

TEXT MINING DAN SENTIMENT ANALYSIS MEMBANDINGKAN

METODE NAÏVE BAYES (NB)

CLASSIFIER DAN SUPPORT VECTOR MACHINE (SVM)

(Studi Kasus: Data Tanggapan Mengenai Presiden Indonesia Bapak Jokowi dan pemerintahanya di Tahun 2022 melalui Media Sosial *Twitter*)

Skripsi

Untuk memenuhi sebagai persyaratan

Mencapai derajat Sarjana S-1



Disusun oleh:

Alvira Yusnia Pradani

20170140116

**PROGRAM STUDI TEKNOLOGI INFORMASI FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH YOGYAKARTA**

2023

PERNYATAAN

Pernyataan ini saya buat dengan sungguh-sungguh bahwa skripsi ini tidak mengandung plagiarisme dan merupakan hasil pemikiran serta penelitian saya sendiri. Saya telah sepenuhnya mematuhi etika penelitian dan akademik yang berlaku di Universitas Muhammadiyah Yogyakarta.

Jika dikemudian hari terjadi tindakan plagiarisem, saya bersedia untuk diberikan sanksi yang berlaku di Universitas Muhammadiyah Yogyakarta.

Yogyakarta, 22 Mei 2023



Alvira Yusnia P

KATA PENGANTAR

Puji syukur saya panjatkan ke hadirat Allah SWT, karena atas rahmat dan hidayah-Nya, skripsi dengan judul "TEXT MINING DAN SENTIMENT ANALYSIS MEMBANDINGKAN METODE NAÏVE BAYES (NB) CLASSIFIER DAN SUPPORT VECTOR MACHINE (SVM)" dapat terselesaikan dengan baik. Penulisan skripsi ini sebagai salah satu syarat untuk menyelesaikan program sarjana Strata-1 pada Program Studi Informatika, Fakultas Teknik, Universitas Muhammadiyah Yogyakarta. Tak lupa saya ucapkan terima kasih kepada Bapak Haris Setyawan, S.T., M.Eng selaku dosen pembimbing skripsi saya, yang telah memberikan arahan, bimbingan, dan dukungan dalam proses penulisan skripsi ini. Ibu Laila Ma'rifatul Azzizah, S.Kom, M.I.M. yang telah dengan sabar memberikan masukan yang berharga sehingga skripsi ini dapat terselesaikan dengan baik.

Tak lupa juga kami ucapkan terima kasih kepada seluruh dosen pengajar di Program Studi Informatika Fakultas Teknik, yang telah memberikan ilmu pengetahuan dan wawasan yang mendalam kepada saya selama masa perkuliahan. Pengalaman dan pembelajaran yang kami peroleh dari para dosen telah memberikan kontribusi besar dalam penulisan skripsi ini.

Saya juga mengucapkan terima kasih kepada seluruh teman-teman kuliah saya, yang selalu memberikan dukungan, semangat, dan bantuan selama proses penulisan skripsi ini. Kerjasama dan persahabatan yang terjalin di antara saya dengan Regitha, Lulu, Pertiwi Larasati, serta saudara saya mba Anggie Sukma dan Mas Bachtiar serta Mas Rama yang membantu saya mengerjakan skripsi ini. Tak lupa juga doa dari orang tua saya yang sudah membesarkan saya dan sangat berpengaruh terhadap penyelesaian hasil skripsi ini dan mereka juga selalu menyemangati saya untuk bisa menyelesaikan skripsi ini serta yang menjadi salah satu faktor penunjang keberhasilan penulisan skripsi ini.

Selanjutnya, kami juga mengucapkan terima kasih kepada pihak-pihak yang tidak dapat kami sebutkan satu per satu, namun telah memberikan dukungan dan doa dalam penulisan skripsi ini.

Akhir kata, semoga skripsi ini dapat memberikan manfaat dan sumbangan ilmu pengetahuan bagi perkembangan ilmu informatika dan teknologi informasi. Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, oleh karena itu, kritik dan saran yang membangun sangat kami harapkan guna perbaikan di masa yang akan datang.

Yogyakarta, 22 Mei 2023



Alvira Yusnia P

DAFTAR ISI

HALAMAN PENGESAHAN I.....	II
HALAMAN PENGESAHAN II	III
PERNYATAAN.....	IV
KATA PENGANTAR	V
Abstrak.....	VII
Abstract.....	VIII
DAFTAR ISI	IX
DAFTAR GAMBAR	XII
DAFTAR TABLE	1
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan Penelitian	2
1.4 Batasan Masalah	3
1.5 Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA.....	4
2.1 Kajian Pustaka.....	4
2.2 Landasan Teori.....	12
2.2.1 Pre-processing text	13

2.2.2 Term Frequency-Inverse Document Frequency ((TF-IDF)).....	14
2.2.3 Algoritma Navie Bayes	15
2.2.4 Support Vector Machine (SVM)	16
2.2.5 Evaluasi klasifikasi	18
2.2.6 Analisis Sentimen	20
2.2.7 Text Mining.....	21
2.2.8 Twitter	21
2.2.9 Netlytic	22
2.2.10 Visualisasi Data	22
2.2.11 Python.....	23
2.2.12 Pemerintahan	24
2.2.13 Presiden.....	25
BAB III METODE PENELITIAN.....	26
3.1 Populasi dan Sample.....	26
3.2 Variabel Penelitian.....	26
3.3 Metode Pengambilan Data	26
3.4 Tahap penelitian.....	28
3.4.1 Crawling data	28
3.4.2 Labelling	28
3.4.3 Data Cleaning.....	29
3.4.4 Stopword Removal	30
3.4.5 Tokenization	31
3.4.6 Stemming	32
3.4.7 Klasifikasi Naive Bayes	33
3.4.8 Klasifikasi Support Vector Machine (SVM)	34
3.4.9 Evaluasi Hasil	38
3.4.10 Sentiment analisis	38
BAB IV PEMBAHASAN DAN HASIL	40

4.1.1 Cleaning	40
4.1.2 Tahapan kedua..adalah tahap..Remove Stopword.....	40
4.1.3 Proses ekstraksi.....	41
4.1.4 Klasifikasi	41
4.1.5 Uji model.....	41
4.2 Pembahasan dan hasil Pengambilan Data (Scrapping data Twitter)	42
4.3 Pembahasan Labelling Data	44
4.4 Pembahasan Tahapan Proses Preprocessing	47
4.5 Hasil Model evaluasi metode Naïve Bayes.....	51
4.6 Hasil evaluasi metode Support Vector Machine (SVM)	53
BAB V	56
KESIMPULAN	56
5.1 Kesimpulan.....	56
5.2 SARAN.....	56
DAFTAR PUSTAKA	57
LAMPIRAN	59

DAFTAR GAMBAR

Gambar 1 <i>Hyperplane</i> terbaik yang memisahkan class-1 dan class +1.....	17
Gambar2 Alur pengerjaan / flowchart	28
Gambar 3 Flowchart Proses Pembuangan Stopword.....	31
Gambar 4 Proses Tokenization.....	32
Gambar 5 Flowchart Proses Stemming.....	33
Gambar 7 visualisasi kalsifikasi SVM.....	36
Gambar 8 Klasifikasi matlab	37
Gambar 9 Tahap Scrapping menggunakan Netlytic	43
Gambar10 Data hasil crawling data tweet menggunakan Netlyitic	44
Gambar11 Diagram hasil labelling	45
Gambar12 Hasil Accuracy model NB	52
Gambar 13 Hasil Matrix menggunakan metode Navie Bayes	53
Gambar 14 Hasil Accuracy model SVM.....	54
Gambar 15 Hasil Matrix menggunakan metode SVM.....	55

DAFTAR TABLE

Table 1 Mengenai penelitian yang relevan	4
Table 2 Definisi Twitter	26
Table 3 Table proses labelling.....	45
Table 4 Proses Cleaning.....	47
Table 5 Melakukan tahapan Remove Stopword.....	48
Table 6 Tahapan Tokenization	49
Table 7 Proses Stemming.....	49