ANALISIS SENTIMEN TERHADAP SHOPEE AFFILIATE MENGGUNAKAN NAÏVE BAYES CLASSIFIER DAN SELEKSI FITUR PARTICLE SWARM OPTIMIZATION

SRI RAHAYU RAHMASIWI
H1101181039

SKRIPSI

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

ANALISIS SENTIMEN TERHADAP SHOPEE AFFILIATE MENGGUNAKAN NAÏVE BAYES CLASSIFIER DAN SELEKSI FITUR PARTICLE SWARM OPTIMIZATION

SRI RAHAYU RAHMASIWI
H1101181039

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	: ANALISIS SENTIMEN TERHADAP SHOPEE AFFILIATE MENGGUNAKAN NAÏVE BAYES CLASSIFIER DAN SELEKSI FITUR PARTICLE SWARM OPTIMIZATION
Nama Mahasiswa	: Sri Rahayu Rahmasiwi
NIM	: H1101181039
Jurusan/Program Studi	: Sistem Informasi
Tanggal Lulus	: 22 Januari 2025
SK Pembimbing	: No. 3451/UN22.8/TD.06/2023 /Tanggal 14 September 2023
SK Penguji	: No. 5657/UN22.8/TD.06/2024 /Tanggal 31 Desember 2024

Dosen Pembimbing

Pembimbing I	Pembimbing II
Dian Prawira, S.T., M.Eng.	Syahru Rahmayuda, S.Kom., M.Kom
NIP. 198411132015041001	NIDK. 8884370018

Dosen Penguji

Ketua Penguji	Anggota Penguji
Nurul Mutiah, S.T., M.T.	Ibnur Rusi, S.Kom., M.M.
NIP. 198711182015042002	NIP. 198907282019031008

Pimpinan Sidang (merangkap anggota penguji)	Sekretaris Sidang (merangkap anggota penguji)
Dian Prawira, S.T., M.Eng.	Syahru Rahmayuda, S.Kom., M.Kom.
NIP. 198411132015041001	NIDK. 8884370018

Mengesahkan
Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Tanjungpura

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

PERNYATAAN INTEGRITAS AKADEMIK

Saya yang bertanda tangan di bawah ini:

Nama : Sri Rahayu Rahmasiwi

NIM : H1101181039

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 pernyataan ini untuk digunakan sebagaimana mestinya.

Pontianak, 17 Januari 2025



Sri Rahayu Rahmasiwi

NIM H1101181039

ANALISIS SENTIMEN TERHADAP SHOPEE AFFILIATE MENGGUNAKAN NAÏVE BAYES CLASSIFIER DAN SELEKSI FITUR PARTICLE SWARM OPTIMIZATION

Abstrak

Affiliate marketing semakin populer di lingkungan *marketplace*. Salah satu perusahaan yang menggunakan teknik ini sebagai strategi *marketing* adalah Shopee. Dengan mendaftar Program Shopee Affiliate, afiliator akan mendapat komisi dari setiap pembelian produk melalui link yang dibagikan. Program ini berhasil menempati posisi teratas dalam jumlah transaksi dan kepuasan afiliator pada tahun 2023. Meskipun menawarkan keuntungan finansial, terdapat berbagai tanggapan dari pengguna media sosial seperti pada Twitter (X) yang mengeluhkan penyebaran *link* produk Shopee. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap Program Shopee Affiliate, apakah positif, negatif, atau netral melalui data opini pengguna Twitter (X), guna memberikan wawasan mengenai persepsi publik dan mempertahankan posisi pasar. *Crawling data* menggunakan Tweet Harvest pada Google Colaboratory dengan Python dan memperoleh data sebanyak 3523 *tweet*. Metode yang digunakan adalah Naïve Bayes Classifier dan Particle Swarm Optimization (PSO) sebagai teknik seleksi fitur untuk meningkatkan akurasi. Hasil penelitian menunjukkan bahwa opini publik cenderung positif pada Program Shopee Affiliate dengan jumlah data sebesar 40,56%. Pengujian algoritma juga menunjukkan model Naïve Bayes tanpa seleksi fitur mencapai akurasi 72,21% dengan AUC 0,857, sementara penerapan seleksi fitur dengan Particle Swarm Optimization meningkatkan akurasi menjadi 79,66% dan AUC menjadi 0,908. Peningkatan akurasi sebesar 7,45% ini menunjukkan bahwa penggunaan PSO efektif dalam meningkatkan kinerja Naïve Bayes dalam klasifikasi sentimen terkait Program Shopee Affiliate.

Kata Kunci: Analisis Sentimen, Naïve Bayes, Particle Swarm Optimization, Shopee Affiliate, Twitter (X)

SENTIMENT ANALYSIS ON SHOPEE AFFILIATE USING NAÏVE BAYES CLASSIFIER AND PARTICLE SWARM OPTIMIZATION FOR FEATURE SELECTION

Abstract

Affiliate marketing is increasingly popular in the marketplace environment. One company that uses this technique as a marketing strategy is Shopee. Affiliates who join the Shopee Affiliate Program earn a commission for each product purchased through a shared link. The program managed to take the top spot in transaction numbers and affiliate satisfaction in 2023. Despite offering financial benefits, there were various responses from social media users, such as on Twitter (X), who complained about the spread of Shopee product links. To gain insight into public perception and maintain market position, this study analyzes Twitter (X) user opinions to assess public sentiment about the Shopee Affiliate Program—whether positive, negative, or neutral. Crawling data using Tweet Harvest on Google Collaboratory with Python and retrieving 3523 tweets. The methods used are Naïve Bayes Classifier and Particle Swarm Optimization (PSO) as a feature selection technique to improve accuracy. The results show that public opinion on the Shopee Affiliate Program is generally positive, with 40,56% of the data reflecting this sentiment. Algorithm testing revealed that the Naïve Bayes model achieved 72,21% accuracy with an AUC of 0,857. Using feature selection with PSO increased accuracy to 79,66% with an AUC of 0,908. PSO effectively enhances the Naïve Bayes model's performance in sentiment classification for the Shopee Affiliate Program, resulting in a 7,45% increase in accuracy.

Keywords: *Sentiment Analysis, Naïve Bayes, Particle Swarm Optimization, Shopee Affiliate, Twitter (X)*

PRAKATA

Puji dan syukur atas kehadirat Allah SWT atas segala nikmat dan karunia-Nya, penulis dapat menyelesaikan Tugas Akhir yang berjudul "Analisis Sentimen Terhadap Shopee Affiliate Menggunakan Naïve Bayes Classifier dan Seleksi Fitur Particle Swarm Optimization" dengan baik.

Penulisan Tugas Akhir ini tidak terlepas dari bantuan pihak-pihak yang memberi bantuan, kritik, saran, dan dukungan moral maupun materil kepada penulis. Oleh karena itu pada kesempatan ini, penulis ingin mengucapkan terima kasih yang tidak terhingga kepada:

1. Kedua orang tua tercinta dan keluarga besar yang selalu memberikan doa, dukungan, dan motivasi tanpa henti selama proses penulisan
2. Bapak Dian Prawira, S.T., M.Eng. selaku Dosen Pembimbing I yang telah memberikan bimbingan, kritik, dan saran sehingga Tugas Akhir ini dapat diselesaikan.
3. Bapak Syahru Rahmayuda, S.Kom., M.Kom. selaku Dosen Pembimbing II yang telah memberikan bimbingan, kritik, dan saran sehingga Tugas Akhir ini dapat diselesaikan.
4. Ibu Nurul Mutiah, S.T., M.T. selaku Dosen Penguji I yang telah memberikan masukan dan saran pada penelitian Tugas Akhir ini.
5. Bapak Ibnur Rusi, S.Kom., M.M. selaku Dosen Penguji II yang telah memberikan masukan dan saran pada penelitian Tugas Akhir ini.
6. Bapak Prof. Dr. Garuda Wiko, S.H., M.Si. selaku Rektor Universitas Tanjungpura Pontianak.
7. Bapak Prof. Dr. Gusrizal, S.Si., M.Si., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Tanjungpura Pontianak.
8. Bapak dan Ibu Dosen serta Staf Jurusan Sistem Informasi yang telah memberikan ilmu dan pengetahuan selama perkuliahan.
9. Sahabat seperjuangan yaitu Jihadini Rahmi, yang selalu memberi dukungan dan semangat serta menjadi tempat berbagi keluh kesah selama penyusunan Tugas Akhir ini.

10. Sahabat-sahabat penulis yaitu Kharisma Ananda Pertiwi, Rindu Cipta Gita, Dewi Andini, Lipia Rahmadini, dan (Almh.) Andy Nurul Azmy yang senantiasa memberikan dukungan, semangat, bantuan, dan selalu ada selama masa-masa sulit.
11. Rekan-rekan kerja di SMP Negeri 12 Sungai Raya yang telah memberikan dukungan dan pengertian selama penulis menyelesaikan Tugas Akhir ini.
12. Teman-teman Sistem Informasi Angkatan 2018 yang selalu mendukung dan berjuang bersama.
13. Semua pihak yang tidak dapat disebutkan satu per satu yang memberikan dukungan selama proses penyusunan Tugas Akhir ini.

Penulis menyadari dalam pelaksanaan dan penulisan Tugas Akhir ini masih terdapat banyak kekurangan, oleh karena itu penulis mengharapkan kritik dan saran yang membangun demi kesempurnaan penelitian ini. Semoga Tugas Akhir ini dapat bermanfaat bagi semua pihak yang membacanya dan menjadi sumbangan kecil bagi perkembangan ilmu pengetahuan.

Pontianak, 17 Januari 2025

penulis

DAFTAR ISI

Halaman

LEMBAR PENGESAHAN ..... **ii**
PERNYATAAN INTEGRITAS AKADEMIK ..... **iii**
Abstrak ..... **iv**
Abstract ..... **v**
PRAKATA ..... **vi**
DAFTAR ISI ..... **viii**
DAFTAR TABEL ..... **x**
DAFTAR GAMBAR ..... **xi**
DAFTAR PERSAMAAN ..... **xii**
DAFTAR LISTING CODE ..... **xiii**
DAFTAR LAMPIRAN ..... **xvi**

BAB I PENDAHULUAN **1**
1.1 Latar Belakang 1
1.2 Rumusan Masalah 3
1.3 Tujuan Penelitian 3
1.4 Batasan Masalah 3
1.5 Manfaat Penelitian 4
1.6 Sistematika Penulisan 4

BAB II TINJAUAN PUSTAKA **6**
2.1 Data Mining 6
2.2 Analisis Sentimen 7
2.3 Twitter (X) 8
2.4 Tweet Harvest 9
2.5 Pelabelan Data 10
2.6 Program Shopee Affiliate 10
2.7 *Knowledge Discovery in Database* (KDD) 11
2.8 Pre-Processing 13
2.9 TF-IDF 14
2.10 Synthetic Minority Oversampling Technique (SMOTE) 15
2.11 Naïve Bayes Classifier 16
2.12 Seleksi Fitur 17
2.13 Particle Swarm Optimization 18
2.14 Evaluasi Model Klasifikasi 20
2.15 Python 22
2.16 Google Colaboratory 23
2.17 Wordcloud 23
2.18 Penelitian Terdahulu 23

BAB III METODOLOGI PENELITIAN 26
3.1 Kerangka Penelitian..... 26
3.2 Data Penelitian..... 27
3.3 Alur Penelitian..... 27

BAB IV HASIL DAN PEMBAHASAN..... 32
4.1 Pengambilan dan Pelabelan Data 32
4.2 Pengolahan Dataset 41
4.3 Seleksi Fitur..... 74
4.4 Implementasi Klasifikasi Naïve Bayes..... 88
4.5 Evaluasi Model Klasifikasi..... 90
4.6 Analisis..... 106

BAB V PENUTUP..... 109
5.1. Kesimpulan..... 109
5.2. Saran..... 109

DAFTAR PUSTAKA..... 111

LAMPIRAN..... 116

DAFTAR TABEL

Halaman

Tabel 2. 1 Confusion Matrix 20
Tabel 2. 2 Penelitian Terdahulu 24
Tabel 3. 1 Skenario Pembagian Dataset 30
Tabel 4. 1 Perbandingan Tweet Setelah Diterjemahkan 36
Tabel 4. 2 Contoh Hasil Pelabelan Menggunakan Textblob 38
Tabel 4. 3 Hasil Sentimen Menggunakan Textblob 38
Tabel 4. 4 Hasil Sentimen Secara Manual 39
Tabel 4. 5 Membandingkan Hasil Label Texblob dan Label Manual 39
Tabel 4. 6 Hasil Perbandingan Label Texblob dan Label Manual 40
Tabel 4. 7 Contoh Tweet Yang Terdeteksi Sebagai Duplikat 42
Tabel 4. 8 Jumlah Sentimen Setelah Proses Data Duplicate Removal 43
Tabel 4. 9 Perbandingan Hasil Case Folding 43
Tabel 4. 10 Perbandingan Menghapus New Line dan Unicode String 45
Tabel 4. 11 Perbandingan Menghapus Karakter Non-ASCII (Emoji) 46
Tabel 4. 12 Perbandingan Menghapus Mention, Link, dan Hashtag 47
Tabel 4. 13 Perbandingan Menghapus Angka 48
Tabel 4. 14 Perbandingan Menghapus Tanda Baca 49
Tabel 4. 15 Perbandingan Hasil Mengubah Spasi Lebih Dari Satu 50
Tabel 4. 16 Perbandingan Hasil Menghapus Single Character 51
Tabel 4. 17 Perbandingan Hasil Proses Cleansing Data 51
Tabel 4. 18 Perbandingan Hasil Proses Tokenisasi 52
Tabel 4. 19 Perbandingan Hasil Proses Normalization of Slang Word 53
Tabel 4. 20 Perbandingan Hasil Proses Filtering 55
Tabel 4. 21 Perbandingan Hasil Proses Stemming 56
Tabel 4. 22 Data Tweet Pada Dokumen Ke-0 71
Tabel 4. 23 Hasil Skenario Pembagian Dataset 73
Tabel 4. 24 Perbandingan Hasil SMOTE Pada Data Training 74
Tabel 4. 25 Hasil Pengujian Skenario Pembagian Dataset Naive Bayes 91
Tabel 4. 26 Hasil Pengujian Skenario Pembagian Dataset Naive Bayes dan PSO 94
Tabel 4. 27 Hasil Eksperimen Naive Bayes dan PSO Dengan Nilai Parameter Population Size 5-50 95
Tabel 4. 28 Hasil Eksperimen Naive Bayes dan PSO Dengan Nilai Parameter Maximun Number of Iteration 10-100 96
Tabel 4. 29 Hasil Eksperimen Naive Bayes dan PSO Dengan Nilai Parameter Inertia Weight 0,1-1,0 96
Tabel 4. 30 Jumlah Data Per Sentimen 106
Tabel 4. 31 Nilai Akurasi, Presisi, Recall, F1-Score, dan AUC Model Klasifikasi 107
Tabel 4. 32 Nilai Weighted Average Model Klasifikasi 107

DAFTAR GAMBAR

Halaman

Gambar 2. 1 Tahapan Knowledge Discovery in Database 12
Gambar 2. 2 Visualisasi Saat Populasi Mencari Posisi Terbaik 18
Gambar 2. 3 Grafik ROC 22
Gambar 3. 1 Kerangka Kerja Penelitian 26
Gambar 3. 2 Alur Penelitian..... 27
Gambar 4. 1 Hasil Crawling dengan Kata Kunci Shopee Affiliate 34
Gambar 4. 2 Hasil Crawling dengan Kata Kunci Link Shopee Affiliate..... 35
Gambar 4. 3 Dataset Shopee Affiliate Setelah Manipulasi Data 35
Gambar 4. 4 Diagram Persentase Perbandingan Hasil Label 40
Gambar 4. 5 Kamus Normalization of Slang Word..... 53
Gambar 4. 6 Wordcloud Keseluruhan Data 58
Gambar 4. 7 Wordcloud Sentimen Negatif..... 60
Gambar 4. 8 Wordcloud Sentimen Netral..... 61
Gambar 4. 9 Wordcloud Sentimen Positif 63
Gambar 4. 10 Visualisasi Distribusi Sentimen Dalam Bentuk Pie Chart 64
Gambar 4. 11 Visualisasi Frekuensi Kata Untuk Sentimen Negatif..... 66
Gambar 4. 12 Visualisasi Frekuensi Kata Untuk Sentimen Netral..... 67
Gambar 4. 13 Visualisasi Frekuensi Kata Untuk Sentimen Positif 68
Gambar 4. 14 Hasil Proses TF-IDF..... 70
Gambar 4. 15 Hasil Get Feature Names Out..... 71
Gambar 4. 16 Tampilan Hasil TF-IDF Menggunakan Pandas 72
Gambar 4. 17 Output Penskalaan Fitur Dengan MinMaxScaler..... 76
Gambar 4. 18 Confusion Matrix Naive Bayes..... 92
Gambar 4. 19 Classification Report Model Naive Bayes 93
Gambar 4. 20 Confusion Matrix NBC Dengan PSO 98
Gambar 4. 21 Classification Report NBC Dengan PSO..... 99
Gambar 4. 22 Kurva ROC Model Naïve Bayes Classifier..... 102
Gambar 4. 23 Kurva ROC Model NBC Dengan Seleksi Fitur PSO..... 105

DAFTAR PERSAMAAN

Halaman

Persamaan 2.1 Rumus Term Frequency (TF) 15
Persamaan 2.2 Rumus Inverse Document Frequency (IDF)..... 15
Persamaan 2.3 Rumus Bobot TF-IDF..... 15
Persamaan 2.4 Rumus Naïve Bayes Classifier 16
Persamaan 2.5 Rumus Vmap Naïve Bayes Classifier..... 17
Persamaan 2.6 Rumus Probabilitas Prior Naïve Bayes Classifier 17
Persamaan 2.7 Rumus Probabilitas Bersyarat Naïve Bayes Classifier..... 17
Persamaan 2.8 Rumus Kecepatan Awal Particle Swarm Optimization..... 19
Persamaan 2.9 Rumus Memperbarui Kecepatan Partikel PSO..... 19
Persamaan 2.10 Rumus Memperbarui Posisi Partikel PSO 20
Persamaan 2.11 Rumus Accuracy..... 21
Persamaan 2.12 Rumus Presisi 21
Persamaan 2.13 Rumus Recall..... 21
Persamaan 2.14 Rumus F1-Score 21

DAFTAR LISTING CODE

Halaman

Listing Code 4. 1 Persiapan Kebutuhan Untuk Crawling Data 32
Listing Code 4. 2 Crawling Data dengan Kata Kunci Shopee Affiliate 33
Listing Code 4. 3 Crawling Data dengan Kata Kunci Link Shopee Affiliate..... 33
Listing Code 4. 4 Proses Penerjemahan Data Ke Dalam Bahasa Inggris 36
Listing Code 4. 5 Import Textblob..... 37
Listing Code 4. 6 Pengaplikasian Textblob Pada Data 37
Listing Code 4. 7 Mengimport Library Pandas..... 41
Listing Code 4. 8 Menampilkan Data Duplikat 41
Listing Code 4. 9 Menghapus Data Duplikat..... 42
Listing Code 4. 10 Proses Case Folding 43
Listing Code 4. 11 Mengimport String 44
Listing Code 4. 12 Mengimport Regex (Regular Expression)..... 44
Listing Code 4. 13 Menghapus Karakter Spesial Pada Tweet 44
Listing Code 4. 14 Menghapus New Line dan Unicode String 45
Listing Code 4. 15 Menghapus Karakter Non-ASCII (emoji)..... 45
Listing Code 4. 16 Menghapus Mention, Link, dan Hashtag 46
Listing Code 4. 17 Menghapus Angka Pada Data Tweet 47
Listing Code 4. 18 Menghapus Tanda Baca 48
Listing Code 4. 19 Menghapus Spasi Di Awal dan Akhir Kalimat 49
Listing Code 4. 20 Mengubah Spasi Lebih Dari Satu..... 49
Listing Code 4. 21 Menghapus Single Character 50
Listing Code 4. 22 Import Library NLTK dan Fungsi Tokenize..... 51
Listing Code 4. 23 Proses Tokenisasi Data..... 52
Listing Code 4. 24 Proses Normalization of Slang Word..... 53
Listing Code 4. 25 Proses Filtering..... 54
Listing Code 4. 26 Instalasi Library Sastrawi 55
Listing Code 4. 27 Instalasi Swifter..... 55
Listing Code 4. 28 Proses Stemming 55
Listing Code 4. 29 Mengimpor Library Untuk Wordcloud 57
Listing Code 4. 30 Import Dataset Untuk Wordcloud 57
Listing Code 4. 31 Proses Mengekstrak dan Menggabungkan Tweet Untuk Wordcloud..... 57
Listing Code 4. 32 Proses Pembuatan Wordcloud..... 58
Listing Code 4. 33 Pembuatan Wordcloud Untuk Sentimen Negatif 59
Listing Code 4. 34 Pembuatan Wordcloud Untuk Sentimen Netral 61
Listing Code 4. 35 Pembuatan Wordcloud Untuk Sentimen Positif..... 62
Listing Code 4. 36 Pembuatan Visualisasi Distribusi Sentimen..... 64
Listing Code 4. 37 Menghapus Kata Kunci Untuk Visualisasi Frekuensi Kata ... 65
Listing Code 4. 38 Pembuatan Visualisasi Frekuensi Kata Untuk Sentimen Negatif..... 65
Listing Code 4. 39 Pembuatan Visualisasi Frekuensi Kata Untuk Sentimen Netral 67

Listing Code 4. 40 Pembuatan Visualisasi Frekuensi Kata Untuk Sentimen Positif 68
Listing Code 4. 41 Mengimpor TfidfVectorizer 69
Listing Code 4. 42 Pengaplikasian TvidfVectorizer Pada Dataset 69
Listing Code 4. 43 Urutan Kata Dengan Get Feature Names 70
Listing Code 4. 44 Mengubah Tampilan Hasil TF-IDF Dengan Pandas 71
Listing Code 4. 45 Proses Pembagian Data Training dan Data Testing 72
Listing Code 4. 46 Proses Mengimport SMOTE 73
Listing Code 4. 47 Penerapan SMOTE Pada Data Training 74
Listing Code 4. 48 Konversi x_train dan x_test menjadi Data Integer 75
Listing Code 4. 49 Mengubah Bentuk Data Menjadi Array NumPy Pada Data SMOTE 75
Listing Code 4. 50 Mengimpor MinMaxScaler 76
Listing Code 4. 51 Penskalaan Fitur Dengan MinMaxScaler 76
Listing Code 4. 52 Error Rate Function 77
Listing Code 4. 53 Fun Function Pada Seleksi Fitur 77
Listing Code 4. 54 Function Untuk Inisialisasi Posisi 79
Listing Code 4. 55 Function Inisialisasi Kecepatan Partikel 79
Listing Code 4. 56 Function Konversi Biner 80
Listing Code 4. 57 Boundary Function Pada Seleksi Fitur 81
Listing Code 4. 58 Inisialisasi Parameter Pada Function jfs 82
Listing Code 4. 59 Dimensi Pada Seleksi Fitur 83
Listing Code 4. 60 Inisialisasi Posisi dan Kecepatan Pada Seleksi Fitur 83
Listing Code 4. 61 Inisialisasi Variabel Algoritma PSO 84
Listing Code 4. 62 Implementasi Algoritma PSO Untuk Pemilihan Fitur 84
Listing Code 4. 63 Menentukan Subset Fitur Terbaik 86
Listing Code 4. 64 Proses Pemilihan Fitur Dengan Algoritma PSO 87
Listing Code 4. 65 Proses Mengimpor Naive Bayes 88
Listing Code 4. 66 Proses Pemodelan Dengan Naive Bayes Classifier 88
Listing Code 4. 67 Pengaplikasian Data Testing Pada Model Naïve Bayes 88
Listing Code 4. 68 Persiapan Dataset Yang Telah Dilakukan Seleksi Fitur 89
Listing Code 4. 69 Pemodelan Naive Bayes Dengan Seleksi Fitur 89
Listing Code 4. 70 Pengaplikasian Data Testing Pada Model Dengan Seleksi Fitur 89
Listing Code 4. 71 Mengimport Modul Evaluasi Performa Model Klasifikasi 90
Listing Code 4. 72 Import Library Seaborn 91
Listing Code 4. 73 Confusion Matrix Naive Bayes Classifier 91
Listing Code 4. 74 Proses Classification Report Model Naïve Bayes 93
Listing Code 4. 75 Perhitungan Nilai Akurasi Model NBC 94
Listing Code 4. 76 Confusion Matrix NBC Dengan PSO 97
Listing Code 4. 77 Proses Classification Report NBC Dengan PSO 99
Listing Code 4. 78 Perhitungan Nilai Akurasi NBC Dengan PSO 100
Listing Code 4. 79 Proses Mengimpor ROC Curve dan ROC AUC Score 100
Listing Code 4. 80 Prediksi Probabilitas Model NBC 101
Listing Code 4. 81 Konversi Biner Pada Label Data Testing Model NBC 101
Listing Code 4. 82 Menghitung Nilai AUC Setiap Kelas Pada Model NBC 101

Listing Code 4. 83 Pembuatan Kurva ROC Pada Model Naïve Bayes 102
Listing Code 4. 84 Menghitung Nilai AUC Keseluruhan Model NBC 103
Listing Code 4. 85 Prediksi Probabilitas NBC dan PSO 103
Listing Code 4. 86 Konversi Biner Pada Label Data Testing NBC Dan PSO 104
Listing Code 4. 87 Menghitung Nilai AUC Setiap Kelas Pada Model NBC dan PSO 104
Listing Code 4. 88 Pembuatan Kurva ROC Pada Model NBC Dan PSO 104
Listing Code 4. 89 Menghitung Nilai AUC Keseluruhan Model NBC Dan PSO 105

DAFTAR LAMPIRAN

Halaman

Lampiran 1. Daftar Stopwords 116
Lampiran 2. Daftar Normalisasi Slang Words 116

BAB I
PENDAHULUAN

1.1 Latar Belakang

Affiliate marketing semakin banyak diterapkan pada platform digital seperti *e-commerce*, dengan memanfaatkan sistem pemasaran melalui media sosial (*social media marketing*) (Husna, 2023). Jangkauan produk ke pasar menjadi lebih luas dengan sistem ini melalui rekomendasi dan ulasan produk yang dibuat oleh para afiliator. Selain itu dengan memanfaatkan *content creator/influencer* juga menjadi salah satu cara promosi terbaik karena tingginya kepercayaan publik dan biaya yang dikeluarkan relatif rendah dibandingkan dengan cara promosi yang lain (Bognar dkk., 2019). Salah satu perusahaan marketplace di Indonesia yang menggunakan strategi afiliasi ini adalah Shopee. Afiliator yang bergabung pada Program Shopee Affiliate akan mempromosikan produk Shopee pada media sosial mereka dengan membuat konten dan mendapatkan komisi/penghasilan apabila audiens melakukan pembelian produk melalui *custom link* yang dibagikan.

Menurut survey yang dilakukan oleh Snapcart, Shopee Affiliate Program menempati posisi pertama sebagai program afiliasi dengan jumlah transaksi tertinggi (*share of orders*) sebesar 59% pada tahun 2023, dimana Shopee memimpin pasar sebagai program afiliasi yang paling diingat (69%), yang paling sering digunakan (70%), dan komisi yang paling menguntungkan (68%) dibandingkan dengan para pesaingnya. Meskipun telah menjadi program afiliasi terbaik di Indonesia dan menawarkan keuntungan finansial yang tinggi bagi afiliator, terdapat pro dan kontra yang dapat mempengaruhi persepsi publik. Shopee beresiko kehilangan kepercayaan penggunanya, terutama pada media sosial Twitter (X) di mana keluhan terkait penyebaran *link* produk Shopee yang dianggap mengganggu dan tidak relavan dengan konteks sering muncul. Dikutip dari artikel Databoks berdasarkan laporan We are Social dan Hootsuite, Indonesia menempati posisi ke-5 sebagai negara dengan pengguna Twitter (X) terbanyak di dunia pada Januari 2023, yaitu sebanyak 24 juta pengguna. Jumlah yang besar ini menjadikan Twitter (X) media yang berpengaruh dalam membentuk opini publik. Pihak Shopee

dapat menjadikan opini-opini terkait Shopee Affiliate sebagai bahan pertimbangan untuk evaluasi dan perbaikan program melalui analisis sentimen.

Analisis sentimen dapat membantu Shopee dalam memberikan pandangan yang menyeluruh mengenai Program Shopee Affiliate baik secara positif, negatif, maupun netral. Melalui sentimen positif, Shopee dapat memperkuat keunggulan kompetitif dan strategi pemasaran dari program afiliasinya, sementara sentimen negatif dapat mengidentifikasi permasalahan atau kelemahan yang menimbulkan keluhan, dan sentimen netral dapat dimanfaatkan untuk peluang menarik minat audiens menjadi pelanggan potensial. Mengetahui umpan balik dari pelanggan memungkinkan perusahaan untuk mengambil langkah dan melakukan riset pasar terhadap produk dan sifat konsumen untuk pengembangan merk (Sumantiawan dkk., 2023). Data dari analisis sentimen ini dapat dijadikan bahan untuk menjaga reputasi dan meningkatkan penerimaan publik, sehingga Shopee dapat mempertahankan posisinya sebagai pemimpin pasar program afiliasi di Indonesia. Pada penelitian ini metode yang akan digunakan adalah *Naïve Bayes Classifier*. Metode ini dipilih karena klasifikasi data dapat dimodifikasi agar sesuai dengan kebutuhan masing-masing (Gunawan dkk., 2018)

Di sisi lain, Naïve Bayes memiliki sensitivitas yang tinggi terhadap pemilihan fitur dan hasil estimasi probabilitas yang tidak selalu optimal yang berakibat pada rendahnya hasil akurasi menjadi kelemahan bagi metode ini (Hayatin dkk., 2020). Untuk mengatasi masalah tersebut, diperlukan seleksi fitur. Seleksi fitur dapat membantu memilih fitur yang penting, menghilangkan fitur yang tidak diperlukan, memperkecil dimensi fitur, dan mempercepat komputasi (Prastyo dkk., 2022). Salah satu metode seleksi fitur yaitu *Particle Swarm Optimization* (PSO). Metode pencarian berbasis populasi yang ditawarkan oleh PSO sebagai alat pengoptimalan melibatkan individu atau yang disebut partikel. Dengan memperbarui posisi dan kecepatan setiap partikel, PSO akan mencari solusi terbaik. Karena kesederhanaan dan mudah diimplementasikan, serta pendekatan optimasi berkelanjutan, PSO dapat digunakan sebagai metode pemilihan fitur (Dulhare, 2018). Penggunaan PSO untuk mengoptimalkan Naïve Bayes diharapkan dapat meningkatkan akurasi untuk mengatasi kelemahan Naïve Bayes pada klasifikasi sentimen.

Berdasarkan uraian di atas, penelitian ini berupa analisis sentimen terhadap Program Shopee Affiliate pada data *tweet* berbahasa Indonesia yang diambil dari media sosial Twitter (X), yang akan diklasifikasikan menjadi tiga kelas sentimen, yaitu positif, netral, dan negatif. Metode yang digunakan adalah Naïve Bayes Classifier dan dilakukan seleksi fitur dengan Particle Swarm Optimization untuk meningkatkan hasil akurasi model klasifikasi.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka dirumuskan masalah penelitian ini yaitu:

1. Bagaimana cara untuk mengetahui kecenderungan opini dari media sosial Twitter (X) terkait Program Shopee Affiliate?
2. Bagaimana kinerja dari algoritma Naïve Bayes Classifier dengan menggunakan seleksi fitur Particle Swarm Optimization untuk klasifikasi pada data opini Program Shopee Affiliate?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut.

1. Melakukan analisis sentimen pada data opini yang berasal dari Twitter (X) mengenai Program Shopee Affiliate sehingga dapat menentukan respon pengguna Twitter (X) cenderung positif, negatif, atau netral.
2. Melakukan klasifikasi dengan mengimplementasikan algoritma Naïve Bayes Classifier dan seleksi fitur Particle Swarm Optimization pada data opini Twitter (X) terkait Program Shopee Affiliate.

1.4 Batasan Masalah

Adapun batasan masalah yang terdapat pada penelitian ini adalah sebagai berikut.

1. Penelitian dilakukan pada *tweets* yang menggunakan bahasa Indonesia.
2. Klasifikasi menggunakan 3 kategori kelas, yaitu positif, negatif, dan netral.
3. Tweet yang dianalisis menggunakan kata kunci yang berhubungan dengan Shopee Affiliate.
4. *Crawling* data dilakukan dengan menggunakan Tweet Harvest.
5. Data diambil dengan rentang waktu bulan Juli 2023 – Agustus 2023

6. Algoritma klasifikasi yang digunakan adalah Naïve Bayes Classifier dan seleksi fitur menggunakan Particle Swarm Optimization.
7. Jumlah populasi yang digunakan untuk eksperimen mengubah parameter Population Size pada seleksi fitur Particle Swarm Optimization maksimal sebanyak 50 karena keterbatasan sumber daya komputasi.

1.5 Manfaat Penelitian

Adapun manfaat yang diharapkan dari penelitian tugas akhir ini adalah sebagai berikut.

1. Manfaat Teoritis

   Penelitian ini diharapkan dapat dijadikan sebagai bahan referensi dan pengetahuan bagi pembaca maupun penelitian-penelitian selanjutnya mengenai sentimen analisis dan seleksi fitur.
2. Manfaat Praktis
   a. Bagi Penulis

      Hasil penelitian ini diharapkan dapat menambah wawasan dan pengetahuan bagi penulis sendiri mengenai klasifikasi sentimen dengan memanfaatkan algoritma Naïve Bayes dan *feature selection* Particle Swarm Optimization.
   b. Bagi Masyarakat

      Hasil penelitian ini diharapkan dapat meningkatkan kepercayaan masyarakat terhadap Program Shopee Affiliate.
   c. Bagi Perusahaan

      Penelitian ini diharapkan dapat menjadi bahan dan informasi yang berguna dalam pengembangan Program Shopee Affilate sesuai dengan kebutuhan masyarakat dan sejalan dengan perkembangan zaman serta masalah-masalah yang mungkin akan timbul di masa depan.

1.6 Sistematika Penulisan

Untuk memberikan gambaran mengenai pembahasan pada penelitian ini, berikut ini adalah penjelasan singkat dari setiap bab yang terbagi menjadi lima bab.

BAB I PENDAHULUAN

Bab ini berisi latar belakang penelitian, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini berisikan landasan teori yang relavan dengan topik penelitian seperti mengenai analisis sentimen, metode-metode dan *tools* yang digunakan, serta penelitian terdahulu yang dijadikan referensi pada penelitian ini.

BAB III METODOLOGI PENELITIAN

Bab ini menguraikan kerangka kerja penelitian menggunakan *IS Research Framework*, teknik pengumpulan data, dan alur serta langkah-langkah yang digunakan pada penelitian ini.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini berisi penjelasan tentang proses penelitian, seperti pengambilan data, *preprocessing data*, seleksi fitur menggunakan Particle Swarm Optimization, klasifikasi dengan Naïve Bayes Classifier, evaluasi model klasifikasi, visualisasi hasil, dan analisis hasil evaluasi algoritma.

BAB V PENUTUP

Bab Penutup berisi kesimpulan dari hasil penelitian yang telah dilakukan dan saran untuk penelitian selanjutnya.