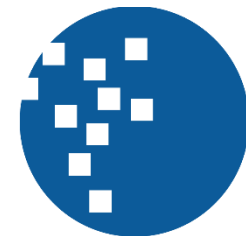


**ANALISIS ULASAN APLIKASI SELLER CENTER
TOKOPEDIA DAN TIKTOK SHOP DI GOOGLE PLAY
SETELAH INTEGRASI LAYANAN DENGAN
PERBANDINGAN LATENT DIRICHLET ALLOCATION
DAN NON-NEGATIVE MATRIX FACTORIZATION**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Vernonia Novianna Putri

00000054293

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

**ANALISIS ULASAN APLIKASI SELLER CENTER
TOKOPEDIA DAN TIKTOK SHOP DI GOOGLE PLAY
SETELAH INTEGRASI LAYANAN DENGAN
PERBANDINGAN LATENT DIRICHLET ALLOCATION
DAN NON-NEGATIVE MATRIX FACTORIZATION**



SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Komputer

Vernonia Novianna Putri

00000054293

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2025

i

Analisis Ulasan Aplikasi Seller Center Tokopedia dan Tiktok Shop di Google..., Vernonia
Novianna Putri, Universitas Multimedia Nusantara

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Vernonia Novianna Putri

Nomor Induk Mahasiswa : 00000054293

Program Studi : Sistem Informasi

Skripsi dengan judul:

Analisis Ulasan Aplikasi Seller Center Tokopedia dan Tiktok Shop di Google Play setelah Integrasi Layanan dengan Perbandingan Latent Dirichlet Allocation dan Non-Negative Matrix Factorization

Merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 28 November 2025



(Vernonia Novianna Putri)

LEMBAR PERNYATAAN PENGGUNAAN BANTUAN KECERDASAN ARTIFISIAL (AI)

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Vernonia Novianna Putri
NIM : 00000054293
Program Studi : Sistem Informasi
Judul Skripsi : Analisis Ulasan Aplikasi Seller Center Tokopedia dan Tiktok Shop di Google Play setelah Integrasi Layanan dengan Perbandingan Latent Dirichlet Allocation dan Non-Negative Matrix Factorization

Dengan ini saya menyatakan secara jujur menggunakan bantuan Kecerdasan Artifisial (AI) dalam pengerjaan Tugas/Project/Tugas Akhir*(coret salah satu) sebagai berikut (beri tanda centang yang sesuai):

- ☐ Menggunakan AI sebagaimana diizinkan untuk membantu dalam menghasilkan ide-ide utama saja
- ☒ ~~Menggunakan AI sebagaimana diizinkan untuk membantu menghasilkan teks pertama saja~~
- ☐ Menggunakan AI untuk menyempurnakan sintaksis dan tata bahasa untuk pengumpulan tugas
- ☒ ~~Karena tidak diizinkan: Tidak menggunakan bantuan AI dengan cara apa pun dalam pembuatan tugas~~

Saya juga menyatakan bahwa:

- (1) Menyerahkan secara lengkap dan jujur penggunaan perangkat AI yang diperlukan dalam tugas melalui Formulir Penggunaan Perangkat Kecerdasan Artifisial (AI)
- (2) Saya mengakui bahwa saya telah menggunakan bantuan AI dalam tugas saya baik dalam bentuk kata, paraphrase, penyertaan ide atau fakta penting yang disarankan oleh AI dan saya telah menyantumkan dalam sitasi serta referensi

Terlepas dari pernyataan di atas, tugas ini sepenuhnya merupakan karya saya sendiri

Tangerang, 28 November 2025
Yang Menyatakan,



Vernonia Novianna Putri

HALAMAN PENGESAHAN

Tugas akhir dengan judul

Analisis Ulasan Aplikasi Seller Center Tokopedia dan Tiktok Shop di Google Play setelah Integrasi Layanan dengan Perbandingan Latent Dirichlet Allocation dan Non-Negative Matrix Factorization

Oleh

Nama : Vernonia Novianna Putri
NIM : 00000054293
Program Studi : Sistem Informasi
Fakultas : Teknik & Informatika

Telah diujikan pada hari Jumat, 19 Desember 2025

Pukul 08.00 s.d 10.00 dan dinyatakan

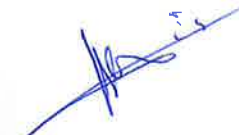
LULUS

Dengan susunan penguji sebagai berikut.

Ketua Sidang

 25/12/2025
Ahmad Faza, S.Kom., M.T.I.
0312019501

Penguji

 25-12-2025
Rudi Sutomo, S.Kom., M.Si., M.Kom.
0222057501

Pembimbing

 25/12/25
Antonius Sony Eko Nugroho, S.T., M.T.
0314058005

Ketua Program Studi Sistem Informasi

 5/126
Ririn Ikana Desanti, S.Kom., M.Kom
0313058001

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama : Vernonia Novianna Putri

NIM : 00000054293

Program Studi : Sistem Informasi

Jenjang : S1

Judul Karya Ilmiah : Analisis Ulasan Aplikasi Seller Center Tokopedia dan Tiktok Shop di Google Play setelah Integrasi Layanan dengan Perbandingan Latent Dirichlet Allocation dan Non-Negative Matrix Factorization

Menyatakan dengan sesungguhnya bahwa saya bersedia* (pilih salah satu):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
- ☒ Hanya dapat diakses secara internal Universitas Multimedia Nusantara
- ☐ Embargo publikasi karya ilmiah dalam kurun waktu 3 tahun.

Tangerang, 22 Desember 2025



(Vernonia Novianna Putri)

* Pilih salah satu

** Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

KATA PENGANTAR

Puji Syukur atas selesainya penulisan tugas akhir “Analisis Ulasan Aplikasi Seller Center Tokopedia dan Tiktok Shop di Google Play setelah Integrasi Layanan dengan Perbandingan Latent Dirichlet Allocation dan Non-Negative Matrix Factorization” yang dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer (S.Kom.) pada Program Studi Sistem Informasi, Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan tugas akhir ini, sangatlah sulit bagi saya untuk menyelesaikan tugas akhir ini. Oleh karena itu, saya mengucapkan terima kasih kepada

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Universitas Multimedia Nusantara.
4. Bapak Antonius Sony Eko Nugroho, S.T., M.T., sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.

Semoga karya ilmiah ini memberikan wawasan baru dan dapat digunakan untuk sebaik-baiknya.

Tangerang, 28 November 2025



Vernonia Novianna Putri

ANALISIS ULASAN APLIKASI SELLER CENTER TOKOPEDIA DAN TIKTOK SHOP DI GOOGLE PLAY SETELAH INTEGRASI LAYANAN DENGAN PERBANDINGAN LATENT DIRICHLET ALLOCATION DAN NON-NEGATIVE MATRIX FACTORIZATION

Vernonia Noviana Putri

ABSTRAK

Integrasi layanan seller center TikTok Shop dan Tokopedia pada 8 April 2025 menimbulkan berbagai tanggapan dari pengguna yang sebelumnya beroperasi secara terpisah di kedua platform. Ulasan aplikasi di Google Play melonjak naik dari 8.366 ulasan pada 6 bulan sebelum integrasi menjadi 13.446 ulasan setelah 6 bulan. Penelitian ini menganalisis sentimen berdasarkan aspek atau topik dominan yang muncul pada data ulasan pada kedua periode tersebut melalui pemodelan topik dan menganalisis sentimen pada topik tersebut untuk meningkatkan aspek atau fitur aplikasi dari *feedback* pengguna aplikasi. Penelitian ini menerapkan pendekatan CRISP-DM pada analisis sentimen serta pemodelan topik (*topic modeling*) Latent Dirichlet Allocation (LDA) dan Non-negative Matrix Factorization (NMF) menggunakan metode evaluasi *coherence score c_v* dan *u_mass*. Corpus dokumen menggunakan pembobotan kata TFIDF dan BoW dengan CountVectorizer. Algoritma Support Vector Machine (SVM) digunakan untuk menganalisis sentimen berdasarkan representasi topik yang dihasilkan. SMOTE dan GridSearch digunakan untuk meningkatkan performa model analisis sentimen. Hasil evaluasi menunjukkan bahwa NMF menghasilkan nilai *coherence cv* lebih tinggi dari LDA pada kedua *feature extraction* TFIDF dan Count Vectorizer. Pemodelan topik NMF dengan Count Vectorizer menghasilkan 12 jumlah topik dengan nilai *c_v* tertinggi 0.744 setelah POS tagging sehingga meningkatkan tahap ekstraksi aspek ABSA. Kombinasi pemodelan topik dan sentimen analysis menghasilkan analisis yang lebih dalam dengan menunjukan aspek atau fitur yang memiliki mayoritas sentimen tertentu.

Kata kunci: LDA, NMF , *Topic Modeling*

SENTIMENT ANALYSIS OF THE TOKOPEDIA AND TIKTOK SHOP SELLER CENTER APPLICATIONS ON GOOGLE PLAY AFTER INTEGRATION OF SERVICES BY COMPARING LATENT DIRICHLET ALLOCATION AND NON-NEGATIVE MATRIX FACTORIZATION

Vernonia Novianna Putri

ABSTRACT (English)

The integration of TikTok Shop and Tokopedia seller center services on April 8, 2025, elicited various responses from users who previously operated separately on both platforms. App reviews on Google Play surged from 8,366 reviews in the 6 months prior to integration to 13,446 reviews in the 6 months following integration. This study analyzes sentiment based on dominant aspects or topics emerging from review data across both periods through topic modeling and examines sentiment within those topics to enhance app features or aspects based on user feedback. This study applied the CRISP-DM approach to topic modeling (topic modeling) Latent Dirichlet Allocation (LDA) and Non-negative Matrix Factorization (NMF) using the coherence score c_v and u_mass evaluation methods. The document corpus created from TFIDF and BoW word weighting with CountVectorizer. The Support Vector Machine (SVM) algorithm is used to analyze sentiment based on the generated topic representation. SMOTE and GridSearch are used to improve the performance of the sentiment analysis model. The evaluation results show that NMF produces a higher coherence c_v value than LDA in both TF-IDF and Count Vectorizer feature extraction. NMF with Count Vectorizer obtained 12 topics with a c_v 0.744 value after POS tagging, thereby improving the ABSA extraction aspect. The combination of topic modeling and sentiment analysis produced a more in-depth analysis by showing aspects or features that had a majority of certain sentiments.

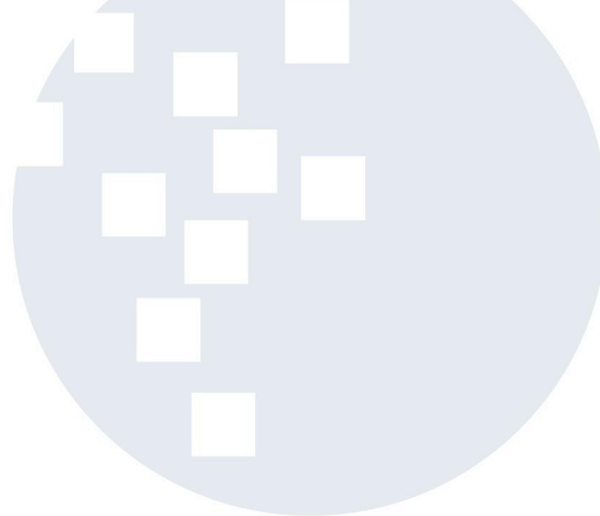
Keywords: LDA, NMF, Topic Modeling

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PERSETUJUAN	iv
HALAMAN PENGESAHAN	v
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	vi
KATA PENGANTAR.....	vii
ABSTRAK	viii
<i>ABSTRACT (English)</i>	ix
DAFTAR ISI.....	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	7
1.3 Batasan Masalah	7
1.4 Tujuan dan Manfaat Penelitian	8
1.4.1 Tujuan Penelitian	8
1.4.2 Manfaat Penelitian	8
1.5 Sistematika Penulisan	9
BAB II LANDASAN TEORI	10
2.1 Penelitian Terdahulu	10
2.2 Teori tentang Topik Skripsi	17
2.2.1 Tokopedia dan Tik tok Shop Seller Center	17
2.2.2 Text Classification	18
2.2.3 Topic Modeling.....	18
2.2.4 Analisis Sentimen	19
2.2.5 Aspect-based Sentiment Analysis	19
2.3 Teori tentang Framework/Algoritma yang digunakan	20
2.3.1 Machine Learning	20
2.3.2 Non-negative Matrix Factorization (NMF)	21

2.3.3	Latent Dirichlet Allocation (LDA).....	21
2.3.4	Support Vector Machine (SVM).....	22
2.3.5	CRISP-DM.....	23
2.4	Teori tentang tools/software yang digunakan	24
2.4.1	Google Colab	24
2.4.2	Python	25
BAB III	METODOLOGI PENELITIAN	26
3.1	Gambaran Umum Objek Penelitian.....	26
3.2	Metode Penelitian.....	26
3.3	Teknik Pengumpulan Data	33
3.4	Variabel Penelitian.....	35
3.5	Teknik Analisis Data.....	35
BAB IV	ANALISIS DAN HASIL PENELITIAN	39
4.1	Business Understanding	39
4.1.1	Analisa Proses Bisnis	39
4.1.2	<i>User Requirements</i>	41
4.2	Data Understanding.....	41
4.2.1	Data Collection	41
4.2.2	Data Description.....	45
4.3	Data Preparation.....	50
4.3.1	Text Cleaning	50
4.3.2	Normalisasi dan stopwords removal.....	52
4.3.3	PoS Tagging.....	54
4.3.4	Stemming	55
4.3.5	Labeling Sentimen	56
4.4	Modeling	57
4.5	Evaluation.....	61
4.5.1	Coherence score.....	62
4.5.2	Ekstraksi Fitur dan Sentiment analysis.....	76
4.6	Deployment.....	82
4.7	Hasil dan Pembahasan.....	85
BAB V	SIMPULAN DAN SARAN	91

5.1 Simpulan.....	91
5.2 Saran.....	92
DAFTAR PUSTAKA.....	93
LAMPIRAN.....	100
Lampiran A Turnitin Similarity Report.....	101
Lampiran B Form Konsultasi Bimbingan	102
Lampiran C Formulir Penggunaan Perangkat Kecerdasan Artifisial (AI)	103



UMN
 UNIVERSITAS
 MULTIMEDIA
 NUSANTARA

DAFTAR TABEL

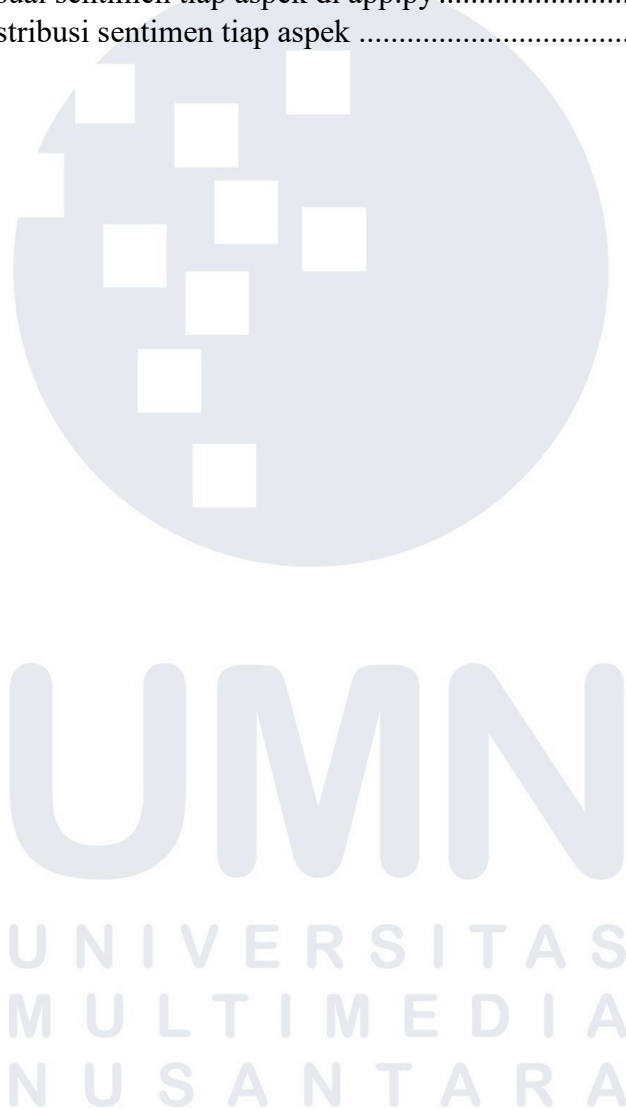
Tabel 2.1 Penelitian Terdahulu	10
Tabel 3.1 Perbandingan Framework <i>Data Mining</i>	27
Tabel 3.2 Nama kolom tabel data hasil <i>scraping</i>	34
Tabel 4.1 Struktur data ulasan.....	45
Tabel 4.2. Hasil data ulasan setelah text cleaning.....	51
Tabel 4.3 Hasil ulasan setelah stopwords removal.....	53
Tabel 4.4 Hasil ulasan setelah PoS Tagging	54
Tabel 4.5 Hasil data ulasan setelah stemming dan tokenization.....	56
Tabel 4.6 Jumlah <i>term</i> yang pada corpus.....	64
Tabel 4.7 Perbandingan hasil nilai coherence score <i>c_v</i> dan <i>u_mass</i> LDA+TFIDF terhadap <i>n_component</i> tertinggi.....	65
Tabel 4.8 Perbandingan hasil nilai coherence score <i>c_v</i> dan <i>u_mass</i> LDA+Count Vectorizer terhadap <i>n_component</i> tertinggi.....	66
Tabel 4.9 Perbandingan hasil nilai coherence score <i>c_v</i> dan <i>u_mass</i> NMF+TFIDF terhadap <i>n_component</i> tertinggi.....	68
Tabel 4.10 Perbandingan hasil nilai coherence score <i>c_v</i> dan <i>u_mass</i> NMF+Count Vectorizer terhadap <i>n_component</i> tertinggi.....	70
Tabel 4.11 Perbandingan coherence score <i>c_v</i> sebelum POS tagging dan setelah POS tagging.	86
Tabel 4.12 Aspek dominan dan frekuensi pada topik.....	87
Tabel 4.13 Perbandingan hasil akurasi, presisi, recall, dan f1-score sentiment analysis pada aspek	89

DAFTAR GAMBAR

Gambar 1.1 Pertumbuhan <i>e-commerce global sales</i>	1
Gambar 1.2 Nilai transaksi <i>e-commerce</i> di Indonesia	2
Gambar 1.3 Data market share sistem operasi perangkat seluler di Indonesia Desember 2024 - November 2025 oleh StatCounter[25].....	4
Gambar 2.1 NMF matrix[64]	21
Gambar 2.2 Latent Dirichlet Allocation Model[60]	22
Gambar 3.1 Alur Penelitian.....	29
Gambar 4.1 Fitur pada aplikasi TikTok dan Tokopedia Seller Center	40
Gambar 4.2 Contoh data dalam bentuk ulasan aplikasi	42
Gambar 4.3 Import library dan inisiasi variabel	43
Gambar 4.4 Fungsi reviews untuk mengambil ulasan	44
Gambar 4.5 fungsi label periode	45
Gambar 4.6 File CSV ulasan aplikasi	45
Gambar 4.7 Tren jumlah ulasan perbulan sebelum dan sesudah integrasi <i>seller center</i>	47
Gambar 4.8 Distribusi rating pada periode sebelum integrasi aplikasi <i>seller center</i>	48
Gambar 4.9 Wordcloud pada periode sebelum integrasi	48
Gambar 4.10 Distribusi rating pada periode setelah integrasi aplikasi <i>seller center</i>	49
Gambar 4.11 Wordcloud pada periode setelah integrasi	49
Gambar 4.12 Library yang digunakan untuk text cleaning.....	50
Gambar 4.13 Fungsi <code>basic_cleaning</code> untuk text cleaning.....	51
Gambar 4.14 Tahap stopword removal.....	53
Gambar 4.15 Tahap stopword removal.....	53
Gambar 4.16 Tahap stemming	55
Gambar 4.17 Tahap stemming	55
Gambar 4.18 Distribusi sentimen.....	57
Gambar 4.19 pustaka (library) yang digunakan.....	57
Gambar 4.20 Kode Vectorization	58
Gambar 4.21 Kode Latent Dirichlet Allocation (LDA).....	59
Gambar 4.22 Kode Non-Negative Matrix Factorization (NMF)	59
Gambar 4.23 Menghasilkan aspek berdasarkan bobot kata pada topik	60
Gambar 4.24 Kode split train-test	60
Gambar 4.25 Kode definisi parameter grid.....	61
Gambar 4.26 Kode train model classification.....	61
Gambar 4.27 Potongan kode untuk coherence score dengan pembobotan kata <code>CountVectorizer</code>	63
Gambar 4.28 Potongan kode untuk coherence score dengan pembobotan kata TFIDF.....	63

Gambar 4.29 Hasil Coherence score c_v dan u_mass terhadap k=2 hingga k=20 LDA dan TFIDF.....	64
Gambar 4.30 Hasil Coherence score c_v dan u_mass terhadap k=2 hingga k=20 LDA dan TFIDF dengan POS tagging.....	64
Gambar 4.31. Hasil Coherence score c_v dan u_mass terhadap k=2 hingga k=20 LDA dan CountVectorizer	66
Gambar 4.32. Hasil Coherence score c_v dan u_mass terhadap k=2 hingga k=20 LDA dan CountVectorizer dengan POS tagging	66
Gambar 4.33 Hasil Coherence score c_v dan u_mass terhadap k=2 hingga k=20 NMF dan TFIDF	68
Gambar 4.34 Hasil Coherence score c_v dan u_mass terhadap k=2 hingga k=20 NMF dan TFIDF setelah POS tagging.....	68
Gambar 4.35 Hasil Coherence score c_v dan u_mass terhadap k=2 hingga k=20 NMF dan CountVectorizer.....	70
Gambar 4.36 Hasil Coherence score c_v dan u_mass terhadap k=2 hingga k=20 NMF dan CountVectorizer setelah POS tagging	70
Gambar 4.37 Hasil coherence u_mass pada LDA (kiri) dan NMF (kanan).....	72
Gambar 4.38 Hasil coherence u_mass pada LDA (kiri) dan NMF (kanan) setelah PoS	73
Gambar 4.39 Hasil coherence c_v LDA (kiri) dan NMF (kanan)	74
Gambar 4.40 Hasil coherence c_v LDA (kiri) dan NMF (kanan) setelah POS tagging.....	74
Gambar 4.41 Perbandingan TFIDF dan CountVect terhadap coherence c_v LDA sebelum (kiri) dan setelah POS tagging (kanan).....	75
Gambar 4.42 Perbandingan TFIDF dan CountVect terhadap coherence c_v NMF sebelum (kiri) dan setelah POS tagging (kanan).....	75
Gambar 4.43 Top 10 words dalam TFIDF dan n_component = 20 sebelum POS tagging.....	77
Gambar 4.44 Top 10 words dalam Count Vectorizer dan n_component = 17 sebelum POS tagging.	77
Gambar 4.45 Top 10 words dalam TFIDF dan n_component = 4 setelah POS tagging.....	77
Gambar 4.46 Top 10 words dalam TFIDF dan n_component = 12 setelah POS tagging.....	78
Gambar 4.47 Frekuensi tertinggi term unigram dan bigram pada corpus setelah POS tagging	79
Gambar 4.48 Confusion matrix dan classification report aspek performa_toko dan produk	80
Gambar 4.49 Confusion matrix dan classification report aspek pesanan_pengiriman dan integrasi layanan.....	80
Gambar 4.50 Confusion matrix dan classification report aspek keuangan dan pendaftaran_pengaturan_akun	81

Gambar 4.51 Hasil akurasi, presisi, recall, dan f1-score rata-rata pada data sebelum SMOTE.....	81
Gambar 4.52 Hasil akurasi, presisi, recall, dan f1-score rata-rata pada data setelah SMOTE	82
Gambar 4.53 Tampilan web deployment.....	83
Gambar 4.54 Potongan script app.py	84
Gambar 4.55 Potongan script preprocessing_module.....	84
Gambar 4.56 Visual sentimen tiap aspek di app.py	85
Gambar 4.57 Distribusi sentimen tiap aspek	90



DAFTAR LAMPIRAN

Lampiran A Turnitin Similarity Report	101
Lampiran B Form Konsultasi Bimbingan	102
Lampiran C Formulir Penggunaan Perangkat Kecerdasan Artifisial (AI)	103



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA