



**ANALISIS SENTIMEN MULTI-CLASS PADA SOSIAL MEDIA  
MENGUNAKAN METODE LONG SHORT-TERM MEMORY (LSTM)**

*TUGAS AKHIR*

Yuliya Astari  
41516120007

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA**

**UNIVERSITAS JAKARTA**

2021

**MERCU BUANA**



**ANALISIS SENTIMEN MULTI-CLASS PADA SOSIAL MEDIA  
MENGUNAKAN METODE LONG SHORT-TERM MEMORY (LSTM)**

*Tugas Akhir*

Diajukan Untuk Melengkapi Salah Satu Syarat  
Memperoleh Gelar Sarjana Komputer

Oleh:  
Yuliy Astari  
41516120007

PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA  
JAKARTA  
2021

UNIVERSITAS  
**MERCU BUANA**

## LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41516120007

Nama : Yuliya Astari

Judul Tugas Akhir : Analisis Sentimen *Multi-Class* Pada Sosial Media  
Menggunakan Metode *Long Short-Term Memory* (LSTM)

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

Jakarta, 25 Februari 2021



Yuliya Astari



UNIVERSITAS  
MERCU BUANA



## SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Yuliya Astari  
NIM : 41516120007  
Judul Tugas Akhir : Analisis Sentimen *Multi-Class* Pada Sosial Media Menggunakan Metode *Long Short-Term Memory* (LSTM)

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS Jakarta, 25 Februari 2021  
MERCU BUANA



Yuliya Astari



## SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Yuliya Astari  
NIM : 41516120007  
Judul Tugas Akhir : Analisis Sentimen *Multi-Class* Pada Sosial Media Menggunakan Metode *Long Short-Term Memory* (LSTM)

Menyatakan bahwa :

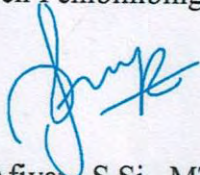
1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis		Status	
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi		Diajukan	✓
		Jurnal Nasional Terakreditasi	✓		
		Jurnal International Tidak Bereputasi		Diterima	
		Jurnal International Bereputasi			
Disubmit/dipublikasikan di :	Nama Jurnal	: Jurnal Linguistik Komputasional			
	ISSN	: 2621-9336			
	Link Jurnal	: <a href="http://inacl.id/journal/index.php/jlk/issue/view/6">http://inacl.id/journal/index.php/jlk/issue/view/6</a>			
	Link File Jurnal Jika Sudah di Publish				

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan

Demikian pernyataan ini saya buat dengan sebenarnya.

Mengetahui  
Dosen Pembimbing TA

  
Afyati, S.Si., MT

Jakarta, 25 Februari 2021



Yuliya Astari

## LEMBAR PERSETUJUAN

NIM : 41516120007  
Nama : Yuliya Astari  
Judul Tugas Akhir : Analisis Sentimen *Multi-Class* Pada Sosial Media  
Menggunakan Metode *Long Short-Term Memory*  
(LSTM)

Tugas Akhir ini telah diperiksa dan disetujui

Jakarta, 7 Januari 2021,

Menyetujui,



(Afiyati, S.Si., MT)  
Dosen Pembimbing


UNIVERSITAS  
MERCU BUANA

## LEMBAR PERSETUJUAN PENGUJI

NIM : 41516120007  
Nama : Yuliya Astari  
Judul Tugas Akhir : Analisis Sentimen *Multi-Class* Pada Sosial Media Menggunakan Metode *Long Short-Term Memory* (LSTM)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 8 februari 2021



(Dr. Leonard Goeirmanto, ST, M.Sc)




## LEMBAR PERSETUJUAN PENGUJI

NIM : 41516120007  
Nama : Yuliya Astari  
Judul Tugas Akhir : Analisis Sentimen *Multi-Class* Pada Sosial Media Menggunakan Metode *Long Short-Term Memory* (LSTM)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 8 februari 2021



(Anis Cherid SE, MTI)

UNIVERSITAS  
MERCU BUANA



## LEMBAR PERSETUJUAN PENGUJI

NIM : 41516120007  
Nama : Yuliya Astari  
Judul Tugas Akhir : Analisis Sentimen *Multi-Class* Pada Sosial Media  
Menggunakan Metode *Long Short-Term Memory*  
(LSTM)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 8 Februari 2021



(Desi Ramayanti, S.Kom, MT)

## LEMBAR PENGESAHAN

NIM : 41516120007  
Nama : Yuliya Astari  
Judul Tugas Akhir : Analisis Sentimen *Multi-Class* Pada Sosial Media  
Menggunakan Metode *Long Short-Term Memory* (LSTM)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 8 Februari 2021

Menyetujui,



(Afiyati, S.Si., MT)  
Dosen Pembimbing

Mengetahui,



(Diky Firdaus, S.Kom, MM)

Koord. Tugas Akhir Teknik Informatika



(Desi Ramayanti, S.Kom, MT)

Ka. Prodi Teknik Informatika



## ABSTRAK

Nama : Yuliy Astari  
NIM : 41516120007  
Pembimbing TA : Afiyati, S.Si., MT  
Judul : Analisis Sentimen *Multi-Class* Pada Sosial Media Menggunakan Metode *Long Short-Term Memory* (LSTM)

Perkembangan teknologi khususnya dalam internet dan media sosial dapat menjadikan subjek penelitian yang sangat penting dalam memperoleh informasi, karena banyaknya jumlah informasi di dalam sebuah teks yang terdapat pada sosial media. Dalam beberapa tahun terakhir, telah terjadi peningkatan penelitian terhadap analisis sentimen pada teks ulasan maupun cuitan agar dapat mengetahui polaritas yang dihasilkan dari sosial media. Masih sedikit penelitian yang menerapkan metode *deep learning* dengan algoritma *Long Short-Term Memory* (LSTM) untuk menganalisa sentimen *multiclass* pada teks berbahasa Indonesia. Penelitian ini bertujuan untuk menganalisa emosi bersifat positif maupun negatif pada teks sosial media menggunakan pendekatan klasifikasi informasi yang ada pada teks dan membaginya menjadi 8 kelas yang berbeda dengan menggunakan metode LSTM. *Dataset* langsung diambil dan dikumpulkan dari tulisan pengguna pada media sosial. Dalam pengujian metodel LSTM ini dihasilkan perhitungan dari nilai akurasi, *exactness*, *review*, *f-measure*. Pada hasil pengolahan metode LSTM menunjukkan cukup baik dengan uji coba sebanyak 5 kali dengan nilai akurasi tertinggi 91,9% dan nilai rata-rata dari *multiclass* mendapatkan hasil 89,45%.

Kata kunci:

analisis sentimen, klasifikasi, long short-term memory, sosial media, multi-class

## ABSTRACT

Name : Yuliyi Astari  
Student Number : 41516120007  
Counsellor : Afiyati, S.Si., MT  
Title : *Multi-Class Sentiment Analysis on Social Media  
Using the Long Short-Term Memory (LSTM)  
Method*

*Technological developments, especially in the internet and social media, could be a very important research subject in obtaining information, because of the large amount of information in a text found on social media. In recent years, there has been an increase in research about sentiment analysis on text reviews and tweets in order to determine the polarity generated by social media. There are still few studies that apply the deep learning method with the Long Short-Term Memory (LSTM) algorithm to analyze multiclass sentiments in Indonesian-language texts. This study aims to analyze positive and negative emotions in social media texts using the information classification approach in the text and dividing them into 8 different classes using the LSTM method. The dataset is directly taken and collected from users' posts on social media. In testing the LSTM method, the calculation of the accuracy, exactness, review, f-measure values is generated. The results of the processing of the LSTM method show quite well with 5 trials with the highest accuracy value of 91.9% and the average value of multiclass getting 89.45% results.*

*Key words:*

*sentiment analysis, classification, long short-term memory, social media, Multi-Class*



## KATA PENGANTAR

Puji dan syukur dipanjatkan Kepada Allah SWT yang telah melimpahkan rahmat, hidayah dan karunia-Nya, shalawat dan salam semoga senantiasa tercurahkan kepada Nabi Muhammad SAW. Sehingga penulis dapat menyelesaikan tugas akhir yang berjudul “Analisis Sentimen *Multi-Class* pada Sosial Media Menggunakan Metode *Long Short-Term Memory (LSTM)*” tepat pada waktunya. Tugas akhir ini disusun untuk memenuhi salah satu syarat memperoleh gelar sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Dalam penyusunan tugas akhir ini penulis mendapatkan banyak bantuan dan bimbingan dari berbagai pihak. Untuk itu disampaikan rasa terima kasih kepada:

1. Ibu Afiyati, S.Si., MT, selaku dosen pembimbing tugas akhir yang telah meluangkan waktunya untuk memberikan arahan dan bimbingan dalam menyusun tugas akhir ini hingga selesai.
2. Ibu Desi Ramayanti, S.Kom., MT, selaku Ka. Prodi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.
3. Bapak Diky Firdaus, S.Kom., MM, selaku Koordinator Tugas Akhir Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.
4. Kedua Orang Tua yang selalu memberikan dukungan dan doa kepada penulis.
5. Teman-teman yang selalu menyemangati dan memberi motivasi kepada penulis selama pelaksanaan tugas akhir ini.
6. Pihak – Pihak yang tidak dapat saya sebutka satu persatu, atas bantuan dalam penyusunan tugas akhir ini.

Akhir kata, penulis berharap tugas akhir ini dapat bermanfaat bagi pembaca guna menambah pengetahuan dan wawasan.

Jakarta, 19 Februari 2021

Yuliya Astari

Penulis

xii

## DAFTAR ISI

HALAMAN SAMPUL.....	i
HALAMAN JUDUL .....	i
LEMBAR PERNYATAAN ORISINALITAS .....	ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR... iii	
SURAT PERNYATAAN LUARAN TUGAS AKHIR.....	iv
LEMBAR PERSETUJUAN .....	v
LEMBAR PERSETUJUAN PENGUJI .....	vi
LEMBAR PENGESAHAN .....	ix
ABSTRAK .....	x
ABSTRACT.....	xi
KATA PENGANTAR.....	xii
DAFTAR ISI.....	xiii
NASKAH JURNAL .....	1
KERTAS KERJA.....	6
BAB 1. LITERATUR REVIEW .....	7
BAB 2. SOURCE CODE .....	10
BAB 3. DATASET.....	20
BAB 4. TAHAPAN EKSPERIMEN.....	26
BAB 5. HASIL SEMUA EKSPERIMEN.....	29
BAB 6. DAFTAR PUSTAKA.....	32
LAMPIRAN DOKUMEN HAKI.....	33
LAMPIRAN SCAN KTP.....	35
LAMPIRAN KORESPONDENSI .....	36
LAMPIRAN CURRICULUM VITAE.....	37



## NASKAH JURNAL



Jurnal  
Linguistik Komputasional

JLK  
Vol 3, No 2  
Januari  
2021

# Analisis Sentimen Multi-Class pada Sosial Media menggunakan metode *Long Short-Term Memory* (LSTM)

Yuliy Astari<sup>#1</sup>, Afiyati<sup>\*2</sup>, Saddam Wahib Rozaqi<sup>#3</sup>

<sup>#\*\*</sup>Teknik Informatika, Ilmu Komputer, Universitas Mercu Buana

Jl. Raya Meruya Selatan, Kembangan, Jakarta, 11650

<sup>1</sup>41516120007@student.mercubuana.ac.id

<sup>2</sup>afiyati.reno@mercubuana.ac.id

<sup>3</sup>41516120084@student.mercubuana.ac.id

**Abstrak**—Perkembangan teknologi khususnya dalam internet dan media sosial dapat menjadikan subjek penelitian yang sangat penting dalam memperoleh informasi, karena banyaknya jumlah informasi di dalam sebuah teks yang terdapat pada sosial media. Dalam beberapa tahun terakhir, telah terjadi peningkatan penelitian terhadap analisis sentimen pada teks ulasan maupun cuitan agar dapat mengetahui polaritas yang dihasilkan dari sosial media. Masih sedikit penelitian yang menerapkan metode deep learning dengan algoritma *Long Short-Term Memory* (LSTM) untuk menganalisa sentimen multiclass pada teks berbahasa Indonesia. Penelitian ini bertujuan untuk menganalisa emosi bersifat positif maupun negatif pada teks sosial media menggunakan pendekatan klasifikasi informasi yang ada pada teks dan membaginya menjadi 8 kelas yang berbeda dengan menggunakan metode LSTM. Dataset langsung diambil dan dikumpulkan dari tulisan pengguna pada media sosial. Dalam pengujian metode LSTM ini dihasilkan perhitungan dari nilai akurasi, exactness, review, f-measure. Pada hasil pengolahan metode LSTM menunjukan cukup baik dengan uji coba sebanyak 5 kali dengan nilai akurasi tertinggi 91,9% dan nilai rata-rata dari multiclass mendapatkan hasil 89,45%.

**Kata kunci** — *analisa sentimen, teks, klasifikasi, Long Short-Term Memory, sosial media, multiclass*

## I. PENDAHULUAN

Pada era digital seperti ini kebanyakan orang melakukan kegiatan komunikasi melalui media sosial [1]. Para pengguna media sosial dapat menuliskan opini, pendapat atau pengungkapan perasaan yang biasanya diungkapkan secara verbal [2]. Penelitian tersebut juga menyatakan bahwa setiap tulisan atau kicauan seseorang pada media sosial seringkali juga berisi tentang informasi ataupun kondisi emosi seseorang dalam menulis sebuah cuitan [3]. Bahkan ketika seseorang yang tidak secara khusus menuliskan tentang status emosi pribadi mereka, tetap saja tulisan pada media sosial dapat mencerminkan suasana hati penulis. Dengan demikian setiap

postingan dapat dianggap sebagai tempat menyalurkan suasana hati atau emosi yang instan [2].

Penelitian ini memfokuskan perhatian pada setiap sentimen, perasaan atau emosi. Sentimen bisa digambarkan sebagai emosi, keputusan, pendapat atau ide-ide sebagai sumber informasi utama. Informasi berupa teks ini digunakan untuk menganalisa konsep ekspresi sentimen dan menghitung klasifikasinya [3]. Untuk mengetahui sifat emosi itu sendiri diperlukan sebuah metode pada machine learning yang dapat digunakan untuk mendeteksi ataupun klasifikasi sentiment [4]. Pada perinsipnya analisis sentimen merupakan bagian dari Natural Language Processing (NLP) dan Machine Learning [5]. Cara kerjanya adalah dengan melakukan klasifikasi terhadap kata positif, negatif dan netral. Sebuah penelitian telah menganalisa pendapat seseorang, jenis sentimen, evaluasi, penilaian, sikap, dan emosi terhadap entitas produk, jasa, organisasi, individu, masalah, peristiwa atau suatu topik [3].

Sebelumnya ada banyak penelitian mengenai analisis sentimen juga sudah pernah dilakukan yaitu untuk melakukan review novel menggunakan metode *Long Short-Term Memory* (LSTM). Berdasarkan hasil pengujian yang dilakukan memperlihatkan bahwa metode LSTM memiliki hasil akurasi 72.85%, presisi 73%, recall 72%, dan f-measure 72%. Penelitian tersebut membandingkan LSTM dengan hasil akurasi dari metode Naïve Bayes dengan nilai akurasi 67,88%, presisi 69%, recall 68%, dan f-measure 68%. Penelitian ini menunjukkan performansi metode LSTM lebih baik daripada metode Naïve Bayes [6]. Pada penelitian selanjutnya yaitu melakukan klasifikasi teks multilebel pada artikel berita menggunakan metode *Long Short-Term Memory* (LSTM), mendapatkan hasil akurasi tertinggi model kelima dengan 95,38% sedangkan nilai rata-rata presisi, recall, dan F1-score 95% [7]. dan yang terakhir penelitian yang terkait adalah Multiclass deteksi beban mental menggunakan *Long Short-Term Memory* dengan mendapatkan akurasi rata-rata convolutional neural network (CNN) and LSTM 87,45% dan 89,31% masing-masing [8].

Oleh karena itu penulis melakukan penelitian untuk menganalisa emosi yang ditimbulkan dalam bentuk positif maupun negatif pada penggunaan sosial media, dengan pendekatan klasifikasi data ke dalam 8 kelas yang berbeda. Dengan menggunakan metode *Long Short-Term Memory*, dan implementasi dengan menganalisis akurasi klasifikasi dan presisi. Pengolahan data menggunakan bahasa pemrograman Python. Penelitian ini diharapkan dapat menjadikan penelitian lanjutan untuk membantu bidang ilmu psikologi dengan membuat aplikasi antar muka yang menganalisa emosi dari setiap kalimat pengguna pada media sosial.

## II. METODOLOGI PENELITIAN

Penelitian ini dilakukan dengan mengumpulkan sumber bacaan dari penelitian-penelitian lain dengan urutan pekerjaan dapat dilihat pada Gambar 1.



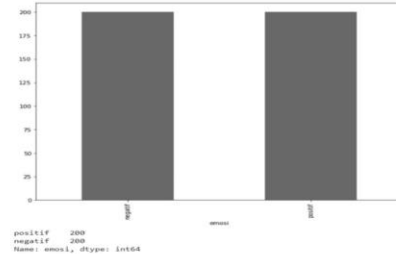
Gambar.1 Proses Metodologi Penelitian

### A. Studi Kasus

Pada penelitian ini mengambil studi kasus yang bersumber dari bidang ilmu komputer yang mengambil metode yang akan digunakan dan bidang ilmu psikologi untuk mengetahui emosi yang ada pada data tulisan pengguna sosial media. Penelitian-penelitian yang dikumpulkan adalah penelitian yang memfokuskan pada klasifikasi sentimen emosi positif dan negatif serta mengklasifikasikan data tulisan pengguna sosial media ke dalam 8 kelas sentimen yang berbeda, dan menghasilkan model dari metode *Long Short-Term Memory* (LSTM) yang paling optimal.

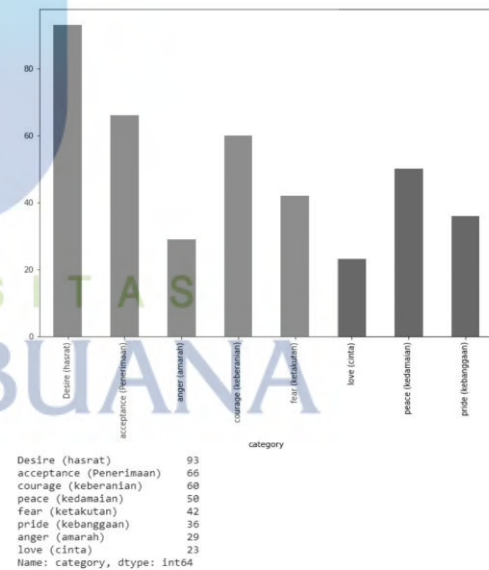
### B. Analisa Data

Dari data tulisan sosial media yang dapat dikumpulkan berjumlah total 400 tulisan yang selanjutnya data tulisan ini diberikan pelabelan secara manual dengan pembagian beris 200 kelas positif dan 200 kelas negatif dapat dilihat pada Gambar 2. Jumlah data yang sama ini diharapkan tidak memberikan masalah pada keseimbangan jumlah data.



Gambar.2 Diagram Batang Jumlah Data

Selanjutnya data positif dan negatif kemudian akan diklasifikasikan menjadi 8 kelas kategori yaitu: (1) Kelas positif *courage* (keberanian), *acceptance* (penerimaan), *love* (cinta) dan *peace* (kedamaian); (2) Kelas negatif *pride* (kebanggaan), *anger* (amarah), *desire* (hasrat) dan *fear* (ketakutan). Jumlah kelas emosi yang terdiri dari 8 kelas dapat dilihat perbandingan jumlah datanya pada Gambar 3. Jumlah data kelas emosi yang tidak sama ini tidak akan memberikan masalah pada keseimbangan jumlah data karena pada penelitian ini hanya ditujukan untuk menemukan akurasi sentimen yang ada pada dataset yang ada.



Gambar.3 Diagram Batang Jumlah Emosi

Pada Tabel 1 dapat dilihat contoh data tulisan dari media sosial yang telah diberi label kelas positif maupun negatif.

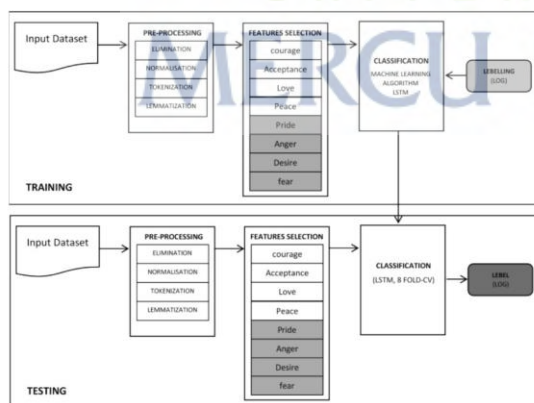


TABEL I  
DATASET CUITAN SOSIAL MEDIA

No.	Teks	Kategori	Emosi
1.	Perdebatan sengit dalam hati, antara mempertahankan EGO The Wachter, dan nalar, seperti saling tarik dan saling menguasai.	acceptance (penerimaan)	Positif
2.	Ikut kuliah transendensi semalam ikut kena tepuk, kerasa sakit dan klepek2, .	acceptance (penerimaan)	Positif
3.	Spiderman aja berani jualan pete. Laku pula. Ahhh. . kreatif itu memang harus out the box atau create the new bo	courage (keberanian)	Positif
...	...	...	...
400.	Saya pernah nggak di rumah, temen2 main ke rumah, ditemani anak kecil yang mirip saya, bahkan sebaju2 yang saya pakai hari itu mirip	fear (ketakutan)	Negatif

### C. Perancangan

Pekerjaan yang dilakukan pada penelitian ini adalah klasifikasi multi emosi dengan menggunakan metode LSTM. Gambar 4 menunjukkan diagram tahapan-tahapan perancangan penelitian untuk mengklasifikasi data yang telah dikumpulkan.



Gambar.4 Diagram Perancangan Penelitian

### III. IMPLEMENTASI

Penelitian ini menggunakan bahasa pemrograman Python untuk mengolah data, dengan mengikuti tahapan-tahapan berikut ini.

#### A. Pelabelan Data

Didalam klasifikasi hal yang utama yang harus dilakukan adalah pelabelan pada data yang benar dan akurat. Pada penelitian ini dilakukan pelabelan secara manual untuk 8 kelas yang berbeda-beda.

#### B. Data Pre-Processing

Data pre-processing adalah sebuah tahapan yang sangat penting dimana dalam melakukan proses klasifikasi data teks. Tujuan dilakukannya *data pre-processing* yaitu untuk menghilangkan *noise* atau data yang tidak dibutuhkan seperti emoji, url, angka dan untuk menyeragamkan bentuk data kata sehingga data kata dapat menjadi lebih bersih sebelum diolah pada proses selanjutnya [9].

Dari data yang didapatkan masih banyak ditemukan kata-kata yang tidak berstruktur sehingga dibutuhkan pemerosesan atau pembersihan data yang dilakukan mencakup tahapan-tahapan sebagai berikut:

- (1) mengubah semua huruf kapital menjadi huruf kecil;
- (2) menghapus data yang terganakan;
- (3) menghilangkan symbol, angka, string ASCII, dan tanda baca seperti seperti [ “TM “, “© “, “SM “ ];
- (4) memproses transformasi untuk membentuk normalisasi satu kata atau proses *lemmatization*;
- (5) tokenisasi membentuk kata-kata, frase atau elemen bermakna lainnya.

#### C. Feature Selection

*Highlight Selection* atau *Feature Reduction* sebuah tahapan yang umum dilakukan pada preprocessing, bertujuan dapat memilih include yang sangat berpengaruh dan mengesampingkan sebuah garis besar yang tidak berpengaruh dalam melakukan suatu pemodelan atau penganalisaan suatu informasi.

#### D. Long Short-Term Memory (LSTM)

*Long Short-Term Memory* (LSTM) diperkenalkan pertamakali pada tahun 1997 yang dikenalkan oleh Hochreiter dan Schmidhuber [10]. LSTM adalah sebuah jenis dari Recurrent Neural Network (RNN). LSTM sendiri dapat menemukan lapisan tersembunyi dari setiap sel dan dirancang untuk menyimpan informasi sel sebelumnya. Metode LSTM digunakan dengan cara mengklasifikasi data secara jangka panjang dengan menyimpan pada sel memori.

Hingga penelitian ini telah dilakukan banyak dilakukan oleh para peneliti agar dapat mengembangkan Metode LSTM [11], pada metode LSTM sendiri memiliki empat komponen utama yaitu : *Input Gate* , koneksi berulang, *forget gate* dan *output gate* [12].

Berikut ini tahapan awal dalam membangun model LSTM adalah dengan mengidentifikasi suatu informasi yang tidak diperlukan lagi dan akan dihilangkan dari sebuah sel informasi. Pada proses dalam mengidentifikasi dan mengecualikan suatu



sel informasi ini diputuskan oleh fungsi *sigmoid*, yang mengambil suatu output dari model LSTM ( $h_{t-1}$ ) pada waktu berikut ( $t-1$ ) dan pada input ( $x_t$ ) memiliki waktu  $t$ . Selain itu, fungsi *sigmoid* menentukan bagian mana dari output lama yang harus dihilangkan. Fungsi itu disebut dengan *Forget gate* di mana  $f_t$  adalah *vector* dengan nilai mulai dari 0 hingga 1, sesuai dengan setiap angka dalam keadaan sel ( $C_{t-1}$ ) [12].

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

Pada tahapan,  $\sigma$  berfungsi dalam *sigmoid*. Pada suatu lapisan *sigmoid* ini akan menghasilkan angka antara 0 dan 1, yang menggambarkan seberapa banyak setiap komponen harus dilewati. Pada nilai 0 mengartikan sebagai tidak ada suatu informasi yang akan diteruskan dan pada nilai 1 mengartikan semua informasi akan diteruskan [12], jika keluarannya 1 maka 1 sebagai semua data yang akan disimpan dan kebalikannya pada keluaran 0 maka semua data akan dibuang [11]. Rumus sebagai berikut :

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (1.1)$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1.2)$$

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (1.3)$$

Pada *input gate* memiliki dua gates yang akan dilaksanakan, pertama akan ditentukan nilai yang akan diperbaharui dengan menggunakan sebuah fungsi aktivasi *sigmoid* [11]. Selanjutnya mengatur sebuah nilai dari input untuk diperbarui pada *cell state*, dapat dihitung dengan persamaan (1.1). *Forget gate* mengatur informasi mana yang akan dihapus dari *cell state*, dapat dihitung dengan persamaan (1.2) dan output gate mengatur nilai yang akan dihasilkan output sesuai dengan input dan memori pada *cell state*, dapat dihitung dengan persamaan (1.3).

Dimana  $i_t$  adalah sebuah *input gate*,  $f_t$  adalah sebuah fungsi *forget gate*, dan  $o_t$  adalah sebuah *output gate*.  $\sigma$  adalah fungsi *sigmoid*,  $W$  dan  $b$  adalah matriks bobot dan bias, masing-masing dari keadaan sel. ( $h_{t-1}$ ) merupakan *concealed state* yaitu memori pada jaringan yang terdapat di *cell state* yang merupakan nilai output sebelum orde ke  $t$  dan  $x_t$  merupakan input pada orde ke  $t$ .

$$N_t = \tanh \cdot (W_n \cdot [h_{t-1}, x_t] + b_n) \quad (1.4)$$

$$C_t = C_{t-1} \cdot f_t + N_t \cdot i_t \quad (1.5)$$

Fungsi *tanh* dalam persamaan (1.4) memberikan suatu bobot pada nilai-nilai yang dilewati, memutuskan suatu tingkat kepentingannya (-1 ke 1). Dua nilai dikalikan untuk memperbarui status sel baru. Memori baru ini kemudian ditambahkan ke memori lama ( $C_{t-1}$ ) yang menghasilkan ( $C_t$ ). Parameter ( $C_{t-1}$ ) dan ( $C_t$ ) adalah keadaan sel pada waktu ( $t-1$ ) dan ( $t$ ), sedangkan  $W$  dan  $b$  adalah matriks bobot serta bias, masing-masing, dari keadaan sel.

$$h_t = O_t \cdot \tanh(C_t) \quad (1.6)$$

Pada tahapan terakhir terlihat pada persamaan di rumus (1.6) nilai output ( $h_t$ ) yang didasarkan pada sel output ( $O_t$ ) merupakan versi yang difilter. Pada awal lapisan *sigmoid* memutuskan bagian sel mana yang membuatnya menjadi output. Setelah itu pada output dari gerbang *sigmoid* ( $O_t$ ) dikalikan dengan nilai-nilai baru yang dibuat pada lapisan *tanh* dari keadaan sel ( $C_t$ ), dengan nilai berkisar antara -1 dan 1 [12].

#### E. Hasil Pembahasan

Hasil uji pada klasifikasi sentimen dari model yang digunakan, telah diukur nilai akurasi, presisi, recall, dan f-measure dari metode LSTM. Dari jumlah data yang digunakan sebanyak 400 data, yang telah diberikan pelabelan manual dan menghasilkan 200 data sentiment positif dan 200 data sentiment negatif. Pada data pelatihan yang digunakan adalah 80% dari total data, yang diproses dengan algoritma *k-fold*. Sementara 20% dari total data digunakan sebagai data pengujian.

Uji klasifikasi dilakukan dengan membandingkan setiap teks tulisan pada sosial media yang telah dilabeli secara manual dengan hasil perhitungan pada metode LSTM. Semakin besar jumlah tulisan yang sesuai dengan label yang diberikan secara manual diawal, semakin tinggi nilai akurasi, presisi, recall, dan f-measure yang diperoleh. Tabel II memperlihatkan hasil klasifikasi menggunakan metode LSTM.

TABEL II  
HASIL UJI TERNARY

Model	Train Acc	Train Loss	Val Acc	Val Loss	Epoch
LSTM	0.9849	0.0896	0.6600	0.6928	20
	0.9950	0.0095	0.6550	0.9949	40
	1.0000	0.0039	0.6600	0.8360	60
	1.0000	0.0016	0.6300	0.2596	80
	1.0000	0.0010	0.6600	1.1996	100

TABEL III  
CONFUSION MATRIX TERNARY

Fold Pertama				
Kelas	Precision	Recall	F1-Score	Support
Negatif	0.62	0.76	0.68	96
Positif	0.72	0.57	0.63	104
Fold Kedua				
Kelas	Precision	Recall	F1-Score	Support
Negatif	0.62	0.72	0.67	96
Positif	0.70	0.60	0.64	104
Fold Ketiga				
Kelas	Precision	Recall	F1-Score	Support
Negatif	0.63	0.69	0.66	96
Positif	0.69	0.63	0.63	104
Fold Keempat				
Kelas	Precision	Recall	F1-Score	Support
Negatif	0.59	0.74	0.66	96
Positif	0.69	0.53	0.60	104

Kelas	Fold Kelima			
	Precision	Recall	F1-Score	Support
Negatif	0.62	0.73	0.67	96
Positif	0.70	0.60	0.65	104

Setelah melakukan lima kali uji coba, maka akurasi yang dihasilkan dapat dilihat pada Tabel III. Dengan mengubah *epoch* yang berbeda-beda, terlihat hasil akurasi yang paling baik pada data training 98% dan testing 66% dengan melakukan sebanyak 20 kali *epoch*.

TABEL IV  
HASIL UJI COBA MULTICLASS

Model	Train Acc	Train Loss	Val Acc	Val Loss	Epoch
LSTM	0.5729	1.3887	0.2400	2.0309	20
	<b>0.9196</b>	<b>0.5163</b>	<b>0.2550</b>	<b>2.1675</b>	<b>40</b>
	0.9950	0.1345	0.2400	2.6501	60
	0.9900	0.0440	0.2450	2.5974	80
	0.9950	0.0284	0.2150	2.6929	100

TABEL V  
CONFUSION MATRIX MULTICLASS

Kelas	Precision	Recall	F1-Score	Support
Desire (hasrat)	0.32	0.30	0.31	43
acceptance (Penerimaan)	0.42	0.15	0.22	33
courage (keberanian)	0.17	0.07	0.10	14
peace (kedamaian)	0.17	0.10	0.12	31
fear (ketakutan)	0.14	0.11	0.12	19
pride (kebanggaan)	0.00	0.00	0.00	10
anger (amarah)	0.20	0.03	0.06	30
love (cinta)	0.50	0.05	0.09	20

Hasil klasifikasi 8 kelas dapat dilihat pada Tabel IV dimana diperoleh nilai akurasi untuk data training 91,9% dan data testing 25,50% dengan melakukan sebanyak 40 kali epoch. Sedangkan untuk 8 kelas multi matrix diperoleh hasil yang dapat dilihat pada Tabel V. Klasifikasi 8 kelas terlihat nilai presisi untuk emosi *acceptance* sebesar 42% dan nilai recall tertinggi pada emosi *desire* yaitu 30% dan nilai F-score tertinggi diperoleh pada emosi *desire* yaitu 31%. Dengan menggunakan jumlah masing-masing data yang diterangkan pada bagian column terakhir atau dicolumn *support* ini dapat memperjelas bagian pada setiap elemen jumlah data.

#### IV. KESIMPULAN

Penelitian analisis sentimen dimana data diklasifikasikan ke dalam 8 kelas yang berbeda telah dilakukan. Hasil yang diperoleh menunjukkan beberapa potensi akurasi pada data ternary dengan nilai akurasi yang baik pada data training 98% dan data testing 66%, dengan melakukan sebanyak 20 kali *epoch*. Sedangkan untuk hasil akurasi ujicoba multiclass dengan 8 label yang berbeda menghasilkan data training dengan nilai 91,9% dan data testing dengan nilai akurasi 25,5%, dengan melakukan sebanyak 40 kali *epoch*, ini memberikan hasil bahwa penelitian ini mendapat nilai rata-rata akurasi dari multiclass 89,45%.

#### REFERENSI

- [1] R. Habibi, D. B. Setyohadi, and E. Wati, "Analisis Sentimen Pada Twitter Mahasiswa Menggunakan Metode Backpropagation," *J. Inform.*, vol. 12, no. 1, pp. 103–109, 2016, doi: 10.21460/inf.2016.121.462.
- [2] F. R. Hartono, Y. A. Sari, and P. P. Adikara, "Pembangkitan Aturan Pengenalan Emosi Pada Twitter Menggunakan Metode Fuzzy-C Means," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 10, pp. 3258–3264, 2018.
- [3] Y. Cahyono, "Analisis Sentiment pada Sosial Media Twitter Menggunakan Naïve Bayes Classifier dengan Feature Selection Particle Swarm Optimization dan Term Frequency," *J. Inform. Univ. Pamulang*, vol. 2, no. 1, p. 14, 2017, doi: 10.32493/informatika.v2i1.1500.
- [4] M. Suhasini and S. Badugu, "Two Step Approach for Emotion Detection on Twitter Data," *Int. J. Comput. Appl.*, vol. 179, no. 53, pp. 12–19, 2018, doi: 10.5120/ijca2018917350.
- [5] R. Abidin, E. Sedyono, and S. Y. joko Prasetyo, "Analisis Data Status Di Facebook Untuk Merekomendasikan Produk Makanan Ringan Menggunakan TF-IDF Dan Document Weighting," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 9, no. 2, pp. 977–984, 2018, doi: 10.24176/simet.v9i2.2518.
- [6] M. A. Nurrohmat and A. SN, "Sentiment Analysis of Novel Review Using Long Short-Term Memory Method," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 13, no. 3, p. 209, 2019, doi: 10.22146/ijccs.41236.
- [7] W. K. Sari, D. P. Rini, R. F. Malik, and I. S. B. Azhar, "Klasifikasi Teks Multilabel pada Artikel Berita Menggunakan Long Short-Term Memory dengan Word2Vec," vol. 1, no. 10, pp. 276–285, 2017.
- [8] U. Asgher *et al.*, "Enhanced Accuracy for Multiclass Mental Workload Detection Using Long Short-Term Memory for Brain-Computer Interface," *Front. Neurosci.*, vol. 14, no. June, pp. 1–19, 2020, doi: 10.3389/fnins.2020.00584.
- [9] M. A. Fatmawati, "Klasifikasi Keluhan Menggunakan Metode Support Vector Machine (SVM) (Studi Kasus : Akun Facebook Group iRaise Helpdesk)," *J. CoreIT*, vol. 3, no. 1, pp. 24–30, 2017, doi: 10.1007/s00766-002-0156-7.
- [10] L. Wiranda and M. Sadikin, "Penerapan Long Short Term Memory Pada Data Time Series Untuk Memprediksi Penjualan Produk Pt. Metiska Farma," *J. Nas. Pendidik. Tek. Inform.*, vol. 8, no. 3, pp. 184–196, 2019.
- [11] M. Wildan, P. Aldi, and A. Aditsania, "Analisis dan Implementasi Long Short Term Memory Neural Network untuk Prediksi Harga Bitcoin," *e-Proceeding Eng.*, vol. 5, no. 2, pp. 3548–3555, 2018.
- [12] D. Li and J. Qian, "Text sentiment analysis based on long short-term memory," in *2016 1st IEEE International Conference on Computer Communication and the Internet, ICCCI 2016*, 2016, pp. 471–475, doi: 10.1109/CCCI.2016.7778967.

## KERTAS KERJA

### Ringkasan

Hasil dari penelitian ini dengan judul Analisis Sentimen *Multi-Class* pada Sosial Media Menggunakan Metode *Long Short-Term Memory* (LSTM). Menganalisa suatu emosi yang ditimbulkan dalam bentuk positif maupun negatif pada penggunaan sosial media, dengan pendekatan klasifikasi data ke dalam 8 kelas yang berbeda. Dengan menggunakan metode *Long Short-Term Memory*, dan implementasi dengan menganalisis akurasi klasifikasi dan presisi. Pengolahan data menggunakan bahasa pemrograman *Python*. Penelitian ini diharapkan dapat menjadikan penelitian lanjutan untuk membantu bidang ilmu psikologi dengan membuat aplikasi antar muka yang menganalisa emosi dari setiap kalimat pengguna pada media sosial.

