

ANALISIS SENTIMEN KOMENTAR TRAILER FILM OPPENHEIMER PADA YOUTUBE MENGGUNAKAN GRU DAN *FASTTEXT EMBEDDING*

AYSUKA ANSARI



**DEPARTEMEN ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Analisis Sentimen Komentar Trailer Film Oppenheimer Pada Youtube Menggunakan GRU dan *Fasttext Embedding*” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juli 2024

Aysuka Ansari
G6401201087

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



ABSTRAK

AYSUKA ANSARI. Analisis Sentimen Komentar Trailer Film Oppenheimer pada YouTube menggunakan GRU dan *Fasttext Embedding*. Dibimbing oleh TOTO HARYANTO dan FIRMAN ARDIANSYAH.

Pada era internet saat ini, trailer umumnya diunggah melalui platform media sosial seperti YouTube. Melalui platform ini, penonton dapat mengakses serta memberikan sentimen mereka terhadap trailer film dengan mudah. Analisis sentimen terhadap komentar trailer film perlu dilakukan agar diperoleh trailer yang disukai oleh penonton sehingga minat penonton untuk menonton film tersebut dapat meningkat. Penelitian ini bertujuan mengembangkan model *Natural Language Processing* (NLP) dengan metode *word embedding* FastText dan arsitektur *Gated Recurrent Unit* (GRU) untuk analisis kecenderungan sentimen komentar di dalam trailer film Oppenheimer. Berdasarkan hasil pelatihan dan evaluasi diperoleh model TextBlob VADER sebagai model terbaik dengan *accuracy* bernilai 0,93. Hasil analisis kecenderungan sentimen pada masing-masing trailer film menggunakan model ini menghasilkan pengamatan bahwa trailer 2 lebih diminati penonton dibandingkan trailer 1. Selain itu, penelitian ini juga telah berhasil mengembangkan aplikasi *web app* sederhana untuk membantu proses analisis sentimen.

Kata kunci: analisis sentimen, *fasttext word embedding*, *gated recurrent unit*, trailer film, youtube

ABSTRACT

AYSUKA ANSARI. Sentiment Analysis of Oppenheimer Movie Trailer Comments on YouTube using GRU and FastText Embedding. Supervised by TOTO HARYANTO and FIRMAN ARDIANSYAH.

In today's internet era, trailers are commonly uploaded on social media platforms like YouTube where viewers can easily access them and express their opinions. Sentiment analysis for a movie trailer is needed in order to create a trailer that is liked by the viewer so that it can increase the viewer's interest in watching the movie. This research aims to develop a Natural Language Processing (NLP) model using the FastText word embedding method and Gated Recurrent Unit (GRU) architecture for analyzing sentiment tendencies in comments on the Oppenheimer's movie trailers. Based on the training and evaluation results, the TextBlob VADER model was found to be the best model with an accuracy of 0,93. The sentiment trend analysis on each movie trailer using this model showed that trailer 2 was more favored by viewers compared to trailer 1. Additionally, this research successfully developed a simple web app to assist in the sentiment analysis process.

Keywords: *fasttext word embedding*, *gated recurrent unit*, movie trailer, sentiment analysis, youtube

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2024
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

ANALISIS SENTIMEN KOMENTAR TRAILER FILM OPPENHEIMER PADA YOUTUBE MENGGUNAKAN GRU DAN *FASTTEXT EMBEDDING*

AYSUKA ANSARI

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Ilmu Komputer

**DEPARTEMEN ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Penguji pada Ujian Skripsi:
Dr. Mushthofa S.Kom., M.Sc.



Judul Skripsi : Analisis Sentimen Komentar Trailer Film Oppenheimer Pada Youtube Menggunakan GRU dan *Fasttext Embedding*

Nama : Aysuka Ansari
NIM : G6401201087

@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Disetujui oleh

Pembimbing 1:
Dr. Toto Haryanto, S.Kom, M.Si.



Pembimbing 2:
Firman Ardiansyah, S.Kom, M.Si.



Diketahui oleh

Ketua Departemen Ilmu Komputer:
Dr. Sony Hartono Wijaya, S.Kom., M.Kom.
19810809 200812 1 002



Tanggal Ujian:
6 Juni 2024

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah *subhanaahu wa ta'ala* atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Desember 2023 sampai bulan Mei 2024 ini ialah analisis sentimen, dengan judul “Analisis Sentimen Komentar Trailer Film Oppenheimer Pada Youtube Menggunakan GRU dan *Fasttext Embedding*”.

Penulis menyadari banyak pihak yang telah memberikan dukungan dan bantuan selama menyelesaikan tugas akhir ini. Pada kesempatan ini, penulis ingin menyampaikan terima kasih kepada:

- a. Orang tua penulis, Bapak Muswar dan Ibu Lismawati yang selalu memberikan dukungan dan doa kepada penulis.
- b. Saudara kandung penulis, Fuad dan Zila yang senantiasa memberikan semangat dan doa kepada penulis.
- c. Bapak Dr. Toto Haryanto S.Kom., M.Si. dan Bapak Firman Ardiansyah, S.Kom, M.Si. selaku dosen pembimbing skripsi penulis yang senantiasa memberikan dukungan dan arahan selama penelitian.
- d. Bapak Dr. Mushthofa S.Kom., M.Sc. selaku dosen penguji yang telah memberikan masukan untuk tugas akhir penulis.
- e. Fadil selaku sahabat penulis yang selalu mendoakan dan mendukung proses pengerjaan tugas akhir ini.
- f. Zikri dan Danil selaku rekan satu tim rsolv3 yang selalu mendukung proses pengerjaan tugas akhir satu sama lain.
- g. Bintang, Surya, Daffa, Dzakiyyah, dan Zahra selaku teman penulis yang selalu memberikan bantuan dan dukungan disaat penulis mengalami kesulitan dalam pengerjaan tugas akhir ini.
- h. Mahasiswa program studi S1 Ilmu Komputer angkatan 57 yang telah memberikan bantuan dan semangat kepada penulis selama perkuliahan.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Mei 2024

Aysuka Ansari



DAFTAR ISI

ABSTRAK	ii
DAFTAR TABEL	ix
DAFTAR GAMBAR	ix
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	3
1.4 Manfaat	3
1.5 Ruang Lingkup	3
TINJAUAN PUSTAKA	4
2.1 Film Oppenheimer	4
2.2 <i>Natural Language Processing (NLP)</i>	4
2.3 Analisis Sentimen	4
2.4 <i>Valence Aware Dictionary for Sentiment Reasoner (VADER)</i>	5
2.5 TextBlob	5
2.6 <i>Synthetic Minority Oversampling Technique (SMOTE)</i>	5
2.7 <i>Random Downsampling</i>	7
2.8 <i>Word Embedding</i>	7
2.9 <i>Continuous Bag of Words (CBOW)</i>	7
2.10 <i>FastText Algorithm</i>	8
2.11 <i>Activation Function</i>	9
2.12 <i>Recurrent Neural Network (RNN)</i>	9
2.13 <i>Gated Recurrent Unit (GRU)</i>	10
2.14 <i>Confusion Matrix</i>	12
III METODE	14
3.1 Data Penelitian	14
3.2 Tahapan Penelitian	14
3.3 Lingkungan Pengembangan Sistem	27
IV HASIL DAN PEMBAHASAN	29
4.1 Persiapan Data	29
4.2 Pembentukan Model	38
4.3 Evaluasi Model	43
4.4 Analisis Sentimen Trailer	47
4.5 <i>Model Deployment</i>	47
V SIMPULAN DAN SARAN	49
5.1 Simpulan	49
5.2 Saran	49
DAFTAR PUSTAKA	50
RIWAYAT HIDUP	54

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

DAFTAR TABEL

1	Struktur <i>confusion matrix</i>	12
2	Spesifikasi <i>endpoint</i> pada <i>web server</i>	27
3	Contoh data komentar di dalam dataset	29
4	Contoh komentar dengan panjang ≤ 2	29
5	Contoh komentar dengan panjang ≥ 1600	30
6	Contoh komentar terurut berdasarkan rata-rata panjang kata	30
7	Perbandingan data sebelum dan sesudah pra-proses	33
8	Hasil validasi pelabelan	35
9	Hasil identifikasi topik	36
10	Sebaran data pada setiap partisi	37
11	Konfigurasi <i>hyperparameter</i> model FastText	37
12	Hasil <i>embedding matrix</i>	38
13	Detail arsitektur model GRU	38
14	Konfigurasi <i>hyperparameter</i> model GRU	38
15	Hasil evaluasi performa model	45
16	Hasil evaluasi lanjutan model	46

DAFTAR GAMBAR

1	Algoritma SMOTE (Chawla <i>et al.</i> 2002)	6
2	Arsitektur model CBOW (Qiu <i>et al.</i> 2020)	8
3	Arsitektur FastText (Joulin <i>et al.</i> 2017)	9
4	Arsitektur RNN (Setianingsih 2023)	10
5	Arsitektur GRU (Caniago <i>et al.</i> 2021)	11
6	Diagram alur penelitian	14
7	Diagram alur pengumpulan data	15
8	Diagram alur pelabelan VADER	18
9	Diagram alur pelabelan TextBlob	19
10	Alur validasi hasil pelabelan data	19
11	Diagram alur implementasi SMOTE	21
12	Diagram alur implementasi <i>random downsampling</i>	21
13	Diagram alur proses pembuatan <i>embedding matrix</i>	22
14	Diagram alur pemecahan dataset	23
15	Arsitektur <i>deep learning</i> model (Zulqarnain <i>et al.</i> 2019)	24
16	Diagram alur detail pembuatan setiap model	25
17	Arsitektur <i>web app</i>	26
18	Diagram 15 <i>stopwords</i> teratas	31
19	Diagram 10 <i>bigram</i> teratas	31
20	Diagram 10 <i>trigram</i> teratas	32
21	Diagram hasil identifikasi bahasa	32
22	Hasil pelabelan data	34
23	<i>Confusion matrix</i> hasil pelabelan TextBlob dan VADER	34
24	Hasil proses <i>balancing</i>	36

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

25	Histori pelatihan model VADER	39
26	Histori pelatihan model TextBlob	40
27	Histori pelatihan model TextBlob VADER	42
28	<i>Confusion matrix</i> setiap model	44
29	Hasil analisis sentimen	47
30	Tampilan halaman <i>home web app</i>	48
31	Hasil prediksi model	48

@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.