

**ANALISIS SENTIMEN PENGGUNA YOUTUBE MENGENAI
ANALOG SWITCH OFF (ASO) MENGGUNAKAN WORD
EMBEDDING DAN METODE LONG SHORT-TERM MEMORY
(LSTM)**

SKRIPSI

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



Disusun Oleh:

MOCHAMAD SUHRI AINUR RIFKY

19082010041

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"

JAWA TIMUR

SURABAYA

2023

LEMBAR PENGESAHAN

ANALISIS SENTIMEN PENGGUNA YOUTUBE MENGENAI *ANALOG SWITCH OFF (ASO)* MENGGUNAKAN *WORD EMBEDDING* DAN METODE *LONG SHORT-TERM MEMORY (LSTM)*

Disusun Oleh:


Mochamad Suhri Ainur Rifky
19082010041


**Telah disetujui mengikuti Ujian Negara Lisan Gelombang Juli Periode
2023 pada Tanggal 20 Juli 2023**

Menyetujui,

Dosen Pembimbing 1


Dosen Pembimbing 2


Amalia Anjani A., S.Kom., M.Kom.
NIP. 19920817 201803 2 001


Reisa Permatasari, S.T., M.Kom.
NIP. 19920514 202203 2 007

Mengetahui,

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


Agung Brastama Putra, S.Kom., M.Kom.
NIP. 19851124 202121 1 003



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

KETERANGAN REVISI

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

Nama : Mochamad Suhri Ainur Rifky
NPM : 19082010041
Program Studi : Sistem Informasi


Telah mengerjakan revisi Ujian Negara Lisan Skripsi pada tanggal 20 Juli dengan judul:

ANALISIS SENTIMEN PENGGUNA YOUTUBE MENGENAI ANALOG SWITCH OFF (ASO) MENGGUNAKAN WORD EMBEDDING DAN METODE LONG SHORT-TERM MEMORY (LSTM)


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

Surabaya, 20 Juli 2023

Dosen penguji yang memeriksa revisi:

1. Agung Brastama P., S.Kom., M.Kom. {  }
NIP. 19851124 2021211 003

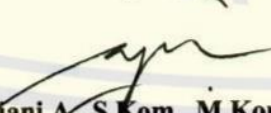
2. Siti Mukaromah, S.Kom., M.Kom. {  }
NIP. 19810704 2021212 011

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

Mengetahui,

Dosen Pembimbing 1

Dosen Pembimbing 2


Amalia Arjani A., S.Kom., M.Kom.
NIP. 19920812 2018032 001


Reisa Permatasari, S.T., M.Kom.
NIP. 19920514 2022032 007

SKRIPSI

ANALISIS SENTIMEN PENGGUNA YOUTUBE MENGENAI *ANALOG SWITCH OFF (ASO)* MENGGUNAKAN *WORD EMBEDDING* DAN METODE *LONG SHORT-TERM MEMORY (LSTM)*

Disusun Oleh:

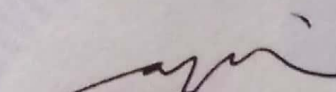
MOCHAMAD SUHLAINUR RIEKY

19082010041

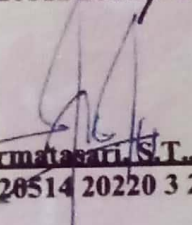
Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Skripsi Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Pembangunan Nasional "Veteran" Jawa Timur Pada Tanggal 20 Juli 2023

Pembimbing:

1.

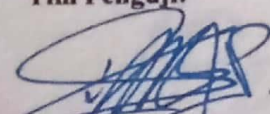

Amalia Anjani A., S.Kom., M.Kom.
NIP. 19920812 201803 2 001

2.

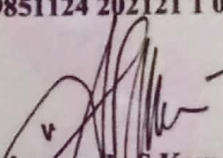

Reisa Permatasari, S.T., M.Kom.
NIP. 19920514 20220 3 2007

Tim Penguji:

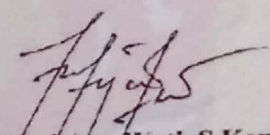
1.


Agung Brastama P., S.Kom., M.Kom.
NIP. 19851124 202121 1 003

2.


Siti Mukaromah, S.Kom., M.Kom.
NIP. 19810704 202121 2 011

3.


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

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


Dr. Ir. Nevirina Hendrasarie, MT.
NIP. 19681126 199403 2 001



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

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Mochamad Suhri Ainur Rifky

NPM : 19082010041

Program Studi : Sistem Informasi

Menyatakan bahwa Judul Skripsi / Tugas Akhir sebagai berikut:

ANALISIS SENTIMEN PENGGUNA YOUTUBE MENGENAI *ANALOG SWITCH OFF (ASO)* MENGGUNAKAN *WORD EMBEDDING* DAN METODE *LONG SHORT-TERM MEMORY (LSTM)*

Bukan merupakan plagiat dari Skripsi / Tugas Akhir / Penelitian orang lain dan juga bukan merupakan Produk / Hasil Karya yang saya beli dari orang 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 Universitas Pembangunan Nasional "Veteran" Jawa Timur maupun di Institusi Pendidikan lain. Jika ternyata dikemudian hari pernyataan terbukti benar, maka Saya bertanggung jawab penuh dan siap menerima segala konsekuensi, termasuk pembatalan ijazah dikemudian hari.

Surabaya, 20 Juli 2023

Format Saya,



Mochamad Suhri Ainur Rifky

19082010041

Judul : ANALISIS SENTIMEN PENGGUNA YOUTUBE
MENGENAI ANALOG SWITCH OFF (ASO)
MENGUNAKAN WORD EMBEDDING DAN
METODE LONG SHORT-TERM MEMORY (LSTM)

Pembimbing 1 : Amalia Anjani Arifiyanti, S.Kom, M.Kom.

Pembimbing 2 : Reisa Permatasari, S.T, M.Kom.

ABSTRAK

Analog Switch Off (ASO) atau program migrasi dari televisi analog ke televisi digital merupakan program yang dikeluarkan oleh Kementerian Komunikasi dan Informatika di Indonesia. Beberapa masyarakat memberikan tanggapan dan opini yang berbeda-beda di komentar YouTube mengenai ASO. Ada yang memberikan komentar bersetimen positif atau netral. Namun, ada juga yang memberikan komentar bersetimen negatif. Analisis sentimen adalah proses yang dilakukan secara otomatis dalam mempelajari, mengambil, dan mengolah data tekstual untuk memperoleh informasi dan melihat tanggapan atau opini mengenai sebuah masalah atau objek menuju opini positif, netral atau negatif. dengan demikian analisis sentimen dapat dijadikan acuan dalam pengambilan keputusan suatu organisasi, meningkat suatu pelayanan, maupun sebagai *review* suatu produk. Analisis sentimen dilakukan menggunakan *word embedding* dengan Word2Vec, serta pengklasifikasian sentimen menggunakan metode *Long Short-Term Memory* (LSTM). Hasil uji validasi dilakukan menggunakan data baru sebanyak 168, didapatkan 63 sentimen negatif dari 70 data yang bernilai benar, 52 sentimen netral dari 57 data bernilai benar, dan 59 sentimen positif dari 71 data yang bernilai benar. Sehingga presentase *accuracy* model adalah 87.78%.

Kata kunci: Analisis Sentimen, *Word Embedding*, *Long Short-Term Memory*.

KATA PENGANTAR

Segala puji syukur penulis panjatkan atas kehadiran Allah SWT karena atas berkat dan hidayah dari Allah SWT penulis mampu menyelesaikan skripsi yang berjudul "Analisis Sentimen Pengguna Youtube Mengenai *Analog Switch Off* (ASO) Menggunakan *Word Embedding* Dan Metode *Long Short-Term Memory* (LSTM)" sebagai syarat untuk menyelesaikan program studi Sistem Informasi di Universitas Pembangunan Nasional "Veteran" Jawa Timur.

Dalam upaya menyelesaikan skripsi ini, tidak dapat dilepaskan dari dukungan yang diterima dari berbagai pihak, dan dalam kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada:

1. Kedua orang tua yang senantiasa doa dan restunya tidak pernah terputus, sehingga penulis terus termotivasi dalam menyelesaikan skripsi ini.
2. Bapak Nur Cahyo Wibowo, S.Kom, M.Kom. selaku dosen wali yang selalu memberikan motivasi dan arahan selama menjalani pendidikan di Sistem Informasi UPN "Veteran" Jawa Timur.
3. Bu Amalia Anjani Arifiyanti, S.Kom, M.Kom dan Ibu Reisa Permatasari, S.T, M.Kom. selaku dosen pembimbing skripsi yang telah membimbing dengan sabar, memberikan ilmu berharga serta motivasi yang sangat bermanfaat sehingga penulis dapat menyelesaikan skripsi ini.
4. Seluruh bapak dan ibu dosen program studi Sistem Informasi UPN "Veteran" Jawa Timur yang telah memberikan banyak ilmu pengetahuan yang bermanfaat selama masa perkuliahan;

5. Terima kasih kepada Eric, Jobas, Hilman, Arsyah, Catur, dan orang-orang baik yang senantiasa meluangkan waktunya untuk membantu saya dalam pengerjaan skripsi ini.
6. Terima kasih kepada Sistem Informasi Angkatan 2019 (SOLASIFO) yang telah memberikan dukungan dan berjuang bersama mulai dari masa orientasi maba hingga saat ini.
7. Terima kasih kepada Razan, Rofi, Robert, Meutya, dan Catherine yang selalu memberikan dukungan, motivasi, hiburan dan candaan “ikan kebahagiaan” sehingga saya selalu bahagia ketika mengerjakan skripsi ini.
8. Terima kasih kepada Google Developers dan rekan kerja di Jobhun, yang telah mendorong dan memberi motivasi untuk penyelesaian skripsi ini.
9. Terima kasih kepada mentor Bangkit dan cohort CC-69 yang menemani saya tiap minggu selama pengerjaan skripsi ini.
10. Terima kasih kepada seluruh teman-teman yang selalu mendukung dan memberikan masukan dan semangat hingga saat ini.

Semoga Allah SWT membalas kebaikan pihak-pihak diatas. Dalam pengerjaan skripsi ini penulis menyadari masih banyak sekali kekurangan. Semoga skripsi ini bermanfaat bagi semua pihak.

Surabaya, Juli 2023

Penulis

DAFTAR ISI

ABSTRAK	i
KATA PENGANTAR	ii
DAFTAR ISI.....	iv
DAFTAR TABEL.....	vii
DAFTAR GAMBAR	viii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	5
1.4 Tujuan.....	5
1.5 Sistematika Penulisan.....	6
BAB II TINJAUAN PUSTAKA.....	8
2.1 Dasar Teori	8
2.1.1 YouTube.....	8
2.1.2 <i>Text Mining</i>	8
2.1.3 Analisis Sentimen.....	9
2.1.4 <i>Preprocessing Text</i>	9
2.1.5 Vader Lexicon.....	10
2.1.6 TextBlob.....	11
2.1.7 <i>Word Embedding</i>	11
2.1.8 Word2vec	12
2.1.9 <i>Long - Short Term Memory (LSTM)</i>	12
2.1.10 <i>Confusion Matrix</i>	16
2.1.11 Pandas	16
2.1.12 Keras	16
2.1.13 Flask	17
2.2 Penelitian Terdahulu.....	17
BAB III METODOLOGI PENELITIAN.....	20
3.1 Studi Literatur.....	20
3.2 Analisis Kebutuhan Sistem.....	21

3.2.1	Kebutuhan Data.....	21
3.2.2	Kebutuhan <i>Software</i> dan <i>Hardware</i>	21
3.3	Pengumpulan Data Komentar.....	22
3.4	Pelabelan Data.....	22
3.5	Pembagian Data.....	22
3.6	Perancangan Model.....	22
3.3.1.	<i>Preprocessing</i> Data.....	23
3.3.2.	Pembagian Dataset untuk Training dan Testing Model LSTM.....	25
3.3.3.	<i>Word Embedding</i> dengan Word2vec.....	25
3.3.4.	Encoding menggunakan Keras Tokenizer.....	25
3.3.5.	Klasifikasi Sentimen Model LSTM.....	25
3.3.6.	Evaluasi Model.....	26
3.3.7.	Penyimpanan Model.....	26
3.7	Penerapan Model.....	26
3.8	Validasi Model.....	26
3.9	Pembahasan.....	27
BAB IV HASIL DAN PEMBAHASAN.....		28
4.1	Implementasi Kebutuhan.....	28
4.2.1	Kebutuhan Data.....	28
4.2.2	Kebutuhan <i>Hardware</i> dan <i>Software</i>	28
4.2	Pengumpulan Data Komentar.....	29
4.3	Pelabelan Data.....	31
4.4	Pembagian Data.....	39
4.5	Perancangan Model.....	40
4.5.1	<i>Preprocessing</i> Data.....	40
4.5.2	Pembagian Dataset untuk Training dan Testing Model LSTM.....	48
4.5.3	<i>Word Embedding</i> dengan Word2vec.....	49
4.5.4	<i>Encoding</i> menggunakan Keras <i>Tokenizer</i>	50
4.5.5	Klasifikasi Sentimen Model LSTM.....	51
4.5.6	Evaluasi Model.....	59
4.5.7	Simpan Model.....	65
4.6	Penerapan Model.....	66

4.7	Validasi Model	67
4.8	Pembahasan	68
BAB V KESIMPULAN DAN SARAN.....		75
5.1	Kesimpulan.....	75
5.2	Saran.....	75
.DAFTAR PUSTAKA		76
LAMPIRAN		81

DAFTAR TABEL

Tabel 4.1 Contoh Hasil Pengumpulan Data.....	30
Tabel 4.2 Data Setelah Dilabeli Vader Lexicon dan TextBlob.....	33
Tabel 4.3 Data Setelah Dilabeli Vader Lexicon, TextBlob, dan Manual	35
Tabel 4.4 Data <i>Final</i> Setelah Dilabeli.....	38
Tabel 4.5 Hasil Proses <i>Cleaning</i>	42
Tabel 4.6 Hasil Proses <i>Case Folding</i>	43
Tabel 4.7 Hasil Proses <i>Stopword Removal</i>	45
Tabel 4.8 Hasil Proses <i>Stemming</i>	47

DAFTAR GAMBAR

Gambar 2.1 Arsitektur LSTM	13
Gambar 3.1 Alur Metode Penelitian	20
Gambar 3.2 Perancangan Model	23
Gambar 4.1 <i>Source Code</i> Proses Pengumpulan Data	29
Gambar 4.2 <i>Source Code</i> Menerjemahkan Data.....	31
Gambar 4.3 <i>Source Code</i> Pelabelan Vader Lexicon.....	32
Gambar 4.4 <i>Source Code</i> TextBlob	32
Gambar 4.5 <i>Source Code</i> Menyimpan Hasil Pelabelan.....	33
Gambar 4.6 <i>Source Code</i> Menyimpan <i>Final Data</i>	38
Gambar 4.7 Total Data Tiap Label	39
Gambar 4.8 <i>Source Code</i> Pembagian Data	40
Gambar 4.9 <i>Source Code</i> Proses <i>Cleaning</i>	41
Gambar 4.10 <i>Source Code</i> Proses <i>Case Folding</i>	43
Gambar 4. 11 <i>Source Code</i> Proses <i>Stopword Removal</i>	45
Gambar 4.12 <i>Source Code</i> Proses <i>Stemming</i>	47
Gambar 4.13 <i>Source Code</i> Proses Pembagian Dataset	48
Gambar 4.14 <i>Source Code</i> Proses Pembuatan Word2Vec.....	49
Gambar 4.15 <i>Source Code</i> Proses <i>Word Embedding</i>	50
Gambar 4.16 <i>Source Code</i> Proses <i>Encodig</i>	50
Gambar 4.17 <i>Source Code</i> Proses Pembuatan Model LSTM	51
Gambar 4.18 <i>Source Code</i> Proses <i>Compile</i> Model LSTM	52
Gambar 4.19 <i>Source Code</i> Proses <i>Training</i>	53
Gambar 4.20 Persebaran Label Dataset <i>Training</i>	53

Gambar 4.21 <i>Training and Validation Accuracy</i> Skenario 1.....	54
Gambar 4.22 <i>Training and Validation Loss</i> Skenario 1	55
Gambar 4.23 <i>Source Code</i> Proses <i>Oversampling</i>	56
Gambar 4.24 <i>Training and Validation Accuracy</i> Skenario 2.....	57
Gambar 4.25 <i>Training and Validation Loss</i> Skenario 2	57
Gambar 4.26 <i>Source Code</i> dan Hasil Evaluasi Skenario 1	59
Gambar 4.27 <i>Source Code</i> dan Hasil Evaluasi Skenario 1	62
Gambar 4.28 <i>Source Code</i> Proses Menyimpan Model	65
Gambar 4.29 <i>Source Code</i> Penerapan Model	66
Gambar 4.30 Hasil Pengembangan <i>Website</i>	67
Gambar 4.31 Proses Validasi Model.....	68
Gambar 4.32 Diagram Persebaran Data per Sentimen.....	69
Gambar 4.33 Grafik Persebaran Data per Minggu.....	69
Gambar 4.34 Top 25 Kata Sentimen Negatif.....	70
Gambar 4.35 Top 25 Kata Sentimen Netral.....	70
Gambar 4.36 Top 25 Kata Sentimen Positif	71
Gambar 4.37 <i>Word Cloud</i> Sentimen Negatif	71
Gambar 4.38 <i>Word Cloud</i> Sentimen Netral	72
Gambar 4.39 <i>Word Cloud</i> Sentimen Positif.....	72