*, dan *F1-score* di atas 99%. Hasil ini mengindikasikan bahwa model BiGRU dengan *word embedding* GloVe merupakan kombinasi terbaik dalam penelitian ini untuk klasifikasi berita hoaks berbahasa Indonesia.

Kata kunci: Berita Hoaks, Klasifikasi Teks, BiLSTM, BiGRU, FastText, GloVe

ABSTRACT

The spread of fake news on digital media has increased significantly and has caused negative impacts on society, such as declining public trust, misinformation, and potential threats to social stability. Therefore, this study aims to analyze and compare the performance of Indonesian fake news classification models using Bidirectional Long Short-Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (BiGRU) with FastText and GloVe word embedding feature extraction methods. The fake and non-fake news datasets were processed through several preprocessing stages, including case folding, text cleaning, abbreviation normalization, stopword removal, stemming, and tokenization to prepare the data for model training. Furthermore, the models were trained using hyperparameter tuning and 5-Fold Cross Validation, and evaluated using accuracy, precision, recall, F1-score, and confusion matrix metrics. The experimental results show that all models achieved excellent performance. The BiLSTM model with FastText achieved an accuracy of 98.90%, while BiLSTM with GloVe reached 98.72%. The BiGRU models demonstrated superior performance, achieving 99.02% accuracy with FastText and 99.27% Akurasi with GloVe, accompanied by precision, recall, and F1-score values above 99%. These findings indicate that the BiGRU model combined with GloVe word embedding is the best-performing approach in this study for Indonesian fake news classification.

Keyword: Fake News, Text Classification, BiLSTM, BiGRU, FastText, GloVe