

**ANALISIS SENTIMEN MASYARAKAT TERHADAP
PENYELENGAARAN PILKADA PADA TWITTER
MENGUNAKAN ALGORITMA
NAIVE BAYES CLASSIFIER**

SKRIPSI



disusun oleh

Rio Triyanto

19.21.1337

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

**ANALISIS SENTIMEN MASYARAKAT TERHADAP
PENYELENGAARAN PILKADA PADA TWITTER
MENGUNAKAN ALGORITMA
NAIVE BAYES CLASSIFIER**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



disusun oleh

Rio Triyanto

19.21.1337

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN MASYARAKAT TERHADAP
PENYELENGARAAN PILKADA PADA TWITTER**

**MENGGUNAKAN ALGORITMA
NAIVE BAYES CLASSIFIER**

yang dipersiapkan dan disusun oleh

**Rio Triyanto
19.21.1337**

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 21 Januari 2021

Dosen Pembimbing,

Windha Mega Pradnya D, M.Kom.

NIK. 190302185

PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN MASYARAKAT TERHADAP
PENYELENGAARAN PILKADA PADA TWITTER
MENGUNAKAN ALGORITMA
NAIVE BAYES CLASSIFIER**

yang dipersiapkan dan disusun oleh

Rio Triyanto

19.21.1337

telah dipertahankan di depan Dewan Penguji
pada tanggal 22 Februari 2021

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Windha Mega Pradnya D, M.Kom.

NIK. 190302185

Norhikmah, M.Kom.

NIK. 190302245

Mei P Kurniawan, M.Kom.

NIK. 190302187

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 7 Maret 2021

DEKAN FAKULTAS ILMU KOMPUTER

Krisnawati, S.Si., M.T.

NIK. 190302028

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, 13 Februari 2021



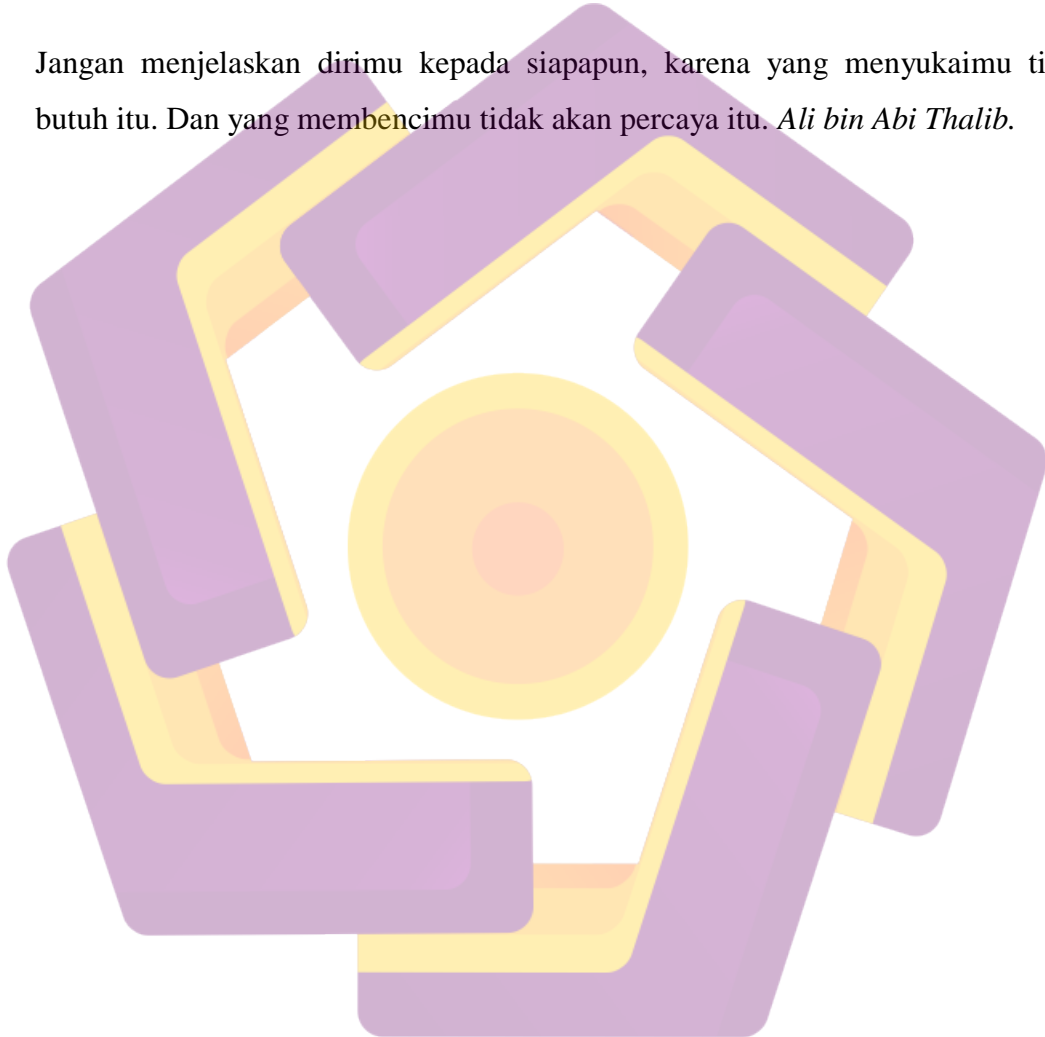
Rio Triyanto

NIM. 19.21.1337

MOTTO

Waktu bagaikan pedang. Jika kamu tidak memanfaatkannya dengan baik, maka ia akan memanfaatkanmu. *HR. Muslim*

Jangan menjelaskan dirimu kepada siapapun, karena yang menyukaimu tidak butuh itu. Dan yang membencimu tidak akan percaya itu. *Ali bin Abi Thalib.*



PERSEMBAHAN

Puji syukur atas kehadiran Allah SWT yang senantiasa memberikan rahmat dan hidayah-Nya, sehingga Skripsi dengan baik dan tepat pada waktunya. Skripsi ini saya dedikasikan kepada pihak-pihak yang telah berperan penting dalam masa perkuliahan hingga penyelesaian Skripsi, pihak-pihak tersebut adalah:

1. Allah SWT yang telah memberikan kemudahan dan kelancaran dalam penyusunan Skripsi ini.
2. Bapak Mudiyono, Ibu Rukinah yang telah memberikan doa dan dukungan tiada hentinya.
3. Bapak Sudarmawan, MT., selaku Ketua Program Studi Informatika.
4. Ibu Windha Mega Pradnya D, M.Kom., selaku dosen pembimbing yang telah membantu dan membimbing dengan penuh kesabaran dalam memberikan motivasi beserta saran-saran yang membangun.
5. Nur Jihan yang bersedia direpotkan tempat dan waktunya selama penyelesaian Skripsi.
6. Teman-teman S1 Informatika Transfer yang telah berjuang bersama dan memberikan dukungan.
7. Semua pihak yang tidak dapat disebutkan satu-persatu, kami ucapkan terima kasih.

KATA PENGANTAR

Alhamdulillah segala puji syukur kami panjatkan atas kehadiran Allah SWT karena dengan limpahan rahmat dan hidayah-Nya, penulis dapat menyelesaikan Skripsi yang berjudul “Analisi Sentimen Masyarakat Terhadap Penyelenggaraan Pilkada pada Twitter Menggunakan Algoritma Naive Bayes Classifier”. Skripsi ini dibuat dengan tujuan memenuhi persyaratan kelulusan jenjang Sarjana – Program Studi Informatika Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.

Dalam pengerjaannya, penulis mendapatkan saran, bimbingan, dorongan serta keterangan-keterangan yang berasal dari beberapa pihak. Sehingga hal tersebut memberikan pengalaman yang tidak bisa diukur dengan materi. Oleh karenanya, saya ucapkan banyak terima kasih kepada seluruh pihak yang telah membantu menyelesaikan Skripsi ini. Khususnya kepada yang terhormat

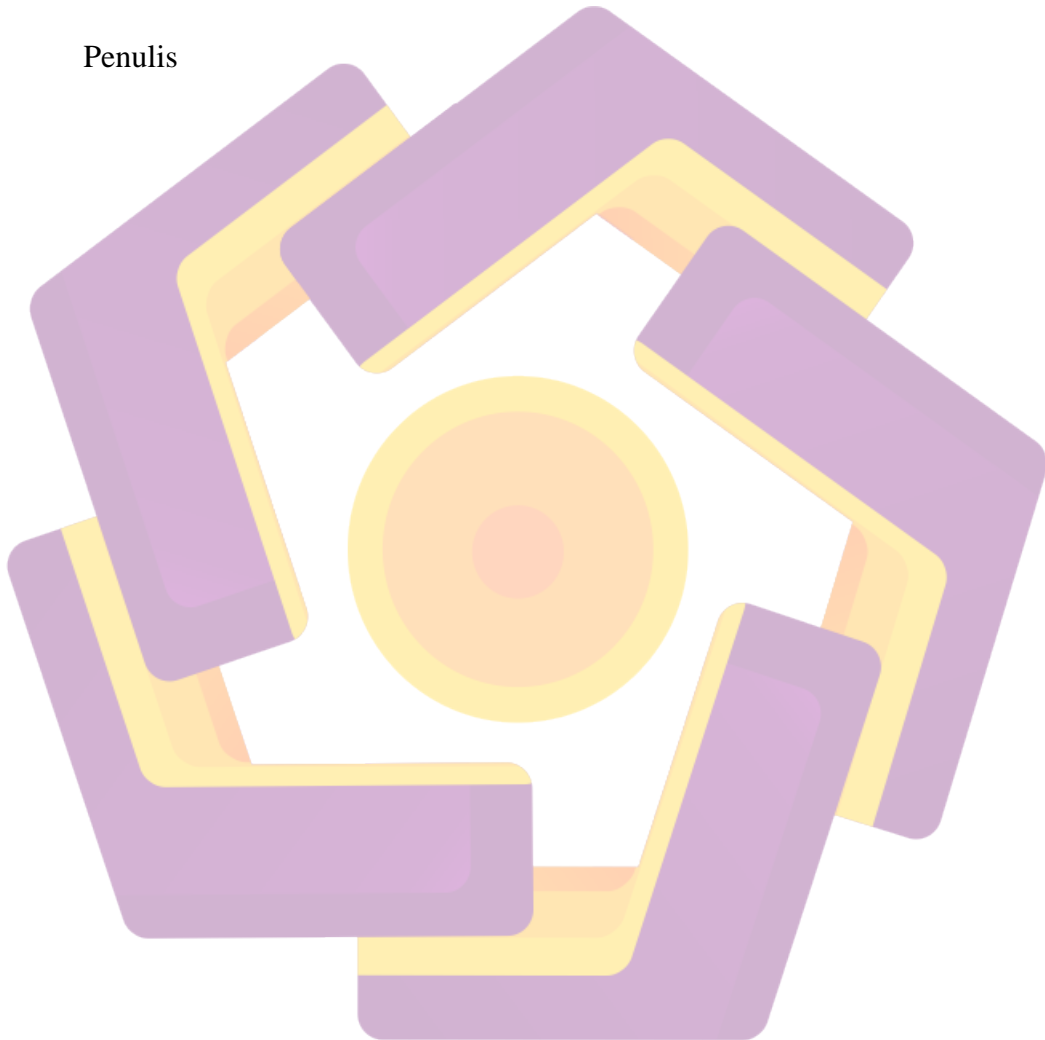
1. Bapak Prof., Dr., M. Suyanto, M.M., selaku Rektor Universitas AMIKOM Yogyakarta.
2. Ibu Krisnawati, S.Si., M.T., selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Bapak Sudarmawan, MT., selaku Ketua Program Studi Informatika.
4. Ibu Windha Mega Pradnya D, M.Kom., selaku Dosen Pembimbing.
5. Para dosen penguji yang bersedia menguji dan memberikan nilai dengan bijaksana.
6. Serta semua kerabat dekat dan rekan-rekan seperjuangan yang tak bisa kami sebutkan satu-persatu.

Pada proses penulisan Skripsi ini, penulis sadar bahwa masih terdapat banyak kelemahan dan kekurangan. Untuk itu penulis memohon maaf yang sebesar-besarnya atas hal tersebut.

Akhir kata, semoga Skripsi ini dapat memberikan manfaat bagi masyarakat luas, organisasi mahasiswa, institusi pendidikan dan khususnya bagi penulis sendiri.

Yogyakarta, 7 Februari 2021

Penulis



DAFTAR ISI

JUDUL	i
PERSETUJUAN	ii
PENGESAHAN	iii
PERNYATAAN	iv
MOTTO	v
PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
INTISARI	xvii
ABSTRACT	xviii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Maksud dan Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Metodologi Penelitian	3
1.6.1 Metode Pengumpulan Data	3
1.6.2 Metode Analisis	3
1.6.3 Metode Perancangan	4
1.6.4 Metode Implementasi	4
1.6.5 Metode Pengujian	4

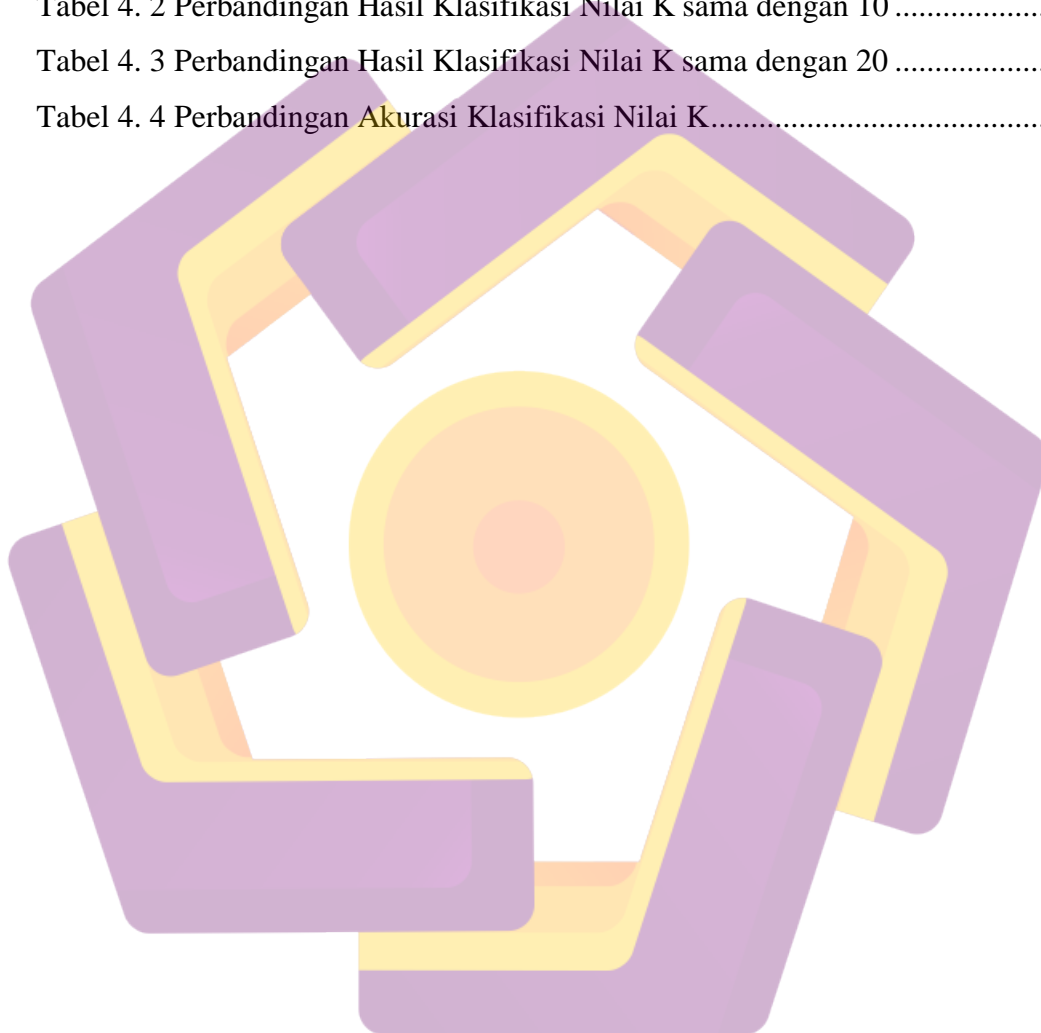
1.7	Sistematika Penulisan.....	5
BAB II	LANDASAN TEORI.....	6
2.1	Tinjauan Pustaka.....	6
2.2	Dasar Teori.....	11
2.2.1	Data Mining.....	11
2.2.2	<i>Text Mining</i>	12
2.2.3	Analisis Sentimen.....	12
2.2.4	Seleksi Fitur Information Gain.....	12
2.2.5	Twitter.....	13
2.2.6	<i>Naive Bayes Classifier (NBC)</i>	13
2.2.7	<i>Python</i>	14
2.2.8	<i>Jupyter Lab</i>	14
2.2.9	<i>Tweepy</i>	14
2.2.10	PHP.....	15
2.2.11	<i>Codeigniter</i>	15
2.2.12	<i>Postgresql</i>	15
2.2.13	JSON.....	15
2.2.14	HTML/CSS.....	15
2.2.15	<i>Javascript</i>	16
2.2.16	<i>K-Fold Cross Validation</i>	16
BAB III	ANALISIS DAN PERANCANGAN.....	16
3.1	Deskripsi Umum.....	16
3.2	Crawling Data.....	16
3.3	Input Dataset.....	17
3.4	Pelabelan.....	18
3.5	Preprocessing.....	19
3.5.1	Case Folding.....	20
3.5.2	Cleansing.....	21

3.5.3	Tokenizing	22
3.5.4	Normalization	23
3.5.5	Stopword Removal	25
3.6	Seleksi Fitur Information Gain.....	27
3.7	Klasifikasi Naive Bayes	29
3.8	Analisis Sistem.....	32
3.8.1	Analisi Kebutuhan Non-Fungsional	32
3.8.2	Analisis Kebutuhan Fungsional	33
3.8.3	Rancangan Database	47
3.8.4	Desain Tampilan	48
BAB IV	IMPLEMENTASI DAN PEMBAHASAN	55
4.1	Pembuatan Database	55
4.1.1	Pembuatan Database Postgresql	55
4.1.2	Pembuatan Table.....	55
4.2	Perancangan Sistem dan Implementasi.....	57
4.2.1	Pengumpulan Data.....	57
4.2.2	Pembuatan Projek Baru	57
4.2.3	Pengaturan Codeigniter dan Database	58
4.2.4	Dashboard	59
4.2.5	Master Data.....	60
4.2.6	Kamus Stopword Removal	61
4.2.7	Kamus Normalisasi.....	63
4.2.8	Preprocessing	65
4.2.9	Seleksi Fitur	66
4.2.10	K-Fold Cross Validation.....	68
4.2.11	Setting Seleksi Fitur.....	68
4.2.12	Klasifikasi Navie Bayes.....	69
4.3	Pengujian.....	70

4.3.1	Pengujian dengan K sama dengan 5	70
4.3.2	Pengujian dengan K sama dengan 10	74
4.3.3	Pegujian dengan K samadengan 20	77
4.4	Hasil	82
BAB 5 PENUTUP.....		82
5.1	Kesimpulan	82
5.2	Saran.....	82
DAFTAR PUSTAKA		83
LAMPIRAN A CONTROLLER HALAMAN DASHBOARD		85
LAMPIRAN B CONTROLLER HALAMAN MASTER		88
LAMPIRAN C CONTROLLER HALAMAN STOPWORD REMOVAL.....		90
LAMPIRAN D CONTROLLER HALAMAN KAMUS NORMALISASI		93
LAMPIRAN E CONTROLLER HALAMAN PREPROSESSING		96
LAMPIRAN F CONTROLLER HALAMAN SELEKSI FITUR		99
LAMPIRAN G CONTROLLER HALAMAN PENGISIAN NILAI K		101
LAMPIRAN H CONTROLLER HALAMAN PENGISIAN NILAI FITUR.....		103
LAMPIRAN I CONTROLLER HALAMAN KLASIFIKASI NAIVE BAYES		104
LAMPIRAN J CONTROLLER HALAMAN PENGUJIAN		107

DAFTAR TABEL

Tabel 2. 1 Perbedaan Penelitian Terdahulu.....	7
Tabel 4. 1 Perbandingan Hasil Klasifikasi Nilai K sama dengan 5	73
Tabel 4. 2 Perbandingan Hasil Klasifikasi Nilai K sama dengan 10	77
Tabel 4. 3 Perbandingan Hasil Klasifikasi Nilai K sama dengan 20	80
Tabel 4. 4 Perbandingan Akurasi Klasifikasi Nilai K.....	81

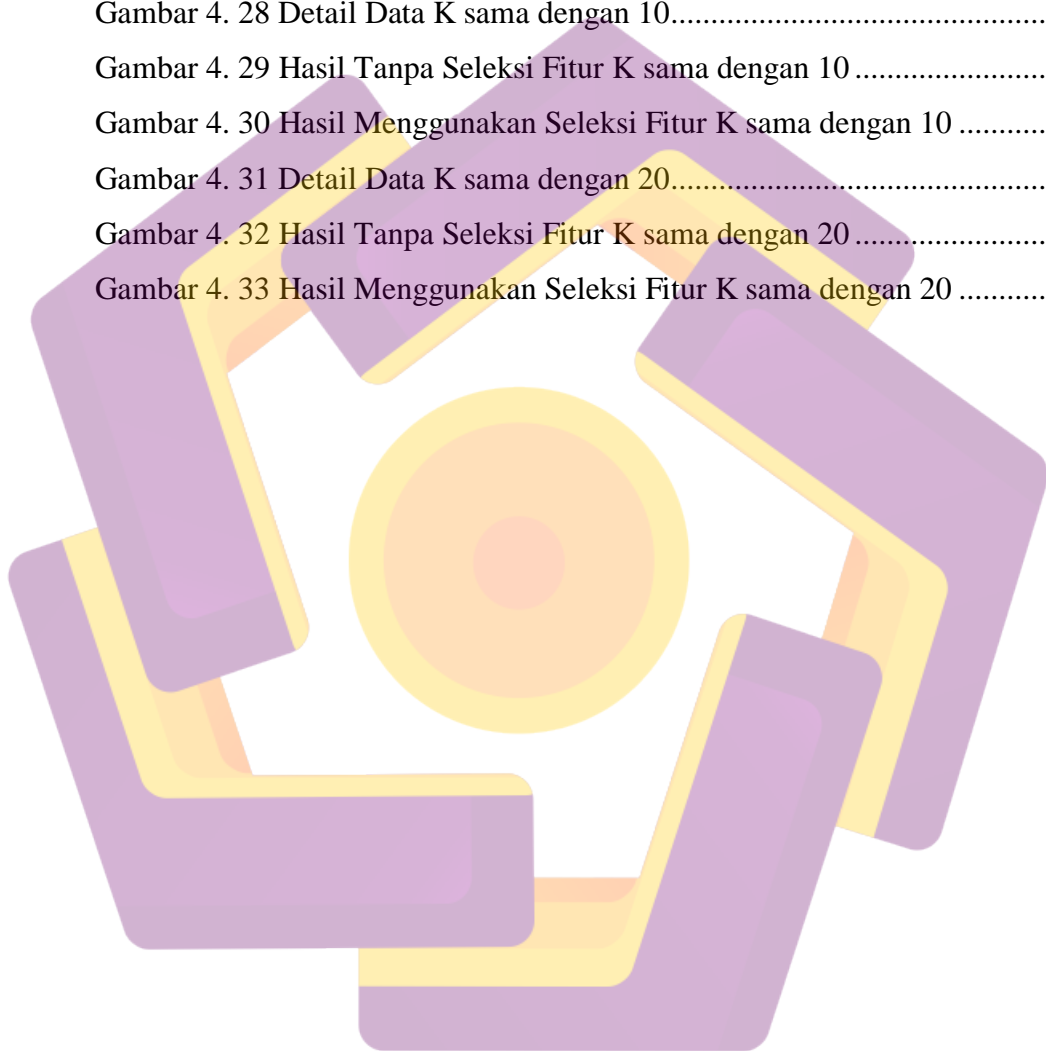


DAFTAR GAMBAR

Gambar 3. 1 Gambaran Umum Sistem.....	16
Gambar 3. 2 <i>Crawling</i> Data.....	17
Gambar 3. 3 Input Dataset.....	18
Gambar 3. 4 Pelabelan.....	19
Gambar 3. 5 Preprocessing.....	20
Gambar 3. 6 Tahapan Case Folding.....	21
Gambar 3. 7 Tahapan Cleansing.....	22
Gambar 3. 8 Tahapan Tokenizing.....	23
Gambar 3. 9 Tahapan Normalization.....	24
Gambar 3. 10 Tahapan Stopword Removal.....	26
Gambar 3. 11 Pembagian Data.....	27
Gambar 3. 12 Tahapan Seleksi Fitur Information Gain.....	28
Gambar 3. 13 Diagram Alur Perhitungan Naive Bayes.....	30
Gambar 3. 14 Pemisahan Kata.....	31
Gambar 3. 15 Perhitungan Naive Bayes.....	32
Gambar 3. 16 Use Case Diagram.....	33
Gambar 3. 17 Activity Diagram Input Dataset.....	37
Gambar 3. 18 Activity Diagram Klasifikasi Data.....	38
Gambar 3. 19 Activity Diagram Create Kamus.....	39
Gambar 3. 20 Activity Diagram Update Kamus.....	40
Gambar 3. 21 Activity Diagram Delete Kamus.....	41
Gambar 3. 22 Activity Diagram Preprocessing.....	42
Gambar 3. 23 Activity Diagram Perhitungan Information Gain.....	43
Gambar 3. 24 Activity Diagram Klasifikasi Naive Bayes.....	44
Gambar 3. 25 Sequence Diagram Input Dataset.....	45
Gambar 3. 26 Sequence Diagram Klasifikasi Data.....	45
Gambar 3. 27 Sequence Diagram CRUD Kamus Stopword Removal.....	46
Gambar 3. 28 Sequence Diagram Preprocessing.....	46

Gambar 3. 29 Sequence Diagram Perhitungan Information Gain	47
Gambar 3. 30 Sequence Diagram Klasifikasi Naive Bayes.....	47
Gambar 3. 31 ERD.....	48
Gambar 3. 32 Halaman Masker Data.....	49
Gambar 3. 33 Halaman Klasifikasi Data	50
Gambar 3. 34 Halaman Kamus Stopword Removal.....	51
Gambar 3. 35 Halaman Preprosessing.....	52
Gambar 3. 36 Halaman Perhitungan Information Gain.....	53
Gambar 3. 37 Halaman Analisis Sentimen.....	54
Gambar 4. 1 Login Server Database.....	55
Gambar 4. 2 Website Codeigniter.....	58
Gambar 4. 3 Pengaturan URL.....	58
Gambar 4. 4 Pengaturan Database.....	59
Gambar 4. 5 Halaman Dashboard.....	60
Gambar 4. 6 Hasil Klasifikasi Dashboard	60
Gambar 4. 7 Data Master.....	61
Gambar 4. 8 Kamus Stopword Removal	61
Gambar 4. 9 Tambah Kata Stopword Removal.....	62
Gambar 4. 10 Ubah Stopword Removal.....	62
Gambar 4. 11 Hapus Stopword Removal	63
Gambar 4. 12 Tampilan Kamus Normalisasi.....	63
Gambar 4. 13 Tambah Kamus Normalisasi.....	64
Gambar 4. 14 Ubah Kamus Normalisasi	64
Gambar 4. 15 Hapus Kamus Normalisasi.....	65
Gambar 4. 16 Tampilan Preprosessing.....	65
Gambar 4. 17 Konfirmasi Preprosessing.....	66
Gambar 4. 18 Tampilan Seleksi Fitur	67
Gambar 4. 19 Mater Kata Seleksi Fitur	67
Gambar 4. 20 Hasil Seleksi Fitur.....	68
Gambar 4. 21 Pengaturan Nilai K.....	68
Gambar 4. 22 Pengaturan Seleksi Fitur	69

Gambar 4. 23 Tampilan Klasifikasi.....	70
Gambar 4. 24 Konfirmasi Klasifikasi.....	70
Gambar 4. 25 Detail Data K sama dengan 5.....	71
Gambar 4. 26 Hasil Tanpa Seleksi Fitur K sama dengan 5.....	72
Gambar 4. 27 Hasil Menggunakan Seleksi Fitur K sama dengan 5.....	73
Gambar 4. 28 Detail Data K sama dengan 10.....	74
Gambar 4. 29 Hasil Tanpa Seleksi Fitur K sama dengan 10.....	75
Gambar 4. 30 Hasil Menggunakan Seleksi Fitur K sama dengan 10.....	76
Gambar 4. 31 Detail Data K sama dengan 20.....	78
Gambar 4. 32 Hasil Tanpa Seleksi Fitur K sama dengan 20.....	78
Gambar 4. 33 Hasil Menggunakan Seleksi Fitur K sama dengan 20.....	79



INTISARI

Pandemi corona atau covid-19 di Indonesia sudah berlangsung selama lebih dari 6 bulan. Dimulai pada bulan Maret 2020, dan sampai saat ini masyarakat yang terkonfirmasi terinfeksi covid 19 semakin meningkat, bahkan di beberapa daerah mengalami kenaikan drastis. Berbagai tindakan daerah pemerintah telah dilakukan, baik pelaksanaan PSBB (Pembatasan Sosial Berskala Besar), peraturan ketat agar masyarakat mematuhi protocol kesehatan, bahkan pemerintah telah memberikan bantuan social kepada masyarakat terdampak pandemic.

Tetapi dari beberapa kebijakan pemerintah ini yang paling menjadi sorotan saat ini adalah tetap dilaksanakannya PILKADA serentak untuk beberapa wilayah di Indonesia. Meskipun akan dilakukan penjagaan dan diharuskan mematuhi protocol kesehatan pilkada akan membuat kerumunan dan berpotensi penularan covid.

Dengan kebijakan yang kontroversial tersebut pasti masyarakat mempunyai opini dan pandangan masing-masing. Oleh karena itu peneliti akan membuat analisis sentiment terhadap kebijakan pemerintah yang tetap akan melaksanakan PILKADA serentak. Dengan focus pada media social twitter, penelitian ini akan menggunakan algoritma Naïve Bayes Classifier. Hasil dari penelitian ini diharapkan pemerintah dapat mengevaluasi kebijakannya dan lebih bijaksana lagi untuk membuat kebijakan agar elemen masyarakat non pemerintahan suaranya juga dapat didengar.

Kata Kunci: klasifikasi, analisis sentiment, Naïve Bayes Classifier, NBC, seleksi fitur, Information Gain

ABSTRACT

The corona or covid-19 pandemic in Indonesia has been going on for more than 6 months. Starting in March 2020, and until now people who have been confirmed infected with Covid 19 are increasing, even in some areas it has increased dramatically. Various local government actions have been carried out, including the implementation of the PSBB (Large-Scale Social Restrictions), strict regulations so that people comply with health protocols, and even the government has provided social assistance to communities affected by the pandemic.

However, of these government policies the most highlighted at the moment is the simultaneous implementation of regional elections for regional elections in Indonesia. Although safeguards will be carried out and required to comply with election health protocols, it will create crowds and have the potential for covid transmission.

With this controversial policy, it is certain that the public will have their own opinions and views. Therefore, researchers will make a sentiment analysis on government policies that will continue to carry out simultaneous regional elections. With a focus on social media twitter, this study will use the Naïve Bayes Classifier algorithm. The results of this study are expected that the government can evaluate its policies and be wiser in making policies so that their voices can also be heard from non-government elements of society.

Keywords: classification, sentiment analysis, Naïve Bayes Classifier, NBC, seleksi futur, Informastion Gain