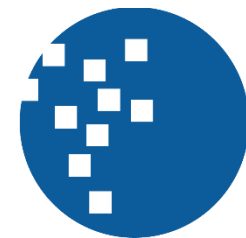


**SENTIMEN ANALISIS KEPUASAN PENGGUNA APLIKASI
TRAVEL AGENCY MENGGUNAKAN ALGORITMA
MACHINE LEARNING DAN OPTIMASI *HYPERPARAMETER*
*TUNING***



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

Skripsi

Fathurrahman Bagaskara

00000048164

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

2025

**SENTIMEN ANALISIS KEPUASAN PENGGUNA APLIKASI
TRAVEL AGENCY MENGGUNAKAN ALGORITMA
MACHINE LEARNING DAN OPTIMASI HYPERPARAMETER
TUNING**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

Skripsi

Diajukan sebagai Salah Satu Syarat untuk Memperoleh

Gelar Sarjana Komputer

Fathurrahman Bagaskara

00000048164

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS TEKNIK INFORMATIKA

UNIVERSITAS MULTIMEDIA NUSANTARA

TANGERANG

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Fathurrahman Bagaskara

Nomor Induk Mahasiswa : 00000048164

Program studi : Sistem informasi

Skripsi dengan judul:

”Sentimen Analisis Kepuasan Pengguna Aplikasi *Travel Agency* Menggunakan Algoritma *Machine Learning* dan Optimasi *Hyperparameter Tuning*”

merupakan hasil karya saya sendiri bukan plagiat dari karya 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 skripsi maupun dalam penulisan laporan skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas Akhir yang telah saya tempuh.

Tangerang, 15 Mei 2025



Fathurrahman Bagaskara

HALAMAN PERSETUJUAN

SENTIMEN ANALISIS KEPUASAN PENGGUNA APLIKASI *TRAVEL AGENCY* MENGGUNAKAN ALGORITMA *MACHINE LEARNING* DAN OPTIMASI *HYPERPARAMETER TUNING*

Oleh

Nama : Fathurrahman Bagaskara

NIM : 00000048164

Program Studi : Sistem Informasi

Fakultas : Teknik dan Informatika

Telah disetujui untuk diajukan pada

Sidang Ujian Skripsi Universitas Multimedia Nusantara

Tangerang, 4 Juni 2025

Pembimbing



Rudi Sutomo, S.Kom., M.Si., M.Kom.f
0222057501

Ketua Sistem Informasi



Ririn Ikana Desanti, S. Kom., M. Kom.
313058001

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

HALAMAN PENGESAHAN

Sentimen Analisis Kepuasan Pengguna Aplikasi *Travel Agency* Menggunakan Algoritma *Machine Learning* dan Optimasi *Hyperparameter Tuning*.

Oleh

Nama : Fathurrahman Bagaskara
NIM : 00000048164
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika

Telah diujikan pada hari Selasa, 10 Juni 2025

Pukul 15.00 s.d 17.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut.

Ketua Sidang



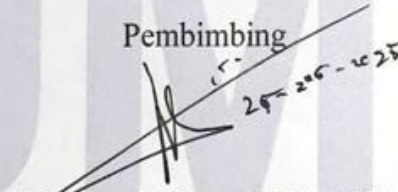
Melissa Indah Fianty, S.Kom., M.MSI
0313019201

Penguji



Dinar Ajeng Kristiyanti, S.Kom, M.Kom
0330128801

Pembimbing



Rudi Sutomo, S.Kom., M.Si., M.Kom.
0222057501

Ketua Program Studi Sistem Informasi



Ririn Ikana Desanti, S.Kom., M.Kom
0313058001

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama : Fathurrahman Bagaskara
NIM : 00000048164
Program Studi : Sistem Informasi
Jenjang : S1
Judul Karya Ilmiah : SENTIMEN ANALISIS KEPUASAN
PENGGUNA APLIKASI *TRAVEL AGENCY*
MENGUNAKAN ALGORITMA
MACHINE LEARNING DAN OPTIMASI
HYPERPARAMETER TUNING

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, 25 November 2023

Yang menyatakan,



Fathurrahman Bagaskara

KATA PENGANTAR

Puji Syukur atas selesainya penulisan Skripsi ini dengan judul “ **SENTIMEN ANALISIS KEPUASAN PENGGUNA APLIKASI TRAVEL AGENCY MENGGUNAKAN ALGORITMA MACHINE LEARNING DAN OPTIMASI HYPERPARAMETER TUNING**” dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Strata 1 Jurusan Sistem Informasi pada Fakultas Teknik Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai dengan penyusunan tugas akhir ini, sangatlah sulit bagi saya untuk menyelesaikan tugas akhir ini. Saya mengucapkan terima kasih kepada

1. Bapak Dr. Andrey Andoko, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng Niki Prastomo, S.T., M.Sc selaku Dekan Fakultas Teknik & Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti S.Kom., M.Kom., selaku Ketua Program Studi Universitas Multimedia Nusantara.
4. Bapak Rudi Sutomo, S.Kom., M.Si., M.Kom. 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.
6. Kepada Gracia Jane yang telah memberikan support emosional.

Semoga karya ilmiah ini dapat bermanfaat dan menjadi sumber informasi bagi para pembaca.

Tangerang, 6 Juni 2025



Fathurrahman Bagaskara

SENTIMEN ANALISIS KEPUASAN PENGGUNA APLIKASI TRAVEL AGENCY MENGGUNAKAN ALGORITMA MACHINE LEARNING DAN OPTIMASI HYPERPARAMETER TUNING

Fathurrahman Bagaskara

ABSTRAK

Pertumbuhan aplikasi *Online Travel Agent* (OTA) seperti Traveloka, Tiket.com, dan Agoda mendorong kebutuhan untuk memahami sentimen pengguna secara akurat. Komparasi lima algoritma klasifikasi machine learning diperlukan untuk mengetahui model yang paling efektif dalam menganalisis opini publik, karena setiap algoritma memiliki karakteristik dan sensitivitas berbeda terhadap data teks ulasan. Analisis sentimen menjadi penting untuk mengevaluasi persepsi pengguna dan meningkatkan kualitas layanan digital.

Penelitian ini mengimplementasikan lima algoritma Naive Bayes, K-Nearest Neighbor (KNN), Decision Tree, Random Forest, dan Support Vector Machine (SVM) dalam empat skenario: *baseline*, *tuning hyperparameter*, *SMOTE*, serta kombinasi *SMOTE + tuning*. Seluruh proses mengikuti tahapan *Knowledge Discovery in Databases* (KDD), mencakup *preprocessing*, transformasi, pemodelan, dan evaluasi performa menggunakan metrik akurasi. Pelabelan sentimen dilakukan menggunakan pendekatan lexicon terhadap data ulasan dari Google Play Store.

Hasil penelitian menunjukkan bahwa algoritma SVM secara konsisten memiliki akurasi tertinggi, dengan performa terbaik pada Traveloka (92%). *Tuning hyperparameter* memberikan dampak positif pada sebagian besar model, sedangkan teknik SMOTE hanya efektif pada algoritma tertentu dan bahkan menurunkan akurasi pada KNN. Traveloka juga memiliki proporsi sentimen positif tertinggi dibandingkan Tiket.com dan Agoda, sehingga mengindikasikan persepsi pengguna yang relatif lebih baik.

Kata kunci: Analisis Sentimen, *Machine Learning*, *Online Travel Agent*

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

SENTIMENT ANALYSIS OF USER SATISFACTION TRAVEL AGENCY APPLICATION USING MACHINE LEARNING ALGORITHM AND HYPERPARAMETER TUNING OPTIMIZATION

Fathurrahman Bagaskara

ABSTRACT

The rapid growth of Online Travel Agent (OTA) applications such as Traveloka, Tiket.com, and Agoda has highlighted the need to accurately understand user sentiment. Comparing five machine learning classification algorithms is essential to determine the most effective model for analyzing public opinion, as each algorithm has different characteristics and sensitivity to textual review data. Sentiment analysis plays a crucial role in evaluating user perception and improving digital service quality.

This study implements five algorithms—Naive Bayes, K-Nearest Neighbor (KNN), Decision Tree, Random Forest, and Support Vector Machine (SVM)—across four scenarios: baseline, hyperparameter tuning, SMOTE, and a combination of SMOTE and tuning. The process follows the Knowledge Discovery in Databases (KDD) framework, including preprocessing, data transformation, modeling, and performance evaluation using accuracy metrics. Sentiment labeling was performed using a lexicon-based approach on user review data collected from the Google Play Store.

The results show that SVM consistently achieved the highest accuracy, with its best performance on Traveloka (92%). Hyperparameter tuning positively impacted most models, while the SMOTE technique was only effective on certain algorithms and significantly reduced accuracy in KNN. Traveloka also recorded the highest proportion of positive sentiment compared to Tiket.com and Agoda, indicating relatively better user perception.

Keywords: Sentiment Analysis, Machine Learning, Online Travel Agent

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	v
KATA PENGANTAR.....	vi
ABSTRAK.....	vii
<i>ABSTRACT</i>	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	xvi
DAFTAR LAMPIRAN	xix
DAFTAR RUMUS	xx
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah.....	5
1.4 Tujuan dan Manfaat Penelitian.....	6
1.4.1 Tujuan Penelitian.....	6
1.4.2 Manfaat Penelitian.....	6
1.5 Sistematika Penulisan.....	7
BAB II LANDASAN TEORI.....	8
2.1 Penelitian Terdahulu	8
2.2 Tinjauan Teori.....	11
2.2.1 Analisis Sentimen.....	11
2.2.2 Teorema Bayes.....	12
2.2.3 Text Mining.....	13
2.2.4 Confusion Matrix	14
2.2.5 Web Scraping.....	16
2.2.6 TF - IDF.....	17
2.2.7 E- Commerce.....	17

2.2.9 SMOTE	18
2.2.10 Hyperparameter Tuning	18
2.3 Framework & Algoritma	19
2.3.1 Knowledge Discovery in Databases (KDD)	19
2.3.2 Naïve Bayes	20
2.3.3 K-Nearest Neighbors (KNN)	22
2.3.4 Decision Tree	23
2.3.5 SVM	23
2.3.6 Random Forest	24
2.4 Tools	25
2.4.1 Python	25
2.4.2 Google Colab	26
BAB III METODE PENELITIAN	27
3.1 Gambaran Umum Objek Penelitian	27
3.2 Metode Penelitian	27
3.2.1 Metode KDD	28
3.2.2 Alur Penelitian	33
3.3 Variabel Penelitian	34
3.3.1 Variabel Independen	34
3.3.2 Variabel Dependen	35
3.4 Teknik Pengumpulan Data	35
3.5 Teknik Analisis Data	37
BAB IV ANALISIS DAN HASIL PENELITIAN	40
4.1 <i>Data Selection</i>	40
4.2 <i>Preprocessing (Prapemrosesan Data)</i>	45
4.2.1 Pemeriksaan <i>Missing Value</i>	45
4.2.3 Pembersihan Simbol dan Elemen	49
4.2.4 <i>Case Folding</i>	50
4.2.5 Normalisasi Teks	52
4.2.6 <i>Stopwords Removal</i>	53
4.2.7 <i>Tokenization</i>	54

4.2.8 <i>Stemming</i>	55
4.2.9 Labeling Sentimen Lexicon (Kamus Bahasa Indonesia)	59
4.2.10 <i>Word Cloud</i>	66
4.3 <i>Data Transformation</i>	70
4.4 <i>Data Mining</i>	73
4.4.1 Hasil Pemodelan Analisis Sentimen menggunakan <i>Machine Learning</i> Konvensional Pada Ulasan Traveloka.....	74
4.4.2 Hasil Pemodelan Analisis Sentimen menggunakan <i>Machine Learning</i> dengan Optimasi Tuning Hyperparameter Pada Ulasan Traveloka	80
4.4.3 Hasil Pemodelan Analisis Sentimen menggunakan <i>Machine Learning</i> dengan Teknik SMOTE Pada Ulasan Traveloka.....	86
4.4.4 Hasil Pemodelan Analisis Sentimen menggunakan <i>Machine Learning</i> dengan Teknik SMOTE dan Tuning Pada Ulasan Traveloka	92
4.4.5 Hasil Pemodelan Analisis Sentimen menggunakan <i>Machine Learning</i> Konvensional Pada Ulasan Tiket.com.....	98
4.4.6 Hasil Pemodelan Analisis Sentimen menggunakan <i>Machine Learning</i> dengan Optimasi Tuning Hyperparameter Pada Ulasan Tiket.com	104
4.4.7 Hasil Pemodelan Analisis Sentimen menggunakan <i>Machine Learning</i> dengan Teknik SMOTE Pada Ulasan Tiket.com	110
4.4.8 Hasil Pemodelan Analisis Sentimen menggunakan <i>Machine Learning</i> dengan Teknik SMOTE dan Tuning Pada Ulasan Tiket.com	116
4.4.9 Hasil Pemodelan Analisis Sentimen menggunakan <i>Machine Learning</i> Konvensional Pada Ulasan Agoda	121
4.4.10 Hasil Pemodelan Analisis Sentimen menggunakan <i>Machine Learning</i> dengan Optimasi Tuning Hyperparameter Pada Ulasan Traveloka.....	127
4.4.11 Hasil Pemodelan Analisis Sentimen menggunakan <i>Machine Learning</i> dengan Teknik SMOTE Pada Ulasan Traveloka.....	133
4.4.12 Hasil Pemodelan Analisis Sentimen menggunakan <i>Machine Learning</i> dengan Teknik SMOTE dan Tuning Pada Ulasan Agoda	138
4.5 <i>Evaluation</i>	144
4.5.1 Evaluasi Perbandingan Akurasi Baseline Model	144

4.5.2 Evaluasi Penerapan Tuning Hyperparameter dan Penanganan Imbalance Data.....	145
4.5.3 Evaluasi Sentimen Pengguna Berdasarkan Aplikasi OTA	148
4.6 Hasil dan Diskusi.....	149
BAB V SIMPULAN DAN SARAN.....	151
5.1 Simpulan.....	151
5.2 Saran	152
DAFTAR PUSTAKA.....	153
LAMPIRAN.....	166



DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu.....	9
Tabel 2. 2 Kelebihan dan Kekurangan Naive Bayes	21
Tabel 2. 3 Kelebihan dan Kekurangan KNN	22
Tabel 2. 4 Kelebihan dan Kekurangan Decision Tree	23
Tabel 2. 5 Kelebihan dan Kekurangan SVM	24
Tabel 2. 6 Kelebihan dan Kekurangan Random Forest	25
Tabel 3. 1 KDD vs CRISP DM	27
Tabel 3. 2 Variabel Dataset	29
Tabel 3. 3 Google Colab vs Selenium.....	35
Tabel 3. 4 Isi Dataset Scraping Traveloka.....	36
Tabel 3. 5 Model algoritma	39
Tabel 4. 1 Kolom yang diambil saat scraping data.....	45
Tabel 4. 2 Sebelum dan Sesudah Case Folding Traveloka	51
Tabel 4. 3 Sebelum dan Sesudah Case Folding Tiket.....	51
Tabel 4. 4 Sebelum dan Sesudah Case Folding Agoda	52
Tabel 4. 5 Normalisasi Teks Traveloka, Tiket.com dan Agoda	52
Tabel 4. 6 Kata kata yang dihapus	53
Tabel 4. 7 Hasil Stemming Traveloka.....	56
Tabel 4. 8 Hasil Stemming Tiket.com.....	57
Tabel 4. 9 Hasil Stemming Agoda	58
Tabel 4. 10 Kalimat Positif dan Negatif untuk Lexicon	59
Tabel 4. 11 Hasil Sentimen Berdasarkan Kata Traveloka	60
Tabel 4. 12 Jumlah sentimen Traveloka.....	61
Tabel 4. 13 Hasil sentimen berdasarkan kata Tiket.com	63
Tabel 4. 14 Jumlah sentimen Tiket.com.....	63
Tabel 4. 15 Hasil Sentimen Berdasarkan Kata Agoda.....	65
Tabel 4. 16 Jumlah Sentimen Agoda	65
Tabel 4. 17 Data Latih dan Uji Ketiga Aplikasi.....	71
Tabel 4. 18 Before After SMOTE Traveloka	72
Tabel 4. 19 Before After SMOTE Tiket.com	72
Tabel 4. 20 Before After SMOTE Agoda.....	73
Tabel 4. 21 Evaluasi Baseline Naive Bayes Traveloka	75
Tabel 4. 22 Evaluasi Baseline Decision Tree Traveloka	76
Tabel 4. 23 Evaluasi Baseline Random Forest Traveloka	77
Tabel 4. 24 Evaluasi Baseline KNN Traveloka	78
Tabel 4. 25 Evaluasi Baseline SVM Traveloka	79
Tabel 4. 26 Evaluasi Baseline Tuning Naive Bayes Traveloka.....	81
Tabel 4. 27 Evaluasi Baseline Tuning Decision Tree Traveloka.....	82
Tabel 4. 28 Evaluasi Baseline Tuning Random Forest Traveloka.....	83
Tabel 4. 29 Evaluasi Baseline Tuning KNN Traveloka.....	85
Tabel 4. 30 Evaluasi Baseline Tuning KNN Traveloka.....	86

Tabel 4. 31 Evaluasi SMOTE Naive Bayes Traveloka.....	88
Tabel 4. 32 Evaluasi SMOTE Decision Tree Traveloka.....	89
Tabel 4. 33 Evaluasi SMOTE Random Forest Traveloka.....	90
Tabel 4. 34 Evaluasi SMOTE KNN Traveloka.....	91
Tabel 4. 35 Evaluasi SMOTE SVM Traveloka.....	92
Tabel 4. 36 Evaluasi SMOTE Tuning Traveloka Naive Bayes	94
Tabel 4. 37 Evaluasi SMOTE Tuning Decision Tree Traveloka	95
Tabel 4. 38 Evaluasi SMOTE Tuning Random Forest Traveloka	96
Tabel 4. 39 Evaluasi SMOTE Tuning KNN Traveloka	97
Tabel 4. 40 Evaluasi SMOTE Tuning SVM Traveloka.....	98
Tabel 4. 41 Evaluasi Baseline Naive Bayes Tiket.com	100
Tabel 4. 42 Evaluasi Baseline Decision Tree Tiket.com	101
Tabel 4. 43Evaluasi Baseline Random Forest Tiket.com	102
Tabel 4. 44 Evaluasi Baseline KNN Tiket.com	103
Tabel 4. 45 Evaluasi Baseline KNN Tiket.com	104
Tabel 4. 46 Evaluasi Baseline Tuning Naive Bayes Tiket.com.....	105
Tabel 4. 47 Evaluasi Baseline Tuning Decision Tree Tiket.com	107
Tabel 4. 48 Evaluasi Baseline Tuning Random Forest Tiket.com.....	108
Tabel 4. 49 Evaluasi Baseline Tuning Tiket.com	109
Tabel 4. 50 Evaluasi Baseline Tuning SVM Tiket.com	110
Tabel 4. 51 Evaluasi SMOTE Naive Bayes Tiket.com	111
Tabel 4. 52 Evaluasi SMOTE Decision Tree Tiket.com	113
Tabel 4. 53 Evaluasi SMOTE Random Forest Tiket.com.....	114
Tabel 4. 54 Evaluasi SMOTE KNN Tiket.com	115
Tabel 4. 55 Evaluasi SMOTE SVM Tiket.com	116
Tabel 4. 56 Evaluasi SMOTE Tuning Naive Bayes Tiket.com.....	117
Tabel 4. 57 Evaluasi SMOTE Tuning Decision Tree Tiket.com.....	118
Tabel 4. 58 Evaluasi SMOTE Tuning Random Forest Tiket.com.....	119
Tabel 4. 59 Evaluasi SMOTE Tuning KNN Tiket.com.....	120
Tabel 4. 60 Evaluasi SMOTE Tuning SVM Tiket.com.....	121
Tabel 4. 61 Evaluasi Baseline Naive Bayes Agoda.....	123
Tabel 4. 62 Evaluasi Baseline Naive Bayes Agoda.....	124
Tabel 4. 63 Evaluasi Baseline Random Forest Agoda.....	125
Tabel 4. 64 Evaluasi Baseline KNN Agoda.....	126
Tabel 4. 65Evaluasi Baseline SVM Agoda	127
Tabel 4. 66 Evaluasi Baseline Tuning Naive Bayes Agoda.....	128
Tabel 4. 67Evaluasi Baseline Tuning Decision Tree Agoda.....	129
Tabel 4. 68 Evaluasi Baseline Tuning Random Forest Agoda.....	131
Tabel 4. 69 Evaluasi Baseline Tuning KNN Agoda.....	132
Tabel 4. 70 Evaluasi Baseline Tuning SVM Agoda.....	133
Tabel 4. 71 Evaluasi SMOTE Naive Bayes Agoda.....	134
Tabel 4. 72 Evaluasi SMOTE Decision Tree Agoda.....	135
Tabel 4. 73 Evaluasi SMOTE Random Forest Agoda.....	136

Tabel 4. 74 Evaluasi SMOTE KNN Agoda.....	137
Tabel 4. 75 Evaluasi SMOTE SVM Agoda.....	138
Tabel 4. 76 Evaluasi SMOTE Tuning Naive Bayes Agoda.....	140
Tabel 4. 77 Evaluasi SMOTE Tuning Decision Tree Agoda.....	141
Tabel 4. 78 Evaluasi SMOTE Tuning Random Forest Agoda.....	142
Tabel 4. 79 Evaluasi SMOTE Tuning KNN Agoda.....	143
Tabel 4. 80 Evaluasi SMOTE Tuning SVM Agoda.....	144
Tabel 4. 81 Akurasi Baseline Ketiga Aplikasi	145
Tabel 4. 82 Evaluasi Penerapan Tuning dan SMOTE (Akurasi).....	146
Tabel 4. 83 Distribusi Sentimen Pengguna Berdasarkan Aplikasi OTA.....	148
Tabel 4. 84 Perbandingan Hasil Penelitian dengan Penelitian Terdahulu	149



DAFTAR GAMBAR

Gambar 1. 1 Survey Top Situs Perjalanan di Indonesia[9].	2
Gambar 1. 2 Top Apps Ranking[10].	2
Gambar 2. 1 Metodologi PRISMA	8
Gambar 2. 2 Confusion Matrix[42].	14
Gambar 2. 3 Web Scraping	16
Gambar 2. 4 Metode KDD[61].	19
Gambar 3. 1 Framework KDD[91]	28
Gambar 3. 2 Flowchart Tahap <i>Preprocessing</i>	30
Gambar 3. 3 Flowchart Tahap <i>Transformation</i> .	30
Gambar 3. 4 Flowchart Tahap <i>Data Mining</i>	31
Gambar 3. 5 Flowchart Tahap <i>Evaluation</i>	32
Gambar 3. 6 Alur Penelitian	34
Gambar 3. 7 kolom yang diambil untuk dataset.	36
Gambar 4. 1 install <i>google-play-scraper</i>	40
Gambar 4. 2 import library pandas, numpy dan google play scraper	40
Gambar 4. 3 Input url link Traveloka	41
Gambar 4. 4 Tabel Ulasan Pengguna Traveloka	42
Gambar 4. 5 Jumlah data yang didapat	42
Gambar 4. 6 Pratinjau Data Ulasan Traveloka (User, Rating, Tanggal, dan Ulasan)	43
Gambar 4. 7 Data Ulasan Traveloka yang Diurutkan Berdasarkan Tanggal	43
Gambar 4. 8 Pemilihan Kolom Data Ulasan Traveloka	43
Gambar 4. 9 Input URL Link Tiket.com	44
Gambar 4. 10 Input URL Link Agoda	44
Gambar 4. 11 Check Missing Value Traveloka	46
Gambar 4. 12 Check Missing Value Tiket.com	46
Gambar 4. 13 Check Missing Value Agoda	47
Gambar 4. 14 Pengecekan Duplikat Traveloka	48
Gambar 4. 15 Pengecekan Duplikat Tiket.com	48
Gambar 4. 16 Pengecekan Duplikat Agoda	48
Gambar 4. 17Pembersihan Simbol dan Elemen Dataset Traveloka	49
Gambar 4. 18 Pembersihan Simbol dan Elemen Dataset Tiket.com	49
Gambar 4. 19 Pembersihan Simbol dan Elemen Dataset Agoda	50
Gambar 4. 20 Case Folding Traveloka	50
Gambar 4. 21 Case Folding Tiket.com	51
Gambar 4. 22 Case Folding Agoda	52
Gambar 4. 23 Coding tokenized traveloka	54
Gambar 4. 24 Coding Tokenized Tiket.com	55
Gambar 4. 25 Coding Tokenized Agoda	55
Gambar 4. 26 Coding Stemming Traveloka	56
Gambar 4. 27 Coding Stemming Tiket.com	57

Gambar 4. 28 Coding Stemming Agoda	58
Gambar 4. 29 Bar Chart Traveloka	61
Gambar 4. 30 Pie Chart Traveloka	62
Gambar 4. 31 Bar Chart Tiket.com	64
Gambar 4. 32 Bar Chart Tiket.com	64
Gambar 4. 33 Wordcloud positif Traveloka	66
Gambar 4. 34 Wordcloud Negatif Traveloka	67
Gambar 4. 35 Wordcloud positif Tiket.com	68
Gambar 4. 36 Wordcloud negatif Tiket.com	68
Gambar 4. 37 Wordcloud positif Agoda	69
Gambar 4. 38 Wordcloud negatif Agoda	69
Gambar 4. 39 Splitting Data 70 30	70
Gambar 4. 40 Declare SMOTE diawal	71
Gambar 4. 41 Naive Bayes Baseline Traveloka	75
Gambar 4. 42 Decision Tree Baseline Traveloka	76
Gambar 4. 43 Random Forest Baseline Traveloka	77
Gambar 4. 44 KNN Baseline Traveloka	78
Gambar 4. 45 SVM Baseline Traveloka	79
Gambar 4. 46 Naive Bayes Baseline Tuning Traveloka	81
Gambar 4. 47 Decision Tree Baseline Tuning Traveloka	82
Gambar 4. 48 Random Forest Baseline Tuning Traveloka	83
Gambar 4. 49 KNN Baseline Tuning Traveloka	84
Gambar 4. 50 SVM Baseline Tuning Traveloka	86
Gambar 4. 51 Naive Bayes SMOTE Traveloka	87
Gambar 4. 52 Decision Tree SMOTE Traveloka	88
Gambar 4. 53 Random Forest SMOTE Traveloka	90
Gambar 4. 54 KNN SMOTE Traveloka	90
Gambar 4. 55 SVM SMOTE Traveloka	91
Gambar 4. 56 Naive Bayes SMOTE Tuning Traveloka	93
Gambar 4. 57 Decision Tree SMOTE Tuning Traveloka	94
Gambar 4. 58 Random Forest SMOTE Tuning Traveloka	95
Gambar 4. 59 KNN SMOTE Tuning Traveloka	97
Gambar 4. 60 SVM SMOTE Tuning Traveloka	98
Gambar 4. 61 Naive Bayes Baseline Tiket.com	99
Gambar 4. 62 Decision Tree Baseline Tiket.com	101
Gambar 4. 63 Random Forest Baseline Tiket.com	102
Gambar 4. 64 KNN Baseline Tiket.com	103
Gambar 4. 65 SVM Baseline Tiket.com	104
Gambar 4. 66 Naive Bayes Baseline Tuning Tiket.com	105
Gambar 4. 67 Decision Tree Baseline Tuning Tiket.com	106
Gambar 4. 68 Random Forest Baseline Tuning Tiket.com	108
Gambar 4. 69 KNN Baseline Tuning Tiket.com	109
Gambar 4. 70 SVM Baseline Tuning Tiket.com	110

Gambar 4. 71 Naive Bayes SMOTE Tiket.com.....	112
Gambar 4. 72 Decision Tree SMOTE Tiket.com.....	113
Gambar 4. 73 Random Forest SMOTE Tiket.com.....	114
Gambar 4. 74 KNN SMOTE Tiket.com.....	115
Gambar 4. 75 SVM SMOTE Tiket.com.....	115
Gambar 4. 76 Naive Bayes SMOTE Tuning Tiket.com.....	117
Gambar 4. 77 Decision Tree SMOTE Tuning Tiket.com.....	118
Gambar 4. 78 Random Forest SMOTE Tuning Tiket.com.....	119
Gambar 4. 79 KNN SMOTE Tuning Tiket.com.....	120
Gambar 4. 80 SVM SMOTE Tuning Tiket.com.....	121
Gambar 4. 81 Naive Bayes Baseline Agoda.....	122
Gambar 4. 82 Decision Tree Baseline Agoda.....	123
Gambar 4. 83 Random Forest Baseline Agoda.....	124
Gambar 4. 84 KNN Baseline Agoda.....	125
Gambar 4. 85 SVM Baseline Agoda.....	126
Gambar 4. 86 Naive Bayes Baseline Tuning Agoda.....	128
Gambar 4. 87 Decision Tree Baseline Tuning Agoda.....	129
Gambar 4. 88 Random Forest Baseline Tuning Agoda.....	130
Gambar 4. 89 KNN Baseline Tuning Agoda.....	131
Gambar 4. 90 SVM Baseline Tuning Agoda.....	132
Gambar 4. 91 Naive Bayes SMOTE Agoda.....	134
Gambar 4. 92 Decision Tree SMOTE Agoda.....	135
Gambar 4. 93 Random Forest SMOTE Agoda.....	136
Gambar 4. 94 KNN SMOTE Agoda.....	137
Gambar 4. 95 SVM SMOTE Agoda.....	138
Gambar 4. 96 Naive Bayes SMOTE Tuning Agoda.....	139
Gambar 4. 97 Decision Tree SMOTE Tuning Agoda.....	140
Gambar 4. 98 Random Forest SMOTE Tuning Agoda.....	141
Gambar 4. 99 KNN SMOTE Tuning Agoda.....	142
Gambar 4. 100 SVM SMOTE Tuning Agoda.....	144

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR LAMPIRAN

Lampiran 1 Turnitin.....	166
Lampiran 2 Turnitin.....	167
Lampiran 3 Turnitin.....	168
Lampiran 4 Turnitin.....	169
Lampiran 5 Form Bimbingan Skripsi.....	170
Lampiran 6 Dataset Traveloka	171
Lampiran 7 Dataset Agoda.....	174
Lampiran 8 Dataset Tiket.com	178



DAFTAR RUMUS

(2.1) Rumus Teorema Bayes	12
(2.2) Rumus Accuracy	15
(2.3) Rumus Precision	15
(2.4) Rumus Recall	15
(2.5) Rumus F1-Score	16
(2.6) Rumus TF	17
(2.7) Rumus IDF	17
(2.8) Rumus Teorema Bayes pada Naïve Bayes	21

UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA