

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

Yaasmiin Nuhaa Asa Putri

Universitas Pamulang, Tangerang Selatan  
e-mail: yaasmiinnuhaa@gmail.com

### 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 Naive 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

## 1. PENDAHULUAN

Satu Persen merupakan suatu *startup* yang bergerak pada bidang Pendidikan di Indonesia dengan tujuannya, yaitu memberikan pengajaran hidup bagi setiap individu agar mampu mencapai kehidupan yang baik. Dalam menjalankan misinya Satu Persen lebih mengacu pada kesehatan mental, pengembangan diri, hingga edukasi *life-skills*. Untuk menjalankan misinya mereka juga menawarkan berbagai layanan, seperti membuka konsultasi Bersama mentor, psikolog, *workshop*, dan kelas *online*.

Pertama kali teretusnya ide-ide konten youtube Satu Persen ialah dimulai dari *Channel* Youtube pribadi dari *Founder* Satu Persen sendiri, yaitu adalah Ifandi Khainur Rahim. Video-video yang dia bagikan menarik minat pengguna Youtube hingga menjadi awal mula Satu Persen didirikan secara resmi, sebagai *startup* oleh Evan dan Rizky pada pertengahan tahun 2019. Saat ini Satu Persen terus berkembang dalam mengatasi permasalahan *mental health* hingga *self-development*. Hal ini dapat dilihat dari konten Youtube Satu Persen yang dapat menjangkau semua kalangan yang ingin memiliki *growth mindset*. Cara penyampaian setiap konten Satu Persen dikemas dengan ringan dan menarik ditambah tema yang diambil selalu disesuaikan dengan fakta yang selama ini dekat disekitar kita. Salah satu kontennya yang menarik perhatian dibagikan tanggal 1 Agustus 2023 di youtube mengenai “Guru di Indonesia Gak Dihargai! Ini Alasannya”. Konten tersebut telah ditonton sebanyak 132.346 ribu kali. Hasilnya memberikan banyak pemikiran mengenai apa yang sedang terjadi di negeri ini. Fakta bahwa seorang guru yang dipandang sebagai pahlawan tanpa tanda jasa memberi celah besar terhadap keberhargaan dari profesi tersebut.

Tanggung jawab besar yang guru emban sebagai pendidik bagi anak-anak di Indonesia hanya diberikan nilai yang tidak sebanding atas usaha yang mereka lakukan. Terkait peran guru yang begitu besar bagi pendidikan di Nusantara, guru juga diharuskan untuk menjaga profesionalitasnya sebagai seorang pendidik yang terdidik maksudnya ialah guru dapat menguasai bidang keilmuan serta pengetahuan yang sesuai pada bidangnya [1]. Sebagai guru banyak hal juga yang harus dikuasai, perihal itu konten satu persen 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.

Pemilihan konten dengan tema guru yang selama ini selalu dibicarakan membuat banyak pihak ikut meramaikan kolom komentar dari konten ‘Satu Persen’ ini, dengan mendapat 1020 komentar dan 5,1 ribu menyukai videonya. Terdapat hal-hal yang menurut mereka, tidak sesuai dengan apa yang terjadi sebenarnya. Apalagi, sekarang telah ada bentuk wacana mengenai pengkualifikasian guru berdasarkan pada peraturan baru Menteri Pendidikan Nadien Makarim, bahwa sekolah dapat mengakses calon guru yang ingin mengajar disekolahnya, setelah terkonfirmasi direkrut sekolah, guru akan secara otomatis ditetapkan sebagai pegawai negeri melalui sistem *marketplace* guru.

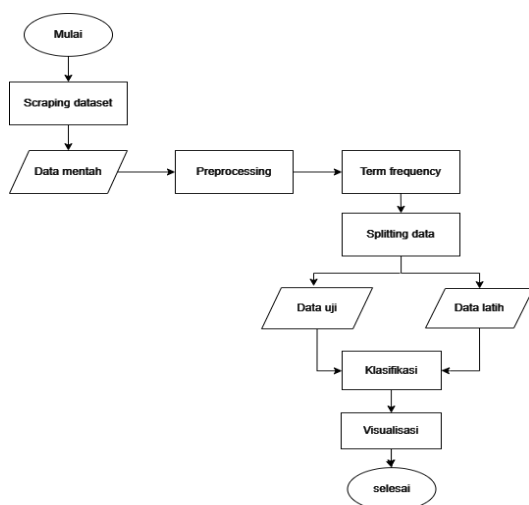
Analisis sentimen berguna agar suatu bentuk informasi mengenai alasan dari guru tidak dihargai dapat ditemukan melalui kumpulan data yang tidak terstruktur pada kolom komentar youtube Satu Persen, sehingga dari data tersebut akan menghasilkan pengaruh respon yang dapat dipertimbangkan untuk kemajuan guru dimasa depan. Analisis berguna agar kita yang tidak memahami konten yang ditampilkan dapat lebih mencerna melalui lengkapnya informasi nilai yang ada. Tetapi tanpa dilakukan tahapan analisis akan timbul keresahan mengenai nilai-nilai yang sebenarnya perlu diperhatikan tersebut. Stigma yang bermunculan mengenai guru dituangkan pada konten Youtube Satu Persen “Guru di Indonesia Gak Dihargai! Ini Alasannya” menghasilkan sudut pandang yang berbeda sehingga sulit untuk ditemukan hasil yang positif untuk menjelaskan alasan ketidakhargaan guru di Indonesia tersebut [2].

Tanggapan yang diperoleh dari konten youtube Satu Persen ini berasal dari berbagai kalangan dan nantinya akan dibagi menjadi tiga, yaitu positif, negatif, dan juga netral sebagai bentuk klasifikasi untuk meningkatkan pemahaman opini publik terhadap faktor-faktor yang mempengaruhi ketidakhargaan guru di Indonesia. Berdasarkan alasan tersebut penulis melihat berbagai komentar yang ada dan mendefinisikan untuk tanggapan yang sifatnya positif lebih mengarah pada reaksi mereka dalam menyetujui bahwa guru perlu meningkatkan kompetensi sehingga dengan keahliannya mampu untuk memberikan pembelajaran yang baik. Pada mereka yang memberi tanggapan bahwa generalisasi tentang guru pada konten tersebut tidak sesuai maka masuk dalam tanggapan yang bersifat negatif. Tanggapan yang merasa baik-baik saja pada masalah ketidakhargaan pada guru tentang masalah gaji termasuk sifatnya netral.

Kumpulan tanggapan diperoleh semua berdasarkan komentar dari youtube Satu Persen yang telah dilakukan *scraping* dengan menggunakan instant data scraper. Hasil yang didapat menjadi 1020 komentar yang diperoleh dari bulan 1 Agustus 2023 hingga 5 Agustus 2024, komentar tersebut akan diolah dengan analisis sentimen metode *Naïve Bayes*. Metode *Naïve Bayes* sendiri merupakan bentuk metode yang cepat dengan tingkat keakuratannya yang tinggi. Tujuan dilakukannya penelitian tersebut, ialah agar penulis dapat melihat berbagai komentar sebagai analisis yang bentuknya dapat menghindari prasangka yang tidak dilandasi pada fakta dan bukti, perlakuan generalisasi terhadap anggota tertentu, dan menemukan solusi dalam mengatasi stigma terhadap guru di Indonesia.

## 2. METODE

Pada proses penelitian nantinya akan dibuat proses langkah menjadi pembagian tertentu yaitu *scraping*, *cleaning*, *labelling*, *case folding*, *filtering/stopwords*, *tokenizing*, *stemming*, *term frequency*, *splitting data*, *klasifikasi*, dan *visualisasi word cloud*.



Gambar 2.1 Kerangka Pemikiran

Tahap pertama adalah *Scraping*, ini merupakan tahap yang akan digunakan untuk mengambil data dengan pengumpulan suatu data dari *database*, sebelum data tersebut diolah atau dianalisis. Pada pengumpulan yang dilakukan adalah mengunduh data dari komentar pada youtube[3]. Selanjutnya *Preprocessing* yang meliputi *cleaning*, *labelling*, *case folding*, *filtering*, *tokenizing*, dan *stemming*.

a. *Cleaning* adalah Tahap pembersihan, misalnya karakter selain huruf jadi para prosesnya akan

dihapus URL dari setiap komentar. Jadi tujuan untuk *cleaning* ini adalah membersihkan suatu komentar.

- b. *Labelling* merupakan tahap penentuan label dari setiap komentar dalam dokumen. Tujuannya memberikan bobot pengelompokkan polaritas sehingga menghasilkan sentimen positif, sentimen negatif, dan sentimen netral [4].
- c. *Case Folding* merupakan tahap pengubahan huruf yang terjadi misalkan terdapat huruf yang diubah menjadi huruf kecil semua. Langkahnya dengan memeriksa setiap karakter lalu nantinya akan ditemukan huruf kapital yang langsung dijadikan huruf kecil.
- d. *Filtering* adalah setiap kata akan muncul pada kategorinya dan pada kata - kata yang tidak bermakna untuk pengkategorian dokumen, nantinya kata- kata tersebut akan dibuang.
- e. *Tokenizing* adalah Tahap pemrosesan untuk memisahkan kalimat menjadi beberapa bagian kata yang biasa disebut dengan *token*. Tahap ini dilakukan agar mempermudah pada bagian proses selanjutnya [5].
- f. *Stemming* merupakan pengubahan dari kata berimbuhan menjadi kata dasar. bertujuan untuk menghemat waktu pengklasifikasian [6].

Tahap *Term Frequency*, bentuk dari proses perhitungan mengenai jumlah term pada suatu dokumen. Nantinya akan terjadi pembobotan atas setiap kata yang akan muncul di sebuah dokumen [7]. Kemudian tahap *Splitting data* merupakan pembagian data pada dokumen menjadi dua bagian, yaitu dinamakan set pelatihan dan set pengujian. Tujuannya agar pengembangan model analisis sentimen dapat bekerja dengan baik dalam memprediksi sentimen data. Data akan dibagi menjadi 3 bagian untuk dibandingkan bagian manakah yang memiliki tingkat akurasi yang paling tinggi. Pembagian data akan dilakukan pada 2 akun, yaitu akun youtube Satu Persen dan akun pembandingnya youtube Berita Satu. Pada bagian pertama data akan dibagi menjadi 2 set, yaitu Data latih dan Data uji.

Setelah data dibagi dengan *Splitting data*, data akan diklasifikasi dengan Metode klasifikasinya *Naïve Bayes Classifier*. Setelah proses pelatihan selesai, model akan diuji menggunakan data uji untuk mengukur seberapa baik model dapat mengenali dan mengklasifikasikan data baru. Metode klasifikasinya *Naïve Bayes Classifier* sebuah algoritma yang sederhana namun efektif untuk klasifikasi, terutama pada tugas-tugas

Metode klasifikasi *naïve bayes* berada direntang nilai yang sangat baik artinya metode mampu untuk mendapatkan jumlah ketepatan pada kelas pengujian ketiga[10].

**Tabel 3.2** Berita Satu

Pembagian	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>
70% : 30%	88.53%	83.29%	85.56%	76.31%
<b>80% : 20%</b>	90.32%	79.61%	87.48%	87.65%
90% : 10%	87.09%	63.06%	80.48%	79.94%

Pada akun pembanding dari youtube Berita Satu yang berjudul “Mengapa Guru di Indonesia Digaji Sangat Rendah?” mendapat hasil *scraping* sebanyak 948 komentar kemudian berkurang setelah proses *cleaning* menjadi 927 komentar, hasil *labelling* terbagi ke dalam 3 bagian, yaitu 134 positif, 17 negatif, 776 netral. Mengurangi jumlah data yang dianalisis membuat pengklasifikasian menjadi lebih efisien dan efektif pada metode pengklasifikasian *naïve bayes*. Dipilihnya data menggunakan salah satu metode pemilihan fitur yang sering digunakan, yaitu *document frequency*, metode ini telah dilakukan setelah proses *stemming* selesai[9].

Pengujian data akan menggunakan *confusion matriks* sebagai bentuk evaluasi model sehingga akan menghasilkan nilai *accuracy*, *precision*, *recall*, dan *f1 Score*. Tiga pengujian akan dilakukan mulai dari pembagian 70% dengan 30%, 80% dengan 20%, dan 90% dengan 10%. Pengujian data dilakukan pada 2 akun, yaitu akun youtube Satu Persen dan akun pembandingnya Berita Satu.

**Gambar 3.1** *Word Cloud*

Tampilan *word cloud* didominasi ketiga topik dibawah ini yang diperoleh dari data sentimen positif pada konten ketidakhargaan terhadap guru di Indonesia pada media youtube Satu Persen. Hasil pengumpulan topik didapatkan dari *Latent Dirichlet Allocation* (LDA) sebagai pustaka pemodelan topik juga pendistribusian probabilitas dalam dokumen sehingga bisa mendapatkan komentar relevan. Berikut merupakan topik dan komentar relevan tersebut:

a. Topik 1  
[guru, 'gaji', 'rendah', 'ulang', 'pikir', 'sistem',  
'butuh', 'kompetensi', 'kerja', 'naik']  
Pertimbangan diambil dari kumpulan komentar  
relevan:

Gaji guru, terutama guru honorer, seringkali berada di bawah standar kelayakan hidup, yang mengakibatkan kesulitan finansial bagi mereka. Perlu ada kebijakan untuk meningkatkan gaji guru agar setara dengan profesi lain yang memiliki fungsi sosial penting.

b. Topik 2

['guru', 'sekolah', 'didik', 'ajar', 'swasta"emang', 'bahas', 'banyak', 'udah', 'kaya']

Pertimbangan diambil dari kumpulan komentar relevan:

Sistem pendidikan saat ini membutuhkan perbaikan untuk memastikan bahwa guru memiliki kompetensi yang tinggi. Bisa dicapai melalui pelatihan berkelanjutan dan insentif yang mendorong peningkatan kompetensi profesional.

c. Topik 3

['guru', 'gak', 'honorer', 'pns', 'sistem', 'kayak', 'org', 'kalo', 'gitu', 'sd']

Pertimbangan diambil dari kumpulan komentar relevan:

Birokrasi yang berbelit sering menjadi penghalang bagi guru untuk mendapatkan hak-hak mereka. Perlu adanya kebijakan yang lebih terbuka dan transparan dalam menyediakan kesempatan karier yang lebih baik bagi guru, termasuk pengangkatan guru honorer menjadi PNS.

#### 4. KESIMPULAN

Dari perolehan data youtube Satu Persen yang didapat melalui proses *scraping* sebanyak 1020 komentar yang melalui tahap *cleaning* lalu berkurang 1014 komentar kemudian dilanjutkan tahap *labelling*, *casefolding*, *tokenizing*, *stopwords*, dan *stemming*. Berdasarkan data keseluruhan hasil terbagi menjadi 136 sentimen positif, 61 sentimen negatif, 817 sentimen netral. Dilihat dari hasil dapat dikatakan publik lebih banyak berkomentar netral mengenai ketidakhargaan guru.

Evaluasi model dan *naïve bayes* digunakan sebagai metode klasifikasi dari analisis sentimen komentar publik youtube satu persen. Hasil akurasi tertinggi mencapai 93.13% untuk pengujian kelas ketiga dengan pembagian 90%:10%, artinya model sangat handal untuk mengidentifikasi sentimen. Pada akun pembanding dari youtube Berita Satu didapatkan hasil *scraping* sebanyak 948 komentar, kemudian berkurang setelah proses *cleaning* menjadi 927 komentar. Hasil *labelling* terbagi ke dalam tiga bagian, yaitu 134 positif, 17 negatif, dan 776 netral. Hasil akurasi tertinggi ada pada pembagian kedua sebesar 80%:20%, dengan

metode *naïve bayes* yang mencapai akurasi 90.32%. Nilai akurasi ini masuk ke dalam rentang nilai klasifikasi performa yang sangat baik. Perbedaan akurasi ini terjadi pada perbandingan yang berbeda karena karakteristik data yang berbeda, yang mempengaruhi performa model dalam menguji.

Perolehan hasil sentimen positif sebanyak 136 komentar ditampilkan dengan *word cloud* untuk mendapatkan topik dominan dari konten ketidakhargaan terhadap guru di Indonesia pada media youtube satu persen. Sebanyak 3 topik yang didapat melalui pengumpulan data menggunakan pustaka pemodelan topik bernama *Latent Dirichlet Allocation* (LDA), Pustaka ini mampu mengambil data sebagai pertimbangan bagi kemajuan guru di Indonesia berdasarkan beberapa komentar relevan yang cocok dengan setiap topik yang diambil.

#### 5. DAFTAR PUSTAKA

- [1] S. Susetyo, R. Basuki, and N. Noermanzah, "Peningkatan Profesionalisme Guru Bahasa Indonesia di Kabupaten Musirawas Sumatera Selatan Melalui Pelatihan Menulis Artikel Jurnal Ilmiah," ABDI: Jurnal Pengabdian dan Pemberdayaan Masyarakat, vol. 2, no. 1, pp. 28–34, Jan. 2020, doi: 10.24036/abdi.v2i1.35.
- [2] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "Analisis Sentimen Aplikasi Ruang Guru DI Twitter Menggunakan Algoritma Klasifikasi," Jurnal Teknoinfo, vol. 14, no. 2, p. 115, Jul. 2020, doi: 10.33365/jti.v14i2.679.
- [3] D. Oleh, A. Mondaref, and J.: 17523166, "Analisis Sentimen Pada Media Sosial Instagram Klub Persija Jakarta Menggunakan Metode Naive Bayes," Yogyakarta, 2023.
- [4] S. Mujahidin, "Impelementasi Analisis Sentimen Masyarakat Mengenai Kenaikan Harga BBM Pada Komentar Youtube Dengan Metode Gaussian Naive Bayes," vol. 10, 2022, [Online]. Available: <http://ejournal.unp.ac.id/index.php/voteknika/index>
- [5] L. Setiyani, M. Wahidin, D. Awaludin, and S. Purwani, "Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Data Mining Naïve Bayes: Systematic Review," Faktor Exacta, vol. 13, no. 1, p. 35, Jun. 2020, doi: 10.30998/faktorexacta.v13i1.5548.
- [6] M. I. Ahmadi, D. Gustian, and F. Sembiring, "Analisis Sentiment Masyarakat terhadap Kasus Covid-19 pada Media Sosial Youtube dengan Metode Naive bayes," 2021.
- [7] M. Diki Hendriyanto, A. A. Ridha, and U. Enri, "Analisis Sentimen Ulasan Aplikasi Mola Pada Google Play Store Menggunakan Algoritma

- Support Vector Machine,” Journal of Information Technology and Computer Science (INTECOMS), vol. 5, no. 1, Jun. 2022.
- [8] A. W. V. Hutabarat, N. L. S. S. Adnyani, and K. Suryadi, “Analisis Sentimen Data Ulasan Pengguna MyPertamina di Twitter dengan Metode Text Mining,” Jurnal Rekayasa Sistem Industri, vol. 13, no. 1, pp. 145–154, Apr. 2024, doi: 10.26593/jrsi.v13i1.6958.145-154.
- [9] A. Bijaksana, P. Negara, H. Muhandi, and I. M. Putri, “Analisis Sentimen Maskapai Penerbangan Menggunakan Metode Naive Bayes dan Seleksi Fitur Information Gain,” vol. 7, no. 3, pp. 599–606, Apr. 2020, doi: 10.25126/jtiik.202071947.
- [10] A. I. Sang, E. Sutoyo, and I. Darmawan, “Analisis Data Mining Untuk Klasifikasi Data Kualitas Udara DKI Jakarta Menggunakan Algoritma Decision Tree dan Support Vector Machine,” Bandung, Oct. 2021. Accessed: Jul. 09, 2024. [Online].