



ANALISIS SENTIMEN PRODUK *FASHION* LOKAL PADA *MARKETPLACE* DENGAN KLASIFIKASI *SUPPORT VECTOR MACHINE (SVM)*

TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat
untuk Memperoleh Gelar Sarjana Komputer pada
Program Studi Sistem Informasi



Oleh:

NADILA ANDRIKA

12050322165



**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU**

2024

- Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSETUJUAN

**ANALISIS SENTIMEN PRODUK *FASHION* LOKAL PADA
MARKETPLACE DENGAN KLASIFIKASI *SUPPORT VECTOR*
MACHINE (SVM)**

TUGAS AKHIR

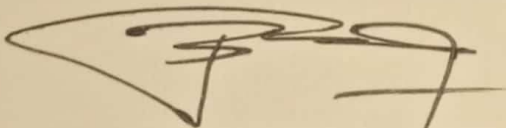
Oleh:

NADILA ANDRIKA

12050322165

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 08 Juli 2024

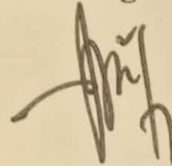
Ketua Program Studi



Eki Saputra, S.Kom., M.Kom.

NIP. 198307162011011008

Pembimbing



Fitriani Muttakin, S.Kom., M.Cs.

NIP. 198606122020122014

LEMBAR PENGESAHAN

ANALISIS SENTIMEN PRODUK *FASHION* LOKAL PADA MARKETPLACE DENGAN KLASIFIKASI *SUPPORT VECTOR* *MACHINE* (SVM)

TUGAS AKHIR

Oleh:

NADILA ANDRIKA

12050322165

Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau
di Pekanbaru, pada tanggal 14 Juni 2024

Pekanbaru, 14 Juni 2024

Mengesahkan,

Ketua Program Studi



Eki Saputra, S.Kom., M.Kom
NIP. 198307162011011008

DEWAN PENGUJI:

Ketua : Syaifullah, SE., M.Sc.

Sekretaris : Fitriani Muttakin, S.Kom., M.Cs.

Anggota 1 : Siti Monalisa, ST., M.Kom.

Anggota 2 : Saide, S.Kom., M.Kom., M.I.M., Ph.D.

Lampiran Surat :

Nomor : Nomor 25/2021

Tanggal : 10 September 2021

LEMBAR PERNYATAAN

Saya yang bertandatangan di bawah ini:

Nama : NADILA ANDRIKA

NIM : 12050322165

Tempat/Tgl. Lahir : MUARA LABUH, 7 NOVEMBER 2000

Fakultas/Pascasarjana : SAINS DAN TEKNOLOGI

Prodi : SISTEM INFORMASI

Judul Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya*:

Analisis Sentimen Produk Fashion Lokal Pada Marketplace dengan
Klasifikasi Support Vector Machine (SVM)

Menyatakan dengan sebenar-benarnya bahwa:

1. Penulisan ~~Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya*~~ dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu ~~Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya*~~ saya ini, saya nyatakan bebas dari plagiat.
4. Apa bila dikemudian hari terbukti terdapat plagiat dalam penulisan ~~Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya*~~ saya tersebut, maka saya bersedia menerima sanksi peraturan perundang-undangan.

Demikianlah Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 12 Juli 2024

Yang membuat pernyataan


NADILA ANDRIKA

NIM: 12050322165

*Pilih salah satu sesuai jenis karya tulis



LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum, dengan ketentuan bahwa hak cipta ada pada peneliti. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan atas izin peneliti dan harus dilakukan mengikuti kaedah dan kebiasaan ilmiah serta menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin tertulis dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan dapat meminjamkan Tugas Akhir ini untuk anggotanya dengan mengisi nama, tanda peminjaman, dan tanggal pinjam pada *form* peminjaman.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah,
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan di dalam daftar pustaka.

Pekanbaru, 08 Juli 2024

Yang membuat pernyataan,

NADILA ANDRIKA

NIM. 12050322165

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah,
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



LEMBAR PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Dengan menyebut nama Allah yang maha pengasih lagi maha penyayang

Assalamu'alaikum Warahmatullahi Wabarakatuh.

Dengan penuh rasa syukur dan bangga, Laporan Tugas Akhir ini kupersembahkan kepada Allah SWT, sumber segala kekuatan dan inspirasi, atas segala rahmat dan hidayah-Nya yang tiada henti, serta karunia kesehatan dan kesabaran yang memungkinkan penyelesaian Tugas Akhir ini. Kepada kedua orang tua tercinta, Ayah dan Ibu, yang telah memberikan cinta, do'a, dukungan tanpa batas, serta pengorbanan tanpa henti. Tanpa kalian, perjalanan ini tidak akan mungkin terwujud. Terima kasih juga kepada keluarga besar yang selalu memberikan semangat, kebahagiaan, dan dukungan moral yang luar biasa. Terima kasih atas kebersamaan dan perhatian kalian.

Saya juga menyampaikan rasa terima kasih yang mendalam kepada dosen pembimbing, yang telah memberikan bimbingan, ilmu, serta motivasi yang sangat berarti dalam proses penyelesaian Tugas Akhir ini. Terima kasih atas kesabaran dan arahan yang diberikan. Ucapan terima kasih juga saya sampaikan kepada teman-teman dan sahabat-sahabat tercinta, terima kasih atas kebersamaan, dukungan moral, serta bantuan yang kalian berikan selama masa perkuliahan hingga penyelesaian Tugas Akhir ini. Dengan segala kerendahan hati, saya mengharapkan saran dan kritik yang membangun untuk penyempurnaan laporan ini di masa yang akan datang.

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



KATA PENGANTAR

Alhamdulillah Rabbil 'Alamin, bersyukur kehadiran Allah *Subhanahu Wa Ta'ala* atas segala rahmat dan karunia-Nya sehingga peneliti dapat menyelesaikan Tugas Akhir ini. *Shalawat* serta salam kita ucapkan kepada Nabi Muhammad *Shalallahu 'Alaihi Wa Sallam* dengan mengucapkan *Allahumma Sholli'Ala Sayyidina Muhammad Wa'Ala Ali Sayyidina Muhammad*. Tugas Akhir ini dibuat sebagai salah satu syarat untuk mendapatkan gelar Sarjana Komputer di Program Studi Sistem Informasi Universitas Islam Negeri Sultan Syarif Kasim Riau.

Pada penulisan Tugas Akhir ini, terdapat beberapa pihak yang sudah berkontribusi dan mendukung peneliti baik berupa materi, moril, dan motivasi. Peneliti ingin mengucapkan banyak terima kasih kepada:

1. Bapak Prof. Dr. Hairunas, M.Ag sebagai Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Hartono, M.Pd sebagai Dekan Fakultas Sains dan Teknologi.
3. Bapak Eki Saputra, S.Kom., M.Kom sebagai Ketua Program Studi Sistem Informasi.
4. Ibu Siti Monalisa, ST., M.Kom sebagai Sekretaris Program Studi Sistem Informasi sekaligus Dosen Penguji I Tugas Akhir yang telah memberikan arahan, kritik, dan saran dalam penyelesaian Tugas Akhir ini.
5. Bapak Tengku Khairil Ahsyar, S.Kom., M.Kom sebagai Ketua Laboratorium Program Studi Sistem Informasi yang telah memberikan masukan dan saran selama perkuliahan.
6. Bapak Syaifullah, SE., M.Sc sebagai Dosen Pembimbing Akademik yang telah banyak memberikan arahan, masukan, nasihat, motivasi, serta menjadi pendengar yang baik selama perkuliahan dan proses penyelesaian perkuliahan ini.
7. Ibu Fitriani Muttakin, S.Kom., M.Cs sebagai Dosen Pembimbing Tugas Akhir yang telah memberikan bimbingan, motivasi, dan semangat dalam menyelesaikan Tugas Akhir ini.
8. Bapak Saide, S.Kom., M.Kom., M.I.M., Ph.D sebagai Dosen Penguji II Tugas Akhir yang telah memberikan arahan, kritik, dan saran dalam penyelesaian Tugas Akhir ini.
9. Bapak Ibu Dosen Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau yang telah memberikan ilmu yang bermanfaat serta memberikan semangat dan motivasi kepada peneliti.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

10. Terima kasih kepada kedua orang tua peneliti yaitu Ayahanda Hendri Musfa dan Ibunda Marniati tercinta yang telah memberikan dukungan, do'a, perhatian, kasih sayang, dan semangat yang tiada hentinya kepada peneliti dalam menyelesaikan Tugas Akhir ini.

11. Terima kasih kepada Himpunan Mahasiswa Program Studi Sistem Informasi yang telah memberikan kesempatan untuk peneliti bergabung dan berkontribusi pada kegiatan-kegiatan mahasiswa.

12. Terima kasih kepada sahabat-sahabat tercinta peneliti yang telah mendukung, mendampingi, dan berjuang bersama menuntut ilmu dalam mengejar mimpi.

13. Semua pihak yang tidak dapat peneliti sebutkan satu persatu pada kesempatan ini yang telah banyak membantu dalam penyelesaian Tugas Akhir ini.

Pengerjaan laporan ini terdapat banyak kesalahan dan kekurangan. Oleh karenanya, kritik dan saran yang membangun sangat diharapkan untuk kesempurnaan Laporan Tugas Akhir ini. Untuk itu, dapat menghubungi peneliti melalui *email* di 12050322165@students.uin-suska.ac.id atau nadiilaandrika@gmail.com. Semoga laporan ini dapat memberikan sesuatu yang bermanfaat bagi siapa saja yang membacanya. *Aamiin*.

Pekanbaru, 08 Juli 2024

Peneliti,

NADILA ANDRIKA

NIM. 12050322165



ANALISIS SENTIMEN PRODUK *FASHION* LOKAL PADA MARKETPLACE DENGAN KLASIFIKASI *SUPPORT VECTOR MACHINE* (SVM)

NADILA ANDRIKA

NIM: 12050322165

Tanggal Sidang: 14 Juni 2024

Periode Wisuda:

Program Studi Sistem Informasi

Fakultas Sains dan Teknologi

Universitas Islam Negeri Sultan Syarif Kasim Riau

Jl. Soebrantas, No. 155, Pekanbaru

ABSTRAK

Perkembangan bisnis digital saat ini merambat hingga *marketplace* seperti Shopee telah menjadi platform populer untuk pembelian produk *fashion* lokal dengan memberikan fitur *review* pelanggan. Jumlah data *review* pengunjung Shopee yang terus bertambah, menyebabkan *seller* akan membutuhkan waktu yang lama untuk membaca setiap *review* secara menyeluruh pada halaman toko yang berada di Shopee. Selain itu, *rating* yang diberikan belum bisa menggambarkan opini dari pelanggan terhadap sepatu lokal tersebut. Penelitian ini bertujuan untuk mengetahui opini pelanggan terhadap produk *fashion* lokal di Shopee dengan menggunakan analisis sentimen dan Algoritma *Support Vector Machine* (SVM). Data *review* pelanggan dari toko *brand* sepatu lokal di Shopee dikumpulkan dan dianalisis menggunakan Algoritma SVM dengan pelabelan menggunakan Metode *Inset Lexicon Based*. Hasil analisis menunjukkan bahwa *review* pelanggan terhadap produk *fashion* lokal di Shopee mempunyai nilai positif dengan tingkat akurasi SVM sebesar 91,58%. Penelitian ini memberikan kontribusi dalam memahami persepsi pelanggan terhadap produk *fashion* lokal di Shopee. Hasil penelitian ini dapat membantu *seller* dalam mengelola ulasan pelanggan dengan efektif, meningkatkan interaksi dengan pelanggan, dan menarik pelanggan baru. Penelitian ini menunjukkan bahwa analisis sentimen dengan Algoritma SVM dapat menjadi alat yang efektif untuk memahami opini pelanggan terhadap produk *fashion* lokal di Shopee. Hasil penelitian ini dapat memberikan manfaat bagi *seller* dan konsumen dalam meningkatkan kualitas produk dan layanan.

Kata Kunci: Analisis Sentimen, *Customer Review*, Shopee, *Support Vector Machine* (SVM)

- Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



SENTIMENT ANALYSIS OF LOCAL FASHION PRODUCTS ON MARKETPLACE WITH SUPPORT VECTOR MACHINE (SVM) CLASSIFICATION

NADILA ANDRIKA
NIM: 12050322165

Date of Final Exam: June 14th 2024
Graduation Period:

Department of Information System
Faculty of Science and Technology
State Islamic University of Sultan Syarif Kasim Riau
Soebrantas Street, No. 155, Pekanbaru

ABSTRACT

The development of digital business has now spread to marketplaces such as Shopee has become a popular platform for purchasing local fashion products by providing customer review features. The number of Shopee visitor review data continues to grow, causing sellers to take a long time to read each review thoroughly on the shop page on Shopee. In addition, the rating given cannot describe the customer's opinion of the local shoes. This research aims to find out customer opinions on local fashion products on Shopee using sentiment analysis and the Support Vector Machine (SVM) algorithm. Customer review data from local shoe brand stores on Shopee was collected and analyzed using the SVM algorithm with labeling using the Inset Lexicon Based method. The analysis results show that customer reviews of local fashion products on Shopee have a positive value with an SVM accuracy rate of 91,58%. This research contributes to understanding customer perceptions of local fashion products at Shopee. The results of this study can help sellers manage customer reviews effectively, increasing interaction with customers, and attracting new customers. This research shows that sentiment analysis with the SVM algorithm can be an effective tool for understanding customer opinions of local fashion products on Shopee. The results of this study can provide benefits for sellers and consumers in improving product and service quality.

Keywords: *Customer Reviews, Sentiment Analysis, Shopee, Support Vector Machine (SVM)*

- Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



DAFTAR ISI

LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL	iv
LEMBAR PERNYATAAN	v
LEMBAR PERSEMBAHAN	vi
KATA PENGANTAR	vii
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI	xi
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xiv
DAFTAR SINGKATAN	xv
1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Tujuan	3
1.5 Manfaat	3
1.6 Sistematika Penulisan	3
2 LANDASAN TEORI	5
2.1 Analisis Sentimen	5
2.2 <i>Marketplace</i>	6
2.3 Shopee	6
2.4 <i>Customer Review</i>	7
2.5 <i>Text Mining</i>	7
2.6 <i>Lexicon Based</i>	8

- Hak Cipta Dilindungi Undang-Undang
- Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.7	Pembobotan TF-IDF	8
2.8	<i>Support Vector Machine (SVM)</i>	9
2.9	<i>Confusion Matrix</i>	9
2.10	Penelitian Terdahulu	11
3	METODOLOGI PENELITIAN	14
3.1	Tahap Perencanaann	14
3.2	Pengumpulan Data	15
3.3	<i>Preprocessing</i>	15
3.4	Validasi	17
3.5	Hasil	17
3.6	Tahap Dokumentasi	17
4	HASIL DAN ANALISA	18
4.1	Identifikasi Masalah	18
4.2	Pengumpulan Data	18
4.3	<i>Cleaning Data</i>	19
4.4	<i>Tokenize Data</i>	20
4.5	Normalisasi Data	21
4.6	<i>Stopwords Removal</i>	22
4.7	<i>Stemming Data</i>	23
4.8	<i>Labeling Data</i>	24
4.9	<i>Data Clean</i>	25
4.10	Perhitungan Nilai Positif dan Negatif	26
4.11	Perhitungan TF-IDF	27
4.12	<i>Wordcloud Data Pada Data Training</i>	27
4.13	<i>Confusion Matrix Data</i>	29
4.14	Klasifikasi Data Menggunakan SVM	29
4.15	Hasil Analisa Sentimen	30
5	PENUTUP	33
5.1	Kesimpulan	33
5.2	Saran	33

DAFTAR PUSTAKA

LAMPIRAN A DATA PENELITIAN

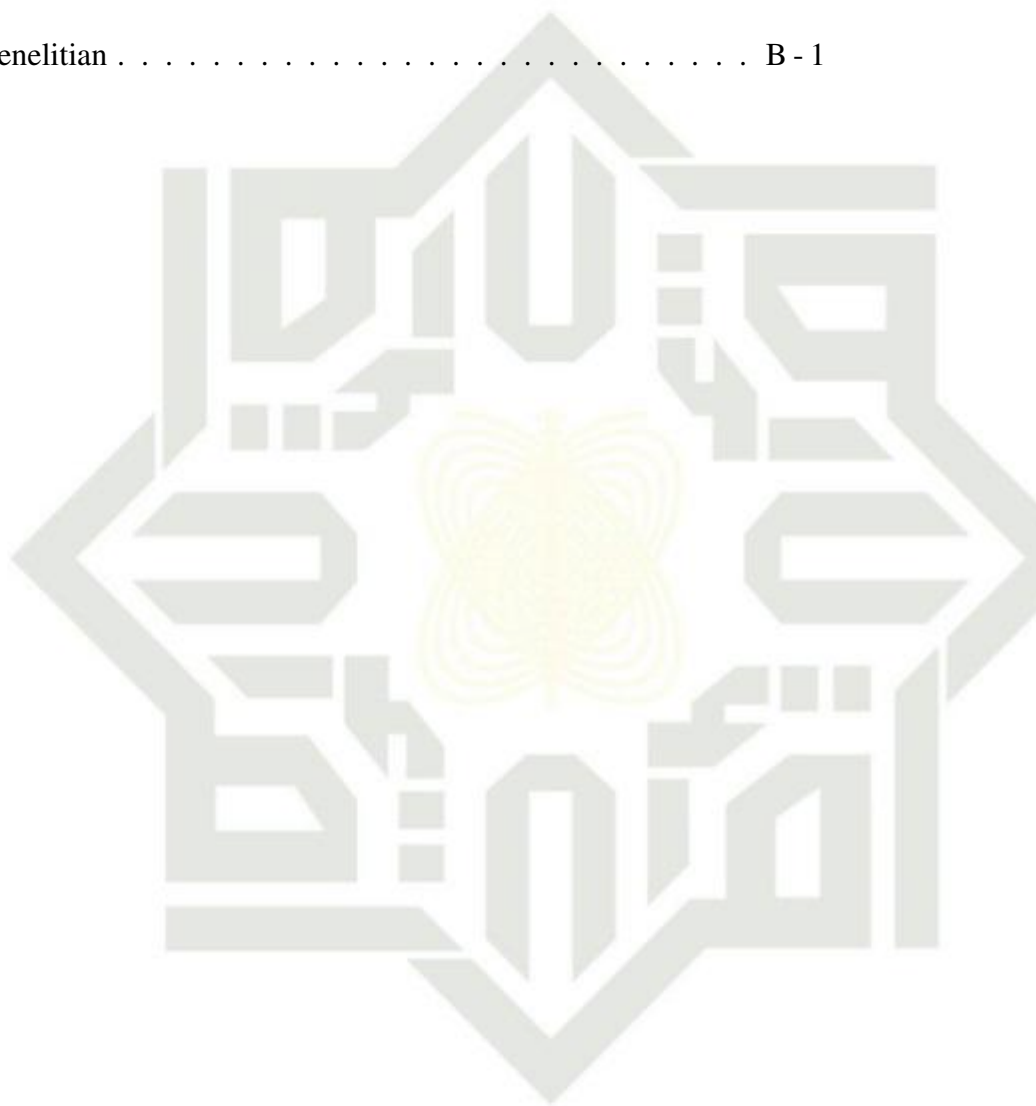
A - 1

LAMPIRAN B PRODUK PENELITIAN

B - 1

DAFTAR GAMBAR

© Hak Cipta Taamlik UIN Suska Riau		
2	<i>Ilustrasi Predictive Class</i>	10
3.1	Metodologi Penelitian	14
4.1	Jumlah Persentase <i>Comment</i> Positif dan Negatif	26
4.2	<i>Wordcloud</i> Sentimen Bernilai Positif	28
4.3	<i>Wordcloud</i> Sentimen Bernilai Negatif	28
4.4	<i>Confusion Matrix</i> SVM Pada Data <i>Training</i>	29
B.1	Produk Penelitian	B - 1



UIN SUSKA RIAU

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR TABEL

21	Penelitian Terdahulu	11
4.1	Tampilan Data Awal	18
4.2	Data <i>Cleaning</i>	20
4.3	Data <i>Tokenize</i>	21
4.4	Data <i>Normalisasi</i>	22
4.5	Data <i>Stopwords Removal</i>	23
4.6	Data <i>Stemming</i>	23
4.7	Data <i>Labeling</i>	24
4.8	Data <i>Clean</i>	25
4.9	Hasil Perhitungan TF-IDF	27
4.10	Hasil Prediksi Data Menggunakan Algoritma SVM	30
A.1	Data Penelitian	A - 1

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



DAFTAR SINGKATAN

DF	: <i>Document Frequency</i>
e-WOM	: <i>Electronic Word of Mouth</i>
FN	: <i>False Negative</i>
FP	: <i>False Positive</i>
IDF	: <i>Inverse Document Frequency</i>
ISNC	: <i>Information System Networking Club</i>
ISOC	: <i>Information System Organization Culture</i>
KNN	: <i>K-Nearest Neighbors</i>
NLP	: <i>Natural Language Processing</i>
TD	: <i>Term Document</i>
TF	: <i>Term Frequency</i>
TF-IDF	: <i>Term Frequency-Inverse Document Frequency</i>
TI	: <i>Teknologi Informasi</i>
TN	: <i>True Negative</i>
TP	: <i>True Positive</i>
TPR	: <i>True Positive Rate</i>
SVM	: <i>Support Vector Machine</i>
UMKM	: <i>Usaha Mikro Kecil dan Menengah</i>

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan bisnis saat ini berlangsung dengan cepat berkat adopsi teknologi yang semakin meningkat, terdapat hubungan yang semakin erat antara kemampuan perusahaan dalam memanfaatkan teknologi dan kemampuannya dalam menerapkan strategi bersaing untuk mencapai tujuan bisnis (Sakti, Sukaris, dan Saepuloh, 2020). Kemajuan yang cepat dalam penggunaan internet telah menjadikan Indonesia sebagai pasar yang menjanjikan dalam penerapan bisnis digital. *Marketplace* adalah platform daring yang mempermudah transaksi jual beli dari beragam toko (Rohman, Musyarofah, Utami, dan Raharjo, 2020). Sebagian besar *marketplace* menyediakan fitur *review* produk dimana pembeli yang sudah melakukan pembelian pada *marketplace* diminta untuk memberikan penilaian mereka (Bangsa, Priyanta, dan Suyanto, 2020).

Salah satu *marketplace* yang banyak dikunjungi yaitu Shopee. Berdasarkan informasi dari Sea Group pada tahun 2022 terdapat 27.000 pelaku Usaha Mikro Kecil dan Menengah (UMKM) yang berasal dari Indonesia, Malaysia, dan beberapa negara lainnya (<https://www.sea.com/sustainability>). Dengan begitu, tentunya terdapat banyak toko yang menjual *fashion* yang berada di Shopee baik lokal maupun impor. Produk *fashion* lokal juga sangat menarik perhatian masyarakat Indonesia dalam berbelanja di Shopee (Gunawan dkk., 2023). Produk *fashion* lokal yang ditawarkan kepada masyarakat dengan harga yang murah meskipun dengan kualitas barang dibawah produk original sangat menarik animo masyarakat dalam menggunakan produk tersebut. Oleh karena itu, *review* dari *customer* diperlukan untuk mengetahui respon *customer* terhadap produk tersebut dalam meningkatkan kepercayaan *customer*. Analisis sentimen dapat membantu dalam memahami kelebihan dan kelemahan dari produk sepatu lokal serta meningkatkan klasifikasi *review* dari produk tersebut (Agustiono dan Darmawan, 2023).

Analisis sentimen merupakan salah satu area penelitian dalam pengolahan bahasa alami, komputasi linguistik, dan *text mining* yang bertujuan secara komputasional untuk menganalisis pendapat, penilaian, dan emosi yang terkandung dalam teks terkait entitas, peristiwa, dan atribut. Dalam analisis sentimen, tujuannya adalah mengelompokkan polaritas yang terdapat dalam teks, termasuk dalam dokumen, kalimat, atau bahkan dalam aspek-aspek tertentu, untuk menentukan apakah pendapat yang terkandung dalam teks tersebut bersifat positif maupun negatif (Nugraha, Harani, dan Habibi, 2020). Analisis sentimen dapat juga diolah menggu-

Hak Cipta Dilindungi Undang-Undang

© Hak Cipta Dilindungi Undang-Undang

Sejarah Islam di Riau Sultan Syarif Kasim II

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



nakan data yang didapatkan pada *review* yang diberikan oleh *customer* Shopee.

Pengunjung pada toko *brand* sepatu lokal yang berada di Shopee juga mengalami perubahan posisi pada Shopee dari tahun ke tahun dengan berbagai jenis *review*. Jumlah data *review* pengunjung Shopee yang terus bertambah, menyebabkan *seller* akan membutuhkan waktu yang lama untuk membaca setiap *review* secara menyeluruh pada halaman toko yang berada di Shopee. Selain itu, *rating* yang diberikan belum bisa menggambarkan opini dari *customer* terhadap sepatu lokal tersebut. *Brand image* yang baik akan membentuk opini yang baik dari *customer* tentang barang dan jasa, yang diharapkan akan mendorong mereka untuk melakukan pembelian, dan sebaliknya.

Hal ini mendorong peneliti untuk melakukan analisis sentimen pada *review* pengguna. Penelitian ini meneliti tanggapan konsumen untuk produk *fashion* lokal yang tersedia di Shopee. Tujuan penelitian ini untuk mengetahui bagaimana opini *customer* terhadap produk *fashion* lokal yang berada pada toko-toko di Shopee (Hidayah, Permanasari, dan Wijayanti, 2019). Penelitian ini menggunakan Algoritma Klasifikasi *Support Vector Machine* (SVM) dalam menentukan nilai akurasi. Proses klasifikasi menggunakan SVM dikarenakan algoritma ini memiliki tingkat akurasi lebih baik dibandingkan algoritma lainnya.

Menurut Idris, Mustofa, dan Salihi (2023) pada penelitiannya menunjukkan bahwa analisis sentimen dengan menggunakan Algoritma *Support Vector Machine* (SVM) terbukti dapat dikatakan sebagai algoritma klasifikasi yang baik dengan perolehan nilai akurasi sebesar 98% dan *F1-Score* 0,98 atau sebesar 98%. Selain itu, menurut Dey dkk. (2020) menganalisis bahwa sentimen konsumen terhadap produk Amazon menggunakan Algoritma *Naive Bayes* dan SVM. Keduanya menghasilkan akurasi yang lebih tinggi daripada teknik konvensional, dengan SVM mencapai 84% dan *Naive Bayes* mencapai 82,875%. Hasil percobaan tersebut menunjukkan bahwa SVM dapat mempolarisasi umpan balik produk Amazon dengan tingkat akurasi yang lebih tinggi.

Berdasarkan permasalahan di atas, penelitian ini dilakukan dengan menggunakan Algoritma Klasifikasi *Support Vector Machine* (SVM) yang dapat mengklasifikasikan *comment* pengguna pada aplikasi Shopee baik secara positif maupun negatif.

1.2 Perumusan Masalah

Rumusan masalah yang terdapat pada penelitian ini yaitu bagaimana opini *customer* terhadap produk *brand* sepatu lokal pada *marketplace*.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



1.3 Batasan Masalah

Batasan masalah pada penelitian ini yaitu:

1. Sumber data yang digunakan adalah data *review* atau *comment customer* terhadap produk sepatu lokal pada *marketplace* terkhusus Shopee dengan *brand* Geofmax, Compas, Brodo, Nah, dan Saint Barkley.
2. Data *review* atau *comment customer* yang digunakan 4 tahun terakhir dari tahun 2020-2023.
3. Penggunaan *comment* yang diambil adalah yang berbahasa Indonesia.
4. Metode yang digunakan adalah Algoritma *Support Vector Machine* (SVM) sebagai media klasifikasi.

1.4 Tujuan

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Untuk mengetahui opini yang diberikan oleh *customer review* terhadap produk *brand* sepatu lokal yang diteliti pada *marketplace* baik bernilai positif maupun negatif.
2. Untuk mengidentifikasi dan mengevaluasi efektivitas Algoritma SVM dalam menganalisis sentimen produk *brand* sepatu lokal yang diteliti pada *marketplace*.

1.5 Manfaat

Adapun manfaat dari penelitian ini yaitu:

1. Memberikan pengetahuan mengenai tanggapan dari *reviewer* lain terhadap *brand* sepatu lokal yang ada di Shopee.
2. Memberikan informasi mengenai nilai terhadap *brand* sepatu lokal yang ada di Shopee menggunakan Algoritma *Support Vector Machine* (SVM).

1.6 Sistematika Penulisan

Sistematika penulisan pada penelitian ini dibagi menjadi 5 (lima) bab. Berikut penjelasan tentang masing-masing bab:

BAB 1. PENDAHULUAN

Didalam bab ini peneliti menjelaskan deskripsi umum dari Tugas Akhir yaitu latar belakang, persamaan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan Laporan Tugas Akhir.

BAB 2. LANDASAN TEORI

Di bab ini membahas mengenai hal-hal dan teori yang menjadi landasan didalam melakukan penulisan Tugas Akhir.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu karya.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim II



BAB 3. METODOLOGI PENELITIAN

Di bab ini berisi tentang pelaksanaan penelitian Tugas Akhir yang memuat penjabaran meliputi perencanaan, pengumpulan data, tahapan penelitian, proses pengelolaan data, serta analisis data.

BAB 4. HASIL DAN ANALISA

Di bab ini berisi tentang penjelasan mengenai hasil proses pengumpulan data dan hasil analisis sentimen mengenai *customer review* terhadap suatu produk sepatu pada *marketplace*.

BAB 5. PENUTUP

Di bab ini berisi kesimpulan dan saran dari penelitian Tugas Akhir yang telah dilakukan mengenai analisis sentimen *customer review* terhadap suatu produk sepatu pada *marketplace*.

Hak Cipta Dilindungi Undang-Undang

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim II

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



BAB 2

LANDASAN TEORI

2.1 Analisis Sentimen

Analisis sentimen yang juga dikenal sebagai penambangan opini merupakan cabang khusus dalam NLP yang berperan penting dalam mengevaluasi pendapat individu terhadap beragam isu yang dibagikan di berbagai platform media sosial dan forum *online* (Mukherjee dkk., 2021). Analisis sentimen umumnya bisa dikerjakan dengan metode leksikon atau pendekatan pembelajaran mesin yang terpantau. Pada metode berbasis *Lexicon*, evaluasi sentimen dilakukan dengan menilai kata dan frasa dalam teks sesuai dengan makna semantiknya untuk menentukan orientasi dokumen teks (Suhartono dkk., 2023).

Jenis tingkatan klasifikasi analisis sentimen menurut Urva dkk (2023) terbagi 3 yaitu:

1. Tingkat Dokumen *Document Level*: Jenis analisis sentimen yang pertama disebut dengan *opinion mining*, di mana setiap dokumen diperiksa untuk menentukan polaritasnya (positif, negatif, atau netral) berdasarkan satu topik, seperti review tentang satu produk. Hal ini memiliki keuntungan yaitu dapat menentukan polaritas fitur tertentu. Kekurangannya yaitu tidak memungkinkan untuk menentukan preferensi orang.
2. Tingkat Kalimat *Sentence Level*: Analisis sentimen tingkat kalimat memerlukan evaluasi setiap kalimat untuk melihat apakah kalimat tersebut memiliki pandangan yang positif, negatif, atau netral. Pernyataan subjektif mewakili pendapat subjektif pengguna. Sedangkan kalimat netral menyampaikan informasi faktual tanpa opini. Analisis sentimen tingkat kalimat memiliki batasan tertentu, tetapi salah satu manfaatnya adalah kategorisasi subjektivitas dan objektivitas.
3. Tingkat Aspek/Ciri *Aspect/Feature Level*: Tingkat analisis sentimen yang terakhir dikenal sebagai tingkat aspek atau tingkat fitur, berkonsentrasi pada elemen atau fitur tertentu dalam teks atau frasa. Untuk membantu meningkatkan kualitas produk atau layanan, kita dapat mengetahui pendapat pengguna tentang setiap fitur atau entitas pada tingkat ini. Tingkat ini memiliki keuntungan karena dapat menentukan fungsi yang paling dihargai konsumen sekaligus meningkatkan layanan *customer*. Kekurangannya adalah tingkat analisis data ini membutuhkan lebih banyak waktu dan usaha.

Hak Cipta Dilindungi Undang-Undang

© Hak cipta dilindungi UIN Suska Riau

State Islamic University of Sultan Syarif K

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



2.2 Marketplace

Istilah "market" mengacu pada strategi untuk mempromosikan dan menjual barang dan jasa secara *online*. Meskipun demikian, *place* adalah tempat. Oleh karena itu, dapat disimpulkan bahwa *marketplace* adalah sebuah lokasi atau area yang digunakan untuk penjualan dan pemasaran barang dan jasa melalui internet. Atau dengan kata lain, *marketplace* adalah situs web atau aplikasi *online* yang memudahkan toko-toko yang berbeda untuk membeli dan menjual barang. Sebenarnya, ide di balik *marketplace* ini sangat mirip dengan pasar konvensional (Rahmayanti, 2023).

2.3 Shopee

Shopee adalah satu-satunya *marketplace* yang pada awalnya diluncurkan oleh Garena, yang namanya telah berubah menjadi SEA Group. Shopee merupakan *marketplace* yang digunakan oleh penjual *online* di Indonesia. Shopee pertama kali muncul di Indonesia pada tahun 2015 dan dengan cepat mendapatkan popularitas di kalangan masyarakat umum. Namun, aplikasi yang paling banyak diunduh dan populer pada tahun lalu adalah *marketplace* Shopee, yang mendapat peringkat sangat tinggi. Saat ini, Usaha Mikro Kecil dan Menengah (UMKM) menggunakan *marketplace* Shopee. Shopee mendapatkan keuntungan dari bisnisnya dengan mengoptimalkan segmen *mobile* karena hal itu dapat membuat transaksi jual beli menjadi lebih mudah (Sulistiyawati dan Widayani, 2020). Salah satu produk yang sering diminati dan dicari pada Shopee yaitu sepatu.

Sepatu merupakan salah satu produk yang paling diinginkan untuk dibeli oleh semua kalangan. Saat ini, setiap keluarga mungkin telah menggunakan sepatu untuk bekerja, sekolah, atau bahkan untuk kegiatan sehari-hari (Prayoga dan Handoko, 2023). Ada banyak produk lokal berkualitas tinggi yang tidak sebegitu baik produk dari luar yang menyebabkan lebih banyak konsumen lokal yang lebih memilih produk lokal daripada produk impor (Nasution, Sembiring, dan Elhanafi, 2023).

Beberapa produk lokal yang sering diminati masyarakat yaitu Compass yang didirikan pada tahun 1998 di kota Bandung, Indonesia. Bisnis ini dimulai oleh Bapak Kahar Gunawan, yang ingin membangun kembali industri sepatu berdasarkan visi yang ditinggalkan oleh almarhum ayahnya, Bapak Kahar Setiadi (Indrawan, Mulyadi, dan Furkan, 2022). Kedua, produk *brand* Saint Barkley yang didirikan pada tahun 2012 di Bandung, Indonesia, Saint Barkley adalah merek lokal yang telah mendapatkan popularitas di industri *fashion* (Arianto, Yogaswara, dan Sagita, 2019).

Produk *brand* lokal ketiga, Brodo yang didirikan pada tahun 2010 di Ban-

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



dung, Jawa Barat oleh CEO bernama Yukka. Perusahaan *brand* sepatu Brodo bertujuan membuat sepatu *stylish*, berkualitas tinggi, dengan harga terjangkau (Ramadhani dan Setiawati, 2022). Keempat, Geoff Max merupakan merek lokal yang banyak digunakan oleh para selebriti baik di dalam maupun di luar negeri. Geoffmax ini berasal dari daerah Bandung, Jawa Barat. Geoff Max pertama kali diperkenalkan oleh Yusuf Ramdhani dan Fauzan Efwanda pada tahun 2012 (Arif, 2022). Kelima, Nah Project adalah bisnis lokal asal Bandung yang memproduksi pakaian dan aksesoris dengan kualitas *Global Standard* dalam berbagai gaya dan model. Sejak didirikan pada tahun 2017, Nah Project telah berkembang menjadi pilihan terbaik untuk anak-anak dan orang dewasa (Azis dan Sofyan, 2021).

2.4 Customer Review

Online Customer Review adalah jenis *electronic Word of Mouth* (e-WOM) yang berfokus pada konten yang dibuat oleh pengguna yang diposting di situs web publik atau pribadi. *Online Customer Review* adalah berbagai macam *comment*-positif, negatif, dan umum-yang dibuat tentang produk, layanan, atau barang lain oleh konsumen dan dibagikan kepada konsumen lain dalam format yang terstruktur sebelum dipublikasikan di situs web independen *Customer Review* (Siti Nuraeni dan Irawati, 2021).

2.5 Text Mining

Text Mining merupakan metode untuk mengekstraksi informasi berharga dari teks yang sering kali melibatkan analisis linguistik dan statistik yang mendalam. *Text Mining* melibatkan berbagai langkah, seperti mengurai teks menggunakan penjabaran linguistik, menghilangkan elemen yang tidak relevan, dan menyimpan hasilnya dalam *database* (Dharmendra, Kusuma, Dewi, dkk., 2023). *Text Mining* adalah proses membangun pengetahuan terstruktur secara otomatis dari sejumlah besar teks yang tidak terstruktur. Teknologi ini dapat mengumpulkan banyak data besar yang kemudian diproses untuk memberikan penawaran harga dari kumpulan teks. Beberapa aplikasi *text mining* berfokus pada analisis sentimen media sosial, ekstraksi informasi untuk mengklasifikasikan teks, dan pemrosesan teks untuk memadatkan teks yang panjang menjadi satu atau dua paragraf. Perkembangan dalam penelitian dan pengembangan ekstraksi informasi otomatis dari dokumen adalah contoh penggunaan spesifik dari teknologi *text mining* (Hermawan, Jowensen, Junaedi, dkk., 2023).

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



2.6 *Lexicon Based*

Lexicon adalah metode yang digunakan dalam analisis sentimen untuk pelabelan. Metode ini menentukan polaritas positif atau negatif dengan mencocokkan setiap kata dalam dataset dengan kata-kata dalam leksikon. Dalam contoh ini, data *comment* yang sudah dipreprocessing akan dicocokkan dengan *Lexicon InSet*. Nilai polaritas total kemudian dihitung dengan menjumlahkan bobot dari kata-kata yang cocok dengan *Lexicon*, baik negatif maupun positif. Setelah itu, setiap *comment* dilabelkan menurut nilai polaritas totalnya: positif (polaritas lebih besar dari 0), netral (polaritas lebih rendah dari 0), atau negatif (polaritas lebih rendah dari 0). Dengan menggunakan diagram dan *wordcloud*, hasil pelabelan ini divisualisasikan pada tahap analisis data eksplorasi (Aisy dan Prasetyo, 2023).

Penelitian ini menggunakan *Lexicon InSet* karena telah terbukti berguna untuk menganalisis sentimen pada data berbahasa Indonesia. *InSet Lexicon* juga dikenal sebagai Indonesia Sentimen *Lexicon*, terdiri dari 3.609 kata positif dan 6.609 kata negatif dalam bahasa Indonesia, dan setiap kata memiliki bobot nilai atau skor polaritas dari -5 hingga +5. Tujuan *InSet Lexicon* adalah untuk menemukan dan mengkategorikan opini tertulis menjadi opini positif atau negatif, yang kemudian dapat digunakan untuk menilai persepsi publik tentang topik ataupun produk tertentu (Musfiroh, Khaira, Eko, Utomo, dan Suratno, 2021). Perhitungan yang digunakan dalam menentukan nilai bobot dapat dilihat pada Persamaan 2.1.

$$\text{NilaiSentimen} = \sum (\text{bobot kata} \times \text{frekuensi kata}) \quad (2.1)$$

Keterangan:

Bobot kata: bobot nilai sentimen kata dari *InSet Lexicon*.

Frekuensi kata: frekuensi kemunculan kata dalam teks.

2.7 Pembobotan TF-IDF

Term Frequency-Inverse Document Frequency (TF-IDF) digunakan untuk menentukan bobot setiap kata yang diambil. Dalam pencarian informasi, pendekatan ini biasanya digunakan untuk menghitung istilah yang sering muncul. Model *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF) digabungkan dalam model pembobotan TF-IDF. *Inverse Document Frequency* (IDF) digunakan untuk menghitung term yang muncul di beberapa dokumen (*comment*), yang dianggap paling sering muncul dan dianggap tidak penting. *Term Frequency* (TF) adalah prosedur untuk menentukan jumlah kemunculan *Term* dalam satu dokumen (Luqyana, Cholissodin, dan Perdana, 2018).

Metode TF-IDF menggabungkan dua konsep: seberapa sering sebuah kata



muncul dalam sebuah dokumen dan seberapa jarang kata tersebut muncul di seluruh koleksi dokumen. Pertama, melakukan perhitungan pada nilai TF untuk setiap kata dengan bobot awal 1, dan kemudian dengan menggunakan Persamaan 2.2.

$$IDF(Word) = \log \left(\frac{td}{df} \right) \quad (2.2)$$

Keterangan:

TD: *Term Document* (Jumlah Dokumen).

DF: *Document Frequency* (Jumlah Frekuensi Kata).

IDF(*Word*) adalah metrik IDF yang diterapkan pada setiap kata yang akan dicari, di mana *td* adalah jumlah total dokumen yang ada dan *df* adalah jumlah kata yang muncul di seluruh dokumen (Deolika, Kusriani, dan Luthfi, 2019).

2.8 Support Vector Machine (SVM)

Menurut Putri dkk (2022) *Support Vector Machine* adalah metode yang paling umum untuk meregresi atau mengklasifikasikan data berdasarkan data sebelumnya, dan pekerjaan pemodelan diawasi secara ketat. SVM adalah contoh pengklasifikasi non-probabilistik, linier, dan biner. Untuk membangun model linier atau hiper-bidang yang paling efektif untuk mengklasifikasikan data, SVM menggunakan batas keputusan (*Decision Boundaries*) yang akan menghasilkan klasifikasi dari perbandingan data-data. SVM mendeteksi *Hyperplane* dengan cara meningkatkan batas antar baris (Margin) agar memiliki kemampuan generalisasi yang kuat terhadap data yang masuk. *Hyperplane* mengubah data menjadi dua kelas yang identik dengan menggabungkannya sedemikian rupa sehingga jarak antara mereka dengan objek lain dengan kelas yang sama terluar (berdekatan) dan peringatan hitam atau putih. SVM mengacu pada data objek-objek yang sekarang tersedia yang paling dekat dengan *Hyperplane* (Agustina, Subanti, dan Zukhronah, 2021).

2.9 Confusion Matrix

Dalam *machine learning*, *Confusion Matrix* adalah alat untuk mengevaluasi kinerja model klasifikasi. Matriks ini menunjukkan jumlah *True Positives*, *True Negatives*, *False Positives* dan *False Negatives*. Dasar awal *Confusion Matrix* yang dilakukan dalam menghitung *matric* dapat dilihat seperti Gambar 2.1 (Nuraliza, Pratiwi, dan Hamami, 2022).

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

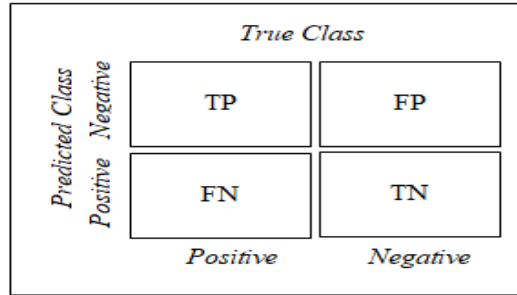
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak cipta dilindungi undang-undang UIN Suska Riau

State Islamic University of Sultan Syarif Kasim II



- Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar 2.1. Ilustrasi Predictive Class

Keterangan:

- TP: *True Positive* (Jumlah prediksi benar kelas positif).
- TN: *True Negative* (Jumlah prediksi benar kelas negatif).
- FN: *False Negative* (Kelas asli positif diprediksi negatif).
- FP: *False Positive* (Kelas asli negatif diprediksi positif).

- *Accuracy*

Indikator berapa banyak kelas yang diprediksi oleh model dengan benar. Ini adalah model yang paling sering digunakan, namun ketika berhadapan dengan data yang tidak seimbang, ini dapat menyebabkan interpretasi kinerja yang bias. Rumus menghitung akurasi dapat dilihat pada Persamaan 2.3.

$$Akurasi = \frac{TN + TP}{TN + FP + TP + FN} \quad (2.3)$$

- *Precision*

Ukuran seberapa akurat model mengidentifikasi kelas positif sebagai kelas yang (benar-benar) positif. Akan tetapi, statistik ini tidak dapat menunjukkan dengan jelas berapa banyak kelas positif (aktual) yang memiliki hasil prediksi yang akurat. Rumus menghitung metrik presisi ini dapat dilihat pada Persamaan 2.4.

$$Presisi = \frac{TP}{TP + FP} \quad (2.4)$$

- *Recall (True Positive Rate) (TPR)*

Statistik yang menampilkan seberapa baik kelas positif (aktual) diramalkan sebagai kelas positif. Namun, kemampuan model untuk meramalkan kelas positif sebagai kelas positif tidak ditangkap oleh statistik ini sebagai kelas

positif (nyata). Rumus menghitung metrik ini terdapat pada Persamaan 2.5.

$$Recall = \frac{TP}{TP + FN} \quad (2.5)$$

• *F1-Score*

Sebuah ukuran yang memperhitungkan kekurangan dari *recall* dan akurasi dalam penilaian kinerja kelas positif dengan menghitung rata-rata harmonik dari keduanya. Namun, karena dua ukuran pertama hanya berkaitan dengan kelas positif, *F1-Score* juga tidak dapat secara tepat mewakili bagaimana kinerja kelas negatif dinilai. kelas yang tidak diinginkan. Menerapkan versi berbobot yang mempertimbangkan kelas lengkap dan distribusinya, bagaimanapun, menyelesaikan semua masalah ini. Metrik itu sendiri memiliki dapat dilihat pada Persamaan 2.6.

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{2TP}{2TP + FP + FN} \quad (2.6)$$

2.10 Penelitian Terdahulu

Tabel 2.1. Penelitian Terdahulu

No	Judul	Penulis	Tahun	Metode	Hasil/Analisa
1	Analisis Sentimen <i>Review Customer</i> Restoran Menggunakan Algoritma <i>Support Vector Machine</i> Dan <i>K-Nearest Neighbor</i>	Amalia, U-maidah, dan Mayasari (2021)	2021	Algoritma <i>Support Vector Machine</i> Dan KNN	Penerapan Algoritma SVM lebih baik dibandingkan dengan menggunakan algoritma KNN dengan hasil akurasi SVM Sebesar 81,92% serta hasil AUC sebesar 0,918. Sedangkan pada algoritma KNN memiliki hasil <i>accuracy</i> sebesar 59,03% dan hasil AUC sebesar 0,590.

Hak Cipta Dilindungi Undang-Undang

- Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu
 - Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
- Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif K

Tabel 2.1. Penelitian Terdahulu (Tabel Lanjutan...)

No	Judul	Penulis	Tahun	Metode	Hasil/Analisa
2	Twitter <i>Sentiment Analysis</i> Terhadap Pengguna <i>e-Commerce</i> Menggunakan <i>Text Mining</i>	Hakim, Marini, Atmaja, Haryadi, dan Suwaryo (2021)	2021	Algoritma <i>Support Vector Machine</i> Dan <i>Naive Bayes</i>	<i>Machine Learning</i> dengan Algoritma <i>Support Vector Machine</i> (Svm) memiliki hasil yang lebih baik dibandingkan dengan menggunakan <i>Naive Bayes</i> dalam mengklasifikasikan data di situs <i>e-Commerce</i> seperti Tokopedia atau Bukalapak yang diperoleh melalui media sosial seperti Twitter yang dibuktikan dengan nilai rata-rata akurasi SVM lebih tinggi dibandingkan dengan Algoritma <i>Naive Bayes</i> .
3	Perbandingan Metode <i>Support Vector Machine</i> Dan <i>Decision Tree</i> Untuk Analisis Sentimen <i>Review Comment</i> Pada Aplikasi <i>Transportasi Online</i>	Rokhman, Berlilana, dan Arsi (2021)	2021	Algoritma <i>Support Vector Machine</i> Dan <i>Decision Tree</i>	Hasil perbandingan analisis sentimen <i>review comment</i> pada aplikasi transportasi <i>online</i> dengan menerapkan metode SVM lebih baik dibandingkan dengan metode <i>Decision Tree</i> dengan nilai akurasi sebesar 90,20%.
4	Review Analisis Produk <i>Marketplace Online</i> Pada Algoritma <i>Support Vector Machine</i>	Permana dan Sahara (2021)	2021	Algoritma <i>Support Vector Machine</i> (SVM)	Pengklasifikasian data dengan menggunakan Algoritma SVM menghasilkan nilai akurasi cukup baik dengan nilai sebesar 75,92%. Dengan demikian, data tersebut dapat membantu pembeli dalam menemukan pasar yang dapat dipercaya.
5	Komparasi Metode K-NN, <i>Support Vector Machine</i> Dan <i>Random Forest</i> Pada <i>e-Commerce</i> Shopee	Suryanto, Martias, dkk. (2021)	2021	Algoritma K-NN, <i>Support Vector Machine</i> Dan <i>Random Forest</i>	Perbandingan klasifikasi menggunakan 3 metode yaitu KNN, SVM Dan <i>Random Forest</i> dengan perolehan nilai akurasi masing-masing sebesar KNN 89,0%, <i>Random Forest</i> 83,0%, dan SVM 89,4%. Hal ini membuktikan bahwa Algoritma SVM memiliki nilai akurasi lebih baik dibandingkan Algoritma KNN dan <i>Random Forest</i> .

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim II

Tabel 2.1. Penelitian Terdahulu (Tabel Lanjutan...)

No	Judul	Penulis	Tahun	Metode	Hasil/Analisa
6	<i>Sentiment Analysis of Community Opinion on Online Store in Indonesia on Twitter using Support Vector Machine (SVM)</i>	Syahputra (2021)	2021	Algoritma Support Vector Machine (SVM)	Algoritma <i>Multiclass (One Vs Rest) Support Vector Machine</i> berhasil mengklasifikasikan tweet dari akun Shopee, Tokopedia, Bukalapak, dan JDId Online Shop ke dalam tiga kelas (negatif, netral, dan positif) dengan akurasi tertinggi di atas 75%. Penggunaan kernel Sigmoid memberikan akurasi terbaik pada dataset Tokopedia dan Bukalapak, sementara kernel Linear lebih baik pada dataset JDId. Hal ini menunjukkan bahwa SVM dapat menjadi pilihan yang efektif dalam analisis sentimen tweet dari berbagai akun toko online.

Berdasarkan Tabel 2.1 maka dapat disimpulkan bahwa Algoritma SVM secara konsisten menghasilkan kinerja yang lebih baik jika dibandingkan dengan beberapa Algoritma klasifikasi lainnya, seperti KNN, *Naive Bayes*, dan *Decision Tree*, dalam berbagai konteks yang membutuhkan analisis sentimen dan klasifikasi data. SVM dapat memberikan wawasan yang berharga bagi *customer* dan bisnis dalam memahami preferensi konsumen dan kebutuhan pasar secara lebih akurat.

Hak Cipta Dilindungi Undang-Undang

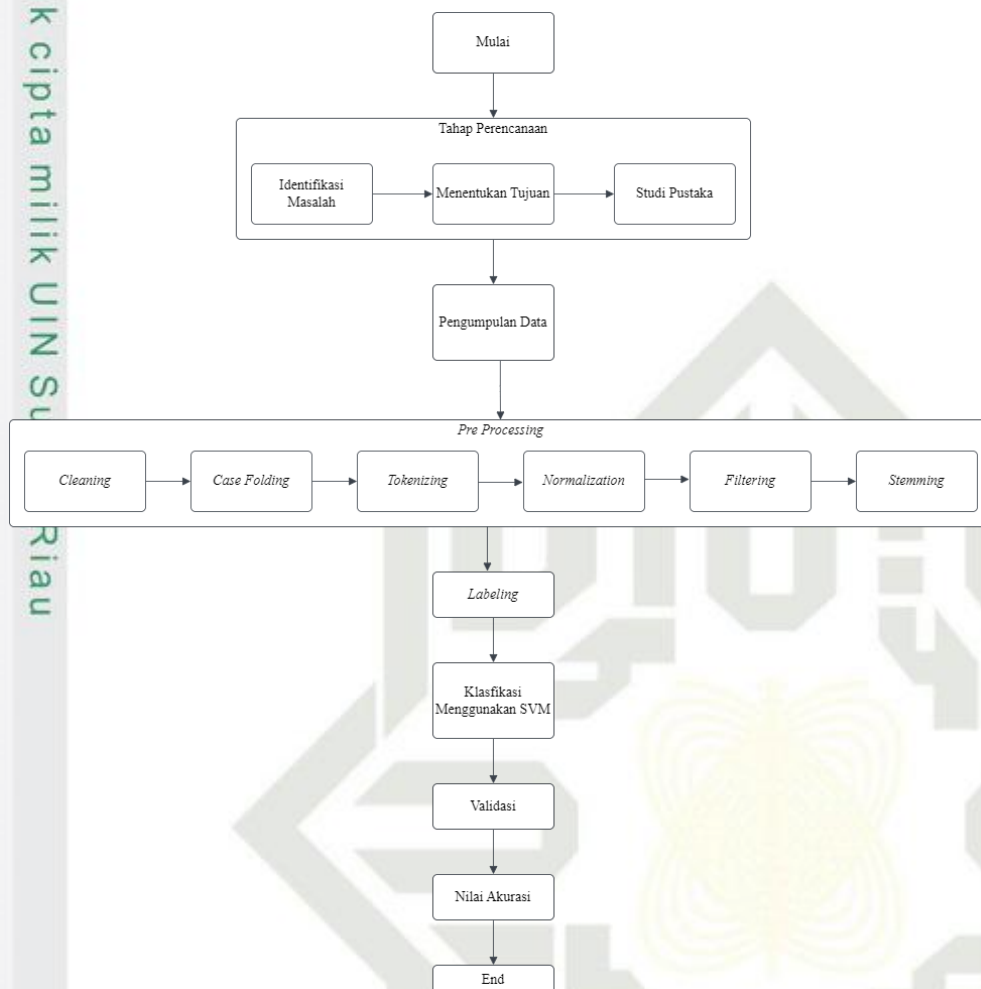
Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif K

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB 3

METODOLOGI PENELITIAN



Gambar 3.1. Metodologi Penelitian

Rangkaian langkah-langkah yang dilakukan dalam menyelesaikan penelitian ini berdasarkan tahapan metodologi sesuai dengan Gambar 3.1. Adapun tahapan metodologi yang dilakukan pada penelitian ini adalah sebagai berikut:

3.1 Tahap Perencanaann

Tahap perencanaan pada penelitian ini dibagi menjadi 3 tahapan yaitu sebagai berikut:

1. Identifikasi Masalah

Tahapan awal pada penelitian yaitu dengan melakukan identifikasi permasalahan yang terdapat pada topik penelitian yang akan dilakukan tersebut.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2. Menentukan Tujuan

Agar sasaran pada Tugas Akhir ini dapat tercapai, langkah berikutnya adalah menentukan tujuan Tugas Akhir, sehingga fokus penulisan dapat lebih terarah. Adapun tujuan pada penelitian Tugas Akhir ini yaitu untuk mengetahui opini yang diberikan oleh *customer review* terhadap produk *brand* sepatu lokal yang diteliti pada *marketplace*.

3. Studi Pustaka

Studi pustaka merujuk pada pengumpulan data sebagai fondasi teoritis yang kuat untuk menyusun laporan dan menganalisis masalah. Dalam penelitian ini, studi pustaka dilakukan dengan menggunakan literatur dari perpustakaan sebagai dasar penelitian lapangan, termasuk buku teks dan *e-book*, jurnal internasional dan nasional yang relevan dengan topik penelitian, skripsi, serta beberapa situs web resmi sebagai dukungan untuk penelitian ini.

3.2 Pengumpulan Data

Dalam pengumpulan data diambil pada setiap *review* yang diberikan *customer* pada masing-masing toko *fashion local* yang berada di *marketplace*. Pengambilan data dilakukan sejak bulan September.

3.3 Preprocessing

Preprocessing dilakukan untuk mendapatkan data yang bersih dan akurat. Langkah-langkah melakukan *preprocessing* yaitu *cleaning*, *case folding*, *tokenizing*, *transforming*, *stemming*, dan *stopword removal*.

1. *Cleaning*

Cleaning adalah proses menghapus semua karakter dari *review customer* yang bukan abjad untuk mengidentifikasi karakter atau simbol yang tidak dipahami atau tidak memiliki arti penting dalam analisis kalimat (Syah dan Witanti, 2022).

2. *Case Folding*

Case folding umumnya dikenal sebagai penulisan huruf dengan mengubah setiap huruf menjadi *Lowercase* (huruf kecil). Hanya huruf A hingga Z yang dapat ditulis ulang, semua karakter lainnya akan hilang (Syah dan Witanti, 2022).

3. *Tokenizing*

Tokenizing merupakan metode pengkodean *string* tertentu berdasarkan penggunaan kata. Hasil dari proses *cleaning* dan *case folding* dilanjutkan dengan tahap *tokenizing*, dimana kalimat *review Customer* diubah menjadi

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim II



- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

kata per kata (Syah dan Witanti, 2022). Tokenisasi adalah proses mengubah judul dokumen menjadi kata kunci atau karakter yang sesuai dengan situasi. Dalam hal apa pun, kalimat dan kata ditulis ulang sesuai dengan spasi yang ada dalam dokumen. Tokenisasi juga mengharuskan semua karakter diubah menjadi huruf kecil dan semua teks diubah menjadi huruf kecil (Sulistio dan Handojo, 2022).

4. *Normalization*

Normalisasi teks merupakan langkah yang tepat dalam memanfaatkan sejumlah besar teks informal dengan efisien. Normalisasi teks melibatkan mengubah teks yang tidak standar menjadi bentuk standar yang umum, sehingga memungkinkan penggunaan metode NLP tradisional dengan metode pemrosesan biasa. Hal ini juga meningkatkan pembangunan model klasifikasi dengan menggunakan data yang telah distandarisasi, bukan data dalam bentuk aslinya yang belum distandarisasi (Mehmood, Essam, Shafi, dan Malik, 2020).

5. *Filtering*

Filtering adalah proses mengidentifikasi kata-kata kunci dari hasil *Tokenizing*, atau yang lebih dikenal dengan proses mengurangi kata-kata kunci menjadi kata-kata yang sesuai dengan konteks. Algoritma yang paling umum digunakan untuk melakukan *filtering* adalah *stopword removal* (Syah dan Witanti, 2022). *Stopword Removal* adalah metode untuk menghilangkan kata-kata umum yang tidak memiliki arti dan tidak diperlukan. *Stopwords* dengan frekuensi kemunculan yang tinggi, seperti kata "dan", "atau", "tapi", "sehingga", dan lain-lain, merupakan ciri-ciri *Stopwords* yang umum. Tujuan dari *Stopwords Removal* adalah untuk menghilangkan kata-kata yang tidak berguna bagi sistem dan tidak memiliki arti, untuk meningkatkan kinerja dan kecepatan proses. *Stopwords* termasuk "penentu", "konjungsi", "proposisi", dan "sejenisnya" (Sulistio dan Handojo, 2022).

6. *Stemming*

Stemming merupakan suatu proses mengubah kata menjadi bentuk dasarnya. Proses ini dilakukan dengan menghapus imbuhan-imbuhan pada kata sehingga kata tersebut kembali ke bentuk dasarnya (Ulgasesa, Negara, dan Tursina, 2022).

7. *Labeling*

Pelabelan berfungsi sebagai dasar dari data mining dengan menggunakan metode SVM, data yang telah didapatkan harus didefinisikan lebih teliti sebagai kalimat yang mengandung nilai positif atau negatif. Setiap *review cus-*



tomor harus memiliki label yang menandakan bahwa *review* tersebut positif atau negatif.

8. Klasifikasi Menggunakan Algoritma SVM

Pada tahap ini, *comment* akan dikumpulkan dalam tiga kelas: positif, negatif, dan universal. Algoritma yang digunakan dalam analisis ini adalah *Support Vector Machine* (SVM). Setelah data diklasifikasikan, kualitas model yang telah dihasilkan dengan tetap mempertahankan akurasi yang dibutuhkan akan dievaluasi.

3.4 Validasi

Tahap ini dengan menggunakan *Confusion Matrix*, yang akan menentukan tingkat akurasi, pengujian dalam penelitian ini (Hasri dan Alita, 2022).

3.5 Hasil

Hasil penelitian data mining yang menggunakan metode SVM menghasilkan sebuah skor untuk proses analisis sentimen produk *fashion* lokal dengan menggunakan *library* Google Colab. Dokumen yang dihasilkan masuk ke dalam salah satu dari tiga kategori penilaian, yaitu penilaian positif, penilaian negatif, dan penilaian netral. Nilai akurasi digunakan untuk meningkatkan akurasi model ketika mengklasifikasikan kelas tertentu sebagai "benar".

3.6 Tahap Dokumentasi

Tahapan ini berisikan dokumentasi dari setiap tahapan yang telah dilakukan serta lampiran sebagai bahan pendukung dalam penelitian ini. Adapun hasil dokumentasi pada penelitian ini dapat dilihat pada Lampiran A dan Lampiran B.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



DAFTAR PUSTAKA

- Agustina, D. A., Subanti, S., dan Zukhronah, E. (2021). Implementasi text mining pada analisis sentimen pengguna twitter terhadap marketplace di indonesia menggunakan algoritma support vector machine. *Indonesian Journal of Applied Statistics*, 3(2), 109–122.
- Agustiono, W., dan Darmawan, A. K. (2023). E-marketplace selection for indonesian msmes: A sentiment analysis approach combining lexicon-based and support vector machine. Dalam *2023 ieee 9th information technology international seminar (itis)* (hal. 1–5).
- Aisy, S. R., dan Prasetyo, B. (2023). Sentiment analysis of the tpks law on twitter using inset lexicon with multinomial naïve bayes and support vector machine based on soft voting. *Recursive Journal of Informatics*, 1(2), 93–101.
- Amalia, B. S., Umaidah, Y., dan Mayasari, R. (2021). Analisis sentimen review pelanggan restoran menggunakan algoritma support vector machine dan k-nearest neighbor. *SITEKIN: Jurnal Sains, Teknologi dan Industri*, 19(1), 28–34.
- Arianto, N. F., Yogaswara, B., dan Sagita, B. H. (2019). Perancangan positioning produk sneakers geoff max bandung berdasarkan perceptual mapping dengan menggunakan metode multidimensional scalling (mds). *eProceedings of Engineering*, 6(2).
- Arif, R. A. (2022). *Pengaruh lifestyle, brand image dan product quality terhadap keputusan pembelian pada sepatu lokal merek geoff max di kabupaten kebumen* (Unpublished doctoral dissertation). Universitas Putra Bangsa.
- Aulia, M., dan Hermawan, A. (2023). Analisis perbandingan algoritma svm, naïve bayes, dan perceptron untuk analisis sentimen ulasan produk tokopedia. *Jurnal Media Informatika Budidarma*, 7(4), 1850–1859.
- Azis, A. N., dan Sofyan, A. (2021). Hubungan antara daya tarik iklan instagram nah project dengan keputusan pembelian konsumen. *Jurnal Riset Manajemen Komunikasi*, 17–22.
- Bangsa, M. T. A., Priyanta, S., dan Suyanto, Y. (2020). Aspect-based sentiment analysis of online marketplace reviews using convolutional neural network. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 14(2), 123–134.
- Deolika, A., Kusriani, K., dan Luthfi, E. T. (2019). Analisis pembobotan kata pada klasifikasi text mining. (*JurTI*) *Jurnal Teknologi Informasi*, 3(2), 179–184.
- Dey, S., Wasif, S., Tonmoy, D. S., Sultana, S., Sarkar, J., dan Dey, M. (2020). A

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu

b. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

©Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim II



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

comparative study of support vector machine and naive bayes classifier for sentiment analysis on amazon product reviews. Dalam *2020 international conference on contemporary computing and applications (ic3a)* (hal. 217–220).

Dharmendra, I. K., Kusuma, I. G. N. A., Dewi, I. A. M. C., dkk. (2023). Implementasi text mining untuk klasifikasi opini alumni pada perguruan tinggi. *Jurnal Teknologi Informasi dan Komputer*, 9(3).

Gunawan, L., Anggreainy, M. S., Wihan, L., Lesmana, G. Y., Yusuf, S., dkk. (2023). Support vector machine based emotional analysis of restaurant reviews. *Procedia Computer Science*, 216, 479–484.

Hakim, A. R., Marini, D., Atmaja, U., Haryadi, D., dan Suwaryo, N. (2021). Twitter sentiment analysis terhadap pengguna e-commerce menggunakan text mining. *SNTEM Semin. Nas. Teknol. Energi dan Miner*, 1, 1227–1237.

Hasri, C. F., dan Alita, D. (2022). Penerapan metode naïve bayes classifier dan support vector machine pada analisis sentimen terhadap dampak virus corona di twitter. *Jurnal informatika dan rekayasa perangkat lunak*, 3(2), 145–160.

Hermawan, A., Jowensen, I., Junaedi, J., dkk. (2023). Implementasi text-mining untuk analisis sentimen pada twitter dengan algoritma support vector machine. *JST (Jurnal Sains Dan Teknologi)*, 12(1), 129–137.

Hidayah, I., Permasari, A. E., dan Wijayanti, N. W. (2019). Sentiment analysis on product review using support vector machine (svm). Dalam *2019 5th international conference on science and technology (icst)* (Vol. 1, hal. 1–4).

Idris, I. S. K., Mustofa, Y. A., dan Salihi, I. A. (2023). Analisis sentimen terhadap penggunaan aplikasi shopee menggunakan algoritma support vector machine (svm). *Jambura Journal of Electrical and Electronics Engineering*, 5(1), 32–35.

Indrawan, M. R., Mulyadi, M., dan Furkan, L. M. (2022). Pengaruh celebrity endorser dan etnosentrisme konsumen terhadap minat beli sepatu lokal “compass” pada mahasiswa s1 di indonesia. *Jurnal Riset Pemasaran*, 1(1), 30–42.

Joseph, V. R., dan Vakayil, A. (2022). Split: An optimal method for data splitting. *Technometrics*, 64(2), 166–176.

Luqyana, W. A., Cholissodin, I., dan Perdana, R. S. (2018). Analisis sentimen cyberbullying pada komentar instagram dengan metode klasifikasi support vector machine. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(11), 4704–4713.

Mehmood, K., Essam, D., Shafi, K., dan Malik, M. K. (2020). An unsupervised lexical normalization for roman hindi and urdu sentiment analysis. *Information*



Processing & Management, 57(6), 102368.

Mukherjee, P., Badr, Y., Doppalapudi, S., Srinivasan, S. M., Sangwan, R. S., dan Sharma, R. (2021). Effect of negation in sentences on sentiment analysis and polarity detection. *Procedia Computer Science*, 185, 370–379.

Musfirroh, D., Khaira, U., Eko, P., Utomo, P., dan Suratno, T. (2021). Sentiment analysis of online lectures in indonesia from twitter dataset using inset lexicon analisis sentimen terhadap perkuliahan daring di indonesia dari twitter dataset menggunakan inset lexicon. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1, 24–33.

Nasution, M. A. A., Sembiring, B. O., dan Elhanafi, A. M. (2023). Penerapan metode weight product dalam menentukan merek sepatu lokal yang paling diminati. *Djtechno: Jurnal Teknologi Informasi*, 4(2), 534–544.

Nugraha, F. A., Harani, N. H., dan Habibi, R. (2020). *Analisis sentimen terhadap pembatasan sosial menggunakan deep learning*. Kreatif.

Nuraliza, H., Pratiwi, O. N., dan Hamami, F. (2022). Analisis sentimen imdb film review dataset menggunakan support vector machine (svm) dan seleksi feature importance. *Jurnal Mirai Management*, 7(1), 1–17.

Permana, R. A., dan Sahara, S. (2021). Review analisis produk marketplace online pada algoritma support vector machine. *Jurnal Ilmiah Informatika*, 6(1), 50–58.

Prayoga, S. H., dan Handoko, W. T. (2023). Kontruksi sosial network analysis untuk menilai kepopuleran brand sepatu di twitter. *Jurnal Sistem Komputer dan Informatika (JSON)*, 5(2), 205–214.

Putri, A. J., Syafira, A. S., Purbaya, M. E., dan Purnomo, D. (2022). Analisis sentimen e-commerce lazada pada jejaring sosial twitter menggunakan algoritma support vector machine. *Jurnal TRINISTIK: Jurnal Teknik Industri, Bisnis Digital, dan Teknik Logistik*, 1(1), 16–21.

Rahmayanti, N. P. (2023). Pengaruh marketplace dan pembayaran digital terhadap tingkat penjualan umkm di kota banjarmasin. *Al-Kalam: Jurnal Komunikasi, Bisnis dan Manajemen*, 10(1), 28–36.

Ramadhani, M. R., dan Setiawati, S. D. (2022). Analisis strategi eco-branding brodo indonesia melalui proyek lestari (studi kasus pada brodo indonesia). *eProceedings of Management*, 9(4).

Rofiqi, M. A., Fauzan, A. C., Agustin, A. P., dan Saputra, A. A. (2019). Implementasi term-frequency inverse document frequency (tf-idf) untuk mencari relevansi dokumen berdasarkan query. *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, 1(2), 58–64.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Rohman, A. N., Musyarofah, R. L., Utami, E., dan Raharjo, S. (2020). Natural language processing on marketplace product review sentiment analysis. Dalam *2020 2nd international conference on cybernetics and intelligent system (i-coris)* (hal. 1–5).

Rokhman, K. A., Berlilana, B., dan Arsi, P. (2021). Perbandingan metode support vector machine dan decision tree untuk analisis sentimen review komentar pada aplikasi transportasi online. *Journal of Information System Management (JOISM)*, 2(2), 1–7.

Sakti, T. A., Sukaris, S., dan Saepuloh, A. (2020). The effect of perceived risk, consumer lifestyle and online trust on the purchase intention of fashion products in instagram social media. *Innovation Research Journal*, 1(2), 133–144.

Siti Nuraeni, Y., dan Irawati, D. (2021). The effect of online customer review, quality product, and promotion on purchasing decision through shopee marketplace (a case study of ubsi college student). *Jurnal Ilmiah Manajemen*, 9(4), 439–450.

Suhartono, D., Purwandari, K., Jeremy, N. H., Philip, S., Arisaputra, P., dan Parmonangan, I. H. (2023). Deep neural networks and weighted word embeddings for sentiment analysis of drug product reviews. *Procedia Computer Science*, 216, 664–671.

Sulistio, H. G., dan Handojo, A. (2022). Aspect-based sentiment analysis pada ulasan ecommerce dengan metode support vector machine untuk mendapatkan informasi sentimen dari beberapa aspek. *Jurnal Infra*, 10(2), 429–433.

Sulistiyawati, E. S., dan Widayani, A. (2020). Marketplace shopee sebagai media promosi penjualan umkm di kota blitar. *Jurnal Pemasaran Kompetitif*, 4(1), 133.

Suryanto, S., Martias, M., dkk. (2021). Komparasi metode k-nn, support vector machine dan random forest pada e-commerce shopee. *INSANtek*, 2(1), 15–21.

Syah H., dan Witanti, A. (2022). Analisis sentimen masyarakat terhadap vaksinasi covid-19 pada media sosial twitter menggunakan algoritma support vector machine (svm). *Jurnal Sistem Informasi Dan Informatika (Simika)*, 5(1), 59–67.

Syahputra, H. (2021). Sentiment analysis of community opinion on online store in indonesia on twitter using support vector machine algorithm (svm). Dalam *Journal of physics: Conference series* (Vol. 1819, hal. 012030).

Ulgasesa, R., Negara, A. B. P., dan Tursina, T. (2022). Pengaruh stemming terhadap performa klasifikasi sentimen masyarakat tentang kebijakan new nor-

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



mal. *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, 10(3), 286–293.

Urva, G., Albanna, I., Sungkar, M. S., Gunawan, I. M. A. O., Adhicandra, I., Ramadhan, S., ... others (2023). *Penerapan data mining di berbagai bidang: Konsep, metode, dan studi kasus*. PT. Sonpedia Publishing Indonesia.

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif K

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



UIN SUSKA RIAU



LAMPIRAN A

DATA PENELITIAN

Tabel A.1. Data Penelitian

No	ID Barang	Date	Comment	Rating
0	Saint Barkley 15	25/01/2023 18:58	Mantep sepatunya kokoh tapi terlalu ngepres masuknya jadi susah kalo mau pakek.	5
1	Saint Barkley 15	30/08/2022 18:52	Paketan udh sampeee, sehari sampeee. Sepatunyya bagus, jahitan rapih. Bahan suede, mantul lahh. Tengkyu	5
2	Saint Barkley 15	19/06/2021 17:37	Pengirimannya cepett bangettttt, ga nyampe sehari, adminnya juga ramah	5
3	Saint Barkley 15	15/07/2022 15:32	Pengiriman cepat, sepatu juga bagus, cantik, anak saya sangat suka sepatunya, terima kasih shopee.	5
4	Saint Barkley 15	13/10/2022 16:03	Ukuran pas, mantap, bagus, oke, nice	5
5	Saint Barkley 15	01/03/2023 20:02	Mantapp lah harga kualitas terjangkau kualitas jg tanggung buat tempur	5
6	Saint Barkley 15	27/07/2022 18:40	Maaf fot/vid melenceng pokonya sebagai itu sepatunya gapernah gagal beli di toko ini... Nextime jadi langganan	5
7	Saint Barkley 15	22/12/2021 17:58	Pengemasannya cepet. Sepatu sesuai emg SB kesukaan suami saya	5
...	5
7296	Brodo	27/05/2022 15:19	Alhamdulillah syudah sampe barangnya mantab bangett the best pokoknya wah nggak nyangka juga dapet kaus kaki, makasih. Mawa	5
7297	Brodo	11/01/2023 10:45	Barang bagus...toko amanah..trims	5
7298	Brodo	09/06/2022 14:13	Pengiriman cepat pengemasan rapi barang lokal punya mantap original pokoknya mantap	5
7299	Brodo	04/12/2022 19:53	Bagus. Tapi kurirnya lama kalo kirim	4
7300	Brodo	30/07/2022 15:32	Pengiriman nya doank yang lama, barangnya oke, yang Laen dpt kaos kaki cuman saya doank gw ga dpt gpp lah	4
7301	Brodo	12/04/2022 07:16	Ini sepatu keren banget, anakku suka banget, nyampe nya juga cepet.. Next pasti order lagi. Maaf foto tdk sesuai..	5

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LAMPIRAN B

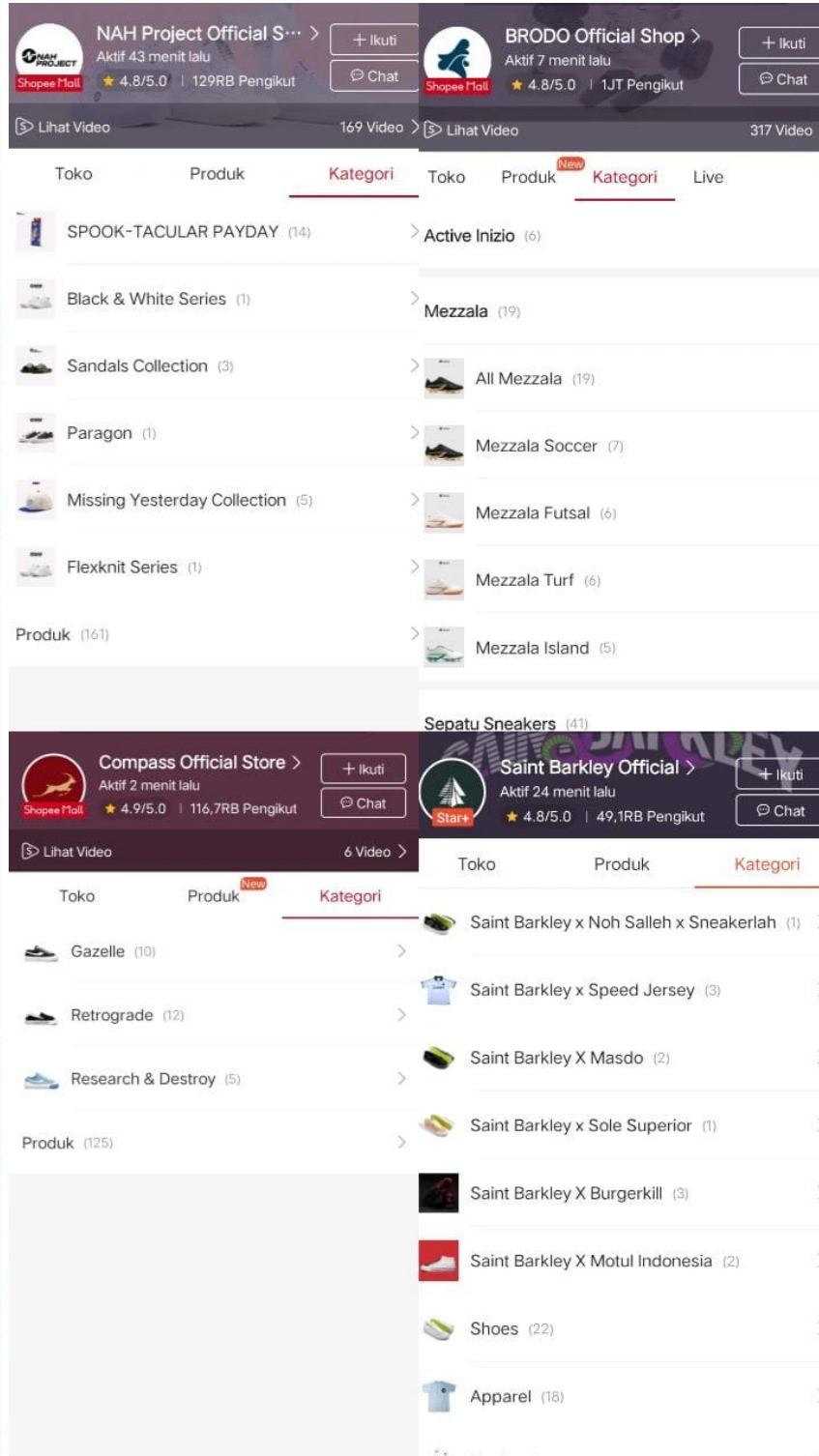
PRODUK PENELITIAN

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim II

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar B.1. Produk Penelitian



DAFTAR RIWAYAT HIDUP



Nadila Andrika lahir di Muara Labuh, pada tanggal 07 November 2000. Peneliti merupakan anak tunggal dari bapak Hendri Musfa dan ibu Marniati. Peneliti menyelesaikan Sekolah Dasar di SD Negeri 001 Bandar Buat, Padang. Lalu pindah sekolah dan lulus di SD Negeri 147 Pekanbaru pada tahun 2013. Setelah itu peneliti melanjutkan pendidikan tingkat SMP di SMP Negeri 23 Pekanbaru dan lulus pada tahun 2016. Kemudian, setelah menyelesaikan pendidikan di SMP Negeri 23 Pekanbaru, pada tahun 2016 tersebut peneliti melanjutkan pendidikan tingkat SMK di SMK Muhammadiyah 2 Pekanbaru pada tahun 2019. Peneliti tidak langsung melanjutkan pendidikan ke Universitas, selama *gapyear* peneliti mengikuti berbagai *workshop* dan pelatihan kompetensi. Setelah itu peneliti pun melanjutkan pendidikan dengan mendaftar ke berbagai Universitas yang ada di Riau. Alhamdulillah pada tahun 2020 peneliti diterima menjadi mahasiswa di Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau melalui jalur Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN). Selama menjadi mahasiswa, peneliti aktif berorganisasi internal kampus serta seminar-seminar yang ada dikampus. Peneliti tergabung dalam Himpunan Mahasiswa Sistem Informasi, *Information System Organization Culture (ISOC)*, dan *Information System Networking Club (ISNC Research)*, serta Peneliti pernah melakukan Kerja Praktek di Laboratorium Prodi Sistem Informasi yang kemudian menjadi bagian dari asisten laboratorium. Akhir kata peneliti mengucapkan rasa syukur yang tak terhingga serta ribuan terimakasih atas bantuan dari seluruh pihak yang terkait sehingga peneliti bisa menyelesaikan pendidikan S-1 dalam waktu 8 semester setelah berhasil menyelesaikan Tugas Akhir dengan judul "Analisis Sentimen Prooduk *Fashion* Lokal Pada *Marketplace* Dengan Klasifikasi *Support Vector Machine (SVM)*".

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.