

**OPTIMASI ANALISIS SENTIMEN DAN REKOMENDASI
REVIEW TAMAN SAFARI DENGAN ALGORITMA
MACHINE LEARNING DAN TERM FREQUENCY**



SKRIPSI

Bramantyo Wishnu Mukti

00000061245

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**OPTIMASI ANALISIS SENTIMEN DAN REKOMENDASI
REVIEW TAMAN SAFARI DENGAN ALGORITMA
MACHINE LEARNING DAN TERM FREQUENCY**



Diajukan sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Komputer

Bramantyo Wishnu Mukti

00000061245

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

2025

i

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Bramantyo Wishnu Mukti

Nomor Induk Mahasiswa : 00000061245

Program Studi : Sistem Informasi

Skripsi dengan judul:

**OPTIMASI ANALISIS SENTIMEN DAN REKOMENDASI REVIEW
TAMAN SAFARI DENGAN ALGORITMA MACHINE LEARNING DAN
TERM FREQUENCY**

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, 25 Juni 2025



Bramantyo Wishnu Mukti

HALAMAN PENGESAHAN

Skripsi dengan judul

OPTIMASI ANALISIS SENTIMEN DAN REKOMENDASI REVIEW TAMAN SAFARI DENGAN ALGORITMA MACHINE LEARNING DAN TERM FREQUENCY

Oleh

Nama : Bramantyo Wishnu Mukti
NIM : 00000061245
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika

Telah diujikan pada hari Selasa, 10 Juni 2025

Pukul 13.00 s.d 15.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut.

Ketua Sidang

Penguji

Ir. Raymond Sunardi Octama, M.CIS
0328046803

Pembimbing

Suryasari, S.Kom., M.T.
0323088301

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 : Bramantyo Wishnu Mukti
NIM : 00000061245
Program Studi : Sistem Informasi
Jenjang : S1
Judul Karya Ilmiah : OPTIMASI ANALISIS SENTIMEN DAN
REKOMENDASI REVIEW TAMAN
SAFARI DENGAN ALGORITMA
MACHINE LEARNING DAN TERM
FREQUENCY

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 Juni 2025



Bramantyo Wishnu Mukti

KATA PENGANTAR

Puji dan syukur saya panjatkan kepada Tuhan Yang Maha Esa atas rahmat dan karunia-Nya, sehingga saya dapat menyelesaikan skripsi yang berjudul “Optimasi Analisis Sentimen dan Rekomendasi Review Taman Safari dengan Algoritma Machine Learning dan Term Frequency.” Skripsi ini dibuat untuk memenuhi tugas akhir dan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana Strata 1 di Program Studi Sistem Informasi pada Fakultas Teknik dan Informatika, Universitas Multimedia Nusantara. Saya menyadari bahwa tanpa bimbingan, arahan, dan bantuan dari berbagai pihak, penulisan skripsi ini akan sangat sulit untuk diselesaikan. Oleh karena itu, pada kesempatan ini saya ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Dr. Andrey Andoko, M.Sc., Ph.D. selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, ST, M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. 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.

Semoga skripsi ini dapat memberikan kontribusi positif dalam pengembangan analisis sentimen dan rekomendasi untuk meningkatkan pengalaman pengunjung di Taman Safari serta menjadi referensi bagi penelitian selanjutnya.

Tangerang, 25 Juni 2025



Bramantyo Wishnu Mukti

OPTIMASI ANALISIS SENTIMEN DAN REKOMENDASI REVIEW TAMAN SAFARI DENGAN ALGORITMA MACHINE LEARNING DAN TERM FREQUENCY

ABSTRAK

Latar belakang penelitian ini berfokus pada semakin tingginya minat masyarakat terhadap destinasi wisata serta pentingnya memahami persepsi pengunjung dalam meningkatkan pengalaman mereka di Taman Safari Indonesia. Seiring dengan meningkatnya penggunaan platform digital seperti *Google Maps*, ulasan pengunjung menjadi sumber data yang berharga untuk analisis sentimen. Namun, banyaknya data yang terdapat dalam ulasan ini menjadi sulit untuk diproses dan dianalisis secara efektif, sehingga diperlukan pendekatan sistematis untuk mengidentifikasi sentimen dan faktor.

Metode yang diterapkan dalam penelitian ini menggunakan pendekatan *Knowledge Discovery in Databases* (KDD). Proses ini mencakup pengumpulan, pemrosesan, dan analisis data menggunakan teknologi *machine learning*. Untuk mengategorikan ulasan menjadi tiga kelas positif, negatif, dan netral dan lima algoritma *machine learning*, yaitu *Logistic Regression*, *Naïve Bayes*, *Support Vector Machine* (SVM), *Random Forest*, dan *K-Nearest Neighbors* (KNN), diterapkan. Penelitian ini bertujuan untuk menentukan algoritma mana yang memiliki kinerja terbaik dalam analisis sentimen serta menganalisis tema utama yang muncul dari ulasan pengunjung.

Hasil penelitian menunjukkan bahwa algoritma SVM memberikan kinerja terbaik dengan akurasi mencapai 94% dan *F1-score* sebesar 0.97, yang menunjukkan efektivitasnya dalam menangani data dengan kompleksitas tinggi. Analisis menunjukkan bahwa mayoritas ulasan bersifat positif dan menyoroti kepuasan pengunjung terhadap kualitas atraksi, pelayanan, interaksi dengan hewan, dan kebersihan fasilitas. Kesimpulan dari penelitian ini menegaskan pentingnya menerapkan *Term Frequency* dalam pengembangan rekomendasi untuk memberikan informasi yang relevan kepada pengunjung dalam memilih destinasi wisata sesuai minat mereka.

Kata kunci: Analisis Sentimen, *Term Frequency*, *Google Maps*, *Machine Learning*, Taman Safari Indonesia

OPTIMIZATION OF SENTIMENT ANALYSIS AND REVIEW RECOMMENDATIONS FOR TAMAN SAFARI USING MACHINE LEARNING ALGORITHMS AND TERM FREQUENCY

ABSTRACT

This research focuses on the increasing public interest in tourist destinations and the importance of understanding visitor perceptions to enhance their experience at Taman Safari Indonesia. With the growing use of digital platforms such as Google Maps, visitor reviews have become a valuable source of data for sentiment analysis. However, the vast amount of data present in these reviews often makes it challenging to process and analyze effectively, necessitating a systematic approach to identify sentiments and influencing factors.

The methodology applied in this research utilizes the Knowledge Discovery in Databases (KDD) approach. This process encompasses data collection, processing, and analysis using machine learning technologies. To categorize the reviews into three classes—positive, negative, and neutral—five machine learning algorithms, namely Logistic Regression, Naïve Bayes, Support Vector Machine (SVM), Random Forest, and K-Nearest Neighbors (KNN), are implemented. This study aims to determine which algorithm performs best in sentiment analysis and to analyze the main themes emerging from visitor reviews.

The results indicate that the SVM algorithm exhibits the best performance, achieving an accuracy of 94% and an F1-score of 0.97, demonstrating its effectiveness in handling highly complex data. The analysis reveals that the majority of reviews are positive, highlighting visitor satisfaction regarding the quality of attractions, service, interaction with animals, and cleanliness of facilities. The conclusion of this research emphasizes the importance of employing Term Frequency method in the development of recommendation to provide relevant information to visitors when choosing tourist destinations that align with their interests.

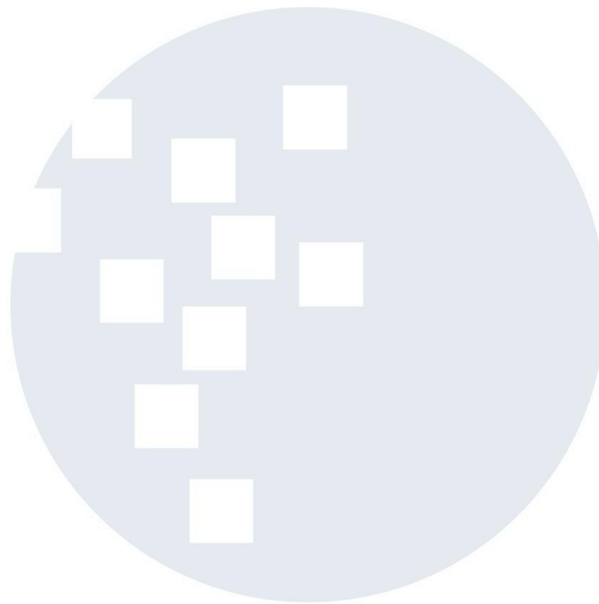
Keywords: *Term Frequency, Google Maps, Machine Learning, Sentiment Analysis, Taman Safari Indonesia.*

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
KATA PENGANTAR.....	v
ABSTRAK	vi
<i>ABSTRACT</i>	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	xi
DAFTAR RUMUS	xii
DAFTAR GAMBAR.....	xiii
DAFTAR	xv
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	10
2.1 Penelitian Terdahulu	10
2.2 Teori Penelitian	17
2.2.1 Analisis Sentimen	17
2.2.2 Taman Safari Indonesia	18
2.2.3 Ulasan Google Maps	18
2.2.4 Term Frequency	19
2.2.5 Penggunaan Data Ulasan untuk Rekomendasi dan Analisis Sentimen	19
2.2.6 Tujuan dan Manfaat Analisis Sentimen	21
2.3 Framework dan Algoritma Penelitian	21
2.4 Tools dan Software Penelitian	31

2.4.1 Jupyter Notebook.....	31
2.4.2 Apify.....	32
2.4.3 CSV Editor	32
2.4.4 Google Maps Extractor	33
BAB III METODOLOGI PENELITIAN	34
3.1 Gambaran Umum Objek Penelitian.....	34
3.2 Metode Penelitian.....	35
3.3 Teknik Pengumpulan Data	43
3.4 Teknik Analisis Data.....	48
BAB IV ANALISIS DAN HASIL PENELITIAN	50
4.1 Selection	50
4.2 Preprocessing.....	54
4.2.1 Missing Value	54
4.2.2 Memeriksa Duplikat	57
4.2.3 Tokenisasi	57
4.2.6 Analisis Sentimen dengan TextBlob	64
4.3 Data Transformation	67
4.4 Data Mining.....	68
4.4.1 <i>Support Vector Machine</i> (SVM)	69
4.4.2 <i>Logistic Regression</i>	72
4.4.3 Naïve Bayes.....	75
4.4.4 <i>Random Forest</i>	78
4.4.5 <i>K-Nearest Neighbors</i> (KNN)	81
4.4.6 Hasil Evaluasi	83
4.7 Term Frequency	88
4.8 Hasil Penelitian.....	94
4.8.1 Proses KDD.....	94
4.8.2 Analisis Sentimen	95
4.8.3 Evaluasi Model	95
4.8.4 Rekomendasi Berdasarkan Term Frequency.....	95
4.9.1 Dukungan Terhadap Penelitian Terdahulu	98
BAB V SIMPULAN DAN SARAN	99

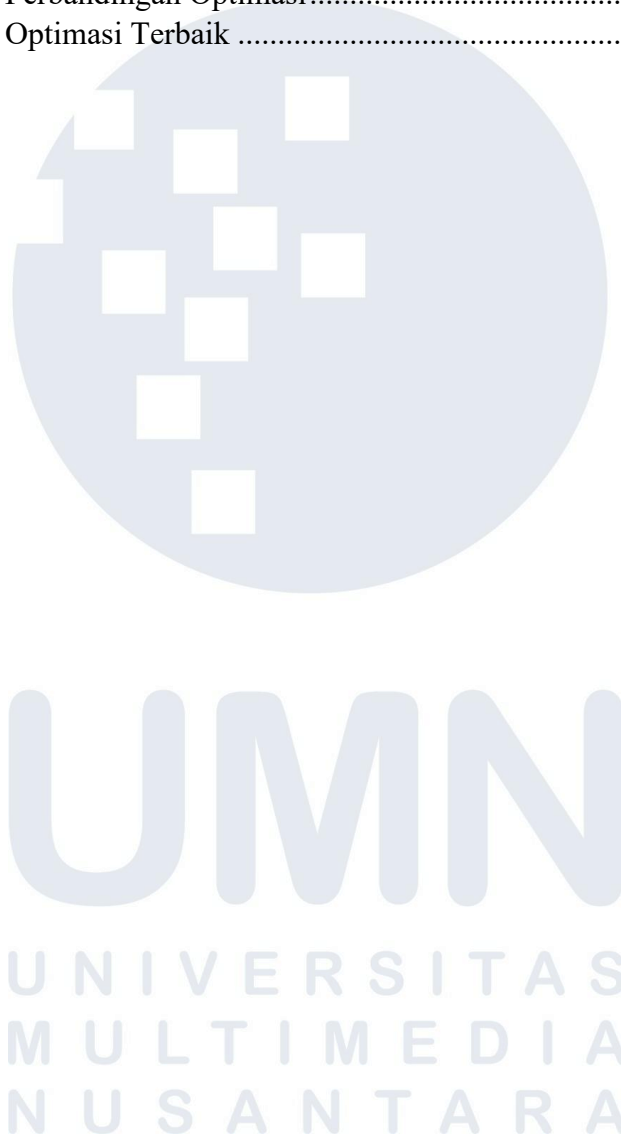
5.1 Simpulan.....	99
5.2 Saran.....	101
DAFTAR PUSTAKA.....	102
LAMPIRAN.....	108



UMN
 UNIVERSITAS
 MULTIMEDIA
 NUSANTARA

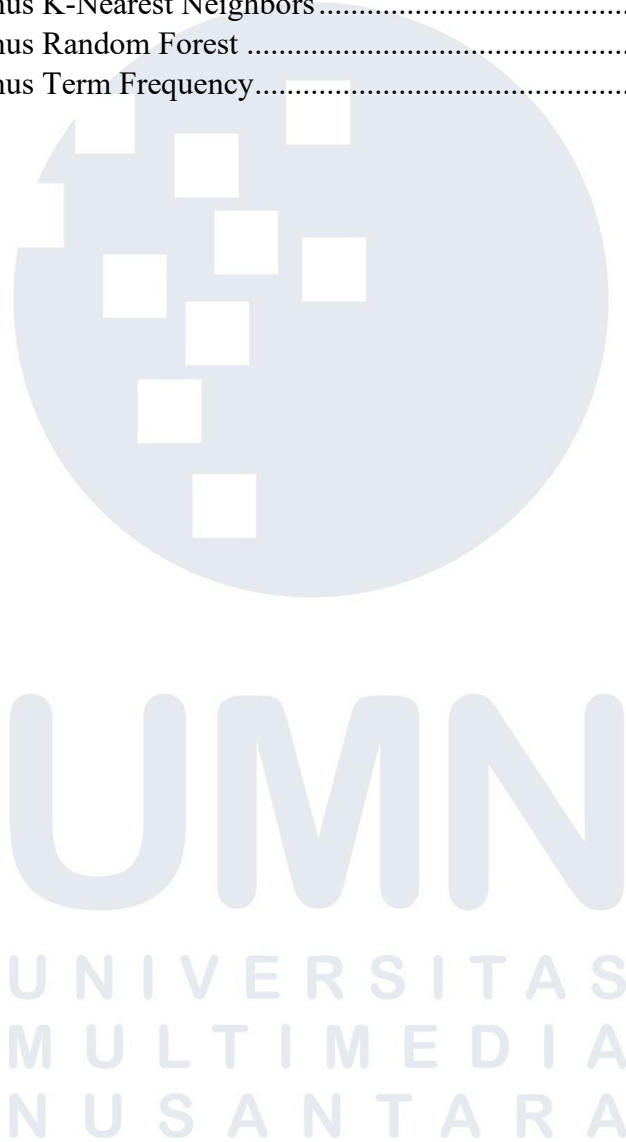
DAFTAR TABEL

Tabel 2. 1 Tabel Penelitian Terdahulu	10
Tabel 3. 1 Tabel Metode Penelitian	35
Tabel 3. 2 Tabel Teknik Pengumpulan Data.....	43
Tabel 4. 1 Tabel Stopwords	59
Tabel 4. 2 Tabel Perbandingan Algoritma	84
Tabel 4. 3 Tabel Perbandingan Optimasi.....	85
Tabel 4. 4 Tabel Optimasi Terbaik	86



DAFTAR RUMUS

Rumus 2. 1 Rumus Support Vector Machine.....	24
Rumus 2. 2 Rumus Fungsi Tujuan untuk Support Vector Machine	25
Rumus 2. 3 Rumus dengan Kendala Support Vector Machine.....	25
Rumus 2. 4 Rumus Logistic Regression	26
Rumus 2. 5 Rumus Naive Bayes.....	27
Rumus 2. 6 Rumus K-Nearest Neighbors.....	28
Rumus 2. 7 Rumus Random Forest	29
Rumus 2. 8 Rumus Term Frequency.....	30



DAFTAR GAMBAR

Gambar 3. 1 Metode Knowledge Discovery in Databases (KDD)	38
Gambar 3. 2 Tools Apify	45
Gambar 3. 3 Insert Data di Apify	46
Gambar 3. 4 Tools Google Maps Extractor	47
Gambar 3. 5 Input Data Google Maps Extractor	48
Gambar 4. 1 Tampilan CSV Editor	50
Gambar 4. 2 Hasil Setelah Penghapusan Kolom	51
Gambar 4. 3 DataFrame Sebelum Proses Filter	52
Gambar 4. 4 Kode untuk Penerapan Filter	53
Gambar 4. 5 DataFrame Setelah Proses Filter	53
Gambar 4. 6 Kode untuk Mencari Missing Value	54
Gambar 4. 7 Output dari Kode Missing Value	54
Gambar 4. 8 Kode untuk Menghapus Missing Value	55
Gambar 4. 9 Data yang Dihapus pada Penghapusan Missing Value	55
Gambar 4. 10 Total Data setelah Penghapusan Missing Value	56
Gambar 4. 11 Kode untuk Memeriksa Duplikat	57
Gambar 4. 12 Kode untuk Tokenisasi	57
Gambar 4. 13 Output setelah Proses Tokenisasi	58
Gambar 4. 14 Tampilan Data setelah Proses Tokenisasi	58
Gambar 4. 15 Kode untuk Penerapan Stopwords	60
Gambar 4. 16 Output dari Kode Stopwords	61
Gambar 4. 17 Hasil DataFrame setelah Output Kode Stopwords	62
Gambar 4. 18 Kode untuk Stemming	62
Gambar 4. 19 Hasil Output dari Kode Stemming	63
Gambar 4. 20 Tampilan Data setelah Penerapan Stemming	63
Gambar 4. 21 Kode untuk Penerapan Wordcloud	64
Gambar 4. 22 Output Hasil Wordcloud	64
Gambar 4. 23 Kode untuk Visualisasi Distribusi Sentimen	65
Gambar 4. 24 Output Visualisasi Distribusi	66
Gambar 4. 25 Output setelah Penerapan Kode Analisis Sentimen	67
Gambar 4. 26 Kode untuk Data Transformation	67
Gambar 4. 27 Kode untuk Training dan Test Data	68
Gambar 4. 28 Penerapan Kode Support Vector Machine	69
Gambar 4. 29 Penerapan Kode untuk Optimasi SMOTE pada Support Vector Machine	70
Gambar 4. 30 Penerapan Kode untuk Optimasi GridsearchCV pada Support Vector Machine	71
Gambar 4. 31 Penerapan Kode Logistic Regression	72
Gambar 4. 32 Penerapan Kode untuk Optimasi SMOTE pada Logistic Regression	73

Gambar 4. 33 Penerapan Kode untuk Optimasi GridsearchCV pada Logistic Regression	74
Gambar 4. 34 Penerapan Kode Naive Bayes	75
Gambar 4. 35 Penerapan Kode untuk Optimasi SMOTE pada Naïve Bayes	76
Gambar 4. 36 Penerapan Kode untuk Optimasi GridsearchCV pada Naïve Bayes	77
Gambar 4. 37 Penerapan Kode Random Forest	78
Gambar 4. 38 Penerapan Kode untuk Optimasi SMOTE pada Random Forest ...	79
Gambar 4. 39 Penerapan Kode untuk Optimasi GridsearchCV pada Random Forest.....	80
Gambar 4. 40 Penerapan Kode K-Nearest Neighbors	81
Gambar 4. 41 Penerapan Kode untuk Optimasi SMOTE pada K-Nearest Neighbors	82
Gambar 4. 42 Penerapan Kode untuk Optimasi GridsearchCV pada K-Nearest Neighbors	83
Gambar 4. 43 Kode Transform TF-IDF pada Dataset	89
Gambar 4. 44 Kode untuk Implementasi Penerapan Term Frequency	90
Gambar 4. 45 Daftar Hewan pada Kode Penerapan Term Frequency	91
Gambar 4. 46 Output TF-IDF Kode Penerapan Term Frequency	92
Gambar 4. 47 Visualisasi Hasil Term Frequency	93
Gambar 4. 48 Peta Taman Safari Indonesia.....	96



DAFTAR LAMPIRAN

Lampiran A Turnitin Similarity Report	108
Lampiran B Form Konsultasi Bimbingan.....	111

