

# Analisa *Tweet* Mahasiswa untuk Deteksi Gejala Depresi dengan Penerapan *Natural Language Processing*

Monica Yoshe Dhinora <sup>1\*</sup>, Evangs Mailoa <sup>2</sup>

<sup>1\*,2</sup>Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana, Kota Salatiga, Provinsi Jawa Tengah, Indonesia.

Email: 672021032@student.uksw.edu <sup>1\*</sup>, evangs.mailoa@uksw.edu <sup>2</sup>

## Histori Artikel:

*Dikirim* 10 Maret 2025; *Diterima dalam bentuk revisi* 20 April 2025; *Diterima* 6 Mei 2025; *Diterbitkan* 10 Mei 2025. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STMIK Indonesia Banda Aceh.

## Abstrak

Isu kesehatan mental semakin menyita perhatian, dimana depresi menjadi faktor utama yang sering dikaitkan dengan tingginya angka bunuh diri akibat gangguan psikologis. Mahasiswa disebut sebagai kelompok rentan mengalami depresi dan kecemasan, yang dapat dipicu oleh berbagai faktor. Disisi lain, cara individu mengungkapkan diri di media sosial, khususnya pada platform X (Twitter) yang menawarkan kebebasan berekspresi, dinilai mampu mencerminkan kondisi kesejahteraan mental seseorang. Penelitian ini bertujuan untuk mengeksplorasi analisis tweet mahasiswa dengan pendekatan Natural Language Processing (NLP) guna mendeteksi gejala depresi melalui pola kebahasaan. Data dikumpulkan melalui teknik crawling dengan kata kunci seperti “depresi”, “stres”, dan “burnout”, yang menghasilkan 24.167 tweet dalam rentang Januari hingga Maret 2025. Sebanyak 8.308 tweet tersisa setelah dilakukan pembersihan data. Pelabelan sentimen dengan Inset Lexicon menunjukkan 68,1% atau sejumlah 5.663 tweet berlabel negatif, yang menunjukkan kecenderungan mahasiswa dalam menggunakan platform X sebagai media untuk meluapkan emosi negatif. Model Random Forest yang diintegrasikan dengan ekstraksi fitur TF-IDF mencapai akurasi 87,51% yang menunjukkan kemampuannya dalam mengatasi bias kelas mayoritas (negatif) dan menangkap kompleksitas morfologi bahasa informal. Implikasi penelitian mendorong pengembangan sistem pemantauan digital deteksi gejala depresi mahasiswa secara proaktif. Keterbatasan leksikon dalam memuat kosakata informal (slang) menjadi rekomendasi studi lanjutan untuk meningkatkan akurasi analisis.

**Kata Kunci:** Deteksi Depresi; Mahasiswa; Tweet; Natural Language Processing; Random Forest; TF-IDF.

## Abstract

Mental health issues are increasingly gaining attention, with depression being a primary factor linked to high suicide rates caused by psychological disorders. College students are identified as a vulnerable group to depression and anxiety, which can be triggered by various factors. Meanwhile, individuals self-expression on social media, especially on platform X (Twitter), which offers freedom of expression, is considered reflective of one's mental well-being. This study aims to explore the analysis of college students tweets using a Natural Language Processing (NLP) approach to detect depressive symptoms through linguistic patterns. Data was collected via crawling techniques using keywords such as “depression”, “stress”, and “burnout” resulting in 24,167 tweets from January to March 2025. After data cleaning, 8,308 tweets remained. Sentiment labeling using the Inset Lexicon shows that 68.1% (5,663 tweets) were labeled negative, reflecting college students tendency to use platform X as a medium to express negative emotions. The Random Forest model integrated with TF-IDF feature extraction achieved 87.51% accuracy, demonstrating its capability to address majority class bias (negative) and capture the morphological complexity of informal language. The implications of the research encourage the development of a digital monitoring system for the proactive detection of college college students depression symptoms. The lexicon's limitations in incorporating informal vocabulary (slang) becomes a recommendation for further research to enhance analysis accuracy.

**Keyword:** Depression Detection; College Student; Tweet; Natural Language Processing; Random Forest; TF-IDF.

## 1. Pendahuluan

Kemajuan teknologi kian mengubah cara masyarakat dalam bersosialisasi, dengan media sosial sebagai pendorong utama terjadinya interaksi (Cahyono, 2021). Melalui media sosial, pengguna dapat saling terhubung untuk bertukar informasi, hingga tergabung dalam komunitas global (Fahma *et al.*, 2024). Media sosial mampu memberikan ruang kebebasan bagi setiap orang untuk mengekspresikan diri (Stepanus Angga *et al.*, 2023). Hal ini menjadikannya sarana yang memegang peranan penting dalam berbagai aspek kehidupan manusia (Nur Erma *et al.*, 2022). Menurut laporan “Digital 2024” oleh *We Are Social* dan *Data Reportal* pada Januari 2024, terdapat 139 juta pengguna aktif media sosial di Indonesia, atau sekitar 49,9% dari total 279 juta jiwa penduduk. Rata-rata, masyarakat Indonesia menghabiskan 7 jam 38 menit per hari di internet, dengan 3 jam 11 menit di antaranya digunakan untuk mengakses media sosial. Data ini menunjukkan, penggunaan media sosial di Indonesia pada tahun 2024 nyaris mencapai separuh dari total waktu penggunaan internet. Berdasarkan skema klasifikasi, media sosial terbagi menjadi enam kategori, yaitu *Collaborative Projects*, *Blogs and Microblogs*, *Content Communities*, *Social Networking Sites*, *Virtual Game Worlds*, dan *Virtual Social Worlds* (Cahyono, 2021). Di antaranya, X (Twitter) menjadi platform *microblogging* yang populer dalam hal kebebasan berekspresi (Nur & Rusly, 2023). Dengan sekitar 24,69 juta pengguna X di Indonesia pada awal tahun 2024, sebagian besar penggunaanya berada dalam rentang usia 18 hingga 25 tahun (Tamaraya & Ubaedullah, 2021). Di mana mayoritasnya adalah kelompok remaja, yang di dominasi oleh mahasiswa (Rizqi & Heriyanto, 2023). Isu kesehatan mental semakin disorot. Depresi menjadi faktor paling umum ditemui, dari banyaknya kasus bunuh diri akibat gangguan psikologis (Distina, 2019). Dalam berbagai penelitian, mahasiswa disebut sebagai kelompok rentan mengalami depresi dan kecemasan, yang dapat dipicu oleh berbagai faktor, seperti tuntutan akademik, keluarga, sosial, hingga tantangan personal (Santoso *et al.*, 2022)(Putra *et al.*, 2023). Namun, seringkali seseorang merasa lega setelah mengekspresikan diri melalui *tweet* yang dibagikan. Sehingga pola pengungkapan ini dinilai mampu mencerminkan kondisi kesejahteraan mental seseorang (Zaskya *et al.*, n.d.). Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk memberikan informasi terkait distribusi sentimen yang terkandung dalam *tweet* mahasiswa, serta menguji efektivitas model *Random Forest* dalam mengidentifikasi *tweet* bermuatan gejala depresi, sebagai dasar pengembangan sistem deteksi dini yang adaptif terhadap dinamika bahasa pada *tweet* mahasiswa.

Pada penelitian yang berjudul “Klasifikasi Teks untuk Mendeteksi Depresi dan Kecemasan Pada Pengguna Twitter Berbasis *Machine Learning*”, dilakukan perbandingan algoritma klasifikasi teks dengan metode *machine learning* untuk mengidentifikasi gejala depresi dan kecemasan berdasarkan postingan *tweet*. Hasil menunjukkan, algoritma *Random Forest* berhasil mencapai akurasi tertinggi sebesar 96%, dibandingkan algoritma *Decision Tree*, *Naive Bayes*, dan *K-Nearest Neighbor* (KNN) (Rahayu *et al.*, 2023). Keunggulan ini disebabkan oleh kemampuan *Random Forest* dalam menangani kompleksitas teks yang implisit melalui mekanisme *ensemble learning*, di mana *multiple decision tree* dilatih secara paralel pada subset data acak (*bagging*) untuk mengurangi varians dan overfitting. Setiap *tree* melakukan pemilihan fitur acak (*feature randomness*) dan memilih *split node* berdasarkan kriteria *entropy* untuk mengukur ketidakhomogenan kelas. Pada kasus data tidak seimbang, *Random Forest* secara intrinsik menerapkan pembobotan kelas (*class weighting*) dengan memberi penalti lebih besar pada kesalahan klasifikasi kelas minoritas (*tweet* positif), sehingga dapat meningkatkan nilai *recall*. Berbeda dengan *Naive Bayes* yang kurang optimal dalam menangani dependensi antar fitur, serta kinerja KNN yang menurun pada vektor TF-IDF berdimensi tinggi, *Random Forest* dinilai tetap mempertahankan stabilitasnya dengan reduksi dimensionalitas melalui seleksi fitur acak. Temuan ini menjadi dasar pemilihan algoritma *Random Forest* dalam penelitian saat ini, terutama untuk menangani data tidak seimbang seperti dominasi *tweet* yang mengandung emosi negatif, serta mengatasi kompleksitas morfologi bahasa Indonesia seperti kata berafiks yang sering digunakan mahasiswa dalam berekspresi. Dalam penelitian lain berjudul “Kemungkinan Depresi dari Postingan pada Sosial Media”, teknik *Natural Language Processing* (NLP) dan algoritma *Support Vector Machine* (SVM) diterapkan pada analisis sentimen terhadap 12.402 *tweet* yang berhasil dikumpulkan. Proses berfokus pada identifikasi emosi

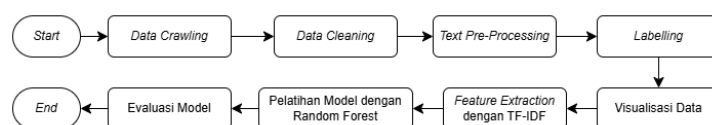
negatif dan kemarahan sebagai indikator depresi (Mutmainah, 2022). Tahap ekstraksi fitur dilakukan menggunakan kamus LIWC, topik LDA dan fitur *N-gram* (bigram) yang dikombinasikan dengan TF-IDF. Pada penelitian ini, teknik *N-gram* berperan dalam menangkap konteks kata berurutan, seperti frasa emosional negatif yang sering muncul pada ekspresi mahasiswa. Sementara TF-IDF (*Term Frequency-Inverse Document Frequency*) memberikan bobot pada kata berdasarkan frekuensi kemunculan dalam dokumen (*term frequency*) dan kelangkaan kata dalam seluruh korpus (*inverse document frequency*). Keunggulan TF-IDF terletak pada kemampuannya dalam mengidentifikasi kata kunci signifikan secara kontekstual. Hasil menunjukkan bahwa kombinasi tersebut secara efektif mampu mengklasifikasikan pola linguistik depresi dengan akurasi mencapai 95.56% pada algoritma SVM. Berdasarkan kesesuaiannya, teknik kombinasi *N-gram* dan TF-IDF dijadikan acuan dalam pengklasifikasian pola linguistik pada penelitian yang dilakukan saat ini.

Sebuah penelitian berjudul “*Depression Detection on Social Media Twitter Using XLNet Method*” membahas tentang penggunaan metode XLNet, yaitu teknik NLP berbasis *transformer* untuk mengidentifikasi gejala depresi melalui analisis *tweet*. Metode ini mengadopsi *Permutation Language Modeling* (PLM) yang memungkinkan pembelajaran konteks bidirectional dengan mempertimbangkan semua permutasi urutan kata, sehingga efektif dalam menangkap dependensi jarak jauh dan ambiguitas semantik pada frasa informal seperti kosakata slang dan dialek. Penerapan arsitektur *two-stream self-attention* pada XLNet dengan menggabungkan *content stream* (akses konteks penuh) dan *query stream* (berfokus pada posisi tanpa kebocoran informasi target) terbukti unggul dalam memodelkan ekspresi emosional implisit di media sosial, yang mencapai akurasi rata-rata hingga 93,33% setelah optimasi *hyperparameter* (Apriliani & Maharani, 2023). Namun, penerapan XLNet dalam penelitian ini memiliki keterbatasan, seperti kebutuhan sumber daya komputasi tinggi untuk pelatihan dan penyempurnaan model (*fine-tuning*), yang menjadi kendala di lingkungan dengan infrastruktur terbatas. Selain dinilai rentan mengalami *overfitting* pada data berskala kecil, XLNet kurang mampu dalam menangani struktur morfologi kompleks bahasa Indonesia, seperti kata berafiks dan slang, yang berisiko mengabaikan nuansa kontekstual penting dalam identifikasi depresi. Meskipun pendekatan yang digunakan berbeda, penelitian ini memiliki landasan serupa dalam pemanfaatan teknik NLP untuk mengidentifikasi pola kebahasaan terkait depresi melalui analisis teks. Hal ini menjadikannya referensi penting bagi pengembangan model lebih lanjut dalam penelitian yang dilakukan saat ini.

Berbagai penelitian sebelumnya telah menunjukkan efektivitas *Natural Language Processing* (NLP) dan algoritma *machine learning* dalam mengidentifikasi tanda-tanda depresi melalui analisis konten teks di media sosial. Temuan ini memperkuat landasan teoretis bahwa pendekatan analisis linguistik berbasis NLP dapat menjadi solusi potensial untuk deteksi dini gejala depresi. Namun, sebagian besar penelitian tersebut masih bersifat umum dan belum menyoroti kelompok tertentu seperti mahasiswa, sebagai populasi yang diketahui rentan mengalami depresi akibat tuntutan akademik maupun sosial (Santoso *et al.*, 2022)(Putra *et al.*, 2023). Oleh karena itu, penelitian ini dirancang untuk mengatasi celah tersebut dengan membatasi sasaran analisis secara khusus pada kelompok mahasiswa sebagai subjek utama.

## 2. Metode Penelitian

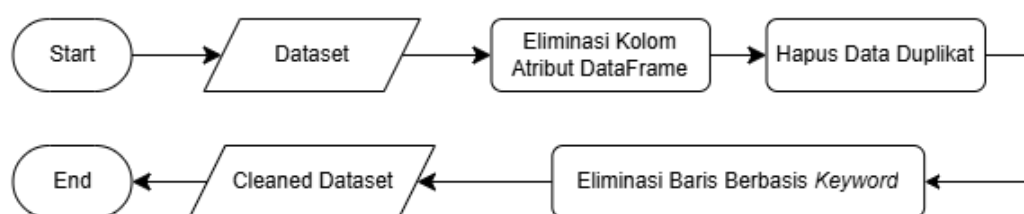
Penelitian ini melibatkan beberapa tahapan sistematis, yang dirancang guna memastikan setiap proses berjalan sesuai dengan prosedur yang telah ditetapkan, seperti yang diilustrasikan dalam diagram alir berikut.



Gambar 1. Alur Proses Penelitian

### 2.1 Data Crawling

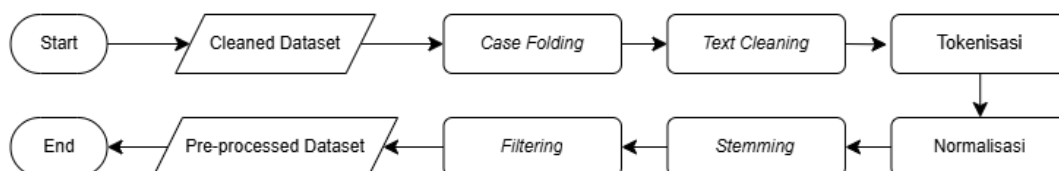
Analisis *Natural Language Processing* (NLP) dalam penelitian ini bertujuan mengidentifikasi pola linguistik pada *tweet* mahasiswa terkait gejala awal depresi. Pada Gambar 1, tertera langkah-langkah yang diimplementasikan dalam penelitian ini. Tahap pertama mencakup proses pengumpulan data, dengan menerapkan teknik *crawling* menggunakan *Authenticated Web Scraping*, yang memungkinkan peneliti untuk mengekstraksi *tweet* berdasarkan parameter kata kunci yang ditentukan. Kata kunci seperti “mahasiswa”, “kuliah”, “depresi”, “stres”, “tekanan”, “capek” dan “burnout”, dipilih berdasarkan relevansinya dengan isu kesehatan mental, yang mengarah pada indikasi gejala depresi. Sebanyak 24.167 data *tweet* berhasil dikumpulkan melalui proses *crawling* dengan rentang waktu pengambilan data pada Januari hingga Maret 2025, untuk memperoleh sampel yang dapat merepresentasikan variasi data aktual dan relevan dengan ekspresi emosi mahasiswa.



Gambar 2. Alur Kerja Tahapan *Data Cleaning*

### 2.2 Data Cleaning

Tahap pembersihan data dilakukan untuk menjaga validitas dan kelayakan dataset sebelum diproses lebih lanjut. Pada Gambar 2, dijelaskan langkah pertama dalam tahap ini adalah mengeliminasi kolom-kolom atribut dari *DataFrame* yang tidak diperlukan dalam proses analisis teks. Langkah ini bertujuan untuk menghilangkan data yang mengandung *noise*, menghindari bias dari fitur non-tekstual, dan menjadikan konten teks (*tweet*) sebagai variabel utama dalam proses analisis. Langkah berikutnya adalah identifikasi dan penghapusan data duplikat pada kolom *Text*, yang muncul akibat *retweet*, *re-post*, juga aktivitas bot. Adanya rekaman data identik secara berulang dapat memicu bias dalam analisis, yang berpotensi mempengaruhi kinerja model, karena ketidakseimbangan distribusi data. Untuk meningkatkan kualitas dataset, proses ini memastikan setiap entri data bersifat unik dan representatif. Tahap akhir dari pembersihan data adalah proses eliminasi baris berbasis kata kunci (*keywords*) yang dianggap sebagai *noise* (informasi tidak relevan). Proses ini dilakukan dengan mendefinisikan daftar kata kunci seperti “*toktokers*”, “*geoya*”, “*hide*”, dan kata lain yang muncul dalam konteks non-emosional, seperti promosi atau diskusi diluar topik. Setiap *tweet* yang mengandung kata kunci tersebut dalam bentuk kata utuh (bukan *substring*), akan diidentifikasi menggunakan *regular expression* (*regex*) dengan pola `\b` (*word boundary*). Setelah kata kunci terdeteksi, nilai kolom teks pada baris tersebut diubah menjadi string kosong (“”) dan setiap baris tersebut akan dihapus dari dataset. Setelah melalui serangkaian proses pembersihan data, dari total 24.167 data *tweet* yang terkumpul, sebanyak 8.308 data (34,38%) berhasil memenuhi kriteria relevansi dan membentuk dataset akhir yang siap dianalisis lebih lanjut.



Gambar 3. Alur Kerja Tahapan *Text Pre-processing*

### 2.3 Text Pre-processing

Pada Gambar 3, diilustrasikan langkah-langkah yang diterapkan dalam tahap pra-pemrosesan teks. Tahapan ini dilakukan untuk mentransformasi data *tweet* menjadi lebih terstruktur dan memastikan format teks sesuai dengan kebutuhan analisis (Fachriza & Artikel, 2023). Data *tweet* mahasiswa diolah melalui beberapa tahapan utama. Dimulai dengan penerapan *case folding* untuk mengonversi seluruh karakter menjadi huruf kecil, untuk menyeragamkan format teks. Teks dibersihkan dari karakter khusus, angka, tautan, spasi berlebih, dan karakter tunggal tak bermakna pada tahap *text cleaning*. Pada tahap tokenisasi, teks dipecah menjadi unit leksikal (token), dengan memanfaatkan fungsi *word\_tokenize* dari *library* NLTK (*Natural Language Toolkit*). Kata tidak baku dinormalisasi ke bentuk baku menggunakan kamus *leksikon*. Proses dilanjutkan dengan *stemming* menggunakan *library* Sastrawi untuk mereduksi kata ke bentuk dasar. Tahap terakhir dalam *pre-processing* adalah *filtering*, yaitu menghapus *stop words* atau kata-kata umum yang tidak memberikan banyak makna dalam analisis sentimen. Proses ini menggunakan *library* NLTK untuk mendapatkan daftar *stop words* bahasa Indonesia, serta diperluas dengan menambahkan *stop words* yang relevan dengan konteks penelitian, guna menghilangkan kata penghubung dan partikel yang tidak informatif. Secara keseluruhan, proses ini menghasilkan sekumpulan token yang siap diolah, untuk mengidentifikasi pola-pola linguistik yang mengindikasikan adanya gejala depresi.

### 2.4 Labelling

Proses pelabelan sentimen dilakukan dengan mencocokkan setiap kata dalam teks terhadap *Inset Lexicon*, yang terdiri dari *leksikon* positif dan *leksikon* negatif. Setiap kata yang terdaftar dalam *leksikon* positif akan menambah skor sentimen sesuai bobot yang telah ditetapkan, sementara kata dalam *leksikon* negatif akan mengurangi skor berdasarkan bobotnya. Total skor sentimen dihitung berdasarkan akumulasi seluruh bobot kata yang teridentifikasi (Musfiroh *et al.*, 2021). Jika hasil skor akhir bernilai nol atau positif ( $\geq 0$ ), teks diberi label “*Positive*”, sedangkan skor bernilai negatif ( $< 0$ ) akan menghasilkan label “*Negatif*” (Musfiroh *et al.*, 2021).

### 2.5 Analisis N-gram

Analisis *N-gram* merupakan metode yang mengidentifikasi rangkaian kata berurutan (*n* unit leksikal) dalam teks untuk mengungkap pola bahasa kontekstual, terutama pada *tweet* mahasiswa (Zhu *et al.*, 2022). Pada penelitian ini, trigram (3-gram) digunakan untuk mengekstraksi frasa yang merepresentasikan konteks emosional kompleks terkait gejala depresi. Pemilihan trigram didasarkan pada kemampuannya menangkap hubungan semantik antar kata secara lebih utuh dibandingkan unigram (1-gram) atau bigram (2-gram). Hal ini memungkinkan deteksi ekspresi tidak langsung seperti metafora atau keluhan tersirat, yang sulit teridentifikasi melalui analisis kata tunggal. Hasil analisis *N-gram* menjadi dasar identifikasi pola linguistik terkait depresi, seperti kombinasi kata yang mengindikasikan ketertekanan, kelelahan mental, hingga keputusan, yang sering muncul dalam komunikasi informal di media sosial.

### 2.6 Feature Extraction dengan TF-IDF

Tahap ekstraksi fitur bertujuan untuk mengonversi data teks *tweet* yang telah melalui tahap *pre-processing* ke dalam bentuk numerik untuk diproses oleh model *machine learning*. Dalam penelitian ini, metode TF-IDF (*Term Frequency-Inverse Document Frequency*) digunakan untuk menghitung bobot tiap kata berdasarkan frekuensi kemunculan kata dalam dokumen (*term frequency*) dan kelangkaan kata di seluruh korpus (*inverse document frequency*) (Gifari *et al.*, 2022). Kata yang sering muncul dalam sebuah dokumen tetapi jarang ditemukan di dokumen lain akan memiliki bobot tinggi, yang menandakan kepentingannya dalam merepresentasikan keunikan dokumen tersebut. Implementasi dilakukan dengan memanfaatkan *TfidfVectorizer* dari *library* *scikit-learn*, di mana parameter seperti *max\_features* (membatasi jumlah fitur), *min\_df* dan *max\_df* (mengatur frekuensi minimal dan maksimal kata dalam dokumen), serta *stop\_words* (menghapus kata umum bahasa Indonesia) diatur untuk mengoptimalkan seleksi fitur. Objek TF-IDF dilatih pada data latih untuk membangun kosakata dan menghitung

bobot, kemudian diaplikasikan pada data latih dan uji melalui metode *transform*, sehingga menghasilkan matriks numerik yang disimpan dalam *DataFrame*.

## 2.6 Pelatihan Model dengan *Random Forest*

Pemilihan *Random Forest* sebagai algoritma pelatihan model dalam penelitian ini didasarkan pada kemampuannya dalam mengatasi ketidakseimbangan kelas melalui pembobotan otomatis (*class weighting*), sebagaimana didukung oleh penelitian sebelumnya yang memperoleh akurasi 96% pada data serupa, sekaligus untuk mengatasi keterbatasan sumber daya yang dapat menghambat penggunaan model kompleks seperti XLNet (Rahayu *et al.*, 2023)(Apriliani & Maharani, 2023). Pada implementasinya, model *Random Forest* dari *library scikit-learn* dikonfigurasi dengan parameter seperti *n\_estimators=200* (jumlah *decision tree*), *criterion='entropy'* (optimasi *split* berdasarkan entropi), dan *random\_state=0* (konsistensi replikasi). Proses pelatihan menggunakan vektor TF-IDF ( $X_{train}$ ) dan label sentimen ( $y_{train}$ ), sementara evaluasi dilakukan pada data uji ( $X_{test}, y_{test}$ ) untuk memvalidasi kemampuan generalisasi model. Pendekatan *ensemble learning* ini dipilih untuk meningkatkan akurasi klasifikasi sekaligus mengurangi risiko *overfitting* melalui agregasi hasil prediksi dari *multiple decision trees*.

## 2.7 Evaluasi Model

Kinerja model *Random Forest* yang diintegrasikan dengan TF-IDF dievaluasi melalui tiga metrik utama, yaitu *confusion matrix*, *classification report*, dan akurasi. *Confusion matrix* digunakan untuk menampilkan distribusi prediksi benar (*True Positive/Negative*) dan salah (*False Positive/Negative*) pada tiap kelas sentimen. Sementara *classification report* memberikan detail analisis berupa *precision* (ketepatan prediksi), *recall* (kemampuan deteksi kasus sebenarnya), dan *F1-score* (keseimbangan *precision-recall*). Nilai akurasi dihitung sebagai persentase total prediksi benar terhadap keseluruhan data. Untuk visualisasi yang intuitif, *confusion matrix* ditampilkan dalam bentuk *heatmap* dengan gradasi warna yang merepresentasikan frekuensi klasifikasi. Selanjutnya, dilakukan optimasi *hyperparameter* menggunakan *Grid Search* pada model *Logistic Regression* untuk memperoleh kombinasi parameter terbaik. Proses ini menguji variasi nilai *regularization strength (C)* dan jenis *regularization (penalty)* melalui pelatihan model dengan *10-fold cross-validation*. Setiap kombinasi parameter dievaluasi berdasarkan akurasi tertinggi yang dihasilkan pada data latih, dimana parameter optimal tersebut digunakan untuk membangun model akhir.

# 3. Hasil dan Pembahasan

## 3.1 Hasil

Analisis terhadap 8.318 *tweet* mahasiswa menghasilkan dua temuan utama. Pertama, hasil menunjukkan bahwa distribusi *tweet* dengan sentimen negatif mendominasi sebanyak 68,1% (5.663 *tweet*), hampir dua kali lipat lebih banyak dibandingkan *tweet* positif yang hanya 31,9% (2.655 *tweet*). Dimana temuan ini menunjukkan kecenderungan mahasiswa dalam menggunakan platform X sebagai media untuk meluapkan emosi negatif. Kedua, model *Random Forest* yang diintegrasikan dengan ekstraksi fitur TF-IDF untuk mengklasifikasikan *tweet* mahasiswa berhasil mencapai akurasi 87,51% pasca dilakukannya optimasi *hyperparameter*. Hal ini menyoroti potensi *Random Forest* dalam menangani ketidakseimbangan data, seperti dominasi pada kelas negatif yang berpotensi menyebabkan bias model terhadap kelas mayoritas. Pada penelitian ini, proses pengambilan data (*crawling*) dari platform X menghasilkan 24.167 *tweet* yang dikumpulkan dalam rentang waktu Januari hingga Februari 2025. Secara umum, kumpulan *tweet* ini menggambarkan kehidupan sosial, akademik, serta politik di kalangan mahasiswa Indonesia.

Tabel 1. Dataset Primer

id	conversation_id_str	created_at	favorite_count	Text	...
0	1890763475871281480	Sat Feb 15 14:01:55 +0000 2025	5349	Hari ini kami mahasiswa UI merasa resah dengan...	...
1	1891056327016804523	Sun Feb 16 09:25:37 +0000 2025	512	damn besok kuliah <a href="https://t.co/lfw2ZQoo9O">https://t.co/lfw2ZQoo9O</a>	...
2	1891099083038724281	Sun Feb 16 12:15:30 +0000 2025	100	Koleksi toktokers yang lgi viral <a href="https://t.co/...">https://t.co/...</a>	...
3	1889685319617839305	Wed Feb 12 14:37:43 +0000 2025	19241	gw rasa kuliah bukan passion gw passion gw bac...	...
4	1890702508181332270	Sat Feb 15 09:59:40 +0000 2025	9558	lu pada stres gasi ngeliat dk casual mode ceo ...	...
...	...	...	...	...	...
24165	1892031775888556213	Wed Feb 19 02:01:42 +0000 2025	15	w stress banget (masih pagi ini ajg) <a href="https://t...">https://t...</a>	...
24166	1892032599658258618	Wed Feb 19 02:04:58 +0000 2025	0	Kemarin pas scrolling LinkedIn kebetulan nemu ...	...

Pada Tabel 1, divisualisasikan struktur dataset yang terdiri dari 15 kolom atribut, meliputi informasi teknis dan interaksi pengguna. Kolom *conversation\_id\_str* berperan sebagai pengenal unik setiap *tweet*, sementara *created\_at* merekam waktu publikasi konten. Konten utama *tweet* tercatat dalam kolom *Text*, sedangkan indikator keterlibatan audiens seperti jumlah suka ditampilkan pada kolom *favorite\_count* dan frekuensi penyebaran *tweet* pada *retweet\_count*. Kolom pendukung seperti *location* merekam lokasi geografis pengguna, dan *image\_url* merujuk pada tautan media visual yang dilampirkan

Tabel 2. Cleaned Dataset

id	Text
0	Hari ini kami mahasiswa UI merasa resah dengan...
1	damn besok kuliah <a href="https://t.co/lfw2ZQoo9O">https://t.co/lfw2ZQoo9O</a>
3	gw rasa kuliah bukan passion gw passion gw bac...

4	lu pada stres gasi ngeliat dk casual mode ceo ...
5	like i struggle to bayar ukt kuliah but im not...
...	...
23164	w stress banget (masih pagi ini ajg) https://t...
23165	Kemarin pas scrolling LinkedIn kebetulan nemu ...

Tabel 2 menampilkan visualisasi data yang telah melalui proses pembersihan. Tahapan ini berhasil mentransformasi struktur dataset menjadi lebih terfokus. Dengan menyisakan kolom *Text*, sejumlah 14 kolom atribut dihapus untuk menghindari bias analisis akibat metadata teknis. Sebanyak 15.788 data duplikat pada kolom *Text* dihapus untuk mencegah overfitting, hingga tersisa 8.379 baris data unik. Proses dilanjutkan dengan eliminasi 61 baris *tweet* yang memuat kata kunci tidak relevan, untuk meminimalkan *noise* dan memusatkan analisis pada pola linguistik yang berkaitan dengan indikator gejala depresi. Hasil akhir dari tahapan ini berupa dataset terkurasi, dengan 8.318 data teks yang telah terstandarisasi.

Tabel 3. Hasil Keluaran Proses Case Folding

id	Text Case Folding
0	hari ini kami mahasiswa ui merasa resah dengan...
1	damn besok kuliah https://t.co/lfw2zqoo9o
3	gw rasa kuliah bukan passion gw passion gw bac...
4	lu pada stres gasi ngeliat dk casual mode ceo ...
5	like i struggle to bayar ukt kuliah but im not...
...	...
23164	w stress banget (masih pagi ini ajg) https://t...
23165	kemarin pas scrolling linkedin kebetulan nemu ...

Proses *case folding* pada penelitian ini dilakukan menggunakan fungsi *lower* dari *library Pandas* untuk mengubah seluruh karakter pada kolom *Text* ke dalam huruf kecil. Hasilnya disimpan dalam kolom baru bernama *Text Case Folding*. Hasil keluaran divisualisasikan pada Tabel 3. Tahapan ini berhasil menghilangkan variasi kapitalisasi pada data *tweet*, sehingga meminimalkan risiko perbedaan interpretasi selama analisis.

Tabel 4. Hasil Keluaran Proses Text Cleaning

id	Text Cleaning
0	hari ini kami mahasiswa ui merasa resah dengan...
1	damn besok kuliah
3	gw rasa kuliah bukan passion gw passion gw bac...
4	lu pada stres gasi ngeliat dk casual mode ceo ...
5	like struggle to bayar ukt kuliah but im not ...
...	...
23164	stress banget masih pagi ini ajg
23165	kemarin pas scrolling linkedin kebetulan nemu ...

Data yang telah melalui tahapan *text cleaning* divisualisasikan pada Tabel 4. Proses ini bertujuan untuk memurnikan data *tweet* dari elemen tak bermakna, guna memastikan konsistensi sebelum dilakukan analisis NLP. Tahapan diawali dengan penghapusan karakter spesial menggunakan fungsi *remove\_tweet\_special*, untuk menghapus *tab*, *newline*, *URL*, *mention*, *hashtag*, dan karakter non-ASCII, dengan kombinasi metode *replace*, *encode*, serta *regular expression (regex)*. Dengan memanfaatkan *re.sub*, angka dihapus menggunakan fungsi *remove\_number* untuk menghindari bias numerik dalam analisis semantik. Tanda baca juga dieliminasi melalui *remove\_punctuation* dengan memanfaatkan *str.maketrans*

dari modul *string*, yang secara efektif menyaring simbol non-verbal. Sementara *remove\_whitespace\_LT* dan *remove\_whitespace\_multiple* berperan untuk menormalisasi spasi menggunakan *strip* dan ekspresi reguler, untuk menghapus spasi berlebih. Tahap akhir melibatkan kombinasi fungsi *remove\_singl\_char* dan *re.sub* untuk menghilangkan karakter tunggal tak bermakna, seperti “a” atau “x”, serta memastikan hanya kata utuh yang dipertahankan. Seluruh rangkaian fungsi tersebut diimplementasikan pada kolom *Text Case Folding* menggunakan metode *apply* dari *library Pandas*, sehingga menghasilkan kolom *Text Cleaning* yang terbebas dari *noise* seperti tautan, emotikon, dan simbol tidak bermakna. Proses *text cleaning* berhasil mengoptimalkan konsistensi format data, dengan transformasi teks menjadi lebih terstruktur, penggunaan spasi yang konsisten, serta penyaringan kata tak bermakna. Pada Tabel 4, dimana keluaran teks telah bersih dari elemen redundan. Fungsi *library re* dan *string* terbukti optimal dalam mengelola pola rumit pada data tidak terstruktur, salah satunya melalui kapabilitas ekspresi reguler yang secara adaptif mendeteksi pola dinamis seperti *URL* dan *mention* dalam teks. Kombinasi kedua *library* ini memungkinkan identifikasi presisi terhadap elemen non-tekstual, serta memperkuat konsistensi data untuk analisis lebih lanjut.

Tabel 5. Hasil Keluaran Proses *Text Tokenizing*

id	Text Tokenizing
0	[hari, ini, kami, mahasiswa, ui, merasa, resah...
1	[damn, besok, kuliah]
3	[gw, rasa, kuliah, bukan, passion, gw, passion...
4	[lu, pada, stres, gasi, ngeliat, dk, casual, m...
5	[like, struggle, to, bayar, ukt, kuliah, but, ...
...	...
23164	[stress, banget, masih, pagi, ini, ajg]
23165	[kemarin, pas, scrolling, linkedin, kebetulan,...

Proses tokenisasi dalam penelitian ini menggunakan fungsi *word\_tokenize* dari *library NLTK* (*Natural Language Toolkit*) untuk memecah teks hasil proses pembersihan (*text cleaning*) menjadi unit kata individu. Tahapan diawali dengan mengunduh *Punkt Tokenizer* sebagai pendukung dan mendefinisikan fungsi *wrapper word\_tokenize\_wrapper* yang mengaplikasikan tokenisasi pada seluruh data di kolom *Text Cleaning* menggunakan metode *apply* dari *Pandas*. Hasilnya, setiap teks dalam dataset diubah menjadi bentuk *list* token yang tersimpan dalam kolom baru bernama *Text Tokenizing*, seperti divisualisasikan pada Tabel 5.

Tabel 6. Data Kamus Kata Baku

No	Slang	Formal
1	woww	wow
2	aminn	amin
3	met	selamat
4	netaas	menetas
5	keberpa	keberapa
...	...	...
15180	udh	sudah
15181	bgt	banget

Proses normalisasi dilakukan untuk menyelaraskan kata tidak baku dalam *tweet* mahasiswa ke bentuk baku, dengan memanfaatkan kamus eksternal *kamus\_katabaku.xlsx*, yang divisualisasikan pada Tabel 3. Tahap ini dimulai dengan pembacaan kamus kata baku melalui *library Pandas*, yang dikonversi menjadi *dictionary* dengan pasangan *key-value* (kata tidak baku - kata baku). Fungsi *normalized\_term()* diterapkan pada data hasil *tokenizing* untuk mengganti setiap kata tidak baku sesuai entri kamus.

Misalnya, kata “ga” dinormalisasi menjadi “tidak”, untuk meningkatkan kesesuaian dengan analisis berbasis model bahasa.

Tabel 7. Hasil Keluaran Proses *Text Normalization*

id	Text Normalization
0	[hari, ini, kami, mahasiswa, ui, merasa, resah...
1	[damn, besok, kuliah]
3	[gue, rasa, kuliah, bukan, passion, gue, passi...
4	[lu, pada, stres, gasi, melihat, dk, casual, m...
5	[like, struggle, tapi, bayar, ukt, kuliah, but...
...	...
23164	[stres, banget, masih, pagi, ini, ajg]
23165	[kemarin, pas, scrolling, linkedin, kebetulan,...

Hasil normalisasi disimpan ke dalam kolom baru sebagai *Text Normalization*, yang divisualisasikan pada Tabel 7. Implementasi tahap ini efektif dalam menstandarisasi variasi penulisan informal seperti singkatan atau *typo*, yang umum ditemui di media sosial. Untuk meminimalkan *noise* sebelum analisis lebih lanjut, tahap normalisasi dinilai penting dalam meningkatkan akurasi deteksi pola linguistik terkait gejala depresi. Mengingat konsistensi kosakata berpengaruh signifikan terhadap validitas interpretasi fitur teks.

Tabel 8. Hasil Keluaran Proses *Text Stemming*

id	Text Stemming
0	[hari, ini, kami, mahasiswa, ui, rasa, resah, ...
1	[damn, besok, kuliah]
3	[gue, rasa, kuliah, bukan, passion, gue, passi...
4	[lu, pada, stres, gas, lihat, dk, casual, mode...
5	[like, struggle, tapi, bayar, ukt, kuliah, but...
...	...
23164	[stres, banget, masih, pagi, ini, ajg]
23165	[kemarin, pas, scrolling, linkedin, betul, nem...

Proses *stemming* menghasilkan teks terstandarisasi yang disimpan ke dalam kolom *Text Stemming*, seperti divisualisasikan pada Tabel 8. Tahapan ini dilakukan menggunakan *library Sastrawi* untuk bahasa Indonesia, yang didukung dengan *swifter* untuk optimasi komputasi. Tahap awal melibatkan inialisasi *stemmer* melalui *StemmerFactory*, untuk menghasilkan objek yang mampu mengidentifikasi kata dasar. Setiap kata unik dari kolom *Text Normalization* diekstraksi ke dalam kamus *term\_dict* dan diproses menggunakan teknik *stemming* untuk menyederhanakan variasi kata ke bentuk dasar. Misalnya, kata “merasakan” atau “tertekan” direduksi menjadi “rasa” dan “tekan”, guna mengurangi redundansi leksikal. Hasil *stemming* diterapkan pada seluruh dataset melalui fungsi *get\_stemmed\_term*, sehingga menghasilkan kolom *Text Stemming* yang berisikan teks terstandarisasi.

Tabel 9. Hasil Keluaran Proses *Text Filtering*

id	Text Filtering
0	[mahasiswa, ui, resah, kondisi, bangsa, akhira...
1	[damn, besok, kuliah]
3	[gue, kuliah, passion, gue, passion, gue, baca...
4	[lu, stres, gas, lihat, dk, casual, mode, ceo,...
5	[like, struggle, bayar, ukt, kuliah, but, im, ...
...	...

23164	[stres, banget, pagi, ajg]
23165	[kemarin, pas, scrolling, linkedin, nemu, post...

Divisualisasikan pada Tabel 9, hasil proses *filtering* menunjukkan reduksi signifikan pada redundansi kata, dengan *output* berupa token tematik yang lebih merepresentasikan konteks emosional. Tahapan ini dilakukan untuk menghilangkan kata tak bermakna (*stop words*) dari korpus teks, sehingga fokus analisis tertuju pada pada *leksikon* inti yang mengindikasikan gejala depresi. Tahap ini mengintegrasikan tiga sumber *stop words*, yakni daftar bawaan NLTK untuk bahasa Indonesia, penambahan manual kosakata slang seperti “yg”, “dg”, “dgn”, “ny”, “kalo”, dan “krn” serta ekstraksi dari file eksternal *stopwordbahasa.txt*. Seluruh kata tersebut dikonversi ke bentuk set guna mengoptimasi kecepatan pencarian selama proses eliminasi. Implementasi fungsi *stopwords\_removal* menggunakan pendekatan *list comprehension* berhasil menghasilkan kolom *Text Filtering* yang memuat token teks terdokumentasi tanpa *noise* linguistik. Proses ini melibatkan eliminasi partikel non-esensial yang memungkinkan algoritma mengidentifikasi *leksikon* bernilai diagnostik secara lebih akurat.

Tabel 10. Data *Leksikon Positive*

No	Word	Weight
1	hai	3
2	kasih sayang	5
3	ekstensif	3
4	paripurna	1
5	detail	2
...	...	...
3608	suka cita	4
3609	filantropi	2

Tabel 11. Data *Leksikon Negative*

No	Word	Weight
1	putus tali gantung	-2
2	gelebah	-2
3	gobar hati	-2
4	sesal	-4
5	isak	-5
...	...	...
6608	stres	-4
6609	senewen	-4

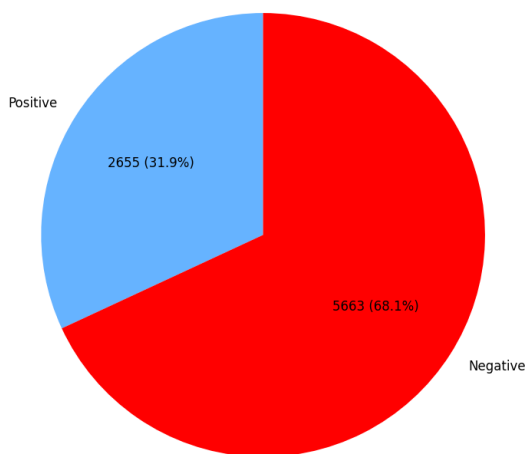
Proses *labelling* dalam penelitian ini mengadopsi *Inset Lexicon* sebagai dasar analisis sentimen berbahasa Indonesia. Tahap diawali dengan pemuatan dua kamus, *Lexicon Positive* divisualisasikan pada Tabel 10 dan *Lexicon Negative* pada Tabel 11, yang dikonversi menjadi *dictionary* untuk memudahkan pencarian kata. Setiap kata dalam *tweet* yang telah melalui tahap *pra-pemrosesan* diiterasi dan dicocokkan dengan entri pada kedua kamus. Skor polaritas dihitung melalui akumulasi nilai sentimen kata. Jika sebuah kata terdapat dalam kamus positif, nilai skor akan bertambah. Sebaliknya, kata yang termasuk dalam kamus negatif akan mengurangi skor sebesar bobot yang telah ditetapkan. Label sentimen ditentukan berdasarkan skor akhir dengan kriteria “*Positive*” untuk skor yang bernilai nol atau lebih ( $\geq 0$ ), sementara “*Negative*” untuk skor yang bernilai di bawah nol ( $< 0$ ).

Tabel 12. Hasil Keluaran Proses *Labelling*

id	Text Filtering	Polarity Score	Indonesia Sentiment
0	[mahasiswa, ui, resah, kondisi, bangsa, akhira...]	-7	Negative
1	[damn, besok, kuliah]	-1	Negative
3	[gue, kuliah, passion, gue, passion, gue, baca...]	1	Positive
4	[lu, stres, gas, lihat, dk, casual, mode, ceo,...]	-6	Negative
5	[like, struggle, bayar, ukt, kuliah, but, im, ...]	-6	Negative
...	...	...	...
23164	[stres, banget, pagi, ajg]	-7	Negative
24166	[kemarin, pas, scrolling, linkedin, nemu, post...]	-1	Negative

Tabel 12 memvisualisasikan hasil implementasi *labelling*, yang menghasilkan dua variabel baru, yaitu *Polarity Score* yang memuat nilai numerik dan *Indonesia Sentiment* sebagai label kategori. Distribusi label yang terbentuk menunjukkan kecenderungan sentimen mahasiswa di media X, seperti kata-kata yang mengindikasikan tekanan atau keputusan, dimana hasil ini menjadi dasar identifikasi pola linguistik terkait indikasi gejala depresi.

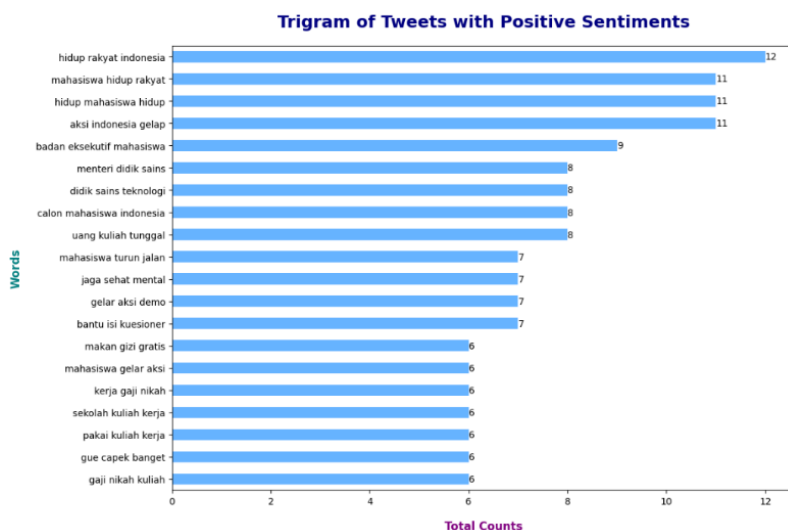
Distribution of Tweets Sentiment



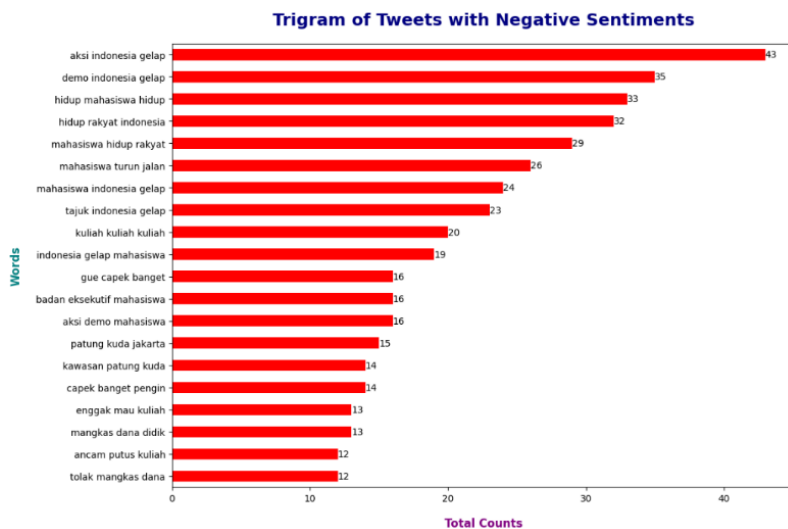
Gambar 4. Hasil Visualisasi Distribusi Sentimen *Tweet* dengan *Pie Chart*

Ditunjukkan pada Gambar 4, distribusi sentimen divisualisasikan melalui diagram lingkaran (*pie chart*), yang dibangun dengan algoritma statistik deskriptif untuk merepresentasikan proporsi kelas sentimen. Proses diawali dengan penghitungan frekuensi kedua sentimen menggunakan metode *value\_counts()*. Berdasarkan analisis terhadap 8.318 data teks, hasil menunjukkan dominasi signifikan sentimen negatif sebesar 68,1% (5.663 *tweet*), atau hampir dua kali lipat lebih tinggi dibandingkan sentimen positif yang hanya mencapai 31,9% (2.655 *tweet*). Penggunaan warna kontras (merah) untuk negatif dan penempatan persentase secara eksplisit memperjelas interpretasi distribusi yang tidak seimbang. Ketimpangan ini mengindikasikan adanya kecenderungan mahasiswa dalam menggunakan

platform X sebagai media untuk meluapkan emosi negatif. Analisis *N-gram* pada *tweet* mahasiswa dilakukan untuk mengidentifikasi pola linguistik berkaitan dengan gejala depresi. Analisis ini bertujuan untuk memetakan frasa-frasa kontekstual yang sering kali tidak terdeteksi melalui analisis kata tunggal. Kemampuan trigram dalam menangkap konteks kalimat secara utuh dapat mengurangi risiko ambiguitas dalam memaknai ekspresi emosional yang terkandung dalam *tweet*. Proses analisis *N-gram* diimplementasikan dengan memanfaatkan *library* NLTK untuk menyaring *stopwords* bahasa Indonesia. *CountVectorizer* dari *library* *scikit-learn* digunakan untuk menghitung frekuensi kemunculan trigram. Tahapan analisis mencakup pembersihan teks dari *stopwords*, transformasi korpus *tweet* menjadi representasi numerik berbasis trigram, penghitungan frekuensi kemunculan trigram, serta pengurutan dan visualisasi 20 trigram dominan pada kelas sentimen positif dan negatif.



Gambar 5. Hasil Visualisasi Trigram *Tweet* dengan Sentimen Positif



Gambar 6. Hasil Visualisasi Trigram *Tweet* dengan Sentimen Negatif

Hasil analisis pada Gambar 3 dan 4 menunjukkan perbedaan pola linguistik antara kedua sentimen. Pada sentimen positif, trigram seperti “hidup rakyat indonesia” dan “gelar aksi demo” mengindikasikan bentuk dukungan sosial. Sementara analisis trigram negatif mengungkap dominasi frasa bermuatan ketidakpuasan sosial seperti “aksi indonesia gelap”, “hidup mahasiswa hidup”, dan

“tolak mangkas dana”. Disusul dengan frasa seperti “gue capek banget”, “capek banget pengen”, “enggak mau kuliah”, dan “ancam putus kuliah” yang mengarah pada indikasi ketertekanan dan kelelahan mental sebagai gejala awal depresi. Frekuensi trigram negatif yang lebih tinggi menunjukkan kecenderungan mahasiswa lebih vokal dalam mengekspresikan tekanan psikologis melalui metafora atau kritik sosial. Temuan ini memperkuat hipotesis bahwa analisis trigram pada media sosial mampu mengungkap gejala depresi yang tersirat dalam ekspresi linguistik, terutama terkait dengan ungkapan kelelahan, protes, dan perasaan terisolasi. Pemisahan dataset (*data splitting*) dalam penelitian ini menggunakan algoritma *stratified random sampling* dengan rasio 80:20, di mana 80% data dialokasikan sebagai data latih (*training data*) dan 20% sebagai data uji (*testing data*) untuk mempertahankan distribusi kelas asli pada subset data. Teknik ini dipilih untuk mencegah bias akibat ketidakseimbangan kelas antar sentimen. Parameter *random\_state=0* berperan dalam menjamin konsistensi pembagian data yang dapat direproduksi. Proses pembagian menghasilkan 6.654 sampel data latih dan 1.664 data uji, dengan validasi distribusi menggunakan *value\_counts* untuk menunjukkan proporsi identik antara data latih yang memuat 68% negatif dan 32% positif, serta data uji dengan 68% negatif dan 32% positif. Pendekatan ini meminimalkan risiko overfitting sekaligus memastikan model teruji pada data yang merepresentasikan populasi asli.

#### Kode Program 1. *Feature Extraction* dengan TF-IDF

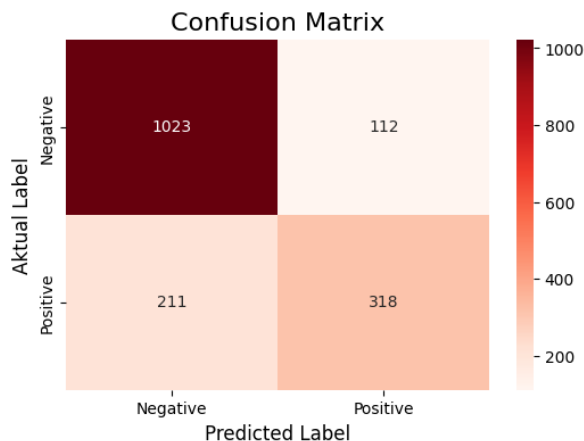
```
from nltk.corpus import stopwords
from sklearn.feature_extraction.text import TfidfVectorizer

tfidf_vectorizer = TfidfVectorizer(max_features=2500, min_df=7, max_df=0.8,
stop_words=stopwords.words('indonesian'))
response = tfidf_vectorizer.fit_transform(X_train)
```

Kode program 1 merupakan bagian krusial dari proses ekstraksi fitur menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) pada kerangka *Natural Language Processing* (NLP), untuk mengubah data teks mentah dari *tweet* menjadi representasi numerik yang dapat diproses oleh algoritma *machine learning*. Tahapan diawali dengan konfigurasi *TfidfVectorizer* menggunakan parameter seperti pembatasan maksimum 2.500 fitur dominan (*max\_features*), penghapusan kata yang muncul di kurang dari tujuh dokumen (*min\_df=7*), eliminasi kata yang terdapat pada lebih dari 80% dokumen (*max\_df=0.8*), serta penyaringan *stopwords* bahasa Indonesia untuk meminimalkan gangguan semantik. Proses ini menghasilkan matriks TF-IDF yang merepresentasikan bobot penting setiap kata dalam korpus, di mana nilai setiap elemen matriks mencerminkan signifikansi kata tersebut baik dalam konteks dokumen tunggal maupun keseluruhan dataset. Matriks TF-IDF dikonversi ke bentuk *DataFrame* untuk mempermudah analisis eksplorasi dan identifikasi kata kunci yang mengarah pada indikator depresi, seperti ekspresi emosional.

Transformasi serupa diaplikasikan pada data uji menggunakan model vektorizer yang sama, guna memastikan konsistensi input sebelum tahap pemodelan. Pendekatan TF-IDF terbukti efektif dalam menyederhanakan kompleksitas teks dengan mempertahankan kata-kata informatif sekaligus meningkatkan performa algoritma klasifikasi seperti *Random Forest*. Melalui teknik ini, model mampu mempelajari pola linguistik spesifik dalam *tweet*, seperti ungkapan kesedihan, kecemasan, atau keputusan, yang menjadi dasar identifikasi gejala depresi. Dengan kata lain, integrasi TF-IDF tidak hanya mengonversi data tekstual menjadi struktur terukur, tetapi juga memperkaya akurasi deteksi melalui analisis kontekstual yang mendalam terhadap dinamika bahasa informal di media sosial. Proses pelatihan model dalam penelitian ini menggunakan algoritma *Random Forest* dari *library sklearn.ensemble* dengan parameter *n\_estimators=200*, *criterion='entropy'*, dan *random\_state=0* untuk memastikan konsistensi hasil. Data latih berupa vektor TF-IDF (*X\_train*) dan label sentimen (*y\_train*) digunakan untuk membangun model klasifikasi. Hasil evaluasi awal menunjukkan akurasi sebesar 80,59% dengan *confusion matrix* mengidentifikasi 318 *True Positive* (*tweet* depresi terprediksi benar) dan 1.023 *True Negative* (*tweet* non-depresi terprediksi benar). Namun, terdapat 211 *False Negative* yang mengindikasikan kegagalan model dalam mendeteksi 40% kasus depresi, diduga akibat

ketidakseimbangan data. *Precision* kelas *Positive* mencapai 74%, tetapi *recall* yang rendah (60%) menegaskan perlunya optimasi untuk meningkatkan sensitivitas terhadap kelas minoritas.



Gambar 7. Visualisasi *Confusion Matrix*

Gambar 7 menampilkan visualisasi *confusion matrix* untuk mengevaluasi kinerja model dalam mengklasifikasikan *tweet* depresi (*Positive*) dan non-depresi (*Negative*). Matriks ini dibangun menggunakan fungsi `sklearn.metrics.confusion_matrix`, yang mengelompokkan hasil prediksi ke dalam empat kategori, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Berdasarkan kategori tersebut, metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score* dihitung untuk mengukur efektivitas model dalam mendeteksi pola linguistik terkait gejala depresi. Analisis ini bertujuan untuk mengidentifikasi kekuatan dan kelemahan model, khususnya dalam konteks ketidakseimbangan data dan prioritas deteksi kasus depresi. Metrik evaluasi dihitung sebagai berikut.

- 1) *Accuracy* menggambarkan sejauh mana prediksi model sesuai dengan data aktual. Akurasi dihitung sebagai rasio total prediksi benar, terhadap keseluruhan data.

$$Accuracy = \frac{TP + TN}{(TP + FP + FN + TN)} = \frac{318 + 1023}{(318 + 112 + 211 + 1023)} = \frac{1.341}{1.664} \approx 0,8059$$

Hasil evaluasi menunjukkan bahwa model mampu mengklasifikasikan 80.59% *tweet* dengan benar. Hasil akurasi diperoleh dari perbandingan antara jumlah prediksi benar (*True Positive + True Negative*) dengan total data yang diuji. Namun, tingginya akurasi ini lebih disebabkan oleh banyaknya data kategori *Negative* yang berhasil diprediksi dengan tepat, yaitu mencapai 1.023 kasus. Sehingga sensitivitas model dalam mendeteksi kelas *Positive* masih tergolong kurang optimal.

- 2) *Precision* mengukur ketepatan model dalam memprediksi suatu kelas tertentu, dengan membandingkan jumlah prediksi positif yang benar (*True Positive*) dengan total prediksi positif yang dibuat (*True Positive + False Positive*).

$$Precision_{Positive} = \frac{TP}{(TP + FP)} = \frac{318}{(318 + 112)} = \frac{318}{430} \approx 0,74$$

Model menghasilkan nilai *precision* sebesar 74% untuk kelas *Positive*. Sehingga, dari seluruh *tweet* yang diprediksi sebagai *Positive*, hanya 74% yang benar-benar positif. Sisanya, 26% merupakan ketidaksesuaian prediksi.

$$Precision_{Negative} = \frac{TN}{(TN + FN)} = \frac{1.023}{(1.023 + 211)} = \frac{1.023}{1.234} \approx 0,83$$

Sedangkan untuk kelas *Negative*, nilai *precision* mencapai 83%, menunjukkan bahwa model cukup akurat dalam mengidentifikasi *tweet* non-depresi, meskipun terdapat 17% ketidaksesuaian.

- 3) *Recall* (Sensitivitas) menilai kemampuan model dalam menangkap seluruh data positif yang sebenarnya ada. Nilai dihitung sebagai rasio prediksi positif benar (*True Positive*) terhadap total data yang sebenarnya positif (*True Positive + False Negative*).

$$Recall_{Positive} = \frac{TP}{(TP + FN)} = \frac{318}{(318 + 211)} = \frac{318}{529} \approx 0,60$$

Nilai *recall* untuk kelas *Positive* mengindikasikan bahwa model hanya berhasil mendeteksi 60% dari *tweet* depresi, sementara 40% lainnya tidak teridentifikasi.

$$Recall_{Negative} = \frac{TN}{(TN + FP)} = \frac{1.023}{(1.023 + 112)} = \frac{1.023}{1.135} \approx 0,90$$

Sebaliknya, nilai *recall* untuk kelas *Negative* mencapai 90%, yang menandakan bahwa model mampu mengenali *tweet* non-depresi dengan baik, meskipun terdapat 10% ketidaksesuaian.

- 4) *F1-Score* merupakan ukuran yang menggabungkan presisi dan *recall* secara harmonis, sehingga memberikan gambaran keseimbangan antara ketepatan prediksi (*precision*) dan cakupan deteksi (*recall*).

$$F1 - Score_{Positive} = 2 \times \frac{Precision_{Pos} \times Recall_{Pos}}{Precision_{Pos} + Recall_{Pos}} = 2 \times \frac{0.74 \times 0.60}{0.74 + 0.60} = \frac{0.888}{1.34} \approx 0,66$$

Perhitungan *F1-Score* menghasilkan nilai 66% untuk kelas *Positive*, menunjukkan kompromi antara *presisi* dan *recall* yang kurang optimal untuk mendeteksi *tweet* depresi.

$$F1 - Score_{Negative} = 2 \times \frac{Precision_{Neg} \times Recall_{Neg}}{Precision_{Neg} + Recall_{Neg}} = 2 \times \frac{0.83 \times 0.90}{0.83 + 0.90} = \frac{1.494}{1.73} \approx 0,86$$

Sementara untuk kelas *Negative*, nilai *F1-Score* mencapai 86%, menandakan performa yang lebih stabil dalam mengklasifikasikan *tweet* non-depresi.

Evaluasi model menggunakan *classification report* mengungkap kinerja lebih baik pada kelas *Negative* dibandingkan kelas *Positive*. Hal ini disebabkan adanya ketidakseimbangan data yang didominasi oleh *tweet* bermuatan depresi. Untuk mencapai akurasi tertinggi pada data yang diberikan, dilakukan *hyperparameter tuning* pada model *Logistic Regression* menggunakan teknik *Grid Search*, untuk mencari kombinasi terbaik dari parameter *C* dan *penalty*. Hasil *tuning* menunjukkan peningkatan akurasi terbaik mencapai 87,51% yang artinya meningkat 6,92% dibandingkan model awal. Temuan ini membuktikan bahwa proses *hyperparameter tuning* dapat mengoptimalkan kemampuan model dalam mengklasifikasikan *tweet* secara lebih tepat.

### 3.2 Pembahasan

Penelitian ini bertujuan untuk menganalisis *tweet* mahasiswa guna mendeteksi gejala depresi dengan menggunakan teknik Pemrosesan Bahasa Alami (*Natural Language Processing/NLP*) melalui pemanfaatan model *Random Forest* dan fitur *TF-IDF*. Berdasarkan hasil analisis terhadap 8.318 *tweet* yang telah diproses, ditemukan bahwa 68,1% *tweet* berisi sentimen negatif, yang menunjukkan kecenderungan mahasiswa untuk menggunakan platform X (*Twitter*) sebagai saluran untuk

mengekspresikan emosi negatif terkait kehidupan akademik dan sosial mereka. Fenomena ini sejalan dengan temuan yang diungkapkan oleh Cahyono (2021) yang menunjukkan peran penting media sosial dalam membentuk perubahan sosial, termasuk dalam hal ekspresi perasaan pribadi dan kesejahteraan mental. Seiring dengan meningkatnya perhatian terhadap masalah kesehatan mental di kalangan mahasiswa, khususnya depresi yang sering dipicu oleh tekanan akademik, sosial, dan personal, penggunaan media sosial menjadi salah satu cara bagi individu untuk mengungkapkan perasaan mereka (Distina, 2019). Dalam penelitian ini, model Random Forest yang dipadukan dengan ekstraksi fitur menggunakan TF-IDF menunjukkan akurasi 87,51% dalam mendeteksi sentimen negatif dan positif pada tweet mahasiswa, yang menunjukkan kemampuannya dalam mengatasi ketidakseimbangan data, terutama dominasi sentimen negatif. Temuan ini mendukung penelitian sebelumnya yang menunjukkan keunggulan model Random Forest dalam menangani data tidak seimbang (Rahayu *et al.*, 2023). Selain itu, analisis trigram pada tweet mahasiswa mengungkapkan pola linguistik yang mengindikasikan gejala depresi, seperti frasa “capek banget” dan “enggak mau kuliah,” yang sering muncul dalam komunikasi mahasiswa yang merasa tertekan. Hal ini sejalan dengan temuan dari Mutmainah (2022), yang menunjukkan bahwa analisis sentimen berbasis NLP dapat membantu mengidentifikasi gejala depresi pada individu yang mengungkapkan perasaan mereka melalui media sosial. N-gram, khususnya trigram, terbukti efektif dalam menangkap hubungan semantik antar kata yang sering kali tidak terdeteksi oleh unigram atau bigram (Zhu *et al.*, 2022), yang sangat penting dalam memahami konteks emosional dalam ekspresi mahasiswa.

Penting untuk dicatat bahwa salah satu keterbatasan utama dalam penelitian ini adalah ketergantungan pada leksikon yang terbatas dalam menangani kosakata informal atau slang yang digunakan oleh mahasiswa. Penggunaan bahasa tidak baku ini seringkali menjadi tantangan dalam analisis sentimen pada media sosial (Fachriza & Artikel, 2023). Oleh karena itu, pengembangan leksikon yang lebih inklusif, yang mencakup berbagai variasi bahasa informal yang umum digunakan oleh mahasiswa, akan meningkatkan akurasi dalam mendeteksi gejala depresi secara lebih tepat. Penelitian lanjutan yang melibatkan kolaborasi dengan ahli psikologi dan pakar bahasa juga dapat memperkaya pemahaman terhadap ekspresi emosional yang lebih relevan dengan konteks kesehatan mental (Putra *et al.*, 2023). Penelitian ini menunjukkan bahwa teknik Pemrosesan Bahasa Alami (NLP), khususnya yang diterapkan pada analisis tweet mahasiswa, dapat menjadi alat yang efektif dalam mendeteksi gejala depresi. Penggunaan model Random Forest yang diintegrasikan dengan TF-IDF memungkinkan analisis yang lebih akurat meskipun menghadapi tantangan ketidakseimbangan data dan penggunaan bahasa informal. Implikasi praktis dari temuan ini adalah bahwa model ini dapat digunakan untuk membangun sistem deteksi dini gejala depresi mahasiswa yang proaktif, yang akan bermanfaat bagi lembaga pendidikan dan layanan kesehatan mental untuk memberikan intervensi lebih cepat dan tepat.

#### 4. Kesimpulan

Penelitian ini berhasil menganalisis ekspresi linguistik pada *tweet* mahasiswa untuk mendeteksi gejala depresi melalui pendekatan *Natural Language Processing* (NLP). Hasil analisis terhadap 8.318 *tweet* menunjukkan distribusi sentimen di dominasi oleh sentimen negatif sebesar 68,1% (5.663 *tweet*), dibandingkan sentimen positif 31,9% (2.655 *tweet*). Analisis lebih lanjut terhadap trigram sentiment negatif mengungkap frasa seperti “gue capek banget”, “capek banget pengen”, “enggak mau kuliah”, dan “ancam putus kuliah” mengarah pada indikasi ketertekanan dan kelelahan mental sebagai gejala awal depresi. Sehingga temuan ini mencerminkan kecenderungan penggunaan platform X (Twitter) oleh mahasiswa sebagai sarana untuk meluapkan emosi negatif. Model *Random Forest* yang diintegrasikan dengan ekstraksi fitur TF-IDF mencapai akurasi 87,51% setelah optimasi *hyperparameter*, membuktikan kemampuannya dalam mengatasi bias kelas mayoritas (negatif) dan menangkap kompleksitas morfologi bahasa informal. Penelitian ini memperkuat dasar pengembangan sistem deteksi dini yang adaptif terhadap dinamika bahasa pada *tweet* mahasiswa.

Penelitian ini dapat memberikan kontribusi praktis bagi lembaga pendidikan maupun layanan kesehatan mental untuk merekomendasikan integrasi model *Random Forest* berbasis NLP ke dalam sistem pemantauan digital deteksi gejala depresi mahasiswa secara proaktif. Namun, penelitian ini memiliki keterbatasan, seperti cakupan kamus *Inset Lexicon* yang masih terbatas dalam memuat kosakata informal (slang) terkait kesehatan mental, yang kerap digunakan dalam percakapan mahasiswa. Guna meningkatkan akurasi analisis, penelitian selanjutnya perlu melibatkan kolaborasi multidisiplin dengan ahli psikologi dan pakar bahasa dalam menyusun korpus *leksikon* khusus yang mencakup kosakata slang serta ekspresi emosional yang lebih relevan. Hal ini bertujuan untuk menciptakan basis data yang lebih adaptif terhadap dinamika bahasa informal dan konteks kesehatan mental kontemporer, sehingga model deteksi depresi dapat mengidentifikasi pola linguistik secara lebih holistik dan kontekstual.

## 5. Ucapan Terima Kasih

Artikel ini disusun sebagai tugas akhir dari penyelesaian Program Studi Strata I Teknik Informatika di Universitas Kristen Satya Wacana. Penulis ingin menyampaikan rasa terima kasih kepada semua pihak yang telah terlibat dalam penyelesaian penelitian ini. Kepada dosen pembimbing yang telah memberikan saran, bimbingan, dan motivasi berharga sejak tahap perencanaan hingga penyusunan artikel. Juga kepada tim reviewer yang telah memberikan masukan selama proses revisi. Terakhir, apresiasi disampaikan kepada keluarga dan sahabat yang senantiasa mendampingi dan memberikan segala bentuk dukungan moral, spiritual, serta keterlibatannya dalam diskusi konstruktif yang memperkaya kualitas penelitian ini. Semoga karya ini dapat bermanfaat bagi pengembangan ilmu pengetahuan dan masyarakat luas.

## 6. Daftar Pustaka

- Angga, S., Poa, A. A., & Rikardus, F. (2023). Etika Komunikasi Netizen Indonesia di Media Sosial sebagai Ruang Demokrasi dalam Telaah Ruang Publik Jürgen Habermas. *Jurnal Filsafat Indonesia*, 6(3), 384-393. <https://doi.org/10.23887/jfi.v6i3.59229>.
- Apriliani, F., & Maharani, W. (2023). Depression Detection on Social Media Twitter using XLNet Method. *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 8(1), 172-180. <https://doi.org/10.29100/jupi.v8i1.3345>.
- Cahyono, A. S. (2016). Pengaruh media sosial terhadap perubahan sosial masyarakat di Indonesia. *Publiciana*, 9(1), 140-157.
- Distina, P. P., & Kumail, M. H. (2019). Pengembangan dimensi psychological well-being untuk pengurangan risiko gangguan depresi. *Jurnal dakwah dan pengembangan sosial kemanusiaan*, 10(1), 39-59.
- Erma, H. N., Santika, D., Nurhasanah, I., & Lestari, I. (2022). Dampak Media Sosial Terhadap Konflik Di Masyarakat. *EDUSOSHUM: Journal of Islamic Education and Social Humanities*, 2(1), 12-20. <https://doi.org/10.52366/edusoshum.v2i1.33>.
- Fachrizza, M., & Munawar, M. (2023). Analisis Sentimen Kalimat Depresi Pada Pengguna Twitter Dengan Naive Bayes, Support Vector Machine, Random Forest. *KOMPUTEK*, 7(2), 49-58.

- Fahma, F., & Safitri, D. (2024). Dinamika Identitas Budaya dalam Era Globalisasi: Tantangan dan Kesempatan Media Sosial terhadap Budaya Masyarakat Lokal: Dynamics of Cultural Identity in the Era of Globalization: Challenges and Opportunities for Social Media on Local Community Culture. *Jurnal Intelek Dan Cendekiawan Nusantara*, 1(3), 3675-3682.
- Gifari, O. I., Adha, M., Freddy, F., & Durrand, F. F. S. (2022). Film Review Sentiment Analysis Using TF-IDF and Support Vector Machine. *Journal of Information Technology*, 2(1), 36-40. <https://doi.org/10.46229/jifotech.v2i1.330>.
- Musfiroh, D., Khaira, U., Utomo, P. E. P., & Suratno, T. (2021). Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon: Sentiment Analysis of Online Lectures in Indonesia from Twitter Dataset Using InSet Lexicon. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1(1), 24-33. <https://doi.org/10.57152/malcom.v1i1.20>.
- Mutmainah, S. (2022). Kemungkinan Depresi Dari Postingan Pada Sosial Media. *Jurnal Sains, Nalar, dan Aplikasi Teknologi Informasi*, 1(2), 82-88. <https://doi.org/10.20885/snati.v1i2.11>.
- Putra, D. E., Nelwati, N., & Fernandes, F. (2023). Hubungan depresi, stres akademik dan regulasi emosi dengan ide bunuh diri pada mahasiswa. *Jurnal Keperawatan Jiwa*, 11(3), 689-706.
- Rahayu, K., Fitria, V., Septhya, D., Rahmadden, R., & Efrizoni, L. (2023). Klasifikasi Teks untuk Mendeteksi Depresi dan Kecemasan pada Pengguna Twitter Berbasis Machine Learning: Text Classification for Detecting Depression and Anxiety among Twitter Users based on Machine Learning. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 3(2), 108-114. <https://doi.org/10.57152/malcom.v3i2.780>.
- Rizqi, B. R. T., & Heriyanto, H. (2023). Penyebaran Informasi melalui Thread Berita di Twitter oleh Mahasiswa S-1 Program Studi Ilmu Perpustakaan Universitas Diponegoro. *Anuva: Jurnal Kajian Budaya, Perpustakaan, dan Informasi*, 7(3), 515-528. <https://doi.org/10.14710/anuva.7.3.515-528>.
- Rusly, N. F., Qoni'ah Nur Wijayanti, S. I., & Ikom, M. (2023). PERILAKU SELF DISCLOSURE PADA KALANGAN REMAJA MELALUI MEDIA SOSIAL TWITTER. *Jurnal Media Akademik (JMA)*, 1(1).
- Santoso, N. N., Maulita, Y., & Khair, H. (2022). Sistem Pakar Diagnosa Tingkat Depresi Mahasiswa Tingkat Akhir Dengan Menggunakan Metode Certainty Factor. *Bull. Multi*, 1, 732-740.
- Tamaraya, A., & Ubaedullah, D. (2021). Dampak Penggunaan Twitter Terhadap Pengungkapan Diri Mahasiswa. *INTERAKSI PERADABAN: Jurnal Komunikasi dan Penyiaran Islam*, 1(1).
- Zaskya, M., Boham, A., & Lotulung, L. J. H. (2021). Twitter sebagai media mengungkapkan diri pada kalangan milenial. *Acta Diurna Komunikasi*, 3(1).
- Zhu, E., Zhang, J., Yan, J., Chen, K., & Gao, C. (2022). N-gram MalGAN: Evading machine learning detection via feature n-gram. *Digital communications and networks*, 8(4), 485-491.