

SKRIPSI

**ANALISIS SENTIMEN PENGGUNA APLIKASI TRAVELOKA MENGGUNAKAN
METODE *MACHINE LEARNING***



Disusun oleh

RAYKA AGUSTASYA RIPHA

20180140017

PROGRAM STUDI TEKNOLOGI INFORMASI

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH YOGYAKARTA

2022

SKRIPSI

**ANALISIS SENTIMEN PENGGUNA APLIKASI TRAVELOKA MENGGUNAKAN
METODE *MACHINE LEARNING***

Disusun Untuk Memenuhi Sebagian Syarat Memperoleh
Derajat Sarjana Teknologi Informasi pada Fakultas Teknik
Universitas Muhammadiyah Yogyakarta



Disusun oleh

RAYKA AGUSTASYA RIPHA

20180140017

PROGRAM STUDI TEKNOLOGI INFORMASI

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH YOGYAKARTA

2022

HALAMAN PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama : Rayka Agustasya Ripha

NIM : 20180140017

Program Studi : Teknologi Informasi

Fakultas : Fakultas Teknik

Jenis Karya : Skripsi

Judul Karya : Analisis Sentimen Pengguna Aplikasi Traveloka Menggunakan Metode Machine Learning

Menyatakan dengan benar dan tanpa paksaan bahwa :

1. Karya ini adalah karya sendiri dengan arahan dan bimbingan dosen pembimbing dan merupakan sebagian hasil dari penelitian di Universitas Muhammadiyah Yogyakarta dengan :

Judul Penelitian : Analisis Sentimen Pengguna Aplikasi Traveloka Menggunakan Metode Machine Learning

Ketua Penelitian : Slamet Riyadi, S.T., M.Sc., Ph.D.

2. Karya ini tidak memuat hasil karya orang lain kecuali acuan dan kutipan yang telah disebutkan sumbernya.
3. Karya ini belum pernah diajukan untuk memperoleh gelar akademik (sarjana, magister, dan doctor) di Universitas Muhammadiyah Yogyakarta atau institusi lainnya.
4. Demi pengembangan ilmu pengetahuan, saya menyetujui memberikan hak kepada dosen pembimbing dan Universitas Muhammadiyah Yogyakarta untuk menyimpan, menggunakan, dan mengelola karya ini dan perangkat lainnya (jika ada) serta mempublikasikannya dalam bentuk lain baik itu semua maupun sebagian dengan tetap mencantumkan nama saya.

Yogyakarta, 12 Oktober 2022

Yang membuat pernyataan



10000
METERAI
TEMPER
3218EAKX0580508
Rayka Agustasya Ripha

KATA PENGANTAR

Assalamualaikum Warahmatullahi Wabarakatuh

Alhamdulillah Robbil Alamin, puji dan syukur karena dengan segala rahmat dan karunia-Nya lah penulis dapat menyelesaikan penyusunan skripsi yang berjudul "Analisis Sentimen Pengguna Aplikasi Traveloka Menggunakan Metode *Machine Learning*". Tidak lupa shalawat serta salam semoga senantiasa terlimpah curahkan kepada Nabi Muhammad SAW, kepada keluarganya, para sahabatnya, hingga pada umatnya hingga akhir zaman, amin.

Dalam penyusunan dan penulisan skripsi ini tidak terlepas dari banyaknya bantuan dan bimbingan serta dukungan dari berbagai pihak. Oleh karena itu dalam kesempatan ini penulis dengan senang hati menyampaikan terima kasih kepada :

1. Allah *Subhanahu Wa Ta'ala* yang telah memberikan petunjuk, kemudahan, keistiqomahan dan segala nikmat-Nya.
2. Ir. Asroni, S.T., M.Eng. selaku Ketua Program Studi Teknologi Informasi Universitas Muhammadiyah Yogyakarta.
3. Bapak Slamet Riyadi, S.T., M.Sc., Ph.D. selaku dosen pembimbing 1 yang telah meluangkan waktu untuk memberikan arahan kepada penulis sehingga laporan skripsi dapat terselesaikan dengan baik.
4. Bapak Cahya Damarjati, S.T. M. Eng., Ph.D. selaku dosen pembimbing 2 yang meluangkan waktu untuk memberikan arahan kepada penulis serta memberikan solusi kepada penulis.
5. Seluruh Dosen beserta Staf Program Studi Teknologi Informasi Universitas Muhammadiyah Yogyakarta yang telah memberikan ilmu dan wawasan terhadap penulis selama melakukan proses studi.

6. Seluruh anggota keluarga penulis terutama Bapak Riptono, Ibu Hapipah, Mba Wulan dan Kaka Pebi yang telah memberikan support penuh dalam melaksanakan tugas akhir.
7. Eka Agustina dan Fadia Rani yang sudah mendengarkan keluh kesah dan memberikan support dalam segala hal.
8. Teman-teman Angkatan 18 dan terutama untuk ghian, Elvira, isty, andang, hirdan, ari, akbar, hafiz, nanda dan alm hanif serta yang tidak bisa disebutkan serta teman satu bimbingan yang sudah memberikan support kepada penulis.
9. Dan pihak-pihak lainnya yang disini tidak dapat saya sebutkan satu per satu, yang telah memberikan bantuan berupa masukan, arahan, nasehat, dan lain sebagainya kepada penulis, sehingga penulis dapat mengerjakan penyusunan skripsi ini hingga selesai. Terima kasih untuk semua dukungan dan doa yang telah diberikan, semoga Allah membalasnya dengan yang lebih baik.

Yogyakarta, 12 Oktober 2022

Penulis



Rayka Agustasya Ripha

DAFTAR ISI

HALAMAN PERSETUJUAN	i
HALAMAN PERSETUJUAN	ii
HALAMAN PERNYATAAN.....	iii
KATA PENGANTAR.....	iv
DAFTAR ISI.....	vi
DAFTAR GAMBAR.....	viii
DAFTAR TABEL	ix
INTISARI	x
ABSTRACT	xi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	3
1.5 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	5
2.1 Tinjauan Pustaka	5
2.2 Landasan Teori	8
2.2.1 Analisis Sentimen	8
2.2.2 Traveloka	9
2.2.3 Dataset.....	9
2.2.4 Text Mining	10
2.2.5 Python	10
2.2.6 Google Colaboratory.....	10
2.2.7 RapidMiner	14
2.2.8 Naïve Bayes.....	15
2.2.9 Support Vector Machine	16
2.2.10 K-Nearest Neighbor	18
2.2.11 Cross Validation.....	18
BAB III METODOLOGI PENELITIAN	21
3.1 Alat	21
3.2 Tahapan Penelitian	21
3.3 Studi Literatur	22

3.4	Pengumpulan Data.....	22
3.4.1	Pembersihan Data.....	25
3.4.2	Labelling Data	27
3.5	Pre-Processing Data.....	29
3.6	Implementasi Algoritma.....	29
BAB IV HASIL DAN PEMBAHASAN		31
4.1	Pre-processing Data	31
4.1.1	Case Folding	31
4.1.2	Tokenization	32
4.1.3	Stopword Removal.....	33
4.1.4	Stemming	34
4.2	Implementasi Algoritma.....	34
4.2.1	Pemodelan Text Mining.....	35
4.2.2	Klasifikasi Data	36
4.2.3	Hasil Pengujian	40
BAB V KESIMPULAN DAN SARAN		48
5.1	Kesimpulan.....	48
5.2	Saran	48
DAFTAR PUSTAKA.....		50
LAMPIRAN.....		53

DAFTAR GAMBAR

Gambar 2.1 Python.....	10
Gambar 2.2 Koneksi Google Colab dengan Google Drive.....	11
Gambar 2.3 Membuat Google Colab dari Google Drive.....	12
Gambar 2.4 Tampilan awal Notebook Google Colab.....	12
Gambar 2.5 Tampilan Lembar Google Colab Notebook.....	13
Gambar 2.6 Tampilan Notebook Setting.....	13
Gambar 2.7 Tampilan Kode untuk koneksi Google Drive ke Google Colab.....	13
Gambar 2.8 Tampilan izin akses ke Google Drive.....	14
Gambar 2.9 Tampilan Google Colab sudah terkoneksi dengan Google Drive.....	14
Gambar 2.10 Codingan sederhana Hello World.....	14
Gambar 2.11 Kurva Naive Bayes Model.....	15
Gambar 2.12 Hyperplane Support Vector Machine.....	17
Gambar 2.13 Skema 10 Fold Cross Validation.....	19
Gambar 3.1 Alur Penelitian.....	22
Gambar 3.2 Komentar atau Ulasan Pengguna Aplikasi Traveloka.....	23
Gambar 3.3 Sisipkan Sel Kode.....	23
Gambar 3.4 Running Install Google Play Scraper.....	23
Gambar 3.5 Running Import Package Pandas dan Numpy.....	24
Gambar 3.6 Running Scraper Result.....	24
Gambar 3.7 Running Result Scraper Reviews.....	24
Gambar 3.8 Running File Download Crawling.txt.....	24
Gambar 3.9 Save atau Download File Crawling Data.....	25
Gambar 3.10 Data Mentah dari Hasil Crawling Data.....	25
Gambar 3.11 Data mentah yang diproses dari format TXT dan diubah ke format .XLSX.....	26
Gambar 3.12 Data Komentar setelah melewati proses manual pada Data Mentah.....	26
Gambar 3.13 Penghapusan Simbol pada Dataset Menggunakan Replace.....	26
Gambar 3.14 Menghilangkan White Space atau Spasi yang Tidak Perlu.....	26
Gambar 3.15 Dataset Duplikat Sebelum dilakukan Remove Duplicate.....	27
Gambar 3.16 Dataset Duplikat Setelah dilakukan Remove Duplicate.....	27
Gambar 3.17 Proses Pembersihan Data di RapidMiner.....	27
Gambar 3.18 Hasil labelling berdasarkan atribut kelas sentimen.....	28
Gambar 3.19 Distribusi Labelling Data.....	28
Gambar 4.1 Operator RapidMiner.....	35
Gambar 4.2 Dataset Labelling.....	36
Gambar 4.3 Operator Read Excel disambungkan dengan Operator Nominal to Text.....	36
Gambar 4.4 Operator Cross Validation disambungkan dengan Result.....	37
Gambar 4.5 Operator Naive Bayes di Lembar Training Cross Validation.....	37
Gambar 4.6 Operator SVM (LibSVM) di Lembar Training Cross Validation.....	38
Gambar 4.7 Operator k-NN di Lembar Training Cross Validation.....	38
Gambar 4.8 Parameter di Cross Validation.....	38
Gambar 4.9 Parameter di Support Vector Machine (LibSVM).....	39
Gambar 4.10 Parameter di K-Nearest Neighbor (k-NN).....	39
Gambar 4.11 Button Run pada RapidMiner.....	40
Gambar 4.12 Hasil Uji Performance Vector dengan Metode Naive Bayes.....	40
Gambar 4.13 Hasil Uji Performance Vector dengan Metode Support Vector Machine.....	40
Gambar 4.14 Hasil Uji Performance Vector dengan Metode K-Nearest Neighbor.....	41
Gambar 4.15 Hasil Labelling dan Prediksi Sentimen.....	42

DAFTAR TABEL

Tabel 2. 1 Kernel Hyperplane Support Vector Machine.....	18
Tabel 4.1 Hasil Proses Case Folding	31
Tabel 4. 2 Hasil dari Tokenization	32
Tabel 4.3 Hasil dari proses Stopword Removal	33
Tabel 4.4 Hasil dari Proses Stemming	34
Tabel 4.5 Hasil Accuracy Performance Vector yang diperoleh.....	41
Tabel 4.6 Analisis Ulasan Positif.....	42
Tabel 4.7 Analisis Ulasan Netral.....	43
Tabel 4.8 Analisis Ulasan Negatif	44
Tabel 4.9 Hasil F1-Score Naive Bayes	46
Tabel 4.10 Hasil F1-Score Support Vector Machine	46
Tabel 4.11 Hasil F1-Score K-Nearest Neighbor	47