

SKRIPSI

**ANALISIS SENTIMEN PADA REVIEW SPOTIFY
MENGUNAKAN MACHINE LEARNING**

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

Universitas Muhammadiyah Yogyakarta



Disusun oleh

NUR AFANDI MUJIB

20180140110

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH YOGYAKARTA**

2023

PERNYATAAN ORIGINALITAS

Saya yang bertanda tangan dibawah ini :

Nama : Nur Afandi Mujib

NIM : 20180140110

Program Studi : Teknologi Informasi

Fakultas : Fakultas Teknik

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis benar-benar merupakan hasil karya saya sendiri dan belum diajukan dalam bentuk apapun kepada perguruan tinggi mana pun. Sumber Informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain tidak disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir Skripsi ini.

Apabila kemudian hari terbukti atau dapat dibuktikan Skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Yogyakarta, 17 Maret 2023

Yang membuat pernyataan



Nur Afandi Mujib

MOTTO

“Barang siapa yang menghendaki kehidupan dunia maka wajib baginya memiliki ilmu, dan barang siapa yang menghendaki kehidupan akhirat, maka wajib baginya memiliki ilmu, dan barang siapa menghendaki keduanya wajib baginya memiliki ilmu”, (HR. Turmudzi).

KATA PENGANTAR

Puji syukur kehadirat Allah SWT, karena atas segala limpahan rahmat, hidayah, dan inayah-Nya, penulis dapat menyelesaikan penyusunan skripsi dengan judul " Analisis Sentimen Pada *Review* Spotify Menggunakan *Machine Learning*"

Penulisan skripsi ini bertujuan untuk memenuhi salah satu persyaratan dalam menyelesaikan pendidikan tinggi di Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Muhammadiyah Yogyakarta. Penulis menyadari bahwa dalam penyusunan skripsi ini, tidak terlepas dari bantuan, dukungan, dan dorongan dari berbagai pihak yang telah membantu selama proses penyusunan skripsi ini.

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Allah *Subhanahu Wa Ta'ala* yang telah memberikan petunjuk, kemurahan, keistiqomahan dan segala nikmat-Nya.
2. Teristimewa kepada kedua orang tua. Ibu dan ayah beserta keluarga yang tersayang dan terkasih yang telah memberikan do'a restu, semangat, motivasi, pengertian yang tulus baik spiritual dan materi yang tak terhingga sampai saya menuntut ilmu menempuh gelar Sarjana.
3. Bapak Cahya Damarjati, S.T. M. Eng., Ph.D. sebagai Ketua Program Studi Teknologi Informasi Universitas Muhammadiyah Yogyakarta yang telah memberikan kesempatan untuk menyusun skripsi ini.
4. Bapak Cahya Damarjati, S.T. M. Eng., Ph.D. dan Bapak Asep Setiawan S.Th.I., M.Ud. Selaku Dosen Pembimbing Skripsi yang telah memberikan bimbingan, meluangkan waktu, pikiran, perhatian serta ketelitian dalam penulisan Skripsi ini.
5. Semua dosen beserta karyawan Program Studi Teknologi Informasi Universitas Muhammadiyah Yogyakarta.
6. Teman-teman angkatan 2018 khususnya Cahaya Karunia Rizki, Muhammad Zulqoniun, Ricki Irawan, Muammar Rosy Rasyidik, David Prasetya, dan Sandy

Eka Yudha yang senantiasa memberikan dukungan, motivasi, dan tantuan selama menjalani perkuliahan.

7. Semua pihak yang memberikan bantuan baik mental maupun spiritual yang tidak bisa penulis sebutkan satu persatu.

Penulis menyadari bahwa Skripsi ini masih jauh dari sempurna, oleh karena itu kritik dan saran yang bersifat membangun selalu diharapkan. Semoga Skripsi ini bermanfaat bagi kita semua.

Yogyakarta, 17 Maret 2023

Yang membuat pernyataan



Nur Afandi Mujib

DAFTAR ISI

| | |
|---|-------------|
| HALAMAN PENGESAHAN I | ii |
| HALAMAN PENGESAHAN II | iii |
| PERNYATAAN ORIGINALITAS | iv |
| MOTTO | v |
| KATA PENGANTAR | vi |
| DAFTAR ISI | viii |
| DAFTAR GAMBAR | x |
| DAFTAR TABEL | xii |
| INTISARI | xiii |
| ABSTRACT | xiv |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Rumusan Masalah | 3 |
| 1.3 Tujuan Penelitian | 3 |
| 1.4 Manfaat Penelitian | 3 |
| 1.5 Struktur Penulisan | 4 |
| 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 Spotify | 9 |
| 2.2.3 Data Mining | 9 |
| 2.2.4 Data Set | 10 |
| 2.2.5 Machine Learning | 10 |
| 2.2.6 Python | 10 |
| 2.2.7 Google Colaboratory | 10 |
| 2.2.8 Support Vector Machine | 11 |

| | |
|--|-----------|
| 2.2.9 Logistic Regression | 12 |
| 2.2.10 Cross Validation | 13 |
| 2.2.11 Confusion Metrics..... | 14 |
| 2.2.11.1 Precision & Recall | 15 |
| 2.2.11.2 Accuracy | 16 |
| 2.2.11.3 F1-Score..... | 16 |
| BAB III METODE PENELITIAN..... | 17 |
| 3.1 Tempat dan Waktu Penelitian | 17 |
| 3.2 Alat..... | 17 |
| 3.3 Tahapan Penelitian | 18 |
| 3.4 Studi Literatur | 19 |
| 3.5 Pre-Processing Data | 19 |
| 3.6 Pembersihan Data | 19 |
| 3.7 Transformasi Data | 19 |
| BAB IV HASIL DAN PEMBAHASAN | 20 |
| 4.1 Pengambilan Data | 20 |
| 4.2 Pre-Processing Data | 21 |
| 4.2.1 Data Selection..... | 21 |
| 4.2.2 Data Cleaning (Pembersihan Data) | 22 |
| 4.2.3 Data Transformation (Transformasi Data)..... | 24 |
| 4.3 Implementasi Algoritma | 24 |
| 4.3.1 Klasifikasi Data..... | 25 |
| 4.3.2 Hasil Pengujian | 40 |
| 4.3.3 Analisis Proses dan Hasil Pengujian..... | 40 |
| BAB V KESIMPULAN DAN SARAN | 47 |
| 5.1 Kesimpulan | 47 |
| 5.2 Saran..... | 47 |
| DAFTAR PUSTAKA | 49 |
| LAMPIRAN..... | 52 |

DAFTAR GAMBAR

| | |
|---|----|
| Gambar 2. 1 Curva sederhana Support Vector Machine..... | 11 |
| Gambar 2. 2 Diagram sigmoid function | 13 |
| Gambar 2. 3 Contoh Confusion Metrics | 14 |
| Gambar 2. 4 Dua class Confusion Matrix | 15 |
| Gambar 3. 1 Tahapan penelitian..... | 18 |
| Gambar 4. 1 Pengambilan data di Kaggle.Com | 20 |
| Gambar 4. 2 Table data review spotify | 21 |
| Gambar 4. 3 Data sebelum melalui proses seleksi data..... | 22 |
| Gambar 4. 4 Data yang dibersihkan..... | 23 |
| Gambar 4. 5 Data yang sudah bersih | 23 |
| Gambar 4. 6 Proses Import Library Python | 25 |
| Gambar 4. 7 Proses read data csv | 25 |
| Gambar 4. 8 Hasil proses read data csv | 26 |
| Gambar 4. 9 Proses memanggil data kolom rating | 26 |
| Gambar 4. 10 Hasil proses memanggil data kolom rating | 27 |
| Gambar 4. 11 Proses pembagian data training dan data testing..... | 27 |
| Gambar 4. 12 Hasil dari proses pembagian data training dan data testing | 27 |
| Gambar 4. 13 Pre-Processing data training dan data testing..... | 28 |
| Gambar 4. 14 Proses Pre-Processing data training dan data testing | 28 |
| Gambar 4. 15 Proses metode klasifikasi..... | 29 |
| Gambar 4. 16 Hasil proses metode klasifikasi..... | 29 |
| Gambar 4. 17 Membuat prediksi menggunakan 2 metode | 30 |
| Gambar 4. 18 Mengevaluasi peforma dari 2 metode | 30 |
| Gambar 4. 19 Hasil mengevaluasi peforma dari 2 metode..... | 30 |
| Gambar 4. 20 Proses Grind Search mencari parameter terbaik | 31 |
| Gambar 4. 21 Hasil Grind Search pada metode Support Vector Machine..... | 31 |
| Gambar 4. 22 Hasil Grind Search pada metode Logistic Regression | 31 |
| Gambar 4. 23 mencetak parameter terbaik dari SVM & LR | 32 |
| Gambar 4. 24 Hasil mencetak parameter terbaik dari SVM & LR..... | 32 |
| Gambar 4. 25 Proses menghitung akurasi terbaik SVM & LR | 32 |
| Gambar 4. 26 Hasil Proses menghitung akurasi terbaik SVM & LR | 32 |
| Gambar 4. 27 Hasil akurasi dari klasifikasi SVM..... | 33 |
| Gambar 4. 28 Hasil akurasi dari klasifikasi LR..... | 33 |
| Gambar 4. 29 Code membuat heatmap visualisasi SVM | 34 |
| Gambar 4. 30 Heatmap visualisasi pada SVM..... | 35 |

| | |
|---|-----------|
| Gambar 4. 31 Code membuat heatmap visualisasi LR | 35 |
| Gambar 4. 32 Heatmap visualisasi pada LR | 36 |
| Gambar 4. 33 Representasi HTML SVM..... | 37 |
| Gambar 4. 34 Hasil Representasi HTML SVM..... | 37 |
| Gambar 4. 35 Representasi HTML LR..... | 38 |
| Gambar 4. 36 Hasil Representasi HTML LR..... | 38 |
| Gambar 4. 37 Proses prediksi kata “but” | 39 |
| Gambar 4. 38 Proses prediksi kata “worst” | 39 |
| Gambar 4. 39 Proses prediksi kata “love” | 39 |
| Gambar 4. 40 Hasil akurasi SVM & LR | 40 |

DAFTAR TABEL

| | |
|---|-----------|
| Tabel 4. 1 Atribut yang telah diseleksi | 22 |
| Tabel 4. 2 Transformasi data | 24 |
| Tabel 4. 3 Analisis ulasan algoritma SVM | 40 |
| Tabel 4. 4 Analisis ulasan algoritma LR..... | 42 |