

SKRIPSI

**ANALISIS SENTIMEN APLIKASI IFLIX DI GOOGLE PLAY
STORE MENGGUNAKAN ALGORITMA INDOBERT**



Disusun oleh :

RYANDA PRATAMA PUTRA WIBISONO

20200140015

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH YOGYAKARTA**

2024

PERNYATAAN ORIGINALITAS

Saya yang bertanda tangan dibawah ini

Nama : Ryanda Pratama Putra Wibisono

NIM : 20200140015

Program Studi : Teknologi Informasi

Fakultas : Fakultas Teknik

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis benar-benar merupakan hasil karya saya sendiri dan belum diajukan dalam bentuk apapun kepada perguruan tinggi mana pun. Sumber Informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain tidak disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir Skripsi ini.

Apabila kemudian hari terbukti atau dapat dibuktikan Skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Yogyakarta, 05 oktober 2024

Yang membuat pernyataan



Ryanda Pratama Putra Wibisono

KATA PENGANTAR

Segala puji syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan rahmat dan karunia-Nya sehingga laporan Skripsi yang berjudul “Analisis Sentimen Aplikasi Iflix Di Google Play Store Menggunakan Algoritma IndoBERT” dapat terselesaikan dengan baik. Penulis menyadari dalam penyusunan laporan Skripsi ini masih banyak kekurangan dan tidak lepas dari bantuan orang lain. Penulis ingin mengucapkan terima kasih kepada yang terhormat :

1. Allah *Subhanahu Wa Ta'ala* yang telah memberikan petunjuk, kemurahan, keistiqomahan dan segala nikmat-Nya.
2. Bapak, Ibu dan semua keluarga besar saya yang telah memberikan doa, semangat, dan dukungan dalam penyelesaian skripsi ini.
3. Bapak Aris Widyo Nugroho, S.T., M.T., Ph.D. selaku Dekan Fakultas Teknik Universitas Muhammadiyah Yogyakarta yang telah memberikan kesempatan untuk menyusun skripsi ini.
4. Bapak Dr. Reza Giga Isnanda, S.T., M.Sc. selaku Ketua Program Studi Teknik Informasi Universitas Muhammadiyah Yogyakarta yang telah memberikan kesempatan untuk menyusun skripsi ini.
5. Bapak Ir. Asroni, S.T., M.Eng. dan Bapak Cahya Damarjati, S.T. M. Eng., Ph.D. selaku Dosen Pembimbing Skripsi yang telah membimbing, meluangkan waktu, dan memberi arahan sehingga laporan Skripsi dapat terselesaikan dengan baik.
6. Teman-teman angkatan 2020 khususnya, David Prastyawan, Muhammad Fahri Mubaroq, Prayoga Saputra, Ilham Fadhillah, Muhammad zulqoniun, Sandy Eka Yudha yang senantiasa memberikan dukungan, motivasi, dan bantuan selama menjalani perkuliahan.
7. Semua pihak yang memberikan bantuan baik mental maupun spiritual yang tidak bisa penulis sebutkan satu persatu.
8. Penulis menyadari sepenuhnya bahwa didalam pembuatan laporan Skripsi masih terdapat banyak kekurangan. Untuk itu saran dan kritik

yang membangun dari semua pihak sangat diharapkan oleh penulis.

Harapan penulis adalah semoga laporan Sk

9. ripsi ini dapat memberikan manfaat kepada pembaca.

Yogyakarta, 05 oktober 2024



Ryanda Pratama Putra Wibisono

DAFTAR ISI

HALAMAN PERSETUJUAN	i
HALAMAN PENGESAHAN	ii
PERNYATAAN ORIGINALITAS	iii
KATA PENGANTAR	iv
DAFTAR ISI	vi
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
INTISARI	xii
ABSTRACT	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah	5
1.4 Tujuan Tugas Akhir	6
1.5 Manfaat Tugas Akhir	6
1.6 Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI	8
2.1 Tinjauan Pustaka	8
2.2 Dasar Teori	10
2.2.1 Analisis Sentimen	10
2.2.2 Iflix	11
2.2.3 Deep Learning	11
2.2.4 Python	12
2.2.5 Google Colaboratory	12
2.2.6 BERT (Bidirectional Encoder Representations from Transformers)	13
2.2.7 IndoBERT	18
2.2.8 Cross Validation	19
2.2.9 Confusion Metrics	20
2.2.10 Accuracy	21
2.2.11 Precesion dan Recall	21

2.2.12 F1-Score.....	23
BAB III METODE TUGAS AKHIR	24
3.1 Tempat dan waktu penelitian	24
3.2 Alat	24
3.3 Tahapan Penelitian	25
3.4 Studi Literatur	25
3.5 Pengumpulan Data	26
3.6 Pre-Processing Data.....	26
3.7 Transformasi Data	28
BAB IV HASIL DAN PEMBAHASAN	29
4.1 Pengambilan Data.....	29
4.2 Pre-Processing Data.....	31
4.2.1 Data Selection.....	31
4.2.2 Data Cleaning (Pembersihan Data)	33
4.2.3 Case Folding.....	34
4.2.4 Tokenizing	35
4.2.5 Normalization	35
4.2.6 Stopword Removal	36
4.2.7 Stemming/Lemmatization.....	37
4.2.8 ToSentence (konversi ke kalimat).....	38
4.2.9 Kamus Lexicon	38
4.3.0 Labelisasi Data.....	39
4.3.1 Data Transformation (Transformasi Data)	43
4.3 Implementasi Algoritma.....	45
4.3.1 Klasifikasi Data.....	45
4.3.2 Hasil Pengujian.....	60
4.3.3 Analisis Proses dan Hasil Pengujian.....	68
BAB V KESIMPULAN DAN SARAN	74
5.1 Kesimpulan.....	74
5.2 Saran	75
DAFTAR PUSTAKA	69

LAMPIRAN	72
-----------------------	-----------

DAFTAR TABEL

Tabel 2. 1 Daftar Penelitian Terdahulu.....	8
<i>Tabel 4. 1 Atribut yang telah diseleksi</i>	<i>32</i>
<i>Tabel 4. 2 Hasil proses cleaning</i>	<i>33</i>
<i>Tabel 4. 3 Hasil proses case folding.....</i>	<i>34</i>
<i>Tabel 4. 4 Hasil proses tokenizing</i>	<i>35</i>
<i>Tabel 4. 5 Hasil proses normalization.....</i>	<i>36</i>
<i>Tabel 4. 6 Hasil Proses stopword removal.....</i>	<i>36</i>
<i>Tabel 4. 7 Proses stemming dan lemmatization.....</i>	<i>37</i>
<i>Tabel 4. 8 Hasil proses tosentence</i>	<i>38</i>
<i>Tabel 4. 9 Hasil proses penggunaan kamus lexicon</i>	<i>39</i>
<i>Tabel 4. 10 Hasil Labelisasi data</i>	<i>40</i>
<i>Tabel 4. 11 Hasil Proses sebelum labelisasi dataset.....</i>	<i>40</i>
<i>Tabel 4. 12 Hasil proses sesudah labelisasi dataset</i>	<i>41</i>
<i>Tabel 4. 13 Hasil perbandingan jumlah label manual dan lexicon</i>	<i>42</i>
<i>Tabel 4. 14 Hasil uji coba pre-training dan fine-tuning</i>	<i>61</i>
<i>Tabel 4. 15 Jumlah kata sentimen negative berdasarkan frekuensi</i>	<i>62</i>
<i>Tabel 4. 16 Jumlah kata sentimen positive berdasarkan frekuensi.....</i>	<i>64</i>
<i>Tabel 4. 17 Jumlah kata sentimen positive berdasarkan frekuensi.....</i>	<i>65</i>
Tabel 4. 18 Hasil Classification Report	68

DAFTAR GAMBAR

Gambar 2. 1	Arsitektur Transformers	14
Gambar 2. 2	Representasi Input BERT	15
Gambar 2. 3	Alur Proses Pre-training dan Fine-tuning	17
Gambar 2. 4	Contoh dari Confusion Matrix	21
Gambar 2. 5	Dua class Confusion Matrix	22
Gambar 3. 1	Tahapan Penelitian	25
Gambar 4. 1	Hasil Aplikasi Iflix pada google play store	30
Gambar 4. 2	Hasil ulasan Iflix pada google play store	30
Gambar 4. 3	Hasil proses scraping	30
Gambar 4. 4	Hasil proses data iflix yang telah di ambil	31
Gambar 4. 5	Data sebelum melalui proses seleksi data	32
Gambar 4. 6	Hasil perbandingan jumlah label manual dan lexicon	42
Gambar 4. 7	Hasil jumlah kesalahan dalam pelabelan	42
Gambar 4. 8	Hasil Transformasi Data	44
Gambar 4. 9	Hasil jumlah distribusi sentimen	44
Gambar 4. 10	Proses install library python	46
Gambar 4. 11	Proses import library python pertama	46
Gambar 4. 12	Proses import library python kedua	46
Gambar 4. 13	Proses read data csv dan hasil proses read data csv	47
Gambar 4. 14	Proses penyimpanan data dalam bentuk tsv	47
Gambar 4. 15	Hasil proses data splitting	48
Gambar 4. 16	Hasil proses inialisasi common functions	49
Gambar 4. 17	Proses pre-trained model IndoBERT	50
Gambar 4. 18	Hasil pre-trained model IndoBERT	51
Gambar 4. 19	Hasil proses tokenisasi dan encoding teks	51
Gambar 4. 20	Proses Inialisasi Model dan Optimizer	52
Gambar 4. 21	Proses Inialisasi Dataset dan Data Loader	52
Gambar 4. 22	Hasil Proses Klasifikasi Sentimen	53
Gambar 4. 23	Proses uji coba pre-trained klasifikasi sentimen pertama .	53
Gambar 4. 24	Proses uji coba pre-trained klasifikasi sentimen kedua	54

<i>Gambar 4. 25 Proses uji coba pre-trained klasifikasi sentimen ketiga</i>	<i>54</i>
<i>Gambar 4. 26 Proses training dan evaluasi.....</i>	<i>55</i>
<i>Gambar 4. 27 Hasil training dan evaluasi.....</i>	<i>55</i>
<i>Gambar 4. 28 Hasil proses evaluasi model setelah training</i>	<i>56</i>
<i>Gambar 4. 29 Hasil uji coba fine-tuning klasifikasi sentimen pertama...</i>	<i>57</i>
<i>Gambar 4. 30 Hasil uji coba fine-tuning klasifikasi sentimen kedua</i>	<i>57</i>
<i>Gambar 4. 31 Hasil uji coba fine-tuning klasifikasi sentimen ketiga</i>	<i>57</i>
<i>Gambar 4. 32 Hasil wordcloud sentimen positive</i>	<i>58</i>
<i>Gambar 4. 33 Hasil wordcloud sentimen neutral.....</i>	<i>58</i>
<i>Gambar 4. 34 Hasil wordcloud sentimen negative</i>	<i>58</i>
<i>Gambar 4. 35 Hasil diagram confusion matrix.....</i>	<i>59</i>
<i>Gambar 4. 36 Hasil prediksi classification report</i>	<i>60</i>
<i>Gambar 4. 37 Worldcloud sentimen negative</i>	<i>62</i>
<i>Gambar 4. 38 Worldcloud sentimen positive</i>	<i>63</i>
<i>Gambar 4. 39 Worldcloud sentimen neutral.....</i>	<i>65</i>
<i>Gambar 4. 40 Hasil confusion matrix</i>	<i>67</i>