



**ANALISIS SENTIMEN PENGGUNA TWITTER  
TERHADAP PSBB DI JAKARTA  
MENGUNAKAN METODE *NAÏVE BAYES CLASSIFIER***

**SKRIPSI**

**AZINI FAUZIA PUTRI  
1710511075**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA  
FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI INFORMATIKA  
2022**



**ANALISIS SENTIMEN PENGGUNA TWITTER  
TERHADAP PSBB DI JAKARTA  
MENGUNAKAN METODE *NAÏVE BAYES CLASSIFIER***

**SKRIPSI**

**Diajukan Sebagai Salah Satu Syarat Untuk Mendapatkan Gelar Sarjana  
Komputer**

**AZINI FAUZIA PUTRI**

**1710511075**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA  
FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI INFORMATIKA**

**2022**

## PERNYATAAN ORISINALITAS

Tugas skripsi ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Azini Fauzia Putri

NIM : 1710511075

Tanggal : 12 Januari 2022

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 12 Januari 2022

Yang Menyatakan,



(Azini Fauzia Putri)

## PERNYATAAN PERSETUJUAN PUBLIKASI

Sebagai civitas academica Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini :

Nama : Azini Fauzia Putri  
NIM : 1710511075  
Fakultas : Ilmu Komputer  
Program Studi : Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non Eksklusif (*Non-Exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

### ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PSBB DI JAKARTA MENGUNAKAN METODE *NAÏVE BAYES CLASSIFIER*

Dengan hak bebas royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan kata (basis data), merawat dan mempublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada Tanggal : 12 Januari 2022

Yang Menyatakan,



(Azini Fauzia Putri)

## LEMBAR PENGESAHAN

Dengan ini dinyatakan bahwa skripsi berikut

Nama : Azini Fauzia Putri  
NIM : 1710511075  
Program Studi : Informatika  
Judul : Analisis Sentimen Pengguna Twitter Terhadap PSBB di Jakarta Menggunakan Metode *Naïve Bayes Classifier*

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



**DR. Didit Widiyanto, S.Kom., M.Si.**

Penguji I



**Nurul Chamidah, S.Kom., M.Kom.**

Penguji II



**Hj. Ernawati, S.Kom., M.Si.**

Pembimbing I



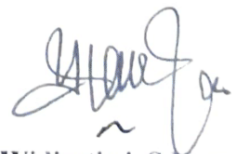
**Anita Muliawati, S.Kom., MTI.**

Pembimbing II



**Dr. Ermatita, M.Kom.**

Dekan



**Yuni Widiastiwi, S.Kom., M.Si.**

Ketua Program Studi

Ditetapkan di : Jakarta

Tanggal Pengesahan : 12 Januari 2022



**ANALISIS SENTIMEN PENGGUNA TWITTER  
TERHADAP PSBB DI JAKARTA  
MENGUNAKAN METODE *NAÏVE BAYES CLASSIFIER***

**AZINI FAUZIA PUTRI**

**ABSTRAK**

PSBB atau Pembatasan Sosial Berskala Besar adalah salah satu tindakan pencegahan dalam penyebaran pandemi COVID-19 yang dilakukan oleh pemerintah. Penerapan PSBB berlangsung hampir di seluruh wilayah Indonesia, salah satunya di Provinsi DKI Jakarta. Setelah PSBB berakhir, kegiatan pembatasan dinamakan PSBB Transisi, di mana adanya kelonggaran dalam beraktivitas dan sejumlah fasilitas umum dibuka dengan memperhatikan protokol kesehatan yang berlaku. Kemudian setelah pemberlakuan PPKM (Pemberlakuan Pembatasan Kegiatan Masyarakat) Jawa dan Bali oleh pemerintah pusat, diberlakukan PSBB ketat di Jakarta yang terus mengalami perubahan peraturan sesuai situasi dan kondisi masyarakat. Pada penelitian ini, dilakukan analisis sentimen masyarakat mengenai PSBB di Jakarta melalui media sosial Twitter dengan metode *Naïve Bayes Classifier*. Data penelitian ini yaitu *tweet* yang didapat dari Twitter menggunakan *keyword* “PSBB DKI Jakarta” yang diambil pada tanggal 1 Februari-31 Maret 2021. Hasil akhir penelitian dengan *oversampling* ini adalah nilai *accuracy* senilai 0.8, nilai *recall* senilai 0.9318, dan nilai *specificity* senilai 0.27.

**Kata kunci:** Analisis Sentimen, *tweet*, *Naïve Bayes Classifier*.

**SENTIMENT ANALYSIS TWITTER USER ABOUT  
LARGE SCALE SOCIAL RESTRICTIONS IN JAKARTA  
USING *NAÏVE BAYES CLASSIFIER* METHOD**

**AZINI FAUZIA PUTRI**

**ABSTRACT**

PSBB or Large-Scale Social Restrictions is one of the measures to prevent spreading COVID-19 pandemic by the government. The implementation of PSBB takes place in almost all parts of Indonesia, one of which is in the DKI Jakarta. After the PSBB ends, the Transitional PSBB begins, where there is leeway in general activities and activities that are opened by observing health protocols. Then after the implementation of the Java and Bali PPKM by the central government, a strict PSBB was imposed in Jakarta which continued to experience changes in regulations accord to situation and public conditions. In this study, an analysis of public sentiment regarding the PSBB in Jakarta was carried out through social media Twitter with Naïve Bayes Classifier method. Data for this research are tweets obtained from Twitter using the keyword "PSBB DKI Jakarta" from 1<sup>st</sup> February-31<sup>th</sup> March 2021. Final result of this research with oversampling is an accuracy value is 0.8, a recall value is 0.9318, and specificity value is 0.8.

**Keywords:** Sentiment Analysis, tweet, *Naïve Bayes Classifier*.

## KATA PENGANTAR

Puji syukur dipanjatkan kepada Allah SWT yang memberikan rahmat dan sehingga penulis mampu menyelesaikan skripsi berjudul “Analisis Sentimen Pengguna Twitter Terhadap PSBB di Jakarta Menggunakan Metode *Naïve Bayes Classifier*”. Penulis berterima kasih kepada :

1. Ibu Iin Ernawati, S.Kom., M.Si. selaku dosen pembimbing I.
2. Ibu Anita Muliawati, S.Kom., M.TI. selaku dosen pembimbing II.
3. Ibu Dr. Ermatita, M.Kom. selaku Dekan Fakultas Ilmu Komputer.
4. Ibu Yuni Widiastiwi, S.Kom., M.Si. selaku Ketua Prodi Informatika juga Dosen Pembimbing Akademik.
5. Seluruh dosen dan pegawai pada Fakultas Ilmu Komputer.
6. Teman-teman Informatika angkatan 2017.
7. Keluarga penulis yang selalu memotivasi untuk menyelesaikan skripsi.
8. Farhana Nabila, Geyessella Manik, dan Hersa Maghdalena selaku sahabat yang selama mendukung dan memberikan semangat.

Penulis juga mengucapkan terima kasih untuk seluruh pihak yang membantu penulisan skripsi ini.

Jakarta, 12 Januari 2022



Penulis



## DAFTAR ISI

HALAMAN AWAL .....	i
JUDUL .....	ii
PERNYATAAN ORISINALITAS .....	iii
PERNYATAAN PERSETUJUAN PUBLIKASI .....	iv
LEMBAR PENGESAHAN .....	v
ABSTRAK .....	vi
ABSTRACT .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI .....	ix
DAFTAR GAMBAR .....	xii
DAFTAR TABEL .....	xiii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Tujuan Penelitian .....	3
1.4 Batasan Masalah .....	3
1.5 Luaran yang diharapkan .....	3
1.6 Manfaat Penelitian .....	3
1.7 Sistematika Penulisan .....	4
BAB II TINJAUAN PUSTAKA .....	5
2.1 Analisis Sentimen .....	5
2.2 Twitter .....	5
2.3 PSBB (Pembatasan Sosial Berskala Besar) .....	6
2.4 <i>Text Mining</i> .....	8
2.5 <i>Text Preprocessing</i> .....	9
2.5.1 <i>Cleaning</i> .....	9
2.5.2 <i>Case Folding</i> .....	9
2.5.3 <i>Tokenization</i> .....	9

2.5.4	Normalisasi Bahasa.....	9
2.5.5	<i>Filtering</i> .....	10
2.5.6	<i>Stemming</i> .....	10
2.6	<i>Term Weighting</i> .....	10
2.6.1	<i>Term Frequency – Invers Document Frequency (TF-IDF)</i> .....	10
2.7	<i>Naïve Bayes Classifier</i> .....	11
2.8	Evaluasi .....	12
2.9	Penelitian Terkait .....	14
BAB III METODOLOGI PENELITIAN.....		16
3.1	Kerangka Pikir.....	16
3.1.1	Menentukan Topik .....	17
3.1.2	Identifikasi Masalah .....	17
3.1.3	Studi Pustaka.....	17
3.1.4	Pengumpulan Data .....	17
3.1.5	<i>Text Preprocessing</i> Tahap 1.....	17
3.1.6	Pelabelan Data.....	18
3.1.7	<i>Text Preprocessing</i> Tahap 2.....	20
3.1.8	Pembobotan Kata .....	20
3.1.9	Klasifikasi .....	20
3.1.10	Evaluasi .....	21
3.1.11	Visualisasi .....	21
3.2	Perangkat Penelitian .....	21
3.3	Jadwal Penelitian .....	22
BAB IV HASIL DAN PEMBAHASAN .....		23
4.1	Pengumpulan Data .....	23
4.2	<i>Text Preprocessing</i> Tahap 1 .....	24
4.3	Pelabelan Data .....	28
4.4	<i>Text Preprocessing</i> Tahap 2 .....	31
4.4.1	<i>Cleaning</i> .....	32
4.4.2	<i>Case Folding</i> .....	33

4.4.3	<i>Tokenization</i> .....	33
4.4.4	Normalisasi Bahasa.....	34
4.4.5	<i>Filtering</i> .....	35
4.4.6	<i>Stemming</i> .....	37
4.5	Pembobotan Kata .....	37
4.6	Klasifikasi.....	40
4.6.1	Proses Pelatihan ( <i>Training</i> ).....	41
4.6.2	Proses Pengujian ( <i>Testing</i> ).....	45
4.7	Evaluasi .....	49
4.8	Visualisasi .....	52
BAB V PENUTUP.....		54
5.1	Kesimpulan.....	54
5.2	Saran.....	54
DAFTAR PUSTAKA .....		56
DAFTAR RIWAYAT HIDUP.....		59
LAMPIRAN.....		60
Lampiran 1 Kamus <i>Slangword</i> .....		61
Lampiran 2 Daftar <i>Stopword</i> .....		63
Lampiran 3 <i>Similarity Index</i> Tugas Akhir .....		66

## DAFTAR GAMBAR

Gambar 2.1 Alur <i>Text Mining</i> .....	8
Gambar 3.1 <i>Flowchart</i> Tahapan Penelitian .....	16
Gambar 4.1 Kode <i>Crawling</i> Data .....	23
Gambar 4.2 Hasil Pengumpulan Data.....	24
Gambar 4.3 Hasil Penghilangan <i>Field</i> Tidak Dibutuhkan.....	24
Gambar 4.4 <i>Tweet</i> Dengan Tanggal Tidak Diperlukan .....	25
Gambar 4.5 Contoh 1 Pada <i>Tweet</i> Duplikat.....	25
Gambar 4.6 Contoh 2 Pada <i>Tweet</i> Duplikat.....	26
Gambar 4.7 Contoh 3 Pada <i>Tweet</i> Duplikat.....	26
Gambar 4.8 Contoh 4 Pada <i>Tweet</i> Duplikat.....	27
Gambar 4.9 Contoh <i>Tweet</i> Data Inkonsisten.....	27
Gambar 4.10 <i>Wordcloud</i> Sentimen Positif PSBB di Jakarta .....	52
Gambar 4.11 <i>Wordcloud</i> Sentimen Negatif PSBB di Jakarta.....	53
Gambar 4.12 Perbandingan Sentimen Pada PSBB di Jakarta.....	53

## DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i> .....	12
Tabel 3.1 Tabel Skala Nilai <i>Kappa</i> .....	20
Tabel 3.2 Jadwal Penelitian.....	22
Tabel 4.1 Contoh Hasil Pelabelan <i>Tweet</i> .....	28
Tabel 4.2 Sampel Data Untuk <i>Kappa Value</i> .....	29
Tabel 4.3 Sampel Data Untuk <i>Text Preprocessing</i> Tahap 2 .....	32
Tabel 4.4 Hasil Sebelum dan Sesudah Proses <i>Cleaning</i> .....	32
Tabel 4.5 Hasil Sebelum dan Sesudah Proses <i>Case Folding</i> .....	33
Tabel 4.6 Hasil Sebelum dan Sesudah Proses <i>Tokenization</i> .....	34
Tabel 4.7 Hasil Sebelum dan Sesudah Proses Normalisasi Bahasa.....	34
Tabel 4.8 Sebelum dan Sesudah Proses <i>Filtering 1</i> .....	35
Tabel 4.9 Hasil Sebelum dan Sesudah Proses <i>Filtering 2</i> .....	36
Tabel 4.10 Hasil Sebelum dan Sesudah Proses <i>Stemming</i> .....	37
Tabel 4.11 Sampel Dokumen Untuk TF-IDF .....	38
Tabel 4.12 Hasil Perhitungan TF-IDF .....	39
Tabel 4.13 Pembagian Data Latih dan Data Uji .....	40
Tabel 4.14 Data Latih.....	41
Tabel 4.15 Nilai TF-IDF Data Latih .....	41
Tabel 4.16 Nilai Probabilitas Data Latih .....	44
Tabel 4.17 Data Uji.....	45
Tabel 4.18 Hasil <i>Preprocessing</i> Data Uji .....	45
Tabel 4.19 Hasil Pembobotan Kata Data Uji.....	46
Tabel 4.20 <i>Confusion Matrix</i> Hasil Evaluasi .....	49
Tabel 4.21 Penerapan Metode SMOTE Pada Data Latih .....	50
Tabel 4.22 <i>Confusion Matrix</i> Hasil Evaluasi Dengan <i>Oversampling</i> .....	51