

TUGAS AKHIR

**ANALISIS SENTIMEN BERBASIS ASPEK PADA
ULASAN APLIKASI VIDIO DI GOOGLE PLAY STORE
MENGUNAKAN INDOBERT DAN LDA**



RAAFI ALHADI

20110025

**PROGRAM STUDI S1 SAINS DATA
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO
2024**

TUGAS AKHIR

**ANALISIS SENTIMEN BERBASIS ASPEK PADA
ULASAN APLIKASI VIDIO DI GOOGLE PLAY STORE
MENGUNAKAN INDOBERT DAN LDA**

**ASPECT-BASED SENTIMENT ANALYSIS ON VIDIO
APP REVIEWS ON GOOGLE PLAY STORE USING
INDOBERT AND LDA**

Disusun Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Sains Data



RAAFI ALHADI

20110025

**PROGRAM STUDI S1 SAINS DATA
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO**

2024

HALAMAN PERSETUJUAN PEMBIMBING

**ANALISIS SENTIMEN BERBASIS ASPEK PADA
ULASAN APLIKASI VIDIO DI GOOGLE PLAY STORE
MENGUNAKAN INDOBERT DAN LDA**

**ASPECT-BASED SENTIMENT ANALYSIS ON VIDIO
APP REVIEWS ON GOOGLE PLAY STORE USING
INDOBERT AND LDA**

Dipersiapkan dan Disusun Oleh:

RAAFI ALHADI

20110025

Fakultas Informatika

Institut Teknologi Telkom Purwokerto

Pada Tanggal: 10 Juni 2024

Pembimbing Utama,



Siti Khomsah, S.Kom., M.Cs.
NIDN. 0517108101

HALAMAN PENGESAHAN TUGAS AKHIR

**ANALISIS SENTIMEN BERBASIS ASPEK PADA
ULASAN APLIKASI VIDIO DI GOOGLE PLAY STORE
MENGUNAKAN INDOBERT DAN LDA**

**ASPECT-BASED SENTIMENT ANALYSIS ON VIDIO
APP REVIEWS ON GOOGLE PLAY STORE USING
INDOBERT AND LDA**

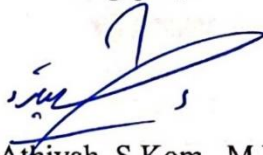
Disusun Oleh:

RAAFI ALHADI

20110025

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tugas Akhir pada Hari
Senin, Tanggal 19 Juni 2024

Penguji 1,



Ummi Athiyah, S.Kom., M.Kom.
NIDN. 0621129001

Penguji 2,



Atika Ratna Dewi, S.Si., M.Sc.
NIDN. 0615128703

Pembimbing Utama,



Siti Khomsah, S.Kom, M.Cs.
NIDN. 0517108101

Dekan,



Auliya Burhanuddin, S.Si., M.Kom.
NIK. 19120008

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertandatangan di bawah ini,

Nama Mahasiswa : Raafi Alhadi

NIM : 20110025

Program Studi : S1 Sains Data

Menyatakan bahwa Tugas Akhir dengan judul berikut:

ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN APLIKASI VIDIO DI GOOGLE PLAY STORE MENGGUNAKAN INDOBERT DAN LDA

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 tanggung jawab Saya, bukan tanggung jawab 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, 10 Juni 2024

Yang Menyatakan,



Raafi Alhadi

DAFTAR ISI

HALAMAN PERSETUJUAN PEMBIMBING	iii
HALAMAN PENGESAHAN TUGAS AKHIR.....	iv
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
ABSTRAK.....	xii
<i>ABSTRACT</i>	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Pertanyaan Penelitian	3
1.4 Batasan Masalah.....	4
1.5 Tujuan Penelitian.....	4
1.6 Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Kajian Pustaka.....	5
2.2 Dasar Teori	22
2.2.1 <i>Opinion Mining</i>	22
2.2.2 Analisis Sentimen	23
2.2.3 <i>Exploratory Data Analysis (EDA)</i>	23
2.2.4 <i>Aspect Based Sentiment Analysis (ABSA)</i>	24
2.2.5 <i>Bidirectional Encoder Representations from Transformers (BERT)</i> ...	27
2.2.6 <i>IndoBERT</i>	29
2.2.7 <i>Latent Dirichlet Allocation (LDA)</i>	31
BAB III METODOLOGI PENELITIAN.....	35
3.1 Subjek dan Objek Penelitian	35
3.2 Alat dan Bahan Penelitian	35

3.2.1	Alat Penelitian	35
3.2.2	Bahan Penelitian.....	35
3.3	Diagram Alir Penelitian	36
3.3.1	Tahap Pemahaman Masalah.....	36
3.3.2	Tahap Pengumpulan Data	37
3.3.3	Pra-pemrosesan Data.....	38
3.3.4	Tahap <i>Exploratory Data Analysis (EDA)</i>	44
3.3.5	Tahap Ekstraksi Aspek.....	46
3.3.6	Tahap Interpretasi Hasil	47
BAB IV HASIL DAN PEMBAHASAN		49
4.1	<i>Data Preparation</i>	49
4.2	<i>Data Labeling</i>	56
4.3	<i>Exploratory Data Analysis (EDA)</i>	56
4.4	Ekstraksi Aspek.....	60
BAB V KESIMPULAN DAN SARAN.....		65
5.1	Kesimpulan.....	65
5.2	Saran.....	65
DAFTAR PUSTAKA		67
LAMPIRAN.....		75

DAFTAR TABEL

Tabel 2.1 Studi Literatur	11
Tabel 3.1 Spesifikasi Variabel Dataset Penelitian	38
Tabel 3.2 Contoh <i>Case Folding</i>	40
Tabel 3.3 Contoh <i>Removing Punctuation</i>	40
Tabel 3.4 Contoh <i>Removing Whitespace</i>	40
Tabel 3.5 Contoh <i>Handling Emoji and Emoticon</i>	41
Tabel 3.6 Contoh <i>Normalizing Slang and Misspelled</i>	41
Tabel 3.7 Contoh <i>Removing Stopword</i>	41
Tabel 3.8 Contoh Perubahan menjadi Token <i>IndoBERT</i>	43
Tabel 3.9 Contoh Penambahan Token [CLS] dan [SEP]	43
Tabel 3.10 Contoh Penyesuaian Panjang Kalimat	43
Tabel 3.11 Contoh Pemisahan <i>Subword</i>	43
Tabel 3.12 Contoh Penentuan <i>Attention Mask</i>	44
Tabel 3.13 Perbandingan Polaritas Opini Pengguna Aplikasi Vidio	45
Tabel 4.1 Pengecekan Data per Bulan	49
Tabel 4.2 Hasil Pengumpulan Data Aplikasi Vidio di <i>Google Play Store</i>	51
Tabel 4.3 Hasil Pemilihan Fitur	52
Tabel 4.4 Hasil <i>Case Folding</i> dan <i>Removing Punctuation</i>	52
Tabel 4.5 Hasil Normalisasi Kata Slang dan Salah Eja	53
Tabel 4.6 Hasil Penghapusan <i>Stopword</i>	54
Tabel 4.7 Hasil Pelabelan Data	56
Tabel 4.8 Kata dengan Frekuensi Terbanyak yang Sering Muncul per Sentimen	58

DAFTAR GAMBAR

Gambar 2.1 Tingkatan <i>Opinion Mining</i>	22
Gambar 2.3 Arsitektur <i>Transformer</i> [33].....	28
Gambar 2.4 Representasi Input <i>BERT</i> [31].....	29
Gambar 2.2 Hubungan Variabel dalam Proses Generatif <i>LDA</i>	32
Gambar 3.1 Diagram Alir Penelitian	36
Gambar 3.2 Seleksi Fitur.....	39
Gambar 3.3 Alur Pembersihan Data	39
Gambar 4.1 Distribusi Data Hasil Pembersihan	55
Gambar 4.2 Hasil Undersampling Data	55
Gambar 4.3 Trend Polaritas Sentimen Pengguna Setiap Tahun (2016-2022)	57
Gambar 4.4 Polaritas Sentimen Sentimen Pengguna Keseluruhan (2016-2022).....	58
Gambar 4.5 Panjang Ulasan.....	60