

SKRIPSI

**ANALISIS KOMPARATIF ALGORITMA RNN, LSTM, DAN
BILSTM PADA TEKS PENDEK BERBAHASA INDONESIA**



Muhammad Ramdany Juliansyah

1512622006

Intelligentia - Dignitas

**PROGRAM STUDI S-1 PENDIDIKAN TEKNIK
INFORMATIKA DAN KOMPUTER**

FAKULTAS TEKNIK

UNIVERSITAS NEGERI JAKARTA

2026

LEMBAR PENGESAHAN TUGAS AKHIR
BENTUK : SKRIPSI

Nama Mahasiswa : Muhammad Ramdany Juliansyah
NIM : 1512622006
Program Studi/Fakultas : Pendidikan Teknik Informatika dan Komputer /
Teknik
Judul Penelitian : Analisis Komparatif Algoritma RNN LSTM dan
BILSTM Pada Teks Pendek Berbahasa Indonesia

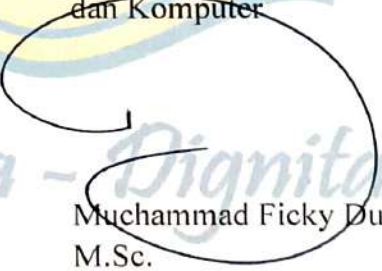
ini telah dipertahankan dihadapan dewan penguji dan dinyatakan **Lulus** / ~~Tidak Lulus~~ * Ujian Tugas Akhir oleh Tim Penguji pada Tanggal 13 Juli 2026
Tim Penguji,

Ketua Penguji	Nur Elah, S.Kom., M.T., NUPTK. 2736769670230442	
Anggota Penguji (Penguji Ahli 1)	Fitrah Izul Falaq, M.Pd NIDN. 0729019801	
Anggota Penguji (Penguji Ahli 2)	Bambang Satrio, S. Pd., M. Pd. NUPTK. 474777567613022	
Anggota Penguji (Pembimbing 1)	Dr. Widodo, S. Kom., M. Kom. NIDN. 0025037206	
Anggota Penguji (Pembimbing 2)	Neng Ayu Herawati, S. Pd., M. T. NUPTK. 125277667723015	

Mengetahui,
Dekan Fakultas Teknik


Prof. Dr. Neneng Siti
Siti Ambarwati, S.Si.,
Apt., M.Si
NIDN. 0029027204

Koordinator Program Studi
Pendidikan Teknik Informatika
dan Komputer


Muchammad Ficky Duskarnaen,
M.Sc.

NIDN. 0024097304

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Skripsi ini merupakan karya asli dan belum pernah diajukan untuk mendapatkan gelar akademik sarjana, baik di Universitas Negeri Jakarta maupun di Perguruan Tinggi lain.
2. Skripsi ini belum dipublikasikan kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
3. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Universitas Negeri Jakarta.

Jakarta, 22 Juli 2026



Muhammad Ramdany Juliansyah
No. Reg 1512622006



KEMENTERIAN PENDIDIKAN TINGGI, SAINS DAN TEKNOLOGI
UNIVERSITAS NEGERI JAKARTA
PERPUSTAKAAN DAN KEARSIPAN

Jalan Rawamangun Muka Jakarta 13220

Telepon/Faksimili: 021-4894221

Laman: lib.unj.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Muhammad Ramdany Juliansyah
NIM : 1512622006
Fakultas/Prodi : Teknik/Pendidikan Teknik Informatika dan Komputer
Alamat email : muhammad_1512622006@mhs.unj.ac.id

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan dan Kearsipan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

☒ Skripsi ☐ Tesis ☐ Disertasi ☐ Lain-lain (.....)

yang berjudul :

Analisis Komparatif Algoritma RNN LSTM dan BILSTM Pada Teks Pendek Berbahasa

Indonesia

Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan dan Kearsipan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara **fulltext** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 7 Agustus 2026

Penulis

(Muhammad Ramdany Juliansyah)

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT atas segala rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis Komparatif Algoritma RNN, LSTM, dan BiLSTM Pada Teks Pendek Berbahasa Indonesia" dengan baik dan tepat waktu. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Pendidikan pada Program Studi Pendidikan Teknik Informatika dan Komputer, Fakultas Teknik, Universitas Negeri Jakarta.

Penulis menyadari bahwa terselesaikannya skripsi ini tidak lepas dari bantuan, bimbingan, dukungan, dan doa dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Bapak Muhammad Ficky Duskarnaen, S.T., M.Sc., selaku Koordinator Program Studi Pendidikan Teknik Informatika dan Komputer. Terima kasih atas arahan dan perhatian yang telah Bapak berikan kepada seluruh mahasiswanya untuk menyelesaikan pendidikan tepat waktu.
2. Bapak Dr. Widodo, S.Kom., M.Kom., selaku Dosen Pembimbing I yang dengan penuh keikhlasan, ketelitian, dan kesabaran telah meluangkan waktu, tenaga, dan pikiran dalam memberikan bimbingan, arahan, serta masukan yang sangat berharga bagi penulis selama proses penyusunan skripsi ini.
3. Ibu Neng Ayu Herawati, S.Pd., M.T., selaku Dosen Pembimbing II yang dengan penuh perhatian, ketulusan, dan dedikasi telah meluangkan waktu, tenaga, dan pikiran dalam memberikan bimbingan, arahan, serta motivasi yang sangat berarti bagi penulis selama proses penyusunan skripsi ini.
4. Seluruh Dosen dan Staf Program Studi Pendidikan Teknik Informatika dan Komputer, atas ilmu pengetahuan dan bantuan yang telah diberikan selama masa perkuliahan.
5. Kedua orang tua tercinta dan keluarga besar, Bapak Hari dan Ibu Suryani terima kasih atas doa, kasih sayang, dukungan moral maupun materiil, dan

semangat yang tiada henti diberikan kepada penulis di sepanjang perjalanan akademik ini.

6. Sri Wulandari yang sudah menemani dan membantu di sepanjang perjalanan perkuliahan serta penyusunan skripsi ini.
7. Teman-teman seperjuangan di Program Studi Pendidikan Teknik Informatika dan Komputer angkatan 2022, yang telah memberikan lingkungan yang suportif, semangat, bantuan, dan hiburan selama masa studi.
8. Semua pihak yang tidak dapat penulis sebutkan satu per satu, yang telah turut berkontribusi dalam penyelesaian skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari semua pihak demi perbaikan dan penyempurnaan skripsi ini. Semoga skripsi ini dapat memberikan manfaat bagi perkembangan ilmu pengetahuan, khususnya di bidang *Natural Language Processing* (NLP), kecerdasan buatan (*deep learning*), dan pemrosesan teks berbahasa Indonesia.

Intelligentia - Dignitas
Jakarta, 1 Juli 2026



Muhammad Ramdany Juliansyah
No. Reg 1512622006

ABSTRAK

Muhammad Ramdany Juliansyah, Analisis Komparatif Algoritma RNN LSTM dan BiLSTM Pada Teks Pendek Berbahasa Indonesia. Dosen Pembimbing: Dr. Widodo, S.Kom., M.Kom. dan Neng Ayu Herawati, S.Pd., M.T. Program Studi Pendidikan Teknik Informatika dan Komputer. Fakultas Teknik. Universitas Negeri Jakarta. 2026.

Teks pendek di media sosial memiliki karakteristik keterbatasan konteks dan penggunaan bahasa tidak baku yang menyulitkan proses klasifikasi sentimen. Hingga saat ini, belum ada penelitian yang secara komprehensif membandingkan performa algoritma *Recurrent Neural Network* (RNN), *Long Short-Term Memory* (LSTM), dan *Bidirectional LSTM* (BiLSTM) khusus pada teks pendek berbahasa Indonesia, sehingga tidak tersedia acuan empiris untuk pemilihan model yang optimal. Oleh karena itu, penelitian ini bertujuan untuk menganalisis hasil dan membandingkan performa ketiga model tersebut dalam mengklasifikasikan sentimen terkait judi online pada tahun 2025. Penelitian ini menggunakan data berupa tweet berbahasa Indonesia yang dikumpulkan dari platform X menggunakan kata kunci "judol" atau "judi online". Data tersebut diproses melalui tahapan *text preprocessing* dan pelabelan berbasis leksikon (InSet) yang menghasilkan 11.664 data bersih, lalu direpresentasikan ke dalam bentuk vektor numerik menggunakan algoritma Word2Vec dengan arsitektur *Skip-gram*. Evaluasi performa model diukur berdasarkan tingkat akurasi, F1-score, serta analisis efisiensi komputasi yang mencakup waktu pelatihan, utilisasi CPU, dan RAM. Hasil penelitian menunjukkan bahwa model BiLSTM mencapai performa klasifikasi tertinggi dengan akurasi 68,80% dan rata-rata makro F1-Score sebesar 0,68, diikuti secara ketat oleh model LSTM dengan akurasi 67,55%, sedangkan model RNN mengalami kegagalan (*underfitting*) dengan akurasi 55,98%. Dari segi efisiensi komputasi, terdapat perbandingan terbalik (*trade-off*) yang sangat signifikan. Meskipun BiLSTM memperoleh akurasi tertinggi, peningkatan performanya tergolong tidak signifikan karena hanya terpaut sekitar 1,25% dibandingkan model LSTM. Di sisi lain, biaya komputasi yang harus dibayar oleh BiLSTM jauh lebih besar, yaitu waktu pelatihan terlama (± 910 detik berbanding ± 330 detik pada LSTM) dan konsumsi memori RAM tertinggi (1.000 - 1.100 MB berbanding 750 - 800 MB pada LSTM). Kesimpulannya, meskipun BiLSTM mampu memproses rentang semantik dari dua arah dengan akurasi terbaik, model LSTM terbukti jauh lebih efisien dan optimal untuk diterapkan pada klasifikasi teks pendek karena memberikan performa yang hampir setara dengan beban komputasi yang jauh lebih ringan.

Kata Kunci: Analisis Sentimen, BiLSTM, *Deep Learning*, Judi Online, LSTM, RNN, Teks Pendek, Word2Vec.

ABSTRACT

Muhammad Ramdany Juliansyah, *Comparative Analysis of RNN, LSTM, and BiLSTM Algorithms on Indonesian Short Texts*. Supervisors: Dr. Widodo, S.Kom., M.Kom. and Neng Ayu Herawati, S.Pd., M.T. Study Program of Informatics and Computer Engineering Education, Faculty of Engineering, Universitas Negeri Jakarta, 2026

Short texts on social media are characterized by limited context and the use of informal language, which complicates the sentiment classification process. To date, there has been no research that comprehensively compares the performance of Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Bidirectional LSTM (BiLSTM) algorithms specifically on Indonesian short texts, resulting in a lack of empirical guidelines for optimal model selection. Therefore, this study aims to analyze the results and compare the performance of these three models in classifying online gambling sentiments in 2025. This research utilized Indonesian tweet data collected from the X platform using the keywords "judol" or "judi online". The data was processed through text preprocessing and lexicon-based labeling (InSet), yielding 11,664 clean data points, and subsequently represented into numeric vectors using the Word2Vec algorithm with a Skip-gram architecture. Model performance evaluation was measured based on accuracy, F1-score, and computational efficiency analysis covering training time, CPU, and RAM utilization. The results showed that the BiLSTM model achieved the highest classification performance with an accuracy of 68.80% and a macro average F1-Score of 0.68, closely followed by the LSTM model with an accuracy of 67.55%, while the RNN model experienced underfitting with an accuracy of 55.98%. In terms of computational efficiency, there is a highly significant trade-off. Although BiLSTM achieved the highest accuracy, the performance gain is relatively insignificant as it only outperforms LSTM by approximately 1.25%. On the other hand, the computational cost demanded by BiLSTM is drastically higher, requiring the longest training time (± 910 seconds compared to ± 330 seconds for LSTM) and the highest memory consumption (1,000 - 1,100 MB compared to 750 - 800 MB for LSTM). In conclusion, while BiLSTM effectively processes bidirectional semantic ranges to yield the highest accuracy, the LSTM model proves to be significantly more efficient and optimal for short text classification, offering nearly identical performance with a considerably lighter computational load.

Keywords: BiLSTM, Deep Learning, LSTM, Online Gambling, RNN, Sentiment Analysis, Short Text, Word2Vec.

DAFTAR ISI

LEMBAR PENGESAHAN TUGAS AKHIR BENTUK : SKRIPSI	ii
LEMBAR PERNYATAAN	iii
KATA PENGANTAR.....	iv
ABSTRAK.....	vi
ABSTRACT.....	vii
DAFTAR ISI	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah.....	1
1.2 Identifikasi Masalah	5
1.3 Pembatasan Masalah	5
1.4 Perumusan Masalah	6
1.5 Tujuan Penelitian.....	6
1.6 Manfaat Penelitian	7
1.6.1 Manfaat Teoritis	7
1.6.2 Manfaat Praktis	7
BAB II KAJIAN PUSTAKA	9
2.1 Konsep dan Teori	9
2.1.1 Teks Pendek.....	9
2.1.2 <i>Natural Language Processing</i>	10
2.1.3 <i>Deep Learning</i>	10
2.1.4 <i>Recurrent Neural Network</i>	12
2.1.5 <i>Long Short-Term Memory (LSTM)</i>	16
2.1.6 <i>Bidirectional Long Short-Term Memory (BiLSTM)</i>	19
2.1.7 <i>Text Preprocessing</i>	20
2.1.8 Pelabelan Data Lexicon-Based	22
2.1.9 <i>Word Embedding Word2Vec</i>	22
2.1.10 Evaluasi Performa Model.....	25

2.2 Penelitian Relevan.....	28
2.3 Kerangka Berpikir.....	38
BAB III METODOLOGI PENELITIAN.....	41
3.1 Tempat dan Waktu Penelitian.....	41
3.2 Alat dan Bahan Penelitian.....	41
3.2.1 Alat Penelitian.....	41
3.2.2 Bahan Penelitian.....	43
3.3 Langkah-Langkah Penelitian.....	44
3.4 Prosedur Penelitian.....	45
3.4.1 Pengumpulan Data.....	45
3.4.2 Data Integration.....	45
3.4.3 Data Cleaning dan Preprocessing.....	46
3.4.4 Seleksi Data dan Distribusi Label.....	48
3.4.5 Representasi Teks (<i>Word Embedding</i>).....	48
3.4.6 Pembagian Data.....	49
3.4.7 <i>Training</i> dan <i>Tuning</i> Model.....	49
3.5 Teknik Pengumpulan Data.....	52
3.6 Teknik Evaluasi Data.....	53
BAB IV HASIL DAN PEMBAHASAN.....	56
4.1 Hasil Pengumpulan dan Integrasi Data.....	56
4.2 Hasil Data Cleaning dan Text Preprocessing.....	59
4.2.1 Tahap Data Cleaning dan Casefolding.....	60
4.2.2 Tahap Normalisasi.....	60
4.2.3 Tahap Stopword Removal.....	61
4.2.4 Tahap Stemming.....	62
4.2.5 Tokenisasi.....	62
4.3 Hasil Pelabelan Lexicon-based.....	63
4.4 Hasil Word Embedding Word2Vec.....	65
4.5 Skenario Pelatihan Model dan Hyperparameter.....	67
4.5.1 Data Splitting.....	67
4.5.2 Arsitektur Model Deep Learning.....	68
4.5.3 Konfigurasi Hyperparameter.....	69

4.6 Hasil Pelabelan dan Penyeimbangan Data	72
4.6.1 Evaluasi Model Recurrent Neural Network (RNN).....	73
4.6.2 Evaluasi Model Long Short-Term Memory (LSTM).....	76
4.6.3 Evaluasi Model Bidirectional Long Short-Term Memory (BiLSTM) ..	78
4.7 Hasil Analisis Komparatif dan Pembahasan	81
4.7.1 Analisis Komparatif Kinerja Model	81
4.7.2 Analisis Komparatif Konsumsi Sumber Daya Komputasi.....	82
4.7.3 Relevansi Temuan terhadap Analisis Sentimen Judi Online	85
BAB V KESIMPULAN DAN SARAN.....	87
5.1 Kesimpulan	87
5.2 Saran.....	88
DAFTAR PUSTAKA	90
LAMPIRAN.....	95



Intelligentia - Dignitas

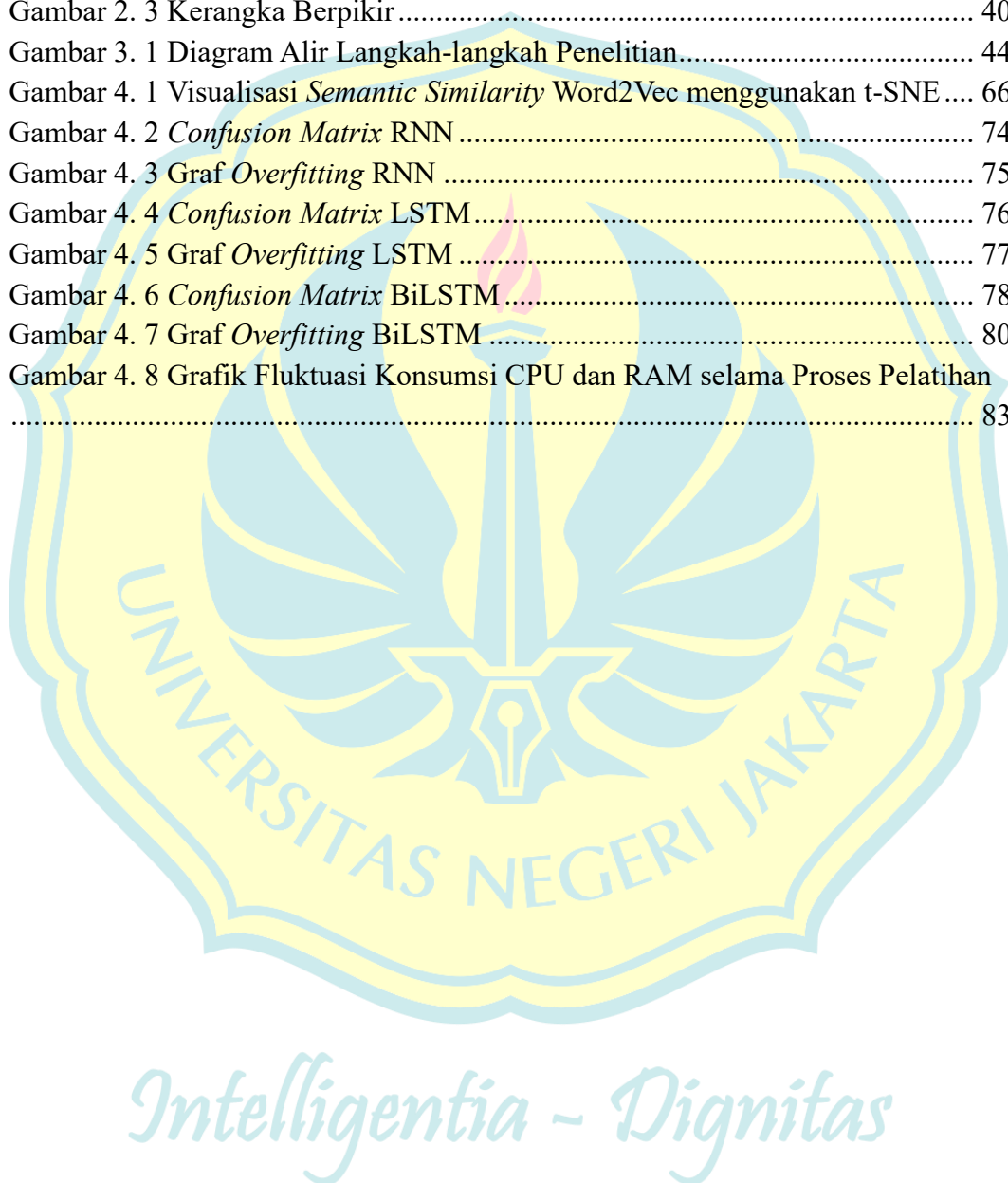
DAFTAR TABEL

Tabel 2. 1 Penelitian Relevan.....	33
Tabel 3. 1 Spesifikasi Desktop	41
Tabel 3. 2 Spesifikasi Laptop	42
Tabel 3. 3 Perangkat Lunak.....	42
Tabel 3. 4 Tabel Data Cleaning dan Preprocessing	47
Tabel 3. 5 Contoh Hyperparameter Model.....	51
Tabel 4. 1 Atribut Dataset Mentah	56
Tabel 4. 2 Ringkasan Dataframe Setelah Seleksi.....	58
Tabel 4. 3 Hasil <i>Data Cleaning</i> dan <i>Case Folding</i>	60
Tabel 4. 4 Hasil Normalisasi Teks.....	61
Tabel 4. 5 Hasil <i>Stopword Removal</i>	61
Tabel 4. 6 Hasil Stemming Teks.....	62
Tabel 4. 7 Hasil Tokenisasi Teks	63
Tabel 4. 8 Distribusi Label Sentimen.....	64
Tabel 4. 9 <i>Hyperparameter</i> Word2Vec.....	65
Tabel 4. 10 Hyperparameter Model Deep Learning.....	69
Tabel 4. 11 Laporan Klasifikasi RNN	74
Tabel 4. 12 Laporan Klasifikasi LSTM.....	77
Tabel 4. 13 Laporan Klasifikasi BiLSTM.....	79
Tabel 4. 14 Ringkasan Hasil Klasifikasi RNN LSTM dan BiLSTM.....	81
Tabel 4. 15 Komparasi Estimasi Konsumsi Sumber Daya Komputasi.....	85

Intelligentia - Dignitas

DAFTAR GAMBAR

Gambar 2.1 Diagram Blok <i>Long Short-Term Memory</i> (LSTM) (Van Houdt, dkk., 2020)	17
Gambar 2. 2 Ilustrasi Perbedaan Arsitektur CBOW dan <i>Skip-Gram</i> (Mikolov, dkk., 2013)	24
Gambar 2. 3 Kerangka Berpikir	40
Gambar 3. 1 Diagram Alir Langkah-langkah Penelitian.....	44
Gambar 4. 1 Visualisasi <i>Semantic Similarity</i> Word2Vec menggunakan t-SNE	66
Gambar 4. 2 <i>Confusion Matrix</i> RNN	74
Gambar 4. 3 Graf <i>Overfitting</i> RNN	75
Gambar 4. 4 <i>Confusion Matrix</i> LSTM.....	76
Gambar 4. 5 Graf <i>Overfitting</i> LSTM	77
Gambar 4. 6 <i>Confusion Matrix</i> BiLSTM	78
Gambar 4. 7 Graf <i>Overfitting</i> BiLSTM	80
Gambar 4. 8 Grafik Fluktuasi Konsumsi CPU dan RAM selama Proses Pelatihan	83



DAFTAR LAMPIRAN

Lampiran 1 Lembar Pernyataan Kelayakan Judul	95
Lampiran 2 Surat Tugas Dosen Pembimbing	96
Lampiran 3 Lembar Konsultasi.....	97
Lampiran 4 Kartu Kehadiran Sempro	101
Lampiran 5 Biodata Peserta Ujian Skripsi	103
Lampiran 6 Surat Permohonan Ujian Skripsi	107
Lampiran 7 Surat Persetujuan Dosen Pembimbing Skripsi	108
Lampiran 8 Surat Persetujuan Dosen Pembimbing Akademik	110
Lampiran 9 Pra-Transkrip Akademik	111
Lampiran 10 Surat Pernyataan Penyelesaian Mata Kuliah Wajib.....	114
Lampiran 11 Dataset Asli	115
Lampiran 12 Dataset Setelah Preprocessing	116
Lampiran 13 Dataset Setelah Preprocess dan Labeling	117



Intelligentia - Dignitas