

SKRIPSI



**SENTIMEN ANALISIS OPINI MASYARAKAT JAKARTA PADA KINERJA
PEMERINTAH JAKARTA TERHADAP ISU TENGGELAMNYA JAKARTA
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE**

**ADITHYA KHARISMA
1910511039**

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA
2023**

SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Komputer



**SENTIMEN ANALISIS OPINI MASYARAKAT JAKARTA PADA KINERJA
PEMERINTAH JAKARTA TERHADAP ISU TENGGELAMNYA JAKARTA
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE**

**ADITHYA KHARISMA
1910511039**

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA
2023**

PERNYATAAN ORISINALITAS

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

Nama : Adithya Kharisma

NIM : 1910511039

Tanggal : 27 Juni 2023

Judul Skripsi : **Sentimen Analisis Opini Masyarakat Jakarta pada Kinerja Pemerintah Jakarta Terhadap Isu Tenggelamnya Jakarta Menggunakan Algoritma *Support Vector Machine***

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

Jakarta, 27 Juni 2023

Yang Menyatakan,



Adithya Kharisma

**PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH
UNTUK KEPENTINGAN AKADEMIS**

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

Nama : Adithya Kharisma
NIM : 1910511039
Fakultas : Ilmu Komputer
Program Studi : Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan karya ilmiah saya kepada Universitas Pembangunan Nasional “Veteran” Jakarta Hak Bebas Royalti Non-Eksklusif (*Non-Exchange Royalty Free Right*) untuk dipublikasikan dengan judul:

Sentimen Analisis Opini Masyarakat Jakarta pada Kinerja Pemerintah Jakarta Terhadap Isu Tenggelamnya Jakarta Menggunakan Algoritma Support Vector Machine

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional “Veteran” Jakarta berhak menyimpan, mengalih media atau memformatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis atau pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 27 Juni 2023

Yang Menyatakan,



Adithya Kharisma

LEMBAR PENGESAHAN

Tugas Akhir ini diajukan oleh:

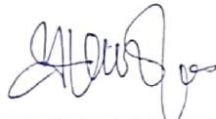
Nama : Adithya Kharisma
NIM : 1910511039
Program Studi : S1 Informatika
Judul Tugas Akhir : Sentimen Analisis Opini Masyarakat Jakarta pada
Kinerja Pemerintah Jakarta Terhadap Isu Tenggelamnya
Jakarta Menggunakan Algoritma *Support Vector Machine*

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



Dr. Ernatita, M.Kom.

Penguji I



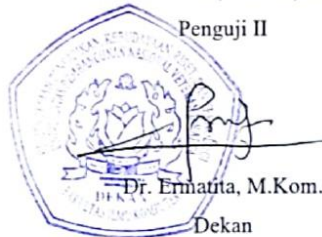
Yuni Widiastiwati, S.Kom, M.Si.

Penguji II




Iin Ernawati, S.kom, M.Si.

Pembimbing



Dr. Ernatita, M.Kom.

Dekan



Dr. Widya Cholil, M.I.T.

Kepala Program Studi

Ditetapkan di : Jakarta
Tanggal Ujian : Selasa, 27 Juni 2023



Abstrak

Berita mengenai isu tenggelamnya Jakarta sudah bukan menjadi hal yang baru lagi. Isu tenggelamnya Jakarta menimbulkan berbagai banyak opini masyarakat Indonesia terutama masyarakat Jakarta. Opini tersebut menimbulkan pro dan kontra di kalangan masyarakat, Opini tersebut dapat dijadikan sebagai ukuran untuk melihat kepuasan masyarakat terhadap kinerja pemerintah Jakarta, karena opini tersebut sangatlah banyak, hal ini menimbulkan permasalahan yaitu bagaimana membuat opini tersebut menjadi sebuah data yang bermanfaat untuk model *support vector machine*, oleh karena itu dilakukan analisis sentimen untuk melihat kepuasan masyarakat terhadap kinerja pemerintah. Analisis sentimen yang dilakukan memiliki 2 label yaitu label positif dan label negatif. Sentimen tersebut kemudian diolah menggunakan *Natural Language Processing* atau *text mining* dan dibuat menjadi 3 buah dataset yaitu dataset awal, dataset *undersampling* dan dataset *oversampling*. Kemudian ketiga dataset tersebut dijadikan data untuk pembuatan model *support vector machine* dengan parameter *cost* dan kernel yang berbeda, setelah dilakukan pembuatan model kemudian dilakukan evaluasi menggunakan *confussion matrix*. Hasil evaluasi tertinggi dari 3 model dataset tersebut adalah dataset *oversampling* dengan nilai rata-rata akurasi sebesar 0,9359, *precision* sebesar 0,9676, *recall* sebesar 0,9003, *specificity* sebesar 0,9703 dan *f1-score* sebesar 0,9326.

Kata Kunci : Sentimen Analisis, Isu Jakarta Tenggelam, *Support Vector Machine*, *Text Mining*, *confussion matrix*.

Abstract

The news about the sinking issue of Jakarta is no longer something new. The sinking issue of Jakarta has sparked various opinions among Indonesian society, especially the people of Jakarta. These opinions have generated pros and cons among the public. Such opinions can be used as a measure to gauge the satisfaction of the people regarding the performance of the Jakarta government. Due to the abundance of opinions, it poses a challenge of how to transform these opinions into useful data for the support vector machine model. Therefore, sentiment analysis is conducted to assess the satisfaction of the people towards the government's performance. The sentiment analysis consists of two labels, namely positive and negative. The sentiments are then processed using Natural Language Processing or text mining and transformed into three datasets: the original dataset, undersampling dataset, and oversampling dataset. These three datasets are used to build support vector machine models with different parameter values for cost and kernel. After building the models, evaluation is performed using a confusion matrix. The highest evaluation result among the three dataset models is the oversampling dataset with an average accuracy of 0.9359, precision of 0.9676, recall of 0.9003, specificity of 0.9703, and F1-score of 0.9326..

Keywords: Sentiment Analysis, Jakarta Sinking Issue, Support Vector Machine, Text Mining, confusion matrix.

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala karunia-Nya sehingga karya ini dapat diselesaikan dengan baik. Judul karya dalam penelitian, yang telah berlangsung sejak Januari 2023 ini adalah “Sentimen Analisis Opini Masyarakat Jakarta Pada Kinerja Pemerintah Jakarta Terhadap Isu Tenggelamnya Jakarta Menggunakan Algoritma *Support Vector Machine*”. Penulis menyadari dalam penyusunan laporan skripsi ini tidak akan selesai tanpa bantuan dari berbagai pihak. Karena itu pada kesempatan ini penulis ingin mengucapkan terimakasih kepada:

1. Ibu Dr. Ermatita, M.Kom selaku Dekan Fakultas Ilmu Komputer.
2. Ibu Dr Widya Cholil, S.Kom, M.I.T. selaku Kepala Program Studi Informatika.
3. Ibu Iin Ernawati, S.Kom., M.Si selaku dosen pembimbing skripsi yang membantu penulis dalam penyusunan skripsi dan memberikan saran dengan baik.
4. Orang tua dan keluarga penulis, yang telah memberikan dukungan dan doa untuk kesuksesan penulis dalam menyelesaikan skripsi.
5. Seluruh teman dekat penulis yang telah memberikan dukungan, doa, dan semangat untuk menyelesaikan skripsi ini, yang tidak dapat penulis sebutkan satu per satu.
6. Teman-teman Program Studi Informatika Angkatan 2019 Fakultas Ilmu Komputer yang telah memberikan dukungan dan doa satu sama lain.
7. Para responden dari berbagai wilayah di provinsi DKI Jakarta
8. Seluruh pihak yang terlibat dalam kelancaran pembuatan skripsi ini dan yang belum disebutkan di atas, penulis ucapkan terimakasih.

Jakarta, 27 Juni 2023

Penulis

DAFTAR ISI

SKRIPSI.....	II
Abstrak.....	VI
<i>Abstract</i>	VII
DAFTAR ISI.....	IX
DAFTAR TABEL.....	XI
DAFTAR GAMBAR.....	XIII
DAFTAR LAMPIRAN.....	XIV
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan Penelitian.....	2
1.4 Manfaat Penelitian.....	2
1.5 Ruang Lingkup.....	3
1.6 Luaran Yang Diharapkan.....	3
1.7 Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA.....	5
2.1 Kuesioner.....	5
2.2 Python.....	6
2.3 <i>Text Mining</i>	7
2.4 Sentimen Analisis.....	8
2.5 <i>Natural Language Processing (NLP)</i>	8
2.6 <i>Preprocessing Text</i>	9
2.7 <i>Imbalance Data</i>	9
2.8 <i>Text Transformation</i>	10
2.9 <i>Support Vector Machine</i>	11
2.10 <i>Evaluasi Model</i>	13
2.11 Penelitian Terdahulu.....	14
BAB III 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 Praproses Data.....	17
3.1.4.1 <i>Exploratory Data</i>	17

3.1.4.2	<i>CaseFolding</i>	17
3.1.4.3	<i>Remove Punctuation</i>	18
3.1.4.4	Tokenizing	18
3.1.4.5	Stopword Removal	18
3.1.4.6	Stemming	18
3.1.5	Pembobotan Kata	18
3.1.6	Pembagian Data Latih dan Data Uji	18
3.1.7	Pembuatan Model <i>Support Vector Machine</i>	19
3.1.8	Evaluasi Confussion Matrix	19
3.2	Alat Bantu Penelitian	19
3.3	Jadwal Penelitian	20
BAB IV HASIL DAN PEMBAHASAN		21
4.1	Data	21
4.2	Praproses Data	- 24 -
4.3	<i>Text Transformation</i>	- 41 -
4.4	Proses Klasifikasi	- 44 -
4.5	Evaluasi Model <i>Support Vector Machine</i>	- 49 -
BAB V PENUTUP		- 58 -
5.1	Kesimpulan	- 58 -
5.2	Saran	- 59 -
DAFTAR PUSTAKA		60
RIWAYAT HIDUP		62
LAMPIRAN		63

DAFTAR TABEL

Tabel 2.1 Hasil Prediksi confusion matrix.....	13
Tabel 2.2 <i>Confussion Matrix</i>	13
Tabel 3.1 Tahapan Perencanaan.....	19
Tabel 4.1 Data Kuesioner Sentimen Masyarakat Jakarta	22
Tabel 4.2 Penyeleksian Kolom Data.....	24
Tabel 4.3 <i>Shape Data</i>	25
Tabel 4.4 Informasi data dan tipe data.....	25
Tabel 4.5 Data hitung duplikat.....	25
Tabel 4.6 Persebaran Label Sentimen pada Data.....	28
Tabel 4.7 Persebaran Label Sentimen pada Data Setelah Balancing Data	29
Tabel 4.8 Tabel Contoh Case Folding	32
Tabel 4.9 Tabel Tanda Baca Punctuation	33
Tabel 4.10 Contoh Hapus Tanda Baca.....	33
Tabel 4.11 Tabel Contoh Tokenization.....	35
Tabel 4.12 Tabel Stoplist Sastrawi	36
Tabel 4.13 Tabel Tambahan Stoplist	38
Tabel 4.14 Tabel Contoh Stopword Removal.....	38
Tabel 4.15 Tabel Stemming dataset.....	40
Tabel 4.16 Dokumen Teks TF-IDF.....	42
Tabel 4.17 Tabel Perhitungan TF-IDF.....	42
Tabel 4.18 Pembagian Data Uji dan Data Latih	44
Tabel 4.19 Parameter Model SVM Data Awal	45
Tabel 4.20 Parameter Model SVM Data Undersampling	46
Tabel 4.21 Parameter Model SVM Data Oversampling	47
Tabel 4.22 Contoh dataset 2 fitur.....	48
Tabel 4.23 Confusion Matrix Dataset Awal perbandingan 80:20	50
Tabel 4.24 Confusion Matrix Dataset Awal perbandingan 70:30	50
Tabel 4.25 Confusion Matrix Dataset Undersampling perbandingan 80:20	51
Tabel 4.26 Confusion Matrix Dataset Undersampling perbandingan 70:30	51
Tabel 4.27 Confusion Matrix Dataset Oversampling perbandingan 80:20	52
Tabel 4.28 Confusion Matrix Dataset Oversampling perbandingan 70:30	53
Tabel 4.29 Hasil Evaluasi Dataset awal perbandingan 80:20.....	54
Tabel 4.30 Hasil Evaluasi Dataset Undersampling perbandingan 80:20.....	54
Tabel 4.31 Hasil Evaluasi Dataset Oversampling perbandingan 80:20.....	55

Tabel 4.32 Hasil Evaluasi Dataset awal perbandingan 70:30.....	56
Tabel 4.33 Hasil Evaluasi Dataset Undersampling perbandingan 70:30.....	56
Tabel 4.34 Hasil Evaluasi Dataset Oversampling perbandingan 70:30.....	57

DAFTAR GAMBAR

Gambar 2.1 Proses dari Text Mining (Data Flair, 2018)	7
Gambar 2.2 Hyperplane Support Vector Machine (Husada, H. C., 2021)	11
Gambar 2.3 Visual Kernel Trik SVM (Medium.com, 2021).....	11
Gambar 3.1 Tahapan Penelitian	15
Gambar 4.1 Visual Responden berdasarkan Gender	26
Gambar 4.2 Visual Responden Berdasarkan Asal Kota	26
Gambar 4.3 Visual Responden mengetahui jakarta tenggelam	27
Gambar 4.4 Visual Kepuasan Responden Secara Umum	27
Gambar 4.5 Kepuasan Responden pada Isu Jakarta Tenggelam	28
Gambar 4.6 Wordcloud Dataset.....	30
Gambar 4.7 WordCloud Sentimen Positif	30
Gambar 4.8 WordCloud Sentimen Negatif.....	31
Gambar 4.9 Worcloud praproses sentimen positif.....	41
Gambar 4.10 Worcloud praproses sentimen negatif.....	41

DAFTAR LAMPIRAN

Lampiran	62
LAMPIRAN 1 Data Kuesioner	63
LAMPIRAN 2 PERBANDINGAN HASIL EVALUASI MODEL	65
LAMPIRAN 3 Hasil Turnitin	66