

TUGAS AKHIR

**ANALISIS SENTIMEN MENGGUNAKAN
WORD2VEC DAN SUPPORT VECTOR MACHINE
(STUDI KASUS: SENTIMEN PELANGGAN SICEPAT)**



DESI NATALIA BR. SIDAURUK

17102077

PROGRAM STUDI S1 TEKNIK INFORMATIKA

FAKULTAS INFORMATIKA

INSTITUT TEKNOLOGI TELKOM PURWOKERTO

2021

TUGAS AKHIR

**ANALISIS SENTIMEN MENGGUNAKAN
WORD2VEC DAN SUPPORT VECTOR MACHINE
(STUDI KASUS: SENTIMEN PELANGGAN SICEPAT)**

**SENTIMENT ANALYSIS USING
WORD2VEC AND SUPPORT VECTOR MECHINE
(CASE STUDY: CUSTOMER SENTIMENT SICEPAT)**

Disusun Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Komputer



DESI NATALIA BR. SIDAURUK

17102077

PROGRAM STUDI S1 TEKNIK INFORMATIKA

FAKULTAS INFORMATIKA

INSTITUT TEKNOLOGI TELKOM PURWOKERTO

2021

Lembar Pengesahan Pembimbing

**ANALISIS SENTIMEN MENGGUNAKAN
WORD2VEC DAN SUPPORT VECTOR MACHINE
(STUDI KASUS: SENTIMEN PELANGGAN SICEPAT)**

**SENTIMENT ANALYSIS USING
WORD2VEC AND SUPPORT VECTOR MACHINE
(CASE STUDY: CUSTOMER SENTIMENT SICEPAT)**

Dipersiapkan dan Disusun Oleh
DESI NATALIA BR. SIDAURUK

17102077

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tugas Akhir
Pada hari Senin, 1 September 2021

Pembimbing I,



8/9
Siti Khomsah

(Siti Khomsah, S.Kom., M.CS)

NIDN. 0517108101

Pembimbing II,



Atika Ratna Dewi

(Atika Ratna Dewi, S.Si., M.Sc.)

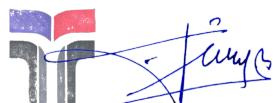
NIDN. 0615128703

Tugas Akhir ini diterima sebagai salah satu persyaratan

untuk memperoleh gelar Sarjana Komputer

Tanggal 1 September 2021

Skripsi/Tugas Akhir ini sudah diujikan dan dinyatakan sah
tanpa tanda tangan pembimbing dan penguji.
Purwokerto,
Dekan Fakultas Informatika
INSTITUT TEKNOLOGI TELKOM PURWOKERTO



Auliya Burhanuddin, S.Si., M. Kom
NIK. 0630058202

Kaprodi,

Auliya Burhanuddin, S.Si., M. Kom

NIK. 0630058202

Lembar Penetapan Pengaji

**ANALISIS SENTIMEN MENGGUNAKAN
WORD2VEC DAN SUPPORT VECTOR MACHINE
(STUDI KASUS: SENTIMEN PELANGGAN SICEPAT)**

**SENTIMENT ANALYSIS USING
WORD2VEC AND SUPPORT VECTOR MACHINE
(CASE STUDY: CUSTOMER SENTIMENT SICEPAT)**

Dipersiapkan dan Disusun Oleh

DESI NATALIA BR. SIDAURUK

17102077

Tugas Akhir Telah diuji dan Dinilai Panitia Pengaji Program

Studi S1 Teknik Informatika

Fakultas Informatika Institut Teknologi Telkom Purwokerto

Pada Tanggal 1 September 2021

Ketua

Pengaji


Siti Khomsah, S.Kom., M.CS

NIDN. 0517108101

Pengaji I,



Faisal Dharma Adhinata, S.Kom., M.Cs

NIDN. 0607079301

Pengaji II,



Paradise, S.Kom., M.Kom.

NIDN. 0624059501

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang Bertandatangan di bawah ini,

Nama Mahasiswa : Desi Natalia BR. Sidauruk

NIM : 17102077

Program Studi : S1 Teknik Informatika

Menyatakan bahwa Tugas Akhir dengan judul berikut:

ANALISIS SENTIMEN MENGGUNAKAN WORD2VEC DAN SUPPORT VECTOR MACHINE (STUDI KASUS: SENTIMEN PELANGGAN SICEPAT)

Dosen Pembimbing Utama : Siti Khomsah, S.Kom., M.CS

Dosen Pembimbing Pendamping : Atika Ratna Dewi, S.Si., M.Sc.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Institut Teknologi Telkom Purwokerto maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan, dan penelitian Saya Sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing.
3. Dalam Karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggungjawab Saya, bukan tanggungjawab Institut Teknologi Telkom Purwokerto.
5. Pernyataan ini Saya buat dengan sesungguhnya, apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka Saya bersedia menerima Sanksi Akademik dengan pencabutan gelar yang sudah diperoleh serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Purwokerto, 26 Agustus 2021,

Yang Menyatakan,


(Desi Natalia BR. Sidauruk)

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Tuhan Yang Maha Esa atas karuniaNya dan atas segala limpahan rahmat Nya, sehingga penulis dapat menyelesaikan laporan penyusunan sidang tugas akhir ini dengan baik.

Penulis mengucapkan terima kasih kepada pihak yang telah membantu dalam penelitian ini, untuk itu penulis dalam kesempatan ini mengucapkan terima kasih kepada:

1. Bapak Lukman Sidauruk dan Ibu Erika Situmeang selaku kedua orang tua tercinta atas segala bantuan, bimbingan, dorongan serta doa yang diberikan kepada penulis selama penyusunan skripsi.
2. Kakak Ester Anggreani Sidauruk, A.Md..T. kedua adik saya Enzela BR Sidauruk dan Farel Suta Manggala Sidauruk terima kasih selalu mengingatkan penulis untuk semangat, selalu mendoakan, dan bantuan sehingga dapat menyelesaikan skripsi ini.
3. Bapak Dr. Arfianto Fahmi, S.T., M.T., IPM selaku Rektor Institut Teknologi Telkom Purwokerto.
4. Ibu Dr. Tenia Wahyuningrum, S.Kom., M.T. selaku Dekan Fakultas Informatika Institut Teknologi Telkom Purwokerto.
5. Bapak Auliya Burhanuddin, S.Si., M. Kom selaku Kaprodi Fakultas Informatika Institut Teknologi Telkom Purwokerto.
6. Ibu Siti Khomsah, S.Kom., M.Cs. selaku Dosen Pembimbing pertama yang telah memberikan bimbingan, arahan, dorongan kepada penulis sehingga penyusunan skripsi ini berjalan dengan baik.
7. Ibu Atika Ratna Dewi, S.Si., M.Sc. selaku Dosen Pembimbing kedua sekaligus Wali Dosen atas segala bimbingan dan saran kepada penulis sehingga penulis dapat menyelesaikan skripsi ini.
8. Untuk Om Pdt. MM Sihombing, S.Th., M.Th dan Tante terimakasih atas dukungan dan doa yang dipanjatkan untuk penulis.
9. Kepada Saudara-saudara terkasih Pemuda/I GPDI Ajibarang yang senantiasa memberikan semangat dan doa selama proses pembuatan skripsi. Penulis sangat

bersyukur boleh dipertemukan dengan saudara-saudara yang begitu luar biasa seperti kalian.

10. Teruntuk teman dekat saya selama perkuliahan, Dilla Febria Elisty, S.Kom, Eka Yuni Arsih, dan Aisyah Ayu Wulandari Terimakasih buat canda, tawa, suka, dan duka bersama kalian, penulis sangat bersyukur mengenal pribadi-pribadi unik seperti kalian. Kiranya hubungan pertemanan yang sudah dijalani selama empat tahun dapat terus dijalin sekalipun jarak akan menjauhkan kita. Semoga Kebaikan dan Kesuksesan menemani perjalanan kita ke depan.
11. Seluruh rekan-rekan saya kelas B Prodi S1 Teknik Informatika angkatan 2017 yang selama ini saling menyemangati dan akhirnya satu per satu lulus sidang serta tak lupa juga buat teman-teman yang lainnya yang tidak dapat disebutkan satu per satu namanya.
12. Saudara Nathanael Novarian yang telah memberikan semangat serta waktunya selama penulis menyelesaikan skripsi ini.
13. Terakhir, penulis hendak menyapa setiap nama yang tidak dapat penulis cantumkan satu per satu, terima kasih atas doa yang senantiasa mengalir tanpa sepengetahuan penulis. Terima kasih sebanyak-banyaknya kepada orang-orang yang turut bersukacita atas keberhasilan penulis menyelesaikan Skripsi ini. Tuhan Memberkati Kita Semua.

Semoga ketulusan dan kebaikan yang telah diberikan kepada penulis menjadi berkat bagi kita semua. Sebagai manusia biasa, tentunya penulis masih memiliki banyak kekurangan pengetahuan dan pengalaman pada topik yang diangkat dalam Skripsi ini, begitu pula dalam penulisannya yang masih banyak terdapat kekurangan. Penulis berharap dapat bermanfaat bagi para pembaca khususnya jurusan Teknik Informatika.

Purwokerto, 26 Agustus 2021



Desi Natalia BR. Sidauruk

DAFTAR ISI

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR	iv
DAFTAR ISI.....	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL.....	x
ABSTRAK	xi
ABSTRACT	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Pertanyaan Penelitian	3
1.4 Batasan Masalah.....	3
1.5 Tujuan Penelitian.....	3
1.6 Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Kajian Pustaka Peneltian	5
2.2 Tabel Kajian Pustaka.....	8
2.3 Dasar Teori	15
BAB III METODOLOGI PENELITIAN	21
3.1 Subyek dan Obyek Penelitian.....	21
3.2 Alat dan Bahan Penelitian.....	21
3.3 Diagram Alir Penelitian.....	22
BAB IV	33
4.1 Crawling Data.....	33
4.2 Labelling.....	33
4.3 Text Preprocessing	34
4.4 Model Klasifikasi dengan Support Vector Machine	39
4.5 Analisis Hasil	42
BAB V.....	46
5.1 Kesimpulan.....	46
5.2 Saran	46

DAFTAR PUSTAKA	47
----------------------	----

DAFTAR GAMBAR

Gambar 2.1 Model arsitektur <i>continous bags-of-words</i> (CBOW)	18
Gambar 2.2 Model arsitektur Skip-gram	19
Gambar 2.3 Ilustrasi Metode Support Vector Machine	20
Gambar 3.1 Diagram Alir Penelitian	21
Gambar 3.2 Alir Diagram Membangun Model.....	22
Gambar 3.3 Arsitektur Skip-gram.....	24
Gambar 3.4 Contoh Skip-gram tweet Sicepat.....	24
Gambar 3.5 Alir Diagram <i>Support Vector Machine</i>	27
Gambar 4.1 Program dari tahap cleaning.....	34
Gambar 4.2 Program dari tahap filtering	34
Gambar 4.3 Program dari case folding	35
Gambar 4.4 Program dari Stopword Removal.....	36
Gambar 4.5 Proses dari Word2vec.....	37
Gambar 4.6 Contoh Proses Word2vec	37
Gambar 4.7 Proses dari SVM.....	38
Gambar 4.8 Hasil Percobaan Data Testing 50%	41
Gambar 4.9 Hasil Percobaan Data Testing 20%	41
Gambar 4.10 Hasil Percobaan Data Testing 60%	42
Gambar 4.11 Program dari Confusion Matrix	43
Gambar 4.12 Hasil Confusion Matrix SVM Kernel Linear.....	43
Gambar 4.13 Hasil Confusion Matrix SVM Kernel RBF.....	44
Gambar 4.14 Hasil Confusion Matrix SVM Kernel Polynomial.....	44
Gambar 4.15 Hasil Confusion Matrix SVM Kernel Sigmoid.....	45

DAFTAR TABEL

Tabel 3.1 Hasil Encoding Setiap Term dalam Opini	26
Tabel 3.2 Word Output dan Vektor	27
Tabel 3.3 Word Output dan Input Opini a	28
Tabel 3.4 Word Output dan Input Opini b	28
Tabel 3.5 Perhitungan SVM Linier.....	31
Tabel 4.1 Data komentar pengguna sicepat dengan format CSV	33
Tabel 4.2 Contoh pelabelan manual.....	34
Tabel 4.3 Pelabelan	34
Tabel 4.4 Contoh Menghapus username pengguna twitter.....	35
Tabel 4.5 Contoh tweet yang tidak memiliki arti.....	35
Tabel 4.6 Contoh tweet huruf kapital menjadi kecil	36
Tabel 4.7 Contoh Tweet tidak memiliki makna penting.....	37
Tabel 4.8 Model Klasifikasi Support Vector Machine dengan tiga kelas....	39
Tabel 4.9 Kernel Linear Support Vector Machine.....	40
Tabel 4.10 Hasil Pengujian Masing-Masing Kernel	45