

SKRIPSI

**ANALISIS SENTIMEN ATAS UJARAN
KEBENCIAN PADA MEDIA SOSIAL TWITTER
MENGUNAKAN METODE NAIVE BAYES
DENGAN EVALUASI CONFUSION MATRIX**



POLITEKNIK NEGERI BALI

Oleh:

Ngurah Komang Wira Galih Adhistanaya
2015354057

**PROGRAM STUDI SARJANA TERAPAN
TEKNOLOGI REKAYASA PERANGKAT LUNAK
JURUSAN TEKNIK ELEKTRO
POLITEKNIK NEGERI BALI
2024**

LEMBAR PERSETUJUAN UJIAN SKRIPSI**LEMBAR PERSETUJUAN UJIAN SKRIPSI****ANALISIS SENTIMEN ATAS UJARAN
KEBENCIAN PADA MEDIA SOSIAL TWITTER
MENGUNAKAN METODE NAIVE BAYES
DENGAN EVALUASI CONFUSION MATRIX**

Oleh :

Ngurah Komang Wira Galih Adhistanaya
NIM. 2015354057

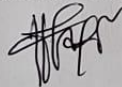
Skripsi ini telah melalui Bimbingan dan Pengujian Hasil, disetujui untuk
diujikan pada Ujian Skripsi
di

Program Sarjana Studi Terapan
Teknologi Rekayasa Perangkat Lunak
Jurusan Teknologi Informasi – Politeknik Negeri Bali

Bukit Jimbaran, 26 Agustus 2024

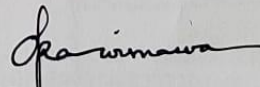
Disetujui Oleh :

Dosen Pembimbing 1:



Putu Indah Ciptayani, S.Kom.,M.Cs.
NIP. 198504132014042001

Dosen Pembimbing 2:



I Putu Oka Wisnawa, S.Kom.,M.T.
NIP. 199011082022031002

LEMBAR PENGESAHAN SKRIPSI

LEMBAR PENGESAHAN SKRIPSI

ANALISIS SENTIMEN ATAS UJARAN KEBENCIAN PADA MEDIA SOSIAL TWITTER MENGUNAKAN METODE NAIVE BAYES DENGAN EVALUASI CONFUSION MATRIX

Oleh :

Ngurah Komang Wira Galih Adhistanaya
NIM. 2015354057

Skripsi ini sudah melalui Ujian Skripsi pada tanggal 23 Agustus 2024,
dan sudah dilakukan Perbaikan untuk kemudian disahkan sebagai Skripsi

di

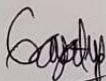
Program Sarjana Studi Terapan
Teknologi Rekayasa Perangkat Lunak
Jurusan Teknologi Informasi – Politeknik Negeri Bali

Bukit Jimbaran, 26 Agustus 2024

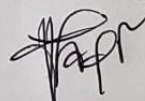
Disetujui Oleh :

Tim Penguji :

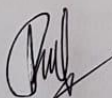
Dosen Pembimbing :



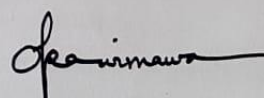
1. Ni Ketut Pradani Gayatri Sarja, S.Kom.,M.Kom.
NIP. 0015069601



1. Putu Indah Ciptayani, S.Kom.,M.Cs.
NIP. 198504132014042001



2. I Putu Astya Prayudha, S.TI.,M.T.
NIP. 199501052023211012



2. I Putu Oka Wisnawa, S.Kom.,M.T.
NIP. 199011082022031002

Diketahui Oleh:

Ketua Jurusan Teknologi Informasi



Prof. Dr. I Nyoman Gede Astya Astawa, ST., M.Kom.)
NIP. 196902121995121001

HALAMAN PERNYATAAN KEASLIAN KARYA SKRIPSI

HALAMAN PERNYATAAN KEASLIAN KARYA SKRIPSI

Saya yang bertanda tangan di bawah ini, menyatakan bahwa Skripsi dengan judul:

“ANALISIS SENTIMEN ATAS UJARAN KEBENCIAN PADA MEDIA SOSIAL TWITTER MENGGUNAKAN METODE NAIVE BAYES DENGAN EVALUASI CONFUSION MATRIX”

adalah asli hasil karya saya sendiri.

Dengan ini saya menyatakan bahwa dalam naskah Skripsi ini tidak terdapat karya orang lain yang pernah diajukan untuk memperoleh gelar di suatu perguruan tinggi, dan atau sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah Skripsi ini, dan disebutkan dalam daftar pustaka.

Apabila saya melakukan hal tersebut di atas, dengan ini saya menyatakan menarik Skripsi yang saya ajukan sebagai hasil karya saya.

Bukit Jimbaran, 29 Agustus 2024

Yang menyatakan



Ngurah Komang Wira Galih Adhistanaya
NIM 2015354057

ABSTRAK

Platform media sosial seperti Twitter, memfasilitasi pengguna dalam bertukar pikiran tentang berbagai topik atau peristiwa secara global. Meskipun Twitter berguna untuk komunikasi dan bebas berpendapat, nyatanya platform ini juga dapat dimanfaatkan untuk penyebaran ujaran kebencian yang memicu permusuhan antar pengguna. Penelitian ini bertujuan untuk melakukan analisis sentimen pada *tweet* guna mengklasifikasikannya sebagai positif atau netral. Analisis ini dilakukan menggunakan metode *naive bayes* yang dikenal karena kesederhanaannya, kecepatan pemrosesannya dan akurasi dalam menanganai dataset yang besar. Namun, karena adanya kebijakan baru Twitter mengenai pembatasan pada pengambilan data sehingga dataset yang digunakan dalam penelitian ini relatif kecil, yang memungkinkan dapat mempengaruhi pada hasil tingkat akurasi dari metode *naive bayes* tersebut. Dataset yang diperoleh menggunakan teknik *crawling* menggunakan pustaka bantuan yaitu *tweet-harvest* dan dibuat dengan pemrograman bahasa Python. Tahapan pengolahan data melibatkan *preprocessing* (mulai dari *cleaning*, *casefolding*, normalisasi, *stopword removal*, *tokenization*, dan *stemming*), *translate*, *labelling*, dan pembobotan pada kata menggunakan metode TF (*Term Frequency*) dan IDF (*Inverse Document Frequency*). Data kemudian dibagi menjadi tiga skenario yaitu 80% data *training* dan 20% data *testing*, 20% data *training* dan 80% data *testing*, serta 50% data *training* dan 50% data *testing*. Evaluasi yang dilakukan menggunakan *confusion matrix*. Hasil metode *naive bayes* memperoleh akurasi sebesar 73% pada skenario 80% data *training* dan 20% data *testing*.

Kata Kunci: analisis sentimen, confusion matrix, naive bayes, Twitter

ABSTRACT

Social media platforms like Twitter are used by people across the world to exchange ideas about different topics or events. Although Twitter can provide communication and impunity for people, it is also open to abuse in the form of hate speech that provokes hostility. This study aims to analyse the sentimental value of tweets with the goal to classify them as either positive or negative. For processing large datasets simply and quickly without sacrificing accuracy, Naive Bayes was employed because of its simplicity, speed and ability to handle huge amounts of data. However, this study's dataset is smallish due to the restrictions placed by Twitter on access to user data; hence these could potentially undermine Naive Bayes' precision levels. The dataset was gotten through crawling using tweet-harvest library which was programmed using Python language for creation purposes. Different processes were carried out including pre-processing such as cleaning up the data (including but not limited to case folding), normalization, removing stop words, tokenization and stemming; translation; labelling; word weighting which happens via TF (Term Frequency) IDF (Inverse Document Frequency). The data was then split into three scenarios: 80% training data versus 20% test data, 20% training data versus 80% test data, and finally 50%-50% training-test split ratios. The evaluation used in this research is confusion matrix. The result indicates highest accuracy at 73% in scenario data split of 80% data training and 20% data testing.

Keywords: *sentiment analysis, confusion matrix, naïve bayes, Twitter*

KATA PENGANTAR

Puji syukur penulis panjatkan kehadapan Ida Sang Hyang Widhi Wasa atau Tuhan Yang Maha Esa karena berkat dan rahmat-Nya sehingga penulis dapat menyelesaikan Laporan Praktik Kerja tepat pada waktunya. Pembuatan karya skripsi ini merupakan mata kuliah yang wajib diambil mahasiswa pada program studi Sarjana Terapan Teknologi Rekayasa Perangkat Lunak, Jurusan Teknologi Informasi, Politeknik Negeri Bali pada semester 8 sebagai syarat untuk kelulusan. Dalam pelaksanaan penyusunan skripsi ini, penulis mendapat bimbingan serta arahan dari banyak pihak dalam mengatasi masalah. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan terima kasih kepada :

1. Bapak I Nyoman Abdi, SE., M.eCom selaku Direktur Politeknik Negeri Bali.
2. Bapak Prof.Dr.I Nyoman Gede Arya Astawa,ST.,M.Kom. selaku Ketua Jurusan Teknologi Informasi Politeknik Negeri Bali.
3. Ibu Ni Gusti Ayu Putu Harry Saptarini,S.Kom.,M.Cs. selaku Ketua Program Studi Teknologi Rekayasa Perangkat Lunak Politeknik Negeri Bali
4. Ibu Putu Indah Ciptayani,S.Kom.,M.Cs. selaku pembimbing pertama yang telah membimbing dalam proses penyelesaian laporan ini.
5. Bapak I Putu Oka Wisnawa,S.Kom.,M.T. selaku pembimbing kedua yang telah membimbing dalam proses penyelesaian laporan ini.
6. Orang tua, saudara, teman-teman serta seluruh pihak yang tidak dapat disebutkan satu persatu yang telah memberikan dukungan baik secara moral maupun material dalam penyelesaian laporan ini.

Akhir kata penulis berharap Tuhan Yang Maha Esa membalas kebaikan semua pihak diatas. Dengan keterbatasan kemampuan yang penulis miliki, penulis menyadari bahwa pembuatan karya skripsi ini masih jauh dari kata sempurna. Oleh karena itu, kritik dan saran yang bersifat membangun sangat diharapkan demi kesempurnaan laporan ini. Semoga laporan ini dapat bermanfaat bagi para pembacanya.

Tabanan, 20 April 2024

Ngurah Komang Wira Galih Adhistanaya

BAB I

PENDAHULUAN

1.1 Latar Belakang

Media sosial merupakan jaringan media yang memungkinkan pengguna dapat bertukar pendapat mengenai suatu hal atau peristiwa tertentu. Media sosial juga menyediakan sarana bagi pengguna untuk berkomunikasi. Dengan media sosial, pengguna dapat menemukan berbagai topik dan tren yang sedang terjadi di berbagai belahan dunia [1].

Twitter sebagai salah satu jenis media sosial berbasis *microblogging* memungkinkan pengguna untuk mengutarakan opini dan pikiran mereka secara terbuka pada topik dan isu-isu yang sedang terjadi [2]. Twitter telah dimanfaatkan untuk keperluan dalam segala aspek. Misal sebagai sarana untuk menyampaikan pendapat, aspirasi, trivia, hiburan ataupun komunikasi yang sifatnya ringan [3]. Aturan yang membatasi jumlah karakter dalam sebuah *tweet* membuat twitter menjadi media sosial yang menarik tetapi juga *challenging*. Pengguna twitter harus memiliki kemampuan literasi yang cukup untuk dapat menuangkan gagasan atau pendapatnya dalam keterbatasan yang telah ditetapkan.

Tetapi disisi lain, twitter ternyata dimanfaatkan sebagai sarana untuk menyebarkan ujaran kebencian oleh tindakan individu atau kelompok yang tidak bertanggungjawab, karena ujaran tersebut memiliki potensi yang sangat merusak karena dampaknya bisa menghambat komunikasi dan toleransi sehingga ketika ada kebencian dan permusuhan, individu akan memiliki kecenderungan untuk menutup diri dari pandangan orang lain dan pada akhirnya lebih memilih pandangan mereka sendiri [4].

Ujaran kebencian itu sendiri terbagi menjadi dua kata yaitu ujaran yang artinya suatu ekspresi atau segala ungkapan yang dapat disampaikan kepada lawan bicara, sedangkan kebencian merupakan suatu emosi yang dapat menciptakan kondisi yang intens dan irasional, jadi ujaran kebencian dapat diartikan sebagai bentuk komunikasi yang dapat mempengaruhi lingkungan sosial atau bentuk komunikasi yang dapat menyerang dalam berbagai aspek. Dengan penelitian ini akan dibahas melalui tahapan yang dilalui untuk melakukan proses analisis sentimen.

Analisis sentimen adalah proses untuk menganalisa dan mengklasifikasi teks atau tweet berdasarkan pandangan subjektif terhadap ujaran tersebut, apakah tweet tersebut tergolong positif atau negatif [4]. Dalam konteks media, analisis sentimen sering digunakan untuk menganalisa data mengenai ulasan, komentar, dan sumber teks lainnya.

Faktanya analisis sentimen ini banyak dikaitkan kedalam media sosial dan dipakai para penelitian untuk mengklasifikasi sentimen yang terkandung dalam kalimat di media sosial. Oleh karena itu, dalam melakukan analisis sentimen, diperlukan suatu algoritma yang mumpuni agar mendapatkan hasil akurasi terbaik. Salah satu metode yang dipakai dalam pengklasifikasian data dan paling sering diterapkan dalam analisis sentimen yaitu Naïve Bayes. Pada dasarnya, Naïve Bayes merupakan salah satu pembelajaran mesin algoritma yang basisnya menggunakan perhitungan probabilitas [5]. Kelebihan dari Naïve Bayes yang akan dipakai yaitu kecepatan dalam pemrosesan dan akurasi yang cukup tinggi bila digunakan pada jumlah data yang banyak [6]. Naïve Bayes adalah salah satu metode yang paling sering diterapkan dalam analisis sentimen karena penggunaannya yang mudah diimplementasikan dengan struktur yang cukup sederhana [7]. Selain itu, penggunaan model algoritma *Naïve Bayes* ini merupakan salah satu metode klasifikasi yang sederhana namun memiliki kemampuan dan keakuratan yang tinggi.

Namun dalam tahap metodologi pada penelitian ini, jumlah data yang ada dalam dataset tidak begitu banyak mengingat aturan kebijakan baru di Twitter dalam hal untuk mencegah penggalan data secara extreme, sehingga pengambilan data dilakukan secara terbatas dan hasil tingkat akurasi tinggi yang didapat belum bisa dipastikan. Pengumpulan data dilakukan dengan teknik *crawling* untuk mengambil dan mengunduh data [8]. Teknik *crawling* yang dilakukan dengan membuat program terlebih dahulu dan kemudian memasukkan kata kunci sesuai kebutuhan peneliti. Berdasarkan uraian diatas peneliti mengusulkan penelitian dengan analisis sentimen menggunakan metode Naïve Bayes untuk klasifikasi atas ujaran kebencian dan sara di kolom tweet supaya bisa menghasilkan hasil keakuratan terbaik serta pembuatan sistem yang dirancang untuk mengklasifikasi sentimen berdasarkan inputan data.

1.2 Perumusan Masalah

Berdasarkan latar belakang permasalahan yang diangkat, maka didapatkan rumusan permasalahan sebagai berikut.

- a. Bagaimana mengumpulkan data dari Twitter?
- b. Varian metode naïve bayes apa yang cocok untuk menghasilkan akurasi tertinggi?

1.3 Batasan Masalah

Ada beberapa batasan yang diimplementasikan dalam penelitian ini, yang tujuannya adalah untuk membatasi fokus penelitian sehingga dapat menghasilkan luaran yang sesuai dengan tujuan penelitian. Adapun Batasan-batasan yang digunakan dalam penelitian ini adalah sebagai berikut.

- a. Penelitian ini hanya berfokus pada implementasi klasifikasi sentimen menggunakan metode *naïve bayes* sehingga tampilan sistem tidak begitu kompleks.
- b. Teknik Crawling yang digunakan belum tentu mendapatkan hasil data yang sempurna dikarenakan adanya susunan kalimat yang tidak teratur dan juga penggunaan bahasa yang kekinian seperti kata yang disingkat, penggunaan kata kekinian serta adanya simbol emoji yang diekstrasi oleh sistem.
- c. Pengambilan data tweet dibatasi dikarenakan kebijakan aturan baru di twitter dalam hal pencegahan tingkat penggalian data dan manipulasi data secara extreme.

1.4 Tujuan

Adapun tujuan penelitian ini adalah :

- a. Mengumpulkan data dari komentar Twitter
- b. Menentukan tipe metode *naïve bayes* yang paling cocok untuk menghasilkan akurasi tertinggi

1.5 Manfaat

- a. Memperdalam pengetahuan di bidang studi mahasiswa
- b. Meningkatkan kepercayaan diri mahasiswa dan mengembangkan keterampilan penelitian yang dapat diterapkan
- c. Membantu memecahkan masalah yang ada di dalam bidang studi mahasiswa

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Metode pembelajaran algoritma *naive bayes* adalah salah satu metode yang mudah diterapkan karena kesederhanaannya walaupun tidak cukup menangani pada data dalam jumlah sangat banyak. Kegunaan metode *naive bayes* dapat diterapkan dalam berbagai tujuan dalam bidang klasifikasi dan analisa data termasuk analisis sentimen. Analisis sentimen menentukan sifat terhadap ujaran kebencian apakah sifatnya netral atau positif.

Kajian penelitian analisis sentimen menggunakan metode *naive bayes* dengan evaluasi *confusion matrix* yang dilakukan pada dataset *tweet* dari hasil perolehan data *crawling* Twitter dengan bantuan Pustaka *tweet-harvest*, terdiri dari 510 sampel dengan melakukan langkah langkah preprocessing seperti *cleaning*, *casefolding*, normalisasi, *stopword removal*, *tokenization* dan *stemming*. Kemudian diberi label dan pembobotan menggunakan TF-IDF dari bantuan pustaka *scikit-learn*. Dan Terakhir menggunakan klasifikasi *multinomial naive bayes* untuk membagi data sampel menjadi tiga skenario data *training* dan data *testing* sebagai berikut : 80 % data *training* dan 20% data *testing*, 20% data *training* dan 80% data *testing*, serta 50% data *training* dan 50% data *testing*. Secara keseluruhan, pengujian skenario data dengan *multinomial naive bayes* yang paling tinggi didapatkan pada 80% data *training* dan 20% data *testing* yang menghasilkan hasil akurasi yang cukup baik yaitu 73% akurasi, 74% presisi, 84% *recall* dan 78% *f1-score*.

5.2 Saran

Saran yang diberikan terkait analisis sentimen pada media sosial twitter menggunakan metode *naive bayes* dengan evaluasi *confusion matrix* yaitu :

1. Saran untuk dosen
 - a. Analisis sentimen pada media sosial twitter menggunakan metode *naive bayes* dengan evaluasi *confusion matrix* disarankan digunakan oleh dosen sebagai bahan ajar dalam suatu bidang studi yang ditempuh untuk meningkatkan pengetahuan dan keterampilan belajar mahasiswa.
 - b. Analisis sentimen pada media sosial twitter menggunakan metode *naive bayes* dengan evaluasi *confusion matrix* dapat diterapkan dengan materi lain yang memiliki karakteristik khusus dengan pustaka tertentu.

2. Saran untuk penelitian lanjutan
 - a. Analisis sentimen pada media sosial twitter menggunakan metode *naïve bayes* dengan evaluasi *confusion matrix* memerlukan perbaikan dan pengembangan lebih lanjut guna terbentuknya kesempurnaan modul.
 - b. Analisis sentimen pada media sosial twitter menggunakan metode *naïve bayes* dengan evaluasi *confusion matrix* memerlukan uji penyebaran secara luas untuk menyempurnakan tahapan penelitian.

DAFTAR PUSTAKA

- [1] N. Ni'matul Rohmah, "Media Sosial Sebagai Media Alternatif Manfaat dan Pemuas Kebutuhan Informasi Masa Pandemi Global Covid 19 (Kajian Analisis Teori Uses And Gratification)," vol. 4, no. 1, pp. 1–16, 2020, [Online]. Available: <https://www.kompas.com/tren/read/2020/03/29/092500765/update-virus-corona-di-dunia-29-maret--662.073-kasus-di-200->
- [2] S. Shinde, P. Ladekar, P. Kumari, and D. M. I. E. T. R. Wardha, "SENTIMENT ANALYSIS ON TWITTER".
- [3] M. Murni, I. Riadi, and A. Fadlil, "Analisis Sentimen HateSpeech pada Pengguna Layanan Twitter dengan Metode Naïve Bayes Classifier (NBC)," *JURIKOM (Jurnal Riset Komputer)*, vol. 10, no. 2, p. 566, Apr. 2023, doi: 10.30865/jurikom.v10i2.5984.
- [4] M. M. Munir, M. A. Fauzi, and R. S. Perdana, "Implementasi Metode Backpropagation Neural Network berbasis Lexicon Based Features dan Bag of Words Untuk Identifikasi Ujaran Kebencian Pada Twitter," 2018. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [5] K. Zuhri, N. Adha, and O. Saputri, "ANALISIS SENTIMEN MASYARAKAT TERHADAP PILPRES 2019 BERDASARKAN OPINI DARI TWITTER MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER," *Bina Darma Conference on Computer Science*, 2020.
- [6] M. Murni, I. Riadi, and A. Fadlil, "Analisis Sentimen HateSpeech pada Pengguna Layanan Twitter dengan Metode Naïve Bayes Classifier (NBC)," *JURIKOM (Jurnal Riset Komputer)*, vol. 10, no. 2, p. 566, Apr. 2023, doi: 10.30865/jurikom.v10i2.5984.
- [7] P. Antinasari, R. Setya Perdana, and M. A. Fauzi, "Analisis Sentimen Tentang Opini Film Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Naive Bayes Dengan Perbaikan Kata Tidak Baku," 2017. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [8] S. Diajukan untuk Memenuhi Persyaratan Memperoleh Gelar Sarjana Komputer Disusun oleh FARHAN RIDHO, "RANCANG BANGUN APLIKASI WEB CRAWLING UNTUK Mencari Harga Barang Termurah dari Berbagai E-Marketplace (Studi Kasus: Tokopedia, Bukalapak, Shopee)."
- [9] M. Syarifuddin, "ANALISIS SENTIMEN OPINI PUBLIK MENGENAI COVID-19 PADA TWITTER MENGGUNAKAN METODE NAÏVE BAYES DAN KNN," *INTI Nusa Mandiri*, vol. 15, no. 1, pp. 23–28, Aug. 2020, doi: 10.33480/inti.v15i1.1347.
- [10] L. Aji Andika and P. Amalia Nur Azizah, "Analisis Sentimen Masyarakat terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier," 2019.
- [11] O. Dwiraswati and K. Nazaruddin Siregar, "ANALISIS SENTIMEN PADA TWITTER TERHADAP PENGGUNAAN ANTIBIOTIK DI INDONESIA DENGAN NAIVE BAYES CLASSIFIER SENTIMENT ANALYSIS ON TWITTER ABOUT THE USE OF ANTIBIOTICS IN INDONESIA WITH NAIVE BAYES CLASSIFIER," 2019. [Online]. Available: www.search.twitter.com
- [12] Q. Bi, K. E. Goodman, J. Kaminsky, and J. Lessler, "What is machine learning? A primer for the epidemiologist," *Am J Epidemiol*, vol. 188, no. 12, pp. 2222–2239, Dec. 2019, doi: 10.1093/aje/kwz189.

- [13] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From Data Mining to Knowledge Discovery in Databases," 1996. [Online]. Available: www.ffly.com/
- [14] J. Jackson, "Data Mining: A Conceptual Overview," *Communications of the Association for Information Systems*, vol. 8, 2002, doi: 10.17705/1cais.00819.
- [15] T. Silwattananusarn, "Data Mining and Its Applications for Knowledge Management : A Literature Review from 2007 to 2012," *International Journal of Data Mining & Knowledge Management Process*, vol. 2, no. 5, pp. 13–24, Sep. 2012, doi: 10.5121/ijdkp.2012.2502.
- [16] J. Eka Sembodo, E. Budi Setiawan, and Z. Abdurahman Baizal, "Data Crawling Otomatis pada Twitter," School of Computing, Telkom University, Sep. 2016, pp. 11–16. doi: 10.21108/indosc.2016.111.
- [17] M. Taboada, "Sentiment Analysis: An Overview from Linguistics," Jan. 14, 2016, *Annual Reviews Inc.* doi: 10.1146/annurev-linguistics-011415-040518.
- [18] A. I. Kadhim, "An Evaluation of Preprocessing Techniques for Text Classification." [Online]. Available: <https://sites.google.com/site/ijcsis/>
- [19] M. Sudhir, M. Gorade, A. Deo, and P. Purohit, "A Study of Some Data Mining Classification Techniques," *International Research Journal of Engineering and Technology*, 2017, [Online]. Available: www.irjet.net
- [20] V. Krishnaiah, G. Narsimha, and N. S. Chandra, "Survey of Classification Techniques in Data Mining," 2014, [Online]. Available: www.ijcaonline.org
- [21] S. F. Rodiyansyah and E. Winarko, "Klasifikasi Posting Twitter Kemacetan Lalu Lintas Kota Bandung Menggunakan Naive Bayesian Classification," *IJCCS*, vol. 6, no. 1, pp. 91–100, 2012, [Online]. Available: <http://bit.ly/mHibqV>
- [22] E. Beauxis-Aussalet and L. Hardman, "Simplifying the Visualization of Confusion Matrix."
- [23] J. Homepage, S. Sahibu, I. Taufik, and P. Studi Sistem Komputer, "MALCOM: Indonesian Journal of Machine Learning and Computer Science Implementation of the Convolutional Neural Network Algorithm for Classifying Types of Organic and Non-Organic Waste Implementasi Algoritma Convolutional Neural Network untuk Klasifikasi Jenis Sampah Organik dan Non Organik," vol. 4, no. 3, pp. 840–852, 2024, doi: 10.57152/malcom.v7i2.1346.
- [24] G. Lindstrom, "Programming with Python," 2005. [Online]. Available: <http://www.python.org/doc/Humor.html#zen>.
- [25] K. Salo, "Comparative study on Python web frameworks: Flask and Django."
- [26] F. Pedregosa FABIANPEDREGOSA *et al.*, "Scikit-learn: Machine Learning in Python Gaël Varoquaux Bertrand Thirion Vincent Dubourg Alexandre Passos PEDREGOSA, VAROQUAUX, GRAMFORT ET AL. Matthieu Perrot," 2011. [Online]. Available: <http://scikit-learn.sourceforge.net>.
- [27] M. Wang and F. Hu, "The application of nltk library for python natural language processing in corpus research," *Theory and Practice in Language Studies*, vol. 11, no. 9, pp. 1041–1049, Sep. 2021, doi: 10.17507/tpls.1109.09.

- [28] E. Loper and S. Bird, "NLTK: The Natural Language Toolkit," 2002. [Online]. Available: <http://nltk.sf.net/>.