

**ANALISIS SENTIMEN PUBLIK TERHADAP TES  
CPNS MELALUI MEDIA TWITER MENGGUNAKAN  
METODE NAÏVE BAYES CLASSIFIER**

**Tugas Akhir**

**Diajukan untuk Memenuhi Salah Satu Persyaratan Mencapai Derajat  
Sarjana Komputer**



Dibuat Oleh:

**Andriel Kiding**

**140707933**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNOLOGI INDUSTRI  
UNIVERSITAS ATMA JAYA YOGYAKARTA  
2021**

# LEMBAR PENGESAHAN

## ANALISIS SENTIMEN PUBLIK TERHADAP TES CPNS MELALUI MEDIA TWITER MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER

Yogyakarta, **Tanggal Bulan Tahun**

**Andriel kiding**  
**140707933**

**Menyetujui,**

**Pembimbing I**

**Pembimbing II**

Patricia Ardanari, S.Si.,M.T.,

Prof. Ir. Suyoto, M.Sc., Ph.D.,

**Penguji I**

**Nama dan Gelar Dosen**

**Penguji II**

**Penguji III**

**Nama dan Gelar Dosen**

**Nama dan Gelar Dosen**

**Mengetahui,**  
**Dekan Fakultas Teknologi Industri**

## HALAMAN PENGESAHAN

Tugas Akhir Berjudul

ANALISIS SENTIMEN PUBLIK TERHADAP TES CPNS MELALUI MEDIA  
TWITTER MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER

yang disusun oleh

Andriel Kiding

140707933

dinyatakan telah memenuhi syarat pada tanggal 27 Januari 2022

### Keterangan

Dosen Pembimbing 1 : Patricia Ardanari, S.Si.,M.T. Telah Menyetujui

Dosen Pembimbing 2 : Prof. Ir. Suyoto, M.Sc., Ph.D. Telah Menyetujui

### Tim Penguji

Penguji 1 : Patricia Ardanari, S.Si.,M.T. Telah Menyetujui

Penguji 2 : Th. Adi Purnomo Sidhi, S.T., M.T. Telah Menyetujui

Penguji 3 : Stephanie Pamela Adithama, S.T., M.T. Telah Menyetujui

Yogyakarta, 27 Januari 2022

Universitas Atma Jaya Yogyakarta

Teknologi Industri

Dekan

ttd.

Dr. A. Teguh Siswantoro, M.Sc.

Dokumen ini merupakan dokumen resmi UAJY yang tidak memerlukan tanda tangan karena dihasilkan secara elektronik oleh Sistem Bimbingan UAJY. UAJY bertanggung jawab penuh atas informasi yang tertera di dalam dokumen ini

## **PERNYATAAN ORISINALITAS & PUBLIKASI ILMIAH**

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Andriel kiding  
NPM : 140707933  
Program Studi : Informatika  
Fakultas : Teknologi Industri  
Judul Penelitian : Analisis Sentiment Publik Terhadap Tes CPNS  
Melalui Media Twiter Menggunakan Metode Naïve Bayes Classifier

Menyatakan dengan ini:

1. Tugas Akhir ini adalah benar tidak merupakan salinan sebagian atau keseluruhan dari karya penelitian lain.
2. Memberikan kepada Universitas Atma Jaya Yogyakarta atas penelitian ini, berupa Hak untuk menyimpan, mengelola, mendistribusikan, dan menampilkan hasil penelitian selama tetap mencantumkan nama penulis.

3. Bersedia menanggung secara pribadi segala bentuk tuntutan hukum atas pelanggaran Hak Cipta dalam pembuatan Tugas Akhir ini.

Demikianlah pernyataan ini dibuat dan dapat dipergunakan sebagaimana mestinya.

Yogyakarta, **Tanggal Bulan Tahun**

Yang menyatakan,

Andriel kiding

140707933



## KATA PENGANTAR

Puji dan syukur penulis haturkan kepada Tuhan Yang Maha Esa karena berkat rahmat dan karunia-Nya penulis dapat menyelesaikan pembuatan tugas akhir “Analisis Sentimen Publik Terhadap Tes CPNS Melalui Media *Twitter* Menggunakan Metode Naïve Bayes Classifier” ini dengan baik.

Penulisan tugas akhir ini bertujuan untuk memenuhi salah satu syarat untuk mencapai derajat sarjana komputer dari Program Studi Informatika, Fakultas Teknologi Industri di Universitas Atma Jaya Yogyakarta.

Penulis menyadari bahwa dalam pembuatan tugas akhir ini penulis telah mendapatkan bantuan, bimbingan, dan dorongan dari banyak pihak. Untuk itu, pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Tuhan Yesus Kristus yang selalu membimbing dalam iman-Nya, memberikan berkat-Nya, dan menyertai penulis selalu.
2. Bapak Dr.A. Teguh Siswantoro, selaku Dekan Fakultas Teknologi Industri, Universitas Atma Jaya Yogyakarta.
3. Ibu Patricia Ardanari, S.Si.,M.T., selaku dosen pembimbing I yang telah membimbing dan memberikan masukan serta motivasi kepada penulis untuk menyelesaikan tugas akhir ini.
4. Bapak Prof. Ir. Suyoto, M.Sc., Ph.D., selaku dosen pembimbing II yang telah membimbing dan memberikan masukan serta motivasi kepada penulis untuk menyelesaikan tugas akhir ini.
5. Mama, yang selalu memberikan dukungan dan doa setiap saat.
6. Om alfi dan asis selaku sensei dan teman seperjuangan penulis dari awal kuliah sampai dengan penulisan tugas akhir.
7. Seluruh teman sukses dan gaje mulai dari Andre, Nelson, Asis, Alfi, Robert, Reno, Cipta, Vendy yang selalu memberikan semangat dan membantu penulis

selama berada dibangku kuliah.

8. Seluruh teman yang selalu memberikan semangat yang tidak penulis sebutkan satu per satu yang selalu memberikan dukungan selama berada dibangku kuliah.

Demikian laporan tugas akhir ini dibuat, dan penulis mengucapkan terima kasih kepada semua pihak. Semoga laporan ini dapat bermanfaat bagi pembaca.

Yogyakarta,



Andriel kiding

140707933

# DAFTAR ISI

JUDUL.....	
iHALAMAN	PENGESAHAN
.....ii	
PERNYATAAN ORISINALITAS & PUBLIKASI ILMIAH .....	iii
HALAMAN PERSEMBAH.....	iv
KATAPENGANTAR .....	v
DAFTAR ISI.....	vii
DAFTARGAMBAR .....	x
DAFTAR TABEL.....	xi
Intisari.....	xii
BAB I.....	1
PENDAHULUAN.....	1
1.1. <b>Error! Bookmark not defined.</b>	
1.2. <b>Error! Bookmark not defined.</b>	
1.3. <b>Error! Bookmark not defined.</b>	
1.4. <b>Error! Bookmark not defined.</b>	
1.5. <b>Error! Bookmark not defined.</b>	
BAB II.....	6
TINJAUAN PUSTAKA.....	6
BAB III.....	11
LANDASAN TEORI.....	11
3.1 <b>Error! Bookmark not defined.</b>	
3.2 <b>Error! Bookmark not defined.</b>	
3.3 <b>Error! Bookmark not defined.</b>	
3.4   3	
3.5 <b>Error! Bookmark not defined.</b>	
3.6 <b>Error! Bookmark not defined.</b>	



3.7	<b>Error! Bookmark not defined.</b>	
3.8	<b>Error! Bookmark not defined.</b>	
3.9	<b>Error! Bookmark not defined.</b>	
3.10	Confusion Matrix.....	16
BAB IV .....		18
METODOLOGI .....		18
4.1.	<b>Error! Bookmark not defined.</b>	
4.2.	<b>Error! Bookmark not defined.</b>	
4.2.1	<i>Data Cleansing</i> .....	20
4.2.2	<i>Stopword Removal</i> .....	22
4.2.3	<i>Stemming</i> .....	23
4.3.	<b>Error! Bookmark not defined.</b>	
4.4.	<b>Error! Bookmark not defined.</b>	
4.4.1	Klasifikasi.....	34
4.4.2	Validasi.....	35
4.5.	<b>Error! Bookmark not defined.</b>	
BAB V.....		38
HASIL DAN PEMBAHASAN.....		38
5.1.	<b>Error! Bookmark not defined.</b>	
5.2.	<b>Error! Bookmark not defined.</b>	
5.2.1.	<i>Data Cleansing</i> .....	42
5.2.2.	<i>Stopword Removal</i> .....	45
5.2.3.	<i>Stemming</i> .....	47
5.3.	<b>Error! Bookmark not defined.</b>	
5.4.	<b>Error! Bookmark not defined.</b>	
5.5.	<b>Error! Bookmark not defined.</b>	

5.6. **Error! Bookmark not defined.**

BAB VI .....63

KESIMPULAN DAN SARAN.....63

6.1. **Error! Bookmark not defined.**

6.2. **Error! Bookmark not defined.**

Daftar Pustaka .....65

LAMPIRAN .....67



## DAFTAR GAMBAR

<b>Gambar 4. 1</b> Alur Proses Analisis Sentimen.....	18
<b>GAMBAR 4. 2</b> ALUR PENGAMBILAN DATA DARI <i>TWITTER</i> .....	19
<b>GAMBAR 4. 3</b> ILUSTRASI PROSES STEMMING .....	23
<b>GAMBAR 4. 4</b> CONTOH HASIL DARI KOLOM SENTIMEN POSITIF .....	28
<b>GAMBAR 4. 5</b> CONTOH HASIL DARI KOLOM SENTIMEN NEGATIF .....	29
<b>GAMBAR 4.6</b> CONTOH HASIL PERHITUNGAN SKOR POSITIF DAN NEGATIF.....	30
<b>GAMBAR 4. 7</b> CONTOH HASIL PERHITUNGAN KELAS POSITIF, NEGATIF DAN NETRAL.....	31
<b>GAMBAR 4. 8</b> JUMLAH DATA POSITIF.....	31
<b>GAMBAR 4. 9</b> JUMLAH DATA NEGATIF .....	32
<b>GAMBAR 4. 10</b> JUMLAH DATA NETRAL .....	32
<b>GAMBAR 4. 11</b> CONTOH JUMLAH TOTAL DATA .....	33
<b>GAMBAR 4. 12</b> DIAGRAM SENTIMEN .....	33
<b>GAMBAR 4. 13</b> ILUSTRASI K- FOLD CROSS VALIDATION .....	34
<b>GAMBAR 4. 14</b> VISUALISASI CONFUSION MATRIX .....	35
<b>GAMBAR 4. 15</b> CONTOH VISUALISASI WORDCLOUD TES CPNS.....	36
<b>Gambar 5. 1</b> Data mentah Tes CPNS.....	39
<b>GAMBAR 5. 3</b> HASIL LABELING DATA TES CPNS.....	50
<b>GAMBAR 5. 5</b> GRAFIK KELAS SENTIMEN.....	52
<b>GAMBAR 5. 6</b> PERSENTASE DARI DATA TWEET TES CPNS.....	57
<b>GAMBAR 5. 8</b> DIAGRAM FREKUENSI KATA PADA DATA TWEET TES CPNS.....	60
<b>GAMBAR 5. 10</b> VISUALISASI WORDCLOUD.....	62

## DAFTAR TABEL

<b>TABEL 2. 1</b> TABEL PERBANDINGAN PENELITIAN.....	9
<b>Tabel 4. 1</b> Ilustrasi Proses Cleansing.....	21
<b>TABEL 4. 2</b> ILUSTRASI PROSES STOPWORDS REMOVAL.....	23
<b>TABEL 4. 3</b> HASIL PROSES STEMMING.....	24
<b>TABEL 4. 4</b> SCORING SENTIMEN.....	25
<b>TABEL 4. 5</b> KAMUS KATA POSITIF.....	26
<b>TABEL 4. 6</b> KAMUS KATA NEGATIF.....	26
<b>Tabel 5. 1</b> Contoh Data Tweet Mentah Tes CPNS.....	40
<b>TABEL 5. 3</b> HASIL PROSES DATA CLEANSING TES CPNS.....	43
<b>TABEL 5. 5</b> DATA HASIL PROSES STOPWORD REMOVAL TES CPNS.....	45
<b>TABEL 5. 7</b> DATA HASIL PROSES STEMMING TES CPNS.....	47
<b>TABEL 5. 9</b> HASIL AKURASI METODE NAÏVE BAYES CLASSIFIER PADA DATA TES CPNS.....	53
<b>TABEL 5. 10</b> MENUNJUKKAN HASIL AKURASI, PRECISION, RECALL, F-MEASURE DAN ROC DENGAN METODE NAÏVE BAYES CLASSIFIER PADA DATA TES CPNS.....	54
<b>TABEL 5. 13</b> FREKUENSI KEMUNCULAN KATA PADA DATA TES CPNS	59

# INTISARI

## ANALISIS SENTIMEN PUBLIK TERHADAP TES CPNS MELALUI MEDIA TWITER MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER

Intisari

Andriel kiding

140707933

Analisis sentiment adalah opini atau pendapat yang mengungkapkan persepsi seseorang dimana mengandung sentimen positif, netral atau negatif yang mana bisa dijadikan sebuah informasi, informasi ini diambil dari media sosial yang mana akan dilihat melalui postingan masyarakat. Media sosial yang akan digunakan sebagai penelitian untuk melakukan analisis adalah *Twitter*, dengan kemudahan dalam penggunaannya dan dengan pertimbangan data yang beragam sebagai alat penentu kebijakan. Berbagai macam opini masyarakat terhadap Tes CPNS merupakan salah satu cara mengetahui berbagai macam opini atau pendapat terhadap penyelenggaraan Tes CPNS tersebut.

Opini-opini pengguna *Twitter* terhadap Tes CPNS akan dianalisa dan menentukan pandangan positif, negatif, dan netral. Dari data yang berisi tentang opini-opini masyarakat mengenai tes CPNS ini, kemudian akan dilakukan sebuah proses *text preprocessing* seperti *Cleansing*, *Stopword removal* dan *Stemming* serta menggunakan sebuah metode *naïve bayes classifier* untuk membantu dalam mengklasifikasikan sebuah data berupa teks. Diharapkan dengan menggunakan metode *naïve bayes classifier*, menghasilkan nilai akurasi yang baik.

Hasil akhir penelitian terhadap pengujian klasifikasi data tweet Tes CPNS dengan data sebanyak 48.310 memperlihatkan bahwa dengan menggunakan metode klasifikasi *Naïve Bayes Classifier* menghasilkan nilai akurasi yang baik sebesar 98.0708% untuk Tes CPNS.

Kata kunci: Tes CPNS, Naïve Bayes Classifier, Analisis Sentimen

Dosen Pembimbing I : Patricia Ardanari, S.Si.,M.T.,

Dosen Pembimbing II : Prof. Ir. Suyoto, M.Sc., Ph.D

Jadwal Sidang Tugas Akhir : Selasa 25 Januari 202