

**LAPORAN AKHIR
RISET KOMPETITIF MAHASISWA**

**ANALISIS SENTIMEN ECO-FRIENDLY PRODUCTS PADA E-
COMMERCE DENGAN ALGORITMA MULTI-LAYER PRECEPTON**

DISUSUN OLEH:

MUHAMMAD KRISTOVER ARMAND / 210601110051

CETRIN APRILIA / 210601110004

SAILA FARIZQIYAH / 220601110042

DOSEN PEMBIMBING

ARI KUSUMASTUTI, M.Si., M.Pd

NIP. 19770521 200501 2 004



**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG**

2025

LEMBAR PENGESAHAN LAPORAN HASIL PENELITIAN

Judul : Analisis Sentimen Eco-friendly Products pada E-Commerce dengan Algoritma Multi-Layer Preceptron
Nama Ketua Peneliti : Muhammad Kristover Armand (210601110051)
Nama Anggota Peneliti : 1. Cetrin Aprilia (210601110004)
2. Saila Farizqiyah (220601110042)
Prodi : Matematika
Lama Kegiatan : 17 April - 10 Juli 2025
Biaya : Rp 1.500.000

Malang, 23 Juni 2025
Ketua Peneliti

Muhammad Kristover Armand
210601110051

Mengetahui
Ketua Prodi Matematika

Mengetahui
Dosen Pembimbing

Dr. Elly Susanti, M.Sc
NIP. 19741129 200012 2 005

Ari Kusumastuti, M.Pd., M.Si
NIP. 19770521 200501 2 004

a.n. Dekan,
Wakil Dekan Bidang Kemahasiswaan & Kerjasama

Dr.Dwi Suheriyanto, M.P

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	ii
DAFTAR ISI.....	iii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	4
BAB II STUDI PUSTAKA	5
2.1 Multi-Layer Preception	5
2.2 Tf-Idf Vectorizer	6
2.3 Smote Oversampling	6
2.4 Confusion Matrix	7
2.5 Natural Language Processing.....	8
2.6 ECO Friendly	9
2.7 Produk dalam E-Commerce.	9
BAB III METODE PENELITIAN	11
3.1 Jenis Penelitian	11
3.2 Data dan Sumber Data.....	11
3.3 Teknik Analisis Data	12
BAB IV PEMBAHASAN.....	13
4.1 Pengumpulan Data	13
4.2 Preprocessing	14
4.2.1 Normalisasi Kata Tidak Baku (<i>Slang Word Normalization</i>)	14
4.2.2 Penghapusan Stopword (<i>Stopword Removal</i>)	14
4.2.3 Penyeragaman Huruf (<i>Case Folding</i>)	15
4.2.4 Penghapusan Angka (<i>Number Removal</i>).....	15
4.2.5 Penghapusan Tanda Baca dan Simbol	15
4.2.6 Penghapusan Spasi Berlebih (<i>Whitespace Cleaning</i>).....	15
4.3 Labeling Lexicon.....	16
4.4 TF-IDF Vectorizer.....	17
4.5 SMOTE Oversampling.....	17
4.6 Label Encoder	17
4.7 Modeling MLP	18
4.8 Evaluasi	19
BAB V PENUTUP	21
5.1 Kesimpulan.....	21
5.2 Saran.....	22
DAFTAR PUSTAKA	23
LAMPIRAN.....	25

BAB I

PENDAHULUAN

1.1 Latar Belakang

Permasalahan pencemaran lingkungan akibat limbah plastik semakin mengkhawatirkan, terutama dengan meningkatnya konsumsi produk sekali pakai yang dijual melalui e-commerce. Berbagai merek mulai memasarkan produk dengan label eco-friendly sebagai respons terhadap meningkatnya kesadaran masyarakat akan keberlanjutan. Namun, tidak semua produk yang diklaim ramah lingkungan benar-benar memenuhi standar keberlanjutan. Praktik greenwashing, yaitu strategi pemasaran yang menyesatkan dengan memberikan kesan bahwa suatu produk lebih ramah lingkungan dari kenyataannya, masih marak terjadi. Hal ini menimbulkan kebingungan bagi konsumen yang ingin berkontribusi dalam mengurangi dampak negatif terhadap lingkungan.

Di sisi lain, e-commerce menyediakan ruang bagi konsumen untuk mengekspresikan pendapat mereka melalui ulasan produk. Ulasan ini dapat menjadi sumber informasi berharga untuk memahami bagaimana konsumen menilai keaslian klaim eco-friendly dari suatu produk. Dengan banyaknya data ulasan yang tersedia, teknik sentiment analysis atau analisis sentimen dapat digunakan untuk mengidentifikasi apakah mayoritas pelanggan memberikan respons positif atau negatif terhadap klaim keberlanjutan suatu produk. Dengan memahami sentimen pelanggan, dapat diketahui apakah sebuah produk benar-benar memiliki dampak lingkungan yang lebih baik atau sekadar memanfaatkan tren pasar.

Salah satu metode yang dapat diterapkan dalam analisis sentimen adalah Algoritma Multi-layer Perceptron (MLP). Algoritma ini, yang merupakan salah satu bentuk dari jaringan saraf tiruan (Artificial Neural Network), bekerja dengan mempelajari pola-pola kompleks dalam data melalui beberapa lapisan pemrosesan (layer). Dalam konteks analisis sentimen, MLP dapat digunakan untuk mengklasifikasikan ulasan pelanggan ke dalam kategori positif, negatif, atau netral dengan memodelkan hubungan non-linear antara kata-kata yang muncul dalam teks ulasan. Keunggulan MLP terletak pada kemampuannya untuk belajar dari data

secara mendalam dan menangkap nuansa sentimen yang mungkin terlewatkan oleh model yang lebih sederhana.

Penelitian ini bertujuan untuk menerapkan Algoritma Multi-layer Perceptron dalam analisis sentimen terhadap produk *eco-friendly* di *e-commerce*. Dengan pendekatan ini, diharapkan dapat diperoleh wawasan yang lebih mendalam mengenai bagaimana pelanggan menilai keaslian klaim *eco-friendly* suatu produk. Hasil penelitian ini tidak hanya membantu konsumen dalam membuat keputusan yang lebih tepat tetapi juga dapat menjadi masukan bagi produsen untuk lebih transparan dalam pemasaran produk ramah lingkungan. Selain itu, temuan dari penelitian ini juga dapat berkontribusi dalam mengurangi praktik *greenwashing* dan mendorong adopsi produk yang ramah lingkungan di pasar *e-commerce*.

Analisis sentimen terhadap produk ramah lingkungan memberikan wawasan mendalam tentang bagaimana konsumen memandang dan merespons inisiatif keberlanjutan. Dengan memahami sentimen ini, kita dapat melihat bahwa kecenderungan konsumen untuk memilih produk 'eco-friendly' bukan hanya tren sesaat, melainkan cerminan dari nilai-nilai yang semakin mengakar dalam masyarakat, termasuk pentingnya kebersihan, kebaikan, dan menjaga kelestarian bumi. Nilai-nilai ini, pada dasarnya, sangat selaras dengan ajaran agama yang mendorong kebaikan dan pelestarian lingkungan berdasarkan ayat berikut (Kemenag, 2025):

وَلَا تُفْسِدُوا فِي الْأَرْضِ بَعْدَ إِصْلَاحِهَا وَادْعُوهُ حَوْفًا وَقَطْمًا إِنَّ رَحْمَتَ اللَّهِ قَرِيبٌ مِّنَ الْمُحْسِنِينَ ﴿٥٦﴾

Artinya: Janganlah kamu berbuat kerusakan di bumi setelah diatur dengan baik. Berdoalah kepada-Nya dengan rasa takut dan penuh harap. Sesungguhnya rahmat Allah sangat dekat dengan orang-orang yang berbuat baik. (QS. Al-Araf: 56)

Berdasarkan tafsir dari kemenag, ditekankan agar tidak berbuat kerusakan di muka bumi setelah penciptaannya yang telah sempurna. Setiap individu dianjurkan untuk memanjatkan doa kepada-Nya dengan penuh rasa takut, yang akan membimbing menuju kekhusyukan dan ketaatan. Bersamaan dengan itu, doa juga harus disertai harapan yang tulus terhadap anugerah serta pengabulan-Nya. Sesungguhnya, rahmat Allah senantiasa dekat bagi mereka yang senantiasa berbuat kebajikan.

Machine learning adalah teknologi yang memungkinkan komputer menjalankan tugas secara mandiri. Dengan mengandalkan pembelajaran berkelanjutan dan peningkatan pengalaman, machine learning membantu komputer memahami tingkat kompleksitas masalah serta menyesuaikan diri dengan kebutuhan yang terus berkembang. Teknologi ini menjadi semakin penting dalam berbagai bidang, karena kemampuannya untuk menganalisis data dalam jumlah besar, mengenali pola, dan membuat keputusan secara otomatis, sehingga meningkatkan efisiensi dan akurasi dalam berbagai aplikasi (Alzubi, dkk., 2018).

Tabel 1.1 Penelitian Terdahulu

Metode	Data	Akurasi
KNN (Adhi, 2021)	Pengguna Aplikasi Bibit	85,14%
Naïve Bayes (Atimi dkk., 2022)	Data Review Produk di Lazada	90,45%
MLP (Makmur dkk., 2024)	Sentimen Penghapusan Skripsi sebagai Tugas Akhir Mahasiswa	94,00%

Berdasarkan tabel 1.1, Penelitian ini memiliki perbedaan dengan penelitian terdahulu dalam beberapa aspek utama. Adhi (2021) yang menggunakan algoritma KNN pada data pengguna pengguna aplikasi Bibit mencapai akurasi 85.14%. Atimi dkk. (2022) dengan algoritma Naïve Bayes pada data review Lazada. Makmur dkk., (2024) menganalisis Sentimen Penghapusan Skripsi sebagai Tugas Akhir Mahasiswa dengan akurasi 94,00% . Penelitian ini juga berkontribusi dengan membandingkan performa MLP dalam ulasan di aplikasi dan algoritma yang berbeda.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah bagaimana penerapan algoritma multi-layer precepton (MLP) dalam analisis sentimen terhadap eco-friendly products pada platform e-commerce?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dijabarkan, tujuan dari penelitian ini adalah menganalisis penerapan algoritma multi-layer preceptron (MLP) dalam klasifikasi sentimen terhadap eco-friendly products pada e-commerce.

1.4 Manfaat Penelitian

Berdasarkan tujuan penelitian yang telah dijabarkan, manfaat dari penelitian ini adalah diharapkan dapat memberikan wawasan lebih mendalam mengenai penerapan algoritma multi-layer preceptron (MLP) dalam analisis sentimen terhadap produk ramah lingkungan pada e-commerce. Selain itu, penelitian ini juga dapat menjadi referensi untuk analisis sentimen berbasis machine learning dalam industri e-commerce.

1.5 Batasan Masalah

Agar penelitian ini tetap terarah pada tujuan yang diinginkan, perlu ditetapkan beberapa pembatasan masalah sebagai berikut:

1. Data yang digunakan produk lock n lock di Tokopedia
2. Jumlah data mencapai 3000 review

BAB II

STUDI PUSTAKA

2.1 Multi-Layer Preceptron

Multi-layer Perceptron (MLP) adalah salah satu jenis jaringan saraf tiruan *feedforward* yang paling umum dan banyak digunakan. Jaringan ini tersusun dari setidaknya tiga lapisan node: satu lapisan input, satu atau lebih lapisan tersembunyi (hidden layers), dan satu lapisan output. Sesuai namanya, 'multi-layer' menunjukkan keberadaan lebih dari satu lapisan (selain lapisan input dan output), khususnya lapisan tersembunyi. MLP bekerja dengan memetakan serangkaian nilai input ke serangkaian nilai output yang sesuai. Proses pembelajaran dalam MLP melibatkan penyesuaian bobot (weights) dan bias melalui algoritma *backpropagation* untuk meminimalkan error antara output yang diprediksi dan output target. Kemampuan MLP untuk mempelajari hubungan non-linear yang kompleks antar data menjadikannya alat yang kuat untuk tugas-tugas seperti klasifikasi dan regresi.

Setelah Multi-layer Perceptron (MLP) berhasil dilatih, model siap digunakan untuk memprediksi kelas atau nilai output berdasarkan input baru. Proses prediksi ini dikenal sebagai *output generation* atau *forward propagation*. Pada tahap ini, input data dimasukkan ke dalam lapisan input jaringan, kemudian sinyal akan dipropagasikan maju melalui setiap neuron di lapisan tersembunyi, dan akhirnya mencapai lapisan output untuk menghasilkan prediksi akhir. Setiap neuron melakukan perhitungan jumlah tertimbang dari inputnya dan menerapkan fungsi aktivasi untuk menghasilkan output yang akan menjadi input bagi lapisan berikutnya (Kosarirad dkk., 2022).

Berikut merupakan perhitungan rumus hidden layer:

$$z_j = \sum_{i=1}^m (w_{ij} \cdot X_i) + b_j$$
$$a_j = f(z_j) = \frac{1}{1 + e^{-z_j}}$$

Berikut merupakan perhitungan Output Layer.

$$z_k = \sum_{j=1}^l (w_{jk} \cdot a_j) + b_k$$

$$\hat{y}_k = f(z_k) = \frac{1}{1 + e^{-z_k}}$$

2.2 Tf-Idf Vectorizer

TF-IDF (*Term Frequency-Inverse Document Frequency*) adalah sebuah ukuran statistik untuk mengevaluasi seberapa penting sebuah kata bagi suatu dokumen dalam sebuah koleksi atau korpus data. Metode ini bekerja dengan menggabungkan dua faktor: *Term Frequency* (TF), yang menghitung seberapa sering sebuah kata muncul di dalam dokumen tersebut, dan *Inverse Document Frequency* (IDF), yang mengukur seberapa unik kata itu di seluruh koleksi dokumen (Liang dkk., 2022).

$$TF - IDF_{ij} = \frac{n_{ij}}{d_j} \cdot \lg \left(\frac{D}{N_i} \right)$$

Keterangan:

n_{ij} : Jumlah kemunculan kata i di dalam dokumen j

d_j : Jumlah total semua kata di dalam dokumen j

D : Jumlah total dokumen dalam koleksi (corpus)

N_i : Jumlah dokumen yang mengandung kata i

2.3 Smote Oversampling

SMOTE (Synthetic Minority Over-sampling Technique) adalah metode *oversampling* yang berbeda dari pendekatan tradisional yang hanya menggandakan data kelas minoritas secara langsung. SMOTE menciptakan sampel baru secara sintetis dengan memanfaatkan interpolasi antara titik-titik data asli dalam kelas minoritas. Dengan cara ini, SMOTE tidak hanya menambah jumlah data kelas minoritas tetapi juga memperluas ruang data sehingga distribusi data menjadi lebih representatif dan beragam. Dengan menggunakan SMOTE, penelitian ini dapat

menghasilkan dataset yang seimbang dan representatif, sehingga memungkinkan analisis yang lebih akurat dan bermakna dalam pengolahan data genetik. (Elreedy dkk., 2024). Berikut merupakan rumus SMOTE oversampling

$$z = x_0 + w(x_1 - x_0)$$

Keterangan:

z : data sintetis baru yang dihasilkan

w : bilangan acak dari distribusi uniform dalam rentang $[0, 1]$

x_0 : titik data asli dari kelas minoritas yang dipilih secara acak

x_1 : salah satu dari *tetangga terdekat* dari x_0

2.4 Confusion Matrix

Confusion matrix adalah tabel yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan hasil prediksi dengan nilai aktual. Matriks ini memberikan gambaran menyeluruh tentang bagaimana model mengklasifikasikan data, termasuk jumlah prediksi yang benar dan salah untuk setiap kategori. Confusion matrix merupakan alat yang berguna dan komprehensif dalam menilai performa model klasifikasi (Kristinic dkk., 2020). Berikut merupakan tabel confusion matrix:

Tabel 2.1 Confusion Matrix

Data	Prediksi Positif	Prediksi Negatif
Aktual Positif	<i>True positive</i> (TP)	<i>False negative</i> (FN)
Aktual Negatif	<i>False positive</i> (FP)	<i>True negative</i> (TN)

Berdasarkan tabel 2.1, *confusion matrix* terdapat:

1. *True positive* (TP): Model memprediksi positif, dan data sebenarnya positif.
2. *False negative* (FN): Model memprediksi negatif, tetapi data sebenarnya positif.
3. *False positive* (FP): Model memprediksi positif, tetapi data sebenarnya negatif.
4. *True negative* (TN): Model memprediksi negatif, dan data sebenarnya negatif.

Dari Confusion Matrix, kita dapat menghitung beberapa metrik evaluasi model:

1. Akurasi

Mengukur seberapa sering model membuat prediksi yang benar.

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN}$$

2. Precision

Mengukur seberapa banyak prediksi positif yang benar.

$$Precision = \frac{TP}{TP + FP}$$

3. Recall

Mengukur model dalam menemukan semua data yang benar-benar positif.

$$Recall = \frac{TP}{TP + FN}$$

4. F1-Score

Rata-rata harmonik dari presisi dan recall.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

2.5 Natural Language Processing

NLP adalah bidang kecerdasan buatan (AI) yang berfokus pada interaksi antara komputer dan bahasa manusia. Tujuan utama NLP adalah memungkinkan komputer untuk memahami, menginterpretasikan, dan menghasilkan bahasa alami secara efektif. Dengan kemampuannya untuk menganalisis teks dalam jumlah besar, NLP telah digunakan secara luas dalam berbagai disiplin ilmu, termasuk, analisis sentimen, serta pengenalan pola dalam teks (Torregrosa dkk., 2023).

Salah satu penerapan utama NLP adalah dalam analisis sentimen, yaitu proses menentukan sikap, emosi, atau opini dalam suatu teks. Analisis sentimen sering digunakan dalam berbagai bidang, seperti bisnis untuk memahami umpan balik pelanggan, politik untuk menganalisis opini publik, dan media sosial untuk mengidentifikasi tren atau reaksi terhadap suatu peristiwa. Teknik yang digunakan dalam analisis sentimen meliputi pendekatan berbasis kamus (lexicon-based), pembelajaran mesin (machine learning), dan deep learning, yang memungkinkan sistem untuk mengenali nuansa emosi dalam teks, baik itu positif, negatif, maupun netral (Jim dkk., 2024).

2.6 ECO Friendly

Konsumen global semakin khawatir tentang dampak negatif lingkungan dari limbah kemasan. *Eco-Friendly* memiliki potensi besar untuk berkontribusi pada pembangunan berkelanjutan (Fernqvist dkk., 2015; Lewis dan Stanley, 2012; Lindh dkk., 2016a; Magnier dan Crie, 2015; Mishra dkk., 2017; Prakash dan Pathak, 2017; Rokka dan Uusitalo, 2008; Taylor dan Villas-Boas, 2016, sebagaimana dikutip dalam (Nguyen et al., 2020). Meskipun kemasan merupakan masalah sosial dan politik, hanya ada penelitian terbatas tentang persepsi konsumen terhadap kemasan *eco-friendly*. Memang, kemasan *eco-friendly* tidak pernah menjadi konsep yang jelas dalam literatur perilaku konsumen (Magnier dan Crie, 2015). Lebih jauh lagi, para peneliti telah menggunakan istilah yang berbeda untuk menunjukkan *eco-friendly*, seperti kemasan *eco-friendly*, kemasan ekologis, kemasan hijau, kemasan berkelanjutan, desain eko, desain untuk lingkungan, dan desain sadar lingkungan (Boks dan Stevels, 2007; Koenig-Lewis et al., 2014; Magnier dan Crie, 2015), yang menyebabkan kebingungan ketika melakukan penelitian.

Penerapan konsep *eco-friendly* di Sekolah Gajahwong ini dinilai masih memiliki potensi untuk dikembangkan sehingga dosen dan tim PKM S-2 Pengkajian Amerika melihat adanya kesempatan untuk berkontribusi dalam masyarakat tersebut dengan membawa *added values* (Karunia et al., 2020). Para peserta juga menggambarkan *eco-friendly* sebagai kemasan yang tidak beracun, mudah terurai saat dibuang, dan paling baik jika dapat terurai secara hayati. Karakteristik sebagian membentuk definisi konsumen tentang produk *eco-friendly* kemasan. Kemasan *eco-friendly* tidak mencemari lingkungan, dapat mudah diolah setelah digunakan dan dapat terurai secara hayati. Kemasan yang ramah lingkungan harus aman dan tidak beracun bagi manusia (Lindh et al., 2016a; Wikstrom et al., 2018, sebagaimana dikutip dalam (Nguyen et al., 2020).

2.7 Produk dalam E-Commerce.

Menurut (Alamin et al., 2023), Perkembangan e-commerce di Indonesia mengalami pertumbuhan yang pesat seiring dengan meningkatnya jumlah pengguna internet. Berdasarkan data terbaru, jumlah pengguna internet di Indonesia

terus meningkat dari tahun ke tahun, yang berdampak pada meningkatnya transaksi jual beli secara daring. E-commerce menjadi pilihan utama bagi masyarakat karena menawarkan kemudahan dalam bertransaksi, efisiensi waktu, serta akses terhadap berbagai produk yang lebih luas dibandingkan dengan toko fisik.

Dalam ekosistem e-commerce, produk yang diperdagangkan sangat beragam, mulai dari barang fisik seperti elektronik, pakaian, dan kebutuhan rumah tangga hingga produk digital seperti perangkat lunak, e-book, dan layanan berbasis langganan. Selain itu, layanan berbasis e-commerce juga mencakup sektor jasa, seperti konsultasi daring dan kursus online. Keberadaan e-commerce tidak hanya menguntungkan konsumen dalam mendapatkan produk dengan lebih mudah, tetapi juga membuka peluang bisnis yang lebih luas bagi pelaku usaha, terutama bagi UMKM yang dapat menjangkau pasar lebih besar tanpa harus memiliki toko fisik.

Salah satu tren yang berkembang dalam e-commerce adalah meningkatnya permintaan terhadap produk berlabel *eco-friendly* (Alamin et al., 2023). Produk ini diklaim lebih ramah lingkungan dengan berbagai inovasi seperti bahan daur ulang, kemasan biodegradable, dan proses produksi yang lebih berkelanjutan. Namun, permasalahan yang muncul adalah banyaknya praktik *greenwashing*, di mana beberapa produsen menggunakan label ramah lingkungan sebagai strategi pemasaran tanpa benar-benar memenuhi standar keberlanjutan. Hal ini membuat konsumen kesulitan dalam membedakan produk yang benar-benar ramah lingkungan dengan yang hanya mengklaim demikian.

Oleh karena itu, ulasan pelanggan menjadi salah satu sumber informasi penting dalam mengevaluasi produk di e-commerce. Konsumen sering kali mengandalkan ulasan dan pengalaman pembeli lain sebelum mengambil keputusan untuk membeli suatu produk. Analisis sentimen berbasis ulasan pelanggan dapat digunakan untuk memahami pola kepuasan atau kekecewaan konsumen terhadap produk *eco-friendly*. Dengan menerapkan metode seperti Algoritma Multi-layer Preceptron, opini pelanggan dapat diklasifikasikan ke dalam sentimen positif, negatif, atau netral, sehingga memberikan gambaran yang lebih objektif terhadap kualitas dan kredibilitas klaim *eco-friendly* suatu produk.

BAB III

METODE PENELITIAN

3.1 Jenis Penelitian

Penelitian ini merupakan studi kuantitatif yang bertujuan untuk menganalisis sentimen ulasan produk ramah lingkungan di e-commerce menggunakan algoritma Multi-Layer Perceptron (MLP). Pendekatan kuantitatif digunakan untuk mengukur aspek numerik seperti akurasi dan performa algoritma dalam mengklasifikasikan sentimen ulasan. Penelitian ini mengeksplorasi berbagai konfigurasi lapisan tersembunyi (hidden layers) dan fungsi aktivasi dalam model MLP untuk menentukan konfigurasi yang paling optimal dalam membedakan sentimen positif, netral, dan negatif dari ulasan produk ramah lingkungan di e-commerce menggunakan bahasa pemrograman Python.

3.2 Data dan Sumber Data

Data yang digunakan dalam penelitian ini berasal dari analisis sentimen ulasan produk Lock & Lock di platform e-commerce Tokopedia. Data tersebut mencakup hingga 3000 ulasan pelanggan yang dikumpulkan untuk mengevaluasi persepsi konsumen terhadap produk ramah lingkungan yang ditawarkan oleh merek tersebut. Setiap ulasan akan dikategorikan ke dalam tiga kelas sentimen utama, yaitu positif, netral, dan negatif, berdasarkan isi dan konteks dari teks ulasan. Setelah diberi label, data disimpan dalam format CSV agar dapat diproses lebih lanjut dalam penelitian ini. Data yang telah dipersiapkan akan digunakan untuk membangun model klasifikasi sentimen menggunakan algoritma Multi-layer Perceptron (MLP), yaitu salah satu jenis jaringan saraf tiruan yang mampu mempelajari pola non-linear kompleks. Berbagai konfigurasi jumlah neuron dan lapisan tersembunyi akan dieksplorasi untuk menentukan arsitektur yang paling optimal dalam membedakan sentimen ulasan pelanggan.

3.3 Teknik Analisis Data

1. Pengumpulan Data
Mengumpulkan data review product lock n lock di tokopedia
2. Preprocessing data
Mempersiapkan teks menjadi lebih siap digunakan untuk analisis sentimen
3. Labeling Lexicon
Memberikan label sentimen (positif, negatif, netral) pada kata-kata berdasarkan kamus leksikon.
4. TF-IDF Vectorizer
Mengubah teks menjadi representasi numerik dengan menghitung frekuensi kata dan invers frekuensi dokumen.
5. SMOTE Oversampling
Meningkatkan jumlah sampel kelas minoritas untuk mengatasi ketidakseimbangan dataset.
6. Label Encoder
Mengubah label kategori menjadi nilai numerik yang dapat diproses oleh model.
7. Modelling MLP
Membangun dan melatih model jaringan saraf tiruan Multi-layer Perceptron untuk klasifikasi.
8. Evaluasi
Evaluasi menggunakan confusion matrix

BAB IV

PEMBAHASAN

4.1 Pengumpulan Data

Data dalam penelitian ini diperoleh dari platform Tokopedia, salah satu e-commerce terbesar di Indonesia. Fokus utama penelitian ini adalah pada produk-produk ramah lingkungan (eco-friendly) dari merek LocknLock. Merek ini dikenal menghasilkan berbagai perlengkapan rumah tangga dengan klaim keberlanjutan, seperti penggunaan material daur ulang, bahan bebas BPA, serta kemasan yang lebih ramah lingkungan.

Proses pengumpulan data dilakukan secara otomatis melalui metode *web scraping*. Teknik ini menggunakan bantuan pustaka Python, yaitu *requests*, *BeautifulSoup*, dan *selenium*. Pustaka *requests* digunakan untuk melakukan permintaan HTTP ke halaman produk, sementara *BeautifulSoup* digunakan untuk mengekstrak dan memarsing konten HTML statis. Adapun *selenium* digunakan untuk mengakses dan menavigasi konten dinamis, terutama bagian ulasan pelanggan yang hanya muncul setelah interaksi pengguna, seperti scroll atau klik tombol “Lihat semua ulasan.”

Melalui proses tersebut, berhasil dikumpulkan sebanyak 3.000 data ulasan pelanggan yang tersebar pada berbagai varian produk LocknLock. Ulasan-ulasan ini berkaitan dengan opini, pengalaman, dan penilaian konsumen terhadap produk yang diklaim sebagai ramah lingkungan. Selain isi ulasan, data yang dikumpulkan juga mencakup nama produk, tanggal ulasan, dan rating bintang apabila tersedia. Namun, untuk keperluan analisis sentimen, fokus utama tetap tertuju pada isi teks ulasan.

Pengumpulan data dilakukan pada bulan April 2025. Setelah proses *scraping* selesai, seluruh data disimpan dalam format CSV (*Comma-Separated Values*) untuk memudahkan pemrosesan pada tahap-tahap analisis selanjutnya, seperti *preprocessing*, pelabelan sentimen, ekstraksi fitur, hingga pelatihan model klasifikasi.

4.2 Preprocessing

Setelah proses pengumpulan data selesai dilakukan, langkah selanjutnya dalam penelitian ini adalah tahapan *preprocessing* atau pra-pemrosesan data teks. Tahapan ini bertujuan untuk membersihkan, menyederhanakan, dan menyamakan format data ulasan agar dapat dianalisis secara efektif pada tahap klasifikasi sentimen selanjutnya. Data teks ulasan yang diperoleh dari platform Tokopedia cenderung bersifat tidak terstruktur, serta mengandung kata tidak baku, singkatan, simbol, angka, dan elemen lain yang dapat mengganggu proses analisis otomatis. Oleh karena itu, dilakukan beberapa tahapan preprocessing sebagai berikut:

4.2.1 Normalisasi Kata Tidak Baku (*Slang Word Normalization*)

Langkah pertama dalam proses preprocessing adalah normalisasi kata, yaitu mengubah kata-kata tidak baku atau slang menjadi bentuk baku dalam Bahasa Indonesia. Ulasan pelanggan sering kali mengandung kata informal seperti “mantep”, “ga/gak”, “gede”, serta campuran istilah asing seperti “happy”, “recommended”, atau “worth it”. Kata-kata tersebut disesuaikan dengan padanan yang lebih baku dan umum dalam Bahasa Indonesia, seperti “mantap”, “tidak”, “besar”, “senang”, “saran” dan “sepadan”. Proses ini dilakukan dengan menyusun kamus konversi manual berdasarkan hasil observasi terhadap data dan dikembangkan secara iteratif selama proses pengolahan.

4.2.2 Penghapusan Stopword (*Stopword Removal*)

Tahap berikutnya adalah menghapus *stopword*, yaitu kata-kata umum yang memiliki frekuensi tinggi tetapi kontribusi rendah terhadap makna sentimen. Contoh dari *stopword* adalah kata “yang”, “di”, “ke”, “dan”, “ini”, serta beberapa kata tambahan yang dianggap tidak relevan dalam konteks ulasan produk. Penghapusan ini dilakukan dengan mengacu pada daftar *stopword* dari pustaka Bahasa Indonesia serta ditambah dengan daftar khusus berdasarkan konteks produk LocknLock yang dikaji.

4.2.3 Penyeragaman Huruf (*Case Folding*)

Case folding adalah proses mengubah semua huruf dalam teks menjadi huruf kecil (lowercase). Langkah ini penting untuk menyeragamkan data, sehingga variasi huruf kapital tidak menyebabkan kata yang sama dianggap berbeda oleh mesin. Misalnya, kata “Bagus” dan “bagus” akan dianggap identik setelah dilakukan case folding.

4.2.4 Penghapusan Angka (*Number Removal*)

Data ulasan sering kali mengandung angka yang tidak relevan terhadap analisis sentimen, seperti ukuran produk, jumlah pembelian, atau penomoran (contoh: “100”, “3pcs”, “1x”). Angka-angka ini dihapus karena umumnya tidak memiliki nilai emosional atau sentimen terhadap kualitas produk.

4.2.5 Penghapusan Tanda Baca dan Simbol

Langkah selanjutnya adalah menghapus tanda baca, simbol khusus, serta karakter non-alfabet lainnya yang tidak berkontribusi pada makna kata. Tanda baca seperti titik, koma, tanda seru, tanda tanya, serta emotikon dihapus agar tidak mengganggu dalam pembentukan fitur teks.

4.2.6 Penghapusan Spasi Berlebih (*Whitespace Cleaning*)

Sebagai langkah akhir dalam preprocessing, dilakukan pembersihan spasi berlebih yang dapat muncul akibat penghapusan elemen-elemen sebelumnya. Spasi yang tidak perlu dapat mengganggu proses tokenisasi atau analisis lanjutan, sehingga dibersihkan untuk menjaga konsistensi format teks. Semua tahapan preprocessing ini dilakukan secara sistematis dengan memanfaatkan pustaka pemrograman Python seperti re (regular expression), pandas, dan Sastrawi. Hasil dari proses preprocessing disimpan ke dalam kolom teks yang baru dan siap digunakan dalam tahap selanjutnya, yaitu ekstraksi fitur dan analisis sentimen menggunakan pendekatan pembelajaran mesin.

4.3 Labeling Lexicon

Tahap ini bertujuan untuk memberikan label sentimen pada setiap data ulasan yang telah melalui proses pra-pemrosesan. Metode yang digunakan adalah pendekatan *lexicon-based*, yaitu klasifikasi sentimen berdasarkan keberadaan kata-kata bermuatan positif atau negatif yang terdapat dalam teks ulasan.

Pendekatan ini mengandalkan dua daftar kata kunci (*lexicon*), yaitu **lexicon positif** dan **lexicon negatif**. Lexicon positif berisi kata-kata yang umumnya merepresentasikan kepuasan pelanggan, seperti "*bagus*", "*puas*", "*awet*", dan "*senang*". Sementara itu, lexicon negatif terdiri dari kata-kata yang mengindikasikan ketidakpuasan atau pengalaman buruk terhadap produk, seperti "*rusak*", "*lambat*", "*bocor*", dan "*cacat*".

Proses labeling dilakukan dengan mengidentifikasi jumlah kemunculan kata-kata positif dan negatif dalam setiap ulasan. Setiap ulasan kemudian dianalisis untuk menentukan dominasi sentimen berdasarkan aturan berikut:

- Jika jumlah kata positif lebih banyak daripada kata negatif, maka ulasan diberi label **positif**;
- Jika jumlah kata negatif lebih banyak daripada kata positif, maka ulasan diberi label **negatif**;
- Jika jumlah kata positif dan negatif sama, atau tidak ditemukan kata dari kedua lexicon, maka ulasan diberi label **netral**.

Seluruh proses ini dilakukan secara otomatis menggunakan fungsi yang dirancang dalam bahasa pemrograman Python. Fungsi tersebut menghitung frekuensi kemunculan kata dalam masing-masing lexicon untuk setiap teks ulasan yang telah dibersihkan sebelumnya. Hasil akhir berupa label sentimen kemudian disimpan dalam kolom baru pada dataset, yang akan digunakan sebagai target dalam proses klasifikasi pada tahap selanjutnya.

Dengan metode ini, sebanyak 3.000 ulasan berhasil diberi label sentimen secara efisien dan sistematis. Untuk menjamin kualitas labeling, dilakukan pula validasi manual pada sebagian data secara acak guna memastikan bahwa pelabelan telah sesuai konteks dan makna sebenarnya dalam ulasan.

4.4 TF-IDF Vectorizer

Setelah melewati tahap preprocessing, teks akan diubah menjadi format numerik menggunakan TF-IDF Vectorizer. Langkah ini penting supaya teks bisa dipahami dan diproses oleh algoritma *machine learning*. Rumus yang digunakan untuk menghitung nilai TF-IDF adalah sebagai berikut:

$$TF - IDF_{ij} = \frac{n_{ij}}{d_j} \cdot \lg \left(\frac{D}{N_i} \right)$$

4.5 SMOTE Oversampling

Setelah berhasil mengubah teks menjadi representasi numerik menggunakan TF-IDF Vectorizer, langkah selanjutnya yang krusial adalah menangani ketidakseimbangan kelas (class imbalance) dalam dataset, jika ada. Untuk mengatasi masalah ini, akan diterapkan teknik SMOTE (Synthetic Minority Oversampling Technique) dengan rumus berikut:

$$z = x_0 + w(x_1 - x_0)$$

Berdasarkan rumus tersebut, didapatkan hasil SMOTE pada tabel berikut:

Tabel 4.1 Hasil SMOTE

Kelas	Sebelum SMOTE	Setelah SMOTE
Negatif	175	2170
Netral	666	2170
Positif	2170	2170

Berdasarkan tabel tersebut, didapatkan bahwa data sebelum SMOTE tidak seimbang yang menyebabkan model machine learning buruk. Setelah SMOTE data seimbang menjadi 2170 data di semua kelas.

4.6 Label Encoder

Setelah proses pelabelan sentimen selesai dan data mengalami penyeimbangan kelas melalui metode SMOTE, langkah selanjutnya adalah mengubah label sentimen dari format kategorikal (string) menjadi format numerik. Proses ini dikenal dengan istilah Label Encoding.

Label Encoding merupakan teknik transformasi data kategori menjadi bentuk numerik agar dapat diproses oleh algoritma pembelajaran mesin yang hanya menerima input dalam bentuk angka. Dalam konteks penelitian ini, tiga label sentimen yang telah dihasilkan dari proses labeling, yaitu "Negatif", "Netral", dan "Positif", dikonversi menjadi representasi angka sebagai berikut:

Tabel 4.2 Hasil Label Encoder

Kelas	Label
Negatif	0
Netral	1
Positif	2

Berdasarkan tabel 4.2, proses encoding dilakukan menggunakan LabelEncoder dari library scikit-learn, yang secara otomatis mengubah setiap label string menjadi bilangan bulat sesuai urutan abjad. Label yang telah di-encode kemudian ditambahkan sebagai kolom baru ke dalam DataFrame hasil oversampling, yang sebelumnya sudah berisi fitur-fitur hasil transformasi TF-IDF.

Langkah ini penting dilakukan sebelum proses pelatihan model klasifikasi, agar setiap sampel data memiliki representasi numerik baik untuk fitur maupun target variabel. Selain itu, penggunaan Label Encoding juga membantu memudahkan proses evaluasi performa model klasifikasi di tahap berikutnya.

4.7 Modeling MLP

Setelah semua tahap preprocessing, termasuk TF-IDF Vectorization dan SMOTE Oversampling, terselesaikan dan data siap, langkah berikutnya adalah melakukan *modelling* menggunakan Multi-Layer Perceptron (MLP).

Berikut merupakan perhitungan rumus hidden layer:

$$z_j = \sum_{i=1}^m (w_{ij} \cdot X_i) + b_j$$

$$a_j = f(z_j) = \frac{1}{1 + e^{-z_j}}$$

Berikut merupakan perhitungan Output Layer.

$$z_k = \sum_{j=1}^l (w_{jk} \cdot a_j) + b_k$$

$$\hat{y}_k = f(z_k) = \frac{1}{1 + e^{-z_k}}$$

4.8 Evaluasi

Proses evaluasi dilakukan untuk mengetahui sejauh mana model klasifikasi sentimen yang dibangun—dalam hal ini menggunakan algoritma MLP—mampu mengelompokkan data ulasan ke dalam kategori sentimen *Negatif*, *Netral*, dan *Positif* dengan akurasi yang tinggi.

Pengujian dilakukan dengan membandingkan hasil prediksi model terhadap data uji (y_{test}) dan label sebenarnya. Salah satu metode evaluasi yang digunakan adalah confusion matrix, yang di jelaskan pada tabel. Tabel berikut merangkum jumlah prediksi benar dan salah untuk masing-masing kelas:

Tabel 4.3 Hasil Confusion Matrix

Kelas	Negatif	Netral	Positif
Negatif	645	0	0
Netral	0	645	0
Positif	17	18	610

Berdasarkan tabel 4.3, model memiliki performa yang sangat baik dalam mengklasifikasikan ulasan dengan sentimen negatif dan netral dengan akurasi sempurna pada kelas tersebut (645 prediksi benar dan tanpa kesalahan klasifikasi). Untuk kelas netral dan positif, meskipun sebagian besar prediksi sudah tepat, masih ditemukan beberapa kesalahan klasifikasi—terutama pada kelas *positif* yang sering kali diklasifikasikan sebagai *netral*.

Untuk mengukur performa model secara lebih komprehensif, digunakan metrik evaluasi lain seperti precision, recall, dan f1-score, yang hasilnya disajikan pada tabel berikut

Tabel 4.4 Hasil Klasifikasi

	Precision	Recall	F1 Score
Negatif	97%	100%	99%
Netral	97%	100%	99%
Positif	100%	95%	97%
Accuracy	98%		

Berdasarkan tabel 4.4, secara umum, model MLP yang digunakan dalam penelitian ini menunjukkan performa klasifikasi yang tinggi, dengan rata-rata *precision* dan *recall* yang memuaskan. Kelebihan utama terlihat pada kemampuan model dalam mendeteksi ulasan negatif dan netral secara konsisten. Namun, perlu dicatat bahwa recall pada kelas positif masih bisa ditingkatkan, mengingat masih terdapat sejumlah prediksi yang salah klasifikasi sebagai netral. Akurasi mencapai 98% disimpulkan bahwa model memiliki hasil yang lebih baik dibandingkan penelitian terdahulu pada tabel (1.1).

Dengan demikian, dapat disimpulkan bahwa model memiliki tingkat keandalan yang baik dalam melakukan analisis sentimen berbasis teks, dan dapat dijadikan sebagai dasar dalam pengembangan sistem klasifikasi ulasan pelanggan secara otomatis.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil evaluasi yang telah dilakukan, model klasifikasi sentimen menggunakan algoritma Multi-Layer Perceptron (MLP) menunjukkan performa yang sangat baik dalam mengelompokkan data ulasan ke dalam tiga kategori sentimen, yaitu Negatif, Netral, dan Positif.

Berdasarkan confusion matrix, model berhasil mengklasifikasikan ulasan dengan sentimen Negatif dan Netral secara sempurna, dengan masing-masing 645 prediksi yang benar dan tanpa kesalahan klasifikasi. Sementara itu, pada kelas Positif, meskipun sebagian besar prediksi sudah tepat (610 benar dari total data), masih ditemukan beberapa kesalahan klasifikasi yang dikategorikan sebagai Netral, yaitu sebanyak 18 kasus, dan 17 kesalahan dari kelas Positif yang diklasifikasikan sebagai Negatif.

Lebih lanjut, evaluasi menggunakan metrik precision, recall, dan f1-score menunjukkan bahwa model memiliki performa klasifikasi yang tinggi secara keseluruhan. Precision dan recall pada kelas Negatif dan Netral mencapai 97%–100%, sedangkan untuk kelas Positif, precision mencapai 100%, namun recall-nya sedikit lebih rendah yaitu 95%. Akurasi keseluruhan model mencapai 98%, yang menunjukkan bahwa model ini sangat andal dan bahkan lebih baik dibandingkan pendekatan-pendekatan sebelumnya sebagaimana diuraikan dalam Tabel 1.1 pada bab sebelumnya.

Dengan demikian, dapat disimpulkan bahwa algoritma MLP mampu memberikan hasil yang akurat, stabil, dan efektif dalam menganalisis sentimen teks ulasan pelanggan. Model ini berpotensi besar untuk dikembangkan lebih lanjut dan diterapkan pada sistem klasifikasi ulasan otomatis untuk mendukung pengambilan keputusan berbasis opini pengguna.

5.2 Saran

Meskipun model MLP telah menunjukkan kinerja yang sangat baik, terdapat beberapa hal yang dapat dipertimbangkan untuk pengembangan selanjutnya:

1. Peningkatan Recall pada Kelas Positif

Model masih menunjukkan kelemahan dalam mengenali ulasan dengan sentimen positif secara menyeluruh, terlihat dari recall yang sedikit lebih rendah (95%). Penambahan data latih yang lebih beragam untuk kelas positif atau penerapan teknik penyeimbangan data seperti oversampling atau SMOTE dapat dipertimbangkan untuk mengatasi hal ini.

2. Eksplorasi Algoritma Alternatif

Untuk mendapatkan perbandingan yang lebih menyeluruh, disarankan untuk mengevaluasi performa algoritma lain seperti LSTM, BERT, atau model berbasis Transformer, yang terbukti efektif dalam menangani teks dan konteks kalimat yang kompleks.

3. Analisis Kesalahan Klasifikasi

Perlu dilakukan analisis mendalam terhadap data yang salah diklasifikasikan, khususnya pada kelas positif yang diprediksi sebagai netral. Hal ini dapat membantu dalam memahami pola kesalahan dan melakukan penyempurnaan pada preprocessing data atau arsitektur model.

4. Implementasi Sistem Nyata

Model yang sudah dikembangkan dapat diimplementasikan ke dalam sistem berbasis web atau aplikasi mobile sebagai fitur klasifikasi sentimen otomatis pada ulasan produk atau layanan, guna memberikan manfaat langsung dalam dunia industri atau bisnis digital.

5. Evaluasi dengan Dataset Lain

Untuk memastikan generalisasi model, sebaiknya dilakukan pengujian pada dataset lain yang memiliki karakteristik berbeda. Hal ini dapat mengukur robust-nya model terhadap variasi bahasa dan konteks yang lebih luas.

Dengan mempertimbangkan saran-saran di atas, diharapkan penelitian ini tidak hanya memberikan hasil yang baik dalam konteks eksperimen, tetapi juga mampu berkontribusi nyata dalam pengembangan sistem klasifikasi opini publik secara otomatis dan andal.

DAFTAR PUSTAKA

- Adhi Putra, A. D. (2021). Analisis sentimen pada ulasan pengguna aplikasi Bibit dan Bareksa dengan algoritma KNN. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 8(2), 636–646. <https://doi.org/10.35957/jatisi.v8i2.962>
- Alamin, Z., Missouri, R., Sutriawan, S., Fathir, F., & Khairunnas, K. (2023). Perkembangan e-commerce: Analisis dominasi Shopee sebagai primadona marketplace di Indonesia. *J-ESA (Jurnal Ekonomi Syariah)*, 6(2), Article 2. <https://doi.org/10.52266/jesa.v6i2.2484>
- Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine learning from theory to algorithms: An overview. *Journal of Physics: Conference Series*, 1142, 012012. <https://doi.org/10.1088/1742-6596/1142/1/012012>
- Atimi, R. L., & Pratama, E. E. (2022). Implementasi model klasifikasi sentimen pada review produk Lazada Indonesia. *Jurnal Sains dan Informatika*, 8(1), 88–96. <https://doi.org/10.34128/jsi.v8i1.419>
- Elreedy, D., Atiya, A. F., & Kamalov, F. (2024). A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning. *Machine Learning*, 113(7), 4903–4923. <https://doi.org/10.1007/s10994-022-06296-4>
- Jim, J. R., Talukder, M. A. R., Malakar, P., Kabir, M. M., Nur, K., & Mridha, M. F. (2024). Recent advancements and challenges of NLP-based sentiment analysis: A state-of-the-art review. *Natural Language Processing Journal*, 6, 100059. <https://doi.org/10.1016/j.nlp.2024.100059>
- Karunia, I., Fataya, I. A., Sari, M. A. K., & Basuki, A. (2020). Penerapan budaya ramah lingkungan pada siswa-siswi Sekolah Gajahwong: Pendidikan alternatif berbasis eco-friendly. *Bakti Budaya*, 3(1), 80. <https://doi.org/10.22146/bb.55504>
- Kosarirad, H., Ghasempour Nejati, M., Saffari, A., Khishe, M., & Mohammadi, M. (2022). Feature Selection and Training Multi-layer Perceptron Neural Networks Using Grasshopper Optimization Algorithm for Design Optimal Classifier of Big Data Sonar. *Journal of Sensors*, 2022, 1–14. <https://doi.org/10.1155/2022/9620555>
- Krstinić, D., Braović, M., Šerić, L., & Božić-Štulić, D. (2020). Multi-label classifier performance evaluation with confusion matrix. *Computer Science & Information Technology*, 10(8), 1–14. <https://doi.org/10.5121/csit.2020.100801>

- Liu, Y., Cheng, Y., Zhang, Z., & Wu, J. (2022). Multi-information fusion fault diagnosis based on KNN and improved evidence theory. *Journal of Vibration Engineering & Technologies*, 10(3), 841–852. <https://doi.org/10.1007/s42417-021-00413-8>
- Liang, M., & Niu, T. (2022). Research on Text Classification Techniques Based on Improved TF-IDF Algorithm and LSTM Inputs. *Procedia Computer Science*, 208, 460–470. <https://doi.org/10.1016/j.procs.2022.10.064>
- Makmur, H., Wulandari, W., Surianto, D. F., & Fajar B, M. (2024). Analisis Sentimen Penghapusan Skripsi sebagai Tugas Akhir Mahasiswa Menggunakan Metode Multi-Layer Perceptron. *Komputika : Jurnal Sistem Komputer*, 13(2). <https://doi.org/10.34010/komputika.v13i2.12402>
- Mutiara Sari, R., & Sutikno, S. (2023). Perbandingan metode KNN dan MKNN untuk deteksi dini diabetes mellitus. *Jurnal Mnemonic*, 6(2), 96–101. <https://doi.org/10.36040/mnemonic.v6i2.6021>
- Nayak, S., Bhat, M., Subba Reddy, N. V., & Ashwath Rao, B. (2022). Study of distance metrics on k-nearest neighbor algorithm for star categorization. *Journal of Physics: Conference Series*, 2161(1), 012004. <https://doi.org/10.1088/1742-6596/2161/1/012004>
- Nguyen, A. T., Parker, L., Brennan, L., & Lockrey, S. (2020). A consumer definition of eco-friendly packaging. *Journal of Cleaner Production*, 252, 119792. <https://doi.org/10.1016/j.jclepro.2019.119792>
- Torregrosa, J., Bello-Orgaz, G., Martínez-Cámara, E., Ser, J. del, & Camacho, D. (2023). A survey on extremism analysis using natural language processing: Definitions, literature review, trends and challenges. *Journal of Ambient Intelligence and Humanized Computing*, 14(8), 9869–9905. <https://doi.org/10.1007/s12652-021-03658-z>
- Zhang, S. (2022). Challenges in KNN classification. *IEEE Transactions on Knowledge and Data Engineering*, 34(10), 4663–4675. <https://doi.org/10.1109/TKDE.2021.3049250>

LAMPIRAN

Lampiran 1 Data

