ANALISIS SENTIMEN TWEET MENGENAI PEDULILINDUNGI MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

SKRIPSI

Diajukan untuk memenuhi persyaratan dalam memperoleh gelar Sarjana Komputer Program Studi Sistem Informasi

Disusun Oleh:

Farhan Setiyo Darusman

18082010044

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAWA TIMUR

SURABAYA

2022

SKRIPSI

ANALISIS SENTIMEN TWEET MENGENAI PEDULILINDUNGI MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

Disusun Oleh:
Farhan Setiyo Darusman
18082010044

Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Skripsi Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Pembangunan Nasional "Veteran" Jawa Timur
Pada Tanggal 11 Maret 2022

Menyetujui,

Pembimbing :

1.

Amalia Anjani Arifiyanti S.Kom., M.Kom
NIP. 19920812 201803 2 001

2.

Seftin Fitri Ana Wati, S.Kom., M.Kom
NPT. 212199 10 320267

Tim Penguji :

Eka Dyar Wahyuni, S.Kom. M.Kom
NIPPPK. 19841201 2021212 005

Dhian Satria Yudha Kartika, S.Kom., M.Kom.
NPT. 201198 60 522249

Anindo Saka Fitri, S.Kom.,M.Kom.
NPT. 212199 30 325268

Mengetahui,
Dekan Fakultas Ilmu Komputer
Universitas Pembangunan "Veteran" Jawa Timur

KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET DAN TEKNOLOGI UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAWA TIMUR FAKULTAS ILMU KOMPUTER

07042021

Dr. Ir. Ni Ketut Sari, M.T
NIP. 19650731 199203 2 001

LEMBAR PENGESAHAN

ANALISIS SENTIMEN TWEET MENGENAI PEDULILINDUNGI MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

Disusun Oleh:
Farhan Setiyo Darusman
18082010044

Telah disetujui mengikuti Ujian Negara Lisan
Periode 2022 pada Tanggal 11 Maret 2022

Menyetujui,

Dosen Pembimbing 1	Dosen Pembimbing 2
Amalia Anjani Arifiyanti S.Kom., M.Kom	Seftin Fitri Ana Wati, S.Kom., M.Kom
NIP. 19920812 2018032 001	NPT. 212199 10 320267

Mengetahui,
Ketua Program Studi Sistem Informasi
Fakultas Ilmu Komputer
Universitas Pembangunan Nasional "Veteran" Jawa Timur

Nur Cahyo Wibowo, S.kom, M.Kom
NIPPPK. 19790317 2021211 002

KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS PEMBANGUNAN NASIONAL
"VETERAN" JAWA TIMUR
Jl. Rungkut Madya Gunung Anyar Surabaya 60294
Telp (031) 8706369,8783189 Fax. (031) 8706372 Website
www.upnjatim.ac.id

KETERANGAN REVISI

Kami yang bertanda tangan di bawah ini menyatakan bahwa mahasiswa berikut:

Nama : Farhan Setiyo Darusman

NPM 18082010044

Program Studi : Sistem Informasi

Telah mengerjakan revisi Ujian Negara Lisan Skripsi pada tanggal 11 Maret 2022 dengan judul :

ANALISIS SENTIMEN TWEET MENGENAI PEDULILINDUNGI MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

Oleh karenanya mahasiswa tersebut diatas dinyatakan bebas revisi Ujian Negara Lisan Skripsi dan diijinkan untuk membukukan laporan SKRIPSI dengan judul tersebut.

Surabaya, 22 Maret 2022

Dosen penguji yang memeriksa revisi:

1. **Eka Dyar Wahyuni, S.Kom. M.Kom** { }
   NIPPPK. 19841201 2021212 005
2. **Dhian Satria Yudha Kartika, S.Kom., M.Kom.** { }
   NPT. 201198 60 522249
3. **Anindo Saka Fitri, S.Kom.,M.Kom.** { }
   NPT. 212199 30 325268

Mengetahui,

Dosen Pembimbing 1	Dosen Pembimbing 2
Amalia Anjani Arifiyanti S.Kom., M.Kom	**Seftin Fitri Ana Wati, S.Kom., M.Kom**
NIP. 19920812 201803 2 001	**NPT. 212199 10 320267**

SURAT PERNYATAAN

Saya, Mahasiswa Sistem Informasi UPN "Veteran" Jawa Timur, yang bertanda tangan di bawah ini :

Nama	: Farhan Setiyo Darusman
NPM	: 18082010044
Program Studi	: Sistem Informasi
Fakultas	: Ilmu Komputer

Menyatakan Bahwa Judul Skripsi / Tugas Akhir Saya Sebagai Berikut :

ANALISIS SENTIMEN TWEET MENGENAI PEDULILINDUNGI MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

Bukan merupakan plagiat dari skripsi / tugas akhir / penelitian orang lain dan juga bukan merupakan produk / software / hasil karya yang saya beli dari pihak lain.

Saya juga menyatakan bahwa Skripsi / Tugas Akhir ini adalah pekerjaan saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di UPN "Veteran" Jawa Timur maupun institusi pendidikan lain.

Jika ternyata dikemudian hari pernyatan ini terbukti tidak benar, maka saya bertanggung jawab penuh dan siap menerima segala konsekuensinya, termasuk pembatalan ijazah dikemudian hari.

Hormat Saya,

Farhan Setiyo Darusman

Judul	:	**ANALISIS SENTIMEN TWEET MENGENAI PEDULILINDUNGI MENGGUNAKAN METODE SUPPORT VECTOR MACHINE**
Penulis	:	**Farhan Setiyo Darusman**
Pembimbing 1	:	**Amalia Anjani Arifiyanti S.Kom., M.Kom**
Pembimbing 2	:	**Seftin Fitri Ana Wati, S.Kom., M.Kom**

ABSTRAK

Pada 23 Agustus 2021, pemerintah mulai memperbolehkan kegiatan di luar rumah dengan syarat penggunaan aplikasi PeduliLindungi. Aplikasi PeduliLindungi memiliki banyak manfaat namun banyak kontroversi timbul di masyarakat. Beragam opini dalam bentuk tweet diutarakan masyarakat, baik itu opini positif dan negatif. Dalam penelitian ini akan dilakukan klasifikasi terhadap tweet tersebut ke dalam dua jenis sentimen yaitu positif dan negatif menggunakan metode *SVM* dan pemilihan fitur dua tahap. Model tersebut memperoleh nilai akurasi sebesar 64,08% dengan 841 fitur dan waktu pemrosesan selama 0,033 detik dengan penggunaan CPU 9,6%. Kurangnya fitur unik pada dataset menyebabkan akurasi yang didapat hanya sebesar 61% hingga 64%. Hal yang membuat Pedulilindungi mendapat sentimen negatif yaitu pengguna Pedulilindungi menganggap Pedulindungi tidak aman, lambat, peladennya sering bermasalah sehingga susah diakses, susah digunakan, serta kode OTP yang tidak kunjung masuk. Masyarakat juga mengeluhkan bagaimana dengan orang berusia lanjut yang tidak mempunyai gawai pintar untuk menjalankan pedulilindungi. Dengan analisis ini, dapat dibuat hasil evaluasi untuk pengembangan aplikasi dan regulasi Pedulilindungi. Model yang dibuat diimplementasikan pada aplikasi berbasis web yang memiliki fitur klasifikasi tweet dan menampilkan visualisasi hasil klasifikasi.

Kata Kunci : *analsis, sentimen, tweet, klasifikasi, model, fitur, SVM, aplikasi.*

DAFTAR ISI

Hal

ABSTRAK vi
KATA PENGANTAR vii
DAFTAR ISI ix
DAFTAR GAMBAR xi
DAFTAR TABEL xii
DAFTAR LAMPIRAN xiii
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah 6
1.3 Batasan Masalah 7
1.4 Tujuan Penelitian 7
1.5 Sistematika Penulisan 7
BAB II TINJAUAN PUSTAKA 10
2.1 Penelitian Terdahulu 10
2.2 Dasar Teori 16
2.2.1 PeduliLindungi 16
2.2.2 Twitter 17
2.2.3 Scraping / Crawling 18
2.2.4 Analisis Sentimen 18
2.2.5 Information Gain 19
2.2.6 Genetic Algorithm 19
2.2.7 Support Vector Machine 20
2.2.8 Evaluasi Model 21
BAB III METODOLOGI PENELITIAN 22
3.1 Pembuatan Model 23
3.1.1 Pengambilan Data 23
3.1.2 Filter Data 24
3.1.3 Data Labeling 24
3.1.4 Data Understanding 24
3.1.5 Data Preprocessing 24
3.1.6 Pembagian Data Latih dan Data Uji 33
3.1.7 Pemilihan Fitur Dua Tahap 33
3.1.8 Klasifikasi Dengan Support Vector Machine 34
3.2 Evaluasi Model 34
3.3 Visualisasi Data Hasil Klasifikasi 35
3.4 Analisis Deskripsi Data 35
3.5 Pembuatan Aplikasi 36
BAB IV HASIL DAN PEMBAHASAN 37
4.1 Pembuatan Model 37
4.1.1 Pengambilan Data 37
4.1.2 Filter Data 38
4.1.3 Data Labeling 40
4.1.4 Data Understanding 42
4.1.5 Data Preprocessing 44
4.1.6 Pembagian Data Latih dan Data Uji 62
4.1.7 Pemilihan Fitur Dua Tahap 63
4.1.8 Klasifikasi Dengan Support Vector Machine 64
4.2 Evaluasi Model 64
4.3 Visualisasi Data Hasil Klasifikasi 66

4.4 Analisis Deskripsi Data 68
4.5 Pembuatan Aplikasi 72
BAB V KESIMPULAN DAN SARAN 76
5.1 Kesimpulan 76
5.2 Saran 77
DAFTAR PUSTAKA 78
LAMPIRAN 81

DAFTAR GAMBAR

Hal

Gambar 2.1 Alur Genetic Algorithm 20
Gambar 3.1 Metode Penelitian 22
Gambar 3.2 Diagram Alir Pembuatan Model 23
Gambar 3.3 Diagram Alir Tahap Preprocessing 25
Gambar 3.4 Diagram Alir Tahap Penerjemahan Emotikon 26
Gambar 3.5 Diagram Alir Mengubah Kata Singkatan 30
Gambar 3.6 Diagram Alir Mengubah Kata Tidak Baku 31
Gambar 3.7 Diagram Alir Stopword Removal 32
Gambar 4.1 Source Code Pengambilan Data 37
Gambar 4.2 Source Code DEA Teks 42
Gambar 4.3 Teks Deskripsi Data 42
Gambar 4.4 Source Code Visualisasi Data 43
Gambar 4.5 Visualisasi Grafik Batang Perbandingan Kelas Sentimen 43
Gambar 4.6 Visualisasi Grafik Garis Tren Kelas Sentimen 43
Gambar 4.7 Source Code Penerjemahan Emoji 45
Gambar 4.8 Source Code Penerjemahan Emotikon 46
Gambar 4.9 Source Code Case Folding 48
Gambar 4.10 Source Code Cleansing 50
Gambar 4.11 Source Code Mengubah Kata Singkatan 53
Gambar 4.12 Source Code Mengubah Kata Tidak Baku 55
Gambar 4.13 Source Code Penerjemahan Bahasa Asing 57
Gambar 4.14 Source Code Remove Stopword 59
Gambar 4.15 Source Code Stemming 61
Gambar 4.16 Source Code Membagi Data Latih dan Data Uji 62
Gambar 4.17 Source Code Pemilihan Fitur 63
Gambar 4.18 Source Code Klasifikasi 64
Gambar 4.19 Source Code Evaluasi Model 64
Gambar 4.20 Evaluasi Model Dengan Pemilihan Fitur Satu Tahap 65
Gambar 4.21 Evaluasi Model Dengan Pemilihan Fitur Dua Tahap 65
Gambar 4.22 Source Code Visualisasi Grafik Batang 66
Gambar 4.23 Visualisasi Grafik Batang Perbandingan Hasil Prediksi 67
Gambar 4.24 Source Code Visualisasi Word Cloud 67
Gambar 4.25 Visualisasi Wordcloud Sentimen Positif 68
Gambar 4.26 Visualisasi Wordcloud Sentimen Negatif 68
Gambar 4.27 Source Code Pemeriksaan Bot 71
Gambar 4.28 Grafik Batang Berdasarkan Sentimen dan Kemungkinan Bot 72
Gambar 4.29 Source Code File PHP 73
Gambar 4.30 Source Code File Python 73
Gambar 4.31 Tampilan Website Halaman Utama 73
Gambar 4.32 Tampilan Webiste Halaman Visualisasi 73

DAFTAR TABEL

Hal

Tabel 2.1 Penelitian Sebelumnya 10
Tabel 3.1 Tabel Emotikon dan Emoji 26
Tabel 3.2 Tabel Confussion Matrix 35
Tabel 4.1 Contoh Data bukan sentimen 38
Tabel 4.2 Contoh Data Sentimen 39
Tabel 4.3 Contoh Data Yang Telah Diberi Label 40
Tabel 4.4 Contoh Data Hasil Penerjemahan Emoji dan Emotikon 46
Tabel 4.5 Contoh Data Hasil Case Folding 48
Tabel 4.6 Contoh Data Hasil Cleansing 51
Tabel 4.7 Contoh Data Hasil Mengubah Kata Singkatan 53
Tabel 4.8 Contoh Data Hasil Mengubah Kata Tidak Baku 55
Tabel 4.9 Contoh Data Hasil Penerjemahan Bahasa Asing 57
Tabel 4.10 Contoh Data Hasil Stopword Removal 59
Tabel 4.11 Contoh Data Hasil Stemming 61
Tabel 4.12 Hasil Klasifikasi Web Aplikasi 74
Tabel 4.13 Black Box Testing Web Aplikasi 75

DAFTAR LAMPIRAN