

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PRESIDEN
JOKOWI MENGGUNAKAN ALGORITMA *RANDOM FOREST***

SKRIPSI



disusun oleh:

Stephen Lambok Marbun

15.11.9275

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2018**



**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PRESIDEN
JOKOWI MENGGUNAKAN ALGORITMA *RANDOM FOREST***

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar sarjana
pada Program Studi Informatika



disusun oleh:
Stephen Lambok Marbun
15.11.9275

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2018**

i



PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PRESIDEN
JOKOWI MENGGUNAKAN ALGORITMA RANDOM FOREST**

yang dipersiapkan dan disusun oleh

Stephen Lambok Marbun

15.11.9275

telah disetujui oleh Dosen Pembimbing Skripsi

pada tanggal 26 September 2018

Dosen Pembimbing,



Mardhiya Hayaty, S.T, M.Kom.

NIK. 190302108

PENGESAHAN

SKRIPSI

ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PRESIDEN JOKOWI MENGGUNAKAN ALGORITMA RANDOM FOREST

yang dipersiapkan dan disusun oleh
Stephen Lambok Marbun

15.11.9275

telah dipertahankan di depan Dewan Penguji
pada tanggal 27 Desember 2018

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Dony Ariyus, M.Kom
NIK. 190302128



Sumarni Adi, S.Kom., M.Cs
NIK. 190302256



Lilis Dwi Farida, S.Kom., M.Eng
NIK. 190302288



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 27 Desember 2018

DEKAN FAKULTAS ILMU KOMPUTER



Krisnawati, S.Si., M.T
NIK. 190302038

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 diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

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

Yogyakarta, 28 Desember 2018

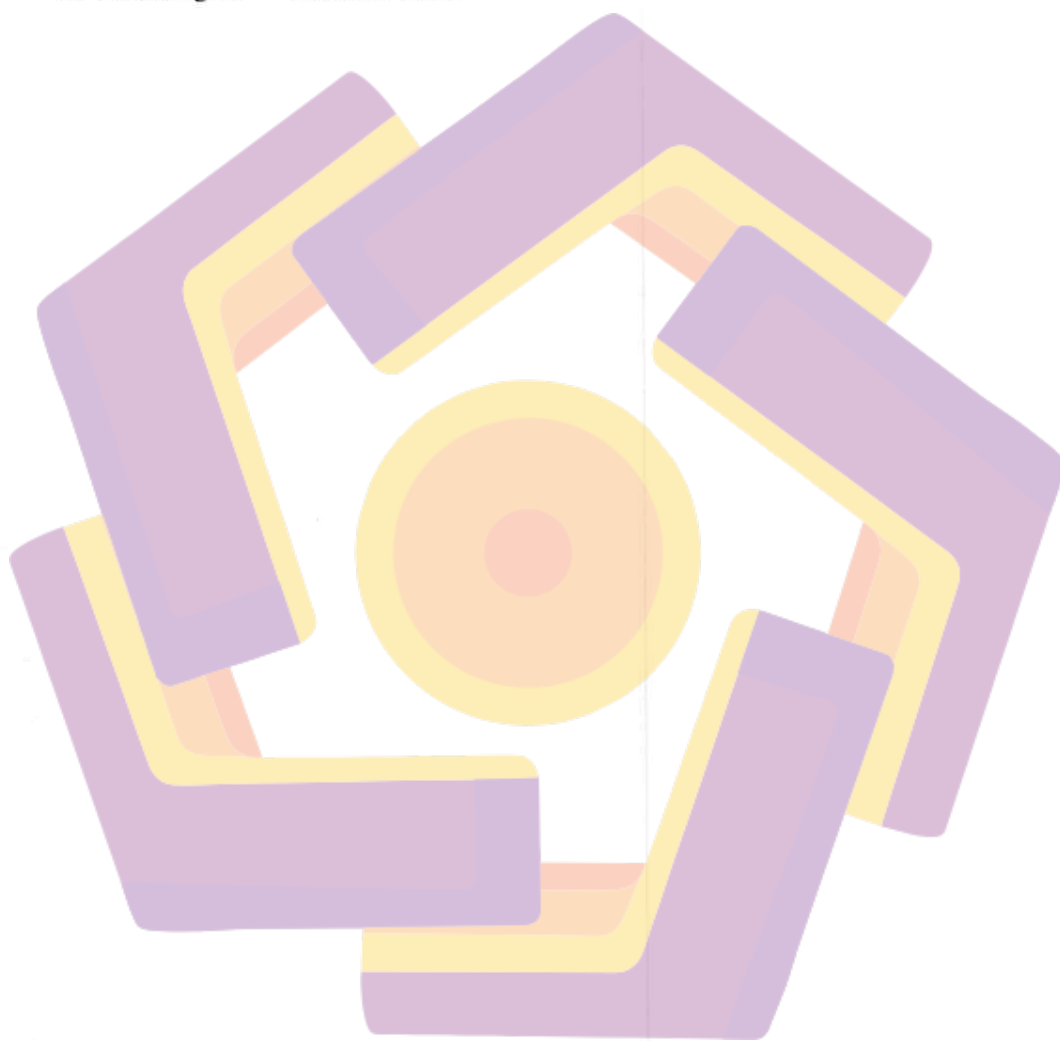


Stephen Lambok Marbun

NIM. 15.11.9275

MOTTO

“Ya Tuhan, bukit batuku, kubu pertahananku dan penyelamatku,
Allahku, gunung batuku, tempat aku berlindung, perisaiku, tanduk keselamatanku,
kota bentengku!” – Mazmur 18:2.



PERSEMBAHAN

Puji Tuhan atas segala berkat yang diberikan Tuhan Yesus sehingga penelitian ini dapat dilakukan dan diselesaikan dengan sebaik-baiknya. Saya juga ucapkan terimakasih untuk dukungan dan bantuan semua pihak yang membantu selesainya penelitian ini.

Skripsi ini saya persembahkan untuk kampus saya, Universitas AMIKOM Yogyakarta.

Untuk orang tua saya, saya ucapkan terimakasih sebesar-besarnya. Semoga saya bisa menjadi anak yang tetap berbakti.

Untuk ibu Mardhiya, yang sudah dengan sabar membimbing saya. Saya ucapkan terimakasih sedalam-dalamnya.

Dan untuk semua pihak yang telah banyak membantu dalam penyusunan skripsi ini yang tidak dapat saya sebutkan satu persatu.

KATA PENGANTAR

Puji syukur penulis haturkan atas berkat dari Tuhan yang telah memberikan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “Analisis Sentimen Pengguna Twitter Terhadap Presiden Jokowi Menggunakan Algoritma Random Forest” dengan sebaik-baiknya

Dengan selesainya skripsi ini, penulis mengucapkan banyak terimakasih kepada:

1. Bapak Prof. Dr. M. Suyanto selaku Rektor Universitas AMIKOM Yogyakarta
2. Bapak Sudarmawan, MT. selaku Dekan Fakultas Sains dan Teknologi, dan Ketua Program Studi S1 Informatika
3. Ibu Mardhiya, S.T., M.Kom. selaku pembimbing yang telah dengan sabar membimbing dan banyak memberikan kontribusi bagi penulis dalam pembuatan skripsi ini.
4. Para Dosen dan Staff Universitas AMIKOM Yogyakarta yang telah banyak memberi bantuan informasi, pengalaman dan kontribusi lainnya selama penulis berkuliah di kampus ini hingga terselesaikannya skripsi ini.
5. Orang tua saya, Robinson Marbun dan Elfridawati Sihombing, dan saudara kandung saya, Samuel Gerry Marbun, Selly Tasya Mauli Marbun, dan Sadrakh Martua Marbun, yang telah memberikan dukungan terbaiknya selama saya hidup.
6. Teman-teman grup “NONGKI” yang sudah menemani penulis selama berkuliah di kampus ini.
7. Utsman Fajar, Wahyu Setiawan, Okti Utari dan Anittabi Miliaty yang sudah banyak memberikan masukan, bantuan, serta ilmu selama pengerjaan skripsi ini.
8. Teman-teman grup “KEMANA LAGI” dan Sondang Granitia Siahaan yang telah banyak membantu mengisi dataset penelitian saya.

Penulis menyadari dengan betul bahwa dalam pembuatan skripsi ini masih banyak sekali kekurangan dan kelemahan. Oleh karena itu penulis berharap

untuk semua pihak yang telah membaca dan memahami penelitian ini untuk dapat menyampaikan kritik dan saran yang membangun sehingga dapat menambah kesempurnaan skripsi ini.

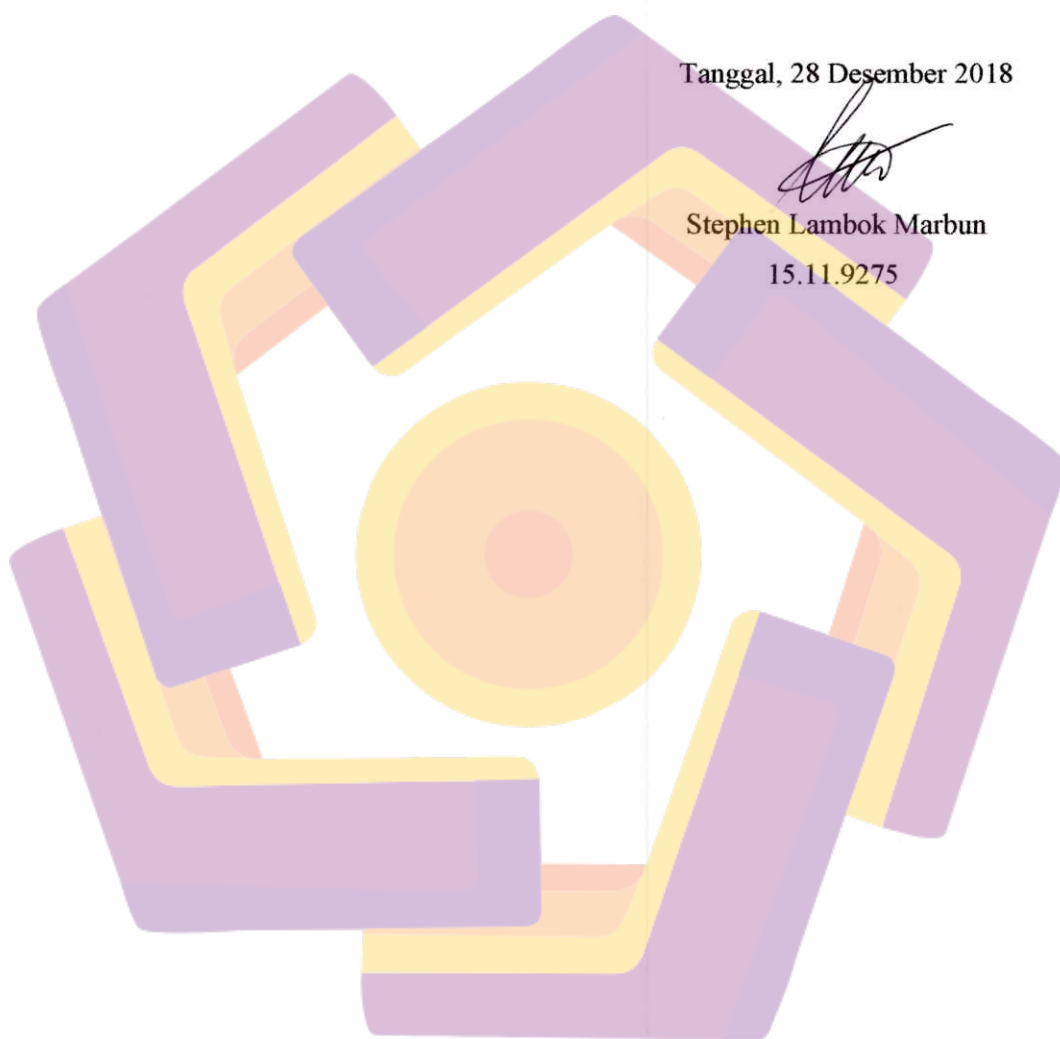
Akhir kata, semoga skripsi ini dapat bermanfaat bagi pihak terkait dan pembaca pada umumnya.

Tanggal, 28 Desember 2018



Stephen Lambok Marbun

15.11.9275



DAFTAR ISI

HALAMAN JUDUL.....	i
PERSETUJUAN.....	ii
PENGESAHAN.....	iii
PERNYATAAN.....	iv
MOTTO.....	v
PERSEMBAHAN.....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
INTISARI.....	xiii
ABSTRACT.....	xiv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	2
1.3. Batasan Masalah.....	2
1.4. Tujuan Penelitian.....	3
1.5. Manfaat Penelitian.....	3
1.6. Metode Penelitian.....	4
1.7. Sistematika Penulisan.....	5
BAB II LANDASAN TEORI.....	7
2.1. Tinjauan pustaka.....	7
2.2. Dasar teori.....	10
2.2.1. <i>Data Mining</i>	10
2.2.2. <i>Supervised Learning</i>	10
2.2.3. <i>Text Mining</i>	11
2.2.4. Analisis Sentimen.....	11
2.2.5. <i>Pre-Processing</i>	11
2.2.6. <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	13
2.2.7. <i>Algoritma Random Forest</i>	14

2.2.8. <i>Decision Tree</i>	16
2.2.9. Evaluasi Klasifikasi.....	17
2.2.10. <i>Flowchart</i>	19
BAB III METODOLOGI PENELITIAN	21
3.1. Alat dan Bahan	21
3.1.1. Alat.....	21
3.1.2. Bahan.....	21
3.2. Langkah Kerja	22
BAB IV HASIL DAN PEMBAHASAN	26
4.1 Deskripsi Hasil dan Pembahasan	26
4.2 Persiapan Data.....	26
4.3 <i>Preprocessing</i> Data	26
4.4 Pembobotan Term	37
4.5 Klasifikasi Data	38
4.6 Evaluasi Hasil.....	39
4.7 Hasil prediksi sentimen Joko Widodo.....	49
BAB V PENUTUP	50
5.1. Kesimpulan.....	50
5.2. Saran.....	50
DAFTAR PUSTAKA	52

DAFTAR TABEL

Tabel 2. 1 Perbandingan Penelitian.....	8
Tabel 2. 2 Confusion Matrix untuk 2 Kelas.....	18
Tabel 2. 3 Flowchart	19
Tabel 4. 1 Hasil Proses Preprocessing	27
Tabel 4. 2 Hasil Proses Tokenisasi	29
Tabel 4. 3 Hasil Proses Normalisasi Kata.....	31
Tabel 4. 4 Hasil Proses Stopword Removal.....	34
Tabel 4. 5 Hasil Proses Stemming	36
Tabel 4. 6 Hasil Confusion Matrix.....	46
Tabel 4. 7 Akurasi Sentimen Joko Widodo	48

DAFTAR GAMBAR

Gambar 3. 1 Flowchart Analisis	22
Gambar 3. 2 Data Tweet	23
Gambar 4. 1 <i>Script</i> Preprocessing	27
Gambar 4. 2 <i>Script</i> Tokenisasi	29
Gambar 4. 3 <i>Script</i> Normalisasi Kata	31
Gambar 4. 4 <i>Script</i> Stopword Removal	33
Gambar 4. 5 <i>Script</i> Stemming	35
Gambar 4. 6 <i>Script</i> TF-IDF	37
Gambar 4. 7 Training Data	38
Gambar 4. 8 <i>Script</i> Testing	38
Gambar 4. 9 Hasil Evaluasi Entropy dan Tree = 10	39
Gambar 4. 10 Hasil Evaluasi Entropy dan Tree = 50	40
Gambar 4. 11 Hasil Evaluasi Entropy dan Tree = 100	40
Gambar 4. 12 Hasil Evaluasi Entropy dan Tree = 250	41
Gambar 4. 13 Hasil Evaluasi Entropy dan Tree = 500	41
Gambar 4. 14 Hasil Evaluasi Entropy dan Tree = 750	42
Gambar 4. 15 Hasil Evaluasi Entropy dan Tree = 1000	42
Gambar 4. 16 Hasil Evaluasi Gini dan Tree = 10	43
Gambar 4. 17 Hasil Evaluasi Gini dan Tree = 50	43
Gambar 4. 18 Hasil Evaluasi Gini dan Tree = 100	44
Gambar 4. 19 Hasil Evaluasi Gini dan Tree = 250	44
Gambar 4. 20 Hasil Evaluasi Gini dan Tree = 500	45
Gambar 4. 21 Hasil Evaluasi Gini dan Tree = 750	45
Gambar 4. 22 Hasil Evaluasi Gini dan Tree = 1000	46
Gambar 4. 23 Grafik Sentimen terhadap Jokowi	49

INTISARI

Analisis sentimen merupakan bagian dari *text mining* yang mendefinisikan opini, perasaan dan sikap dalam bentuk teks. Salah satu media untuk menganalisis adalah Twitter. Di Indonesia, Presiden Indonesia saat ini, Joko Widodo, atau biasa disapa Jokowi, kerap kali menjadi bahan perbincangan orang. Perbincangan berupa tweet tersebut dijadikan bahan untuk menganalisis sentimen.

Pendekatan algoritma *Random Forest* merupakan algoritma yang mengklasifikasikan data berdasarkan target-target tertentu. Dengan menyiapkan data mentah untuk dilakukan *preprocessing* dan *feature extraction*, akan didapat model yang dijadikan alat untuk melakukan analisis sentimen.

Dari hasil penelitian, didapat bahwa hasil akurasi tertinggi pada penelitian ini adalah 57,2% dengan persentase hasil sentimen positif 28,8%, negatif 45,9%, dan netral 25,4%.

Kata kunci : *Random Forest, text mining, Joko Widodo, Twitter*

ABSTRACT

Sentiment analysis is part of the text mining that defines opinions, feelings, and behavior in the form of text. One of the media that used in the analysis is twitter. In Indonesia, the current president Mr.Joko Widodo, or usually called Mr.Jokowi often becomes people favorite discussion topic. Their discussion in the form of tweets is used to do this sentiment analysis.

The Random Forest algorithm approach is an algorithm that classifies data based on certain targets. The Model that used as a tool to do sentiment analysis is obtained by preparing the raw data for the preprocessing and feature extraction.

The result of the research shown that the highest accurate result is 57.2% while the positive, negative, and neutral sentiment percentage are respectively on the level of 28.8%, 45.9%, and 25.4%.

Keywords : *Random Forest, text mining, Joko Widodo, Twitter*