

**ANALISIS SENTIMEN TERHADAP RUU PKS MENGGUNAKAN  
ALGORITMA NAÏVE BAYES CLASSIFIER SEBAGAI BAHAN  
REKOMENDASI BAGI PEMANGKU KEBIJAKAN**

**SKRIPSI**



disusun oleh

**Male Istighosah**

**17.61.0106**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2021**

**ANALISIS SENTIMEN TERHADAP RUU PKS MENGGUNAKAN  
ALGORITMA NAÏVE BAYES CLASSIFIER SEBAGAI BAHAN  
REKOMENDASI BAGI PEMANGKU KEBIJAKAN**

**SKRIPSI**

untuk memenuhi sebagian persyaratan  
mencapai gelar Sarjana  
pada Program Studi Informatika



disusun oleh

**Male Istighosah**

**17.61.0106**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2021**

**PERSETUJUAN**

**SKRIPSI**

**ANALISIS SENTIMEN TERHADAP RUU PKS MENGGUNAKAN  
ALGORITMA NAIVE BAYES CLASSIFIER SEBAGAI BAHAN  
REKOMENDASI BAGI PEMANGKU KEBIJAKAN**

yang dipersiapkan dan disusun oleh

**Male Istighosah**

**17.61.0106**

telah disetujui oleh Dosen Pembimbing Skripsi

pada tanggal

**Dosen Pembimbing**

**Sri Ngudi Wahyuni, S.T., M.Kom**

**NIK. 190302060**

**PENGESAHAN**

**SKRIPSI**

**ANALISIS SENTIMEN TERHADAP RUU PKS MENGGUNAKAN  
ALGORITMA NAIVE BAYES CLASSIFIER SEBAGAI BAHAN  
REKOMENDASI BAGI PEMANGKU KEBIJAKAN**

yang dipersiapkan dan disusun oleh

**Male Istighosah**

**17.61.0106**

telah dipertahankan di depan Dewan Penguji  
pada tanggal

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Sri Ngudi Wahyuni, S.T., M.Kom**

**NIK. 190302060**

**Norhikmah, M.Kom**

**NIK. 190302245**

**Eli Pujastuti, M.Kom**

**NIK. 190302227**

Skripsi ini telah diterima sebagai salah satu persyaratan  
Untuk memperoleh gelar Sarjana Komputer  
Tanggal

**DEKAN FAKULTAS ILMU KOMPUTER**

**Hanif Al Fatta M.Kom**

**NIK.190302096**

## PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi Pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis di acuan dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta,



Maie Istighosah

NIM. 17.61.0106

## MOTTO

*“Learn to enjoy every minute of your life. Be happy now. Don't wait for something outside of yourself to make you happy in the future. Think how really precious is the time you have to spend, whether its at work or with your family.*

*Every minute should be enjoyed and savored”*

-Earl Nightingale

*“You cannot change your future, but you can change your habits, and surely your habits will change your future”*

-Dr. Abdul Kalam

## PERSEMBAHAN

Dengan rasa syukur yang mendalam, dengan telah diselesaikannya skripsi ini, penulis mempersembahkannya untuk:

1. Kedua orang tua yang memberi “kehidupan”, Ibuku tersayang, Tri Sukarti yang tidak pernah lelah menyadarkan kepada putrinya bahwa dunia terlalu keras untuk berpangku tangan kepada orang lain, selalu mengajarkan untuk bekerja keras untuk membuat diri sendiri bangga, dan selalu berdiri sendiri meraih semua yang diinginkan dengan tanganku sendiri. Untuk Bapakku tersayang, R.Biwafiqudin yang mengalirkan semangat juang dan kerja keras tiada henti melawan segala keterbatasan.
2. Kakakku tersayang, Iema Binti Chumaidah selalu mengajarkan bahwa tidak ada yang tidak mungkin dengan keajaiban usaha dan doa, menjadi partner untuk menghadapi sukacita, tawa, air mata dan perselisihan yang dikirim oleh kehidupan, dan tetap berpegangan erat.
3. Orang-orang yang memberi daya hidup, teruslah hidup dalam cinta dan kebaikan.

## KATA PENGANTAR

*Assalamu 'alaikum Wr.Wb.*

*Alhamdulillahirabbil'Alamin.* Segala puji bagi Allah SWT yang telah memberikan rahmat, hidayah, dan ridha-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis Sentiment terhadap RUU PKS Menggunakan Algoritma Naïve Bayes Classifier sebagai Bahan Rekomendasi bagi Pemangku Kebijakan".

Skripsi ini disusun dan diajukan untuk memenuhi syarat perolehan gelar Sarjana Komunikasi (S.Kom) pada Fakultas Ilmu Sosial dan Hukum di Universitas AMIKOM Yogyakarta. Disamping itu, penulisan skripsi ini juga bertujuan untuk memberikan pengetahuan kepada pembaca.

Dalam menyelesaikan skripsi ini, banyak pihak yang telah memberikan bantuan secara moril, nasihat, dan semangat maupun material. Atas segala bantuan yang telah diberikan, penulis ingin menyampaikan doa dan ucapan terimakasih yang tak terhingga kepada:

1. M. Suyanto, Prof, Dr, M.M. selaku rector Universitas AMIKOM Yogyakarta
2. Krisnawati, S.Si., M.T. selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta
3. Sudarmawan, S.T., M.T. selaku ketua JuRUSAN Informatika Universitas AMIKOM Yogyakarta
4. Erik Hadi Saputra, S.Kom., M.Eng. selaku dosen wali selama saya menempuh Pendidikan S1 di Universitas AMIKOM Yogyakarta
5. Sri Ngudi Wahyuni, S.T., M.Kom, selaku dosen pembimbing skripsi yang telah memberikan banyak pengarahan
6. Segenap civitas akademika jurusan Informatika, terutama seluruh dosen, terimakasih atas segenap ilmu dan bimbingannya.



7. Untuk Bapak, Ibu, Mbak Iema dan keluarga yang memberikan dukungan, semangat, dan rumah bagi penulis selama ini.
8. Seluruh pihak yang membantu penulis dalam penyelesaian skripsi ini dan tidak dapat disebutkan satu persatu.
9. *Last but not least, I wanna thank me, I wanna thank me for believing in me, I wanna thank me for doing all this hard work, I wanna thank me for having no days off, I wanna thank me for never quitting, For just being me at all times.*

Atas segala kekurangan dan ketidak sempurnaan skripsi ini, penulis sangat mengharapkan masukan, kritik, dan saran yang bersifat membangun ke arah perbaikan dan penyempurnaan skripsi ini, agar dalam penyusunan karya tulis selanjutnya dapat lebih baik.

Penulis berharap semoga skripsi ini dapat bermanfaat bagi seluruh pihak dan semoga amal baik yang telah diberikan kepada penulis mendapat balasan dari Allah SWT.

*Aamiin ya rabbal'alam.*

*Waalamualaikum Wr.Wb.*

Yogyakarta,

Maie Istighosah  
NIM. 17.61.0106

## DAFTAR ISI

JUDUL .....	i
PERSETUJUAN .....	ii
PENGESAHAN .....	iii
PERNYATAAN .....	<b>Error! Bookmark not defined.</b>
MOTTO .....	v
PERSEMBAHAN .....	vi
KATA PENGANTAR .....	vii
DAFTAR ISI .....	ix
DAFTAR TABEL .....	xii
DAFTAR GAMBAR .....	xiii
DAFTAR ISTILAH .....	xiv
INTISARI .....	xv
<i>ABSTRACT</i> .....	xvi
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah .....	4
1.4 Maksud dan Tujuan Penelitian .....	4
1.5 Manfaat Penelitian .....	5
1.6 Metode Penelitian .....	5
1.6.1 Metode Pengumpulan Data .....	5
1.6.2 Analisis Kebutuhan .....	6
1.7 Sistematika Penulisan .....	7

BAB II LANDASAN TEORI .....	9
2.1 Kajian Pustaka .....	9
2.2 Landasan Teori .....	11
2.2.1 Twitter .....	11
2.2.2 <i>Phyton</i> .....	12
2.2.3 <i>Sentiment Analysis</i> .....	12
2.2.4 <i>Text Mining</i> .....	13
2.2.5 <i>Text Preprocessing</i> .....	13
2.2.6 <i>Naïve Bayes Classifier</i> .....	14
2.2.7 Pembobotan Kata ( <i>Term Weighting</i> ) .....	15
2.2.8 <i>Confusion Matrix</i> .....	16
BAB III METODE PENELITIAN .....	19
3.1 Alur Penelitian .....	19
3.1.1 Metode Pengumpulan <i>Dataset</i> .....	19
3.1.2 Seleksi dan Pelabelan Data .....	19
3.1.3 <i>Preprocessing Data</i> .....	20
3.1.4 Pembobotan Kata ( <i>TF-IDF</i> ) .....	23
BAB IV HASIL DAN PEMBAHASAN .....	38
4.1 Pengumpulan Data ( <i>Crawling</i> ) .....	38
4.2 Pembuatan Data Latih .....	38
4.3 <i>Preprocessing Data</i> .....	39
4.4 Analisis dan Evaluasi .....	42
4.4.1 Analisis .....	42
4.4.2 Pengujian Algoritma Dengan <i>Confusion Matrix</i> .....	43
4.5 Implementasi .....	44

BAB V PENUTUP.....	47
5.1 Kesimpulan.....	47
5.2 Saran.....	47
DAFTAR PUSTAKA .....	49



## DAFTAR TABEL

Tabel 2.1	Kajlan Pustaka.....	10
Tabel 2.2	Confusion Matrix.....	16
Tabel 3.1	Contoh Hasil Cleansing.....	21
Tabel 3.2	Contoh Hasil <i>Case Folding</i> .....	21
Tabel 3.3	<i>Tokenisasi</i> .....	22
Tabel 3.4	<i>Convert Slangword</i> .....	22
Tabel 3.5	<i>Stopword Removal</i> .....	23
Tabel 3.6	<i>Stemming</i> .....	23
Tabel 3.7	Contoh Koleksi Data.....	24
Tabel 3.8	Contoh Hasil <i>Preprocessing Data Tweet</i> dengan Label.....	24
Tabel 3.9	Nilai <i>Term Frequency</i> .....	25
Tabel 3.10	Nilai <i>Term Frequency Neutral</i> .....	28
Tabel 3.11	Nilai <i>Term Frequency Negative</i> .....	30
Tabel 3.12	Nilai <i>Term Frequency Positive</i> .....	32
Tabel 3.13	Nilai <i>TF-IDF</i> .....	36
Tabel 4.1	Komposisi Sentimen Data Training.....	39

## DAFTAR GAMBAR

Gambar 3.1	Alur Sistem.....	19
Gambar 3.2	Preprocessing .....	20
Gambar 4.1	Fungsi Crawling Data .....	38
Gambar 4.2	Fungsi Cleansing.....	40
Gambar 4.3	Fungsi Case Folding .....	40
Gambar 4.4	Fungsi Tokenizing .....	40
Gambar 4.5	Fungsi Convert Slangword .....	41
Gambar 4.6	Fungsi Stopword Removal.....	41
Gambar 4.7	Fungsi Stemming .....	42
Gambar 4.8	Fungsi TF-IDF .....	43
Gambar 4.9	Fungsi Naïve Bayes Classifier.....	43
Gambar 4.10	Fungsi Implementasi Model Klasifikasi.....	45
Gambar 4.11	Hasil Implementasi Model Klasifikasi pada Data Uji.....	45
Gambar 4.12	Data Kata yang Sering Muncul.....	46

## DAFTAR ISTILAH

*Crawling* adalah proses pengumpulan data dari halaman sebuah web untuk mengindeks informasi.

*Library* yaitu sekumpulan kode yang memiliki fungsi-fungsi tertentu yang ditulis dalam bahasa pemrograman, dan dapat dipanggil kedalam program lain.

*Python* adalah bahasa pemrograman interpretatif multiguna dengan kode yang mudah diimplementasikan dapat digunakan di berbagai platform yang berfokus pada keterbacaan kode dan merupakan bahasa populer yang berkaitan dengan *Data Science*.

Probabilitas adalah peluang dari suatu kejadian, terjadi atau tidak dan seberapa besar kemungkinan kejadian tersebut berpeluang untuk terjadi.

Sastrawi *stemmer* adalah python *library* *stemmer* sederhana yang memungkinkan untuk mengubah kata dalam bahasa Indonesia ke kata dasarnya .

*Script* merupakan *code* yang digunakan untuk menerjemahkan setiap perintah dalam website.

## INTISARI

Kekerasan seksual terus meningkat dan bentuk bentuk kekerasan semakin kompleks. KemenPPA mencatat jumlah total kasus kekerasan terhadap perempuan dan anak pada tahun 2021 telah mencapai angka 3112 kasus. UU PKS memang diperlukan untuk menyempurnakan perlindungan hukum terhadap kekerasan seksual terutama terhadap korban perempuan dan anak. RUU PKS harus segera disahkan untuk memaksimalkan perlindungan kepada korban kekerasan seksual.

Sifat media social yang terbuka membuat pengguna bebas untuk mengekspresikan pendapatnya. Twitter merupakan situs jejaring social yang digemari masyarakat di Indonesia. Penelitian ini memanfaatkan sentiment dari pengguna posting di social media Twitter atau lebih dikenal dengan tweet. Menggunakan metode pengumpulan data *crawling*, data yang dianalisis adalah postingan tweet hasil pencarian kata kunci ruu pks, kekerasan seksual, dan tolak ruu pks.

Hasil implementasi analisis sentiment pada data uji dengan hastag #ruupks dengan memanfaatkan model klasifikasi dari data latih menggunakan Naïve Bayes Classifier dan pembobotan TF-IDF didapatkan akurasi hasil klasifikasi sebanyak 85,96%. Dengan hasil akurasi yang didapatkan diharapkan dapat menjadi bahan pertimbangan dan rekomendasi bagi pemangku kebijakan untuk segera mengesahkan RUU PKS.

Kata kunci: analisis sentiment, twitter, naïve bayes classifier, ruupks, tf-idf.



## **ABSTRACT**

*Sexual violence continues to increase, and the forms of violence are increasingly complex. KemenPPA noted that the total number of cases of violence against women and children in 2021 had reached 3112 cases. The UU PKS is indeed needed to improve legal protection against sexual violence, especially against women and child victims. The RUU PKS must be ratified immediately to maximize protection for victims of sexual violence.*

*The open nature of social media makes users free to express their opinions. Twitter is a popular social networking site in Indonesia. This study utilizes the sentiment of users posting on social media Twitter or better known as tweets. Using the crawling data collection method, the data analyzed is posting tweets from the search results for the RUU PKS keyword.*

*The results of the implementation of sentiment analysis on test data with the hashtag #rupks by utilizing a classification model from training data using the Naïve Bayes Classifier and TF-IDF weighting obtained the classification accuracy of 85.96%. With the accurate results obtained, it is hoped that they can be taken into consideration and recommendations for policymakers to immediately ratify the RUU PKS.*

**Keywords:** *sentiment analysis, twitter, naive bayes classifier, rupks, tf-idf.*