



**IMPLEMENTASI ALGORITMA *BIDIRECTIONAL LONG
SHORT-TERM MEMORY* MENGGUNAKAN *GLOBAL
VECTORS* UNTUK KLASIFIKASI *TWEET HATE SPEECH***

SKRIPSI

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Komputer
Program Studi Ilmu Komputer

Oleh:

Andhika Cahya Rizqillah

4611419064

**PROGRAM STUDI SARJANA TEKNIK INFORMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI SEMARANG**

2023

PERSETUJUAN PEMBIMBING

Skripsi berjudul “Implementasi Algoritma *Bidirectional Long Short-Term Memory* Menggunakan *Global Vectors* untuk Klasifikasi *Tweet Hate Speech*” yang disusun oleh

Nama : Andhika Cahya Rizqillah

NIM : 4611419064

Prodi : Teknik Informatika

Telah disetujui untuk diajukan ke sidang ujian skripsi.

Semarang, 18 September 2023

Pembimbing,



Aji Purwinarko, S.Si., M.Cs.

NIP. 198509102015041001

PENGESAHAN PENGUJI

Skripsi berjudul “Implementasi Algoritma *Bidirectional Long Short-Term Memory* Menggunakan *Global Vectors* untuk Klasifikasi *Tweet Hate Speech*” yang disusun oleh






Nama : Andhika Cahya Rizqillah

NIM : 4611419064

Prodi : Teknik Informatika

Telah dipertahankan dalam ujian sidang skripsi pada hari selasa, tanggal 19 september tahun 2023.

Tim Penguji

Ketua Penguji Prof. Dr. Edy Cahyono, M.Si. NIP. 196412051990021001	 KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI UNIVERSITAS NEGERI SEMARANG UNNES
Sekretaris Dr. Alamsyah, S.Si., M.Kom. NIP. 197405172006041001	FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM. 
Penguji 1 Anggyi Trisnawan Putra, S.Si., M.Si. NIP. 198707062014041003	
Penguji 2 Riza Arifudin, S.Pd., M.Cs. NIP. 19800525005011001	
Penguji 3/Pembimbing Aji Purwinarko, S.Si., M.Cs. NIP. 198509102015041001	

PERNYATAAN

Skripsi yang ditulis berjudul “Implementasi Algoritma *Bidirectional Long Short-Term Memory* Menggunakan *Global Vectors* untuk Klasifikasi *Tweet Hate Speech*” merupakan karya ilmiah asli dan bukan hasil plagiasi dari karya ilmiah orang lain. Pendapat atau temuan orang lain yang dikutip di dalam Skripsi ini telah ditulis berdasarkan kode etik ilmiah.

Semarang, 15 September 2023

Yang menyatakan,



Andhika Cahya Rizqillah

NIM. 4611419064

MOTTO DAN PERSEMBAHAN

MOTTO

- Rencanakan dengan akurat, lakukan dengan tekad. (Colin Powell)
- Tidak ada jalan pintas menuju keberhasilan. Itu adalah hasil dari perencanaan, kerja keras, dan dedikasi. (Colin Powell)

PERSEMBAHAN

Skripsi ini saya persembahkan kepada:

- Kedua orang tua saya, Bapak Kurnadi (alm) dan Ibu Ani Irawati yang selalu memberi dukungan, binaan, tempaan, kasih sayang, do'a, serta kesabaran dalam mendidik saya sedari kecil hingga saat ini.
- Untuk saudara saya, Raffaeyza Muthi'illah Alfarezqi yang telah memberikan dukungan bagi penulis.
- Diri saya sendiri, yang sudah mau berusaha, belajar, dan berjuang sampai sekarang serta bertahan ditengah banyaknya terpaan dari segala sisi. Semoga dapat menjadikan pribadi yang lebih kuat untuk kehidupan kedepannya.
- Seluruh dosen Universitas Negeri Semarang yang senantiasa memberikan kepada saya bekal ilmu yang tak ternilai. Khususnya Bapak dan Ibu dosen prodi Teknik Informatika yang saya anggap seperti orang tua saya sendiri ketika berada di kampus.
- Seluruh keluarga yang tidak dapat disebutkan satu persatu yang telah memberikan doa dan dukungan moril selama proses penyusunan skripsi.
- Keluarga The Kams, Amara, Fachrizal, Bharaka, dan Ghamal.
- Sahabat-sahabat saya yang selalu menemani hari-hari saya dalam situasi dan kondisi apapun. Semoga tali persahabatan ini selalu terjaga.
- Orang spesial di hati saya yang sudah memberikan semangat dalam penulisan skripsi ini.
- Almamater Universitas Negeri Semarang.

ABSTRAK

Rizqillah, Andhika Cahya. 2023. Implementasi Algoritma *Bidirectional Long Short-Term Memory* Menggunakan *Global Vectors* untuk Klasifikasi *Tweet Hate Speech*. Skripsi, Program Studi Teknik Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang.

Pembimbing Aji Purwinarko, S.Si., M.Cs.

Kata kunci: klasifikasi teks, *hate speech tweet*, *deep learning*, Bi-LSTM, GloVe.

Tweet adalah pesan singkat di Twitter yang mengandung berbagai informasi dan sering digunakan untuk berkomunikasi dan berpartisipasi dalam percakapan online. Namun, seringkali *tweet* disalahgunakan, seperti dengan *hate speech*, yang melibatkan pesan diskriminatif terhadap individu atau kelompok berdasarkan karakteristik tertentu. *Hate speech* juga sering disertai *abusive language*, yaitu penggunaan bahasa kasar atau merendahkan individu atau kelompok dalam pesan. Untuk mengatasi masalah ini, *text mining* dapat digunakan untuk mengklasifikasikan *tweet* dengan lebih efektif dan akurat. Penelitian ini bertujuan untuk menerapkan algoritma *Bidirectional Long Short-Term Memory* (Bi-LSTM) dengan menggunakan pendekatan metode *Global Vectors for Word Representation* (GloVe) dalam upaya mengklasifikasikan *tweet* yang mengandung *hate speech*. Melalui pendekatan ini, kata-kata yang memiliki makna atau konteks yang serupa akan memiliki representasi vektor yang lebih mendekati satu sama lain secara geometris. Penelitian dilakukan dengan menggunakan *dataset Indonesian multi-label hate speech and abusive language detection*. Hasil eksperimen menunjukkan bahwa implementasi Bi-LSTM dengan GloVe mencapai tingkat akurasi sebesar 86,86% dalam mengklasifikasikan *hate speech* pada *tweet* berbahasa Indonesia. Model ini juga mencapai tingkat *precision* sebesar 90,41%, *recall* sebesar 85,33%, dan *f1-score* sebesar 0,8779. Hasil ini mengindikasikan bahwa penggunaan GloVe sebagai word embedding memberikan kontribusi positif dalam meningkatkan kemampuan model dalam mengidentifikasi dan mengklasifikasikan konten *hate speech* dengan lebih baik.

PRAKATA

Puji syukur saya panjatkan atas kehadiran Allah Subhanahu wa ta'ala yang telah memberikan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat melakukan penelitian dan menyelesaikan skripsi yang berjudul “Implementasi Algoritma *Bidirectional Long Short-Term Memory* Menggunakan *Global Vectors* untuk Klasifikasi *Tweet Hate Speech*” dengan lancar. Penyusunan skripsi ini digunakan untuk memenuhi salah satu persyaratan akademik dalam menyelesaikan Program Sarjana (S1) pada Program Studi Teknik Informatika, Program Studi Teknik Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Semarang.

Penulis menyadari bahwa tanpa bimbingan, motivasi, bantuan, dan dukungan dari berbagai pihak maka skripsi ini tidak mungkin dapat terselesaikan seperti pada saat ini. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih kepada seluruh pihak yang berperan dalam proses penyusunan skripsi ini, diantaranya:

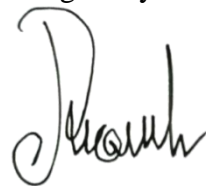
1. Bapak Prof. Dr. S Martono, M.Si., Rektor Universitas Negeri Semarang.
2. Bapak Prof. Dr. Edy Cahyono, M.Si., Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang.
3. Bapak Dr. Alamsyah, S.Si., M.Kom., Koordinator Program Studi Teknik Informatika Universitas Negeri Semarang.
4. Ibu Endang Sugiharti, S.Si., M.Kom. selaku Dosen Wali yang telah memberikan pengarahan selama masa perkuliahan.
5. Bapak Aji Purwinarko, S.Si., M.Cs. selaku Dosen Pembimbing yang telah meluangkan waktu untuk mendampingi dan memberikan arahan penulis dalam proses penyusunan skripsi.
6. Bapak Anggyi Trisnawan Putra, S.Si., M.Si., selaku Dosen Penguji 1 dan Bapak Riza Arifudin, S.Pd., M.Cs., selaku Dosen Penguji 2 yang telah memberikan masukan berupa kritik dan sarannya dalam memperbaiki skripsi ini.
7. Seluruh dosen dan jajaran staf akademik di lingkungan Universitas Negeri Semarang yang telah membantu penulis dalam mengurus dokumen administrasi.

8. Kedua orang tua saya, Bapak Kurnadi (alm) dan Ibu Ani Irawati yang telah memberikan dukungan doa, moril, dan materil yang tiada henti- hentinya.
9. Saudara saya yang telah memberikan semangat dan motivasi bagi penulis dalam menyusun skripsi ini.
10. Sahabat dan kerabat terdekat yang turut memberikan semangat dan dukungan selama proses penyusunan skripsi ini.
11. Teman-teman di Program Studi Teknik Informatika angkatan 2019 Universitas Negeri Semarang yang turut membantu, memberikan dukungan, dan motivasi bagi penulis hingga terselesaikannya skripsi ini.
12. Seluruh pihak yang berperan dalam membantu proses dari awal penyusunan skripsi hingga skripsi ini dapat terselesaikan yang tidak dapat disebutkan satu per satu.

Demikian pengantar yang dapat penulis sampaikan. Segala kelebihan berasal dari Allah Swt. dan banyak kekurangan dari penulis. Oleh karena itu penulis memohon maaf apabila terdapat kekurangan dalam penulisan skripsi ini dan penulis berharap saran serta kritik yang membangun dari pembaca untuk menjadikan skripsi ini lebih baik lagi. Semoga skripsi ini dapat bermanfaat bagi para pembaca. Akhir kata, penulis menyampaikan terima kasih.

Semarang, 15 September 2023

Yang menyatakan,



Andhika Cahya Rizqillah

NIM. 4611419064

DAFTAR ISI

HALAMAN SAMBUL.....	i
PERSETUJUAN PEMBIMBING.....	ii
PENGESAHAN PENGUJI.....	iii
PERNYATAAN.....	iv
MOTTO DAN PERSEMBAHAN	v
ABSTRAK	vi
PRAKATA.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Tujuan Penelitian.....	6
1.4 Batasan Penelitian	6
1.5 Manfaat Penelitian.....	7
1.6 Keaslian Penelitian	7
BAB 2 KAJIAN PUSTAKA.....	10
2.1 Tinjauan Pustaka	10
2.2 Landasan Teori	11
2.2.1 <i>Text Mining</i>	11
2.2.2 <i>Text Classification</i>	13
2.2.3 <i>Supervised Learning</i>	14
2.2.4 <i>Recurrent Neural Network (RNN)</i>	14

2.2.5	<i>Long Short-Term Memory (LSTM)</i>	15
2.2.6	<i>Bidirectional Long Short-Term Memory (Bi-LSTM)</i>	17
2.2.7	<i>Word Embedding</i>	19
2.2.8	<i>Global Vectors for Word Representation (GloVe)</i>	19
2.2.9	<i>Evaluation Model</i>	20
2.2.10	<i>Hate Speech</i>	22
BAB 3 METODE PENELITIAN		24
3.1	Pendekatan dan Desain Penelitian.....	24
3.2	Lokasi Penelitian	24
3.3	Data dan Sumber Data.....	24
3.4	Teknik Pengumpulan Data	25
3.5	Tahapan Penelitian	26
3.5.1	<i>Preprocessing</i>	27
3.5.2	<i>Split data</i>	28
3.5.3	<i>Train Data</i>	29
3.5.4	<i>Test Data</i>	31
3.5.5	<i>Evaluation Model</i>	31
3.6	Teknik Analisis Data	31
BAB 4 HASIL PENELITIAN DAN PEMBAHASAN		33
4.1	Hasil Penelitian.....	33
4.1.1	Hasil Pengumpulan Data.....	33
4.1.2	Hasil Pengolahan Data	35
4.1.3	Hasil <i>Mining</i> Data	41
4.1.4	Hasil Evaluasi Data.....	46
4.1.5	Hasil Implementasi Algoritma	61
4.1.6	Hasil Implementasi Sistem.....	64

4.2 Pembahasan	67
BAB 5 PENUTUP	74
5.1 Simpulan.....	74
5.2 Saran.....	75
DAFTAR PUSTAKA	76
LAMPIRAN.....	94