

**TUGAS AKHIR**

**ANALISIS SENTIMEN MASYARAKAT TERHADAP  
COVID-19 PADA MEDIA SOSIAL *TWITTER*  
MENGUNAKAN METODE *K-NEAREST NEIGHBOR*  
(K-NN) DAN *NAÏVE BAYES***



**OLEH:**

**ARDIANNE LUTHFIKA FAIRUZ**

**NIM. 16102150**

**PROGRAM STUDI S1 TEKNIK INFORMATIKA  
FAKULTAS INFORMATIKA  
INSTITUT TEKNOLOGI TELKOM PURWOKERTO  
2020**

**TUGAS AKHIR**

**ANALISIS SENTIMEN MASYARAKAT TERHADAP  
COVID-19 PADA MEDIA SOSIAL *TWITTER*  
MENGUNAKAN METODE *K-NEAREST NEIGHBOR*  
(K-NN) DAN *NAÏVE BAYES***

***PUBLIC SENTIMENT ANALYSIS OF COVID-19 IN  
TWITTER USING K-NEAREST NEIGHBOR AND  
NAÏVE BAYES METHODS***

Disusun Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Komputer



**ARDIANNE LUTHFIKA FAIRUZ**

**16102150**

**PROGRAM STUDI S1 TEKNIK INFORMATIKA  
FAKULTAS INFORMATIKA  
INSTITUT TEKNOLOGI TELKOM PURWOKERTO**

**2020**

**HALAMAN PENGESAHAN PEMBIMBING**

**ANALISIS SENTIMEN MASYARAKAT TERHADAP COVID-19 PADA  
MEDIA SOSIAL *TWITTER* MENGGUNAKAN METODE *K-NEAREST  
NEIGHBOR (K-NN) DAN NAÏVE BAYES***

Dipersiapkan dan Disusun Oleh :  
**ARDIANNE LUTHFIKA FAIRUZ**  
**16102150**

Telah dilakukan pembimbingan dan Dinyatakan Siap Melaksanakan  
Sidang Ujian Tugas Akhir  
Periode 2020/2021

Pembimbing I

Rima Dias R, S.Kom., M.Kom  
NIDN. 0602039301

Pembimbing II

Nia Annisa F. T, S.Si., M.Sc  
NIDN. 0630049203

Tugas Akhir ini diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer

Ketua Program Studi  
S1 Teknik Informatika

Fahrudin Mukti Wibowo, S.Kom., M.Eng  
NIDN. 0622018403

**HALAMAN PENETAPAN PENGUJI**

**ANALISIS SENTIMEN MASYARAKAT TERHADAP COVID-19 PADA  
MEDIA SOSIAL *TWITTER* MENGGUNAKAN METODE *K-NEAREST  
NEIGHBOR (K-NN) DAN NAÏVE BAYES***

Dipersiapkan dan Disusun Oleh  
**ARDIANNE LUTHFIKA FAIRUZ**  
**16102150**

Tugas Akhir telah diuji dan Dinilai Panitia Penguji  
Program Studi Teknik Informatika

Fakultas Informatika  
Institut Teknologi Telkom Purwokerto  
Pada Tanggal : 20 Januari 2021

Penguji I,

Siti Khomsah, S.Kom., M.Cs  
NIDN. 0517108101

Penguji II,

Agi Prasetiadi, S.T., M.Eng  
NIDN. 0617098802

## HALAMAN PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini,

Nama Mahasiswa : Ardianne Luthfika Fairuz

NIM : 16102150

Program Studi : Teknik Informatika

**Menyatakan bahwa Tugas Akhir dengan judul:  
ANALISIS SENTIMEN MASYARAKAT TERHADAP COVID19 PADA  
MEDIA SOSIAL TWITTER MENGGUNAKAN METODE K-NEAREST  
NEIGHBOR (K-NN) DAN NAÏVE BAYES**

1. Seluruh tulisan dalam Tugas Akhir ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Institut Teknologi Telkom Purwokerto maupun di perguruan tinggi lainnya.
2. Seluruh tulisan dalam Tugas Akhir ini merupakan gagasan, rumusan, dan penelitian Penulis sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen pembimbing dan menjadi tanggung jawab Penulis seutuhnya.
3. Seluruh tulisan dalam Tugas Akhir ini tidak terdapat karya atau pendapat orang lain. Kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam daftar pustaka.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab Penulis, bukan tanggung jawab Institut Teknologi Telkom Purwokerto.
5. Dilarang mengutip sebagian atau seluruh isi Tugas Akhir ini tanpa mencantumkan atau menyebutkan sumbernya dan dilarang menggunakan dan memperbanyak sebagian atau seluruh isi Tugas Akhir dalam bentuk apapun tanpa izin penulis.
6. Pernyataan ini Saya buat dengan sesungguhnya, apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka Saya bersedia menerima Sanksi Akademik dengan pencabutan gelar yang sudah diperoleh serta sanksi lainnya sesuai norma yang berlaku di Perguruan Tinggi.

Purwokerto, 20 Januari 2021



Menyatakan,

  
Ardianne Luthfika Fairuz

## **KATA PENGANTAR**

Puji syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa atas karuniaNya dan atas segala limpahan rahmatNya, sehingga penyusun dapat menyelesaikan laporan proposal penelitian ini dengan baik. Penulis mengucapkan terima kasih kepada pihak yang telah membantu dalam penelitian ini, untuk itu penulis dalam kesempatan ini mengucapkan terima kasih kepada:

1. Kedua Orang Tua dan seluruh keluarga yang selalu mendukung dalam mengambil sebuah keputusan, mendoakan, dan serta memberikan semangat.
2. Bapak Dr. Ali Rokhman, M.Si. selaku Rektor Institut Teknologi Telkom Purwokerto.
3. Ibu Dr. Tenia Wahyuningrum, S.Kom., M.T. selaku Dekan Fakultas Informatika Institut Teknologi Telkom Purwokerto.
4. Bapak Fahrudin Mukti Wibowo, S.Kom., M.Eng. selaku Ketua Program Studi S1 Informatika.
5. Ibu Rima Dias Ramadhani, S.Kom., M.Kom. selaku Dosen Pembimbing pertama yang telah memberikan bimbingan dan pengarahan pada saat penyusunan proposal penelitian.
6. Ibu Nia Annisa Ferani Tanjung, S.Si., M.Sc. selaku Dosen Pembimbing kedua yang sudah memberikan bimbingan dan pengarahan pada saat penyusunan proposal penelitian.
7. Teman-teman mahasiswa seperjuangan Institut Teknologi Telkom Purwokerto yang sudah banyak memberikan dukungan serta kritik dan saran terhadap penelitian ini.

Dalam penyusunan tugas akhir ini penulis berharap semoga tugas akhir ini dapat bermanfaat dan menambah wawasan bagi pembaca.

Purwokerto, 16 Januari 2021

Ardianne Luthfika Fairuz

## DAFTAR ISI

SAMPUL DEPAN .....	i
SAMPUL DALAM.....	ii
HALAMAN PENGESAHAN PEMBIMBING .....	iii
HALAMAN PENETAPAN PENGUJI.....	iv
KATA PENGANTAR .....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR .....	x
ABSTRAK .....	xi
<i>ABSTRACT</i> .....	xii
BAB I .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	6
1.3 Tujuan Penelitian.....	7
1.4 Manfaat Penelitian.....	7
1.5 Batasan Masalah.....	7
BAB II.....	8
2.1 Penelitian Terdahulu.....	8
2.2 Dasar Teori .....	14
2.2.1 COVID-19.....	14
2.2.2 <i>Natural Language Processing (NLP)</i> .....	14
2.2.3 <i>Sentiment Analysis</i> .....	16
2.2.4 <i>Twitter</i> .....	17
2.2.5 Klasifikasi .....	18
2.2.6 <i>Text Preprocessing</i> .....	21
2.2.7 <i>Term Frequency Inverse Document Frequency (TF-IDF)</i> .....	22
2.2.8 <i>K-Fold Cross Validation</i> .....	23
2.2.9 <i>Python</i> .....	23
2.2.10 <i>Scikit-learn</i> .....	23
BAB III .....	25

3.1	Proses Analisis Sentimen .....	25
3.1.1	Pengumpulan Data .....	27
3.1.2	<i>Text Preprocessing</i> .....	27
3.1.3	Pembobotan <i>Term</i> .....	28
3.1.4	Klasifikasi .....	28
3.1.5	Validasi dengan <i>K-Fold Cross Validation</i> .....	29
BAB IV	.....	31
4.1	Pengumpulan Data .....	31
4.1.1	<i>Crawling Data</i> .....	31
4.1.2	<i>Labeling</i> .....	33
4.2	<i>Text Preprocessing</i> .....	34
4.3	Pembobotan TF-IDF.....	37
4.4	Klasifikasi.....	38
4.4.1	<i>Naïve Bayes</i> .....	39
4.4.2	<i>K-Nearest Neighbor</i> .....	42
4.5	Validasi <i>K-Fold Cross Validation</i> .....	45
4.6	Akurasi .....	46
4.7	Analisa dan Pembahasan .....	47
BAB V	.....	50
5.1	Kesimpulan.....	50
5.2	Saran.....	50
DAFTAR PUSTAKA	.....	51



## DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu .....	11
Tabel 2.2 Sampel Data yang Akan Diklasifikasi .....	19
Tabel 4.1 Hasil Crawling Data Twitter .....	31
Tabel 4.2 Data Labeling .....	33
Tabel 4.3 Proses Case Folding .....	35
Tabel 4.4 Proses Tokenizing .....	35
Tabel 4.5 Proses Stopword Removal .....	36
Tabel 4.6 Proses Stemming .....	37
Tabel 4.7 Perhitungan Manual TF-IDF .....	38
Tabel 4.8 Dokumen yang Akan Diklasifikasi .....	39
Tabel 4.9 Dokumen yang Akan Diklasifikasi .....	42
Tabel 4.10 Pembobotan Dokumen TF-IDF .....	42
Tabel 4.11 Perhitungan Cosine Similiarity .....	43
Tabel 4.12 Urutan Hasil Perhitungan .....	45
Tabel 4.13 Cara Kerja K-Fold Cross Validation .....	45
Tabel 4.14 Hasil Akurasi yang Didapatkan K-Nearest Neighbor .....	46

## DAFTAR GAMBAR

Gambar 1.1 Jumlah Kasus Terkonfirmasi COVID-19 [3].....	1
Gambar 1.2 Data Pengguna Internet Dalam Satuan Juta [7] .....	3
Gambar 1.3 Media Sosial yang Sering Digunakan di Indonesia [7].....	3
Gambar 3.1 Proses Analisis Sentimen .....	27
Gambar 3.2 Diagram Alur Naive Bayes dan K-Nearest Neighbor.....	29