

TUGAS AKHIR

PELABELAN SENTIMEN BERBASIS
SEMI-SUPERVISED LEARNING MENGGUNAKAN
DEEP LEARNING



PUJI AYUNINGTYAS

20102122

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO

2024

TUGAS AKHIR

**PELABELAN SENTIMEN BERBASIS
SEMI-SUPERVISED LEARNING MENGGUNAKAN
DEEP LEARNING**

***SENTIMENT LABELING BASED ON SEMI-
SUPERVISED LEARNING USING DEEP LEARNING***

Disusun Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Komputer



PUJI AYUNINGTYAS

20102122

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO
2024**

LEMBAR PERSETUJUAN PEMBIMBING

**PELABELAN SENTIMEN BERBASIS
SEMI-SUPERVISED LEARNING MENGGUNAKAN
DEEP LEARNING**

***SENTIMENT LABELING BASED ON SEMI-SUPERVISED
LEARNING USING DEEP LEARNING***

Dipersiapkan dan Disusun Oleh

PUJI AYUNINGTYAS

20102122

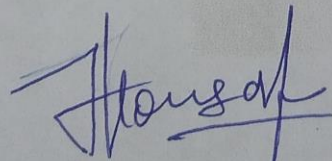
Fakultas Informatika

Institut Teknologi Telkom Purwokerto

Pada Tanggal 21 Mei 2024

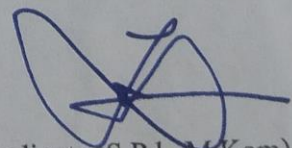
Pembimbing Utama,

Pembimbing Pendamping,



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

NIDN. 0517108101



(Sudianto, S.Pd., M.Kom)

NIDN. 0605049301

HALAMAN PENGESAHAN PENGUJI

**PELABELAN SENTIMEN BERBASIS SEMI
SUPERVISED LEARNING MENGGUNAKAN
DEEP LEARNING**

***SENTIMENT LABELING BASED ON SEMI-
SUPERVISED LEARNING USING DEEP LEARNING***

Dipersiapkan dan Disusun oleh

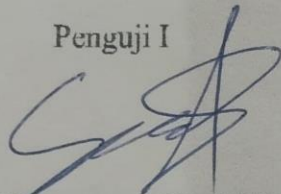
Puji Ayuningtyas

20102122

Telah diujikan dan dipertahankan dalam Sidang Tugas Akhir

Pada tanggal 30 Mei 2024

Penguji I



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

NIDN. 0613109201

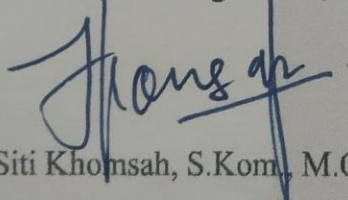
Penguji II



(Annisaa Utami, S.Kom., M.Sc.)

NIDN. 0607079403

Pembimbing Utama,



(Siti Khoimah, S.Kom., M.Cs.)

NIDN. 0517108101

Pembimbing Pendamping,

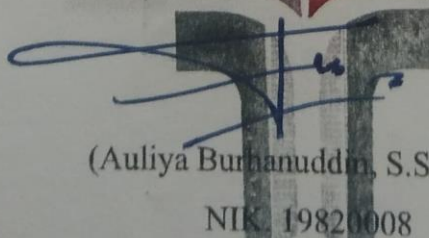


(Sudianto, S.Pd., M.Kom)

NIDN. 0605049301

Mengetahui,

Dekan Fakultas Informatika,



(Auliya Burhanuddin, S.Si., M.Kom.)

NIK. 19820008

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertandatangan di bawah ini,

Nama mahasiswa : Puji Ayuningtyas
NIM : 20102122
Program Studi : Teknik Informatika

Menyatakan bahwa Tugas Akhir dengan judul:

**PELABELAN SENTIMEN BERBASIS SEMI-SUPERVISED LEARNING
MENGUNAKAN DEEP LEARNING**

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

Dosen Pembimbing Pendamping : Sudianto, S.Pd., M.Kom.

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, 20 Mei 2024

Yang menyatakan,



(Puji Ayuningtyas)

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT, atas rahmat dan ridhonya, penulis diberikan kesabaran, kekuatan, dan kelancaran sehingga penulis mampu menyelesaikan tugas akhir ini dengan baik. Tugas akhir ini disusun dalam rangka memenuhi syarat untuk memperoleh gelar sarjana pada Program Studi S1 Teknik Informatika Institut Teknologi Telkom Purwokerto. Penulis menyadari tanpa adanya dukungan moral dan material dari berbagai pihak, terasa sulit bagi penulis untuk dapat menyelesaikan penyusunan tugas akhir ini. Dengan ini, penulis menyampaikan terima kasih kepada:

1. Dr. Tenia Wahyuningrum, S.Kom., M.T. selaku Rektor Institut Teknologi Telkom Purwokerto.
2. Auliya Burhanuddin, S.Si., M.Kom. selaku Dekan Fakultas Informatika.
3. Amalia Beladinna Arifa, S.Pd., M.Cs. selaku Ketua Program Studi S1 Teknik Informatika.
4. Ayah dan Ibu tercinta yang telah memberikan semangat, materi, dan doa yang tidak terputus.
5. Siti Khomsah, S.Kom., M.Cs. selaku dosen pembimbing utama yang telah membimbing penulis menyelesaikan penyusunan naskah tugas akhir dengan baik.
6. Sudianto, S.Pd., M.Kom. selaku Dosen Pembimbing Pendamping yang dengan sabar membimbing dan membantu dalam penyelesaian penyusunan naskah tugas akhir dengan baik.
7. Muhamad Azrino Gustalika, S.Kom., M.Tr.T. selaku dosen wali kelas S1IF-08D, yang memberikan dukungan penuh dan motivasi kepada penulis selama masa perkuliahan, terutama ketika masa penyelesaian tugas akhir.
8. Diri sendiri yang telah berhasil bertahan untuk menuntaskan penyusunan tugas akhir dengan maksimal.
9. Teman-teman terbaik yang sangat berjasa membantu penulis selama masa perkuliahan termasuk ketika masa penyusunan tugas akhir,

Nur Aini L.A., Shofiyana Dhela N.A., Mahda Laina A., Aulia Romadloni N.I., M. Junaedi, Herlinda Sundari, Rizky Ade S., Rayhan Hafidh W., dan Hanugrah Surya Purwaka.

10. Teman-teman dan pihak lain yang membantu penulis berproses selama menempuh perkuliahan yang tidak dapat penulis sebutkan satu persatu.

Penulis sepenuhnya menyadari bahwa tugas akhir ini belum sempurna karena keterbatasan pengalaman dan pengetahuan penulis. Oleh karena itu, kritik dan saran akan sangat membantu untuk mengembangkan kemampuan penulis dalam menyusun hasil penelitian di kemudian hari.

Purwokerto, 30 April 2024

A handwritten signature in black ink, consisting of stylized, overlapping letters and a horizontal line at the bottom.

Puji Ayuningtyas

DAFTAR ISI

TUGAS AKHIR	ii
LEMBAR PERSETUJUAN PEMBIMBING.....	iii
HALAMAN PENGESAHAN PENGUJI.....	iv
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL	xiii
ABSTRAK.....	xiv
ABSTRACT	xv
BAB I.....	2
PENDAHULUAN	2
1.1 Latar Belakang	2
1.2 Rumusan Masalah	3
1.3 Pertanyaan Penelitian	3
1.4 Batasan Masalah.....	3
1.5 Tujuan Penelitian.....	4
1.6 Manfaat Penelitian.....	4
BAB II	5
TINJAUAN PUSTAKA	5
2.1 Penelitian Terdahulu	5
2.2 Landasan Teori	19
2.2.1 <i>Natural Language Processing</i>	19

2.2.2	Analisis Sentimen	22
2.2.3	<i>Semi-Supervised Learning</i>	23
2.2.4	<i>Deep Learning</i>	24
2.2.5	Algoritma <i>Long Short Term Memory</i> (LSTM)	25
2.2.6	Algoritma <i>Gated Recurrent Unit</i> (GRU)	27
2.2.7	VADER Lexicon	29
2.2.8	Confusion Matrix	30
BAB III		32
METODOLOGI PENELITIAN		32
3.1	Subjek dan Objek Penelitian	32
3.2	Bahan dan Alat Penelitian	32
3.3	Diagram Alir Penelitian	32
3.3.1.	Identifikasi Masalah	33
3.3.2.	Studi Literatur	33
3.3.3.	Pengumpulan Data	33
3.3.4.	Perancangan dan Evaluasi Model	36
BAB IV		49
HASIL DAN PEMBAHASAN		49
4.1.	Pemodelan Algoritma	49
4.1.1.	Skema <i>0.002-GRU-50-128</i>	49
4.1.2.	Skema <i>0.001-GRU-50-128</i>	49
4.1.3.	Skema <i>0.001-LSTM-50-512</i>	50
4.1.4.	Skema <i>0.002-LSTM-50-512</i>	51
4.1.5.	Evaluasi Model Algoritma	51
4.2.	Testing Unlabeled Data	56

4.3. Hasil Analisis.....	60
BAB V	62
KESIMPULAN DAN SARAN	62
5.1 Kesimpulan.....	62
5.2 Saran	62
DAFTAR PUSTAKA	64
LAMPIRAN	72

DAFTAR GAMBAR

Gambar 2. 1 Arsitektur CBoW [43].....	21
Gambar 2. 2 Alur <i>Semi-Supervised Learning</i> [50]	23
Gambar 2. 3 Arsitektur <i>Deep Learning</i>	24
Gambar 2. 4 Struktur algoritma LSTM [58].....	26
Gambar 2. 5 Arsitektur Algoritma Gated Recurrent Unit (GRU).....	28
Gambar 3. 1 Rancangan Alur Penelitian.....	33
Gambar 3. 2 Perancangan Model Penelitian.....	36
Gambar 3. 3 Alur Pemrosesan awal Data	36
Gambar 3. 4 <i>Source code</i> menghapus duplikat data.....	37
Gambar 3. 5 <i>Source code</i> menghapus data kosong	37
Gambar 3. 6 Proses <i>Under Sampling</i>	40
Gambar 3. 7 Banyak data OVO sebelum <i>Under Sampling</i>	40
Gambar 3. 8 Banyak data OVO setelah <i>Under Sampling</i>	41
Gambar 3. 9 Banyak data DANA sebelum <i>Under Sampling</i>	41
Gambar 3. 10 Banyak data DANA setelah <i>Under Sampling</i>	41
Gambar 3. 11 Banyak data OVO setelah <i>labelling</i>	42
Gambar 3. 12 Banyak data DANA setelah <i>labeling</i>	42
Gambar 3. 13 <i>Source code</i> penggabungan data.....	43
Gambar 3. 14 Kuantitas Data pada Setiap Selas.....	44
Gambar 3. 15 Hasil Pengambilan Data Uji	44
Gambar 3. 16 Alur Metode <i>Semi-Supervised Learning</i>	45
Gambar 4. 1 Performa Skema 0.002-GRU-50-128	49
Gambar 4. 2 Performa Skema 0.001-GRU-50-128	50
Gambar 4. 3 Performa Skema 0.001-LSTM-50-512	50
Gambar 4. 4 Performa Skema 0.002-LSTM-50-512	51
Gambar 4. 5 Confusion Matrix Pelatihan Algoritma.....	52
Gambar 4. 6 Akurasi Pelatihan Model.....	52
Gambar 4. 7 <i>Precision</i> Pelatihan Model.....	53
Gambar 4. 8 <i>Recall</i> Pelatihan Model.....	54
Gambar 4. 9 F1-Score Pelatihan Model	54

Gambar 4. 10 Rata-Rata Hasil Evaluasi Model Algoritma.....	56
Gambar 4. 11 Rata-Rata Hasil Evaluasi Pelabelan (SSL) Algoritma	58
Gambar 4. 12 Hasil Nilai <i>Confidence Model</i>	58
Gambar 4. 13 Perbandingan Hasil Pelabelan <i>Deep Learning</i>	59
Gambar 4. 14 Hasil pelabelan menggunakan metode VADER	59
Gambar 4. 15 <i>Classification Report</i> Pelabelan Metode VADER	60

DAFTAR TABEL

Tabel 2. 1 Penelitian Sebelumnya.....	10
Tabel 2. 2 Confusion Matrix	30
Tabel 3. 1 Dataset pertama (DANA)	35
Tabel 3. 2 Dataset Kedua (OVO).....	35
Tabel 3. 3 Dataset Ketiga (GoPay)	35
Tabel 3. 4 Perbandingan Banyak Data.....	37
Tabel 3. 5 Hasil <i>cleaning data</i>	38
Tabel 3. 6 Hasil <i>case folding</i>	38
Tabel 3. 7 Hasil Tokenisasi	38
Tabel 3. 8 Hasil <i>slangword removal</i>	39
Tabel 3. 9 Hasil penggabungan dataset.....	43
Tabel 3. 10 Contoh Hasil Word Embedding	45
Tabel 3. 11 Susunan Layer Algoritma <i>Deep Learning</i>	46
Tabel 3. 12 Nilai Hyperparameter Model <i>Deep Learning</i>	46
Tabel 3. 13 Skenario Pemodelan Algoritma LSTM.....	47
Tabel 3. 14 Skenario Pemodelan Algoritma GRU	47
Tabel 4. 1 Tabel Hasil Evaluasi Algoritma LSTM.....	55
Tabel 4. 2 Tabel Hasil Evaluasi Algoritma GRU	55
Tabel 4. 3 Tabel Hasil Percobaan Pelabelan (SSL) Algoritma LSTM.....	56
Tabel 4. 4 Tabel Hasil Percobaan Pelabelan (SSL) Algoritma GRU.....	57
Tabel 4. 5 Contoh Ketidaksesuaian antara Komentar dan Label.....	60