



**ANALISIS SENTIMEN TENTANG UU CIPTA KERJA MENGGUNAKAN  
ALGORITMA *NAÏVE BAYES***

**SKRIPSI**

**I GUSTI NAUFHAL DAFFA ADNYANA  
1710511082**

**INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA  
JAKARTA  
2021**



**ANALISIS SENTIMEN TENTANG UU CIPTA KERJA MENGGUNAKAN  
ALGORITMA *NAÏVE BAYES***

**SKRIPSI**

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

**I GUSTI NAUFHAL DAFFA ADNYANA  
1710511082**

**INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA  
JAKARTA  
2021**

## PERNYATAAN ORISINALITAS

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

Nama : I Gusti Naufhal Daffa Adnyana  
NIM : 1710511082  
Tanggal : 24 Juli 2021

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

Jakarta, 24 Juli 2021

Yang Menyatakan,



(I Gusti Naufhal Daffa Adnyana)

**PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI  
UNTUK KEPENTINGAN AKADEMIS**

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

Nama : I Gusti Naufhal Daffa Adnyana

NIM : 1710511082

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 TENTANG UU CIPTA KERJA MENGGUNAKAN  
ALGORITMA NAÏVE BAYES**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non eksklusif ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta selama waktu yang sudah ditetapkan.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada Tanggal : 24 Juli 2021

Yang menyatakan,



(I Gusti Naufhal Daffa Adnyana)

## LEMBAR PENGESAHAN

Dengan ini dinyatakan bahwa tugas akhir berikut :

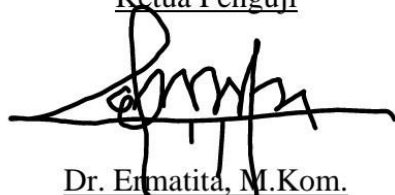
Nama : I Gusti Naufhal Daffa Adnyana  
NIM : 1710511082  
Program Studi : Informatika  
Judul : ANALISIS SENTIMEN TENTANG UU CIPTA  
KERJA MENGGUNAKAN METODE *NAÏVE*  
*BAYES*

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



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

Ketua Penguji



Dr. Ermatita, M.Kom.

Pembimbing I



Dr. Ermatita, M.Kom.

Dekan Fakultas Ilmu Komputer



Nurul Chamidah, S.Kom.,M.Kom

Anggota Penguji



Sarika, S.Kom.,M.Kom.

Pembimbing II



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

Ketua Program Studi Informatika

Ditetapkan di : Jakarta

Tanggal Persetujuan : 23 Juli 2021



# ANALISIS SENTIMEN TENTANG UU CIPTA KERJA MENGGUNAKAN METODE *NAÏVE BAYES*

I Gusti Naufhal Daffa Adnyana

## ABSTRAK

*Twitter* adalah platform yang memungkinkan orang untuk mengekspresikan keinginan, pendapat, dan kritik mereka secara langsung. Pengesahan dan Pengundangan UU Cipta Kerja telah menimbulkan banyak opini yang beragam di masyarakat khususnya para pengguna *twitter*. Tujuan dari penelitian ini adalah untuk memahami opini para pengguna *Twitter* terhadap UU Cipta Kerja melalui tagar #UUCIPTAKERJA. Dalam penelitian ini, opini publik digolongkan sebagai positif dan negatif, kemudian digunakan algoritma *Naive Bayes* untuk mengklasifikasikan *tweet*. Data yang diambil dimulai dari tanggal 5 Oktober - 30 November 2020 menggunakan API yang disediakan oleh pihak *Twitter*. Hasil pengujian yang didapat menggunakan data uji memiliki nilai akurasi sebesar 80,53%, nilai *recall* sebesar 84,78%, dan nilai *specificity* sebesar 73,79%.

**Kata Kunci:** *Twitter*, UU Cipta Kerja, Klasifikasi, *Naive Bayes*.

# **ANALYSIS OF SENTIMENT ABOUT OMNIBUS LAW USING NAÏVE BAYES METHOD**

I Gusti Naufhal Daffa Adnyana

## **ABSTRACT**

Twitter is a platform that allows people to express their wishes, opinions and critiques directly. The ratification and promulgation of the Omnibus Law has generated many diverse opinions in the community, especially Twitter users. The purpose of this study is to understand the opinions of Twitter users on the Omnibus Law through the hashtag #UUCIPTAKERJA. In this study, public opinion is classified as positive and negative, then the Naive Bayes algorithm is used to classify tweets. The data is taken starting from October 5 - November 30, 2020 using the API provided by Twitter. The test results obtained using test data have an accuracy value of 80.53%, a recall value of 84.78%, and a specificity value of 73.79%.

**Keyword:** *Twitter, Omnibus Law, Classification, Naïve Bayes.*

## KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas segala karunia-Nya sehingga Skripsi ini berhasil diselesaikan, tak lupa Sholawat dan salam kepada Nabi besar Muhammad SAW yang telah membawa manusia dari zaman kebodohan menuju zaman terang benderang. Penulis ingin mengucapkan terima kasih kepada:

1. Kedua orang tua penulis, I Gusti Putu Sila Adnyana (ayah) dan Susanti (ibu), serta keluarga yang selalu memberi dorongan dan doa sehingga dapat menyelesaikan skripsi ini.
2. Almarhumah Ibu Dr. Titin Pramiyati S.Kom., M.Si. Ibu Dr. Ermatita M.Kom., dan Ibu Sarika, S.Kom, M.Kom., selaku dosen pembimbing yang telah memberikan saran yang bermanfaat.
3. Ibu Dr. Ermatita, M. Kom. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.
4. Ibu Yuni Widiastiwi, S. Kom., M. Si. selaku Ketua Program Studi Informatika Universitas Pembangunan Nasional Veteran Jakarta.
5. Bapak/Ibu dosen Informatika Universitas Pembangunan Nasional Veteran Jakarta terima kasih atas ilmu-ilmu yang bermanfaat.
6. Kak Tantri Ayu Prasetiarini, yang selalu memberi petunjuk hingga terselesaikannya skripsi ini.
7. Teman-teman Fakultas Ilmu Komputer, rekan-rekan KSM Robotika, dan seluruh pihak terkait yang membantu terselesaikannya skripsi ini yang tidak dapat penulis sebutkan namanya satu persatu, terima kasih atas saran dan dukungan yang telah diberikan.

Akhir kata, semoga skripsi ini dapat bermanfaat bagi para pembacanya.

Jakarta, 1 Juli 2021

Penulis



## DAFTAR ISI

PERNYATAAN ORISINALITAS .....	i
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS .....	ii
LEMBAR PENGESAHAN.....	iii
ABSTRAK .....	iv
ABSTRACT .....	v
KATA PENGANTAR .....	vi
DAFTAR ISI .....	vii
DAFTAR GAMBAR .....	x
DAFTAR TABEL .....	xi
DAFTAR SIMBOL .....	xii
BAB 1 PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	2
1.3 Batasan Masalah .....	3
1.4 Tujuan Penelitian .....	3
1.5 Manfaat Penelitian .....	3
1.6 Luaran yang Diharapkan .....	4
1.7 Sistematika Penulisan.....	4
BAB 2 TINJAUAN PUSTAKA.....	6
2.1 <i>Twitter</i> .....	6
2.2 Analisis Sentimen.....	6
2.3 <i>Data Mining</i> .....	7
2.3.1 <i>Text Mining</i> .....	8
2.4 <i>Text Analysis</i> .....	8
2.5 <i>Twitter Crawling</i> .....	8
2.6 <i>Text Preprocessing</i> .....	9
2.7 <i>Term Frequency – Inverse Document Frequency (TF IDF)</i> .....	10

2.8	<i>Naïve Bayes</i> (NB).....	10
2.9	Evaluasi .....	12
2.10	Penelitian Terkait .....	14
BAB 3 METODOLOGI PENELITIAN .....		16
3.1	Tahapan Penelitian .....	16
3.1.1	Identifikasi Masalah.....	16
3.1.2	Studi Pustaka.....	17
3.1.3	Pengumpulan Data.....	17
3.1.4	Pelabelan Data .....	17
3.1.5	<i>Text Preprocessing</i> .....	19
3.1.5.1	<i>Cleaning</i> .....	20
3.1.5.2	<i>Case Folding</i> .....	20
3.1.5.3	<i>Tokenizing</i> .....	20
3.1.5.4	Normalisasi Bahasa .....	21
3.1.5.5	<i>Filtering</i> .....	21
3.1.6	Pembobotan <i>Term</i> .....	21
3.1.7	Klasifikasi <i>Naïve Bayes</i> .....	22
3.1.8	Evaluasi.....	22
3.1.7	Visualisasi .....	23
3.2	Perangkat Penelitian.....	24
3.3	Jadwal Penelitian.....	24
BAB 4 HASIL DAN PEMBAHASAN .....		26
4.1	Data .....	26
4.1.1	Pelabelan Data .....	26
4.2	<i>Text Preprocessing</i> .....	32
4.2.1	<i>Cleaning</i> .....	32

4.2.2	<i>Case Folding</i> .....	34
4.2.3	<i>Tokenizing</i> .....	35
4.2.4	Normalisasi Bahasa .....	35
4.2.5	<i>Filtering</i> .....	37
4.3	Pembobotan Kata (TF- IDF).....	37
4.4	Klasifikasi <i>Naïve Bayes</i> .....	40
4.4.1	Proses Latih.....	41
4.4.2	Proses Uji .....	47
4.5	Hasil Evaluasi .....	51
4.6	Visualisasi.....	53
BAB 5 PENUTUP .....		59
5.1	Kesimpulan.....	59
5.2	Saran.....	60
DAFTAR PUSTAKA .....		61
DAFTAR RIWAYAT HIDUP .....		64
LAMPIRAN .....		65
Lampiran 1 <i>Similarity Index</i> Tugas Akhir.....		66




## DAFTAR GAMBAR

Gambar 3.1 Bagan Tahapan Penelitian .....	16
Gambar 4.1 Hasil Crawling dengan Format.csv.....	26
Gambar 4.2 <i>Wordcloud</i> Sentimen Positif UU Cipta Kerja Setelah Pengesahan.....	54
Gambar 4.3 <i>Wordcloud</i> Sentimen Negatif UU Cipta Kerja Setelah Pengesahan .....	54
Gambar 4.4 Perbandingan Sentimen UU Cipta Kerja Setelah Pengesahan.....	56
Gambar 4.5 <i>Wordcloud</i> Sentimen Positif UU Cipta Kerja Setelah Pengundangan ...	56
Gambar 4.6 <i>Wordcloud</i> Sentimen Negatif UU Cipta Kerja Setelah Pengundangan ..	56
Gambar 4.7 Perbandingan Sentimen UU Cipta Kerja Setelah Pengundangan .....	56

## DAFTAR TABEL

Tabel 2.1 Contoh Tabel Confusion Matrix .....	13
Tabel 3.1 Tabel Skala Nilai <i>Kappa</i> .....	19
Tabel 3.2 <i>Confussion Matrix</i> .....	23
Tabel 3.3 Jadwal Penelitian.....	25
Tabel 4.1 Contoh Pelabelan Data Tweet.....	27
Tabel 4.2 Sampel Data Tweet Setelah Pengesahan UU Cipta Kerja.....	28
Tabel 4.3 Data Sebelum Proses Cleaning .....	33
Tabel 4.4 Data Hasil Proses Cleaning.....	33
Tabel 4.5 Data Hasil Proses Case Folding .....	34
Tabel 4.6 Data Hasil Proses <i>Tokenizing</i> .....	35
Tabel 4.7 Data Hasil Proses Normalisasi .....	36
Tabel 4.8 Data Hasil Proses <i>Filtering</i> .....	37
Tabel 4.9 Sampel Dokumen .....	38
Tabel 4.10 Perhitungan TF IDF.....	39
Tabel 4.11 Pembagian Data yang Digunakan .....	40
Tabel 4.12 Data Latih .....	41
Tabel 4.13 Nilai TF IDF Data Latih .....	41
Tabel 4.14 Nilai Probabilitas Data Latih.....	45
Tabel 4.15 Tabel Data Uji.....	47
Tabel 4.16 Data Uji Setelah Preprocessing.....	47
Tabel 4.17 Pembobotan Data Uji.....	47
Tabel 4.18 Confussion Matrix Hasil Evaluasi Naïve Bayes .....	51
Tabel 4.19 Penerapan Undersampling Pada Data Latih .....	52
Tabel 4.20 Confussion Matrix Hasil Evaluasi Naïve Bayes dengan Undersampling .....	53

## DAFTAR SIMBOL

<b>Simbol <i>Flowchart</i></b>			
<b>NO</b>	<b>Simbol</b>	<b>Nama Simbol</b>	<b>Keterangan</b>
1		<i>Terminator</i> ( <i>start, end</i> )	Simbol yang menggambarkan dimulainya atau berakhirnya suatu kegiatan berakhir.
2		<i>Flow Line</i>	Simbol yang menggambarkan hubungan dari simbol satu ke simbol lainnya dalam suatu proses kegiatan.
3		<i>Process</i>	Simbol yang menggambarkan suatu pengolahan yang dijalankan.