

**ANALISASENTIMEN PENGGUNA MEDIA  
SOSIALTWITTERTENTANG PPKMDENGAN  
METODEDECISION TREEDAN NAÏVE BAYES**



**SKRIPSI**

diajukan untuk memenuhi salah satu syarat  
memperoleh gelar sarjana komputer

Oleh :

**ARIEF SUTRISNO**  
201769040011

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS YUDHARTA PASURUAN  
2021**

## PERNYATAAN PENULIS

JUDUL : ANALISA SENTIMEN PENGGUNA  
MEDIA SOSIAL *TWITTER* TENTANG  
PPKM DENGAN METODE *DECISION*  
*TREE* DAN *NAÏVE BAYES*


NAMA : ARIEF SUTRISNO

NIM : 201769040011

"Saya menyatakan dan bertanggung jawab dengan sebenarnya bahwa Skripsi ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain mengklaim bahwa Skripsi ini sebagai karyanya, yang disertai dengan bukti-bukti yang cukup, maka saya bersedia untuk dibatalkan gelar Sarjana Komputer saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut".

Pasuruan, 07 Agustus 2021



  
Arief Sutrisno  
Penulis

## PERSETUJUAN SKRIPSI

JUDUL : ANALISA SENTIMEN PENGGUNA  
MEDIA SOSIAL *TWITTER* TENTANG  
PPKM DENGAN METODE *DECISION  
TREE* DAN *NAÏVE BAYES*

NAMA : ARIEF SUTRISNO

NIM : 201769040011

Sripsi ini telah diperiksa dan disetujui

Pasuruan, 07 Agustus 2021

Kaprodi,

Pembimbing



Moch. Lutfi, M. Kom.  
NIP.Y. 0691603004

## PENGESAHAN SKRIPSI

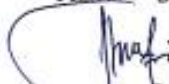
JUDUL : ANALISA SENTIMEN PENGGUNA  
MEDIA SOSIAL *TWITTER* TENTANG  
PPKM DENGAN METODE *DECISION  
TREE* DAN *NAÏVE BAYES*

NAMA : ARIEF SUTRISNO  
NIM : 201769040011

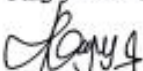

Skripsi ini telah diujikan dan dipertahankan de depan Dewan  
Penguji pada Sidang Skripsi tanggal 7 Agustus 2021. Menurut  
pandangan kami, Skripsi ini memadai dari segi kualitas untuk  
tujuan penganugerahan gelar Sarjana Komputer (S.Kom)

Pasuruan, 07 Agustus 2021

Pembimbing,

  
Moch Lutfi, M. Kom  
NIP. Y. 0691603004  
Penguji Utama,

Penguji Anggota,

   
Cahya Bagus Sanjaya, M. Kom M. Imron Rosadi, M. Kom  
NIP. Y. 0690213127 NIP. Y. 0690213121  
Dekan Fakultas Teknik,



*Skripsi ini kutujukan kepada  
Ayahanda dan Ibunda tercinta,  
serta Adikku tersayang*

## **ABSCTRACT**

*COVID-19 or Coronavirus disease 2019 is currently a pandemic that is spreading rapidly throughout the world. Various handling and policies have been carried out, one of which is the PPKM policy or the Enforcement of Community Work Restrictions issued by the government. PPKM is currently one of the topics of conversation for the public on social media, one of which is social media Twitter. Twitter is one of the microblogging that allows users to write about opinions, comments, and news that discuss issues that occur today. The number of users who post their opinions, it is necessary to have sentiment analysis. The aim is to find out how the public's sentiment towards PPKM is implemented by the government through tweets and comments on the Twitter social media platform. In that case it can be used as a data source used to assess sentiment on Twitter. Automatic classification methods can be used, one of which is the Decision Tree methods and Naïve Bayes Classifier methods. The purpose of this research is to build a system that is automatically able to group each Tweet, and determine the level of accuracy.*

*Text Mining is done in data processing. The first process carried out is preprocessing which consists of case folding, tokenizing, filtering, and stemming. After the preprocessing process is carried out, word weighting will be carried out, then it can be processed using the Dicision Tree methods and Naïve Bayes Classifier methods. The classification stage consists of grouping a data to find a model that has been labeled based on the attributes on each label with the aim of estimating the class of an object whose label is*

*not yet known. After that perform the calculation of accuracy by using the Confusion Matrix. Next, to determine the suitability of the system created, new data is entered that is processed by the system, so that it can determine the data in positive and negative categories.*

*The data from this research is taken from the crawling of the Twitter API. Testing in this study used 80% training data and 20% testing data. The test uses data of 200 tweets consisting of 100 positive tweets and 100 negative tweets. The accuracy results obtained are 85% for Decision Tree and 95% for Naïve Bayes.*

***Keywords:*** *Sentiment Analysis, Covid19, PPKM Policy, Twitter, Text Mining, Word Weighting, Decision Tree, Naïve Bayes Classifier, Confusion Matrix*

# **ANALISASENTIMEN PENGGUNA MEDIA SOSIAL *TWITTER* TENTANG PPKM DENGAN METODE *DECISION TREEDAN NAÏVE BAYES***

Arief Sutrisno

Program Studi Teknik Informatika, Universitas Yudharta  
Pasuruan

## **ABSTRAK**

COVID-19 atau *Coronavirus disease* 2019 saat ini menjadi sebuah pandemik yang menyebar dengan cepat ke seluruh dunia. Berbagai penanganan dan kebijakan sudah dilakukan salah satunya kebijakan PPKM atau Pemberlakuan Pembatasan Kerja Masyarakat yang dikeluarkan pemerintah. PPKM saat ini menjadi salah satu topik perbincangan bagi masyarakat pada media sosial, salah satunya media sosial *Twitter*. *Twitter* merupakan salah satu *microblogging* yang memungkinkan penggunaanya untuk menulis tentang opini, komentar, dan berita yang membahas tentang isu yang terjadi saat ini. Banyaknya pengguna yang melakukan *posting* pendapat mereka, maka perlu adanya analisis sentimen. Tujuannya yaitu untuk mengetahui bagaimana sentimen masyarakat terhadap PPKM yang diterapkan pemerintah melalui *tweet* dan komentar di *platform* media sosial *Twitter*. Dalam hal itu dapat digunakan sebagai sumber data yang digunakan untuk menilai sentimen pada *Twitter*. Cara klasifikasi secara otomatis dapat digunakan, salah satunya menggunakan metode *Decision Treedan Naïve Bayes Classifier*. Tujuan dari penelitian ini adalah membangun sistem yang secara otomatis mampu mengelompokkan setiap *Tweet*, dan mengetahui tingkat akurasinya.



*Text Mining* dilakukan dalam pengolahan datanya. Proses pertama yang dilakukanyaitupreprocessing yang terdiri dari *case folding*, *tokenizing*, *filtering*, dan *stemming*. Setelah dilakukan proses *preprocessing* maka akan dilakukan pembobotan kata, selanjutnya dapat diolah menggunakan metode *Decision Tree* dan *Naïve Bayes Classifier*. Tahap klasifikasi terdiri dari pengelompokan suatu data untuk menemukan suatu model yang sudah diberi label berdasarkan atribut pada tiap label dengan tujuan untuk memperkirakan kelas dari suatu objek yang labelnya belum diketahui. Setelah itu melakukan perhitungan akurasi dengan menggunakan *Confusion Matrix*. Selanjutnya untuk mengetahui kesesuaian sistem yang dibuat, maka dimasukkan data baru yang diproses dengan sistem, sehingga dapat menentukan data dalam kategori positif dan negatif.

Data dari penelitian ini diambil dari *crawling* dari *API Twitter*. Pengujian pada penelitian ini menggunakan data *training* sebanyak 80% dan data *testing* 20%. Pengujian menggunakan data sebanyak 200 *tweet* yang terdiri dari 100 *tweet* positif dan 100 *tweet* negatif. Hasil akurasi yang didapatkan sebesar 85% untuk *Decision Tree* dan 95% untuk *Naïve Bayes*.

**Kata Kunci:** Analisis Sentimen, *Covid19*, Kebijakan PPKM, *Twitter*, *Text Mining*, Pembobotan Kata, *Decision Tree*, *Naïve Bayes Classifier*, *Confusion Matrix*.

## KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT, berkat rahmat dan karunia-Nya akhirnya penulis dapat menyelesaikan skripsi yang berjudul “ANALISA SENTIMEN PENGGUNA MEDIA SOSIAL *TWITTER* TENTANG PPKM DENGAN METODE *DECISION TREE* DAN *NAÏVE BAYES*”.

Tujuan penulis skripsi ini adalah untuk memenuhi persyaratan akademis dalam menyelesaikan studi di program Sarjana Teknik Informatika Universitas Yudharta Pasuruan.

Dalam proses penyusunan skripsi ini penulis memperoleh bimbingan, pengarahan dan masukan dari pihak lainnya. Penulis menyadari bahwa laporan skripsi ini tidak akan terselesaikan dengan baik tanpa adanya bantuan mereka. Sehubungan dengan itu penulis mengucapkan terima kasih dengan sebesar-besarnya kepada:

1. Romo Kyai Sholeh Bahruddin, selaku pembimbing dan Pembina menaungi Universitas Yudharta Pasuruan.
2. Bapak Dr. H. Kholid Murtado, M.HI, selaku Rektor Universitas Yudharta Pasuruan.
3. Bapak Misbach Munir, ST, MT, selaku Dekan Fakultas Teknik Universitas Yudharta Pasuruan.
4. Bapak M. Imron Rosadi, S.Kom, M.Kom, selaku Ketua Program Studi Teknik Informatika Universitas Yudharta Pasuruan.
5. Bapak Moch. Lutfi, S.Kom, M.Kom, selaku Dosen Pembimbing Skripsi yang senantiasa mendampingi dan memberikan masukan.
6. Sujud dan terima kasih yang dalam penulis persembahkan kepada Ibunda dan Ayahanda tercinta, atas dorongan yang kuat, kebijaksanaan dan senantiasa mendo'akan serta tidak pernah bosan memberikan motivasi kepada putranya.

7. Rekan-rekan Teknik Informatika yang senantiasa berbagi ilmu dalam proses perkuliahan selama menjadi mahasiswa.

Penulis menyadari bahwa Skripsi ini masih jauh dari kesempurnaan. Untuk hal itu penulis mengharapkan kritik dan saran dari semua pihak demi mendapatkan kesempurnaan dari laporan ini. Akhir kata semoga laporan ini dapat bermanfaat bagi kita semua.

Pasuruan, Juli 2021

Penulis

## DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>PERNYATAAN PENULIS</b> .....	ii
<b>PERSETUJUAN SKRIPSI</b> .....	Error! Bookmark not defined.
<b>PENGESAHAN SKRIPSI</b> .....	Error! Bookmark not defined.
<b>ABSTRACT</b> .....	vi
<b>ABSTRAK</b> .....	viii
<b>KATA PENGANTAR</b> .....	x
<b>DAFTAR ISI</b> .....	xii
<b>DAFTAR GAMBAR</b> .....	xvi
<b>DAFTAR SOURCECODE</b> .....	xviii
<b>DAFTAR OUTPUT</b> .....	xix
<b>DAFTAR TABEL</b> .....	xx
<b>BAB I</b> .....	Error! Bookmark not defined.
<b>PENDAHULUAN</b> .....	Error! Bookmark not defined.
<b>1.1 Latar Belakang</b> .....	Error! Bookmark not defined.
<b>1.2 Rumusan Masalah</b> ...	Error! Bookmark not defined.
<b>1.3 Tujuan Penelitian</b> ....	Error! Bookmark not defined.
<b>1.4 Manfaat Penelitian</b> ..	Error! Bookmark not defined.
<b>1.5 Batasan Masalah</b> .....	Error! Bookmark not defined.
<b>BAB II</b> .....	Error! Bookmark not defined.

**TINJAUAN PUSTAKA ..... Error! Bookmark not defined.**

**2.1 Penelitian Terkait ..... Error! Bookmark not defined.**

**2.1.1 Analisis Sentimen Data Komentar Sosial Media Facebook Dengan *K-Nearest Neighbor* (Studi Kasus Pada Akun Jasa Ekspedisi Barang J&T Ekspres Indonesia) ..... Error! Bookmark not defined.**

**2.1.2 Analisis Sentimen Pada Twitter Menggunakan Pendekatan *Agglomerative Hierarchical Clustering* Error! Bookmark not defined.**

**2.1.3 Komparasi Algoritma *Text Mining* Untuk Klasifikasi Review Hotel ..... Error! Bookmark not defined.**

**2.1.4 Analisa Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma *Naïve Bayes* ..... Error! Bookmark not defined.**

**2.1.5 Analisa Sentimen Menggunakan *Naïve Bayes* Untuk Melihat Persepsi Masyarakat Terhadap Kenaikan Harga Jual Rokok Pada Media Sosial *Twitter* ..... Error! Bookmark not defined.**

**2.1.6 Analisis Sentimen *Tweet* Menggunakan *Backpropagation Neural Network* ..... Error! Bookmark not defined.**

**2.1.7 Implementasi Sentimen Analisis Komentar Channel Video Pelayanan Pemerintah Di *Youtube* Menggunakan Algoritma *Naïve Bayes* ... Error! Bookmark not defined.**

**2.1.8 *Personality Prediction System from Facebook Users* ..... Error! Bookmark not defined.**

2.1.9 <i>Sentiment Analysis of Social Media Twitter with Case of Anti-LGBT Campaign in Indonesia using Naïve Bayes, Decision Tree, and Random Forest Algorithm</i> .....	Error! Bookmark not defined.
2.1.10 <i>Lung Cancer Survivability Prediction based on Performance Using Classification Techniques of Support Vector Machines, C4.5 and Naive Bayes Algorithms for Healthcare Analytics</i> .....	Error! Bookmark not defined.
2.2 <b>Kajian Teori</b> .....	Error! Bookmark not defined.
2.2.1 <b>Analisis Sentimen</b> ...	Error! Bookmark not defined.
2.2.2 <b>Data Mining</b> .....	Error! Bookmark not defined.
2.2.3 <i>Text Mining</i> .....	Error! Bookmark not defined.
2.2.4 <i>NLP</i> .....	Error! Bookmark not defined.
2.2.5 <b>Klasifikasi</b> .....	Error! Bookmark not defined.
2.2.6 <i>Decision Tree</i> .....	Error! Bookmark not defined.
2.2.7 <i>Naïve Bayes Classifier</i> .....	Error! Bookmark not defined.
2.2.8 <b>Pembobotan Kata</b> ...	Error! Bookmark not defined.
2.2.9 <b>Akurasi</b> .....	Error! Bookmark not defined.
2.2.10 <b>Evaluasi Performa</b>	Error! Bookmark not defined.
<b>BAB III</b> .....	Error! Bookmark not defined.
<b>METODE PENELITIAN</b> .....	Error! Bookmark not defined.
3.1 <b>Kerangka Teori</b> .....	Error! Bookmark not defined.
3.2 <b>Data</b> .....	Error! Bookmark not defined.

3.3 Metodologi Penelitian....	Error! Bookmark not defined.
3.3.1 Tahap Prosedur Penelitian ...	Error! Bookmark not defined.
3.4 Analisa .....	Error! Bookmark not defined.
3.4.1 Kebutuhan Data .....	Error! Bookmark not defined.
3.4.2 Tahapan <i>Text Mining</i> .....	Error! Bookmark not defined.
3.5 Tahap Evaluasi .....	Error! Bookmark not defined.
3.6 Spesifikasi Alat .....	Error! Bookmark not defined.
3.7 Perancangan Penelitian	Error! Bookmark not defined.
3.8 Metode Yang Diusulkan	Error! Bookmark not defined.
BAB IV .....	Error! Bookmark not defined.
IMPLEMENTASI DAN PEMBAHASAN .	Error! Bookmark not defined.
4.1 Kebutuhan Perangkat Sistem .....	Error! Bookmark not defined.
4.2 Kebutuhan Perangkat Keras ( <i>Hardware</i> ) .....	Error! Bookmark not defined.
4.3 Kebutuhan Perangkat Lunak ( <i>Software</i> ) .....	Error! Bookmark not defined.
4.4 Implementasi <i>Crawling Data</i> .....	Error! Bookmark not defined.
4.5 Perhitungan Nilai .....	Error! Bookmark not defined.

4.5.1 Implementasi Proses <i>Load Data Excel</i> .....	Error! Bookmark not defined.
4.5.2 Implementasi Proses <i>Preprocessing</i> .....	Error! Bookmark not defined.
4.5.3 Implementasi <i>TF-IDF</i> .....	Error! Bookmark not defined.
4.5.4 Implementasi <i>Splitting Data</i> .	Error! Bookmark not defined.
4.6 Pengujian Hasil.....	Error! Bookmark not defined.
BAB V .....	Error! Bookmark not defined.
PENUTUP .....	Error! Bookmark not defined.
5.1 Kesimpulan .....	Error! Bookmark not defined.
5.2 Saran.....	Error! Bookmark not defined.
DAFTAR PUSTAKA .....	Error! Bookmark not defined.
LAMPIRAN-LAMPIRAN.....	Error! Bookmark not defined.

## DAFTAR GAMBAR

	Halaman
<b>Gambar 2.1</b> Metode yang diusulkan Penelitian (Salam et al., 2018) .....	6
<b>Gambar 2.2</b> Metode yang diusulkan Penelitian (Yenni Tresnawati, 2017).....	7
<b>Gambar 2.3</b> Metode yang diusulkan Penelitian (Taufik, 2018) .....	8



<b>Gambar 2.4</b> Metode yang diusulkan Penelitian (Muljono et al., 2018).....	8
<b>Gambar 2.5</b> Metode yang diusulkan Penelitian (AFSHOH, 2017) .....	9
<b>Gambar 2.6</b> Metode yang diusulkan Penelitian (Assuja & Saniati, 2016) .....	10
<b>Gambar 2.7</b> Metode yang diusulkan Penelitian (Saputra et al., 2019) .....	11
<b>Gambar 2.8</b> Metode yang diusulkan Penelitian (Tandera et al., 2017).....	12
<b>Gambar 2.9</b> Metode yang diusulkan Penelitian (Fitri et al., 2019) .....	13
<b>Gambar 2.10</b> Metode yang diusulkan Penelitian (Pradeep & Naveen, 2018) .....	14
<b>Gambar 2.11</b> Tahap <i>Knowledge Discovery in Database (KDD)</i> .....	20
<b>Gambar 3.1</b> Kerangka Teori .....	31
<b>Gambar 3.2</b> <i>Tweet</i> Positif .....	32
<b>Gambar 3.3</b> <i>Tweet</i> Negatif.....	32
<b>Gambar 3.4</b> Contoh <i>Tokenizing</i> .....	35
<b>Gambar 3.5</b> Contoh <i>Filtering</i> .....	36
<b>Gambar 3.6</b> Contoh <i>Stemming</i> .....	37
<b>Gambar 3.7</b> <i>TF tweet</i> 1 .....	37

<b>Gambar 3.8</b> <i>TF tweet 2</i> .....	38
<b>Gambar 3.9</b> <i>TF tweet 3</i> .....	38
<b>Gambar 3.10</b> Metode yang diusulkan .....	42
<b>Gambar 4.1</b> Hasil <i>Crawling</i> .....	46
<b>Gambar 4.2</b> Jumlah Data .....	46

## DAFTAR SOURCECODE

	Halaman
<i>Sourcecode 4.1</i> Proses Pencarian Data .....	45
<i>Sourcecode 4.2</i> Proses Analisis Data.....	45
<i>Sourcecode 4.3</i> Proses <i>Load</i> Data.....	47
<i>Sourcecode 4.4</i> Proses <i>Case Folding</i> .....	49
<i>Sourcecode 4.5</i> Proses <i>Tokenizing</i> .....	50
<i>Sourcecode 4.6</i> Proses <i>Filtering</i> .....	52
<i>Sourcecode 4.7</i> Proses <i>Stemming</i> .....	53
<i>Sourcecode 4.8</i> Proses <i>TF-IDF</i> .....	54
<i>Sourcecode 4.9</i> Proses <i>Splitting</i> .....	55
<i>Sourcecode 4.10</i> Proses <i>Confusion Matrix Dicision Tree</i> .....	56
<i>Sourcecode 4.11</i> Proses <i>Confusion Matrix Naïve Bayes Classifier</i> .....	57

## DAFTAR OUTPUT

	Halaman
<i>Output 4.1</i> Hasil <i>Load Data</i> .....	48
<i>Output 4.2</i> Hasil <i>Case Folding</i> .....	49
<i>Output 4.3</i> Hasil <i>Tokenizing</i> .....	51
<i>Output 4.4</i> Hasil <i>Filtering</i> .....	52
<i>Output 4.5</i> Hasil <i>Stemming</i> .....	54
<i>Output 4.6</i> Hasil <i>Confusion Matrix Decision Tree</i> .....	56
<i>Output 4.7</i> Hasil <i>Confusion Matrix Naïve Bayes Classifier</i> ...	58

## DAFTAR TABEL

	Halaman
<b>Tabel 2.1</b> Tabel Penelitian Terdahulu.....	14
<b>Tabel 2.2</b> <i>Confusion Matrix</i> .....	29
<b>Tabel 3.1</b> Menghitung <i>df</i> .....	38
<b>Tabel 3.2</b> Menghitung <i>idf</i> .....	39
<b>Tabel 3.3</b> Hitung bobot ( $W_{ij}$ ) <i>tweet</i> 1.....	40
<b>Tabel 3.4</b> Hitung bobot ( $W_{ij}$ ) <i>tweet</i> 2.....	40
<b>Tabel 3.5</b> Hitung bobot ( $W_{ij}$ ) <i>tweet</i> 3.....	40



