

TUGAS AKHIR

ANALISIS SENTIMEN BERBASIS ASPEK PADA
EDOM PEMBELAJARAN MENGGUNAKAN
METODE CNN DAN WORD2VEC



MUHAMMAD IRFANI

20104037

PROGRAM STUDI REKAYASA PERANGKAT LUNAK
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO

2024

TUGAS AKHIR

**ANALISIS SENTIMEN BERBASIS ASPEK PADA
EDOM PEMBELAJARAN MENGGUNAKAN
METODE CNN DAN WORD2VEC**

**ASPECTS-BASED SENTIMENT ANALYSIS IN
LEARNING EDOM USING CNN AND WORD2VEC
METHODS**



MUHAMMAD IRFANI

20104037

**PROGRAM STUDI REKAYASA PERANGKAT LUNAK
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO**

2024

LEMBAR PERSETUJUAN PEMBIMBING

**ANALISIS SENTIMEN BERBASIS ASPEK PADA
EDOM PEMBELAJARAN MENGGUNAKAN
METODE CNN DAN WORD2VEC**


**ASPECTS-BASED SENTIMENT ANALYSIS IN
LEARNING EDOM USING CNN AND WORD2VEC
METHODS**

Dipersiapkan dan Disusun Oleh

MUHAMMAD IRFANI
20104037

**Fakultas Informatika
Institut Teknologi Telkom Purwokerto
Pada Tanggal: 12 Januari 2024**

Pembimbing Utama,



Siti Khomsah, S.Kom., M.Cs

NIDN. 051710810

LEMBAR PENGESAHAN

**ANALISIS SENTIMEN BERBASIS ASPEK PADA
EDOM PEMBELAJARAN MENGGUNAKAN
METODE CNN DAN WORD2VEC**

**ASPECTS-BASED SENTIMENT ANALYSIS IN
LEARNING EDOM USING CNN AND WORD2VEC
METHODS**

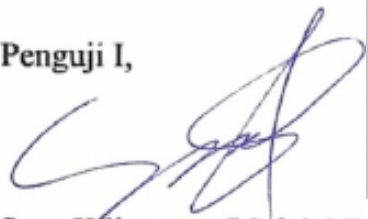
Disusun Oleh:

MUHAMMAD IRFANI

20104037

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tugas Akhir
Pada Jumat, 19 Januari 2024

Penguji I,



Sena Wijayanto, S.Pd., M.T,

NIDN. 0613109201

Penguji II,



Dr. Yogo Dwi Prasetyo, S.Si., M.Si

NIDN. 0101088702

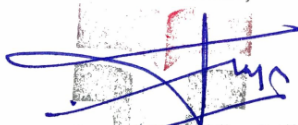
Pembimbing Utama,



Siti Khomsah, S.Kom., M.Cs

NIDN. 051710810

Dekan,



Auliya Burhanuddin, S.Si., M.Kom

NIK: 19820008

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertandatangan di bawah ini,

Nama Mahasiswa : Muhammad Irfani
NIM : 20104037
Program Studi : Rekayasa Perangkat Lunak



Menyatakan bahwa Tugas Akhir dengan judul berikut:
ANALISIS SENTIMEN BERBASIS ASPEK PADA EDOM
PEMBELAJARAN MENGGUNAKAN METODE CNN DAN WORD2VEC

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

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, 30 Januari 2024,

Yang Menyatakan,

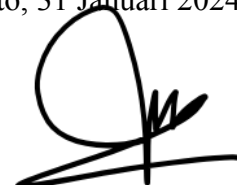


PAB5FAKX778254883
Muhammad Irfani

KATA PENGANTAR

Puji dan syukur penulis sampaikan kepada Allah SWT, yang atas berkat dan rahmat-Nya, penulis berhasil menyelesaikan tugas akhir ini. Penulisan tugas akhir ini merupakan bagian dari syarat untuk meraih gelar sarjana di Program Studi S1 Rekayasa Perangkat Lunak Institut Teknologi Telkom Purwokerto. Penulis menyadari sepenuhnya bahwa tanpa bimbingan dan bantuan dari berbagai pihak, dari awal masa perkuliahan hingga akhirnya, penyelesaian laporan ini akan menjadi suatu tantangan yang sangat sulit bagi penulis. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Allah SWT karena atas izin-Nya penulis dapat menyelesaikan Tugas Akhir
2. Dr. Tenia Wahyuningrum, S.Kom., M.T. Selaku Rektor Institut Teknologi Telkom Purwokerto.
3. Auliya Burhanudin, S.Si., M.Kom. Selaku Dekan Fakultas Informatika.
4. Ariq Cahya Wardhana S.Kom., M.Kom. Selaku Ketua Program Studi S1 Rekayasa Perangkat Lunak
5. Ayah dan ibu yang telah banyak memberikan perhatian, doa dan semangat
6. Siti Khomsah, S.Kom., M.Cs. selaku Dosen Pembimbing utama yang dengan sabar membimbing dan membantu dalam menyelesaikan penyusunan naskah tugas akhir dengan baik.
7. Diri sendiri yang telah berhasil menyelesaikan penyusunan Tugas Akhir ini sampai selesai
8. Teman terbaik saya dan dosen yang tidak bisa disebutkan namanya satu persatu, yang telah menemani dan memberi motivasi selama belajar di Institut Teknologi Telkom Purwokerto.

Purwokerto, 31 Januari 2024



Muhammad Irfani

DAFTAR ISI

TUGAS AKHIR.....	i
LEMBAR PERSETUJUAN PEMBIMBING.....	ii
LEMBAR PENGESAHAN.....	iii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR.....	iv
KATA PENGANTAR.....	v
DAFTAR ISI.....	vi
DAFTAR TABEL.....	viii
DAFTAR GAMBAR.....	ix
DAFTAR LAMPIRAN.....	xi
ABSTRAK.....	xii
ABSTRACT.....	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan 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. Tinjauan Pustaka.....	5
2.2. Landasan Teori.....	12
2.2.1 <i>Sentiment Analysis</i>	12
2.2.2 <i>Aspect Based Sentiment Analysis</i>	12
2.2.3 <i>Convolutional Neural Network (CNN)</i>	13
2.2.3.1 <i>Convolutional Layer (CONV)</i>	13
2.2.3.2 <i>Activation Function</i>	14
2.2.3.3 <i>Pooling Layer</i>	15
2.2.3.4 <i>Flatten</i>	17
2.2.3.5 <i>Fully Connected Layer (FC)</i>	17
2.2.5 Keras.....	18
2.2.6 Word2vec.....	18
2.2.7 <i>Confusion Matrix</i>	19
2.2.8 <i>Oversampling</i>	22
2.2.9 <i>Text to Sequences</i>	25
2.2.10 <i>Padding</i>	26
2.2.11 <i>Imbalance Data</i>	26

BAB III METODOLOGI PENELITIAN.....	28
3.1 Objek dan Subjek Penelitian.....	28
3.2 Alat dan Bahan.....	28
3.2.1 Data.....	28
3.2.2 Spesifikasi Kebutuhan Perangkat Keras.....	28
3.3 Diagram Alir Penelitian.....	30
3.3.1 Perumusan Masalah.....	30
3.3.2 Studi Literasi.....	31
3.3.3 Pengumpulan <i>Dataset</i>	31
3.3.4 <i>Data Cleaning</i>	31
3.3.5 Penanganan Slang <i>Word</i>	32
3.3.6 <i>Data Labelling</i>	32
3.3.7 <i>Data Splitting</i>	33
3.3.8 <i>Word Embedding</i>	33
3.3.9 <i>Oversampling</i>	36
3.3.10 <i>Modelling</i>	38
3.3.11 <i>Evaluation</i>	39
BAB IV HASIL DAN PEMBAHASAN.....	40
4.1 Pengumpulan <i>Dataset</i>	40
4.2 <i>Data Cleaning</i>	41
4.3 Penanganan Slang <i>Word</i>	42
4.4 <i>Data Labelling</i>	44
4.5 <i>Data Split</i>	49
4.6 <i>Word Embedding</i>	52
4.7 <i>Oversampling</i>	54
4.8 <i>Modelling</i>	57
4.9 <i>Evaluation</i>	87
BAB V KESIMPULAN DAN SARAN.....	100
5.1 Kesimpulan.....	100
5.2 Saran.....	100
DAFTAR PUSTAKA.....	101
LAMPIRAN.....	104

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu.....	8
Tabel 3.1 Spesifikasi Kebutuhan Perangkat keras.....	28
Tabel 3.2 Spesifikasi Kebutuhan Perangkat Lunak.....	29
Tabel 3.3 Ilustrasi Penanganan Slang <i>Word</i>	32
Tabel 3.4 Contoh <i>Text to Sequences</i>	34
Tabel 3.5 Contoh penggunaan <i>pad_sequences</i>	35
Tabel 3.6 Algoritma <i>Oversampling</i>	37
Tabel 3.7 Konfigurasi CNN.....	38
Tabel 4.1 Contoh <i>dataset raw</i>	40
Tabel 4.2 Sampel hasil data <i>cleaning</i>	42
Tabel 4.3 Sampel hasil dari penanganan slang <i>word</i>	43
Tabel 4.4 Sampel hasil data <i>labelling sentimen</i>	44
Tabel 4.5 Sampel <i>non-aspect word</i>	45
Tabel 4.6 Sampel hasil mengganti label netral ke label yang lebih cocok.....	46
Tabel 4.7 Jumlah Distribusi Label Sentimen.....	48
Tabel 4.8 Jumlah Distribusi Label Aspek.....	48
Tabel 4.9 Hasil pembagian data untuk model sentimen.....	50
Tabel 4.10 Hasil pembagian data untuk model aspek.....	51
Tabel 4.11 Sampel hasil <i>text to sequences</i>	52
Tabel 4.12 Sampel hasil <i>pad sequences</i>	53
Tabel 4.13 Hasil Latihan Algoritma SMOTE <i>Dataset Sentimen</i>	59
Tabel 4.14 Hasil Latihan Algoritma <i>Random Oversampling Dataset Sentimen</i> ...	62
Tabel 4.15 Hasil Latihan Algoritma ADASYN <i>Dataset Sentimen</i>	65
Tabel 4.16 Hasil Latihan Algoritma SMOTE-NC <i>Oversampling Dataset Sentimen</i>	68
Tabel 4.17 Hasil Latihan Algoritma <i>Borderline SMOTE Dataset Sentimen</i>	71
Tabel 4.18 Hasil Latihan Algoritma <i>Borderline SMOTE Dataset Aspek</i>	74
Tabel 4.19 Hasil Latihan Algoritma <i>Random Oversampling Dataset Aspek</i>	76
Tabel 4.20 Hasil Latihan Algoritma ADASYN <i>Dataset Aspek</i>	79
Tabel 4.21 Hasil Latihan Algoritma SMOTE-NC <i>Dataset Aspek</i>	82
Tabel 4.22 Hasil Latihan Algoritma <i>Borderline SMOTE Dataset Aspek</i>	85
Tabel 4.23 Hasil Evaluasi Akurasi Model Sentimen.....	97
Tabel 4.24 Hasil Evaluasi Akurasi Model Aspek.....	98

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi ABSA.....	13
Gambar 2.2 Ilustrasi dari ketiga jenis <i>pooling layer</i>	16
Gambar 2.3 Ilustrasi <i>Flattening</i>	17
Gambar 2.4 Ilustrasi <i>Word Embedding</i> dengan CBOW.....	19
Gambar 2.5 <i>Confusion Matrix</i>	20
Gambar 2.6 Ilustrasi SMOTE.....	23
Gambar 2.7 Ilustrasi <i>Random Oversampling</i>	23
Gambar 3.1 Diagram Alir Penelitian.....	30
Gambar 3.2 Contoh <i>Dataset</i> EDOM ITTP.....	31
Gambar 3.3 <i>Flowchart</i> Data <i>Labelling</i>	32
Gambar 3.4 Contoh Implementasi CBOW.....	34
Gambar 3.5 Sebelum <i>Oversampling</i>	36
Gambar 3.6 Setelah <i>Oversampling</i>	37
Gambar 4.1 <i>Source code</i> <i>Label Converter</i> <i>Dataset</i> Sentimen.....	49
Gambar 4.2 <i>Source code</i> <i>Label Converter</i> <i>Dataset</i> Aspek.....	50
Gambar 4.3 Hasil <i>word embedding</i>	54
Gambar 4.4 <i>Dataset</i> Sentimen sebelum <i>oversampling</i>	55
Gambar 4.5 <i>Dataset</i> Sentimen sesudah <i>oversampling</i>	55
Gambar 4.6 <i>Dataset</i> Aspek sebelum <i>oversampling</i>	56
Gambar 4.7 <i>Dataset</i> Aspek sesudah <i>oversampling</i>	57
Gambar 4.8 Hasil <i>Learning Curve</i> Algoritma SMOTE <i>Dataset</i> Sentimen.....	62
Gambar 4.9 Hasil <i>Learning Curve</i> Algoritma <i>Random Oversampling</i> <i>Dataset</i> Sentimen.....	64
Gambar 4.10 Hasil <i>Learning Curve</i> Algoritma ADASYN <i>Dataset</i> Sentimen.....	67
Gambar 4.11 Hasil <i>Learning Curve</i> Algoritma SMOTE-NC <i>Dataset</i> Sentimen.....	70
Gambar 4.12 Hasil <i>Learning Curve</i> Algoritma <i>Borderline</i> SMOTE <i>Dataset</i> Sentimen.....	73
Gambar 4.13 Hasil <i>Learning Curve</i> Algoritma SMOTE <i>Dataset</i> Aspek.....	75
Gambar 4.14 Hasil <i>Learning Curve</i> Algoritma <i>Random Oversampling</i> <i>Dataset</i> Aspek.....	78
Gambar 4.15 Hasil <i>Learning Curve</i> Algoritma ADASYN <i>Dataset</i> Aspek.....	81
Gambar 4.16 Hasil <i>Learning Curve</i> Algoritma SMOTE-NC <i>Dataset</i> Aspek.....	84
Gambar 4.17 Hasil <i>Learning Curve</i> Algoritma <i>Borderline</i> SMOTE <i>Dataset</i> Aspek.....	86
Gambar 4.18 Hasil <i>Confusion Matrix</i> <i>Dataset</i> Sentimen dengan SMOTE.....	88
Gambar 4.19 Hasil <i>Confusion Matrix</i> <i>Dataset</i> Sentimen dengan <i>Random</i>	

<i>Oversampling</i>	89
Gambar 4.20 Hasil <i>Confusion Matrix Dataset</i> Sentimen dengan ADASYN.....	89
Gambar 4.21 Hasil <i>Confusion Matrix Dataset</i> Sentimen dengan SMOTE-NC.....	90
Gambar 4.22 Hasil <i>Confusion Matrix Dataset</i> Sentimen dengan <i>Borderline</i> SMOTE.....	91
Gambar 4.23 Hasil <i>Confusion Matrix Dataset</i> Aspek dengan SMOTE.....	92
Gambar 4.24 Hasil <i>Confusion Matrix Dataset</i> Aspek dengan <i>Random</i> <i>Oversampling</i>	93
Gambar 4.25 Hasil <i>Confusion Matrix Dataset</i> Aspek dengan ADASYN.....	94
Gambar 4.26 Hasil <i>Confusion Matrix Dataset</i> Aspek dengan SMOTE-NC.....	95
Gambar 4.27 Hasil <i>Confusion Matrix Dataset</i> Aspek dengan <i>Borderline</i> SMOTE.....	96

DAFTAR LAMPIRAN

LAMPIRAN.....	104
LAMPIRAN 1. Hasil wawancara.....	104