

**ANALISIS SENTIMEN PENGARUH TIKTOKSHOP TERHADAP  
UMKM LOKAL PADA TWITTER MENGGUNAKAN  
ALGORITMA BERT**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**DEWI AYU PRATIWI**

**18.11.2377**

Kepada

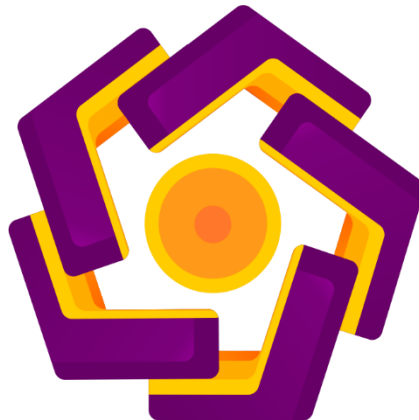
**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2024**

**ANALISIS SENTIMEN PENGARUH TIKTOKSHOP TERHADAP  
UMKM LOKAL PADA TWITTER MENGGUNAKAN  
ALGORITMA BERT**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**DEWI AYU PRATIWI**

**18.11.2377**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA**

**YOGYAKARTA**

**2024**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**ANALISIS SENTIMEN PENGARUH TIKTOKSHOP TERHADAP UMKM  
LOKAL PADA TWITTER MENGGUNAKAN ALGORITMA BERT**

yang disusun dan diajukan oleh

**Dewi Ayu Pratiwi**

**18.11.2377**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 04 Juni 2024

**Dosen Pembimbing,**



**Theopilus Bayu Sasongko, S.Kom, M.Eng**

**NIK. 190302375**



**HALAMAN PENGESAHAN**

**SKRIPSI**

**ANALISIS SENTIMEN PENGARUH TIKTOKSHOP TERHADAP UMKM  
LOKAL PADA TWITTER MENGGUNAKAN ALGORITMA BERT**

yang disusun dan diajukan oleh

**Dewi Ayu Pratiwi**

**18.11.2377**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 24 Juni 2024

**Susunan Dewan Penguji**

**Nama Penguji**

**Kusnawi, S.Kom, M.Eng**

**NIK. 190302112**

**Windha Mega Pradnya D, M.Kom**

**NIK. 190302185**

**Theopilus Bayu Sasongko, S.Kom, M.Eng**

**NIK. 190302375**

**Tanda Tangan**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 24 Juni 2024

**DEKAN FAKULTAS ILMU KOMPUTER**



**Hanif Al Fatta, S.Kom., M.Kom.**

**NIK. 190302096**

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Dewi Ayu Pratiwi  
NIM : 18.11.2377

Menyatakan bahwa Skripsi dengan judul berikut:

**Analisis Sentimen Pengaruh TikTok Shop Terhadap UMKM Lokal Pada Twitter Menggunakan Algoritma BERT**

Dosen Pembimbing : Theopilus Bayu Sasongko, S.Kom, M.Eng

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian **SAYA** sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab **SAYA**, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini **SAYA** buat dengan sesungguhnya, apabila di kemudian hari terdapat **penyimpangan** dan **ketidakbenaran** dalam pernyataan ini, maka **SAYA** bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 24 Juni 2024

Yang Menyatakan,



Dewi Ayu Pratiwi

## HALAMAN PERSEMBAHAN

Alhamdulillahirrobbil ‘alamin, skripsi ini penulis persembahkan untuk:

1. Allah SWT, karena berkat rahmad dan hidayah-Nya, saya masih diberi kesempatan untuk dapat memperjuangkan apa yang perlu diperjuangkan.
2. Kedua Orang Tua, Kakak, Adik, dan Tante Haji tercinta yang telah memberikan motivasi, apresiasi, dan dukungan yang tiada hentinya.
3. Bapak Rektor Universitas Amikom Yogyakarta dan staf ahli, sehingga saya dapat menimba ilmu di universitas ini.
4. Dosen pembimbing, Bapak Theopilus Bayu Sasongko, S.Kom, M.Eng., yang telah membimbing dari awal hingga terselesaikannya skripsi, serta seluruh dosen pengajar Universitas Amikom Yogyakarta khususnya Jurusan Informatika yang telah memberikan ilmu yang bermanfaat hingga pada tahap ini.
5. Semua pihak yang telah memberikan bantuan dan doa sehingga dapat tersusunnya skripsi ini.



## KATA PENGANTAR

Puji Syukur penulis panjatkan kehadirat Allah SWT, yang telah memberikan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Penulisan skripsi ini merupakan bagian dari tugas akhir dalam menyelesaikan pendidikan guna meraih gelar Sarjana Komputer pada Universitas Amikom Yogyakarta.

Penulis ingin menyampaikan ucapan terimakasih yang sebesar-besarnya kepada dosen pembimbing, Bapak Theopilus Bayu Sasongko, S.Kom, M.Eng dan [Dosen Penguji], atas waktu dan masukan yang sangat berharga dalam pengembangan skripsi ini.

Penulis juga mengucapkan terimakasih kepada semua pihak yang turut serta mendukung dan memberikan semangat. Ucapan terimakasih yang tak terhingga juga disampaikan kepada orang tua, keluarga, dan pihak-pihak yang telah memberikan dukungan moral dan doa restu selama proses penyusunan skripsi ini.

Semoga hasil dari skripsi ini dapat memberikan manfaat dan menjadi kontribusi positif dalam pengembangan ilmu pengetahuan dan teknologi. Akhir kata, penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, kritik dan saran yang bersifat membangun sangat diharapkan demi perbaikan pada penelitian mendatang.

Yogyakarta, 24 Juni 2024

Penulis

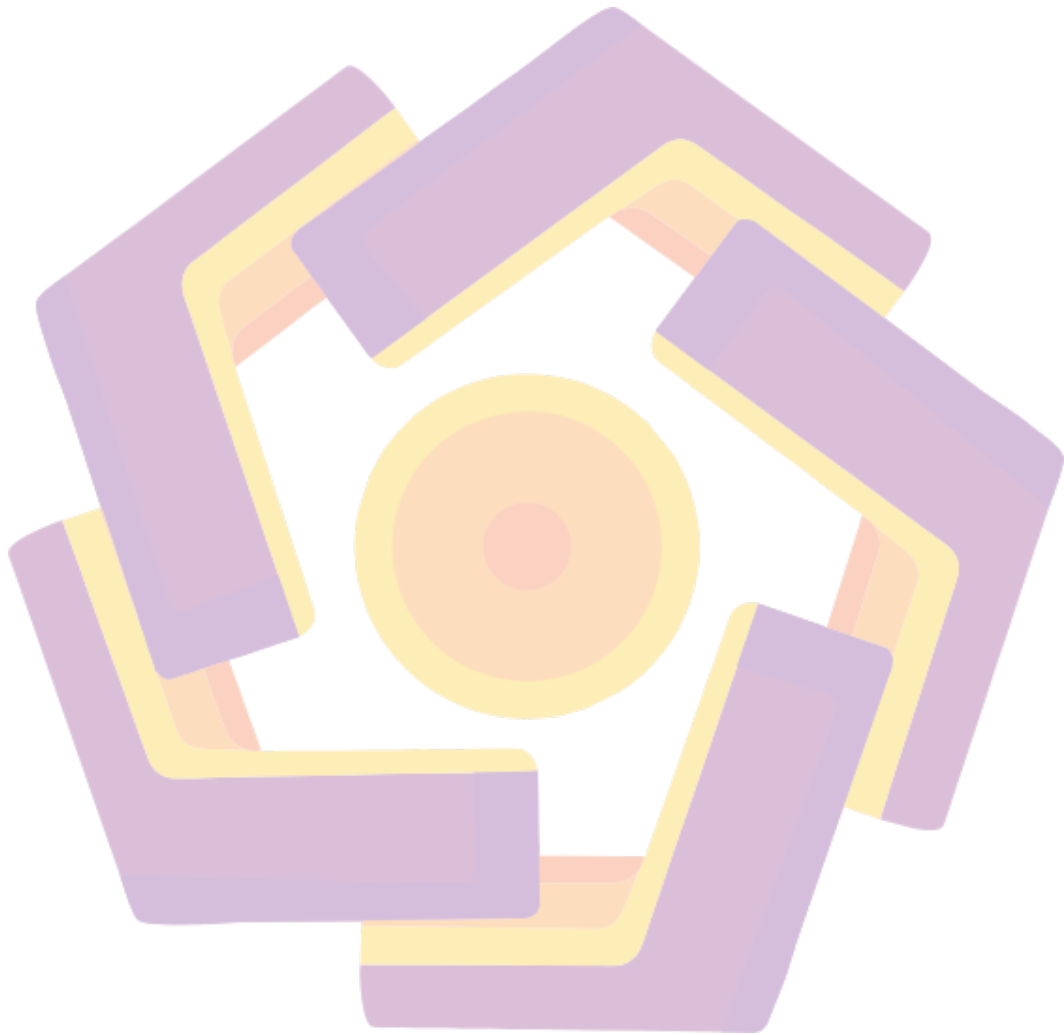
## DAFTAR ISI

|   |    |
|---|----|
| HALAMAN JUDUL .....   | 1  |
| HALAMAN PERSETUJUAN.....  | 2  |
| HALAMAN PENGESAHAN .....  | 3  |
| HALAMAN PERNYATAAN KEASLIAN SKRIPSI .....                       | 4  |
| HALAMAN PERSEMBAHAN .....                                       | 5  |
| KATA PENGANTAR .....  | 6  |
| DAFTAR ISI.....   | 7  |
| DAFTAR TABEL.....   | 11 |
| DAFTAR GAMBAR.....  | 12 |
| DAFTAR LAMPIRAN.....  | 13 |
| DAFTAR LAMBANG DAN SINGKATAN .....                              | 14 |
| DAFTAR ISTILAH .....  | 15 |
| INTISARI .....  | 16 |
| <i>ABSTRACT</i> .....   | 17 |
| BAB I PENDAHULUAN.....  | 1  |
| 1.1 Latar Belakang .....  | 1  |
| 1.2 Rumusan Masalah.....  | 2  |
| 1.3 Batasan Masalah .....                                       | 2  |
| 1.4 Tujuan Penelitian .....                                     | 3  |
| 1.5 Manfaat Penelitian .....                                    | 3  |
| 1.6 Sistematika Penulisan .....                                 | 4  |
| BAB II TINJAUAN PUSTAKA .....                                   | 5  |
| 2.1 Studi Literatur.....  | 5  |
| 2.2 Usaha Mikro, Kecil, dan Menengah (UMKM).....                | 18 |
| 2.2.2 Karakteristik Usaha Mikro, Kecil, dan Menengah (UMKM).... | 20 |
| 2.3 E-Commerce .....  | 20 |
| 2.3.1 Pengertian E-Commerce .....                               | 21 |
| 2.3.2 Keunggulan E-Commerce.....                                | 22 |
| 2.4 Twitter.....  | 22 |



|  |   |           |
|--|---|-----------|
| 2.5                                      | Tiktok.....   | 23        |
| 2.6                                      | Data Mining .....   | 25        |
| 2.6.1                                    | Pengertian Data Mining .....  | 25        |
| 2.6.2                                    | Proses Data Mining.....   | 26        |
| 2.7                                      | Machine Learning .....  | 27        |
| 2.7.1                                    | Pengertian Machine Learning.....  | 27        |
| 2.7.2                                    | Proses Machine Learning.....  | 27        |
| 2.8                                      | <i>Text Mining</i> .....  | 28        |
| 2.8.1                                    | Pengertian <i>Text Mining</i> .....   | 28        |
| 2.8.2                                    | Tahapan <i>Text Mining</i> .....  | 29        |
| 2.9                                      | Pembobotan Kata Berbasis Leksikon (Lexion-based word weighting)<br>29       |           |
| 2.10                                     | Analisis Sentimen .....   | 30        |
| 2.10.1                                   | Pengertian Analisis Sentimen .....  | 30        |
| 2.10.2                                   | Proses Analisis Sentimen.....   | 30        |
| 2.11                                     | Klasifikasi .....   | 31        |
| 2.12                                     | Bidirectional Encoder <i>Representations from Transformers (BERT)</i> ..... | 33        |
| <b>BAB III METODE PENELITIAN .....</b>   |   | <b>37</b> |
| 3.1                                      | Objek Penelitian.....   | 37        |
| 3.2                                      | Alur Penelitian .....   | 37        |
| 3.2.1                                    | Identifikasi Masalah.....   | 39        |
| 3.2.2                                    | Studi Kasus .....   | 39        |
| 3.3.3                                    | Scrapping Data.....   | 39        |
| 3.3.4                                    | Pre-processing.....   | 40        |
| 3.3                                      | Alat dan Bahan.....   | 46        |
| <b>BAB IV HASIL DAN PEMBAHASAN .....</b> |   | <b>48</b> |
| 4.1                                      | Input Dataset .....   | 48        |

|                     |  |    |
|---------------------|--|----|
| 4.2                 | Data Preprocessing.....                | 49 |
| 4.2.1               | Cleaning Text.....                     | 49 |
| 4.2.2               | Case Folding .....                     | 50 |
| 4.2.3               | Tokenizing .....                       | 50 |
| 4.2.4               | Filtering.....                         | 51 |
| 4.2.5               | Stemming .....                         | 52 |
| 4.2.6               | Labelling .....                        | 52 |
| 4.3                 | EDA .....                              | 53 |
| 4.3.1               | Percentage .....                       | 54 |
| 4.3.2               | Visualisasi (Word Cloud).....          | 55 |
| 4.3.2.1             | Positif .....                          | 55 |
| 4.3.2.2             | Neutral.....                           | 56 |
| 4.3.2.3             | Negatif.....                           | 57 |
| 4.4                 | Data Preparation.....                  | 57 |
| 4.4.1               | Splitting.....                         | 57 |
| 4.4.2               | Tokenizer and Encoding Our Data.....   | 58 |
| 4.4.3               | Data Loader.....                       | 58 |
| 4.5                 | Modelling (BERT).....                  | 59 |
| 4.5.1               | Setting Up BERT Pretrained Model ..... | 59 |
| 4.5.2               | Setting Up Optimizer & Scheduler ..... | 60 |
| 4.5.3               | Training BERT.....                     | 60 |
| 4.6                 | Prediction and Evaluation .....        | 61 |
| BAB V PENUTUP ..... |  | 63 |
| 5.1                 | Kesimpulan .....                       | 63 |
| 5.2                 | Saran .....                            | 63 |
| REFERENSI .....     |  | 65 |



## DAFTAR TABEL

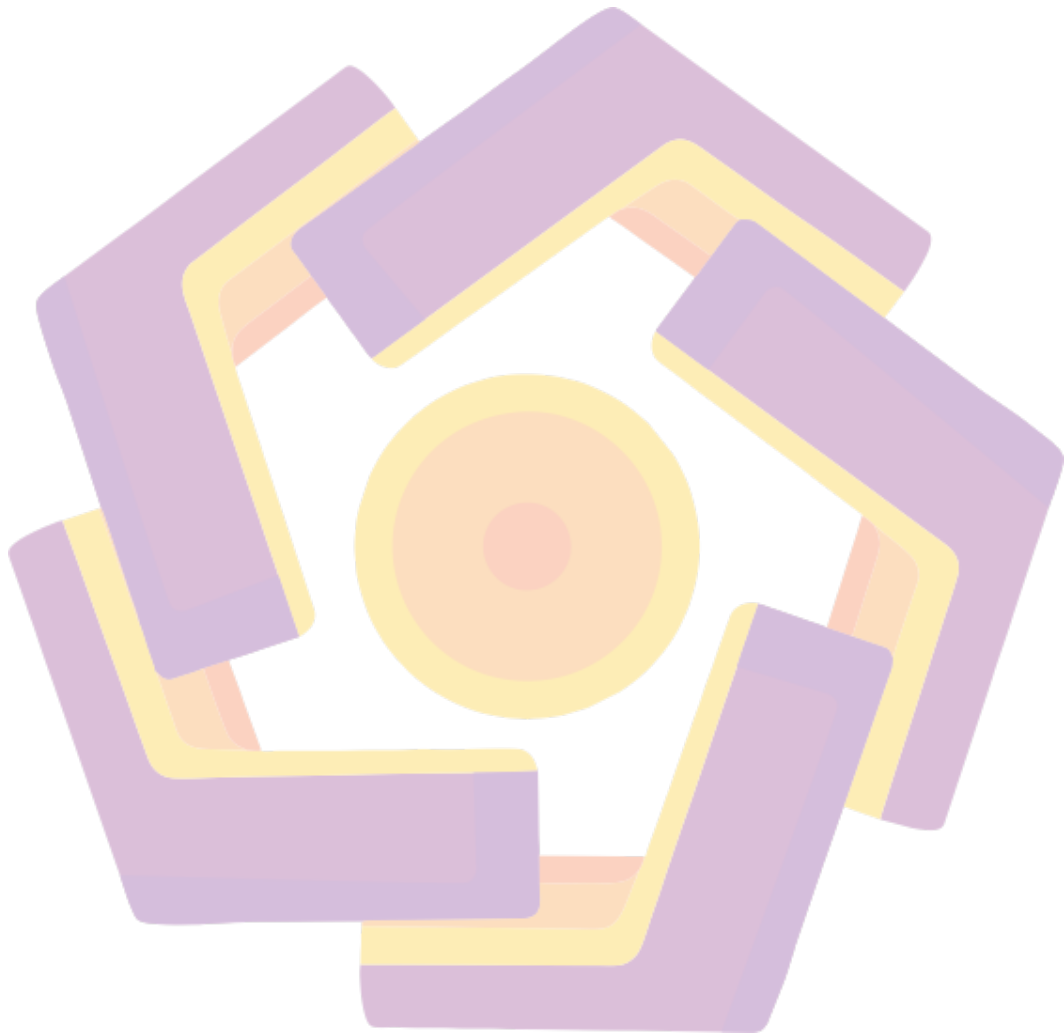
|  |    |
|--|----|
| <b>Tabel 2. 1</b> Keaslian Penelitian.....                 | 9  |
| <b>Tabel 2. 2</b> Matrik 2 x 2.....                        | 32 |
| <b>Tabel 3. 1</b> Perangkat Keras dan Perangkat Lunak..... | 46 |
| <b>Tabel 3. 2</b> Twitter Data.....                        | 47 |
| <b>Tabel 4. 1</b> Hasil Scrapping.....                     | 48 |
| <b>Tabel 4. 2</b> Hasil Cleaning.....                      | 49 |
| <b>Tabel 4. 3</b> Hasil Case Folding.....                  | 50 |
| <b>Tabel 4. 4</b> Hasil Tokenizing.....                    | 51 |
| <b>Tabel 4. 5</b> Hasil Filtering.....                     | 51 |
| <b>Tabel 4. 6</b> Hasil Stemming.....                      | 52 |
| <b>Tabel 4. 7</b> Hasil Labelling.....                     | 53 |
| <b>Tabel 4. 8</b> Jumlah Sentiment Berdasarkan Score.....  | 53 |
| <b>Tabel 4. 9</b> Akurasi Analisis Sentimen.....           | 62 |



## DAFTAR GAMBAR

|  |    |
|--|----|
| <b>Gambar 2. 1</b> Jumlah UMKM di Negara ASEAN (2021) [17]. .....  | 18 |
| <b>Gambar 2. 2</b> 10 Negara dengan Persentase Penggunaan E-Commerce Tertinggi di Dunia (April 2021) [20]..... | 21 |
| <b>Gambar 2. 3</b> 10 Negara dengan Jumlah Pengguna Twitter Terbanyak di Dunia (April 2023) .....              | 23 |
| <b>Gambar 2. 4</b> 10 Negara dengan Jumlah Pengguna TikTok Terbanyak di Dunia (April 2023) .....               | 24 |
| <b>Gambar 2. 5</b> Persentase pengguna Media Sosial untuk Berbelanja .....                                     | 25 |
| <b>Gambar 2. 6</b> Tahapan Proses KDD [31].....  | 26 |
| <b>Gambar 2. 7</b> The Transformer - model architecture [43]. .....  | 34 |
| <b>Gambar 2. 8</b> Model Attention [43]. .....   | 35 |
| <b>Gambar 3. 1</b> Alur Penelitian .....   | 38 |
| <b>Gambar 3. 2</b> Alur Scrapping .....  | 39 |
| <b>Gambar 3. 3</b> Alur Case Folding.....  | 40 |
| <b>Gambar 3. 4</b> Alur Tokenization.....  | 41 |
| <b>Gambar 3. 5</b> Alur Filtering .....  | 42 |
| <b>Gambar 3. 6</b> Alur Stemming .....   | 42 |
| <b>Gambar 3. 7</b> Alur Labelling.....   | 43 |
| <b>Gambar 4. 1</b> Hasil Klasifikasi .....   | 54 |
| <b>Gambar 4. 2</b> Hasil Persentase .....  | 55 |
| <b>Gambar 4. 3</b> Visualisasi kategori positif .....  | 56 |
| <b>Gambar 4. 4</b> Visualisasi kategori netral .....   | 56 |
| <b>Gambar 4. 5</b> Visualisasi kategori negative .....   | 57 |
| <b>Gambar 4. 6</b> Hasil Splitting .....   | 58 |
| <b>Gambar 4. 7</b> Tokenisasi data train.....  | 58 |
| <b>Gambar 4. 8</b> Tokenisasi data validasi.....   | 58 |
| <b>Gambar 4. 9</b> Hasil set-up BERT pre-trained model .....   | 59 |
| <b>Gambar 4. 10</b> Grafik Training Loss .....   | 60 |
| <b>Gambar 4. 11</b> Diagram Confusion Matrix.....  | 61 |

## DAFTAR LAMPIRAN

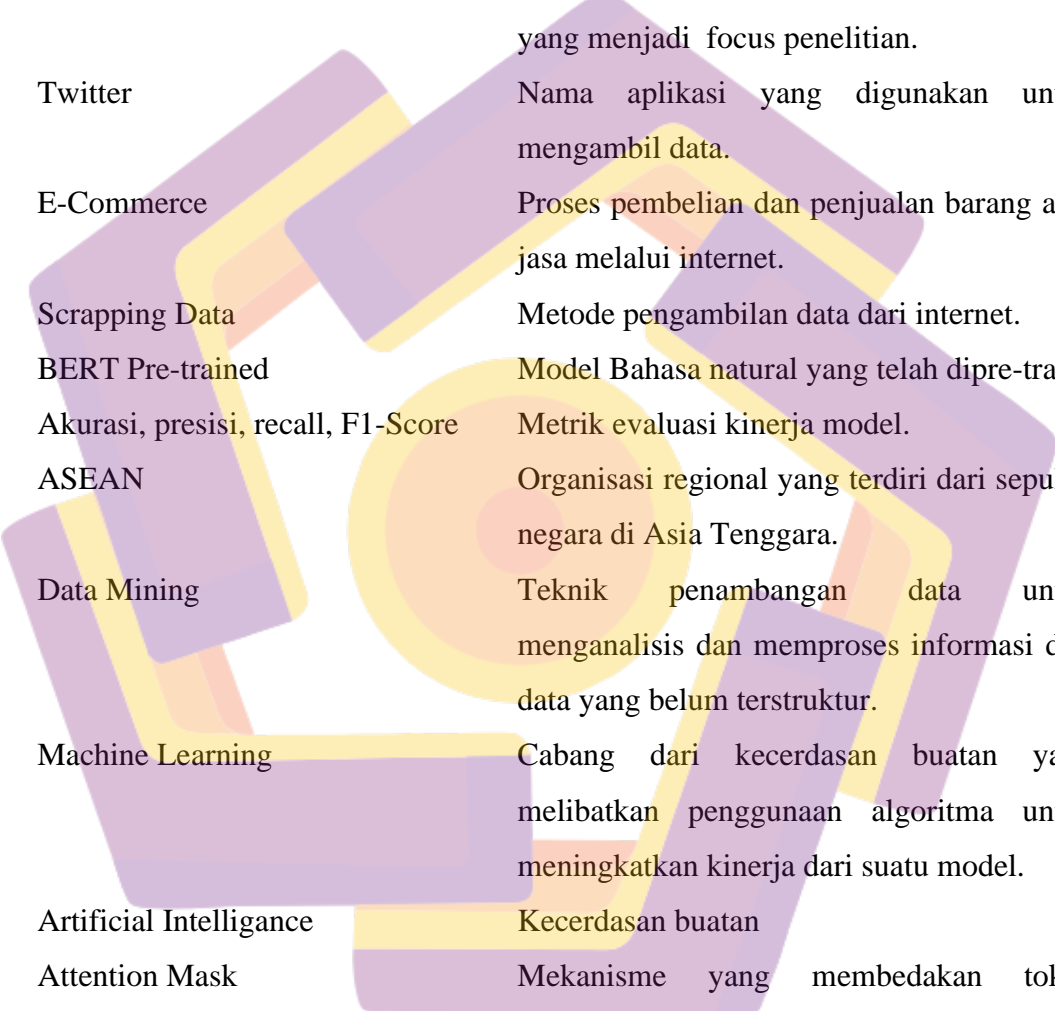


## DAFTAR LAMBANG DAN SINGKATAN



|            |   |
|------------|---|
| BERT       | Bidirectional Encoder Representations from Transformers |
| UMKM       | Usaha Mikro Kecil Menengah                              |
| AI         | Artificial Intelligence                                 |
| ASEAN      | Association of Southeast Asian Nations                  |
| PDB        | Produk Domestik Bruto                                   |
| E-Commerce | Electronic Commerce                                     |
| KDD        | Knowledge Discovery in Databases                        |
| ML         | Machine Learning  |
| NLP        | Natural Language Processing                             |
| TP         | True Positive   |
| TN         | True Negative   |
| FP         | False Positive  |
| FN         | False Negative  |
| API        | Application programming Interface                       |
| URL        | Uniform Resource Locator                                |
| HTML       | HyperText Markup Language                               |
| EDA        | Exploratory Data Analysis                               |

## DAFTAR ISTILAH



|                                    |   |
|------------------------------------|---|
| Tiktok Shop                        | Toko atau bisnis yang memanfaatkan platform TikTok untuk memasarkan dan menjual produk.                         |
| UMKM Lokal                         | Jenis bisnis berskala kecil hingga menengah yang menjadi focus penelitian.                                      |
| Twitter                            | Nama aplikasi yang digunakan untuk mengambil data.  |
| E-Commerce                         | Proses pembelian dan penjualan barang atau jasa melalui internet.   |
| Scrapping Data                     | Metode pengambilan data dari internet.  |
| BERT Pre-trained                   | Model Bahasa natural yang telah dipre-train.  |
| Akurasi, presisi, recall, F1-Score | Metrik evaluasi kinerja model.  |
| ASEAN                              | Organisasi regional yang terdiri dari sepuluh negara di Asia Tenggara.  |
| Data Mining                        | Teknik penambangan data untuk menganalisis dan memproses informasi dari data yang belum terstruktur.            |
| Machine Learning                   | Cabang dari kecerdasan buatan yang melibatkan penggunaan algoritma untuk meningkatkan kinerja dari suatu model. |
| Artificial Intelligence            | Kecerdasan buatan   |
| Attention Mask                     | Mekanisme yang membedakan token padding dari token valid selama proses attention.                               |



## INTISARI

Usaha Mikro, Kecil, dan Menengah (UMKM) memiliki peran penting dalam pertumbuhan ekonomi dan stabilitas. Lonjakan transaksi online, dipermudah oleh platform seperti TikTok Shop, ini menarik minat UMKM lokal untuk mempromosikan produk. Namun, penelitian tentang dampak TikTok Shop terhadap UMKM lokal masih terbatas. Penelitian ini menggunakan analisis data Twitter untuk memahami respons pengguna terhadap TikTok Shop yang dijalankan oleh UMKM lokal, dengan menggunakan algoritma deep learning BERT untuk analisis sentiment. Hasil evaluasi model menunjukkan akurasi total sebesar 79%, dengan *precision* untuk kelas netral (0) mencapai 86%, sedangkan untuk kelas positif (1) dan negatif (2) masing-masing adalah 54% dan 30%. Recall model menunjukkan kemampuan untuk mengidentifikasi data yang relevan, dengan kelas netral memiliki recall 92%, sementara kelas positif dan negatif hanya memiliki recall 38% dan 24%. F-1 Score, yang mencerminkan keseluruhan performa model, menunjukkan nilai tertinggi pada kelas netral 89%, diikuti dengan kelas positif 45% dan negatif 27%. Evaluasi ini menyoroti kebutuhan untuk meningkatkan pengenalan terhadap kelas positif dan negatif guna meningkatkan performa keseluruhan model.

**Kata kunci:** TikTok Shop, UMKM Lokal, Twitter, dan BERT.

## **ABSTRACT**

*Micro, Small, and Medium Enterprises (MSMEs) play a crucial role in economic growth and stability. The surge in online transactions facilitated by platforms like TikTok Shop has sparked interest among local MSMEs to promote their products. However, research on the impact of TikTok Shop on local MSMEs remains limited. This study utilizes Twitter data analysis to understand user responses to TikTok Shop operated by local MSMEs, employing the deep learning BERT algorithm for sentiment analysis. Model evaluation results show an overall accuracy of 79%, with precision for neutral class (0) reaching 86%, while for positive (1) and negative (2) classes it is 54% and 30%, respectively. Model recall demonstrates the ability to identify relevant data, with the neutral class achieving 92% recall, whereas positive and negative classes only achieve 38% and 24% recall, respectively. The F-1 Score, reflecting overall model performance, indicates the highest value in the neutral class at 89%, followed by 45% for positive and 27% for negative classes. This evaluation highlights the need to enhance recognition of positive and negative classes to improve overall model performance.*

**Keyword:** TikTok Shop, Local MSMEs, Twitter, and BERT.