

**KLASIFIKASI UJARAN KEBENCIAN PADA  
MEDIA SOSIAL MENGGUNAKAN METODE  
*LONG SHORT-TERM MEMORY (LSTM)***

**SKRIPSI**



Disusun oleh :  
**CHODIJATUN NAIMAH**  
**1955201000971**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNOLOGI INFORMASI  
UNIVERSITAS MERDEKA PASURUAN**

**2023**

**KLASIFIKASI UJARAN KEBENCIAN PADA  
MEDIA SOSIAL MENGGUNAKAN METODE  
*LONG SHORT-TERM MEMORY (LSTM)***

**SKRIPSI**

Diajukan sebagai salah satu persyaratan dalam menyelesaikan  
program Sarjana Komputer (S.Kom.)

Oleh :

**CHODIJATUN NAIMAH**

**NPM. 1955201000971**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNOLOGI INFORMASI  
UNIVERSITAS MERDEKA PASURUAN**

**2023**

## PERSETUJUAN PEMBIMBING SKRIPSI

### KLASIFIKASI UJARAN KEBENCIAN PADA MEDIA SOSIAL MENGUNAKAN METODE *LONG SHORT-TERM MEMORY* (LSTM)

**Nama** : Chodijatun Naimah

**NPM** : 1955201000971

**Program Studi** : Informatika

Pasuruan, 18 Juli 2023

Telah diperiksa dan disetujui oleh:

**Pembimbing I**



(Dian Ahkam Sani, S.Kom., M.Kom)

NIDN. 0724079202

**Pembimbing II**



(Mohammad Zoqi Sarwani, S.Pd., M.Kom)

NIDN. 0716049101

## PERSETUJUAN DAN PENGESAHAN SKRIPSI

Skripsi oleh Chodijatun Naimah ini  
telah dipertahankan di depan dewan penguji  
pada tanggal 18 Juli 2023

Dewan Penguji

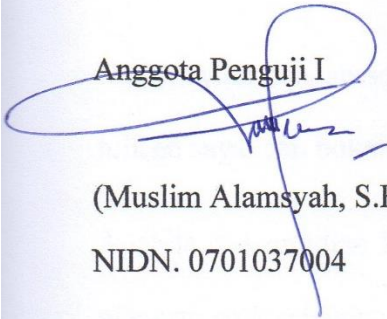
Ketua Penguji



(Nanda Martyan Anggadimas, S.T., M.T.)

NIDN. 0716038604

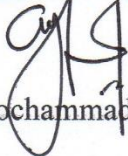
Anggota Penguji I



(Muslim Alamsyah, S.Kom., M.Kom.)

NIDN. 0701037004

Anggota Penguji II



(Mochammad Firman Arif, S.Kom., M.Kom.)

NIDN. 0704029201

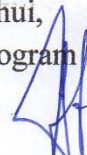
Mengesahkan,  
Dekan Fakultas Teknologi Informatika



(M. Misdrum, S.Kom., M.Kom.)

NIDN. 0717046704

Mengetahui,  
Ketua Program Studi Informatika



(Dian Ahkam Sani, S.Kom., M.Kom.)

NIDN. 0724079202

## PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan di bawah ini:

Nama : Chodijatun Naimah

NPM : 1955201000971

Program Studi : Informatika

Fakultas : Teknologi Informasi

Menyatakan dengan sesungguhnya bahwa **skripsi** yang saya tulis ini benar-benar tulisan saya, dan bukan merupakan plagiasi baik sebagian atau seluruhnya.

Apabila di kemudian hari terbukti atau dapat dibuktikan **bahwa skripsi** ini hasil plagiasi, baik sebagian atau seluruhnya, maka saya bersedia menerima sanksi atas perbuatan tersebut sesuai dengan ketentuan yang berlaku.

Pasuruan, 18 Juli 2023

Yang membuat pernyataan



Chodijatun Naimah  
NPM. 1955201000971

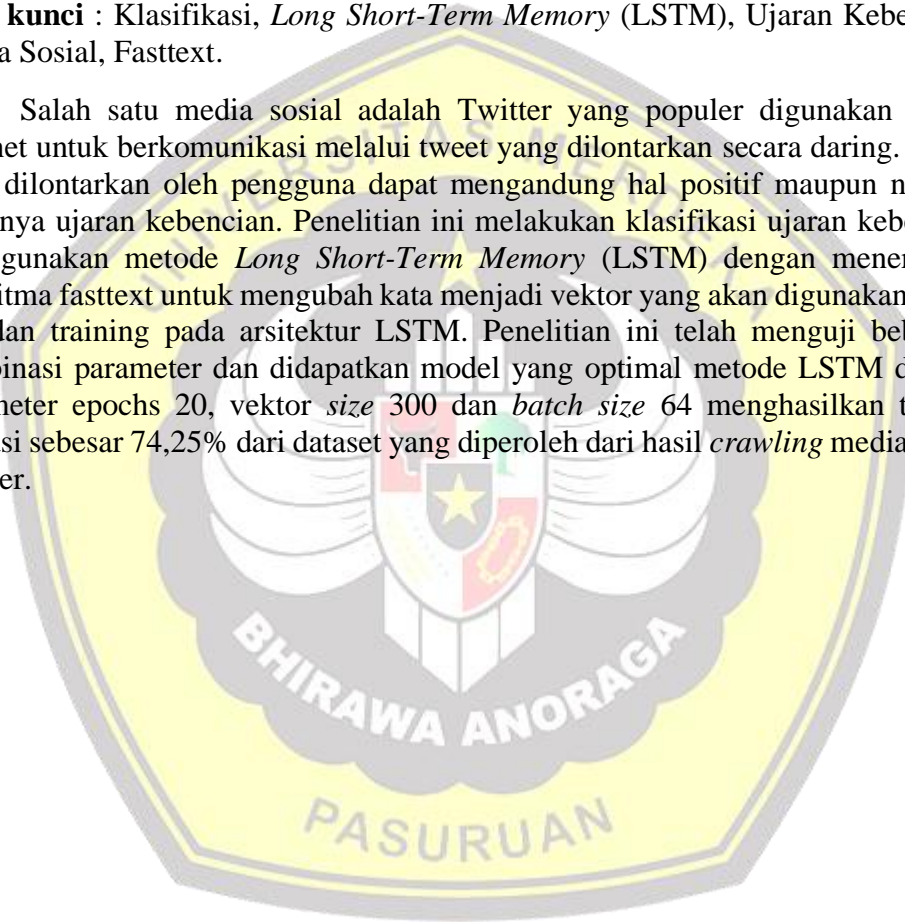
## ABSTRAK

Naimah, Chodijatun. 2023. *Klasifikasi Ujaran Kebencian Pada Media Sosial Menggunakan Metode Long Short-Term Memory (LSTM)*, Program Studi Informatika, Fakultas Teknologi Informasi Universitas Merdeka Pasuruan.

Pembimbing : (I) Dian Ahkam Sani, S.Kom., M.Kom. (II) Mohammad Zoqi Sarwani, S.Pd., M.Kom.

**Kata kunci** : Klasifikasi, *Long Short-Term Memory (LSTM)*, Ujaran Kebencian, Media Sosial, Fasttext.

Salah satu media sosial adalah Twitter yang populer digunakan warga internet untuk berkomunikasi melalui tweet yang dilontarkan secara daring. Tweet yang dilontarkan oleh pengguna dapat mengandung hal positif maupun negatif, misalnya ujaran kebencian. Penelitian ini melakukan klasifikasi ujaran kebencian menggunakan metode *Long Short-Term Memory (LSTM)* dengan menerapkan algoritma fasttext untuk mengubah kata menjadi vektor yang akan digunakan untuk test dan training pada arsitektur LSTM. Penelitian ini telah menguji beberapa kombinasi parameter dan didapatkan model yang optimal metode LSTM dengan parameter epochs 20, vektor *size* 300 dan *batch size* 64 menghasilkan tingkat akurasi sebesar 74,25% dari dataset yang diperoleh dari hasil *crawling* media sosial Twitter.



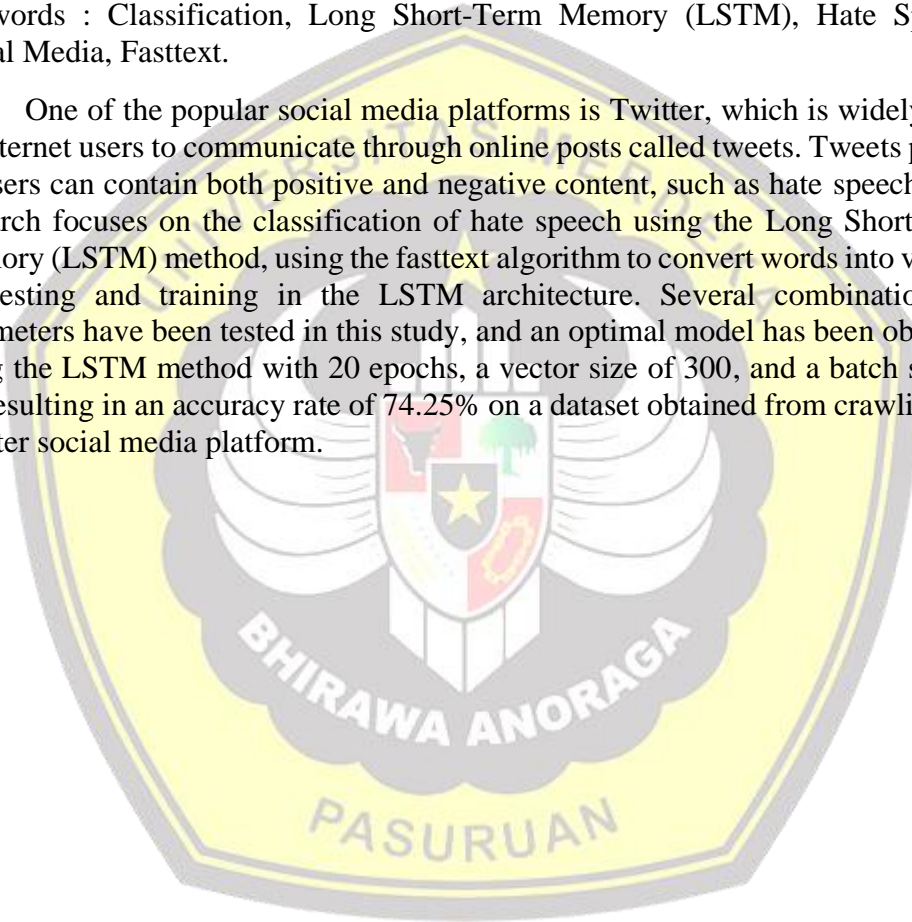
## ABSTRACT

Naimah, Chodijatun. 2023. *Classification of Hate Speech on Social Media using Long Short-Term Memory (LSTM) Method*, Informatics Study Program, Faculty of Information Technology, Merdeka University, Pasuruan.

Advisor : (I) Dian Ahkam Sani, S.Kom., M.Kom. (II) Mohammad Zoqi Sarwani, S.Pd., M.Kom.

Keywords : Classification, Long Short-Term Memory (LSTM), Hate Speech, Social Media, Fasttext.

One of the popular social media platforms is Twitter, which is widely used by internet users to communicate through online posts called tweets. Tweets posted by users can contain both positive and negative content, such as hate speech. This research focuses on the classification of hate speech using the Long Short-Term Memory (LSTM) method, using the fasttext algorithm to convert words into vectors for testing and training in the LSTM architecture. Several combinations of parameters have been tested in this study, and an optimal model has been obtained using the LSTM method with 20 epochs, a vector size of 300, and a batch size of 64, resulting in an accuracy rate of 74.25% on a dataset obtained from crawling the Twitter social media platform.



## KATA PENGANTAR

Bismillahirrohmaanirrokhiiim.

Dengan mengucapkan puji syukur Alhamdulillah penulis panjatkan kehadiran Allah SWT yang telah memberikan rahmat dan karunia-Nya sehingga penulisan Skripsi dengan judul “**KLASIFIKASI UJARAN KEBENCIAN PADA SOSIAL MEDIA MENGGUNAKAN METODE *LONG SHORT-TERM MEMORY (LSTM)***” dapat terselesaikan dengan baik.

Perjalanan panjang telah penulis lalui dalam rangka perampungan penulisan skripsi ini. Banyak hambatan yang dihadapi dalam penyusunannya, namun berkat kehendak dan ridho-Nya penulis berhasil menyelesaikan penyusunan skripsi ini. Karena sesungguhnya dibalik kesulitan pasti ada kemudahan. Skripsi ini ditujukan untuk memenuhi salah satu persyaratan ujian guna memperoleh gelar Sarjana Komputer (S.Kom) pada Program Studi Informatika Fakultas Teknologi Informasi di Universitas Merdeka Pasuruan.

Penulis sepenuhnya menyadari bahwa skripsi ini tidak akan terwujud tanpa adanya bantuan, bimbingan, dan dukungan tersebut penulis mengucapkan rasa terima kasih dan menyampaikan penghargaan setinggi-tingginya kepada :

1. Dr. Ir. Sulistyawati, M.P., selaku Rektor Universitas Merdeka Pasuruan.
2. Muhammad Misdrum, S.Kom., M.Kom., selaku Dekan Fakultas Teknologi Informasi.
3. Dian Ahkam Sani, S.Kom., M.Kom., selaku Ketua Program Studi Informatika.



4. Dian Ahkam Sani, S.Kom., M.Kom., selaku Dosen Pembimbing I dan Mohammad Zoqi Sarwani, S.Pd., M.Kom., selaku Dosen Pembimbing II serta segenap Dosen Universitas Merdeka Pasuruan.
5. Orang tua beserta saudara saya yang selalu memberikan doa agar skripsi saya cepat selesai.
6. Sahabat-sahabat saya yang selalu memberikan dukungan dan motivasi untuk menyelesaikan skripsi ini.
7. Teman-teman FTI Angkatan 2019 yang selalu memberi semangat kepada saya untuk menyelesaikan skripsi ini.
8. Seluruh pihak yang membantu hingga tersusunnya skripsi ini.
9. *Last but not least, I want to thank myself for staying strong no matter what happens.*

Demikian laporan ini disusun, penulis menyadari bahwa masih banyak kekurangan dan jauh dari kata sempurna dalam penyusunan skripsi ini. Penulis berharap laporan skripsi ini dapat bermanfaat bagi semua pihak.

Pasuruan, 18 Juli 2023

Penulis

Chodijatun Naimah

## DAFTAR ISI

<b>ABSTRAK .....</b>	<b>i</b>
<b>KATA PENGANTAR.....</b>	<b>iii</b>
<b>DAFTAR ISI.....</b>	<b>v</b>
<b>DAFTAR TABEL.....</b>	<b>vii</b>
<b>DAFTAR GAMBAR.....</b>	<b>viii</b>
<b>LAMPIRAN.....</b>	<b>ix</b>
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
1.6 Sistematika Penulisan.....	5
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Penelitian terdahulu.....	7
2.2 Teori Dasar yang Digunakan.....	10
<b>BAB III METODOLOGI PENELITIAN</b>	
3.1 Alur Penelitian.....	22
3.2 Studi Literatur.....	23
3.3 Perancangan Sistem.....	25

**BAB IV HASIL DAN PEMBAHASAN**

4.1 Implementasi Sistem.....35

4.2 Pelatihan Model LSTM .....43

4.3 Hasil .....44

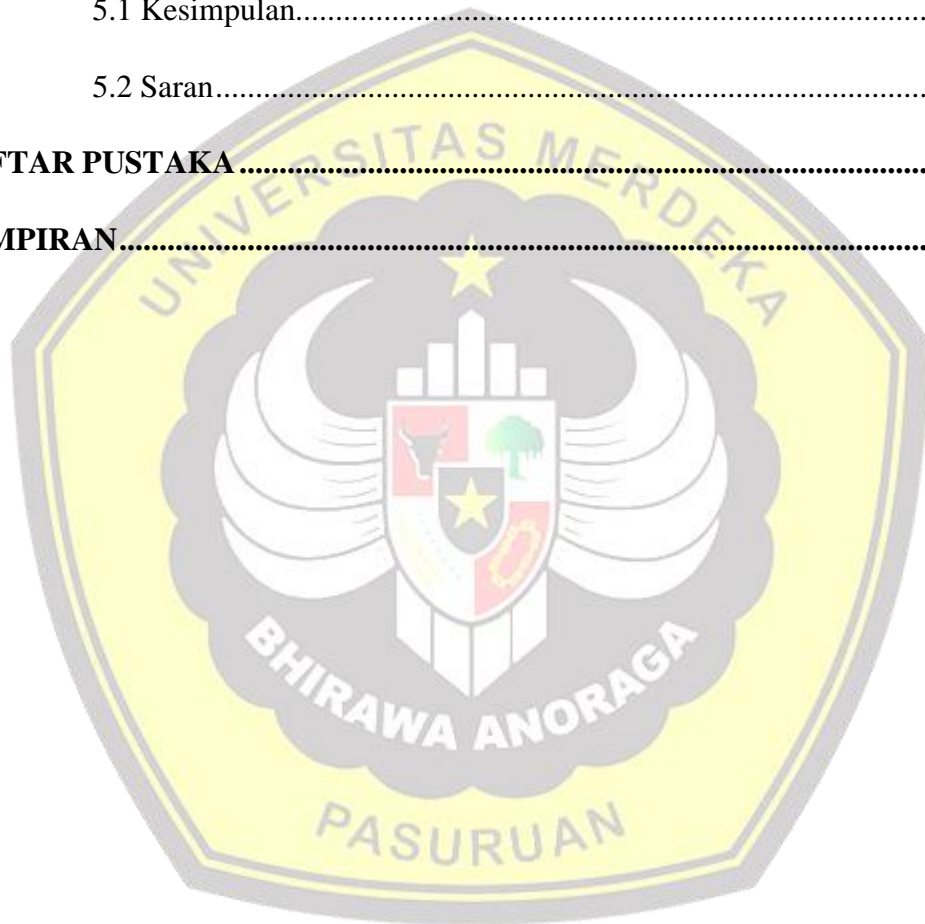
**BAB V PENUTUP**

5.1 Kesimpulan.....49

5.2 Saran.....49

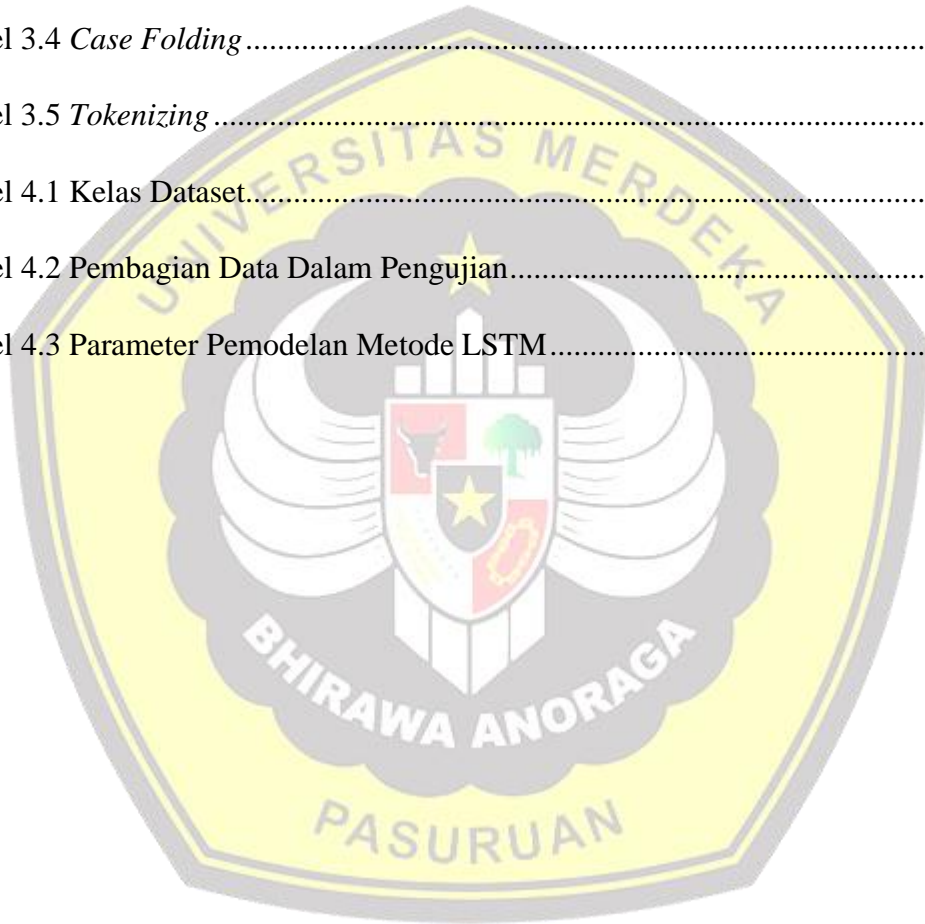
**DAFTAR PUSTAKA.....50**

**LAMPIRAN..... ix**



## DAFTAR TABEL

Tabel 2.1 <i>Flowchart</i> .....	17
Tabel 3.1 Spesifikasi <i>Hardware</i> .....	24
Tabel 3.2 Dataset.....	26
Tabel 3.3 <i>Data Cleaning</i> .....	28
Tabel 3.4 <i>Case Folding</i> .....	29
Tabel 3.5 <i>Tokenizing</i> .....	30
Tabel 4.1 Kelas Dataset.....	35
Tabel 4.2 Pembagian Data Dalam Pengujian.....	36
Tabel 4.3 Parameter Pemodelan Metode LSTM.....	41



## DAFTAR GAMBAR

Gambar 2.1 Arsitektur LSTM.....	12
Gambar 2.2 Arsitektur Sigmoid.....	12
Gambar 2.3 Gerbang Sigmoid.....	13
Gambar 2.4 Tabel <i>Confusion Matrix</i> .....	20
Gambar 3.1 Tahapan Penelitian .....	22
Gambar 3.2 <i>Flowchart</i> Sistem <i>Traning</i> .....	27
Gambar 3.3 <i>Flowchart</i> Sistem <i>Testing</i> .....	27
Gambar 3.4 <i>Flowchart</i> SkipGram dan CBOW Fasttext .....	32
Gambar 3.5 <i>Flowchart</i> LSTM.....	32
Gambar 4.1 <i>WordCloud</i> Dataset.....	35
Gambar 4.2 <i>Flowchart</i> Fasttext.....	37
Gambar 4.3 Parameter Model Fasttext.....	37
Gambar 4.4 Contoh Nilai Vektor Kata ‘Pemilu’ .....	39
Gambar 4.5 Contoh Nilai Vektor Kata ‘Menang’ .....	40
Gambar 4.6 Arsitektur Pemodelan LSTM.....	41
Gambar 4.7 Pelatihan Model LSTM.....	43
Gambar 4.8 Nilai Akurasi .....	44
Gambar 4.9 Hasil Kinerja Metode LSTM Dalam <i>Confusion Matrix</i> .....	44
Gambar 4.10 Hasil Nilai Akurasi Dalam <i>Classification Report</i> .....	46