

**ANALISIS SENTIMEN OPINI PUBLIK TERHADAP ISU KENAIKAN
GAJI ANGGOTA DPR PADA PLATFORM TIKTOK
MENGUNAKAN NAIVE BAYES DAN SUPPORT VECTOR
MACHINE**

SKRIPSI

Oleh:

**CALLISTO CARLOS
NIM. 221110837**



**PROGRAM STUDI S-1 TEKNIK INFORMATIKA
FAKULTAS INFORMATIKA
UNIVERSITAS MIKROSKIL
MEDAN
2026**

**SENTIMENT ANALYSIS OF PUBLIC OPINION TOWARD THE
ISSUE OF SALARY INCREASES FOR MEMBERS OF HOUSE OF
REPRESENTATIVES ON TIKTOK PLATFORM USING NAIVE
BAYES AND SUPPORT VECTOR MACHINE**

UNDERGRADUATE THESIS

By:

**CALLISTO CARLOS
ID NUMBER. 221110837**



**UNDERGRADUATE PROGRAM OF COMPUTER SCIENCE
FACULTY OF INFORMATICS
UNIVERSITAS MIKROSKIL
MEDAN
2026**

LEMBARAN PENGESAHAN

**ANALISIS SENTIMEN OPINI PUBLIK TERHADAP ISU KENAIKAN
GAJI ANGGOTA DPR PADA PLATFORM TIKTOK
MENGUNAKAN NAIVE BAYES DAN SUPPORT VECTOR
MACHINE**

SKRIPSI

Diajukan untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Sarjana
Program Studi S-1 Teknik Informatika

Oleh:

CALLISTO CARLOS
NIM. 221110837

Disetujui Oleh:

Dosen Pembimbing I

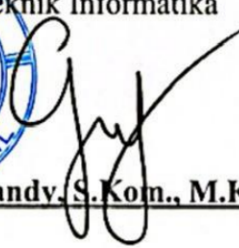

Ng Poi Wong, S.Kom., M.T.I.

Dosen Pembimbing II


Hernawati Gohzali, S.kom., M.T.I.

Medan, 14 Juli 2026

Diketahui dan Disahkan Oleh:

Ketua Program Studi
S-1 Teknik Informatika

Charles Juliandy, S.Kom., M.Kom.

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Teknik Informatika Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Callisto Carlos

NIM : 221110837

Saya telah melaksanakan penelitian dan penulisan Skripsi dengan judul dan tempat penelitian sebagai berikut:

Judul Skripsi : Analisis Sentimen Opini Publik terhadap Isu Kenaikan Gaji Anggota DPR pada Platform TikTok Menggunakan *Naïve Bayes* dan *Support Vector Machine*

Tempat Penelitian : Universitas Mikroskil

Alamat Tempat Penelitian : Jl. M.H Thamrin No.140, Pusat Ps., Kec. Medan Kota, Kota Medan, Sumatera Utara 20212

No. Telp. Tempat Penelitian : (061) 4573767

Sehubungan dengan Skripsi tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Skripsi tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Skripsi saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Skripsi saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 04 Juli 2026

Saya yang membuat pernyataan,



Callisto Carlos

ANALISIS SENTIMEN OPINI PUBLIK TERHADAP ISU KENAIKAN GAJI ANGGOTA DPR PADA PLATFORM TIKTOK MENGGUNAKAN NAIVE BAYES DAN SUPPORT VECTOR MACHINE

ABSTRAK

Perkembangan media sosial menjadikan TikTok sebagai salah satu sarana masyarakat untuk menyampaikan opini terhadap isu publik, termasuk isu kenaikan gaji anggota DPR. Banyaknya komentar yang tidak terstruktur menyebabkan analisis secara manual menjadi kurang efektif sehingga diperlukan analisis sentimen. Penelitian ini bertujuan membangun aplikasi analisis sentimen komentar TikTok menggunakan algoritma Naïve Bayes dan Support Vector Machine (SVM), serta membandingkan kinerja kedua algoritma berdasarkan nilai accuracy, precision, recall, dan F1-score. Penelitian menggunakan 7.000 komentar yang dibagi menjadi 5.600 data latih dan 1.400 data uji. Tahapan penelitian meliputi preprocessing teks, pelabelan menggunakan Lexicon InSet, ekstraksi fitur TF-IDF, penyeimbangan data dengan SMOTE, serta pelatihan dan pengujian model. Hasil penelitian menunjukkan bahwa SVM memperoleh nilai accuracy 87,14%, precision 87,16%, recall 87,14%, dan F1-score 87,15%, lebih baik dibandingkan Naïve Bayes yang memperoleh accuracy 56,57%, precision 55,47%, recall 56,57%, dan F1-score 49,56%. Hasil klasifikasi menunjukkan sentimen negatif mendominasi sebesar 52,71%, diikuti sentimen netral 40,43% dan positif 6,86%. Dengan demikian, algoritma SVM lebih efektif untuk analisis sentimen opini publik terkait isu kenaikan gaji anggota DPR pada platform TikTok.

Kata kunci: analisis sentimen, TikTok, Naïve Bayes, Support Vector Machine, TF-IDF.

ABSTRACT

The rapid growth of social media has made TikTok one of the primary platforms for expressing public opinions on various issues, including the issue of salary increases for members of the House of Representatives (DPR). The large volume of unstructured comments makes manual analysis inefficient, thereby requiring sentiment analysis. This study aims to develop a sentiment analysis application to classify TikTok comments related to the salary increase issue into positive, negative, and neutral sentiments using the Naïve Bayes and Support Vector Machine (SVM) algorithms, as well as to compare their performance based on accuracy, precision, recall, and F1-score. The study employed 7,000 TikTok comments, consisting of 5,600 training data and 1,400 testing data. The research process included text preprocessing, sentiment labeling using the InSet Lexicon, feature extraction using TF-IDF, data balancing with SMOTE, and model training and testing. The results showed that SVM outperformed Naïve Bayes, achieving an accuracy of 87.14%, precision of 87.16%, recall of 87.14%, and F1-score of 87.15%, while Naïve Bayes achieved an accuracy of 56.57%, precision of 55.47%, recall of 56.57%, and F1-score of 49.56%. The sentiment analysis results indicated that negative sentiment dominated with 52.71%, followed by neutral sentiment (40.43%) and positive sentiment (6.86%). Therefore, SVM proved to be more effective than Naïve Bayes for sentiment analysis of public opinion regarding the salary increase issue for DPR members on the TikTok platform.

Keywords: sentiment analysis, TikTok, Naïve Bayes, Support Vector Machine, TF-IDF.

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas rahmat dan karunianya sehingga penulis dapat menyelesaikan penyusunan Skripsi yang berjudul **"Analisis Sentimen Opini Publik terhadap Isu Kenaikan Gaji Anggota DPR pada Platform TikTok Menggunakan *Naïve Bayes* dan *Support Vector Machine*"** dengan baik.

Adapun tujuan penulisan Skripsi ini adalah sebagai salah satu syarat akademik dalam menyelesaikan pendidikan program Sarjana Strata Satu (S-1) pada Program Studi Teknik Informatika Universitas Mikroskil Medan.

Dalam penyusunan Skripsi ini, penulis telah berupaya semaksimal mungkin untuk menghasilkan karya yang terbaik. Skripsi ini tidak akan terselesaikan dengan baik tanpa adanya dukungan, bimbingan, serta doa dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis menyampaikan ucapan terima kasih kepada:

1. Bapak Ng Poi Wong, S.kom., M.T.I., selaku Dosen Pembimbing I sekaligus Dosen Wali.
2. Ibu Hernawati Gohzali, S.Kom., M.TI., selaku Dosen Pembimbing II.
3. Bapak Hardy, S.Kom., M.Sc., Ph.D., selaku Rektor Universitas Mikroskil Medan.
4. Bapak Sunaryo Winardi, S.Kom., M.T., selaku Dekan Fakultas Informatika Universitas Mikroskil Medan.
5. Bapak Carles Juliandy, S.Kom., M.Kom., selaku Ketua Program Studi S-1 Teknik Informatika Universitas Mikroskil Medan.
6. Bapak dan Ibu Dosen Universitas Mikroskil yang telah memberikan ilmunya selama menempuh Pendidikan di Universitas Mikroskil Medan.
7. Orang tua dan keluarga penulis yang senantiasa memberikan doa, dukungan, serta motivasi kepada penulis selama menempuh pendidikan hingga terselesaikannya Skripsi ini.
8. Seluruh pihak yang tidak dapat penulis sebutkan satu per satu, yang telah memberikan dukungan, dan doa selama proses penyusunan Skripsi ini.

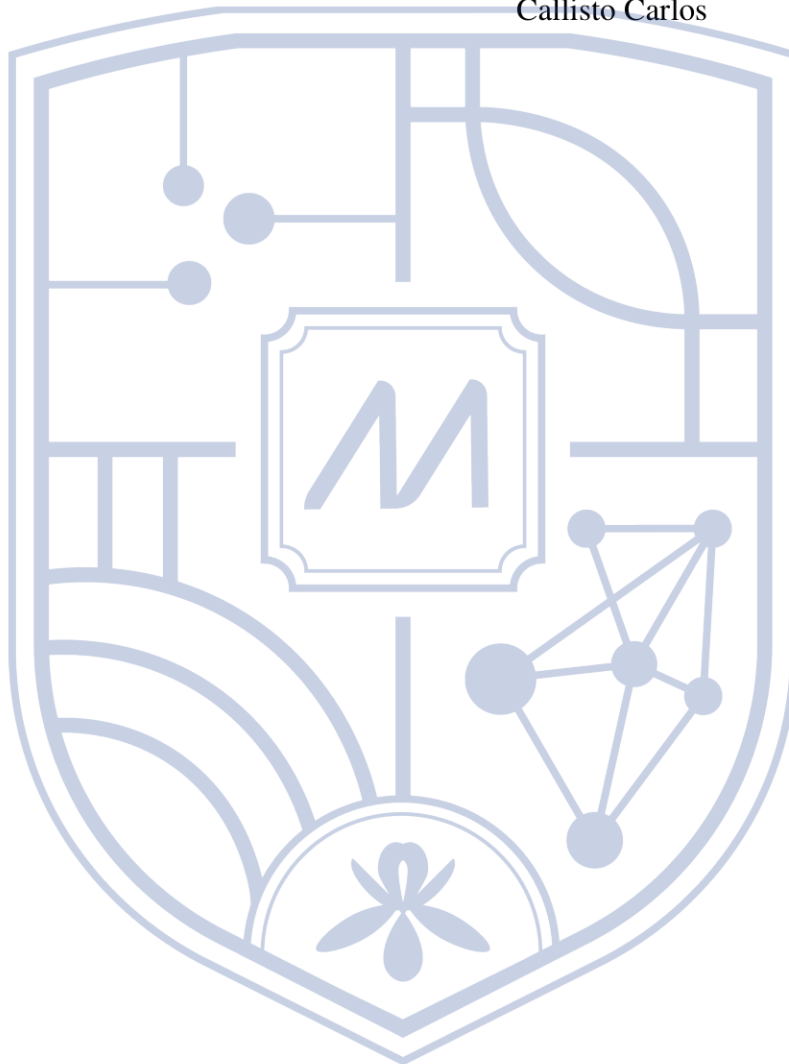
Penulis menyadari bahwa Skripsi ini masih memiliki keterbatasan dan belum sepenuhnya sempurna. Oleh karena itu, penulis mengharapkan saran dan kritik yang bersifat membangun sebagai bahan perbaikan di masa mendatang. Penulis berharap Skripsi ini dapat

memberikan manfaat serta menjadi referensi bagi pengembangan ilmu pengetahuan maupun bagi pihak-pihak yang memerlukannya.

Medan, 04 Juli 2026

Penulis,

Callisto Carlos



DAFTAR ISI

ABSTRAK.....	i
KATA PENGANTAR	ii
DAFTAR ISI	iv
DAFTAR GAMBAR	vi
DAFTAR TABEL	vii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan	3
1.4 Manfaat	3
1.5 Ruang Lingkup	3
BAB II KAJIAN LITERATUR.....	5
2.1 <i>Artificial Intelligence (AI)</i>	5
2.1.1 <i>Machine Learning</i>	5
2.1.2 <i>Natural Language Processing (NLP)</i>	6
2.2 <i>Text Mining</i>	6
2.3 <i>Pembagian Data</i>	7
2.4 <i>Preprocessing</i>	7
2.5 <i>Pelabelan Sentimen</i>	8
2.6 <i>Ekstraksi Fitur</i>	9
2.7 <i>Klasifikasi</i>	10
2.8 <i>Naïve Bayes</i>	11
2.9 <i>Support Vector Machine (SVM)</i>	13
2.10 <i>Analisis Sentimen</i>	14
2.11 <i>Confusion Matrix</i>	16

2.12 Metode Pengembangan <i>Waterfall</i>	19
2.13 <i>Synthetic Minority Oversampling Technique</i> (SMOTE)	20
2.14 Dewan Perwakilan Rakyat (DPR)	21
2.15 Aplikasi TikTok	22
BAB III TAHAPAN PELAKSANAAN	24
3.1 Tahapan Pelaksanaan	24
3.2 Analisis	25
3.2.1 Analisis Proses	25
3.2.2 Analisis Kebutuhan Fungsional	56
3.2.3 Analisis kebutuhan Non-Fungsional	60
3.3 Perancangan	60
BAB IV HASIL DAN PEMBAHASAN	64
4.1 Hasil	64
4.2 Pembahasan.....	68
4.2.1 Persiapan Pengujian	68
4.2.2 Hasil Pengujian	68
BAB V PENUTUP	79
5.1 Kesimpulan	79
5.2 Saran	80
DAFTAR PUSTAKA	81

DAFTAR GAMBAR

Gambar 2. 1 <i>Support Vector Machine</i> [40]	13
Gambar 3. 1 <i>Flowchart</i> proses (a) <i>training</i> , (b) <i>testing</i>	25
Gambar 3. 2 <i>Flowchart</i> Tahapan <i>Preprocessing</i>	26
Gambar 3. 3 <i>Use Case Diagram</i>	57
Gambar 3. 4 Rancangan Tampilan Halaman Utama Aplikasi	61
Gambar 3. 5 Rancangan Tampilan Halaman <i>Training</i>	61
Gambar 3. 6 Rancangan Tampilan Halaman <i>Testing</i>	62
Gambar 3. 7 Lanjutan Rancangan Tampilan Halaman <i>Testing</i>	63
Gambar 4. 1 Tampilan Hasil Halaman Utama	64
Gambar 4. 2 Tampilan Hasil Halaman <i>Training</i>	65
Gambar 4. 3 Tampilan Hasil Halaman <i>Testing</i>	65
Gambar 4. 4 Kode <i>Frontend</i> Aplikasi	66
Gambar 4. 5 Kode <i>pipeline</i> Model	67
Gambar 4. 6 Kode <i>Backend</i> Aplikasi	67

DAFTAR TABEL

Tabel 2. 1 <i>Confusion Matrix</i> Tiga Kelas	16
Tabel 2. 2 Elemen <i>Confusion Matrix</i> Untuk Kelas A	17
Tabel 2. 3 Elemen <i>Confusion Matrix</i> Untuk Kelas B.....	17
Tabel 2. 4 Elemen <i>Confusion Matrix</i> Untuk Kelas C.....	17
Tabel 3. 1 Tahapan <i>Lowercasing</i>	26
Tabel 3. 2 Tahapan <i>Remove URL</i>	27
Tabel 3. 3 Tahapan <i>Remove Mention, Hashtag</i>	27
Tabel 3. 4 Tahapan <i>Remove Punctuation</i>	28
Tabel 3. 5 Tahapan <i>Remove Number</i>	29
Tabel 3. 6 Tahapan <i>Normalize Space</i>	29
Tabel 3. 7 Tahapan <i>Slang Word Normalization</i>	30
Tabel 3. 8 Tahapan <i>Stopword Removal</i>	31
Tabel 3. 9 Tahapan <i>Tokenization</i>	31
Tabel 3. 10 Tahapan <i>Tokenization</i>	32
Tabel 3. 11 Contoh <i>Labeling</i>	33
Tabel 3. 12 Nilai <i>Term Frequency (TF)</i>	34
Tabel 3. 13 Nilai <i>Inverse Document Frequency (IDF)</i>	37
Tabel 3. 14 Hasil perhitungan nilai TF-IDF.....	38
Tabel 3. 15 Sebelum dan Setelah SMOTE.....	41
Tabel 3. 16 Contoh data <i>Training</i>	41
Tabel 3. 17 Contoh data <i>Testing</i>	42
Tabel 3. 18 Frekuensi kemunculan data pada kategori label	43
Tabel 3. 19 Nilai $P(X_i C_j)$	45
Tabel 3. 20 Perbandingan Hasil Nilai PC_jX	48
Tabel 3. 21 Tabel Data Latih SVM	49
Tabel 3. 22 Daftar Fitur Hasil Reduksi Dimensi	50
Tabel 3. 23 Vektor Ringkasan Setiap Dokumen	50
Tabel 3. 24 Pembentukan <i>Classifier One-vs-Rest (OvR)</i> untuk Tiga Kelas	51
Tabel 3. 25 Data untuk SVM1	51
Tabel 3. 26 Data untuk SVM2	52

Tabel 3. 27 Data untuk SVM3	53
Tabel 3. 28 Ringkasan Ketiga <i>Classifier</i>	54
Tabel 3. 29 Nilai fungsi keputusan	55
Tabel 3. 30 Penjelasan <i>Use Case Diagram</i>	56
Tabel 3. 31 Skenario <i>Use Case</i> : Melakukan <i>Preprocessing</i>	57
Tabel 3. 32 Skenario <i>Use Case</i> : <i>Training Data</i>	57
Tabel 3. 33 Skenario <i>Use Case</i> : <i>Testing Data</i>	58
Tabel 3. 34 Kebutuhan Non-Fungsional Sistem	59
Tabel 4. 1 Distribusi Sentimen <i>Naïve Bayes</i>	68
Tabel 4. 2 <i>Confusion Matrix Naïve Bayes</i>	68
Tabel 4. 3 Metrik Evaluasi Setiap Label <i>Naïve Bayes</i>	71
Tabel 4. 4 Metrik Evaluasi Keseluruhan <i>Naïve Bayes</i>	71
Tabel 4. 5 Distribusi Sentimen SVM	72
Tabel 4. 6 <i>Confusion Matrix SVM</i>	72
Tabel 4. 7 Metrik Evaluasi Setiap Label SVM	75
Tabel 4. 8 Metrik Evaluasi Keseluruhan SVM	75
Tabel 4. 9 Perbandingan Hasil Pengujian Kedua Algoritma	76
Tabel 4. 10 Hasil Analisis Sentimen	76