PERBANDINGAN ANALISIS SENTIMEN MENGGUNAKAN METODE SVM, LOGISTIC REGRESSION, DAN LSTM
(Studi Kasus : Data Twitter Menantea)

SITI AULIA RIZKIKA
H1101181023

SKRIPSI

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS TANJUNGPURA
PONTIANAK
2025

PERBANDINGAN ANALISIS SENTIMEN MENGGUNAKAN METODE SVM, LOGISTIC REGRESSION, DAN LSTM (Studi Kasus : Data Twitter Menantea)

SITI AULIA RIZKIKA
H1101181023

Skripsi
Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Komputer pada Program Studi Sistem Informasi



PROGRAM STUDI SISTEM INFORMASI
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS TANJUNGPURA
PONTIANAK
2025

LEMBAR PENGESAHAN

Judul Tugas Akhir	: Perbandingan Analisis Sentimen Menggunakan Metode *SVM, Logistic Regression,* dan *LSTM* (Studi Kasus : Data Twitter Menantea)
Nama Mahasiswa	: Siti Aulia Rizkika
NIM	: H1101181023
Jurusan/Program Studi	: Sistem Informasi
Tanggal Lulus	: 31 Januari 2025
SK Pembimbing	: 85/UN22.8/TD.06/2023/Tanggl 03 Januari 2023
SK Penguji	: 225/UN22.8/TD.06/2025/Tanggal 24 Januari 2025

Dosen Pembimbing

Pembimbing I

Nurul Mutiah, S.T., M.T.
NIP. 198711182015042002

Pembimbing II

Syahru Rahmayuda, S.Kom., M.Kom.
NIDK. 8884370018

Dosen Penguji

Ketua Penguji

Renny Puspita Sari, S.T., M.T.
NIP. 198704182015042001

Anggota Penguji

Ibnur Rusi, S.Kom., M.M.
NIP. 198907282019031008

Pimpinan Sidang
(merangkap anggota penguji)

Nurul Mutiah, S.T., M.T.
NIP. 198711182015042002

Sekretaris Sidang
(merangkap anggota penguji)

Syahru Rahmayuda, S.Kom., M.Kom.
NIDK. 8884370018

Mengesahkan
Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Tanjungpura

Prof. Dr. Gusnedi, S.Si., M.Si.
NIP. 197108022000031001

LEMBAR PENGESAHAN

Judul Tugas Akhir	: Perbandingan Analisis Sentimen Menggunakan Metode *SVM, Logistic Regression,* dan *LSTM* (Studi Kasus : Data Twitter Menantea)
Nama Mahasiswa	: Siti Aulia Rizkika
NIM	: H1101181023
Jurusan/Program Studi	: Sistem Informasi
Tanggal Lulus	: 31 Januari 2025
SK Pembimbing	: 85/UN22.8/TD.06/2023/Tanggl 03 Januari 2023
SK Penguji	: 225/UN22.8/TD.06/2025/Tanggal 24 Januari 2025

Dosen Pembimbing

Pembimbing I	Pembimbing II
Nurul Mutiah, S.T., M.T. NIP. 198711182015042002	Syahru Rahmayuda, S.Kom., M.Kom. NIDK. 8884370018

Dosen Penguji

Ketua Penguji	Anggota Penguji
Renny Puspita Sari, S.T., M.T. NIP. 198704182015042001	Ibnur Rusi, S.Kom., M.M. NIP.198907282019031008
Pimpinan Sidang (merangkap anggota penguji)	Sekretaris Sidang (merangkap anggota penguji)
Nurul Mutiah, S.T., M.T. NIP. 198711182015042002	Syahru Rahmayuda, S.Kom., M.Kom. NIDK. 8884370018

Mengesahkan
Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Tanjungpura

Prof. Dr. Gusrizal, S.Si., M.Si.
NIP. 197108022000031001

PERNYATAAN INEGRITAS AKADEMIK

Saya yang bertanda tangan di bawah ini:

Nama	: Siti Aulia Rizkika
NIM	: H1101181023
Program Studi/ Jurusan	: Sistem Informasi
Fakultas	: Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan bahwa dokumen ilmiah Tugas Akhir yang disajikan ini tidak mengandung unsur pelanggaran integritas akademik sesuai Peraturan Menteri Pendidikan, Kebudayaan, Riset, Dan Teknologi Republik Indonesia Nomor 39 Tahun 2021. Apabila di kemudian hari dokumen ilmiah Tugas Akhir ini mengandung unsur pelanggaran integritas akademik sesuai ketentuan perundangan tersebut, maka saya bersedia menerima sanksi akademik dan/atau sanksi hukum yang berlaku.
Demikian pernyatan ini untuk digunakan sebagaimana mestinya.

Pontianak, 31 Januari 2025

Siti Aulia Rizkika
NIM H1101181023

PERBANDINGAN ANALISIS SENTIMEN MENGGUNAKAN METODE SVM, LOGISTIC REGRESSION, DAN LSTM (Studi Kasus : Data Twitter Menantea)

Abstrak

Menantea merupakan sebuah perusahaan yang bergerak dalam sector *Food and Beverage* dengan menyediakan minuman teh berkualitas tinggi dengan berbagai varian rasa yang inovatif. Dalam upaya untuk memahami preferensi konsumen, penelitian ini melakukan analisis sentimen dengan menggunakan tanggapan konsumen yang ada pada media social *twitter* (X). Media sosial *twitter* (X) sering digunakan untuk memberikan pendapat atau *review* terhadap suatu produk. Penelitian ini bertujuan untuk menganalisis sentimen pengguna *twitter* terhadap Menantea berupa sentiment positif, negatif, dan netral dengan menggunakan algoritma klasifikasi *Support Vector Machine, Logistic Regression,* dan *Long Short Term Memory* untuk melakukan perbandingan algoritma manakah yang memiliki performa terbaik dalam melakukan analisis sentimen. Data yang digunakan pada proses ini diambil menggunakan *Snscrape* pada Google Colaboratory dengan python. Hasil yang diperoleh yaitu pada algoritma *Support Vector Machine* memiliki nilai akurasi sebesar 62%, algoritma *Logistic Regression* sebesar 66%, dan pada algoritma *Long Short Term Memory* sebesar 99%.

Kata Kunci: Analisis Sentimen, Support Vector Machine, Logistic Regression, Long Short Term Memory, Twitter (X), Menantea

COMPARASION OF SENTIMENT ANALYSIS USING SVM, LOGISTIC REGRESSION, AND LSTM (Case Study : Menantea Twitter Data)

Abstract

Menantea is a company operating in the Food and Beverage sector, specializing in high-quality tea beverages with a variety of innovative flavors. In an effort to understand consumer preferences, this research conducts sentiment analysis using consumer responses found on the social media platform Twitter (X). Twitter (X) is frequently used for sharing opinions or reviews about products. This study aims to analyze Twitter users' sentiments towards Menantea, categorizing them into positive, negative, and neutral sentiments, by employing classification algorithms such as Support Vector Machine, Logistic Regression, and Long Short-Term Memory (LSTM) to compare which algorithm performs best in sentiment analysis. The data used in this process was collected using Snscrape on Google Colaboratory with Python. The results indicate that the Support Vector Machine algorithm achieved an accuracy of 62%, the Logistic Regression algorithm reached 66%, and the Long Short-Term Memory algorithm achieved an impressive accuracy of 99%.

Keywords: Sentiment Analysis, Support Vector Machine, Logistic Regression, Long Short Term Memory, Twitter (X), Menantea

KATA PENGANTAR

Alhamdulillah puji syukur peneliti hantarkan kepada Allah Subhanallahu Wa Ta'ala yang telah memberikan petunjuk dan karunia-Nya sehingga peneliti dapat menyelesaikan tugas akhir yang berjudul "Perbandingan Analisis Sentimen Menggunakan Metode SVM, Logistic Regression, dan LSTM" sebagai salah satu syarat untuk menyelesaikan Program Sarjana (S1) Sistem Informasi di Universitas Tanjungpura Pontianak.

Dalam kesempatan ini, peneliti ingin mengucapkan terima kasih kepada berbagai pihak yang telah memberikan dukungan, bantuan, bimbingan, dan nasehat selama penyusunan tugas akhir ini. Pada kesempatan ini peneliti menyampaikan terima kasih yang tulus kepada:

1. Sosok kepala keluarga yang hebat yaitu Ayahanda Juni Hariyadi. Terima kasih atas tetes keringat dalam setiap langkah mencari nafkah untuk memenuhi kebutuhan finansial penulis selama ini dan juga kesabaran menunggu peneliti untuk lulus. Pintu surgaku yaitu Ibunda Siti Rachmaniar. Terimakasih atas kasih sayang, ridho, motivasi, semangat, nasihat dan do'a yang tiada henti diucapkan untuk keberhasilan peneliti dalam menggapai cita-citanya.
2. Saudara saya yaitu Abang, Opal, dan Umay yang senantiasa berjasa atas segala usaha dalam memberikan dukungan secara finansial dan doa selama proses pengerjaan tugas akhir ini.
3. Ibu Nurul Mutiah, S.T., M.T. selaku Pembimbing Utama dan Bapak Syahru Rahmayuda, S.Kom., M.Kom. selaku Pembimbing Pendamping atas segala bimbingan, arahan, serta saran yang diberikan kepada peneliti sehingga tugas akhir ini dapat diselesaikan dengan baik.
4. Renny Puspita Sari, S.T., M.T. selaku Ketua Penguji dan Ibnur Rusi, S.Kom., M.M. selaku Anggota Penguji yang telah memberikan masukan dan saran pada penelitian tugas akhir ini.
5. Bapak dan Ibu dosen Program Studi Sistem Informasi yang telah memberikan ilmu dan pengetahuan selama proses perkuliahan yang tak

ternilai selama peneliti menempuh Pendidikan di jurusan Sistem Informasi Universitas Tanjungpura.

6. Salsabilah Tasya Oktaputri dan Nurul Lita yang selalu memotivasi, mendukung, dan menemani peneliti serta menghibur peneliti saat peneliti sedang berada di titik terendahnya beberapa bulan lalu. Terimakasih atas kehadirannya di hidup peneliti.
7. Ainun Nur Aini, sahabat sejak awal perkuliahan hingga saat ini, sahabat yang telah bersama-sama mengerjakan dan saling menyemangati dalam penyelesaian tugas akhir ini. Terimakasih untuk selalu menemani peneliti dalam suka maupun duka.
8. Teman-teman mahasiswa Sistem Informasi Angkatan 2018, terutama teman-teman residensi yang telah saling membantu dan saling memberikan semangat dalam mengerjakan tugas akhir.
9. Semua pihak yang tidak dapat disebutkan satu per satu, yang telah dengan tulus ikhlas memberikan do'a dan motivasi sehingga dapat terselesaikannya skripsi ini.

Semoga Allah Subhanallahu Wa Ta'ala membalas segala kebaikan yang telah semua pihak berikan dalam penyusunan tugas akhir ini. Penulis menyadari bahwa dalam penyusunan tugas akhir ini masih jauh dari kata sempurna, karena itu segala kritik dan saran yang membangun akan menyempurnakan penulisan tugas akhir ini serta bermanfaat bagi peneliti dan para pembaca.

DAFTAR ISI

Halaman

LEMBAR PENGESAHAN ii
PERNYATAAN INEGRITAS AKADEMIK iii
Abstrak iv
Abstract v
KATA PENGANTAR vi
DAFTAR ISI viii
DAFTAR TABEL x
DAFTAR GAMBAR xi
DAFTAR PERSAMAAN xii
DAFTAR LISTING CODE xiii
BAB I PENDAHULUAN 1
1.1 Pendahuluan 1
1.2 Rumusan Masalah 2
1.3 Tujuan Penelitian 3
1.4 Batasan Masalah 3
1.5 Manfaat Penelitian 4
1.6 Sistematika Penulisan 4
BAB II LANDASAN TEORI 6
2.1 Analisis Sentimen 6
2.2 Data Mining 6
2.3 Menantea 8
2.4 *Text Pre-processing* 9
2.5 *Machine Learning* 10
2.6 *Deep Learning* 11
2.7 *Support Vector Machine* 12
2.8 *Logistic Regression* 13
2.9 *Long Short-Term Memory* 14
2.10 *Google Colaboratory* 15
2.11 *Confusion Matrix* 15
2.12 TF – IDF 17
2.13 Twitter 17

2.14 Penelitian Terdahulu 18
BAB III METODOLOGI PENELITIAN 20
3.1 Kerangka Kerja Penelitian 20
3.2 Data Penelitian 21
3.2.1 Jenis Data 21
3.2.2 Metode Pengumpulan Data 21
3.3 Alur Penelitian 21
BAB IV HASIL DAN PEMBAHASAN 25
4.1 Pengambilan dan Pelabelan Data 25
4.1.1 *Crawling Data Twitter* 25
4.1.2 *Manual Labelling* 27
4.2 Pengolahan Dataset 29
4.2.1 *Pre-Processing Data* 29
4.2.2 Pembobotan Kata (TF-IDF) 40
4.2.3 Pembagian Data Training dan Data Testing 41
4.2.4 Proses SMOTE 43
4.3 Implementasi Klasifikasi 44
4.3.1 Classification Report 44
BAB V PENUTUP 56
5.1 Kesimpulan 56
5.2 Saran 56
DAFTAR PUSTAKA 58

DAFTAR TABEL

Tabel 4.1 Confusion Matrix 16
Tabel 4.2 Kajian Pustaka..... 20
Tabel 4.3 Variable Twitter Crawling 25
Tabel 4.4 Hasil Klasifikasi Manual Data Twit Menantea..... 28
Tabel 4.5 Hasil Sentimen Manual..... 29
Tabel 4.6 Perbandingan Menghapus link..... 29
Tabel 4.7 Twit yang Dihapus..... 30
Tabel 4. 8 Hasil Sentimen setelah Proses Remove Duplicate..... 31
Tabel 4.9 Perbandingan hasil proses Case Folding..... 32
Tabel 4.10 Perbandingan twit pada proses Tokenizing 33
Tabel 4.11 Perbandingan twit pada proses Normalisasi 33
Tabel 4.12 Perbandingan dari proses Stopword..... 34
Tabel 4.13 Perbandingan pada proses Stemming 37
Tabel 4.14 Perbandingan Hasil SMOTE Pada Data Training 44
Tabel 4.15 Perbandingan Nilai Akurasi..... 46
Tabel 4.16 Perbandingan Nilai Akurasi..... 48

DAFTAR GAMBAR

Gambar 2.1 Knowledge Discovery in Database (Fayyad et al., 1996) 7
Gambar 2.2 Akun twitter menantea 9
Gambar 2.3 Struktur neuron buatan 12
Gambar 2.4 Grafik Support Vector Machine Algorithm 13
Gambar 2.5 Cara kerja LSTM (Ahmad et al. 2021) 14
Gambar 3.1 IS Research Framework 20
Gambar 3. 2 Alur Penelitian..... 22
Gambar 3.3 Alur Penelitian (Lanjutan)..... 23
Gambar 4.1 File Data Menantea hasil Twitter Crawling 27
Gambar 4.2 Output Stemming 36
Gambar 4. 3 Output *Wordcloud* keseluruhan..... 37
Gambar 4.4 Output Wordcloud sentiment positif..... 38
Gambar 4.5 Output Wordcloud sentiment netral 39
Gambar 4.6 Output Wordcloud sentiment negative..... 40
Gambar 4.7 Hasil Proses TF-IDF..... 41
Gambar 4.8 Classification Report SVM 45
Gambar 4.9 Classification Report LR..... 47
Gambar 4.10 Confusion Matrix Support Vector Machine..... 50
Gambar 4.11 Confusion Matrix Logistic Regression 52
Gambar 4.12 Confusion Matrix Long Short Term Memory..... 54

DAFTAR PERSAMAAN

Persamaan 2.1 Persamaan Linear Support Vector Machine 13
Persamaan 2.2 Persamaan Linear Support Vector Machine Bernilai 1 ..... .13
Persamaan 2.3 Persamaan Linear Support Vector Machine Bernilai -1 ..... .13
Persamaan 2.4 Persamaan Logistic Regression 14
Persamaan 2.5 Persamaan Forget Gate LSTM 14
Persamaan 2.6 Persamaan Input Gate LSTM 15
Persamaan 2. 7 Persamaan Output Gate LSTM..... 15
Persamaan 2. 8 Persamaan Bobot Bias Sel LSTM 15
Persamaan 2. 9 Persamaan Output LSTM 15
Persamaan 2. 10 Persamaan Nilai Recall..... 16
Persamaan 2. 11 Persamaan Nilai Precision 16
Persamaan 2. 12 Persamaan Nilai Accuracy..... 17
Persamaan 2. 13 Persamaan Nilai F1-Score 17
Persamaan 2. 14 Persamaan TF-IIDF 17

DAFTAR LISTING CODE

Listing Code 4.1 Proses Install Module Snscrape 25
Listing Code 4.2 Proses Import Library Snscrape 26
Listing Code 4.3 Proses Penentuan Rentang Waktu 26
Listing Code 4.4 Pembuatan Data Frame 26
Listing Code 4.5 Proses Scrapping 27
Listing Code 4.6 Mengimport Regex dan String 29
Listing Code 4.7 Menampilkan Duplikasi Data 30
Listing Code 4.8 Menghapus Duplikasi 31
Listing Code 4.9 Proses Case Folding 31
Listing Code 4.10 Proses Import NLTK dan Tokenize 32
Listing Code 4.11 Proses Tokenizing pada Google Colab 32
Listing Code 4.12 Proses Normalisasi 33
Listing Code 4.13 Proses Stopword Removal 34
Listing Code 4.14 Penginstallan Library Sastrawi 35
Listing Code 4.15 Proses Stemming 35
Listing Code 4.16 Proses Stemming (Lanjutan) 36
Listing Code 4.17 Visualisasi Wordcloud keseluruhan 37
Listing Code 4.18 Visualisasi wordcloud sentiment positif 38
Listing Code 4.19 Visualisasi wordcloud sentiment netral 39
Listing Code 4.20 Visualisasi wordcloud sentiment negatif 39
Listing Code 4.21 Mengimpor TfidfVectorizer, ... 40
Listing Code 4.22 Penggunaan TfidfVectorizer 40
Listing Code 4.23 Listing Proses Pembagian Data Training dan 42
Listing Code 4.24 Listing Proses Pembagian Data Training dan 42
Listing Code 4.25 Listing Proses Pembagian Data Training dan 43
Listing Code 4.26 Proses Import SMOTE 43
Listing Code 4.27 Penerapan SMOTE Pada Data Training 43
Listing Code 4.28 Proses Mengimpor Support Vector Machine 44
Listing Code 4.29 Proses Pemodelan dengan Support Vector Machine 44
Listing Code 4.30 Pengaplikasian Data Testing pada Model SVM 44
Listing Code 4.31 Proses Classification Report 45
Listing Code 4.32 Proses Mengimpor Logistic Regression 46
Listing Code 4.33 Proses Pemodelan dengan Logistic Regression 46
Listing Code 4.34 Pengaplikasian Data Testing pada Model LR 46
Listing Code 4.35 Proses Classification Report Logistic Regression 47
Listing Code 4.36 Pemodelan Long Short Term Memory 48
Listing Code 4.37 Model Long Short Term Memory 49
Listing Code 4.38 Mengimpor Modul Untuk Confusion Matrix 49
Listing Code 4.39 Visualisasi Confusion Matrix Support Vector Machine 49
Listing Code 4.40 Visualisasi Confusion Matrix Logistic Regression 51
Listing Code 4.41 Confusion Matrix Long Short Term Memory 53

BAB I

PENDAHULUAN

1.1 Pendahuluan

Perusahaan food and beverage merupakan perusahaan yang bergerak di bidang industri makanan dan minuman. Perusahaan ini memiliki peluang untuk berkembang pesat dalam sektor industri di Bursa Efek Indonesia (BEI). Melihat kegemaran masyarakat Indonesia dalam menikmati makanan siap saji, banyak pelaku bisnis yang memanfaatkan kesempatan ini untuk membangun bisnis baru dibidang food and beverage. Pertumbuhan industri makanan dan minuman mencapai angka 3,57% di triwulan III tahun 2022, dimana angka tersebut memiliki kenaikan sebesar 0,08% dari tahun sebelumnya (Kemenperin, 2022).

Satu diantara perusahaan bisnis food and bevarage yang sedang berkembang saat ini adalah Menantea. Menantea merupakan perusahaan yang memiliki ciri khas produk yaitu modern fusion tea. Modern fusion tea merupakan minuman yang menggunakan komoditas utama teh dengan inovasi penambahan topping yang memiliki tujuan untuk menambah cita rasa yang unik, menarik, dan bervariasi. Satu diantara banyaknya inovasi yang dilakukan adalah dengan menambahkan buah-buahan, boba, cheese cream, jelly, dan lain sebagainya. Inovasi tersebut menarik banyak peminat, sehingga Menantea dapat mendirikan sebanyak 125 gerai di 56 kota yang tersebar di seluruh Indonesia (Yubilanti dkk, 2023). Strategi pemasaran yang dilakukan oleh Menantea adalah dengan memanfaatkan platform digital yaitu media sosial Twitter (X).

Penggunaan media sosial saat ini sangat tinggi, hal tersebut dapat dilihat dari artikel yang dikeluarkan oleh dateportal, mengenai hasil survey yang dilakukan oleh We Are Social pada Februari 2022, Indonesia memiliki 191,4 juta pengguna media sosial. Satu diantara platform media sosial yang banyak digunakan untuk mengekspresikan diri adalah Twitter. Menurut data yang ada pada laporan We Are Social, jumlah pengguna aktif Twitter di Indonesia mencapai 21,2 juta pada July 2022. Kemudahan akses twitter membuat pengguna media sosial tersebut menjadi platform yang sangat banyak digunakan pengguna

untuk bertanya ataupun memberikan pendapat. Sentimen atau opini konsumen yang berasal dari sosial media twitter dapat dimanfaatkan untuk mengetahui feedback dan menjadi bahan evaluasi terhadap produk yang ditawarkan. Pengambilan data yang akan digunakan sebagai dataset, diperoleh dengan menerapkan data mining. Husin dan Mulyaningsih (2015) menyebutkan bahwa, data mining merupakan teknologi yang dapat dimanfaatkan oleh perusahaan dalam menemukan informasi penting dari database yang selama ini hanya digunakan sebagai arsip perusahaan. Oleh karena itu, data mining dapat digunakan dalam melakukan analisis sentimen.

Dataset yang telah diperoleh akan dibagi menjadi beberapa kelas melalui analisis sentimen. Analisis sentimen merupakan suatu studi komputasional yang berasal dari opini orang, sentimen, dan emosi melalui atribut yang dapat diekspresikan dalam bentuk teks (Liu, 2012). Dilihat dari pengertian tersebut, dataset akan dilakukan klasifikasi dengan membagi opini masyarakat menjadi opini positif, netral, dan negatif. Pengolahan data tersebut dapat memanfaatkan algoritma klasifikasi. Algoritma klasifikasi terdiri atas banyak metode, namun yang menjadi fokus pada penelitian ini adalah SVM, Logistic Regression, dan LSTM. Digunakannya metode ini adalah untuk mengetahui perbandingan antara metode Support Vector Machine, Logistic Regression, dan Long Short Term Memory yang lebih tepat dalam melakukan analisis sentimen pada teks pendek yang ada di twitter. Diharapkan dengan dilakukannya analisis sentimen ini dapat membantu perusahaan dalam mengetahui pendapat konsumen terhadap produk yang ditawarkan melalui tiga kelompok yaitu positif, netral, dan negatif untuk memperbaiki produk kedepannya.

Dari latar belakang yang telah dijelaskan, penulis tertarik untuk melakukan penelitian yang berjudulkan “Perbandingan Analisis Sentimen Menggunakan Metode SVM, Logistic Regression, Dan LSTM (Studi Kasus : Data Twitter Menantea)”.

1.2 Rumusan Masalah

Dari latar belakang yang sudah dijelaskan sebelumnya didapatkan rumusan

permasalahan yang ada, yaitu :

1. Bagaimana cara menilai pendapat masyarakat terhadap produk Menantea berdasarkan komentar konsumen pada media sosial Twitter?
2. Bagaimana menentukan metode terbaik yang digunakan untuk menilai ketepatan sentiment analisis pada produk Menantea?

1.3 Tujuan Penelitian

Dari rumusan masalah diatas penelitian ini memiliki tujuan sebagai berikut, yaitu :

1. Melakukan analisis sentiment yang dapat dimanfaatkan untuk menilai umpan balik dan bahan evaluasi perbaikan untuk produk Menantea.
2. Melakukan perbandingan metode Support Vector Machine, Logistic Regression, dan Long Short Term Memory untuk menentukan nilai akurasi dari tiga metode tersebut.

1.4 Batasan Masalah

Agar penelitian ini tidak meluas, maka diberikan batasan masalah sebagai berikut:

1. Data ulasan yang diambil pada penelitian hanya menggunakan bahasa Indonesia.
2. Metode klasifikasi yang digunakan adalah *Support Vector Machine, Logistic Regression,* dan *Long Short Term Memory*.
3. Ulasan konsumen terbagi atas tiga kelas klasifikasi yaitu positif, negatif, dan netral.
4. Data yang akan digunakan pada penelitian diambil dari bulan Januari 2023 sampai bulan Juni 2023.
5. Pengambilan data dilakukan dengan menggunakan *library Snscrape.*
6. Tweet yang diambil menggunakan kata kunci “Menantea” , “Scam Menantea” , dan “Menantea Jerome”.
7. Epoch yang digunakan pada metode klasifikasi *Long Short Term Memroy* adalah 50.

1.5 Manfaat Penelitian

Adapun manfaat yang diharapkan dari penelitian yang dilakukan yaitu:

a. Bagi peneliti

   Dengan dilakukannya penelitian ini, peneliti mengharapkan dapat menerapkan ilmu yang telah didapatkan selama mengikuti perkuliahan dan membuat peneliti lebih mendalami lagi bidang keilmuan yang dipilih khususnya pada bidang data mining.

b. Bagi pembaca

   Dengan adanya penelitian yang dilakukan, peneliti mengharapkan bahwa kedepannya hasil dari penelitian ini dapat menjadi manfaat terhadap pengetahuan pembaca mengenai hasil perbandingan nilai akurasi dari beberapa algoritma yang diterapkan pada opini para pengguna Twitter.

c. Bagi bidang keilmuan

   Dengan dilakukannya penelitian ini, diharapkan dapat menjadi referensi bagi penelitian selanjutnya mengenai analisis sentimen review produk dengan menggunakan klasifikasi *Support Vector Machine, Logistic Regression,* dan *Long Short Term Memory.*

1.6 Sistematika Penulisan

Dalam bagian sistematika penulisan ini peneliti membuat beberapa bab yang ditujukan agar memudahkan peneliti dalam melakukan penelitian secara urut dan sistematis yang dibagi menjadi lima bab yakni sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini menggambarkan isi penelitian secara keseluruhan yang berisikan hal mengenai latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, dan sistematika penulisan tugas akhir "Analisis Sentimen Menggunakan Metode *SVM, Logistic Regression,* dan *LSTM*".

BAB II LANDASAN TEORI

Bab ini membahas mengenai perihal dasar teori yang memiliki relevansi terhadap penelitian seperti analisis sentiment, *Support Vector Machine, Logistic Regression,* dan *Long Short Term* Memory dan tinjauan pustaka sebagai pembanding antara penelitian yang sudah dilakukan sebelumnya terhadap penelitian yang dilakukan saat ini.

BAB III METODOLOGI PENELITIAN

Pada bab ini mendeskripsikan proses penelitian dan tahapan alur penelitian yang akan digunakan oleh peneliti seperti kerangka kerja penelitian menggunakan *Is Research Framework* dan alur penelitian yang dilakukan seperti proses pengambilan data dan pelabelan dataset, pengolahan dataset, implementasi metode *SVM, LR,* dan *LSTM*, dan evaluasi model yang dilakukan agar mendapatkan hasil yang sesuai dengan tujuan penelitian terhadap permasalahan yang ada.

BAB IV HASIL DAN PEMBAHASAN

Di bab ini dilakukan pengambilan dan pelabelan dataset, pengolahan dataset dengan menerapkan *pre-processing*, penerapan model klasifikasi *SVM, Logistic Regression*, dan *LSTM* dengan hasil nilai akurasi dan dilakukan pembahasan metode manakah yang memiliki performa terbaik dalam analisis sentiment.

BAB V PENUTUP

Bab penutup ini berisi tentang kesimpulan dari bab-bab pada penelitian yang telah dibahas sebelumnya dan saran yang diberikan terhadap penelitian yang akan dilakukan selanjutnya.