

**ANALISIS SENTIMEN MASYARAKAT TERHADAP  
HASIL *REAL COUNT* KPU 2024 PADA MEDIA  
SOSIAL INSTAGRAM MENGGUNAKAN  
*NAIVE BAYES* DAN *SUPPORT  
VECTOR MACHINE***

**SKRIPSI**

Oleh :

**Silviana Sevira**

**202010225179**



**PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS BHAYANGKARA JAKARTA RAYA  
2024**

## LEMBAR PERSETUJUAN PEMBIMBING

Judul Tugas Akhir : Analisis Sentimen Masyarakat terhadap Hasil *Real Count* KPU 2024 pada Media Sosial Instagram Menggunakan *Naive Bayes* dan *Support Vector Machine*

Nama Mahasiswa : Silviana Sevira

Nomor Pokok Mahasiswa : 202010225179

Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 27 Juni 2024

Jakarta, 05 Juli 2024

MENYETUJUI,

Pembimbing I



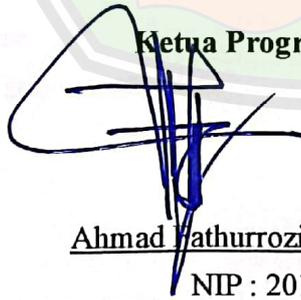
Wowon Priatna, S.T., M.TI.  
NIDN : 0429118007

Pembimbing II



Khairunnisa Fadhillah Ramdhania, S.Si., M.Si.  
NIDN : 0328039201

Ketua Program Studi



Ahmad Athurrozi, S.E., M.M.S.I.  
NIP : 2012486

Program Studi Informatika  
Fakultas Ilmu Komputer  
Universitas Bhayangkara Jakarta Raya  
2024

## LEMBAR PENGESAHAN

Judul Tugas Akhir : Analisis Sentimen Masyarakat terhadap Hasil *Real Count* KPU 2024 pada Media Sosial Instagram Menggunakan *Naive Bayes* dan *Support Vector Machine*

Nama Mahasiswa : Silviana Sevira

Nomor Pokok Mahasiswa : 202010225179

Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 27 Juni 2024

Jakarta, 05 Juli 2024

MENGESAHKAN,

Ketua Tim Penguji : Dr. Herlawati, S.Si., M.M., M.Kom.

NIDN : 0311097302

Penguji I : Hendarman Lubis, S.Kom., M.Kom.

NIDN : 0013077002

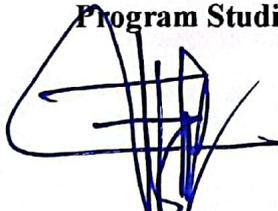
Penguji II : Wowon Priatna, S.T., M.T.I

NIDN : 0429118007

MENGETAHUI,

Ketua

Program Studi Informatika

  
Ahmad Fathurrozi, S.E., M.M.S.I  
NIP. 2012486

Dekan

Fakultas Ilmu Komputer

  
Dr. Dra. Tyastuti Sri Lestari, M.M  
NIP. 1408206



UNIVERSITAS BHAYANGKARA JAKARTA  
RAYA FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI INFORMATIKA

**LEMBAR PERNYATAAN BUKAN PLAGIASI**

Yang bertanda tangan dibawah ini :

Nama : Silviana Sevira

NPM : 202010225179

Program Studi : Informatika

Fakultas : Ilmu Komputer

Judul Tugas Akhir : Analisis Sentimen Masyarakat Terhadap Hasil *Real Count* KPU 2024 Pada Media Sosial Instagram Menggunakan *Naive Bayes* Dan *Support Vector Machine*

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Jakarta, 05 Juli 2024

  
Silviana Sevira.

## ABSTRAK

**Silviana Sevira. 202010225179.** Analisis Sentimen Masyarakat Terhadap Hasil *Real Count* KPU 2024 Pada Media Sosial Instagram Menggunakan *Naïve Bayes* Dan *Support Vector Machine*.

Di Indonesia, sistem pemerintahan yang diterapkan adalah demokrasi, salah satu ciri utamanya adalah pemilihan umum secara berkala. Media sosial, khususnya Instagram, telah menjadi platform utama bagi masyarakat untuk berdiskusi dan menyampaikan opini terkait berbagai isu, termasuk proses pemilu dan hasil real count, terdapat beragam pendapat masyarakat yang mencerminkan kecurigaan terhadap kemungkinan terjadinya kecurangan dalam proses tersebut. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap hasil *Real Count* KPU 2024 berdasarkan komentar di Instagram. Metode yang digunakan adalah algoritma *Naïve Bayes* dan *Support Vector Machine* untuk mengklasifikasikan sentimen menjadi positif dan negatif. Hasil analisis menunjukkan bahwa dari total 1.500 komentar, 865 komentar bersentimen negatif, sementara 635 komentar bersentimen positif. Untuk pengujian model, dataset dibagi menjadi beberapa rasio. Pada rasio 70:30, akurasi *Naïve Bayes* menghasilkan 73% dan *Support Vector Machine* 76%. Pada rasio 80:20, akurasi *Naïve Bayes* sebesar 77% dan *Support Vector Machine* 76%. Pada rasio 90:10, hasil akurasi *Naïve Bayes* sebesar 81% dan *Support Vector Machine* 75%. Hal ini menunjukkan *Naïve Bayes* menghasilkan akurasi lebih tinggi pada rasio 90:10 dan 80:20, sedangkan *Support Vector Machine* menghasilkan akurasi lebih tinggi hanya pada rasio 70:30. Dapat disimpulkan bahwa metode *Naïve Bayes* lebih unggul dengan rasio 90:10 dalam menganalisis sentimen masyarakat terhadap hasil *Real Count* KPU 2024 daripada *Support Vector Machine*.

**Kata Kunci:** Analisis Sentimen, *Real Count* KPU 2024, Instagram, *Naïve Bayes*, *Support Vector Machine*.

## ***ABSTRACT***

**Silviana Sevira. 202010225179. Analysis of Public Sentiment towards KPU 2024**

*Real Count Results on Instagram Social Media Using Naïve Bayes and Support Vector Machine.*

*In Indonesia, the system of government implemented is democracy, one of the main characteristics of which is periodic elections. Social media, especially Instagram, has become the main platform for people to discuss and express opinions on various issues, including the election process and the results of the real count, there are various public opinions that reflect suspicion of possible fraud in the process. This study aims to analyze public sentiment towards the 2024 KPU Real Count results based on comments on Instagram. The methods used are Naive Bayes algorithm and Support Vector Machine to classify sentiment into positive and negative. The analysis results show that out of a total of 1,500 comments, 865 comments have negative sentiments, while 635 comments have positive sentiments. For model testing, the dataset was divided into several ratios. At a 70:30 ratio, the Naïve Bayes accuracy yielded 73% and Support Vector Machine 76%. At a ratio of 80:20, Naïve Bayes accuracy was 77% and Support Vector Machine 76%. At a ratio of 90:10, Naïve Bayes accuracy results in 81% and Support Vector Machine 75%. This shows that Naïve Bayes produces higher accuracy at a ratio of 90:10 and 80:20, while Support Vector Machine produces higher accuracy only at a ratio of 70:30. It can be concluded that the Naive Bayes method is superior to the 90:10 ratio in analyzing public sentiment towards the 2024 KPU Real Count results than the Support Vector Machine.*

**Keywords:** *Sentiment Analysis, Real Count KPU 2024, Insagram, Naive Bayes, Support Vector Machine.*

## LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI

### LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Silviana Sevira  
NPM : 202010225179  
Program Studi : Informatika  
Fakultas : Ilmu Komputer  
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (Non-Exclusive Royalty-Free Right)**, atas karya ilmiah saya yang berjudul : *Analisis Sentimen Masyarakat Terhadap Hasil Real Count KPU 2024 Pada Media Sosial Instagram Menggunakan Naive Bayes Dan Support Vector Machine* beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta  
Pada tanggal : 05 Juli 2024  
Yang Menyatakan

  
METEPAI  
TEMPEL  
B28ALX177535596  
Silviana Sevira

## KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT yang telah melimpahkan rahmat, hidayah, serta karunia-Nya sehingga penulis dapat menyelesaikan penulisan skripsi ini yang berjudul “**Analisis Sentimen Masyarakat Terhadap Hasil Real Count KPU 2024 pada Media Sosial Instagram Menggunakan Naive Bayes dan Support Vector Machine**”.

Dalam penulisan skripsi ini tidak terlepas dari bantuan dan dukungan berbagai pihak yang telah memberikan motivasi, bimbingan, serta dorongan selama proses penulisan berlangsung. Oleh karena itu, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Bapak Irjen Pol (Purn) Dr. Drs. Bambang Karsono, S.H., M.M. selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Ibu Dr. Tyastuti Sri Lestari, S.Si., M.M. selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Bapak Ahmad Fathurrozi, S.E., M.M.S.I. selaku Kepala Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara.
4. Bapak Wowon Priatna, S.T., M.TI. selaku Dosen Pembimbing I dalam penelitian skripsi Program Studi Informatika Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya.
5. Ibu Khairunnisa Fadhillah Ramdhania, S.Si., M.Si. selaku Dosen Pembimbing II penelitian skripsi Program Studi Informatika Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya.

6. Bapak Tri Dharma Putra, S.T., M.Sc. selaku Dosen Pembimbing Akademik Program Studi Informatika Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya.
7. Kepada Bapak Sarwani dan Ibu Erica jantung hatiku, selaku orang tua yang sangat penulis cintai, terimakasih senantiasa memberikan doa, dukungan berupa materi dan kasih sayang yang tiada henti, menjadikan motivasi penulis untuk terus berusaha sehingga dapat menyelesaikan skripsi ini.
8. Mahesa Agung selaku *support system* penulis yang menemani selama penyusunan skripsi ini.
9. Sahabat-sahabat yang saya sayangi yaitu Syafira Cessa, Prilia Hashifah dan Nida Rachmatin yang selalu saling membantu dan berjuang bersama selama perkuliahan 8 semester dari awal hingga akhir.
10. Dan kepada teman-teman setiakku Peelka, Bina dan Nurwati yang senantiasa menghibur dan mendengarkan keluh kesah penulis.

Penulis menyadari bahwa dalam penulisan skripsi ini masih terdapat kekurangan dan keterbatasan. Oleh karena itu, kritik dan saran yang bersifat membangun sangat penulis harapkan guna perbaikan di masa yang akan datang. Penulis berharap semoga skripsi ini dapat memberikan manfaat dan kontribusi yang positif bagi perkembangan ilmu pengetahuan serta dapat menjadi bahan referensi bagi peneliti selanjutnya.

Jakarta, 05 Juli 2024

Hormat Saya



Silviana Sevira

## DAFTAR ISI

<b>LEMBAR PERSETUJUAN PEMBIMBING</b> .....	i
<b>LEMBAR PENGESAHAN</b> .....	ii
<b>LEMBAR PERNYATAAN BUKAN PLAGIASI</b> .....	iii
<b>ABSTRAK</b> .....	iv
<b>ABSTRACT</b> .....	v
<b>LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI</b> .....	vi
<b>KATA PENGANTAR</b> .....	vii
<b>DAFTAR ISI</b> .....	ix
<b>DAFTAR TABEL</b> .....	xiii
<b>DAFTAR GAMBAR</b> .....	xiv
<b>DAFTAR LAMPIRAN</b> .....	xvi
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang .....	1
1.2 Identifikasi Masalah .....	4
1.3 Rumusan Masalah .....	5
1.4 Batasan Masalah .....	5
1.5 Tujuan Penelitian .....	5
1.6 Manfaat Penelitian .....	6
1.7 Sistematika Tugas Akhir .....	6

<b>BAB II TINJAUAN PUSTAKA .....</b>	<b>8</b>
2.1 Penelitian Terdahulu .....	8
2.2 Landasan Teori .....	11
2.2.1 <i>Machine Learning</i> .....	11
2.2.2 <i>Data Mining</i> .....	11
2.2.3 <i>Text Mining</i> .....	12
2.2.4 <i>Knowledge Discovery in Databases (KDD)</i> .....	13
2.2.5 <i>Text Preprocessing</i> .....	15
2.2.6 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i> .....	16
2.2.7 <i>Naïve Bayes</i> .....	17
2.2.8 <i>Support Vector Machine</i> .....	18
2.2.9 <i>Confusion Matrix</i> .....	20
2.2.10 <i>Lexicon Based</i> .....	22
2.2.11 <i>SMOTE</i> .....	23
2.2.12 <i>Word Cloud</i> .....	23
2.2.13 <i>Analisis Sentimen</i> .....	24
2.2.14 <i>Real Count</i> .....	24
2.2.15 <i>Klasifikasi</i> .....	25
2.2.16 <i>Instagram</i> .....	26
2.2.17 <i>Jupyter Notebook</i> .....	26
2.2.18 <i>Python</i> .....	27

<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>28</b>
3.1 Objek Penelitian .....	28
3.2 Kerangka Penelitian .....	29
3.3 Metode Pengumpulan Data .....	31
3.4 Metode Analisis Data .....	31
3.4.1 <i>Data Selection</i> .....	31
3.4.2 <i>Preprocessing</i> .....	32
3.4.3 <i>Labeling</i> .....	34
3.4.4 <i>Transformation</i> .....	34
3.4.5 <i>SMOTE</i> .....	34
3.4.6 <i>Data Mining</i> .....	34
3.4.7 <i>Evaluation</i> .....	36
<b>BAB IV HASIL DAN PEMBAHASAN.....</b>	<b>37</b>
4.1 <i>Data Selection</i> .....	37
4.2 <i>Preprocessing</i> .....	37
4.2.1 <i>Case Folding</i> .....	38
4.2.2 <i>Tokenizing</i> .....	39
4.2.3 <i>Normalize</i> .....	40
4.2.4 <i>Filtering / Stopword Removing</i> .....	42
4.2.5 <i>Stemming</i> .....	43
4.3 <i>Labeling</i> .....	45

4.4	<i>Transformation</i> .....	46
4.5	<i>SMOTE</i> .....	48
4.6	<i>Splitting Data</i> .....	50
4.7	<i>Data Mining</i> .....	50
4.7.1	<i>Naïve Bayes</i> .....	50
4.7.2	<i>Support Vector Machine</i> .....	54
4.8	<i>Evaluation</i> .....	58
4.8.1	<i>Model Naïve Bayes</i> .....	59
4.8.2	<i>Model Support Vector Machine</i> .....	65
4.8.3	<i>Pembahasan</i> .....	72
4.9	<i>Word Cloud</i> .....	73
<b>BAB V PENUTUP</b>	.....	76
5.1	<i>Kesimpulan</i> .....	76
5.2	<i>Saran</i> .....	76
<b>DAFTAR PUSTAKA</b>	.....	78
<b>LAMPIRAN</b>	.....	86

## DAFTAR TABEL

<b>Tabel 2. 1</b> Penelitian Terkait .....	8
<b>Tabel 2. 2</b> Tabel <i>Confusion Matrix</i> .....	21
<b>Tabel 4. 1</b> Hasil Tahap <i>Case Folding</i> .....	38
<b>Tabel 4. 2</b> Hasil Tahapan <i>Tokenizing</i> .....	40
<b>Tabel 4. 3</b> Hasil Tahapan <i>Normalize</i> .....	41
<b>Tabel 4. 4</b> Hasil Tahapan <i>Filtering / Stopword Removing</i> .....	43
<b>Tabel 4. 5</b> Hasil Tahap <i>Stemming</i> .....	44
<b>Tabel 4. 6</b> Contoh Dataset Perhitungan .....	47
<b>Tabel 4. 7</b> Perhitungan Manual <i>TF, DF, IDF</i> .....	48
<b>Tabel 4. 8</b> Perhitungan Manual <i>TF-IDF</i> .....	48
<b>Tabel 4. 9</b> <i>Data Training</i> .....	51
<b>Tabel 4. 10</b> <i>Data Testing</i> .....	54
<b>Tabel 4. 11</b> Contoh Perhitungan Manual <i>Support Vector Machine</i> .....	55
<b>Tabel 4. 12</b> Hasil Perbandingan <i>NB</i> dan <i>SVM</i> Dengan Beberapa Rasio.....	73
<b>Tabel 4. 13</b> Jumlah Kata <i>Wordcloud</i> Positif .....	74
<b>Tabel 4. 14</b> Jumlah Kata <i>Wordcloud</i> Negatif.....	75

## DAFTAR GAMBAR

<b>Gambar 2. 1</b> Tahapan <i>KDD</i> .....	13
<b>Gambar 2. 2</b> Tahapan <i>Text Preprocessing</i> .....	15
<b>Gambar 3. 1</b> Objek penelitian postingan Instagram terkait <i>Real Count 2024</i> .....	28
<b>Gambar 3. 2</b> Kerangka Penelitian.....	29
<b>Gambar 3. 3</b> Tahapan <i>Case Folding</i> .....	32
<b>Gambar 3. 4</b> Tahapan <i>Tokenizing</i> .....	32
<b>Gambar 3. 5</b> Tahapan <i>Normalize</i> .....	33
<b>Gambar 3. 6</b> Tahapan <i>Filtering / Stopword Removing</i> .....	33
<b>Gambar 3. 7</b> Tahapan <i>Stemming</i> .....	33
<b>Gambar 3. 8</b> Tahapan <i>Naïve Bayes</i> .....	35
<b>Gambar 3. 9</b> Tahapan <i>Support Vector Machine</i> .....	36
<b>Gambar 3. 10</b> Tahapan <i>Confusion Matrix</i> .....	36
<b>Gambar 4. 1</b> Data Komentar Instagram.....	37
<b>Gambar 4. 2</b> Contoh Kode <i>Case Folding</i> .....	38
<b>Gambar 4. 3</b> Contoh Kode <i>Tokenizing</i> .....	39
<b>Gambar 4. 4</b> Contoh Kode <i>Text Normalize</i> .....	41
<b>Gambar 4. 5</b> Kamus Text Normalisasi.....	41
<b>Gambar 4. 6</b> Contoh Kode <i>Filtering / Stopword Removing</i> .....	42
<b>Gambar 4. 7</b> Contoh Kode Tahap <i>Stemming</i> .....	44
<b>Gambar 4. 8</b> Contoh Kode <i>Labeling</i> .....	45
<b>Gambar 4. 9</b> Hasil <i>Labeling</i> .....	45
<b>Gambar 4. 10</b> Contoh Kode <i>TF-IDF</i> Per-kalimat.....	46

<b>Gambar 4. 11</b> Hasil <i>TF-IDF</i> Per-kalimat .....	47
<b>Gambar 4. 12</b> Contoh Kode <i>TF-IDF</i> Seluruh Dokumen .....	47
<b>Gambar 4. 13</b> Kode dan Hasil <i>SMOTE</i> .....	49
<b>Gambar 4. 14</b> Hasil dan Grafik Setelah <i>SMOTE</i> .....	49
<b>Gambar 4. 15</b> Kode dan Hasil <i>Splitting Data</i> Rasio 70:30.....	50
<b>Gambar 4. 16</b> Kode dan Hasil <i>Splitting Data</i> Rasio 80:20.....	50
<b>Gambar 4. 17</b> Kode dan Hasil <i>Splitting Data</i> Rasio 90:10.....	50
<b>Gambar 4. 18</b> Contoh Kode <i>Modelling Naïve Bayes</i> .....	51
<b>Gambar 4. 19</b> Contoh Kode <i>Modelling Support Vector Machine</i> .....	55
<b>Gambar 4. 20</b> Hasil Evaluasi <i>Naïve Bayes</i> Rasio 70:30 .....	59
<b>Gambar 4. 21</b> <i>Confusion Matrix</i> Dengan <i>Naïve Bayes</i> Rasio 70:30 .....	60
<b>Gambar 4. 22</b> Hasil Evaluasi <i>Naïve Bayes</i> Rasio 80:20 .....	61
<b>Gambar 4. 23</b> <i>Confusion Matrix</i> Dengan <i>Naïve Bayes</i> Rasio 80:20 .....	62
<b>Gambar 4. 24</b> Hasil Evaluasi <i>Naïve Bayes</i> Rasio 90:10 .....	63
<b>Gambar 4. 25</b> <i>Confusion Matrix</i> Dengan <i>Naïve Bayes</i> Rasio 90:10 .....	64
<b>Gambar 4. 26</b> Hasil Evaluasi <i>SVM</i> Rasio 70:30 .....	66
<b>Gambar 4. 27</b> <i>Confusion Matrix</i> Dengan <i>SVM</i> Rasio 70:30.....	66
<b>Gambar 4. 28</b> Hasil Evaluasi <i>SVM</i> Rasio 80:20 .....	68
<b>Gambar 4. 29</b> <i>Confusion Matrix</i> Dengan <i>SVM</i> Rasio 80:20.....	68
<b>Gambar 4. 30</b> Hasil Evaluasi <i>SVM</i> Rasio 90:10 .....	70
<b>Gambar 4. 31</b> <i>Confusion Matrix</i> Dengan <i>SVM</i> Rasio 90:10.....	70
<b>Gambar 4. 32</b> Tampilan Kata <i>Word Cloud</i> Positif .....	73
<b>Gambar 4. 33</b> Tampilan <i>Word Cloud</i> Negatif .....	74

## DAFTAR LAMPIRAN

<b>Lampiran 1.</b> Plagiarisme.....	87
<b>Lampiran 2.</b> Biodata Mahasiswa.....	88
<b>Lampiran 3.</b> Kartu Bimbingan Mahasiswa.....	89
<b>Lampiran 4.</b> Data Mentah Komentar Instagram.....	92

