

Analisis Sentimen Dosen terhadap Kebijakan Tunjangan Kinerja di Perguruan Tinggi Negeri Menggunakan Algoritma Naive Bayes

(Sentiment Analysis of Lecturers toward the Performance Allowance Policy at Public Universities Using the Naive Bayes Algorithm)

Citra Fathia Palembang^{1*}, Emanuella M. C. Wattimena², Vyarlita Isqama Fataruba³

^{1,2,3}Program Studi Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas Pattimura
Jl. Ir. M. Putuhena, Ambon, 97233, Indonesia

*Corresponding author's e-mail: [*fpchiet@gmail.com](mailto:fpchiet@gmail.com)

Manuscript submitted:
17th February 2025

Manuscript revision:
20th March 2025

Accepted for publication:
29th May 2025

Abstract

This study aims to analyze lecturers' sentiments toward the performance allowance policy at public universities by applying the Naive Bayes algorithm. The data used consisted of 607 records collected through data crawling using Tweet-Harvest; however, after preprocessing, only 259 data points remained. For the initial labeling process, a simple sentiment analysis was performed to classify the data as either positive or negative sentiment. The labeling results showed a dominance of negative sentiment at 94.57%, while positive sentiment accounted for only 5.43%, reflecting a high level of dissatisfaction among lecturers regarding the policy's implementation. The Naive Bayes model was trained using a data split of 80% for training and 20% for testing. The model evaluation demonstrated an accuracy of 83.78%, with an F1-Score of 90.32% for negative sentiment and 50.00% for positive sentiment. The confusion matrix indicated that the prediction of negative sentiments was highly accurate, while the prediction of positive sentiments still needs improvement due to the imbalance in the data distribution. The results of this study indicate that the Naive Bayes algorithm is quite effective in classifying lecturers' opinions regarding the Tukin policy, particularly in detecting negative sentiments.

Keywords: Sentiment Analysis; Lecturers; Naive Bayes; Text Classification; Data Mining.



This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

How to cite this article:

C. Palembang, E. Wattimena, and V. Fataruba, "Analisis Sentimen Dosen Terhadap Kebijakan Tunjangan Kinerja (Tukin) di Perguruan Tinggi Negeri Menggunakan Algoritma Naive Bayes", *algorithm*, vol. 1, no. 1, pp. 21-28, May 2025.

Copyright © 2025

Journal homepage: <https://ojs3.unpatti.ac.id/index.php/algorithm>

Research Article [Open Access](#)

1. PENDAHULUAN

Perguruan tinggi negeri (PTN) di Indonesia telah mengalami berbagai reformasi, salah satunya adalah pemberlakuan sistem Tunjangan Kinerja (Tukin) bagi dosen. Kebijakan ini bertujuan untuk meningkatkan kinerja dosen, baik dalam bidang pendidikan, penelitian, pengabdian kepada masyarakat, maupun administrasi akademik[1]. Tukin dinilai sebagai instrumen motivasional dalam mempercepat tercapainya standar mutu pendidikan tinggi [2]. Namun, implementasi Tukin di kalangan dosen tidak lepas dari berbagai kontroversi. Beberapa penelitian menunjukkan adanya ketidakpuasan terkait ketidakadilan distribusi, ketidakjelasan indikator kinerja, serta beban administratif yang meningkat [3].

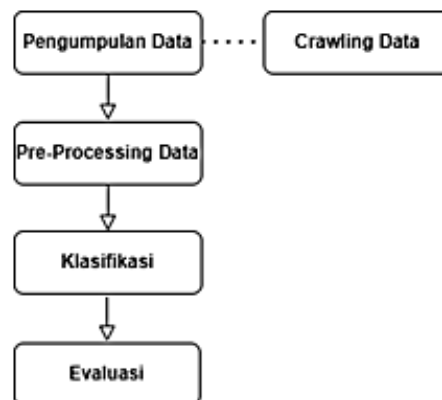
Seiring perkembangan media sosial, sentimen dosen terhadap kebijakan ini kini dapat dipetakan melalui analisis sentimen. Sentimen analisis adalah pendekatan berbasis data yang digunakan untuk mengidentifikasi sikap subjektif terhadap suatu isu [4],[5]. Sentimen analisis diklasifikasi ke dalam kategori positif, negatif, atau netral [6]. Metode yang digunakan dapat berbasis *lexicon* [7], *machine learning* [8], maupun *deep learning* [9].

Beberapa penelitian yang telah dilakukan terkait dengan analisis sentimen diantaranya peneliti yang membahas program kampus merdeka menghasilkan perolehan klasifikasi sentimen positif sebanyak 272 opini dan 229 sentimen negatif dengan rata-rata akurasi 60%, presisi 64%, recall 58% dan f1-score 58 [10], Penelitian analisis sentimen lainnya yang dilakukan terhadap pembelajaran daring dengan metode yang sama menghasilkan 30% sentimen positif, 69% sentimen negatif, dan 1 % netral [11], serta penelitian lainnya yang menggunakan metode yang sama tahun 2023 dengan analisis sentimen terkait dengan pemindahan ibu kota negara baru menghasilkan nilai akurasi sebesar 99,12%, *recall* untuk hasil negatif sebesar 98,37% dan positif 99,71%, kemudian untuk precision untuk hasil negatif sebesar 99,63%, dan 99,72% untuk hasil positif [12].

Dari beberapa penelitian sebelumnya yang telah dilakukan dengan menggunakan metode yang sama yaitu naïve bayes, penelitian ini juga bertujuan untuk menganalisis sentimen dosen terhadap kebijakan Tukin di PTN menggunakan metode yang sama. Metode naïve bayes memiliki kelebihan dibandingkan metode klasifikasi lainnya, yaitu memiliki kecepatan dan akurasi yang tinggi ketika diaplikasikan ke dalam basis data yang besar serta beragam [13].

2. METODE PENELITIAN

Terdapat beberapa tahapan yang dilakukan pada penelitian ini, yaitu pengumpulan data, preprocessing data, klasifikasi, ekstraksi data, evaluasi. Berikut tahapan yang digambarkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.1. Pengumpulan Data

Data yang digunakan yaitu data yang diperoleh dari media sosial x menggunakan teknik Crawling. Teknik Crawling adalah proses yang dilakukan oleh sebuah program komputer untuk mengumpulkan data dari berbagai sumber[14]. Data diambil dari media sosial x pada tanggal 25 April 2025 menggunakan Tweet-Harvest, yang merupakan sebuah alat berbasis playwright yang digunakan untuk mengumpulkan data x secara otomatis menggunakan bahasa pemrograman python yang diintegrasikan ke Google Colab [15]. Data yang diperoleh dari hasil *crawling* dengan kata kunci “tukin dosen” adalah sebanyak 607 Data. Selanjutnya, data mentah tersebut disimpan dalam format CSV agar mudah diolah dan dianalisis. Data ini kemudian melalui tahap pra-pemrosesan, seperti pembersihan teks dan penghapusan duplikasi, sebelum digunakan dalam analisis lebih lanjut.

2.2. Preprocessing Data

Selanjutnya adalah pembersihan data sebelum diklasifikasikan. Data mentah diolah untuk digunakan sebagai model dalam pembelajaran mesin melalui serangkaian tahapan dalam *preprocessing* data berupa *data cleansing*, *tokenization*, *stopwords removal*, dan *normalization* [12] menggunakan aplikasi RapidMiner. RapidMiner adalah platform yang menyediakan lingkungan terintegrasi untuk persiapan data, pembelajaran mesin, pembelajaran dalam penambangan teks, dan analisis prediktif [16].

2.3. TF-IDF

TF-IDF digunakan untuk mengetahui frekuensi dari istilah tertentu pada sebuah kata dalam dokumen (sekumpulan teks yang terstruktur) [10]. Metode ini tidak hanya menghitung banyaknya kemunculan suatu kata (*term frequency*), tetapi juga memperhitungkan seberapa penting kata tersebut dalam keseluruhan dokumen dengan memperhitungkan inverse document frequency. Dengan demikian, kata-kata umum yang sering muncul di banyak dokumen, seperti “dan” atau “yang”, akan memiliki bobot rendah, sedangkan kata-kata spesifik yang lebih jarang muncul akan memiliki bobot lebih tinggi. Pendekatan ini sangat berguna untuk representasi teks sebelum dilakukan analisis atau klasifikasi.

2.4. Klasifikasi

Proses klasifikasi dilakukan untuk mengelompokkan data teks hasil preprocessing ke dalam kategori sentimen positif atau negatif. Klasifikasi dilakukan menggunakan algoritma Naive Bayes yang merupakan salah satu algoritma *supervised learning* yang populer dalam analisis sentimen dan klasifikasi teks. Naive Bayes bekerja berdasarkan prinsip teorema Bayes dengan asumsi independensi antar fitur, sehingga perhitungan probabilitas dapat dilakukan dengan lebih sederhana namun tetap efektif. Keunggulan algoritma ini adalah efisiensi komputasi, kemudahan implementasi, dan performa yang cukup baik bahkan pada dataset yang besar. Oleh karena itu, Naive Bayes sering dijadikan baseline dalam penelitian analisis sentimen.

2.5. Evaluasi

Tahapan terakhir yaitu evaluasi *cross validation*, untuk menilai keakuratan sebuah model. Metode ini membagi data menjadi beberapa lipatan (*fold*), kemudian setiap lipatan secara bergantian digunakan sebagai data uji, sedangkan sisanya sebagai data latih. Teknik ini membantu mengurangi bias dan variansi dalam penilaian performa model, sehingga hasil evaluasi menjadi lebih andal. Selain itu, metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score* juga dapat digunakan untuk memberikan gambaran yang lebih menyeluruh mengenai performa model dalam mengklasifikasikan sentimen.

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Berdasarkan hasil crawling data x dengan kata kunci ‘tukin”, menghasilkan 607 data dan batasan hanya cuitan yang berbahasa indonesia dapat ditunjukkan pada Gambar 2. Data Crawling yang dihasilkan, kemudian disimpan ke dalam file dengan format csv.

index	conversation_id_str	created_at	favorite_count	full_text	id_str	image_url
225	1913906255430910066	Sun Apr 20 11:25:33 +0000 2025	9	@tukin_dosenASN @Kemdikisaintek @prabowo @prabowo benar Pak kawal kemendikisaintek terkait tukin dosen. Untuk memperoleh hak dosen sja harus memenuhi berbagai syarat² gila. Sementara tendik saja tnpa syarat dapat tukin 100%.	1913916945285742993	NaN
226	1913906255430910066	Sun Apr 20 10:43:04 +0000 2025	84	@Kemdikisaintek banget ini pak @prabowo Sudah susah2 berjuang biar tukin dosen tjair sekalnya Perpres udah turun harus diturunkan laai ke	1913906255430910066	https://obs.twimg.com/ext_tw_video_thumb/1913906212472832000/bu...

Gambar 2. Data Crawling

3.2. Preprocessing Data

3.2.1. Data Cleansing

Tahap preprocessing Data, diawali dengan Cleansing atau pembersihan data yaitu pembersihan data tweet / cuitan untuk menghilangkan karakter-karakter tidak penting, seperti; menghilangkan kata yang mengandung mention atau karakter “@”, *hashtag* atau karakter “#”, hyperlink, spasi berlebih, simbol2, emoji. Selain itu, membersihkan cuitan yang berulang atau *duplicate* dan baris-baris yang kosong. Maka data yang sudah dibersihkan terlihat pada Gambar 3 di bawah ini.

Sebelum data cleansing

Row No.	full_text
1	@tukin_dosenASN Jangan nambah aplikasi lagi ya...Tukin harusnya berdasarkan BKD saja..kalau BKD kan syaratnya jelas 12-16 SKS
2	@Sam_Ardi @tok_jos @samdonie @dwioktariyadi @hasyimmah @hayder_imron @alyapedia wkwkwk langsung lupa tukin
3	@opicksz @ganeshuor Bukan masalah penak. Soalnya kalau guru dgn tubel + beasiswa dianggap cuti tukin ga cair lakyo podo ekonomis e karo beasi...
4	Trs kalo kamu blm serti bakal ada tukin yg besarnya setengah dari gapok kamu (-tunjangan) tapi kalo kamu udh serti itu beda cerita cuy bisa sama bes...
5	@Tunniskao @Bonakm_ @daniellsinaga_ @grok Ga semua daerah ada Tukin kecuali pusat.. Yang di Pemda cuma dapat gapok dan kalau ada sertifikasi...
6	@Bonakm_ @daniellsinaga_ @grok Berarti klo sertifikasi 3 bulan sekali dapat 3jutax3bulan= 9juta kak? Maaf berarti bener kan? P3k guru bersertifikasi...
7	@AriefTegarG @ardisatriawan Mirisnya generalisasi PNS=kaya tuh mereka lihatnya dari para pejabatnya atau instansi yang Tukin nya dashyat tapi tutup...
8	@nadaliciasihna @daniellsinaga_ Nominal Tukin belum serti dpt 1/2 gapok Tukin sudah serti gapok penuh
9	@wawam_bejoe @daniellsinaga_ Itu kalau gak salah ya haha ntr coba crosscheck lg barangkali adanya TPP kali ya. Oia kalau P3K ada Tukin alias T...
10	Pengabdian ngadi ngadi with the same logic kita harusnya ga ush gaji presiden dan antek antek menteri nya tukin mereka sendiri bisa buat biaya 1 ka...

Setelah data cleansing

Data	Row No.	full_text
Statistics	1	Jangan nambah aplikasi lagi yaTukin harusnya berdasarkan BKD sajakalau BKD kan syaratnya jelas 1216 SKS
	2	wkwkwk langsung lupa tukin
	3	Bukan masalah penak Soalnya kalau guru dgn tubel + beasiswa dianggap cuti tukin ga cair lakyo podo ekonomis e ka...
Visualizations	4	Trs kalo kamu blm serti bakal ada tukin yg besarnya setengah dari gapok kamu tunjangan tapi kalo kamu udh serti itu ...
	5	Ga semua daerah ada Tukin kecuali pusat Yang di Pemda cuma dapat gapok dan kalau ada sertifikasi yang dibayar s...
	6	Berarti klo sertifikasi 3 bulan sekali dapat 3jutax3bulan= 9juta kak Maaf berarti bener kan P3k guru bersertifikasi 2 gap...
Annotations	7	Mirisnya generalisasi PNS=kaya tuh mereka lihatnya dari para pejabatnya atau instansi yang Tukin nya dashyat tapi tut...
	8	Nominal Tukin belum serti dpt 1/2 gapok Tukin sudah serti gapok penuh
	9	Itu kalau gak salah yah haha ntr coba crosscheck lg barangkali adanya TPP kali ya Oia kalau P3K ada Tukin alias Tunj...
	10	Pengabdian ngadi ngadi with the same logic kita harusnya ga ush gaji presiden dan antek antek menteri nya tukin mer...

Gambar 1: Sebelum dan Sesudah Cleansing

Dari 607 data *crawling*, setelah dilakukan data *cleansing*, tersisa 296 data bersih yang akan digunakan untuk analisis data sentimen. Namun, sebelum melanjutkan ke tahap berikutnya, sebagian data yang sudah kita bersihkan terlebih dahulu akan dilabeli sentimen secara manual, yang terdiri dari sentimen positif dan sentimen negatif, sisanya akan dilakukan secara otomatis, menggunakan algoritma *machine learning*-naive bayes.

3.2.2. Tokenization

Fungsi dari tokenization adalah memecah teks menjadi kata per kata. Berikut adalah hasil dari proses tokenisasi.

Row No.	A	AA	ACC	ADA	AJA	ANJING	APALAGI	APBDP	APC	APH	ASN
1	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0.243	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0

Gambar 4. Hasil tokenization

Berdasarkan Gambar 4. Terdapat ribuan kata yang dipecah. Untuk menganalisis menjadi kata yang bernilai untuk digunakan dalam analisis sentimen, maka kata-kata yang tidak penting, tidak memiliki arti, dan atau kurang dari empat huruf, maka proses stopwords dapat diterapkan.

3.2.3. Stopword

Stopword digunakan untuk menghilangkan kata-kata yang dianggap tidak memiliki makna penting dalam proses pengolahan data review. Kata-kata ini biasanya bersifat umum, seperti kata hubung, kata depan, atau kata ganti, yang tidak memberikan kontribusi berarti terhadap analisis isi.

Setelah proses *stopword* diterapkan, kata-kata yang terdiri dari kurang dari empat huruf berhasil dihapus dari kumpulan data. Hasil dari proses ini ditunjukkan pada Gambar 5, yang memperlihatkan bahwa kata-kata yang termasuk dalam daftar *stopword* telah berhasil dihilangkan dari tabel frekuensi kata.

Dari hasil *stopword* ini, data dapat divisualisasikan dalam bentuk *wordcloud* untuk melihat kata-kata yang paling sering muncul atau paling dominan dalam data. Kata-kata yang sering disebut akan ditampilkan lebih menonjol pada visualisasi *wordcloud*, sehingga memudahkan dalam mengidentifikasi topik atau tema utama dalam kumpulan data.

jidat	adilant	adilnnya	administrasi	ahli	aizawl	ajuan	akal	akalakalan	akeciki	akhirat	akomodasi
	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0.392	0	0	0

Gambar 5. Hasil *Stopword*

Setelah diterapkan proses Stopword, kata-kata yang kurang dari empat huruf sudah berhasil dihapus. Dari hasil stopwords, dapat divisualisasikan ke dalam bentuk wordcloud. Dimana kata-kata yang penyebutannya paling sering atau banyak disebut dapat ditampilkan pada wordcloud berikut.



Gambar 6. Wordcloud

3.3. Pemrosesan TF-IDF

TF-IDF digunakan untuk mengetahui frekuensi dari istilah tertentu pada sebuah kata dalam dokumen (sekumpulan teks yang terstruktur) [10]. Berikut ini adalah hasil dari TF-IDF pada Gambar 7.

Row No.	text	aahmasa	aamiin	abalabal	abba	abdi
133	saianin nrsnya pemerintan oatar cpn...	0	0	0	0	0
134	halah omon tuKin dosen negatif	0	0	0	0	0
135	yahh mudah kategori mempersulit do...	0	0	0	0	0
136	jujur aneh abdi negara ditaruh diman...	0	0	0	0	0.193
137	iyalah tuKin dibayarkan gapunya kelu...	0	0	0	0	0
138	selamat dosen tuKin dianaktirikan ne...	0	0	0	0	0
139	puji syukur pembayaran berdo mem...	0	0	0	0	0
140	aamiin legowo curiga komponen neg...	0	0.528	0	0	0
141	ditempatkan gaji tuKin pendidikan ker...	0	0	0	0	0
142	ditengah polemik tuKin dosen menun...	0	0	0	0	0
143	heran kemendikbud gada tuKin	0	0	0	0	0
144	sokan bilang beda abdi negara kontr...	0	0	0	0	0.357
145	beban kerja melebihi tupoksi utama ...	0	0	0	0	0
146	gada tuKin dosen	0	0	0	0	0

ExampleSet (295 examples, 1 special attribute, 1,717 regular attributes)

Gambar 7. Hasil TF-IDF

3.4. Klasifikasi

Selanjutnya yaitu proses klasifikasi. Data latih digunakan untuk membangun model klasifikasi, sedangkan data uji digunakan untuk mengukur performa model. Model Naive Bayes dibangun berdasarkan prinsip probabilistik, yaitu menghitung probabilitas setiap kategori (positif atau negatif) terhadap fitur yang dimiliki oleh dokumen (kata-kata yang sudah ditransformasikan menggunakan TF-IDF). Data latih kemudian dimasukkan ke dalam model untuk menghitung peluang prior probabilitas dan peluang bersyarat yang diperlukan oleh naive bayes. Model yang sudah dilatih kemudian digunakan untuk memprediksi label sentimen pada data uji. Setiap data uji diklasifikasikan ke dalam kategori sentimen positif dan sentimen negatif berdasarkan prediksi model.

Dataset pada penelitian ini dibagi menjadi dua bagian, yaitu data latih 80% sebanyak 207 data, dan data uji 20% sebanyak 52 data menggunakan ratio 0.2. Diperoleh hasil akurasi sebesar 93% untuk data latih dan 83,88% hasil akurasi untuk data testing, dimana sentimen negatif berjumlah 245 data (94,57%) dan sentimen positif berjumlah 14 data (5,43%).

3.5. Evaluasi

Pada tahap evaluasi, digunakan metode confusion matrix untuk mengetahui akurasi, presisi, dan recall. Berdasarkan confusion matrix, kategori sentimen negatif, diperoleh nilai presisi sebesar 91,11% dan recall sebesar 91,11%, yang menandakan bahwa model sangat efektif dalam mengidentifikasi data dengan sentimen negatif. Sementara itu, untuk sentimen positif, nilai presisi hanya mencapai 33,33%, dan recall sebesar 28,57%, menunjukkan bahwa model mengalami kesulitan dalam mendeteksi sentimen positif secara akurat. Hal ini disebabkan oleh ketidakseimbangan jumlah data positif yang jauh lebih sedikit dibandingkan data negatif dalam dataset.

4. KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa mayoritas dosen di perguruan tinggi negeri menunjukkan sentimen negatif terhadap kebijakan tunjangan kinerja (tukin) yang diterapkan. Dari 259 data hasil preprocessing, sebesar 94,57% di antaranya mengungkapkan ketidakpuasan, sedangkan hanya 5,43% yang menunjukkan sentimen positif. Temuan ini mengindikasikan bahwa kebijakan tukin saat ini dinilai belum optimal dari sudut pandang dosen, baik dalam aspek keadilan, implementasi, maupun dampaknya terhadap kinerja akademik. Bagi pemerintah, hasil ini menjadi masukan penting untuk mengevaluasi dan memperbaiki kebijakan tukin, agar lebih adil, transparan, dan mampu meningkatkan motivasi dosen. Sementara bagi

dosen itu sendiri, hasil ini dapat menjadi landasan objektif untuk mendorong dialog terbuka dan usulan perbaikan kebijakan berbasis data nyata.

REFERENSI

- [1] Humas Kementerian Pendidikan Tinggi, Sains, "No Title," *kemdiktisaintek.go.id*, 2025. <https://kemdiktisaintek.go.id/kabar-dikti/kabar/mendiktisaintek-jelaskan-kebijakan-tunjangan-kinerja-bagi-pegawai-di-lingkungan-kemdiktisaintek/> (accessed Apr. 25, 2025).
- [2] Stephen P. Robbins, *Organizational Behavior*, 10th ed. Prentice Hall, 2003, 2003.
- [3] Kemenpan RB, *Reformasi Birokrasi dalam Praktik*. 2013.
- [4] B. Liu, "Sentiment Analysis: Mining Opinions, Sentiments, and Emotions, Second Edition," *Sentim. Anal. Min. Opin. Sentim. Emot. Second Ed.*, no. May, pp. 1–432, 2020, doi: 10.1017/9781108639286.
- [5] L. L. Bo Pang, "Opinion mining and sentiment analysis," *Found. Trends Inf. Retr.*, vol. 02, no. 1–2, pp. 1–135, 2008.
- [6] E. Cambria, B. Schuller, and X. and C. H. Y, "New Avenues in Opinion Mining and Sentiment Analysis," *IEEE Intell. Syst.*, vol. 28, no. 2, pp. 15–21, 2013, doi: <https://doi.org/10.1109/MIS.2013.30>.
- [7] E. Lunando and A. Purwarianti, "Indonesian social media sentiment analysis with sarcasm detection," *2013 Int. Conf. Adv. Comput. Sci. Inf. Syst. ICACIS 2013*, pp. 195–198, 2013, doi: 10.1109/ICACIS.2013.6761575.
- [8] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," *Proc. 7th Int. Conf. Lang. Resour. Eval. Lr. 2010*, pp. 1320–1326, 2010, doi: 10.17148/ijarccce.2016.51274.
- [9] R. Socher *et al.*, "Recursive deep models for semantic compositionality over a sentiment treebank," *EMNLP 2013 - 2013 Conf. Empir. Methods Nat. Lang. Process. Proc. Conf.*, pp. 1631–1642, 2013.
- [10] I. P. Rahayu, A. Fauzi, and J. Indra, "Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Naive Bayes Dan Support Vector Machine," *J. Sist. Komput. dan Inform.*, vol. 4, no. 2, p. 296, 2022, doi: 10.30865/json.v4i2.5381.
- [11] A. Tanggu Mara, E. Sedyono, and H. Purnomo, "Penerapan Algoritma K-Nearest Neighbors Pada Analisis Sentimen Metode Pembelajaran Dalam Jaringan (DARING) Di Universitas Kristen Wira Wacana Sumba," *Jointer - J. Informatics Eng.*, vol. 2, no. 01, pp. 24–31, 2021, doi: 10.53682/jointer.v2i01.30.
- [12] D. J. Putnarubun and C. Palembang, "Penerapan Metode *Machine learning* - Naive Bayes Pada Analisis Sentimen Pemindahan Ibu Kota Negara Baru," vol. 02, no. 01, pp. 69–78, 2023.
- [13] F. Nurhuda, S. Widya Sihwi, and A. Doewes, "Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier," *J. Teknol. Inf. ITSmart*, vol. 2, no. 2, p. 35, 2016, doi: 10.20961/its.v2i2.630.
- [14] T. Codingstudio.id, "Pengertian Crawling Data," *Codingstudio*, 2024. <https://codingstudio.id/blog/crawling-data-adalah/>.
- [15] R. N. Ilyas, S. Mulyani, S. Wiguna, and M. Ramli, "PENGUNAAN CRAWLING DATA X DENGAN MENGGUNAKAN TWEET," no. September 2022, pp. 851–856, 2024.
- [16] B. G. Sudarsono, M. I. Leo, A. Santoso, and F. Hendrawan, "Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner," *JBASE - J. Bus. Audit Inf. Syst.*, vol. 4, no. 1, pp. 13–21, 2021, doi: 10.30813/jbase.v4i1.2729.