

**ANALISIS STATUS PADA MEDIA SOSIAL TWITTER TERHADAP
PEMILU 2019 MENGGUNAKAN ALGORITMA NAIVE BAYES
CLASSIFIER**

SKRIPSI



disusun Oleh :

**Devy Kumalasari
15.11.8673**

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2019**



**ANALISIS STATUS PADA MEDIA SOSIAL TWITTER TERHADAP
PEMILU 2019 MENGGUNAKAN ALGORITMA NAIVE BAYES
CLASSIFIER**

SKRIPSI

Untuk memenuhi sebagian persyaratan
Mencapai gelar sarjana
Pada Program Studi Informatika



disusun Oleh :
Devy Kumalasari
15.11.8673

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2019**



PERSETUJUAN

SKRIPSI

**ANALISIS STATUS PADA MEDIA SOSIAL TWITTER TERHADAP
PEMILU 2019 MENGGUNAKAN ALGORITMA NAIVE BAYES
CLASSIFIER**

yang dipersiapkan dan disusun oleh

Devy Kumalasari

15.11.8673

telah disetujui oleh Dosen Pembimbing Skripsi

pada tanggal 10 Januari 2019

Dosen Pembimbing,

a.n


Hartatik, S.T., M.CS.

NIK. 190302232

PENGESAHAN

SKRIPSI

ANALISIS STATUS PADA MEDIA SOSIAL TWITTER TERHADAP PEMILU 2019 MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER

yang dipersiapkan dan disusun oleh

Devy Kumalasari

15.11.8673

telah dipertahankan di depan Dewan Penguji
pada tanggal 10 Januari 2019

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Windha Mega Pradnya D, M.Kom.
NIK. 190302185



Erni Seniwati, S.Kom, M.Cs
NIK. 190302231



Anggit Dwi Hartanto, M.Kom
NIK. 190302163



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 7 Februari 2019

DEKAN FAKULTAS ILMU KOMPUTER



Krishawati, S.Si, MT.
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 tanggungjawab saya pribadi.

Yogyakarta, 2 Februari 2019



Devy Kumalasari

NIM. 15.11.8673



MOTTO

“Saat berbuat baik, jangan mengharap balasan yang instan. Pohon pun harus dirawat agar menghasilkan buah yang manis”.

(Devy)

“Maka sesungguhnya bersama kesulitan ada kemudahan. Sesungguhnya bersama kesulitan ada kemudahan. Maka apabila engkau telah selesai (dari sesuatu urusan), tetaplah bekerja keras (untuk urusan yang lain). Dan hanya kepada Tuhanmulah engkau berharap.”

(QS. Al-Insyirah,6-8)

Learn from yesterday,

Live for today,

And hope for tomorrow

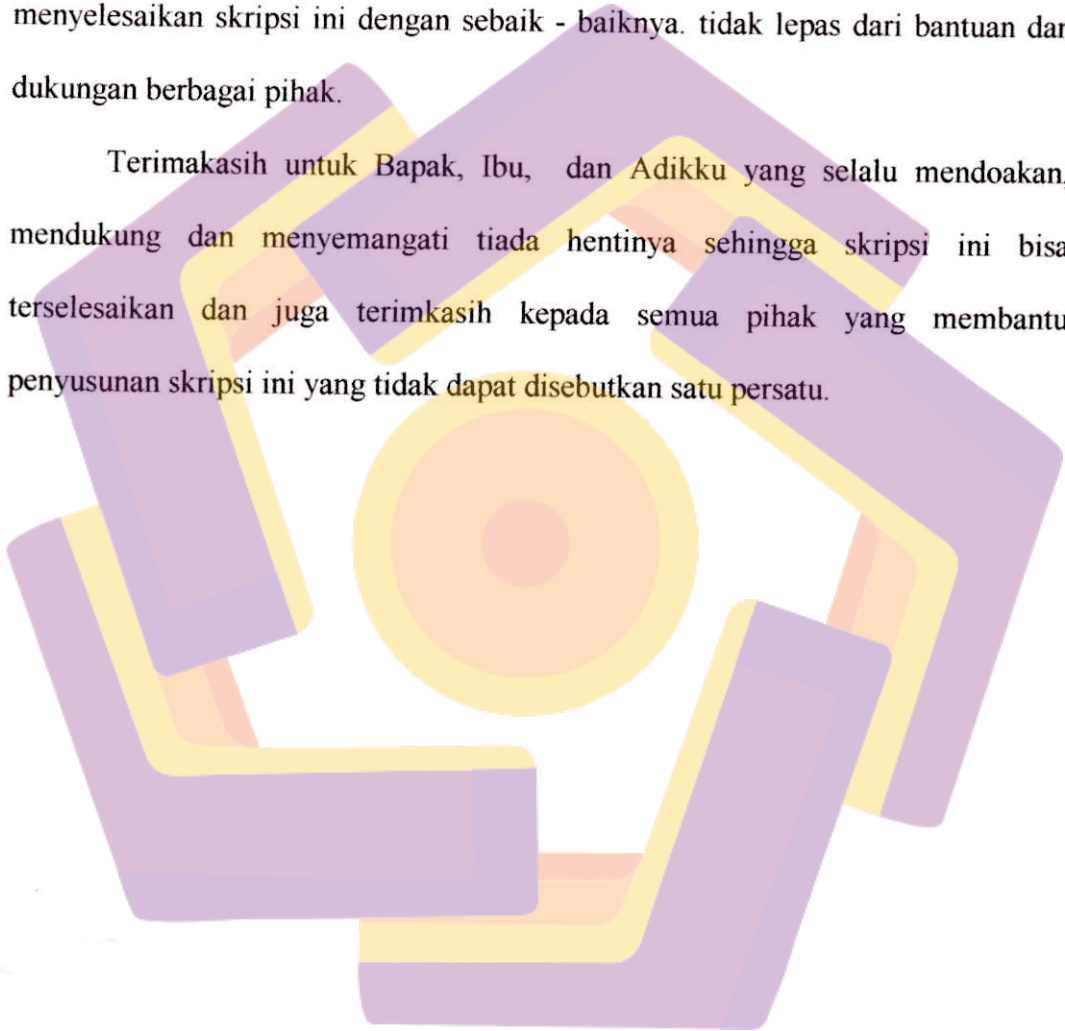
(Albert Einstein)



PERSEMBAHAN

Alhamdulillah segala puji syukur atas berkat rahmat dan karunia Allah SWT yang telah memberikan kemudahan dan kelancaran bagi penulis dalam menyelesaikan skripsi ini dengan sebaik - baiknya. tidak lepas dari bantuan dan dukungan berbagai pihak.

Terimakasih untuk Bapak, Ibu, dan Adikku yang selalu mendoakan, mendukung dan menyemangati tiada hentinya sehingga skripsi ini bisa terselesaikan dan juga terimakasih kepada semua pihak yang membantu penyusunan skripsi ini yang tidak dapat disebutkan satu persatu.



KATA PENGANTAR

Segala puji dan syukur kepada Allah SWT, Dzat yang telah memberikan segenap rahmat, karunia serta hidayah-Nya, khususnya kepada penulis, sehingga dapat menyelesaikan skripsi ini dengan baik. Shala wat serta salam semoga tetap terlimpahkan kepada Nabi Muhammad SAW, yang selalu diharapkan syafaatnya di hari akhir nanti, tak lupa kepada keluarga dan para sahabat atau orang-orang yang mengikuti jejaknya.

Sehubungan dengan selesainya skripsi ini, dengan rendah hati penulis hanya bisa mengucapkan banyak terima kasih kepada semua pihak yang telah membantu, baik berupa moril maupun materiil, terutama kepada:

1. Allah SWT yang selalu setia memberikan petunjuk dan memberikan jalan terang pada saat genting dalam menyelesaikan skripsi ini.
2. Bapak M. Suyanto, Prof., Dr., M.M selaku Rektor Universitas Amikom Yogyakarta.
3. Ibu Hartatik, ST, M.CS selaku dosen pembimbing yang telah memberikan pengarahan bagi penulis serta telah sabar membimbing dalam pembuatan skripsi ini .
4. Bapak I Made Artha Agastya, S.T., M.Kom selaku dosen yang membantu dalam penelitian skripsi ini.
5. Ibu, Bapak, Adik penulis yang selalu setia mendoakan, membimbing, mendukung, sehingga skripsi ini bisa terselesaikan.
6. Assofiq Dwi Kurniawan yang selalu dengan sabar memberikan dukungan dan memotivasi dalam pembuatan skripsi ini.

7. Sri Suryati, Oktavia, dan Adhe Zhaputri yang selalu mendukung, membantu dalam menyelesaikan skripsi ini dari awal perkuliahan sampai sekarang.
8. Dika Mukhanas Putri dan Retno Nurhikmah Prihastanti yang telah menemani saat – saat sulit dalam proses menyelesaikan skripsi ini.
9. Sigit Suryono selaku alumni AMIKOM Yogyakarta yang membantu dalam penelitian skripsi ini.
10. Seluruh Pengurus dan Alumni - Alumni HMIF AMIKOM Yogyakarta.
11. Para Dosen dan Staff Universitas AMIKOM Yogyakarta yang telah membantu memberikan ilmu pengetahuan, pengalaman selama masa perkuliahan.
12. Seluruh Teman Kelas 15 S1 Informatika 03 dan Semua pihak yang telah membantu dan tidak dapat disebutkan satu persatu.

Penulis berharap semoga jasa kebaikan mereka diterima Allah SWT, dan dicatat sebagai amal shaleh.

Lebih dari itu, penulis berharap semoga skripsi ini dapat berguna dan bermanfaat baik bagi penulis maupun bagi semua pihak yang membutuhkannya. Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu kritik dan saran sangat diharapkan demi sempurnanya skripsi ini.

Tanggal, 16 Januari 2019

Devy Kumalasari

15.11.8673

DAFTAR ISI

HALAMAN JUDUL	ii
HALAMAN PERSETUJUAN.....	ii
HALAMAN PERNYATAAN.....	ii
MOTTO	iv
PERSEMBAHAN.....	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiv
INTISARI	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah.....	3
1.4 Maksud dan Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
1.6 Metode Penelitian.....	5
1.7 Sistematika Penulisan.....	7
BAB II LANDASAN TEORI.....	9
2.1 Tinjauan Pustaka	9
2.2 Pengertian Data Mining	11

2.3	Pengertian Sentimen Analysis.....	12
2.4	Preprocessing	13
2.5	Term Frequency	14
2.6	Naive Bayes Classifier	14
2.7	Split Validation	15
2.8	Coffusion Matrix.....	16
2.9	Flowchart.....	18
2.9.1	System Flowchart (Diagram Alir Sistem).....	18
2.9.2	Program Flowchart (Diagram Alir Program).....	20
2.10	Python.....	21
2.11	Natural Language Toolkit (NLTK).....	22
BAB III ANALISIS DAN PERANCANGAN		23
3.1	Analisis Masalah	23
3.2	Solusi yang Dipilih.....	24
3.3	Analisis Kebutuhan	24
3.3.1	Analisis Kebutuhan Fungsional	24
3.3.2	Analisis Kebutuhan Non-Fungsional.....	25
3.4	Analisa Sentimen Klasifikasi Algoritma Naive Bayes Classiffier.....	25
3.4.1	Cleaning data	26
3.4.2	Case Folding	28
3.4.3	Tokenization	29
3.4.4	Steaming	30
3.4.5	Term Frequency.....	32

3.4.6	Naive Bayes Classifier.....	35
3.4.7	Split Validation.....	44
3.5	Perancangan Aplikasi.....	44
3.5.1	Preprocessing Data.....	44
3.5.2	Proses Klasifikasi Naive Bayes.....	46
3.6	Perancangan Data Twitter.....	47
3.7	Perancangan File Training.....	48
BAB IV	IMPLEMENTASI DAN PEMBAHASAN.....	49
4.1	Dekripsi Implementasi.....	49
4.2	Implementasi Pengumpulan Data Tweet.....	49
4.3	Implementasi Pelabelan Data Training.....	51
4.4	Implementasi Perancangan Fungsi dan Sistem.....	51
4.4.1	Mengakses Dataset Training.....	51
4.4.2	Tahap Preprocessing Data.....	52
4.4.3	Pembuatan Fungsi Naive Bayes Classifier.....	53
4.5	Pembuatan Interface.....	54
4.6	Evaluasi Program.....	59
4.6.1	Pengujian dan Pengukuran Skenario.....	60
BAB V	KESIMPULAN.....	70
5.1	Kesimpulan.....	70
5.2	Saran.....	70

DAFTAR TABEL

Tabel 3.1 Data Tweet Sebelum Dilakukan Cleaning Data	26
Tabel 3.2 Data Setelah Dilakukan Cleaning Data	27
Tabel 3.3 Data Tweet Setelah Case Folding.....	28
Tabel 3.4 Hasil Tokenization Dari Case Folding.....	29
Tabel 3.5 Hasil Steaming Dari Tokenizing	30
Tabel 3.6 Feature List	31
Tabel 3.7 Term Positif	32
Tabel 3.8 Term Negatif.....	34
Tabel 3.9 Data Latih	35
Tabel 3.10 Data Latih Positif.....	36
Tabel 3.11 Data Latih Negatif	37
Tabel 3.12 Presentase Pembagian Data	44
Tabel 4.1 Hasil Rata – Rata Pengujian Akurasi.....	68

DAFTAR GAMBAR

Gambar 2.1 System Flowchart (Diagram Alir System).....	19
Gambar 2.2 Program Flowchart (Diagram Alir Program).....	20
Gambar 3.1 Alur Preprocessing.....	45
Gambar 3.2 Alur Naive Bayes Classifier.....	47
Gambar 3.3 Program Akses API Twitter.....	48
Gambar 3.4 File Training.....	48
Gambar 4.1 Script Crawling Data Twitter.....	50
Gambar 4.2 Data hasil crawling.....	50
Gambar 4.3 Data Training.....	51
Gambar 4.4 Script Akses Data Training.....	51
Gambar 4.5 Script Preprocessing Data.....	52
Gambar 4.6 Script Perhitungan Probabilitas Kategori.....	53
Gambar 4.7 Script Perhitungan Probabilitas Setiap Kata.....	53
Gambar 4.8 Script Perhitungan Prediksi.....	54
Gambar 4.9 Halaman Utama.....	55
Gambar 4.10 Halaman Proses Analisis.....	56
Gambar 4.11 Halaman Hasil Klasifikasi Jokowi.....	57
Gambar 4.12 Pie Chart Hasil Klasifikasi Jokowi.....	57
Gambar 4.13 Halaman Hasil Klasifikasi Prabowo.....	58
Gambar 4.14 Pie Chart Hasil Klasifikasi Prabowo.....	59
Gambar 4.15 Pengujian Skenario 1.....	60
Gambar 4.16 Klasifikasi Naive Bayes Skenario 1.....	61

Gambar 4.17 Hasil Akurasi Skenario 1	61
Gambar 4.18 Pengujian Skenario 2	62
Gambar 4.19 Pengujian Klasifikasi Naive Bayes Skenario 2.....	62
Gambar 4.20 Hasil Akurasi Skenario 2	63
Gambar 4.21 Pengujian Skenario 3	64
Gambar 4.22 Pengujian Klasifikasi Naive Bayes Skenario 3.....	64
Gambar 4.23 Hasil Akurasi Pengujian Skenario 3	65
Gambar 4.24 Pengujian Skenario 4	65
Gambar 4.25 Pengujian Klasifikasi Naive Bayes Skenario 4.....	66
Gambar 4.26 Hasil Akurasi Pengujian Skenario 4	66
Gambar 4.27 Pengujian Skenario 5	67
Gambar 4.28 Pengujian Klasifikasi Naive Bayes Skenario 5.....	67
Gambar 4.29 Pengujian Hasil Klasifikasi Skenario 5.....	68

INTISARI

Media Sosial saat ini sudah menjadi keutuhan pokok hampir seluruh lapisan masyarakat. Dari media sosial, masyarakat dapat dengan mudah menemukan informasi yang dicari dan mengetahui informasi terbaru dari belahan dunia. Melalui media sosial, masyarakat dapat menyampaikan opininya melalui status pada media sosial tertentu. Salah satu media sosial yang banyak digunakan oleh masyarakat adalah Twitter.

Menurut kominfo, berdasarkan data PT Bakrie Telecom, Twitter memiliki 19,5 juta pengguna di Indonesia dari total 500 juta pengguna global. Dengan demikian dapat disimpulkan bahwa twitter memiliki dampak baik untuk konsumen maupun produsen dalam mengambil keputusan, saran, opini atau ulasan – ulasan.

Oleh karena itu opini – opini tersebut dapat diimplementasikan untuk analisis sentimen terhadap pemilihan Pemilihan Presiden pada Pemilu 2019 yang bertujuan untuk meningkatkan strategi masing – masing calon agar menjadi lebih baik. Dalam analisis sentimen status pada twitter, data akan di crawling lalu dilakukan preprocessing untuk dilakukan klasifikasi dengan Naive Bayes Classifier.

Kata Kunci : Analisis Sentimen, Media Sosial, Pemilu 2019, Twitter, Preprocessing , Crawling., Naive Bayes Classifier

ABSTRACT

Today's social media has become the main integrity of almost all levels of society. From social media, people can easily find information that is searched for and find out the latest information from around the world. Through social media, people can express their opinions through status on certain social media. One social media that is widely used by the public is Twitter.

According to Kominfo, based on data from PT Bakrie Telecom, Twitter has 19.5 million users in Indonesia out of a total of 500 million global users. Thus it can be concluded that twitter has an impact on both consumers and producers in making decisions, suggestions, opinions or reviews.

Therefore, these opinions can be implemented for the sentiment analysis of the Presidential Election in the 2019 Election which aims to improve the strategies of each candidate to be better. In the status sentiment analysis on twitter, the data will be crawled and then preprocessed to be classified with the Naive Bayes Classifier.

Keywords: *Sentiment Analysis, Social Media, 2019 Election, Twitter, Preprocessing, Crawling., Naive Bayes Classifier*

