

# ANALISIS SENTIMEN PADA EDOM STMIK PRIMAKARA MENGGUNAKAN METODE LONG SHORT TERM MEMORY

I Gusti Bagus Arinata<sup>1)</sup> I Putu Satwika<sup>2)</sup> Nengah Widya Utami<sup>3)</sup>

Program Teknik Informatika <sup>1)2)</sup>

Program Studi Sitem Informasi Akuntansi <sup>3)</sup>

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Primakara, Denpasar, Bali<sup>1)2)3)</sup>

arinata098@gmail.com<sup>(1)</sup> satwika@primakara.ac.id<sup>2)</sup> widya@primakara.ac.id<sup>3)</sup>

## ABSTRACT

*EDOM (Evaluasi Dosen Oleh Mahasiswa) is an instrument for evaluating the performance of lecturers at universities on the STMIK Primakara campus with the aim of obtaining information from students about the performance of lecturers in the teaching process. The EDOM evaluation process at STMIK Primakara still uses the manual method, which is certainly not effective when there is a lot of EDOM data. Sentiment analysis is a part of artificial intelligence that is used to analyze opinions, emotions, judgments and evaluations of an object. With sentiment analysis the evaluation process on EDOM will be easier to do, because the calculations will be assisted by machines. The classification method used in this study is the recurrent neural network which gets an average accuracy of 88.93% in the class 2 scenario that does not use neutral sentiment..*

**Keywords:** EDOM, sentiment analysis, recurrent neural network, LSTM.

## ABSTRAK

EDOM (Evaluasi Dosen Oleh Mahasiswa) merupakan instrumen untuk evaluasi kinerja dosen pada perguruan tinggi di kampus STMIK Primakara dengan tujuan untuk memperoleh informasi dari mahasiswa tentang kinerja dosen dalam proses mengajar. Proses evaluasi EDOM di STMIK Primakara masih menggunakan cara manual, yang tentunya tidak efektif ketika data EDOM yang sangat banyak. Analisis sentimen merupakan bagian ilmu dari kecerdasan buatan yang digunakan untuk menganalisa pendapat, emosi, penilaian dan evaluasi suatu objek. Dengan analisis sentimen proses evaluasi pada EDOM akan lebih mudah dilakukan, karena perhitungan akan dibantu oleh mesin. Metode klasifikasi yang digunakan pada penelitian ini yaitu *recurrent neural network* dengan varian LSTM yang mendapatkan rata – rata akurasi 88.93% pada skenario class 2 yang tidak menggunakan sentimen netral.

**Kata Kunci :** EDOM, analisis sentimen, *recurrent neural network*, LSTM.

## PENDAHULUAN

Persaingan antar perguruan tinggi semakin ke depan akan semakin pesat. Meningkatkan kualitas pelayanan merupakan salah satu cara untuk menjaga kelangsungan hidup dari perguruan tinggi itu tersebut. Dengan meningkatkan pelayanan kepada mahasiswa merupakan salah satu upaya untuk meningkatkan kualitas pelayanan kampus. Kampus STMIK Primakara memiliki sistem evaluasi kinerja dosen yang mengajar dalam

sistem informasi akademik, biasa disebut EDOM (evaluasi dosen mahasiswa). Kritik dan saran dari mahasiswa bisa digunakan untuk peningkatan mutu pendidikan di STMIK Primakara dengan melakukan proses evaluasi dari komentar yang diberikan, yang terdiri dari komentar negatif maupun komentar positif. Analisis sentimen merupakan cabang penelitian dari text mining, yang dapat digunakan untuk mengklasifikasikan dokumen teks berupa opini berdasarkan label sentimen. Untuk menganalisa sebuah sentimen pada

suatu kalimat, memerlukan pengetahuan akan arti atau makna dari tiap rangkaian kata yang menyusun kalimat tersebut, dimana tiap kata yang tersusun secara berurutan akan menghasilkan suatu kalimat yang bermakna. Deret penggunaan kata pada satu kalimat sangat mempengaruhi arti dari kalimat tersebut untuk menentukan apakah kalimat tersebut merupakan sentimen positive, negative atau neutral. Recurrent neural network yang biasa disebut dengan repetitif neural network atau jaringan saraf tiruan merupakan metode deep learning yang pengolahannya disebut secara berulang untuk mengolah data masukan. Data masukan biasanya berupa beberapa data time series, dimana untuk analisis sentimen, kalimat merupakan salah satu contoh data time series yang tersusun dari kata yang terurut sehingga menyusun sebuah kalimat. Maka dari itu metode recurrent neural network sangat cocok digunakan untuk analisis sentimen. Long Short-Term Memory adalah salah satu arsitektur dari RNN. LSTM dibuat oleh Hochreiter & Schmidhuber dan kemudian dikembangkan dan dipromosikan oleh banyak peneliti. Jaringan LSTM juga memiliki modul pemrosesan berulang yang sama dengan RNN.

## TINJAUAN PUSTAKA

### Analisis Sentimen

Analisis sentimen sebagai bagian dari opinion mining adalah bidang penelitian yang digunakan untuk menganalisis pendapat, emosi, penilaian dan evaluasi, sikap seseorang, dan keadaan emosional suatu objek [1]. Tujuan dasar dari analisis sentimen adalah untuk menemukan dan mengklasifikasikan isi teks dalam kalimat, apakah opini yang diungkapkan oleh fitur dokumen, kalimat atau aspek positif, negatif atau netral. Oleh karena itu, analisis sentimen berfokus pada pemrosesan opini yang mengandung polaritas, yang memiliki nilai sentimen positif, negatif, atau netral [2]

### EDOM (Evaluasi Dosen Oleh Mahasiswa)

Evaluasi Dosen Oleh Mahasiswa (EDOM), instrumen atau alat untuk mengevaluasi kinerja dosen di akhir semester [3]. EDOM merupakan rangkaian proses

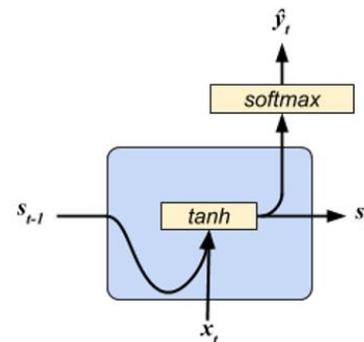
evaluasi pengajaran, tujuannya untuk memperoleh informasi dari mahasiswa tentang kinerja dosen dalam proses mengajar [4].

### Library Pendukung

Tujuan dari penggunaan library untuk mengembangkan sistem analisis sentimen yaitu untuk membantu proses dari analisis sentimen agar efisien dan efektif, karena sudah banyak function yang disediakan oleh library tersebut yang sudah jadi dan tinggal digunakan. Berikut beberapa library yang akan digunakan yaitu: Pandas, Sastrawi, NumPy, NLTK, Scikit-learn, Matplotlib, Keras.

### Recurrent Neural Network

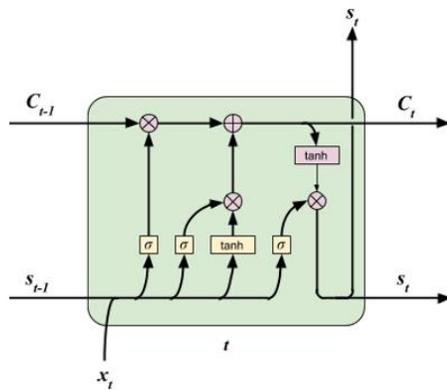
Recurrent Neural Network (RNN) adalah arsitektur jaringan syaraf tiruan, yang pemrosesannya berulang kali dipanggil untuk memproses data input. Data input biasanya berupa beberapa data kontinu (data sekuensial). RNN diklasifikasikan ke dalam kategori deep learning karena data melewati banyak lapisan pemrosesan. Dasar pengembangan arsitektur RNN didasarkan pada cara berpikir bahwa manusia tidak mengambil keputusan tunggal, karena manusia selalu mempertimbangkan informasi yang diterima atau disimpan di masa lalu ketika mengambil keputusan. RNN menyimpan informasi masa lalu dengan mengulangnya dalam arsitekturnya sehingga informasi masa lalu tetap tersimpan [5].



**Gambar 1.** Arsitektur Jaringan RNN

### Long Short-Term Memory

Memori jangka pendek panjang atau yang biasanya disebut Long Short-Term Memory (LSTM) adalah arsitektur jaringan RNN yang dirancang untuk menghindari kendala buruk RNN pada masalah memori jangka panjang. LSTM merupakan evolusi dari arsitektur RNN, pertama kali diusulkan oleh Hochreiter & Schmidhuber pada tahun 1997. Hingga penelitian ini dilakukan, banyak peneliti yang terus mengembangkan arsitektur LSTM di berbagai bidang seperti pengenalan suara dan prediksi [6].



Gambar 2. Arsitektur Jaringan LSTM

### Natural Language Processing

Pemrosesan bahasa alami salah satu bidang kecerdasan buatan yang bertujuan untuk mengelola bahasa alami. Bahasa alami adalah bahasa yang biasa digunakan manusia dalam percakapan sehari-hari. Agar mesin dapat memahami bahasa alami, perlu menggunakan *preprocessing* kata untuk membuat mesin memahami arti kalimat. Melalui proses NLP, mesin dapat memahami respon yang diberikan oleh manusia [7].

### Preprocessing

*Preprocessing* adalah sebuah metode untuk membersihkan dataset komentar yang tidak terstruktur menjadi data yang terstruktur. Biasanya data *preprocessing* dilakukan dengan cara menghilangkan data yang tidak sesuai atau mengubahnya menjadi bentuk yang lebih sederhana dengan tujuan untuk menyatukan dan memperkecil ukuran kata untuk mempermudah proses data latih [8]. Adapun proses *preprocessing* yang dilakukan yaitu:

*Cleansing, Tokenizing, Case folding, Stopword removal, Stemming.*

### Confusion Matrix

Confusion matrix adalah teknik machine learning yang berisi informasi tentang keadaan sebenarnya dari data dan hasil prediksi dari klasifikasi menggunakan model atau metode klasifikasi yang digunakan. Kinerja klasifikasi dievaluasi menggunakan data dalam matriks yang berisi data aktual, yang merupakan hasil prediksi dari metode klasifikasi yang digunakan [1].

### Cross Validation

Cross validation atau validasi silang adalah metode pengambilan sampel ulang yang digunakan untuk mengevaluasi model pembelajaran mesin menggunakan bagian data yang berbeda untuk menguji dan melatih model pada iterasi yang berbeda. Metode ini bisa mengindikasikan seberapa baik model ini memprediksi data yang tidak pernah dilihat sebelumnya sehingga pengujian ini membuat model tidak overfitting [9]. Nilai-nilai *confusion matrix* akan digunakan untuk mencari nilai *accuracy, recall, precision* dan *f1 score* pada hasil kasifikasi nantinya.

## METODE PENELITIAN

### Alat dan Bahan

#### a. Perangkat Keras (Laptop)

- 1 Unit Laptop merk HP
- RAM 16GB DDR 4
- 1TB HDD
- 250GB SSD
- Nvidia GeForce GTX 1660 TI with Max-Q Design 6GB
- Intel Core i7 9<sup>th</sup> Gen

#### b. Perangkat Lunak

- Windows 10 Home 64bit
- Google Chrome Version 95.0.4638.69
- Visual Studio Code
- Anaconda Navigator 4.9.2
- Jupyter Notebook 6.0.1
- Python 3.7.7

- Ms Excel 2016
- VOSviewer 1.6.16

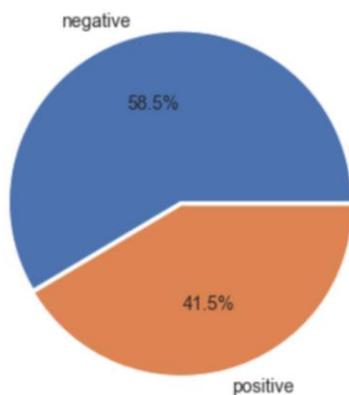
### Teknik Pengumpulan Data

Adapun jenis pengumpulan data yang digunakan dalam penelitian ini sebagai berikut :

1. Observasi (*Observation*)
2. Studi Pustaka (*Literature*)

### Dataset

Dataset yang digunakan pada penelitian ini berjumlah **3.558** data komentar yang telah melalui proses *preprocessing* dan pemberian label dengan label yang digunakan hanya label sentimen positif dan negatif. Untuk jumlah data positif yaitu **41.5%** dari keseluruhan data dan **58.5%** data dengan label negatif. Berikut ilustrasi dari grafik data yang akan digunakan.



**Gambar 3.** Grafik Dataset

### Hyperparameter

Tahap pembuatan model merupakan penyusunan arsitektur model yang dibuat dengan memilih jenis-jenis hyperparameter dan parameter tambahan yang akan digunakan. Peneliti menggunakan library hyperparameter tuning dari library sklearn model\_selection yaitu GridSearchCV. Library ini digunakan untuk mencari hyperparameter terbaik untuk model dari dataset yang akan digunakan.

Dari hasil perhitungan menggunakan library GridSearchCV menggunakan dataset training, menghasilkan hyperparameter terbaik dapat dilihat pada tabel dibawah.

**Tabel 1.** Best Hyperparameter.

Parameter	Value
Epoch	10
Batch size	128
Dropout rate	0.2
Embedim	64
Hiden unit	32
Optimizer	Adam
Learning rate	0.001

## HASIL DAN PEMBAHASAN

### Training dan Testing Model

Pada tahap training dan testing akan menggunakan metode *cross validation* untuk mencari hasil model, dimana perulangan atau *fold* yang akan digunakan yaitu berjumlah 5 perulangan. Dataset akan dimasukan dalam model yang telah dibuat dengan menggunakan *best hyperparameter* dan *cross validation* dan menghasilkan hasil berupa nilai akurasi pada tiap-tiap fold dan akan dicari nilai rata-rata pada seluruh perulangan. Berikut tabel dari hasil model.

**Tabel 2.** Perulangan Nilai *Accuracy*.

Fold	Nilai
1	0.89451477
2	0.89295775
3	0.8915493
4	0.88028169
5	0.88732394

**Tabel 3.** Perulangan Nilai *Precision*.

Fold	Nilai
1	0.89025975
2	0.8884528
3	0.88972156
4	0.87741935
5	0.88463874

**Tabel 4.** Perulangan Nilai *Recall*.

Fold	Nilai
1	0.89086634
2	0.88932412
3	0.88468303
4	0.81887067
5	0.88527501

**Tabel 5.** Perulangan Nilai *f1 Score*.

Fold	Nilai
1	0.89086634
2	0.88888065
3	0.88696395
4	0.87809928
5	0.88494944

**Nilai Mean dan Standar Deviasi Model**

Nilai rata-rata dan standar deviasi atau nilai simpangan baku perlu dicari untuk mengetahui seberapa bagus hasil dari model diambil dari seluruh perulangan atau sebaran data yang dilakukan, makadari itu perlu dicari nilai rata-rata dan nilai simpangan baku. Berikut tabel dari hasil rata-rata dan standar deviasi dari model yang telah melalui proses training, dapat dilihat pada tabel berikut.

**Tabel 6.** Mean dan Standar Deviasi Model.

Nilai	Mean	Stdev
Accuracy	0.88932	0.00571
Precision	0.88609	0.00532
Recall	0.88593	0.00485
f1 Score	0.88595	0.00490

**SIMPULAN**

Dari hasil penelitian analisis sentimen EDOM yang berjumlah **3.558** data yang telah melalui proses *prerocessing* pada tiap data, dapat diimplementasikan dalam analisis sentimen menggunakan model LSTM. Model yang dibentuk merupakan hasil dari perhitungan menggunakan *hyperparameter* terbaik menggunakan *library GridSearchCV* dan digunakan untuk model training yang menghasilkan nilai confusion matrix terbaik yaitu dari skenario label data class 2 yang tidak menggunakan label sentimen netral yaitu dengan nilai rata-rata akurasi sebesar **88.93%**, presisi **88.60%**, recall **88.59%** dan F1 score **88.59%**, dan nilai standar deviasi akurasi **0.5%**, presisi **0.5%**, recall **0.4%** dan F1 score **0.4%**. Dari hasil tersebut dapat ditarik kesimpulan model yang dibuat sudah cukup baik untuk melakukan prediksi. Untuk penelitian kedepannya dapat mengimplementasikan proses *preprocessing* yang lebih *advance* agar model mendapatkan hasil akurasi yang lebih baik lagi.

**DAFTAR PUSTAKA**

[1] B. U. Manalu, "Kinerja Deep Learning Dalam Analisis Sentimen," *Kinerja Deep Learning Dalam Analisis Sentimen*, vol. I, p. 5, 2020.

[2] Lian Ardiani, Herry Sujaini , Tursina, "Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak," *Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak*, vol. II, p. 184, 2020.

[3] U. Indonesia, "Edom," Universitas Indonesia, 2007. [Online]. Available: <https://edom.ui.ac.id/login.php>. [Accessed

- 31 July 2021].
- [4] R. N. Fuadiyati, "Analisis Requirement Sistem Evaluasi Dosen Oleh Mahasiswa (EDOM) Online Uin Syarif Hidayatullah Dengan Goal Oriented Requirement Engineering (Gore) Model," *Analisis Requirement Sistem Evaluasi Dosen Oleh Mahasiswa (EDOM) Online Uin Syarif Hidayatullah Dengan Goal Oriented Requirement Engineering (Gore) Model*, vol. II, p. 10, 2018.
- [5] I. Habibie, "Identifikasi Judul Berita Clickbait Berbahasa Indonesia Dengan Algoritma Long Short Term Memory (LSTM) Recurrent Neural Network," *Identifikasi Judul Berita Clickbait Berbahasa Indonesia Dengan Algoritma Long Short Term Memory (LSTM) Recurrent Neural Network*, vol. II, p. 8, 2018.
- [6] Muhammad Wildan Putra Aldi, Jondri, Annisa Aditsania, "Analisis dan Implementasi Long Short Term Memory Neural Network untuk Prediksi," *Analisis dan Implementasi Long Short Term Memory Neural Network untuk Prediksi*, vol. II, p. 2, 2018.
- [7] F. Syadid, "Analisis Sentimen Komentar Netizen Terhadap Calon Presiden Indonesia 2019 Dari Twitter Menggunakan Algoritma Term Frequency-Invers Document Frequency (Tf-Idf) Dan Metode Multi Layer Preceptron (MLP) Neural Network," *Analisis Sentimen Komentar Netizen Terhadap Calon Presiden Indonesia 2019 Dari Twitter Menggunakan Algoritma Term Frequency-Invers Document Frequency (Tf-Idf) Dan Metode Multi Layer Preceptron (MLP) Neural Network*, vol. II, pp. 10-13, 2019.
- [8] A. Syakuro, "Analisis Sentimen Masyarakat Terhadap E-Commerce Pada Media Sosial Menggunakan Metode Naive Bayes Classifier (NBC) Dengan Seleksi Filer Information Gain (IG)," *Analisis Sentimen Masyarakat Terhadap E-Commerce Pada Media Sosial Menggunakan Metode Naive Bayes Classifier (NBC) Dengan Seleksi Filer Information Gain (IG)*, vol. II, p. 10, 2017.
- [9] N. SEPTIANA, "SISTEM KLASIFIKASI CITRA MAKANAN MENGGUNAKAN REPRESENTASI ANTI TEXTONS DAN K-NEAREST NEIGHBOUR," *SISTEM KLASIFIKASI CITRA MAKANAN MENGGUNAKAN REPRESENTASI ANTI TEXTONS DAN K-NEAREST NEIGHBOUR*, vol. II, p. 18, 2017.