

**ANALISIS SENTIMEN KOMENTAR PUBLIK DI  
YOUTUBE SATU PERSEN MENGENAI  
KETIDAKHARGAAN TERHADAP GURU DI  
INDONESIA DENGAN METODE NAIVE BAYES**

**SKRIPSI**



Oleh:

**YAASMIIN NUHAA ASA PUTRI  
201011402172**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PAMULANG  
TANGERANG SELATAN  
2024**

**ANALISIS SENTIMEN KOMENTAR PUBLIK DI  
YOUTUBE SATU PERSEN MENGENAI  
KETIDAKHARGAAN TERHADAP GURU DI  
INDONESIA DENGAN METODE NAIVE BAYES**

**SKRIPSI**

Diajukan Untuk Melengkapi Salah Satu Syarat  
Memperoleh Gelar Sarjana Komputer



Oleh:

**YAASMIIN NUHAA ASA PUTRI**

**201011402172**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PAMULANG  
TANGERANG SELATAN  
2024**



FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI TEKNIK INFORMATIKA

## LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : YAASMIIN NUHAA ASA PUTRI  
NIM : 201011402172  
Program Studi : Teknik Informatika  
Fakultas : Ilmu Komputer  
Jenjang Pendidikan : Strata 1

Menyatakan bahwa skripsi yang saya buat dengan judul:  
**ANALISIS SENTIMEN KOMENTAR PUBLIK DI YOUTUBE SATU PERSEN  
MENGENAI KETIDAKHARGAAN TERHADAP GURU DI INDONESIA  
DENGAN METODE NAÏVE BAYES**

1. Merupakan hasil karya tulis ilmiah sendiri, bukan merupakan karya yang pernah diajukan untuk memperoleh gelar akademik oleh pihak lain, dan bukan merupakan hasil plagiat.
2. Saya ijinkan untuk dikelola oleh Universitas Pamulang sesuai dengan norma hukum dan etika yang berlaku.

Pernyataan ini saya buat dengan penuh tanggung jawab dan saya bersedia menerima konsekuensi apapun sesuai aturan yang berlaku apabila di kemudian hari pernyataan ini tidak benar.

Tangerang Selatan, 13 Juli 2024

Yaasmiin Nu.  



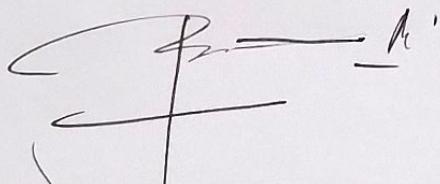

## LEMBAR PERSETUJUAN

NIM : 201011402172  
Nama : YAASMIIN NUHAA ASA PUTRI  
Program Studi : TEKNIK INFORMATIKA  
Fakultas : ILMU KOMPUTER  
Jenjang Pendidikan : STRATA 1  
Judul Skripsi : ANALISIS SENTIMEN KOMENTAR PUBLIK DI YOUTUBE SATU PERSEN MENGENAI KETIDAKHARGAAN TERHADAP GURU DI INDONESIA DENGAN METODE NAÏVE BAYES

Skripsi ini telah diperiksa dan disetujui oleh pembimbing untuk persyaratan sidang skripsi

Tangerang Selatan, 13 Juli 2024

Pembimbing



Drs. Dede Sahrul Bahri, M.Si  
NIDK : 8998070023

Mengetahui,

Ketua Program Studi



Achmad Udin Zailani, S.Kom., M.Kom.  
NIDN : 0429058303



FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI TEKNIK INFORMATIKA

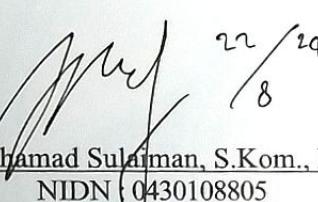
### LEMBAR PENGESAHAN

NIM : 201011402172  
Nama : YAASMIIN NUHAA ASA PUTRI  
Program Studi : TEKNIK INFORMATIKA  
Fakultas : ILMU KOMPUTER  
Jenjang Pendidikan : STRATA 1  
Judul Skripsi : ANALISIS SENTIMEN KOMENTAR PUBLIK DI YOUTUBE SATU PERSEN MENGENAI KETIDAKHARGAAN TERHADAP GURU DI INDONESIA DENGAN METODE NAÏVE BAYES

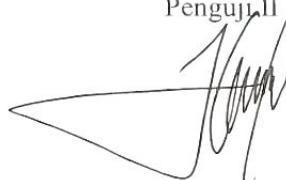
Skripsi ini telah dipertahankan di hadapan dewan pengaji ujian skripsi fakultas Ilmu Komputer, program studi Teknik Informatika dan dinyatakan LULUS.

Tangerang, ..... 23 Agustus ..... 2024

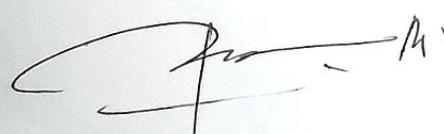
Pengaji I

  
Maulana Muhammad Sulaiman, S.Kom., M.Kom  
NIDN : 0430108805

Pengaji II

  
Iwan Giri Waluyo, S.Kom., M.Kom  
NIDN : 0404049101

Pembimbing

  
Drs. Dede Sahrul Bahri, M.Si  
NUDK : 8998070023

Mengetahui,

Ketua Program Studi

  
Achmad Udin Zailani, S.Kom., M.Kom.  
NIDN : 0429058303

## ***ABSTRACT***

*The role of teachers is crucial in the education of Indonesia's youth, yet the recognition they receive is often not commensurate with their efforts. This issue is discussed in a video by the youtube channel Satu Persen, which addresses various aspects of the teaching profession, including low salaries, the abundance of teachers with low qualifications, and the competency levels of Indonesian teachers. To analyze public sentiment regarding the undervaluation of teachers in Indonesia, the author collected 1020 comments from the video. After cleaning and labelling, 1014 comments were used for analysis. The Naïve Bayes method was employed, yielding a highest accuracy of 93.13% with a 90% training and 10% testing split. A comparative analysis was also conducted on the youtube channel Berita Satu, which discussed the National Education System Bill (RUU Sisdiknas) and low teacher salaries. From 948 comments, reduced to 927 after preprocessing, the highest accuracy achieved was 90.32% with an 80%:20% data split. Word cloud and Latent Dirichlet Allocation (LDA) identified three main topics, including low teacher salaries and inadequate subject mastery among teachers. The analysis shows that neutral sentiment dominates the discussions on both Satu Persen and Berita Satu, with 817 and 776 neutral comments, respectively. This research aims to provide insights into addressing the issues faced by teachers in Indonesia.*

**Keyword:** Public Sentiment, Naïve Bayes, Education Policy, Youtube, Teacher

## ABSTRAK

Peran guru begitu besar bagi Pendidikan di Nusantara, tanggung jawab besar yang guru emban sebagai pendidik bagi anak-anak di Indonesia hanya diberikan nilai yang tidak sebanding atas usaha yang mereka lakukan. Munculnya fenomena mengenai guru diramaikan pada konten Satu Persen yang memberikan diskusi tentang pendekatan guru dimulai dari: fenomena guru digaji rendah, kenapa guru digaji rendah dibandingkan profesi lain, guru: *supply* banyak tapi kualifikasi rendah, nilai kompetensi guru Indonesia, *skill* yang harus dipelajari siswa dan guru. Menganalisis sentimen publik mengenai ketidakhargaan terhadap guru di Indonesia melalui komentar video pada akun youtube Satu Persen. Penulis mendapatkan 1020 komentar yang lalu dibersihkan dan dilabeli sehingga menghasilkan 1014 data. Data ini dianalisis menggunakan metode *naïve bayes*, dengan akurasi tertinggi mencapai 93,13% pada pembagian data 90% untuk pelatihan dan 10% untuk pengujian, pada akun pembandingnya yaitu konten youtube Berita Satu mengenai RUU Sisdiknas dan guru digaji rendah diperoleh komentar 948 lalu berkurang menjadi 927 setelah tahap *preprocessing*, menghasilkan akurasi 90.32% pada pembagian 80%:20%. *Word cloud* dan *Latent Dirichlet Allocation* (LDA) mengidentifikasi 3 topik utama, termasuk mengenai gaji guru rendah dan banyak guru yang kurang dalam penguasaan materi. Analisis menunjukkan bahwa sentimen netral adalah yang paling dominan pada konten Satu Persen dan akun pembanding Berita Satu, dengan 817 komentar Satu Persen dan 776 komentar Berita Satu yang dilabeli sebagai netral, oleh karena itu peneliti menganalisis hasil komentar untuk memberikan pertimbangan dalam mengatasi permasalahan guru di Indonesia.

**Kata Kunci:** Sentimen Publik, *Naïve Bayes*, Kebijakan Pendidikan, Youtube, Guru

## KATA PENGANTAR

Puji Syukur Alhamdulillah atas kehadirat Allah SWT dan berkah karunia-Nya, Penulis dapat menyelesaikan skripsinya yang berjudul “ANALISIS SENTIMEN KOMENTAR PUBLIK DI YOUTUBE SATU PERSEN MENGENAI KETIDAKHARGAAN TERHADAP GURU DI INDONESIA DENGAN METODE NAÏVE BAYES”, dapat menyelesaikan tugas akhir ini dengan tepat waktu sehingga memenuhi syarat untuk menyelesaikan Program Studi Teknik Informatika Jenjang Strata 1 di Universitas Pamulang, Tangerang Selatan.

Pada masa penulisan skripsi berlangsung, penulis menyadari banyak keterbatasan dan masih jauh dari kata sempurna. Dengan penuh kerendahan hati, penulis ingin menyampaikan ucapan terima kasih terhadap berbagai pihak yang sangat membantu dalam penyelesaian skripsi. Penulis sampaikan terima kasih kepada :

1. Bapak **Dr. Pranoto, S.E., M.M.**, selaku ketua Yayasan Sasmita Jaya.
2. Bapak **Dr. E, Nurzaman, M.M., M.Si.**, selaku Rektor Universitas Pamulang.
3. Bapak **Yan Mitha Djaksana, S.Kom., M.Kom.**, selaku Dekan Fakultas Ilmu Komputer Universitas Pamulang.
4. Bapak **Achmad Udin Zailani, S.Kom., M.Kom.**, selaku Ketua Program Studi Teknik Informatika Universitas Pamulang.
5. Bapak **Drs. Dede Sahrul Bahri, M.Si.**, selaku Dosen Pembimbing Skripsi yang selalu memberikan masukan hingga penyelesaian skripsi yang baik.
6. Kepada Ayah, Ibu, Kakak, dan keluarga tercinta yang selalu memberikan dorongan positif, bantuan, dan doa tulus untuk dapat menyelesaikan studi dengan baik.
7. Kepada Teman sekelas 08TPLP016 yang senantiasa selalu membantu selama perkuliahan berlangsung.
8. Kepada Teman rumah Neji, Buku, *Black Kid* yang selalu menemani mengerjakan skripsi hingga larut malam.
9. Kepada Kerabat yang senantiasa membantu dan memberi dukungan.
10. Kepada teman sepebimbangan yang selalu bersemangat dan selalu rajin mengingatkan untuk bimbingan hingga sampai pada bimbingan terakhir.
11. Kepada *Channel* Youtube Satu Persen yang menyediakan tema menarik mengenai guru sehingga dapat dijadikan judul dalam skripsi.
12. Kepada *Channel* Youtube Yuk Nyistem yang menyediakan cara mudah untuk dapat melakukan *Preprocessing* data.

Semoga Allah SWT membalas semua kebaikan dan memberi pahala yang berlimpah atas segala kebaikan dan doa yang diberikan. Penulis juga berharap semoga skripsi ini dapat bermanfaat nantinya untuk penelitian selanjutnya.

Tangerang Selatan, ..... 2024

Yaasmiin Nuhaa Asa Putri

## DAFTAR ISI

<b>LEMBAR PERNYATAAN .....</b>	<b>ii</b>
<b>LEMBAR PERSETUJUAN .....</b>	<b>iii</b>
<b>LEMBAR PENGESAHAN .....</b>	<b>iv</b>
<b>ABSTRACT .....</b>	<b>v</b>
<b>ABSTRAK .....</b>	<b>vi</b>
<b>KATA PENGANTAR .....</b>	<b>vii</b>
<b>DAFTAR ISI .....</b>	<b>ix</b>
<b>DAFTAR GAMBAR .....</b>	<b>xii</b>
<b>DAFTAR TABEL .....</b>	<b>xiv</b>
<b>DAFTAR SIMBOL .....</b>	<b>xvi</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1    Latar Belakang .....	1
1.2    Identifikasi Masalah .....	4
1.3    Rumusan Masalah .....	5
1.4    Batasan Penelitian .....	5
1.5    Tujuan Penelitian.....	6
1.6    Manfaat Penelitian.....	6
1.7    Metodologi Penelitian .....	7
1.8    Sistematika Penulisan.....	9
<b>BAB II TINJAUAN PUSTAKA .....</b>	<b>11</b>
2.1    Penelitian Terkait.....	11
2.2    Landasan Teori .....	15
2.2.1 <i>Text Mining</i> .....	16
2.2.2 <i>Data Mining</i> .....	16
2.2.3    Analisis Sentimen .....	17
2.2.4    Guru .....	18
2.2.5    Youtube .....	20
2.2.6    Konten Youtube.....	22
2.2.7 <i>Channel</i> Satu Persen .....	24
2.2.8    Youtube Berita Satu Mengenai Guru Digaji Sangat Rendah .....	26
2.2.9    Instant Data Scraper .....	28
2.2.10    Python .....	29

2.2.11	Google Colab .....	30
2.2.12	Microsoft Excel.....	31
2.2.13	Metode <i>Naïve Bayes</i> .....	31
2.2.14	<i>Confusion Matrix</i> .....	32
2.3	Kerangka Pemikiran .....	35
2.3.1	<i>Scraping</i> Data Komentar.....	35
2.3.2	Data Mentah.....	35
2.3.3	<i>Preprocessing</i> .....	36
2.3.4	<i>Term Frequency</i> .....	37
2.3.5	<i>Splitting</i> data .....	37
2.3.6	Data Latih.....	37
2.3.7	Data Uji .....	38
2.2.8	Visualisasi <i>Word Cloud</i> .....	38
<b>BAB III METODE PENELITIAN</b>	.....	<b>39</b>
3.1	Analisa Kebutuhan .....	39
3.1.1	Kebutuhan Sistem .....	39
3.2	Perancangan Penelitian.....	41
3.2.1	Data Set .....	41
3.2.2	<i>Cleaning</i> .....	42
3.2.3	<i>Labelling</i> .....	43
3.2.4	<i>Case Folding</i> .....	47
3.2.5	<i>Tokenizing</i> .....	52
3.2.6	<i>Filtering / Stopwords</i> .....	55
3.2.7	<i>Stemming</i> .....	57
3.2.8	<i>Term Frequency</i> .....	60
3.2.9	TF-IDF .....	62
3.3	Perancangan Model Klasifikasi .....	64
3.4	Evaluasi Model Klasifikasi .....	65
<b>BAB IV HASIL DAN PEMBAHASAN</b>	.....	<b>66</b>
4.1	Data Set .....	66
4.2	<i>Preprocessing</i> Data.....	67
4.3	<i>Splitting</i> Data .....	67
4.4	Pengujian Klasifikasi Data .....	70
4.4.1	Pengujian Kelas Pertama Youtube Satu Persen.....	70

4.4.2	Pengujian Kelas Kedua Youtube Satu Persen .....	75
4.4.3	Pengujian Kelas Ketiga Youtube Satu Persen .....	79
4.4.4	Pengujian Kelas Pertama Youtube Berita Satu .....	83
4.4.5	Pengujian Kelas Kedua Youtube Berita Satu .....	87
4.4.6	Pengujian Kelas Ketiga Youtube Berita Satu.....	91
4.5	Hasil Akurasi Pengujian .....	95
4.6	Visualisasi.....	98
<b>BAB V</b>	<b>PENUTUP .....</b>	<b>100</b>
5.1	Kesimpulan.....	100
5.2	Saran.....	101
<b>DAFTAR PUSTAKA.....</b>		<b>103</b>
<b>LAMPIRAN.....</b>		<b>106</b>

## DAFTAR GAMBAR

Gambar 2.1 Gaji Guru Terendah.....	20
Gambar 2.2 Youtube .....	20
Gambar 2.3 Survei Pengguna Youtube.....	21
Gambar 2.4 Konten Youtube .....	22
Gambar 2.5 Konten Youtube Netizen.....	23
Gambar 2.6 <i>Channel</i> Satu Persen .....	24
Gambar 2.8 <i>Channel</i> Metro Tv .....	26
Gambar 2.9 Instant Data Scraper .....	28
Gambar 2.10 Python .....	29
Gambar 2.11 Google Colab.....	30
Gambar 2.12 Microsoft Excel .....	31
Gambar 2.13 <i>Confusion Matriks</i> .....	32
Gambar 2.14 Kerangka Pemikiran.....	35
Gambar 3.1 Data Set .....	41
Gambar 3.2 Tahap <i>Cleaning</i> .....	42
Gambar 3.3 <i>Code</i> Python <i>Cleaning</i> .....	42
Gambar 3.4 <i>Code</i> Python <i>Labelling</i> .....	43
Gambar 3.5 <i>Code</i> Python <i>Case Folding</i> .....	48
Gambar 3.6 Tahap <i>Case Folding</i> .....	48
Gambar 3.7 Tahap <i>Tokenizing</i> .....	52
Gambar 3.8 Tahap <i>Filtering</i> .....	55
Gambar 3.9 Tahap <i>Stemming</i> .....	57
Gambar 3.10 <i>Code</i> Python <i>Term Frequency</i> .....	61
Gambar 3.11 Tahap <i>Term Frequency</i> .....	61
Gambar 3.12 <i>Code</i> Python TF-IDF.....	63
Gambar 4.1 Hasil <i>Preprocessing</i> .....	67
Gambar 4.2 <i>Confusion Matriks</i> Pertama Satu Persen .....	71
Gambar 4.3 <i>Classification Matriks</i> Pertama Satu Persen .....	71
Gambar 4.4 <i>Confusion Matriks</i> Kedua Satu Persen.....	75
Gambar 4.5 <i>Classification Matriks</i> Kedua Satu Persen .....	76

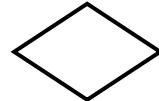
Gambar 4.6 <i>Confusion Matriks</i> Ketiga Satu Persen .....	80
Gambar 4.7 <i>Classification Report</i> Ketiga Satu Persen .....	80
Gambar 4.8 <i>Classification Report</i> Pertama Berita Satu.....	84
Gambar 4.9 <i>Confusion Matriks</i> Pertama Berita Satu.....	84
Gambar 4.10 <i>Confusion Matriks</i> Kedua Berita Satu.....	88
Gambar 4.11 <i>Classification Matriks</i> Kedua Berita Satu.....	88
Gambar 4.12 <i>Confusion Matriks</i> Ketiga Berita Satu .....	92
Gambar 4.13 <i>Classification Report</i> Ketiga Berita Satu .....	92

## DAFTAR TABEL

Tabel 3.1 Perangkat Keras .....	40
Tabel 3.2 Perangkat Lunak .....	40
Tabel 3.3 Hasil <i>Labelling</i> .....	44
Tabel 3.4 Hasil <i>Case Folding</i> .....	49
Tabel 3.5 Hasil <i>Tokenizing</i> .....	53
Tabel 3.6 Hasil <i>Filtering/ Stopword</i> .....	56
Tabel 3.7 Hasil <i>Stemming</i> .....	58
Tabel 3.8 Hasil <i>Term Frequency</i> .....	62
Tabel 3.9 Hasil TF-IDF.....	63
Tabel 4.1 Pembagian Data Pertama Satu Persen .....	68
Tabel 4.2 Pembagian Data Pertama Berita Satu .....	68
Tabel 4.3 Pembagian Data Kedua Satu Persen .....	69
Tabel 4.4 Pembagian Data Kedua Berita Satu .....	69
Tabel 4.5 Pembagian Data Ketiga Satu Persen .....	70
Tabel 4.6 Hasil <i>Precision</i> Pertama Satu Persen .....	72
Tabel 4.7 Hasil <i>Recall</i> Pertama Satu Persen .....	73
Tabel 4.8 Hasil <i>F1 Score</i> Pertama Satu Persen.....	74
Tabel 4.9 Hasil <i>Precision</i> Kedua Satu Persen.....	77
Tabel 4.10 Hasil <i>Recall</i> Kedua Satu Persen.....	78
Tabel 4.11 Hasil <i>F1 Score</i> Kedua Satu Persen .....	79
Tabel 4.12 Hasil <i>Precision</i> Ketiga Satu Persen .....	81
Tabel 4.13 Hasil <i>Recall</i> Ketiga Satu Persen.....	82
Tabel 4.14 Hasil <i>F1 Score</i> Ketiga Satu Persen .....	83
Tabel 4.15 <i>Precision</i> Pertama Berita Satu .....	85
Tabel 4.16 <i>Recall</i> Pertama Berita Satu .....	86
Tabel 4.17 <i>F1 Score</i> Pertama Berita Satu .....	87
Tabel 4.18 <i>Precision</i> Kedua Berita Satu .....	89
Tabel 4.19 <i>Recall</i> Kedua Berita Satu .....	90
Tabel 4.20 Hasil <i>F1 Score</i> Ketiga Satu Persen .....	91
Tabel 4.21 <i>Precision</i> Kedua Berita Satu .....	93

Tabel 4.22 <i>Recall</i> Kedua Berita Satu .....	94
Tabel 4.23 Hasil <i>F1 Score</i> Ketiga Satu Persen .....	95
Tabel 4.24 Keseluruhan Hasil Pengujian Satu Persen .....	96
Tabel 4.25 Keseluruhan Hasil Pengujian Berita Satu .....	96
Tabel 4.26 Batasan Nilai Akurasi .....	97
Tabel 4.27 Kategori Penilaian.....	97

## DAFTAR SIMBOL

No	Gambar	Nama Simbol	Keterangan
1.		<i>Terminator</i>	Digunakan sebagai awalan atau akhiran suatu kegiatan.
2.		<i>Processing</i>	Digunakan sebagai pernyataan suatu proses yang dilakukan komputer.
3.		<i>Decision</i>	Simbol yang menunjukkan dua pilihan kondisi.
4.		<i>Input/output</i>	Untuk menyatakan proses <i>input</i> atau <i>output</i> .
5.		<i>Flow</i>	Untuk menggabungkan satu simbol dengan simbol lain.