

**ANALISIS SENTIMEN TERHADAP LAYANAN INTERNET FIRST  
MEDIA DI TWITTER DENGAN ALGORITMA *DEEP LEARNING* STUDI  
KOMPARASI *CONVOLUTIONAL NEURAL NETWORK* DAN  
*RECURRENT NEURAL NETWORK***

SKRIPSI

Diajukan untuk memenuhi sebagian dari syarat memperoleh gelar Sarjana pada  
Program Studi Sistem Telekomunikasi Universitas Pendidikan Indonesia



Oleh  
Titania Emaniar  
1905725

**PROGRAM STUDI SISTEM TELEKOMUNIKASI  
KAMPUS UPI DI PURWAKARTA  
UNIVERSITAS PENDIDIKAN INDONESIA  
2023**

**ANALISIS SENTIMEN TERHADAP LAYANAN INTERNET FIRST  
MEDIA DI TWITTER DENGAN ALGORITMA *DEEP LEARNING* STUDI  
KOMPARASI *CONVOLUTIONAL NEURAL NETWORK* DAN  
*RECURRENT NEURAL NETWORK***

Oleh

Titania Emaniar

Sebuah skripsi yang diajukan untuk memenuhi salah satu syarat memperoleh gelar  
Sarjana Teknik pada Program Studi Sitem Telekomunikasi

© Titania Emaniar 2023  
Universitas Pendidikan Indonesia  
Agustus 2023

Hak Cipta dilindungi undang-undang.  
Skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian,  
dengan dicetak ulang, difoto kopi, atau cara lainnya tanpa ijin dari penulis

**LEMBAR PENGESAHAN**

**TITANIA EMANIAR**

**ANALISIS SENTIMEN TERHADAP LAYANAN INTERNET FIRST MEDIA  
DI TWITTER DENGAN ALGORITMA *DEEP LEARNING* STUDI  
KOMPARASI *CONVOLUTIONAL NEURAL NETWORK* DAN *RECURRENT  
NEURAL NETWORK***

Disetujui dan Disahkan Oleh:

Pembimbing I,



Dewi Indriati Hadi Putri, S.Pd., M.T.

NIP.920190219906126201

Pembimbing II



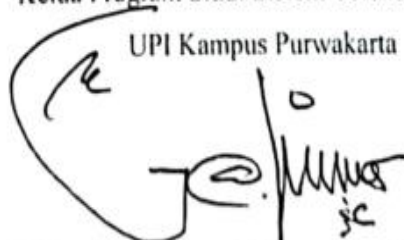
Endah Setvowati, S.T., M.T.

NIP.920190219920908201

Mengetahui,

Ketua Program Studi Sistem Telekomunikasi

UPI Kampus Purwakarta



Galura Muhammad Suranegara, S.Pd., M.T.

NIP.920190219920111101

## SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa skripsi saya dengan judul Analisis Sentimen Terhadap Layanan Internet First Media di Twitter Dengan Algoritma *Deep Learning* Studi Komparasi *Convolutional Neural Network* dan *Recurrent Neural Network* Ini beserta seluruh isinya adalah benar-benar karya saya sendiri. Saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika ilmu yang berlaku dalam masyarakat, Atas pernyataan ini, saya siap menanggung risiko atau sanksi apa bila dikemudian hari ditemukan adanya pelanggaran etika keilmuan atau ada klaim dari pihak lain terhadap keaslian karya saya ini.

Purwakarta, 29 Agustus 2023  
Penulis,



Titania Emaniar

## KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT, Tuhan Yang Maha Esa, atas rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi ini dengan judul "Analisis Sentimen Terhadap Layanan Internet First Media Dengan Algoritma *Deep Learning* Studi Komparasi *Convolutional Neural Network* Dan *Recurrent Neural Network*". Skripsi ini disusun sebagai salah satu syarat untuk meraih gelar Sarjana dari Program Studi Sistem Telekomunikasi, Universitas Pendidikan Indonesia. Solawat serta salam senantiasa tercurahkan kepada Rasulullah Nabi Muhammad SAW, keluarga, sahabat dan semua yang mengikuti ajarannya hingga akhir Zaman. Penulis mengucapkan terima kasih kepada semua pihak yang terlibat dalam penelitian ini.

Penulis menyadari bahwa manusia pada hakikatnya tak luput dari kesalahan dan kekurangan. Oleh karena itu, jika terdapat kesalahan baik dalam aspek penulisan maupun penyajian. Penulis dengan rendah hati menerima saran dan kritik yang membangun dari pembaca, sehingga kesalahan dan kekurangan tersebut dapat diperbaiki dan diambil sebagai pembelajaran untuk kedepannya

Akhir kata, semoga skripsi ini dapat memberikan ilmu pengetahuan dan bermanfaat bagi semua yang membacanya. Aamiin Yaa Rabbal Aalamiin.

Purwakarta, Agustus 2023

Penulis,



Titania Emaniar

## UCAPAN TERIMA KASIH

Puji dan Syukur penulis panjatkan kehadirat Allah SWT, yang senantiasa memberikan rahmat dan ridhanya sehingga penulis dapat menyelesaikan skripsi ini tepat pada waktunya. Penulis menyadari skripsi ini tidak akan selesai tanpa doa, dukungan dan dorongan dari berbagai pihak. Dalam kesempatan ini penulis ingin mengucapkan rasa terima kasih yang sebesar-besarnya kepada:

1. Allah Subhanahu Wa Ta'ala, Tuhan Yang Maha Esa, atas rahmat, berkah dan karunia-Nya yang selalu memberikan kemudahan dan kelancaran selama penyusunan tugas akhir ini.
2. Kedua orang tua penulis, yaitu Bapak Eman dan Ibu Nina yang telah senantiasa memberikan kasih sayang dan dukungan secara moril maupun materil, memotivasi penulis serta selalu mendoakan kebaikan dan kesuksesan kepada penulis sehingga dapat menyelesaikan perkuliahan tepat pada waktunya.
3. Kepada adik saya tercinta yaitu Nahda Rafanda yang selalu ceria dan memberikan dorongan kepada penulis untuk selalu semangat dalam menjalankan perkuliahan dan menyelesaikan penyusunan skripsi penulis ini.
4. Ibu Dewi Indriati Hadi Putri, S.Pd., M.T. selaku Dosen Pembimbing I skripsi penulis yang selalu menyempatkan waktunya untuk membimbing dan memberikan arahan serta masukan kepada penulis selama proses penyusunan skripsi.
5. Ibu Endah Setyowati, S.T., M.T. selaku pembimbing II yang telah menyempatkan waktu untuk membimbing dan memberikan masukan-masukan mengenai kepenulisan kepada penulis dalam penyusunan skripsi ini.
6. Bapak Ahmad Fauzi, S.Si., M.T. selaku dosen wali dan Ketua Program Studi Sistem Telekomunikasi Universitas Pendidikan Indonesia Kampus Purwakarta periode 2019-2023
7. Bapak Galura Muhammad Suranegara, S.Pd., M.T. selaku Ketua Program Studi Sistem Telekomunikasi Universitas Pendidikan Indonesia Kampus Purwakarta. Yang selalu aktif memberikan informasi-informasi terkait penyusunan dan sidang skripsi sehingga membantu kelancaran ini.
8. Kepada Winda Pratiwi yang telah menemani dan berjuang bersama penulis selama 4 tahun dari awal masuk kuliah hingga akhir masa perkuliahan.

9. Kepada teman-teman Sistem Telekomunikasi tahun 2019 yang telah kebersamai dan berjuang bersama selama ini memberikan penulis pengalaman berharga selama masa perkuliahan.
10. Segenap dosen Program Studi Sistem Telekomunikasi yang telah membimbing dan memberikan ilmu kepada penulis selama masa perkuliahan.
11. Dan seluruh pihak yang tidak bisa penulis sebutkan satu persatu yang telah kebersamai penulis selama masa perkuliahan ini.

## ABSTRAK

Internet telah menjadi aspek penting yang tidak terpisahkan dalam kehidupan di era digital saat ini. First Media merupakan salah satu provider Internet yang menyediakan akses Internet dengan kualitas yang baik, namun pengguna First Media tidak signifikan jika dibandingkan dengan provider Internet lain. Oleh karena itu, penelitian ini bertujuan untuk memvalidasi apakah Internet First Media memang memiliki kualitas Internet yang baik dengan cara melakukan analisis sentimen terhadap pengguna layanan First Media di Twitter. Metode analisis sentimen digunakan untuk memahami pandangan, perasaan, dan preferensi pelanggan terhadap layanan First Media. Data yang digunakan adalah tweet di Twitter yang berisi ulasan mengenai layanan Internet First Media. Data ini dianalisis menggunakan metode *Deep Learning*, yaitu *Convolutional Neural Network* (CNN) dan *Recurrent Neural Network* (RNN), dengan penerapan tiga *Feature Extraction* yaitu *FastText*, CBOW, TF-IDF. Hasil klasifikasi menunjukkan sentimen yang lebih banyak yaitu data dalam kelas negatif dari pada data positif. Dan evaluasi kinerja model menunjukkan bahwa seluruh percobaan model CNN memiliki hasil yang lebih baik dari RNN. CNN dengan *FastText* memiliki akurasi 91%, CNN dengan CBOW 93% dan CNN dengan TF-IDF 97%. Sedangkan model RNN dengan *FastText* 79%, RNN dengan CBOW 77% dan TF-IDF 96%.

Kata Kunci: Analisis Sentimen, *Convolutional Neural Network*, *Recurrent Neural Network*, TF-IDF, *FastText*, CBOW.



## ABSTRACT

*The internet has become an important and inseparable aspect of life in today's digital era. First Media is an internet provider that provides good quality internet access, but First Media users are not significant when compared to other internet providers. Therefore, this research aims to validate whether First Media Internet does have good internet quality by conducting a sentiment analysis of First Media service users on Twitter. The sentiment analysis method is used to understand the views, feelings, and preferences of customers towards First Media's services. The data used are tweets on Twitter that contain reviews about First Media's internet services. This data was analyzed using Deep Learning methods, namely Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN), with the application of three Feature Extraction, namely FastText, CBOW, and TF-IDF. The classification results show more sentiment, namely data in the negative class than positive data. And the model performance evaluation shows that in all experiments the CNN model has better results than RNN. CNN with FastText has 91% accuracy, CNN with CBOW 93%, and CNN with TF-IDF 97%. While the RNN model with 79% FastText, RNN with 77% CBOW, and 96% TF-IDF.*

**Keywords:** *Sentiment Analysis, Convolutional Neural Network, Recurrent Neural Network, TF-IDF, FastText, CBOW.*

## DAFTAR ISI

HAK CIPTA .....	ii
LEMBAR PENGESAHAN .....	iii
SURAT PERNYATAAN.....	iv
KATA PENGANTAR .....	v
UCAPAN TERIMA KASIH.....	vi
ABSTRAK .....	viii
ABSTRACT .....	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR .....	xv
DAFTAR ISTILAH .....	xvi
DAFTAR LAMPIRAN .....	xvii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Tujuan Penelitian.....	3
1.4 Batasan Masalah.....	3
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	5
BAB II LANDASAN TEORI .....	6
2.1 Kajian Penelitian yang Relevan .....	6
2.1 Analisis Sentimen.....	10
2.2 Twitter .....	11
2.3 First Media .....	12
2.4 <i>Natural Language Processing</i> .....	12
2.5 <i>Deep Learning</i> .....	13
2.8 <i>Convolutional Neural Network</i> .....	13

2.10	<i>Recurrent Neural Network</i> .....	15
2.11	<i>Feature Extraction</i> .....	16
2.11.1	<i>Word Embedding</i> .....	16
2.11.2	TF-IDF .....	17
2.12	<i>Confusion Matrix</i> .....	18
BAB III METODOLOGI PENELITIAN .....		20
3.1	Jenis Penelitian .....	20
3.2	Objek, Populasi dan Sample Penelitian.....	20
3.5	Alur Penelitian.....	20
3.5.1.	Pengumpulan Data .....	21
3.5.2.	Pra-pemrosesan Data.....	22
3.5.3.	<i>Modeling</i> .....	28
3.5.4.	Evaluasi Model.....	30
3.5.5.	Perangkat Penelitian.....	31
BAB IV TEMUAN DAN PEMBAHASAN.....		32
4.1	Hasil Data <i>Crawling</i> .....	32
4.2	<i>Preprocessing</i> .....	32
4.2.1	Penggabungan Data.....	32
4.2.2	Hasil Hapus Duplikasi Data .....	33
4.2.3	Hasil Penghapusan Data Tidak Relevan .....	34
4.2.4	Hasil Pelabelan.....	36
4.2.5	Hasil Penghapusan URL, Tanda baca dan simbol .....	37
4.2.6	Hasil <i>Tokenizer</i> , <i>Stopword</i> dan <i>Casefolding</i> .....	38
4.3	<i>Modeling</i> .....	38
4.3.1	Hasil <i>Modeling</i> CNN .....	39
4.3.2	Hasil <i>Modeling</i> RNN .....	42

4.4	Analisis dan Interpretasi Hasil .....	47
4.4.1	Analisis Performansi Model.....	47
4.4.2	Hasil Analisis Sentimen .....	49
BAB V SIMPULAN, IMPLIKASI DAN SARAN.....		51
5.1	Kesimpulan.....	51
5.2	Implikasi .....	51
5.3	Saran .....	52
DAFTAR PUSTAKA .....		53
LAMPIRAN .....		55

## DAFTAR TABEL

Tabel 2.1 Penelitian Sebelumnya.....	8
Tabel 2. 2 <i>Confusion Matrix</i> .....	18
Tabel 3. 1 Kata Kunci Pencarian.....	22
Tabel 3. 2 Identifikasi Data Tidak Relevan .....	24
Tabel 3. 3 Pelabelan Data.....	26
Tabel 3.4 Proses Penghapusan Tanda Baca dan Simbol.....	27
Tabel 3. 5 <i>Software</i> yang di Gunakan .....	31
Tabel 4.1 Proses dan Hasil Penggabungan Data.....	33
Tabel 4.2 Proses dan Hasil Penghapusan Duplikasi Data.....	34
Tabel 4.3 Proses dan Hasil Penghapusan Data Tidak Relevan.....	35
Tabel 4.4 Hasil Penghapusan URL, Tanda Baca dan Simbol.....	37
Tabel 4.5 Hasil <i>Tokenizer, Stopword dan Casefolding</i> .....	38
Tabel 4.6 Perbandingan Performa Model .....	47
Tabel 4.7 Kata yang Sering Muncul .....	50

## DAFTAR GAMBAR

Gambar 2.1 Arsitektur <i>Layer CNN</i> , direproduksi dari (Dang dkk., 2020) .....	14
Gambar 2. 2 Arsitektur RNN, direproduksi dari (Faturohman dkk,2020).....	15
Gambar 3.1 Alur Penelitian.....	21
Gambar 3.2 Proses <i>Crawling Data</i> .....	22
Gambar 3.3 Tahapan Pra-pemrosesan Data .....	23
Gambar 3.4 Proses Pelebelan Data .....	25
Gambar 3.5 Alur Modeling CNN.....	28
Gambar 4.1 Hasil Data <i>Crawling</i> .....	32
Gambar 4.2 Data Tidak Relevan yang Berisi Iklan .....	36
Gambar 4.3 Persebaran Data Dengan Label .....	37
Gambar 4.4 <i>Classification Report CNN</i> dengan FastText.....	39
Gambar 4.5 (a) <i>Loss Selama Epoch</i> dan (b) <i>Accuracy Selama Epochs CNN</i> dengan <i>FastText</i> .....	39
Gambar 4.6 <i>Classification Report CNN</i> dengan CBOW.....	40
Gambar 4.7 (a) <i>Loss Selama Epoch</i> dan (b) <i>Accuracy Selama Epochs CNN</i> dengan CBOW .....	41
Gambar 4.8 <i>Classification Report RNN</i> dengan TF-IDF .....	41
Gambar 4. 9 (a) <i>Loss Selama Epoch</i> dan (b) <i>Accuracy Selama Epochs CNN</i> dengan TF-IDF.....	42
Gambar 4.10 <i>Classification Report RNN</i> dengan FastText.....	43
Gambar 4.11 (a) <i>Loss Selama Epoch</i> dan (b) <i>Accuracy Selama Epochs RNN</i> dengan FastText .....	43
Gambar 4.12 <i>Classification Report RNN</i> dengan CBOW.....	44
Gambar 4.13 (a) <i>Loss Selama Epoch</i> dan (b) <i>Accuracy Selama Epochs RNN</i> dengan CBOW .....	45
Gambar 4.14 <i>Classification Report RNN</i> dengan TF-IDF .....	46
Gambar 4.15 (a) <i>Loss Selama Epoch</i> dan (b) <i>Accuracy Selama Epochs RNN</i> dengan TF-ID .....	46
Gambar 4.16 <i>Word cloud Data Postif</i> .....	49
Gambar 4.17 <i>Word cloud Data Negatif</i> .....	50

## DAFTAR ISTILAH

SINGKATAN	NAMA	PEMAKAIAN PERTAMA KALI
NLP	<i>Natural Language Processing</i>	2
CNN	<i>Convolutional Neural Network</i>	2
RNN	<i>Recurrent Neural Network</i>	2
LSTM	<i>Long Short-Term Memory</i>	15
TF-IDF	<i>Term Frequency-Inverse Document Frequency</i>	17

## DAFTAR LAMPIRAN

Lampiran 1.1 <i>Dataset</i> yang digunakan .....	55
Lampiran 1.2 Code Penggabungan Data.....	57
Lampiran 1.3 Hapus Duplikasi Data.....	58
Lampiran 1.4 Penghapusan Data Tidak Relevan .....	58
Lampiran 1.5 Code Pelabelan <i>Transformers</i> .....	59
Lampiran 1.6 Penghapusan URL, Tanda baca dan simbol .....	60
Lampiran 1.7 Lampiran Code <i>Modeling</i> CNN dengan <i>FastText</i> .....	61
Lampiran 1.8 Code <i>Modeling</i> CNN dengan CBOW .....	62
Lampiran 1.9 Code CNN dengan TF-IDF .....	63
Lampiran 1.10 Code <i>Modeling</i> RNN dengan <i>FastText</i> .....	63
Lampiran 1.11 Code RNN dengan CBOW .....	64
Lampiran 1.12 Code RNN dengan TF-IDF .....	65



## DAFTAR PUSTAKA

- Annur. (2023). 10 Negara dengan Jumlah Pengguna Twitter Terbanyak di Dunia [online]. Diakses pada <https://databoks.katadata.co.id/datapublish/2023/02/27/pengguna-twitter-di-indonesia-capai-24-juta-hingga-awal-2023-peringkat-berapa-di-dunia>
- APJII. (2023). Survei APJII Pengguna Internet di Indonesia Tembus 215 Juta Orang [online]. Diakses pada <https://apjii.or.id/berita/d/survei-apjii-pengguna-Internet-di-indonesia-tembus-215-juta-orang>
- Badjrie, S. H., Pratiwi, O. N., & Anggana, H. D. (2021). Analisis Sentimen Review Customer Terhadap Produk Indihome Dan First Media Menggunakan Metode Convolutional Neural Network. *eProceedings of Engineering*, 8(5), 9057.
- Dang, N. C., Moreno-García, M. N., & De la Prieta, F. (2020). Sentiment Analysis Based on *Deep Learning*: A Comparative Study. *Electronics*, 9(3). <https://doi.org/10.3390/electronics9030483>
- Gandhi, U. D., Malarvizhi Kumar, P., Chandra Babu, G., & Karthick, G. (2021). Sentiment Analysis on Twitter Data by Using *Convolutional Neural Network* (CNN) and *Long Short Term Memory* (LSTM). *Wireless Personal Communications*. <https://doi.org/10.1007/s11277-021-08580-3>
- Goularas, D., & Kamis, S. (2019). Evaluation of *Deep Learning* Techniques in Sentiment Analysis from Twitter Data. *2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML)*, 12–17. <https://doi.org/10.1109/Deep-ML.2019.00011>
- Khatri. (2023). Indonesia, Fixed Broadband Experience, June 2023 [online]. Diakses pada <https://www.opensignal.com/2023/06/27/indonesia-fixed-broadband-experience-june-2023>
- Kim, H., & Jeong, Y.-S. (2019). Sentiment Classification Using Convolutional Neural Networks. *Applied Sciences*, 9(11), 2347. <https://doi.org/10.3390/app9112347>
- Madasu, A., & E, S. (2019). *A Study of Feature Extraction techniques for Sentiment Analysis* (arXiv:1906.01573). arXiv. <http://arxiv.org/abs/1906.01573>

**Titania Emaniar, 2023**

**ANALISIS SENTIMEN TERHADAP LAYANAN FIRST MEDIA DI TWITTER DENGAN ALGORITMA DEEP LEARNING STUDI KOMPARASI CONVOLUTIONAL NEURAL NETWORK DAN RECURRENT NEURAL NETWORK**

Universitas Pendidikan Indonesia | [repository.upi.edu](https://repository.upi.edu) | [Perpustakaan.upi.edu](https://perpustakaan.upi.edu)

- Mengistie, T. T., & Kumar, D. (2021). *Deep Learning Based Sentiment Analysis On COVID-19 Public Reviews. 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, 444–449. <https://doi.org/10.1109/ICAIIIC51459.2021.9415191>
- Purnamasari, P. D., Taqiyuddin, M., & Ratna, A. A. P. (2017). Performance comparison of text-based sentiment analysis using *Recurrent Neural Network* and convolutional neural network. *Proceedings of the 3rd International Conference on Communication and Information Processing*, 19–23. <https://doi.org/10.1145/3162957.3163012>
- Rahaman Wahab Sait, A., & Khairi Ishak, M. (2023). *Deep Learning with Natural Language Processing Enabled Sentimental Analysis on Sarcasm Classification. Computer Systems Science and Engineering*, 44(3), 2553–2567. <https://doi.org/10.32604/csse.2023.029603>
- Sadya. (2023). APJII: IndiHome Jadi Fixed Broadband Paling Banyak Dipakai 2023 [online]. Diakses pada <https://dataindonesia.id/digital/detail/apjii-indihome-jadi-fixed-broadband-paling-banyak-dipakai-2023>
- Wattimena, A. D. (2018). Analisis Sentimen Teks Bahasa Indonesia Pada Media Sosial Menggunakan Metode *Convolutional Neural Network* (Studi Kasus: E-Commerce). *Doctoral Dissertation, Institut Teknologi Sepuluh Nopember*.