

**ANALISIS SENTIMEN *TWITTER* MENGGUNAKAN ALGORITMA  
*K-MEANS* MENGENAI DAMPAK PENGHAPUSAN TKA SNBT TAHUN  
2023**

**SKRIPSI**

Diajukan untuk memenuhi sebagian syarat untuk memperoleh gelar  
Sarjana Teknik Program Studi Sistem Telekomunikasi



Oleh:

Riyadh Ahmad Faridz

1909076

**PROGRAM STUDI SISTEM TELEKOMUNIKASI  
KAMPUS UPI DI PURWAKARTA  
UNIVERSITAS PENDIDIKAN INDONESIA  
2024**

**LEMBAR HAK CIPTA**  
**ANALISIS SENTIMEN *TWITTER* MENGGUNAKAN ALGORITMA**  
***K-MEANS* MENGENAI DAMPAK PENGHAPUSAN TKA SNBT TAHUN 2023**

Oleh

**Riyadh Ahmad Faridz**

Sebuah skripsi yang diajukan untuk memenuhi salah satu syarat memperoleh gelar  
Sarjana Teknik pada Program Studi Sistem Telekomunikasi

© **Riyadh Ahmad Faridz 2024**

Universitas Pendidikan Indonesia

Januari 2024

Hak cipta dilindungi undang-undang.

Skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian, dengan dicetak  
ulang, di foto *copy*, atau cara lainnya tanpa izin dari penulis.

**LEMBAR PENGESAHAN  
SKRIPSI**

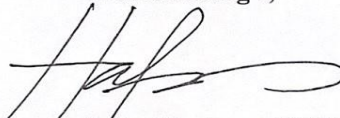
**LEMBAR PENGESAHAN  
DRAFT SKRIPSI**

**RIYADH AHMAD FARIDZ**

**ANALISIS SENTIMEN *TWITTER* MENGGUNAKAN ALGORITMA  
*K-MEANS* MENGENAI DAMPAK PENGHAPUSAN TKA SNBT TAHUN  
2023**

**Disetujui dan Disahkan Oleh Pendamping**

**Pembimbing I,**



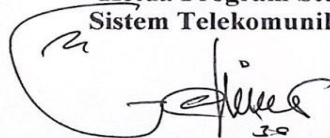
**Hafivyan Patra Pratama, S.ST., M.T**  
NIP. 920190219921224101

**Pembimbing II,**



**Endah Setyowati, S.T., M.T**  
NIP. 920190219920908201

**Mengetahui,  
Ketua Program Studi  
Sistem Telekomunikasi**



**Galura Muhammad Suranegara, S.Pd, M.T**  
NIP. 920190219920111101

## PERNYATAAN ANTI PLAGIARISME

Dengan ini penulis menyatakan bahwa skripsi dengan judul “**Analisis Sentimen *Twitter* Menggunakan Algoritma *K-Means* Mengenai Dampak Penghapusan TKA SNBT Tahun 2023**” ini beserta seluruh isinya adalah benar-benar karya saya sendiri.

Saya tidak melakukan penjiplakan atau pengutipan dengan etika ilmu yang berlaku dalam masyarakat keilmuan. Atas pernyataan ini, saya siap menanggung resiko atau sanksi yang berlaku apabila dikemudian hari ditemukan adanya pelanggaran etika keilmuan atau ada klaim dari pihak lain terhadap keaslian karya saya ini.

Purwakarta, 2 Januari 2024

Penulis,



**Riyadh Ahmad Faridz**

**NIM.1909076**

## KATA PENGANTAR

*Assalammu 'alaikum Warahmatullahi Wabarakatuh*

Puji dan syukur penulis panjatkan kepada kehadirat Allah Subhanahu Wata'ala atas limpahan berkat dan juga karunia-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul “**Analisis Sentimen Twitter Menggunakan Algoritma K-Means Mengenai Dampak Penghapusan TKA SNBT Tahun 2023**” sebagai salah satu syarat kelulusan pada Program studi S1 Sistem Telekomunikasi Universitas Pendidikan Indonesia Kampus Daerah di Purwakarta. Tidak lupa pula Sholawat serta salam kita curahkan kepada baginda Nabi Muhammad Shalallahu 'Alaihi Wasallam pada keluarga, sahabat, serta kita semua sebagai pengikutnya hingga akhir zaman. Penulis ingin menyampaikan terima kasih kepada seluruh pihak yang terlibat dalam proses penyusunan tugas akhir skripsi ini.

Penulis menyadari bahwa dalam mengerjakan skripsi ini masih banyak kesalahan dan kekurangan dalam penyajian maupun dalam penulisan. Oleh karenanya penulis berharap bahwa dalam penulisan skripsi ini diberikan saran dan kritik yang membangun untuk kedepannya dijadikan bahan evaluasi dan pembelajaran bagi penulis kedepannya.

Akhir kata semoga skripsi ini dapat memberikan manfaat untuk ilmu pengetahuan dan bermanfaat bagi yang membacanya.

*Wassalammu 'alaikum Warahmatullahi Wabarakatuh*

Purwakarta, 2 Januari 2024

Penulis,



Riyadh Ahmad Faridz  
NIM.1909076

## UCAPAN TERIMA KASIH

Dalam penyusunan skripsi ini, penulis menyadari bahwa tanpa bantuan, dukungan, dan bimbingan, maka skripsi ini tidak akan terselesaikan. Maka dari itu, dalam kesempatan kali ini penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu dalam proses menyelesaikan skripsi ini.

1. Allah Subhanahu Wa Ta'ala atas limpahan berkah, rahmat dan karunia-Nya, kasih sayang-Nya selama perkuliahan dan penyelesaian skripsi ini. Serta senantiasa memberikan ridho untuk kemudahan dan kelancaran, selalu memberikan kekuatan dalam menghadapi kesulitan dalam pengerjaan skripsi ini. Penulis menjadikannya tempat untuk berdoa dan memohon agar setiap langkah yang dilakukan oleh penulis dapat diridhoi-Nya.
2. Kedua Orang Tua Penulis yaitu, Bapak Ahmad Saepudin dan Ibu Yanti Murdiyanti yang selalu memberi doa dan dukungan, semangat, motivasi, materi, untuk kelancaran masa perkuliahan penulis hingga ditahap pengerjaan skripsi ini.
3. Kepada Kedua Kakak penulis Teh Wulan Mustika dan Mas Muhammad Ricky Kurniawan, yang selalu memberikan masukan kepada penulis dalam hal pengerjaan skripsi, baik berupa dukungan, semangat, dan juga langkah-langkah dalam mengerjakannya.
4. Diri Saya sendiri yang mampu menjalani setiap rintangan dan hambatan selama masa perkuliahan hingga pengerjaan skripsi ini. Saya bangga dengan diri sendiri karena tidak pernah merasa putus asa, selalu memberikan keyakinan untuk bisa menyelesaikan, melawan rasa panik dan cemas dalam pengerjaan skripsi ini dari tahap pencarian topik untuk judul hingga penyusunan skripsi.
5. Bapak Hafiyyan Putra Pratama S.ST., M.T., selaku dosen pembimbing I yang telah menyempatkan waktu, tenaga, dan pikiran untuk memberikan bimbingan, masukan, kritik dan saran, kepada penulis dimulai dari pencarian topik judul hingga penyusunan skripsi ini.
6. Ibu Endah Setyowati S.T., M.T., selaku dosen pembimbing II yang telah menyempatkan waktu, tenaga, dan pikiran untuk memberukan bimbingan,

masukan, kritik, dan saran, kepada penulis dalam hal penulisan agar penyusunan skripsi ini membuahkan hasil yang sebaik-baiknya.

7. Bapak Galura Muhammad Suranegara S.Pd., M.T., selaku Ketua Prodi Sistem Telekomunikasi dan juga selaku dosen wali penulis, yang telah memberikan arahan dan juga informasi-informasi terkait perkuliahan selama penulis duduk di bangku mahasiswa Sistem Telekomunikasi hingga akhir dari perkuliahan penulis di tahap skripsi ini.
8. Bapak Ahmad Fauzi S.Si., M.T. selaku Ketua Prodi Sistem Telekomunikasi periode 2019-2023, yang telah memberikan berbagai masukan terhadap pengerjaan skripsi ini.
9. Kepada sahabat penulis, yang telah banyak memberikan motivasi kepada penulis untuk menyelesaikan penulisan skripsi ini. Selain itu dapat memberikan saran-saran atas keluhan yang penulis sampaikan.
10. Teman-teman penulis, penghuni kos beringin, yang telah memberikan fasilitas kepada penulis untuk menginap dan juga mengerjakan skripsi, selain itu memberikan semangat dan dukungan agar kita semua dapat lulus bersamaan.
11. Teman-teman Mahasiswa Sistem Telekomunikasi Angkatan 2019 yang telah bersama-sama dengan penulis, berjuang bersama dalam keadaan suka maupun duka selama masa perkuliahan di Sistem Telekomunikasi.
12. Seluruh Dosen dan Staff Sistem Telekomunikasi, yang telah memberikan penulis ilmu-ilmu yang belum pernah didapatkan sebelumnya selama masa perkuliahan berlangsung.
13. Kepada seluruh pihak yang tidak dapat disebutkan namanya satu persatu, tanpa mengurangi rasa hormat, penulis haturkan terima kasih yang sebesar-besarnya, hingga penulis dapat menyelesaikan skripsi ini.

## ABSTRAK

### ANALISIS SENTIMEN *TWITTER* MENGGUNAKAN ALGORITMA *K-MEANS* MENGENAI DAMPAK PENGHAPUSAN TKA SNBT TAHUN 2023

Proses analisis sentimen merupakan teknik untuk mengekstrak data berbentuk teks yang diberikan oleh pengguna internet di berbagai sosial media untuk menyampaikan opini dalam setiap permasalahan yang sedang terjadi. Analisis sentimen mengolah data berupa opini, emosi, evaluasi, dan sikap seseorang terhadap orang lain, benda, dan peristiwa. Dalam sebuah studi kasus analisis sentimen penting untuk dilakukan untuk mengetahui opini dari masyarakat sehingga dapat dijadikan bahan evaluasi bagi pihak terkait. Penelitian ini menggunakan metode *clustering* yaitu *K-Means* dibantu dengan kamus pelabelan *vadersentiment* dan *lexicon based*. Tujuan dari penelitian ini untuk mengetahui bagaimana sentimen masyarakat di sosial media *twitter* terkait dampak penghapusan TKA UTBK SNBT tahun 2023, selain itu tujuan dilakukan untuk mengetahui performa algoritma *K-Means* menggunakan bantuan kamus pelabelan *vadersentiment* dan *lexicon based*. Penelitian ini menggunakan 1120 data diambil dari *twitter*. Hasil dari penelitian ini yaitu 72,1% sentimen negatif, dan 27,9% sentimen bernilai positif. Tingkat akurasi tertinggi yang didapat oleh algoritma *K-Means* dengan bantuan kamus pelabelan *lexicon based* sebesar 72%, *precision* 69,2%, dan *recall* 47,3%.

**Kata Kunci:** Analisis Sentimen, *K-Means*, *vadersentiment*, *lexicon based*



**ABSTRACT*****SENTIMENT ANALYSIS REGARDING THE IMPACT OF REMOVING TKA  
SNBT IN 2023 USING K-MEANS***

*Sentiment analysis is a technique employed to extract textual data provided by internet users on various social media platforms, enabling them to convey opinions on ongoing issues. Sentiment analysis processes data encompassing opinions, emotions, evaluations, and attitudes of individuals towards others, objects, and events. In a case study, sentiment analysis is crucial to ascertain public opinions for the purpose of evaluation by relevant parties. This research employs the clustering method known as K-Means, aided by the vadersentiment labeling dictionary and lexicon-based approach. The objective of this study is to uncover the sentiment of the Twitter community regarding the impact of the removal of TKA UTBK SNBT in 2023. Additionally, it seeks to assess the performance of the K-Means algorithm utilizing the vadersentiment labeling dictionary and lexicon-based approach. The study is based on 1120 data points collected from Twitter. The findings indicate that 72.1% of sentiments are negative, while 27.9% are positive. The highest accuracy rate achieved by the K-Means algorithm with the assistance of the lexicon-based labeling dictionary is 72%, with a precision of 69.2% and a recall of 47.3%*

**Keywords:** *Sentiment Analysis, K-Means, vadersentiment, lexicon based*

## DAFTAR ISI

LEMBAR HAK CIPTA .....	i
LEMBAR PENGESAHAN SKRIPSI .....	ii
PERNYATAAN ANTI PLAGIARISME .....	iii
KATA PENGANTAR .....	iv
UCAPAN TERIMA KASIH .....	v
ABSTRAK .....	vii
<i>ABSTRACT</i> .....	viii
DAFTAR ISI .....	ix
DAFTAR TABEL .....	xii
DAFTAR GAMBAR .....	xiv
DAFTAR LAMPIRAN .....	xv
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	2
1.3 Batasan Masalah .....	3
1.4 Tujuan Penelitian .....	3
1.5 Manfaat Penelitian .....	3
1.5.1 Manfaat Teoritis .....	3
1.5.2 Manfaat Praktis .....	4
1.6 Struktur Organisasi Penulisan .....	4
BAB II KAJIAN PUSTAKA .....	5
2.1 Penelitian Relevan .....	6
2.2 Analisis Sentimen .....	8
2.3 <i>Twitter</i> .....	9
2.4 <i>Classification</i> .....	10
2.5 <i>Clustering</i> .....	10
2.6 TF-IDF ( <i>Term Frequency-Inverse Document Frequency</i> ) .....	10
2.7 <i>Wordcloud</i> .....	11
2.8 <i>K-Means</i> .....	11
2.9 <i>Lexicon Based</i> .....	15
2.10 <i>VaderSentiment</i> .....	15
2.11 NLP ( <i>Natural Language Processing</i> ) .....	15

2.12	<i>Text Mining</i> .....	16
2.13	<i>Machine Learning</i> .....	16
2.14	UTBK (Ujian Tulis Berbasis Komputer) .....	17
2.15	SNBT (Seleksi Nasional Berdasarkan Tes).....	17
2.16	<i>Python</i> .....	18
2.17	<i>Google Colaboratory</i> .....	19
2.18	<i>Confusion Matrix</i> .....	19
2.19	WEKA ( <i>Waikato Environment for Knowledge Analysis</i> ) .....	19
BAB III METODE PENELITIAN .....		21
3.1	Jenis Penelitian.....	21
3.2	Prosedur Penelitian.....	21
3.2.1	Alur Pengolahan Data.....	21
3.2.2	Alur Evaluasi Data.....	23
3.3	Data Penelitian .....	25
3.4	Analisis Data.....	26
3.4.1	<i>Accuracy</i> (Akurasi).....	27
3.4.2	<i>Precision</i> (Presisi).....	27
3.4.3	<i>Recall</i> .....	27
3.5	Alat Penelitian.....	28
3.5.1	Kebutuhan Perangkat Lunak ( <i>Software</i> ) .....	28
3.5.2	Kebutuhan Perangkat Keras ( <i>Hardware</i> ).....	28
BAB IV HASIL DAN PEMBAHASAN .....		29
4.1	<i>Crawling Data</i> (Pengambilan Data) .....	29
4.2	Pengujian Data .....	29
4.2.1	<i>Data Cleaning</i> .....	29
4.2.2	<i>Preprocessing</i> .....	31
4.2.3	Visualisasi Kata menggunakan <i>wordcloud</i> . .....	35
4.2.4	Pelabelan Data Latih menggunakan <i>Lexicon Based</i> .....	37
4.2.5	Pembobotan TF-IDF (Term Frequency-Inverse Document Frequency) .....	37
4.2.6	Klasterisasi menggunakan algoritma <i>K-Means</i> .....	38
4.2.7	Pengujian algoritma menggunakan kamus pelabelan.....	49
4.2.8	Perhitungan <i>Accuracy, Precision, Recall</i> pada hasil pelabelan .....	49
BAB V KESIMPULAN .....		56
5.1	Kesimpulan .....	56
5.2	Implikasi Penelitian.....	56

5.3	Saran.....	57
	REFERENSI.....	58
	LAMPIRAN .....	63

## DAFTAR TABEL

Tabel 2.1 Perbandingan objek penelitian dan <i>tools</i> .....	8
Tabel 2.2 Perbandingan penggunaan algoritma <i>K-Means</i> .....	14
Tabel 4.1 Perbandingan proses penghapusan URL .....	30
Tabel 4.2 Perbandingan proses penghapusan tanda baca .....	30
Tabel 4.3 Perbandingan proses penghapusan <i>Username</i> .....	31
Tabel 4.4 Proses perbandingan <i>Transform Case</i> .....	32
Tabel 4.5 Perbandingan proses <i>Stopwords</i> .....	34
Tabel 4.6 Perbandingan proses <i>Stemming</i> .....	35
Tabel 4.7 Kata yang paling sering muncul menggunakan fitur <i>wordcloud</i> .....	36
Tabel 4.8 Hasil TF-IDF pada sample 25 kata .....	38
Tabel 4.9 Data uji untuk proses <i>clustering</i> .....	39
Tabel 4.10 Penentuan <i>centroid</i> awal.....	41
Tabel 4.11 Perhitungan <i>Euclidean Distance cluster 1</i> .....	42
Tabel 4.12 Perhitungan <i>Euclidean Distance cluster 2</i> .....	42
Tabel 4.13 Pembagian objek terhadap setiap <i>cluster</i> .....	43
Tabel 4.14 Hasil <i>centroid</i> baru .....	44
Tabel 4.15 Perhitungan <i>Euclidean Distance</i> Kedua <i>cluster 1</i> .....	44
Tabel 4.16 Perhitungan <i>Euclidean Distance</i> Kedua <i>cluster 2</i> .....	44
Tabel 4.17 Pembagian objek terhadap setiap <i>cluster</i> .....	45
Tabel 4.18 Hasil <i>centroid</i> baru .....	46
Tabel 4.19 Perhitungan <i>Euclidean Distance</i> Ketiga <i>cluster 1</i> .....	46
Tabel 4.20 Perhitungan <i>Euclidean Distance</i> Ketiga <i>cluster 2</i> .....	46
Tabel 4.21 Pembagian objek terhadap setiap <i>cluster</i> .....	47
Tabel 4.22 Hasil akhir pembagian objek menggunakan <i>Euclidean Distance</i> .....	48
Tabel 4.23 Pengujian menggunakan 50 data uji <i>K-Means</i> dengan <i>vadersentiment</i> .....	50
Tabel 4.24 Pengujian menggunakan 100 data uji <i>K-Means</i> dengan <i>vadersentiment</i> .....	50
Tabel 4.25 Pengujian menggunakan 150 data <i>K-Means</i> dengan <i>vadersentiment</i> ..	51
Tabel 4.26 Pengujian menggunakan 50 data <i>K-Means</i> dengan <i>Lexicon Based</i> ....	52
Tabel 4.27 Pengujian menggunakan 100 data <i>K-Means</i> dengan <i>Lexicon Based</i> ..	53

Tabel 4.28 Pengujian menggunakan 150 data <i>K-Means</i> dengan <i>Lexicon Based</i> ..	54
Tabel 4.29 Hasil Pengujian perfomansi algoritma <i>K-Means</i> dengan dua kamus pelabelan .....	55

**DAFTAR GAMBAR**

Gambar 3.1 Alur Pengolahan Data.....	22
Gambar 3.2 Alur Evaluasi Data.....	24
Gambar 4.1 <i>Flowchart Data Cleaning</i> .....	30
Gambar 4.2 <i>Flowchart Preprocessing</i> .....	31
Gambar 4.3 Proses <i>Transform Case</i> .....	32
Gambar 4.4 <i>Import library</i> proses <i>Tokenize</i> .....	33
Gambar 4.5 <i>Coding</i> proses <i>Tokenize</i> .....	33
Gambar 4.6 Hasil dari proses <i>Tokenize</i> .....	33
Gambar 4.7 <i>Import library stopwords</i> .....	34
Gambar 4.8 <i>Import library Stemming</i> .....	34
Gambar 4.9 Hasil visualisasi kata menggunakan fitur <i>wordcloud</i> .....	35
Gambar 4.10 Persentase pelabelan menggunakan kamus <i>Lexicon Based</i> .....	37
Gambar 4.11 Hasil Pembagian pada dua <i>cluster</i> menggunakan aplikasi WEKA .	48
Gambar 4.12 Hasil dari perhitungan iterasi menggunakan aplilasi WEKA .....	48

**DAFTAR LAMPIRAN**

Lampiran 1- <i>Coding Crawling</i> Data .....	63
Lampiran 2 - <i>Coding Cleaning</i> Data dan <i>Stemming</i> .....	64
Lampiran 3 - <i>Coding Visualisasi Wordcloud</i> .....	65
Lampiran 4 - <i>Coding Lexicon Based</i> .....	66
Lampiran 5 - <i>Coding TF-IDF</i> .....	67
Lampiran 6 - <i>Coding K-Means</i> .....	69
Lampiran 7 - <i>Coding Vadersentiment</i> .....	72
Lampiran 8 - Hasil <i>Crawling</i> Data .....	74
Lampiran 9 - Tabel Hasil Pelabelan .....	75
Lampiran 10 - Riwayat Hidup Penulis .....	80



## REFERENSI

- Afrilia, C., & Siahaan, C. (2021). Penggunaan Komunikasi Media Sosial Twitter di Kalangan Remaja di Kecamatan Cibinong, Kabupaten Bogor. *JISIP: Jurnal Ilmu Sosial Dan Ilmu Politik*, 10.
- Ali Imron. (2019). Analisis sentimen terhadap tempat wisata di kabupaten rembang menggunakan metode naive bayes classifier. *Program studi teknik informatika – program sarjana fakultas teknologi industri universitas islam indonesia*.
- Amin, F., Sri Anggraeni, D., & Aini, Q. (2022). Penerapan Metode K-Means dalam Penjualan Produk Souq.Com. *Applied Information System and Management (AISM)*, 5, 7–14.
- Aziz Muslim, M., Prasetyo, B., Laily Harum Mawarni, E., Juli Herowati, A., Mirqotussa'adah, Hardiyanti Rukmana, S., & Nurzahputra, A. (2019). *DATA MINING ALGORITMA C4.5*. Universitas Negeri Semarang.
- Beri, A. (2020). *SENTIMENTAL ANALYSIS USING VADER interpretation and classification of emotions. Towards Data Science*.  
<https://towardsdatascience.com/sentimental-analysis-using-vader-a3415fef7664>
- D Manning, C., Raghavan, P., & Schutze, H. (2009). *An Introduction to Information Retrieval*. Cambridge University Press Cambridge, England.
- Devlin, J., Wei Chang, M., Lee, K., & Toutanova, K. (2019). BERT: Pre-training

of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, 4171–4186.

Fadli, M. (2021). Analisis sentimen komentar masyarakat jakarta terhadap pembatasan sosial berskala besar (psbb) menggunakan algoritma k-means dan algoritma levensthein distance. *Program studi teknik informatika fakultas sains dan teknologi universitas islam negeri syarif hidayatullah*.

Fitri, E., Yuliani, Y., Rosida, S., & Gata, W. (2020). Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine. *TRANSFORMTIKA*, 18. <https://doi.org/journals.usm.ac.id/index.php/transformatika>

Fridom Mailo, F., & Lazuardi, L. (2019). Analisis Sentimen Data Twitter Menggunakan Metode Text Mining Tentang Masalah Obesitas di Indonesia. *Ournal of Information Systems for Public Health*, 4.

Galih Pradana, M. (2020). Penggunaan fitur *wordcloud* dan *document term matrix* dalam *text mining*. *Jurnal Ilmiah Informatika (JIF)*.

Jo Dixon, S. (2022). *Number of monthly active Twitter users worldwide from 1st quarter 2010 to 1st quarter 2019*. Statista. <https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/>

Josen a. Limbong, j., sembiring, i., & dwi hartomo, k. (2022). Analisis klasifikasi sentimen ulasan pada e-commerce shopee berbasis word cloud dengan metode naive bayes dan k-nearest neighbor. *Jurnal teknologi informasi dan ilmu komputer (jtiik)*, 9.

Karsito, & susanti, s. (2019). Klasifikasi kelayakan peserta pengajuan kredit rumah dengan algoritma naïve bayes di perumahan azzura residencia. *Program studi teknik informatika sekolah tinggi teknologi pelita bangsa*.

Mahendrajaya, r., asrofi buntoro, g., & bhanu setyawan, m. (2019). Analisis sentimen pengguna gopay menggunakan metode lexicon based dan support vector machine. *Komputek*.

Moniaga sipahutar, c., fitriana pooerana, a., & nurkinan. (2020). Pengalaman komunikasi curhat anonim bagi followers @18autobase di twitter. *Jurnal lensa mutiara komunikasi program studi ilmu komunikasi universitas singaperbangsa karawang*.

Muliono, r., & sembiring, z. (2019). Data mining clustering menggunakan algoritma k-means untuk klasterisasi tingkat tridarma pengajaran doseN. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 4(2).

Qorib, F. (2020). Pola Konsumsi Media pada Generasi Milenial Kota Malang. *Jurnal Ilmu Komunikasi ULTIMACOMM*, 12.

Rifaih, P. (2023). *Jalan Masuk Perguruan Tinggi Negeri Melalui SNBT*. EduCampus. <https://edufund.co.id/blog/snbt-adalah/>

- Riyanis Putra, M. P., & Nova Wardani, K. R. (2020). Penerapan text mining dalam menganalisis kepribadian pengguna media sosial. *JUTIM (Jurnal Teknik Informatika Musirawas)*.
- Safitri, S. I., Suhery, C., & Bahri, S. (2021). Implementasi algoritma k – means untuk clustering sentimen pada opini kualitas pelayanan jasa penerbangan. *Jurnal Komputer dan Aplikasi*, 09(02), 12.
- Salsabila, N. (2019). Klasifikasi barang menggunakan metode clustering k-means dalam penentuan prediksi stok barang (studi kasus : ukm mar'ah jilbab kediri). *Jurusan teknik informatika fakultas sains dan teknologi universitas islam negeri maulana malik ibrahim malang*.
- Santoso, G. T. (2021). Analisis Sentimen Pada Tweet Dengan Tagar #BPJSRASARENTENIR Menggunakan Metode Support Vector Machine. *Program Studi Teknik Informatika Fakultas Teknik Universitas Islam Riau Pekanbaru*.
- Sefty, W. (2022). *Aturan Baru Seleksi Nasional Masuk Perguruan Tinggi Negeri 2023*. <https://Diskominfo.Kaltimprov.Go.Id/>. <https://diskominfo.kaltimprov.go.id/pendidikan/aturan-baru-seleksi-nasional-masuk-perguruan-tinggi-negeri-2023>
- Soen, G. I. E. (2022). Implementasi Cloud Computing dengan Google Colaboratory Pada Aplikasi Pengolah Data Zoom Participants. *JITU : Journal Informatic Technology And Communication*, 6(1).

Sugiyono, Dr. (2013). *Metode Penelitian Kuantitatif, Kualitatif, Dan RND* (1st ed.). Alfabeta.

Sulaiman, H., & Robia Khaerudin, I. (2021). Simulasi ujian try out utbk-sbmptn online melalui platform web based learning bagi siswa-siswi kelas xii di sma n kota cirebon. *Jurnal pintar abdimaS*, 56–64.

Syuhada, A. S., M Simanullang, A., Sama Lewa, D., & Jeffry Marthin, S. (2021). Makalah pembelajaran mesin ( machine learning ). *Fakultas teknik, Universitas Maritim Raja Ali Haji*.

Teguh Santoso, J. (2016). *Proyek Coding dengan Python*. Yayasan Prima Agus Teknik, Universitas Sains & Teknologi Komputer (Universitas STEKOM).

Wulandari, T. (2022). *Info SBMPTN 2023: Berubah Jadi SNBT 2023, Ini Perbedaan dan Poin Pentingnya*. <https://www.detik.com/edu/seleksi-masuk-pt/d-6487803/info-sbmptn-2023-berubah-jadi-snbt-2023-ini-perbedaan-dan-poin-pentingnya>

Yunita, R. (2019). Aktivitas Pengungkapan Diri Remaja Putri Melalui Sosial Media Twitter. *Universitas Bina Sarana Informatika*, 10.