

**ANALISIS SENTIMEN *CYBERBULLYING* PADA
INSTAGRAM MENGGUNAKAN *TERM FREQUENCY-
INVERSE DOCUMENT FREQUENCY* (TF-IDF) DAN
JARINGAN *BACKPROPAGATION***

SKRIPSI



MASVIYA TUL CHUSNA

**PROGRAM STUDI S1-MATEMATIKA
DEPARTEMEN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS AIRLANGGA**

2022

**ANALISIS SENTIMEN *CYBERBULLYING* PADA
INSTAGRAM MENGGUNAKAN *TERM FREQUENCY-
INVERSE DOCUMENT FREQUENCY* (TF-IDF) DAN
JARINGAN *BACKPROPAGATION***

SKRIPSI



MASVIYA TUL CHUSNA

**PROGRAM STUDI S1-MATEMATIKA
DEPARTEMEN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS AIRLANGGA**

2022

i

**ANALISIS SENTIMEN *CYBERBULLYING* PADA
INSTAGRAM MENGGUNAKAN *TERM FREQUENCY-
INVERSE DOCUMENT FREQUENCY* (TF-IDF) DAN
JARINGAN *BACKPROPAGATION***

SKRIPSI

Sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Sains Bidang Matematika
pada Fakultas Sains dan Teknologi
Universitas Airlangga

Masviya Tul Chusna
NIM. 081611233057

Disetujui oleh:

Pembimbing I,



Auli Damayanti, S.Si., M.Si.
NIP. 197511072003122004

Pembimbing II,



Drs. Edi Winarko, M.Cs.
NIP. 196705141993031001

LEMBAR PENGESAHAN NASKAH SKRIPSI

Judul : Analisis Sentimen *Cyberbullying* pada Instagram
Menggunakan *Term Frequency-Inverse Document
Frequency* (TF-IDF) dan Jaringan *Backpropagation*
Penyusun : Masviya Tul Chusna
NIM : 081611233057
Pembimbing I : Auli Damayanti, S.Si, M.Si.
Pembimbing II : Drs. Edi Winarko, M.Cs.
Tanggal Ujian : 22 Agustus 2022

Disetujui oleh :

Pembimbing I,



Auli Damayanti, S.Si., M.Si.
NIP. 197511072003122004

Pembimbing II



Drs. Edi Winarko, M.Cs.
NIP. 196705141993031001

Mengetahui,

Ketua Departemen Matematika
Fakultas Sains dan Teknologi
Universitas Airlangga



Dr. Herry Suprajitno, M.Si.
NIP. 196804041994031020

PEDOMAN PENGGUNAAN SKRIPSI

Skripsi ini tidak dipublikasikan, tetapi tersedia di perpustakaan dalam lingkungan Universitas Airlangga, diperkenankan dipakai sebagai referensi kepustakaan, tetapi pengutipan harus seizin penyusun dan harus menyebutkan sumbernya sesuai dengan kebiasaan ilmiah. Dokumen skripsi ini merupakan hak milik Universitas Airlangga.

SURAT PERNYATAAN TENTANG ORISINALITAS

Yang bertanda tangan di bawah ini, saya:

Nama : Masviya Tul Chusna
NIM : 081611233057
Program Studi : S-1 Matematika
Fakultas : Sains dan Teknologi
Jenjang : Sarjana (S1)

Menyatakan bahwa saya tidak melakukan tindakan plagiat dalam penulisan skripsi saya yang berjudul:

“Analisis Sentimen *Cyberbullying* pada Instagram Menggunakan *Term Frequency-Inverse Document Frequency (TF-IDF)* dan Jaringan *Backpropagation*”

Apabila suatu saat nanti terbukti melakukan tindakan plagiat, maka saya akan menerima sanksi yang telah ditetapkan.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya.

Surabaya, 22 Agustus 2022



Masviya Tul Chusna
NIM 081611233057

KATA PENGANTAR



Alhamdulillahirabbil'alamin. Dengan menyebut nama Allah SWT yang Maha Pengasih dan Maha Penyayang. Segala puji syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan taufiq, rahmat, hidayah, dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul “**Analisis Sentimen Cyberbullying pada Instagram Menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) dan Jaringan Backpropagation**”.

Keberhasilan penulis dalam menyusun skripsi ini tidak lepas dari bantuan berbagai pihak yang telah memberikan doa, dukungan, bimbingan, dan masukan demi kesempurnaan penulisan ini. Oleh karena itu, pada kesempatan ini penulis mengucapkan terimakasih yang sebesar-besarnya kepada:

1. Prof Dr. Moh. Nasih SE., MT., Ak. selaku Rektor Universitas Airlangga.
2. Prof. Dr. Moh. Yasin, M.Si. selaku Dekan Fakultas Sains dan Teknologi Universitas Airlangga.
3. Dr. Herry Suprajitno, S.Si, M.Si. selaku Ketua Departemen Matematika dan Koordinator Program Studi S1 Matematika Fakultas Sains dan Teknologi Universitas Airlangga dan Dosen wali yang selalu memberikan saran dan motivasi.
4. Auli Damayanti, S.Si, M.Si. selaku dosen pembimbing I dan Drs. Edi Winarko, M.Cs. selaku dosen pembimbing II yang sabar dalam memberikan bimbingan berupa ilmu, saran, dan motivasi kepada penulis.
5. Seluruh dosen dan staf di Fakultas Sains dan Teknologi, khususnya di Departemen Matematika yang telah mengajarkan ilmu yang bermanfaat kepada penulis selama masa perkuliahan.
6. Kedua orang tua tercinta, Bapak Sukardi dan Ibu Irianis yang tiada henti memberikan kasih sayang, motivasi, doa, dan dukungan dalam segala hal sehingga penulis mampu melanjutkan pendidikan yang lebih tinggi.

7. Kakak tersayang yaitu Windi Pri Iriyanti dan Dimas Yuliatmoko beserta saudaraku yang selalu menghibur, memberikan dukungan, motivasi dan doa kepada penulis.
8. Sahabat karibku Reny, Leli, Cici, April dan Wiwit yang selalu membantu dan memberikan dukungan.
9. Sahabat seperjuangan di perkuliahan Jesis, Talitha, Lula, Birgitta, Rada, Ela, Ratna, Bintang, Viktor, Indah dan keluarga besar Matematika 2016 Universitas Airlangga yang telah berjuang bersama, saling membantu dan memberikan dukungan selama menjalani perkuliahan.
10. Seluruh pihak yang tidak dapat disebutkan satu per satu, yang telah terlibat membantu penulis dalam penyelesaian skripsi secara langsung maupun tidak langsung.

Surabaya, 22 Agustus 2022

Penyusun,

Masviya Tul Chusna

Masviya Tul Chusna. 2022. **Analisis Sentimen Cyberbullying pada Instagram Menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) dan Jaringan Backpropagation**. Skripsi ini dibawah bimbingan Auli Damayanti, S.Si., M.Si. dan Drs. Edi Winarko, M.Cs. Departemen Matematika. Fakultas Sains dan Teknologi. Universitas Airlangga, Surabaya.

ABSTRAK

Penulisan skripsi ini bertujuan untuk melakukan analisis sentimen cyberbullying pada Instagram menggunakan TF-IDF dan jaringan backpropagation, sehingga dapat mengklasifikasikan komentar Instagram yang berunsur *bullying* atau tidak. Hasil klasifikasi bersentimen negatif menyatakan komentar berunsur *bullying*, sedangkan komentar yang bersentimen positif menyatakan tidak ada unsur *bullying*. Proses diawali dengan *preprocessing* teks untuk mendapatkan fitur yang akan digunakan pada proses TF-IDF. Kemudian pada proses TF-IDF fitur akan diubah menjadi data numerik. Hasil TF-IDF akan digunakan sebagai data *input* pada proses analisis sentimen menggunakan jaringan *backpropagation*. Data yang digunakan dalam penelitian ini sebanyak 100 data komentar Instagram yang terdiri dari 50 data bersentimen negatif (*bullying*) dan 50 data bersentimen positif. Data tersebut akan dibagi menjadi 2, antara lain 70 data untuk proses pelatihan dan 30 data untuk proses uji validasi. Berdasarkan hasil implementasi program dari proses uji validasi menunjukkan bahwa nilai akurasi dalam menganalisis sentimen *cyberbullying* pada Instagram sebesar 100%.

Kata Kunci: Analisis Sentimen, Instagram, *Cyberbullying*, *Preprocessing*, TF-IDF, Jaringan *Backpropagation*.

Masviya Tul Chusna. 2022. **Cyberbullying Sentiment Analysis on Instagram Using Term Frequency-Inverse Document Frequency (TF-IDF) and Backpropagation Neural Network**. This thesis is supervised by Auli Damayanti, S.Si., M.Si. and Drs. Edi Winarko, M.Cs. Department of Mathematics. Faculty of Science and Technology. Airlangga University, Surabaya.

ABSTRACT

The aim of this research was to analyze the sentiment of cyberbullying on Instagram using TF-IDF and the backpropagation neural network, therefore it can classify Instagram comments that contain bullying or not. The results of the classification with negative sentiments stated that the comments contained bullying elements, while those with positive sentiments stated that there was no bullying element. The process was started with preprocessing text to get features that will be used in the TF-IDF process. Furthermore, while in the TF-IDF process, the features will be converted into numeric data. The results of the TF-IDF will be used as input data in the sentiment analysis process using the backpropagation neural network. The data analyzed in this research were 100 Instagram comments consisting of 50 data with negative sentiments (bullying) and 50 data with positive sentiments. They will be divided into 2, including 70 data for the training process and 30 data for the validation test process. Based on the results of the program implementation from the validation test process, it showed that the accuracy value in analyzing cyberbullying sentiment on Instagram is 100%.

Keywords: Sentiment Analysis, Instagram, Cyberbullying, Preprocessing, TF-IDF, Backpropagation Neural Network.

DAFTAR ISI

	Halaman
LEMBAR JUDUL	i
LEMBAR PERNYATAAN	ii
LEMBAR PENGESAHAN NASKAH SKRIPSI.....	iii
LEMBAR PEDOMAN PENGGUNAAN SKRIPSI.....	iv
SURAT PERNYATAAN TENTANG ORISINALITAS	v
KATA PENGANTAR	vi
ABSTRAK.....	viii
ABSTRACT	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Analisis Sentimen.....	5
2.2 Instagram dan <i>Cyberbullying</i>	6
2.3 <i>Preprocessing</i>	7
2.4 <i>Term Frequency Inverse Document Frequency</i>	7
2.5 Normalisasi Data	8
2.6 Jaringan Saraf Tiruan	9
2.7 Jaringan <i>Backpropagation</i>	9
BAB III METODE PENELITIAN.....	16

BAB IV PEMBAHASAN.....	24
4.1 Data Komentar Instagram	24
4.2 Prosedur Analisis Sentimen <i>Cyberbullying</i> Menggunakan TF-IDF dan Jaringan <i>Backpropagation</i>	24
4.2.1 Proses <i>Preprocessing</i>	25
4.2.1.1 <i>Case Folding</i>	25
4.2.1.2 <i>Tokenizing</i>	26
4.2.1.3 <i>Stopword Removal</i>	26
4.2.1.4 <i>Stemming</i>	27
4.2.2 Proses TF-IDF	28
4.2.3 Proses Normalisasi Data	29
4.2.4 Proses Pelatihan Data	29
4.2.4.1 Inisialisasi Parameter	30
4.2.4.2 <i>Input Data</i> Pelatihan	30
4.2.4.3 Membangkitkan Bobot dan Bias Awal.....	31
4.2.4.4 Proses <i>Feedforward</i>	32
4.2.4.5 Proses <i>Backpropagation of Error</i>	33
4.2.4.6 Proses Pembaruan Bobot dan Bias	33
4.2.4.7 Proses Perhitungan MSE	34
4.2.5 Proses Uji Validasi	35
4.3 Perhitungan Manual Analisis Sentimen <i>Cyberbullying</i> Menggunakan TF-IDF dan Jaringan <i>Backpropagation</i>	35
4.3.1 Proses <i>Preprocessing</i>	35
4.3.1.1 <i>Case Folding</i>	36
4.3.1.2 <i>Tokenizing</i>	36
4.3.1.3 <i>Stopword Removal</i>	37
4.3.1.4 <i>Stemming</i>	37
4.3.2 Proses TF-IDF	38
4.3.3 Proses Normalisasi Data	40
4.3.4 Proses Pelatihan Data	40

4.3.4.1	Inisialisasi Parameter	41
4.3.4.2	<i>Input</i> Data Pelatihan	41
4.3.4.3	Membangkitkan Bobot dan Bias Awal.....	42
4.3.4.4	Proses Uji Kondisi Berhenti	43
4.3.4.5	Proses <i>Feedforward</i>	45
4.3.4.5	Proses <i>Backpropagation of Error</i>	46
4.3.4.6	Proses Pembaruan Bobot dan Bias	48
4.3.4.7	Proses Uji Kondisi Berhenti	49
4.3.5	Proses Uji Validasi	50
4.4	Implementasi Program	52
BAB V PENUTUP.....		55
5.1	Kesimpulan.....	55
5.2	Saran.....	55
DAFTAR PUSTAKA		57
LAMPIRAN		

DAFTAR TABEL

Nomor	Judul Tabel	Halaman
4. 1	Data Teks untuk Perhitungan Manual	36
4. 2	Hasil <i>Case Folding</i>	36
4. 3	Hasil <i>Tokenizing</i>	37
4. 4	Hasil <i>Stopword Removal</i>	37
4. 5	Hasil <i>Stemming</i>	38
4. 6	Perhitungan TF_{wd} dan DF_w	39
4. 7	Hasil Pembobotan Tiap Kata ($TFIDF_{wd}$)	40
4. 8	Hasil Normalisasi Data	41
4. 9	Data Pelatihan <i>Input</i> dan Target	41
4. 10	Bobot dan Bias Awal dari <i>Input Layer</i> dan <i>Hidden Layer</i>	42
4. 11	Bobot dan Bias Awal dari <i>Hidden Layer</i> dan <i>Output Layer</i>	42
4. 12	Hasil Perhitungan untuk Pola 1	44
4. 13	Hasil Perhitungan untuk Pola 2,3, dan 4	44
4. 14	Hasil Perhitungan Proses <i>Feedforward</i>	46
4. 15	Hasil Perhitungan Proses <i>Backpropagation of Error</i>	47
4. 16	Hasil Perhitungan Proses Pembaruan Bobot dan Bias	49
4. 17	Bobot dan Bias Optimal	50
4. 18	Hasil Perhitungan <i>Feedforward</i> untuk Pola ke-5 dan ke-6	51
4. 19	Hasil Pelatihan dan Uji Validasi Jaringan Backpropagation	53

DAFTAR GAMBAR

Nomor	Judul Gambar	Halaman
3. 1	<i>Flowchart</i> Penelitian Analisis Sentimen <i>Cyberbullying</i>	19
3. 2	<i>Flowchart</i> <i>Preprocessing</i> Teks	20
3. 3	<i>Flowchart</i> TF-IDF	21
3. 4	<i>Flowchart</i> Jaringan <i>Backpropagation</i>	22
3. 5	<i>Flowchart</i> Pengujian Jaringan <i>Backpropagation</i>	23
4. 1	Prosedur Analisis Sentimen <i>Cyberbullying</i>	24
4. 2	Prosedur <i>Preprocessing</i>	25
4. 3	Prosedur <i>Case Folding</i>	25
4. 4	Prosedur <i>Tokenizing</i>	26
4. 5	Prosedur <i>Stopword Removal</i>	26
4. 6	Prosedur <i>Stemming</i>	27
4. 7	Prosedur Proses TF-IDF	28
4. 8	Prosedur Normalisasi Data	29
4. 9	Prosedur Pelatihan Data	30
4. 10	Prosedur Inisialisasi Parameter	30
4. 11	Prosedur <i>Input</i> Data Pelatihan	31
4. 12	Prosedur Membangkitkan Bobot dan Bias Awal	31
4. 13	Prosedur <i>Feedforward</i>	32
4. 14	Prosedur <i>Backpropagation of Error</i>	33
4. 15	Prosedur Pembaruan Bobot dan Bias	34
4. 16	Prosedur Perhitungan MSE	34
4. 17	Prosedur Uji Validasi	35
4. 18	Grafik MSE	54

DAFTAR LAMPIRAN

Nomor	Judul Lampiran
--------------	-----------------------

-
- | | |
|---|---------------------------------|
| 1 | Data Komentar Instagram |
| 2 | Hasil <i>Preprocessing</i> Teks |
| 3 | Normalisasi Data TF-IDF |
| 4 | Hasil Bobot dan Bias Optimal |
| 5 | Hasil Uji Validasi |
| 6 | <i>Source Code</i> Program |
| 7 | Tampilan Program |