



# Penerapan LSTM dalam Analisis Sentimen Berbasis Lexicon untuk Meningkatkan Sistem Pemantauan Citra PLN di Platform Digital

Angela Lisanthoni<sup>1</sup>, Ellexia Leonie Gunawan<sup>2</sup>, Chelsea Ayu Adhigiadany<sup>3</sup>, Dwi Arman Prasetya<sup>4</sup>

<sup>1, 2, 3, 4</sup>Sains Data, Universitas Pembangunan Nasional “Veteran” Jawa Timur

<sup>1</sup>[21083010032@student.upnjatim.ac.id](mailto:21083010032@student.upnjatim.ac.id)

<sup>2</sup>[21083010027@student.upnjatim.ac.id](mailto:21083010027@student.upnjatim.ac.id)

<sup>3</sup>[21083010028@student.upnjatim.ac.id](mailto:21083010028@student.upnjatim.ac.id)

<sup>4</sup>[arman.prasetya.sada@upnjatim.ac.id](mailto:arman.prasetya.sada@upnjatim.ac.id)

Corresponding author email: [21083010032@student.upnjatim.ac.id](mailto:21083010032@student.upnjatim.ac.id)

**Abstract:** In the digital era, business competition is increasingly fierce, and digital media has become an important platform for companies to promote products and understand public opinion. This research carries out sentiment analysis to evaluate public opinion towards the PLN during January 2024 with the aim of helping brands identify public sentiment towards their brand, so that they can take the right steps to maintain the brand image in the eyes of consumers. Data is collected from various digital media platforms and labeled using a sentiment lexicon method created specifically to suit the data. When a word is identified as a positive word, the value of the sentence increases by one and vice versa, then all the values of the words in the sentence will be totaled. If the total is positive, then the sentence has a positive sentiment. If the total is negative, then the sentence has a negative sentiment. If the total is 0, then the sentence has a neutral sentiment. Of the 49,432 data analyzed, 16.64% of sentences had negative sentiments, 36.69% were positive, and 26.43% were neutral. In addition, a Long Short-Term Memory (LSTM) based deep learning model with the TensorFlow framework is used to classify sentiment automatically. The LSTM model shows 92% accuracy, with an average precision of 91.67%, recall of 92%, and f1-score of 91.33%.

**Keywords:** Sentiment Analysis, TensorFlow, LSTM

**Abstrak:** Di era digital, persaingan bisnis semakin ketat, dan media digital menjadi platform penting bagi perusahaan untuk mempromosikan produk serta memahami opini publik. Penelitian ini melakukan analisis sentimen untuk mengevaluasi opini publik terhadap brand PLN selama Januari 2024 dengan tujuan untuk membantu brand dalam mengidentifikasi sentimen publik terhadap brand mereka, sehingga dapat mengambil langkah yang tepat untuk mempertahankan citra brand di mata konsumen. Data dikumpulkan dari berbagai platform media digital dan dilabelkan menggunakan metode lexicon sentimen yang dibuat secara khusus menyesuaikan data. Ketika kata teridentifikasi kata positif, maka nilai kalimat bertambah satu dan sebaliknya, kemudian seluruh nilai kata dalam kalimat akan ditotal. Jika total bernilai positif, maka kalimat tersebut bersentimen positif. Jika total bernilai negatif, maka kalimat tersebut bersentimen negatif. Jika total bernilai 0, maka kalimat tersebut bersentimen netral. Dari 49.432 data yang dianalisis, 16,64% kalimat bersentimen negatif, 36,69% positif, dan 26,43% netral. Selain itu, model deep learning berbasis Long Short-Term Memory (LSTM) dengan framework TensorFlow digunakan untuk mengklasifikasikan sentimen secara automasi. Model LSTM menunjukkan akurasi 92%, dengan rata – rata precision 91,67%, recall 92%, dan f1-score 91,33%.

**Kata kunci:** Analisis Sentimen, TensorFlow, LSTM

## I. PENDAHULUAN

Di era digital, persaingan bisnis menjadi semakin sengit. Banyak perusahaan berusaha untuk meningkatkan minat beli konsumen melalui berbagai macam promosi, salah satunya adalah media online atau yang dikenal dengan digital Marketing. Strategi digital marketing diperkirakan akan berkontribusi sebesar 78% terhadap proses pemasaran produk dari suatu brand [1]. Media online seperti Twitter, Instagram, Youtube, Facebook, Tiktok, dan blog menjadi ladang promosi bagi



perusahaan. Hal ini dimulai karena perusahaan menyadari betapa besarnya dampak media sosial bagi penjualan produk maupun jasa. Tak heran, banyak perusahaan yang menunjukkan kekreatifannya dalam pembuatan iklan di media sosial, serta berusaha membangun citra perusahaan agar semakin baik dan dikenal masyarakat luas.

Berdasarkan penelitian yang dilakukan oleh Adelia, dkk, terdapat hubungan positif yang signifikan antara citra brand dengan keputusan pembelian dan niat pembelian [2]. Dengan begitu, mempertahankan citra perusahaan di media sosial menjadi sangat penting. Media online saat ini tidak hanya sebagai media promosi, tetapi juga sebagai media penyaluran pendapat atau kritik dari masyarakat terhadap suatu *brand* atau perusahaan. Hal inilah yang dapat menjadi salah satu tolak ukur citra suatu *brand* tersebut baik atau tidak. Karena, jika hanya aktif mempromosikan *brand* tanpa menghiraukan aspirasi masyarakat, maka *brand* tersebut tidak akan mampu bersaing dengan *brand* lainnya. Para pelaku bisnis wajib hukumnya untuk menjalin hubungan kerja sama yang baik dengan media massa agar mendapat kemudahan dalam menyebarkan berita-berita positif dari brand untuk mempertahankan citra perusahaan [3].

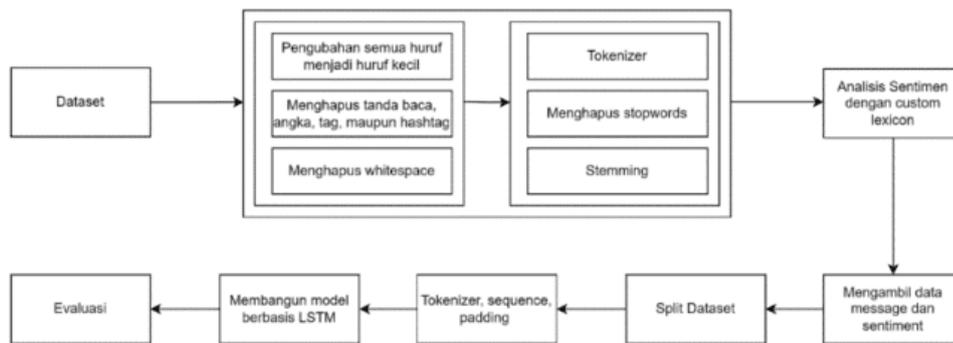
Analisis sentimen merupakan suatu teknik mengekstrak data dari suatu media untuk mendapatkan informasi mengenai sentimen bernilai positif, negatif, ataupun netral. Dalam permasalahan ini, analisis sentimen menjadi alat yang sangat bermanfaat untuk membantu para *brand* dalam memahami cara pandang konsumen terhadap produknya [4]. Analisis sentimen tidak hanya membantu mereka untuk menangkap opini publik, tetapi juga dapat memberikan wawasan untuk evaluasi dan meningkatkan pelayanan publik [5]. Tidak hanya itu, sentimen analisis juga dapat memberikan informasi terkait tanggapan masyarakat terhadap suatu *brand* baik yang bersifat positif, negatif, maupun netral [6]. Tanggapan masyarakat yang teridentifikasi negatif dapat menjadi bahan evaluasi *brand* untuk menyelesaikan keresahan Masyarakat [7]. Adanya tanggapan masyarakat yang positif dapat menjadi acuan dalam aspek-aspek yang disukai konsumen dan terus memberikan yang terbaik. Sedangkan adanya tanggapan masyarakat yang bernilai netral dapat menjadi potensi untuk menggali ide-ide baru dari *feedback* yang diberikan konsumen dan meningkatkan kualitas produk juga layanannya. Dengan pemahaman ini, *brand* dapat mengambil langkah-langkah yang sesuai untuk memperbaiki, mempertahankan, atau meningkatkan citra mereka di depan konsumen.

Dalam mengembangkan sistem pada analisis sentimen, terdapat banyak metode yang bisa digunakan. Syafrizal, dkk melakukan sebuah penelitian berupa membandingkan Algoritma *Naïve Bayes* dengan *K-Nearest Neighbor* dan didapatkan hasil bahwasanya *Naïve Bayes* lebih baik dibandingkan dengan *K-Nearest Neighbor* dengan akurasi sebesar 77,69% [8]. Selain itu, Auliya, dkk melakukan sebuah penelitian berupa membandingkan algoritma LSTM dengan *Naïve Bayes* untuk analisis sentimen. Dalam penelitian tersebut didapatkan hasil bahwasanya algoritma LSTM memiliki kinerja lebih baik dibandingkan dengan *Naïve Bayes* dengan nilai akurasi sebesar 83,33% dan *Naïve Bayes* sebesar 82% [9]. Penelitian lain juga dilakukan oleh Laina, dkk yaitu analisis sentimen menggunakan arsitektur LSTM terhadap fenomena citayam *fashion week* dan didapatkan nilai akurasi model yang cukup baik, yaitu sebesar 88% [10]. Pada tahun 2022, Tri, dkk juga melakukan sebuah analisis layanan pelanggan PT.PLN berdasarkan media sosial twitter dengan menggunakan metode *Naïve Bayes Classifier*. Penelitian tersebut mendapatkan hasil yang cukup memuaskan, yaitu dengan nilai akurasi sebesar 88,24% [11]. Zaini, dkk melakukan kajian terhadap fenomena covid-19 di Twitter menggunakan word embedding dan LSTM dan didapatkan nilai akurasi sebesar 81% [12]. Serta Karlina, dkk juga melakukan kajian analisis sentimen terhadap review film menggunakan

LSTM dengan membandingkan dua optimizer yaitu Adam optimizer mendapatkan akurasi 77.11% dan RMSprop optimizer mendapatkan akurasi 80.07% [13].

Adapun tujuan dari penelitian ini yaitu untuk membantu *brand* dalam mengidentifikasi dan memahami sentimen publik terhadap *brand* mereka, sehingga mereka dapat mengambil langkah-langkah yang tepat untuk memperbaiki atau mempertahankan citra *brand* di mata konsumen. Melalui analisis sentimen ini, *brand* dapat lebih proaktif dalam menanggapi isu-isu yang muncul dan memanfaatkan *feedback* dari konsumen untuk meningkatkan strategi pemasaran dan pelayanan mereka. Dengan demikian, *brand* dapat terus membangun citra yang positif dan mendorong keputusan pembelian yang lebih kuat dari konsumen. Berdasarkan latar belakang permasalahan dan tujuan penelitian, dilakukan kajian dengan judul “Penerapan LSTM dalam Analisis Sentimen Berbasis Lexicon untuk Meningkatkan Sistem Pemantauan Citra PLN di Platform Digital”.

## II. METODE PENELITIAN



Gambar 1. Alur proses penelitian

- *Dataset*

*Dataset* yang digunakan adalah data hasil *crawling* dari beberapa sosial media seperti Twitter, Instagram, Facebook, Youtube, Tiktok, *blog*, dan *News* untuk brand PLN selama periode Januari 2024 dengan ukuran data yang diperoleh adalah 88.690 baris dan 29 kolom. Kolom yang digunakan hanya kolom “message” yang menyimpan data teks hasil *crawling* sehingga ukuran data menjadi 88.690 baris dan 1 kolom. Data didapatkan dari perusahaan PT. Ionesia Solusi Data atau Ivosights.

- *Preprocessing Data*

Beberapa proses *preprocessing data* dilakukan untuk memastikan data yang digunakan dipahami oleh sistem dan mendapatkan hasil yang maksimal. Beberapa proses tersebut diantaranya memastikan tidak ada data N/A atau *null*, menghapus nilai duplikasi, *case folding*, menghapus tanda baca, menghapus angka, menghapus *whitespace*, menghapus jenis *tag* (@), menghapus *hashtag* (#), menghapus *link* yang ada dalam teks, dan lain sebagainya. Data kemudian dilakukan tokenizer, menghapus stopwords, beserta stemming

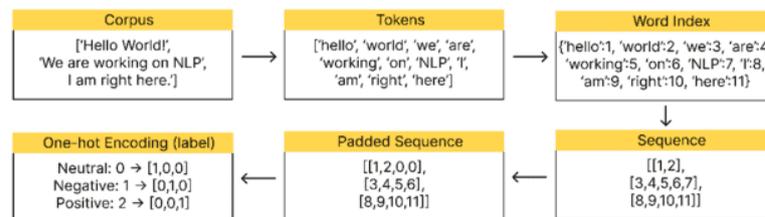
- Analisis Sentimen

Analisis Sentimen adalah salah satu teknik *Natural Language Processing* (NLP) di bidang pembelajaran mesin yang digunakan untuk mengekstrak sentiment maupun emosi dalam teks. Analisis sentimen dapat menentukan apakah sebuah teks mengandung emosi positif, negatif, atau netral. Salah satu pendekatan analisis sentimen adalah menggunakan *Lexicon*. Metode ini

mengacu pada nilai sentimen pada tiap kata dalam kalimat. *Lexicon* adalah suatu *database* yang berisi teks atau *keyword* yang menentukan suatu kata tergolong dalam sentimen positif atau negatif sehingga skor sentimen tiap kalimat memiliki skala dari negatif, nol atau netral, hingga positif [14]. Metode *Lexicon* digunakan untuk melakukan pelabelan pada data sehingga data yang digunakan dalam model LSTM telah terlabeli berdasarkan sentimennya. sehingga data yang digunakan menjadi hanya 49.432 baris dan 6 kolom yaitu kolom “message”, “token”, “full\_text\_token”, “full\_text\_token\_WSW”, “full\_text\_token\_normalized” dan “sentiment\_score”.

- Persiapan data untuk Model *Deep Learning*

Data yang digunakan untuk pembuatan model LSTM hanya kolom “message” dan “sentiment\_score”. Kemudian data dipersiapkan dengan melakukan *split* dataset, tokenisasi, *sequence*, dan *padding*. *Split dataset* bertujuan untuk membagi data menjadi data latih dan data uji. Tokenisasi adalah proses mengubah teks menjadi token-token yang lebih kecil dan mengkonversinya menjadi vektor yang merepresentasikan. Tokenisasi dilakukan kembali dikarenakan menyesuaikan dengan format tokenizer di *framework* TensorFlow. *Sequence* digunakan untuk mengonversi token-token tersebut menjadi urutan numerik yang dapat diproses oleh model. *Padding* dilakukan untuk memastikan bahwa semua urutan memiliki panjang yang sama, sehingga dapat diproses secara *batch* oleh model. Dilakukan juga *one-hot encoding* pada data label yaitu proses pengubahan label menjadi vektor. One-hot encoding perlu dilakukan untuk menghindari kesalahan persepsi model terhadap label dan mempertahankan sifat diskrit dari data kategorikal.



Gambar 2. Ilustrasi proses persiapan data

- Membangun model berbasis LSTM

Model *deep learning* dibangun menggunakan *framework* TensorFlow dengan algoritma *Long Short-Term Memory* (LSTM). Algoritma LSTM dipilih karena kemampuannya dalam menangani data berurutan dan mengenali pola jangka panjang dalam data teks, yang sangat penting dalam analisis sentimen. Arsitektur LSTM dibangun dengan lapisan *input*, lapisan *output*, dan lapisan *hidden* dimana lapisan *hidden* memiliki dengan tiga gerbang utama yaitu *input gate*, *forget gate*, dan *output gate*. *Input gate* berfungsi untuk mengontrol seberapa banyak informasi disimpan. *Forget gate* digunakan untuk mengontrol sejauh mana nilai ditetapkan dalam sel memori. Dan *Output Gate* mengontrol seberapa banyak nilai dalam sel memori yang digunakan untuk menghitung hasil *output*-nya [15]. *Layer* LSTM akan digunakan selama pembuatan model disertai layer lain seperti *embedding layer*, *dense layer*, *global average pooling layer*, dan *dropout layer*.

- Evaluasi Model LSTM



Evaluasi model LSTM dilakukan menggunakan *confusion matrix*, yang membantu menilai performa model dengan mengukur *precision*, *recall*, *F1-score*, dan akurasi [11]. *Precision* menunjukkan seberapa tepat model dalam mengidentifikasi kelas positif dari keseluruhan prediksi positif. *Recall* mengukur seberapa baik model dalam menemukan semua contoh kelas positif dari total kelas positif yang sebenarnya. *F1-score* adalah perbandingan rata - rata dari *precision* dan *recall*, yang memberikan gambaran keseimbangan antara keduanya. Akurasi menunjukkan persentase prediksi yang benar dari keseluruhan prediksi yang dilakukan oleh model.

### III. HASIL DAN PEMBAHASAN

#### III.1. Pre-Processing Data

Data dikumpulkan dari berbagai media digital seperti Twitter, Instagram, Youtube, Facebook, Tiktok, News, dan Blog dalam rentang bulan Januari 2024 dengan beberapa kata kunci yakni “PLN”, “Perusahaan Listrik Negara”, “energiterbarukan”, “plnterdepantujujebt”, “plnmobile”, “gelagarcuan”, “harbelnasPLN”, “promoPLN”, “PLNtambahdaya”, “PLNuntukIndonesia”, “unleashingenergyandbeyond”, “powerbeyondgenerations”, “semuamakinmudah” dan “PLN monopoli” dengan jumlah 88.690 data. Untuk memastikan data sesuai, akan dilakukan *text pre-processing* pada tahap awal menggunakan *Google Colab* yang terdiri dari:

- *Case Folding*

Tahap pertama yang dilakukan adalah mengubah semua huruf kapital menjadi huruf tidak kapital untuk menyetarakan data yang sama.

*Tabel 1. Hasil Case Folding*

index	Sebelum	Sesudah
0	Electrizen, Selamat Tahun Baru 2024!\nRaih mom...	electrizen, selamat tahun baru 2024!\nraih mom...
1	Segenap Dewan Komisaris, Direksi, dan Karyawan...	segenap dewan komisaris, direksi, dan karyawan...
2	Electrizen, Selamat Tahun Baru 2024! Raih mome...	electrizen, selamat tahun baru 2024! raih mome...
3	Electrizen, Wakil Menteri BUMN Kartika Wirjoat...	electrizen, wakil menteri bumn kartika wirjoat...
4	Segenap Dewan Komisaris, Direksi, dan Karyawan...	segenap dewan komisaris, direksi, dan karyawan..

- Penghapusan Duplikat Data dan Nilai Null

Tahap kedua yang dilakukan adalah menghapus data yang terduplikasi beserta nilai *null* jika ditemukan. Sehingga total data yang digunakan dalam proses analisis menjadi 49.432 data. Metode yang digunakan menggunakan fungsi bawaan dari python yaitu “*drop\_duplicates()*” dan “*isnull()*” pada kolom “*message*”.

- *Punctuation and Whitespace Removal*

Tahap ketiga adalah melakukan penghapusan tanda baca termasuk *tag* (@), *link*, *hashtag* (#) dan *whitespace* yang berlebih. Selain itu, juga dilakukan penghapusan angka sehingga data hanya berisi data yang bersih dan siap digunakan untuk dilakukan analisis sentimen. Metode yang digunakan adalah membuat fungsi sendiri menggunakan “*def*” beserta kombinasi dengan library “*re*” atau regular expressions untuk mencari pola yang ingin dihapus.



- *Tokenizing*

Tahap keempat adalah membagi kata – kata dalam kalimat menjadi satuan kata yang terpisah namun masih tersusun dalam kalimat yang sama. Hal ini dilakukan untuk memudahkan proses pembobotan tiap kata yang ada dalam kalimat. Dilakukan juga perhitungan frekuensi untuk tiap kata

*Tabel 2. Hasil Tokenizing*

index	message	Token	full_text_token
0	electrizen selamat tahun baru raih momentum be...	[electrizen, selamat, tahun, baru, raih, momen...	{'electrizen': 1, 'selamat': 1, 'tahun': 3, 'b...
1	segenap dewan komisaris direksi dan karyawan p...	[segenap, dewan, komisaris, direksi, dan, kary...	{'segenap': 1, 'dewan': 1, 'komisaris': 1, 'di...
2	electrizen wakil menteri bumh kartika wirjoatm...	[electrizen, wakil, menteri, bumh, kartika, wi...	'electrizen': 1, 'wakil': 1, 'menteri': 1, 'b...
3	tidak ada yang bisa kembali ke masa lalu untuk...	[tidak, ada, yang, bisa, kembali, ke, masa, la...	{'tidak': 1, 'ada': 1, 'yang': 4, 'bisa': 1, '...
4	kesuksesan dan prestasi adalah bagian dari keh...	[kesuksesan, dan, prestasi, adalah, bagian, da...	{'kesuksesan': 1, 'dan': 3, 'prestasi': 1, 'ad...

- *Stopword Removal*

Tahap kelima adalah menghapus kata – kata umum yang tidak memiliki manfaat pembobotan pada kalimat sebagai contoh “dan”, “tapi”, “kamu”, dan lain sebagainya.

*Tabel 3. Hasil Stopword Removal*

index	message	Token	full_text_token	full_text_token_WSW
0	electrizen selamat tahun baru raih...	[electrizen, selamat, tahun, baru, raih, momen...	{'electrizen': 1, 'selamat': 1, 'tahun': 3, 'b...	[energi, indonesia, electrizen, selamat, raih,...
1	segenap dewan komisaris direksi dan...	[segenap, dewan, komisaris, direksi, dan, kary...	{'segenap': 1, 'dewan': 1, 'komisaris': 1, 'di...	[nusantara, pln, power, sobat, indonesia, sege...
2	electrizen wakil menteri bumh kartika...	[electrizen, wakil, menteri, bumh, kartika, wi...	'electrizen': 1, 'wakil': 1, 'menteri': 1, 'b...	[pln, kelistrikan, listrik, kesiapan, siaga, m...
3	tidak ada yang bisa kembali ke masa lalu...	[tidak, ada, yang, bisa, kembali, ke, masa, la...	{'tidak': 1, 'ada': 1, 'yang': 4, 'bisa': 1, '...	[semoga, mengubah, berusahalah, menciptakan, i...
4	kesuksesan dan prestasi adalah bagian dari keh...	[kesuksesan, dan, prestasi, adalah, bagian, da...	{'kesuksesan': 1, 'dan': 3, 'prestasi': 1, 'ad...	[kesuksesan, prestasi, kehidupan, seiring, ber...

- *Stemming*

Tahap terakhir adalah *stemming* yang berfungsi untuk mengubah kata yang memiliki imbuhan menjadi kata dasar dari kata tersebut. Sebagai contoh: kata “kesuksesan” berubah menjadi “sukses”.



**Tabel 4. Hasil Stemming**

index	message	Token	full_text_token	full_text_token_WS	full_text_normaliz
0	electrizen selamat tahun...	[electrizen, selamat, tahun, baru,...	{'electrizen': 1, 'selamat': 1, 'tahun',...}	[energi, indonesia, electrizen, selamat, raih,...	[energi, indonesia, electrizen, selamat, raih,...
1	segenap dewan...	[segenap, dewan,...	{'segenap': 1, 'dewan': 1, 'komisaris'i...	[nusantara, pln, power, sobat, indonesia, sege...	[nusantara, pln, power, teman, indonesia, sege...
2	electrizen wakil...	[electrizen, wakil, menteri, i...	{'electrizen': 1, 'wakil': 1, 'menteri':...}	[pln, kelistrikan, listrik, kesiapan, siaga, m...	[pln, listrik, listrik, siap, siaga, malam, si...
3	tidak ada yangi...	[tidak, ada, yang, bisa, ...	{'tidak': 1, 'ada': 1, 'yang': 4,'...}	[semoga, mengubah, berusahalah, menciptakan, i...	[semoga, ubah, usaha, cipta, indah, ...
4	kesukses an dan...	[kesuksesan, dan, prestasi, adalah,...	{'kesuksesan' : 1, 'dan': 3, 'prestasi':...}	[kesuksesan, prestasi, kehidupan, seiring, ber...	[sukses, prestasi, kehidupan, seiring, berlalu...

### III.2. Analisis Sentiment

Analisis sentimen menggunakan metode *Lexicon Sentiment* yakni melabeli kalimat menjadi sentimen tertentu berdasarkan jumlah nilai kata yang dikandung dalam kamus bahasa Indonesia yang digunakan. Ketika kata tersebut berada di kamus *lexicon* positif, maka nilai kalimat bertambah satu dan ketika kata tersebut berada di kamus *lexicon* negatif, maka nilai kalimat berkurang satu yang kemudian seluruh nilai kata dalam kalimat akan ditotal. Jika total bernilai positif, maka kalimat tersebut bersentimen positif. Jika total bernilai negatif, maka kalimat tersebut bersentimen negatif. Jika total bernilai 0, maka kalimat tersebut bersentimen netral. Pada analisis ini kamus *Lexicon* yang digunakan merupakan *customize* dimana terdapat kamus *Lexicon* positif dan kamus *Lexicon* negatif. Dari hasil analisis sentimen didapatkan 18.225 kalimat bersentimen negatif yang dikategorikan 1, 18.141 kalimat bersentimen positif yang dikategorikan 2, dan 13.066 kalimat bersentimen netral yang dikategorikan 0.

**Tabel 5. Kamus Lexicon Positif**

Frasa dalam Kamus Lexicon Positif		
'apik':1	'bermanfaat':1	'brilian':1
'kebersihan':1	'kelayakan':1	'membantu':1
'garcep':1	'terbaik':1	'pinter':1
'perbaiki':1	'cegah':1	'salurkan':1

**Tabel 6. Kamus Lexicon Negatif**

Frasa dalam Kamus Lexicon Negatif		
'bajingan':-1	'frustasi':-1	'ga jelas':-1
'jancok': -1	'mahal': -1	'membengkak': -1
'mampus': -1	'bego': -1	'anjlok': -1
'lambat': -1	'bermasalah': -1	'asu': -1



*Tabel 7. Hasil sentiment analysis berbasis lexicon*

index	message	Token	full_text_token	full_text_token_WSW	full_text_normalized	sentiment_score
0	electrizen selamat tahun...	[electrizen, selamat, tahun, baru,...]	{'electrizen': 1, 'selamat': 1, ...}	[energi, indonesia, electrizen, selamat, raih,...]	[energi, indonesia, electrizen, selamat, raih,...]	2
1	segenap dewan...	[segenap, dewan,...]	{'segenap': 1, 'dewan': ...}	[nusantara, pln, power, sobat, indonesia, sege...]	[nusantara, pln, power, teman, indonesia, sege...]	2
2	electrizen wakil...	[electrizen, wakil, menteri, i...]	{'electrizen': 1, 'wakil': ...}	[pln, kelistrikan, listrik, kesiapan, siaga, m...]	[pln, listrik, listrik, siap, siaga, malam, si...]	2
3	tidak ada yang...	[tidak, ada, yang, bisa, ...]	{'tidak': 1, 'ada': 1, ...}	[semoga, mengubah, berusahalah, menciptakan, i...]	[semoga, ubah, usaha, cipta, indah, ...]	2
4	kesuksesan dan...	[kesuksesan, dan, prestasi, adalah,...]	{'kesuksesan': 1, 'dan': 3, ...}	[kesuksesan, prestasi, kehidupan, seiring, ber...]	[sukses, prestasi, kehidupan, seiring, berlalu...]	2

### III.3. LSTM Model

Data hanya diambil kolom ‘message’ dan ‘sentiment score’ dalam pembuatan model *deep learning* berbasis LSTM. Terdapat beberapa langkah yang dilakukan dimana data akan di *convert* menjadi bentuk *array*, kemudian dilakukan *splitting* dengan *ratio* 80:20, proses *tokenizer*, *sequence*, *padding*, serta *one-hot encoding* untuk label. Proses ini dilakukan untuk data *training* maupun data *testing*. Karena data memiliki 3 kelas, maka data label akan dilakukan *one-hot-encoding* yaitu proses mengubah setiap kategori dalam variabel kategorikal menjadi vektor biner. Setiap kategori unik diwakili oleh vektor yang terdiri dari nol dan satu, dengan satu menunjukkan keberadaan kategori tertentu dan nol menunjukkan ketiadaan kategori tersebut. Sebagai contoh, kalimat pertama bersentimen positif, sehingga menunjukkan [0,0,1] dimana 1 berada di posisi ketiga yang artinya bersentimen positif.

*Tabel 8. Hasil Persiapan Data Training*

Feature	Label
array([[ 13, 110, 596, ..., 0, 0, 0], [ 80, 392, 1751, ..., 0, 0, 0], [ 379, 24, 3388, ..., 0, 0, 0], ..., [4410, 48, 450, ..., 433, 419, 1267], [ 261, 977, 122, ..., 0, 0, 0], [ 39, 83, 1090, ..., 0, 0, 0]], dtype=int32)	array([[1., 0., 0.], [0., 1., 0.], [1., 0., 0.], ..., [1., 0., 0.], [0., 1., 0.], [1., 0., 0.]], dtype=float32)

Dalam membuat model *deep learning*, digunakan *framework TensorFlow sequential* dengan algoritma LSTM. Model yang dibentuk memiliki beberapa komponen layer yang terlihat pada tabel 9.



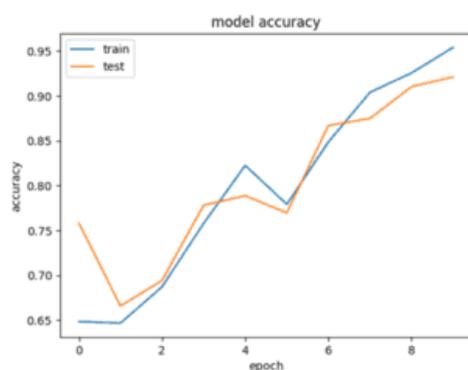
**Tabel 9.** Komponen layer model

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 120, 16)	160000
lstm (LSTM)	(None, 120, 64)	20736
dropout (Dropout)	(None, 120, 64)	0
global_average_pooling1d (GlobalAveragePooling1D)	(None, 64)	0
Dropout_1 (Dropout)	(None, 64)	0
dense (Dense)	(None, 6)	390
dense_1 (Dense)	(None, 3)	21

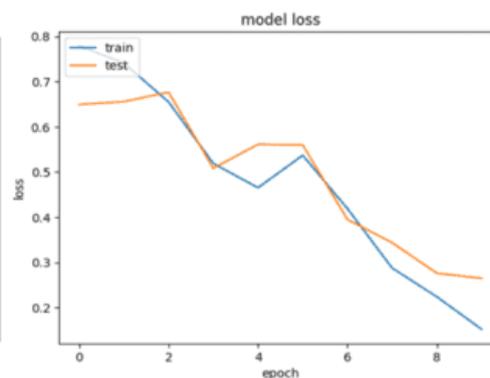
Setelah model dibentuk, akan dilakukan proses training dengan menggunakan fungsi *fit* selama 10 *epochs* untuk mencegah overfitting maupun waktu pemrosesan model yang terlalu lama. Parameter yang ditampilkan berupa loss data training, akurasi data training, loss data test, dan akurasi data test. Hasil Pengujian terlihat pada tabel 10.

**Tabel 10.** Hasil Training Tiap Epoch

Epoch	Loss	Accuracy	val_loss	val_accuracy
1	0.7789	0.6485	0.6493	0.7584
2	0.7409	0.6468	0.6558	0.6659
3	0.6547	0.6872	0.6764	0.6942
4	0.5189	0.7579	0.5083	0.7782
5	0.4653	0.8226	0.5611	0.7887
6	0.5369	0.7792	0.5599	0.7697
7	0.4191	0.8486	0.3951	0.8670
8	0.2876	0.9036	0.3437	0.8748
9	0.2240	0.9256	0.2761	0.9103
10	0.1523	0.9540	0.2650	0.9211



**Gambar 3a.**



**Gambar 3b.**

**Gambar 3.** Grafik hasil training metriks : (3a) *accuracy* (3b) *loss*



### III.4. Evaluasi Model

Dalam evaluasi model, digunakan metode *confusion matrix* untuk menunjukkan tingkat *precision*, *recall*, *f1-score* beserta akurasi. Secara keseluruhan akurasi model mencapai 92% dengan nilai rata – rata *precision*, *recall*, dan *f1-score* secara berturut – turut adalah 91.67%, 92%, dan 91.33%. Nilai *confusion matrix* beserta hasil *classification report* tampak pada tabel 11 dan 12.

**Tabel 11.** Hasil *Confusion Matrix*

		<i>Predicted Class</i>		
		<i>Neutral</i>	<i>Negative</i>	<i>Positive</i>
<i>True Class</i>	<i>Neutral</i>	2284	170	154
	<i>Negative</i>	189	3397	45
	<i>Positive</i>	200	22	3426

**Tabel 12.** Hasil *Classification Report*

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
<i>Class 0</i>	0.85	0.88	0.86
<i>Class 1</i>	0.95	0.94	0.94
<i>Class 2</i>	0.95	0.94	0.94

## IV. KESIMPULAN

Analisis sentimen dilakukan terhadap data digital untuk brand PLN selama periode Januari 2024. Dataset yang digunakan, terdiri dari 88.690 data menjadi 49.432 data setelah proses *pre-processing*, dikumpulkan dari berbagai *platform* sosial media seperti Twitter, Instagram, Facebook, YouTube, TikTok, serta berita dan *blog*. Dalam analisis sentimen menggunakan metode *lexicon sentiment*, ditemukan bahwa 18.225 kalimat (36.87%) bersentimen negatif, 18.141 kalimat (36.7%) bersentimen positif, dan 13.066 kalimat (26.43%) bersentimen netral. Kemudian, dibangun model *deep learning* berbasis *Long Short-Term Memory* (LSTM) menggunakan *framework* TensorFlow. Evaluasi model menggunakan *confusion matrix* menunjukkan akurasi model sebesar 92%, dengan nilai rata-rata *precision*, *recall*, dan *f1-score* masing-masing sebesar 91.67%, 92%, dan 91.33%. Disarankan, mempertimbangkan penggunaan teknik tambahan seperti *fine-tune* model untuk memungkinkan menghasilkan model yang lebih baik.

## UCAPAN TERIMA KASIH

Kami ingin menyampaikan penghargaan yang tulus kepada Tim SENADA atas dedikasi dan kerja keras mereka dalam menyediakan template ini. Kontribusi mereka telah mempermudah proses penulisan artikel ini.

## REFERENSI

1. F. Budi Dewanto, M. Habib Purnama Febrian, M. Raihan Amir, and I. Farida Adi Prawira, “Penerapan SEO Dalam Strategi Pemasaran Perusahaan,” *JURNAL MANEKSI*, vol. 12, no. 4, pp. 709–715, Dec. 2023.
2. A. Trisa and M. R. Roosdhani, “Analisis Hubungan Antara Strategi Pemasaran Media Sosial, Citra Merek, dan Minat Pembelian di Industri Fashion Sweet Mango Jepara,” 2024.



3. D. Nurusyifa, H. Hafiar, and C. Priyatna, “Analisis Sentimen Terhadap Brand Daihatsu Berdasarkan Pemberitaan Online pada Maret 2023,” *JURNAL OF DIGITAL COMMUNICATION SCIENCE*, vol. 2, no. 1, pp. 1–15, Mar. 2024, [Online]. Available: <https://journals.inaba.ac.id/>
4. S. Mutmatinah, Khairunnisa, and Khairunnisa, “Metode Deep Learning LSTM dalam Analisis Sentimen Aplikasi PeduliLindungi,” *Journal of Computers Sciences and Informatics*, vol. 1, pp. 9–19, 2024, doi: 10.34304/scientific.v1i1.231.
5. R. Darman, “ANALISIS SENTIMEN RESPONS TWITTER TERHADAP PERSYARATAN BADAN PENYELENGGARA JAMINAN SOSIAL (BPJS) DI KANTOR PERTANAHAN,” *Widya Bhumi*, vol. 3, no. 2, pp. 113–136, Oct. 2023.
6. L. Ardiani, H. Sujaini, and T. Tursina, “Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak,” *Jurnal Sistem dan Teknologi Informasi (Justin)*, vol. 8, no. 2, p. 183, Apr. 2020, doi: 10.26418/justin.v8i2.36776.
7. K. A. Lubis, M. Theo, A. Bangsa, and A. Yudertha, “ANALISIS SENTIMEN OPINI MASYARAKAT TERHADAP PINDAHNYA IBU KOTA INDONESIA DENGAN MENGGUNAKAN KLASIFIKASI NAÏVE BAYES,” *JURNAL TEKNOINFO*, vol. 18, no. 1, pp. 226–238, 2024, [Online]. Available: <https://ejournal.teknokrat.ac.id/index.php/teknoinfo/index>
8. S. Syafrizal, M. Afdal, and R. Novita, “Analisis Sentimen Ulasan Aplikasi PLN Mobile Menggunakan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, pp. 10–19, Dec. 2023, doi: 10.57152/malcom.v4i1.983.
9. A. I. Rahman, H. Sulistiani, B. H. Miftaq, A. Nurkholis, and Styawati, “Analisis Perbandingan Algoritma LSTM dan Naive Bayes untuk Analisis Sentimen,” *JEPIN: Jurnal Edukasi dan Penelitian Informatika*, vol. 8, no. 2, pp. 299–303, Aug. 2022.
10. L. Farsiah *et al.*, “ANALISIS SENTIMEN MENGGUNAKAN ARSITEKTUR LONG SHORT-TERM MEMORY (LSTM) TERHADAP FENOMENA CITAYAM FASHION WEEK,” *Cyberspace: Jurnal Pendidikan Teknologi informasi*, vol. 6, no. 2, pp. 86–94, Oct. 2022.
11. T. Prasetyo, H. Zakaria, and P. Wiliantoro, “Analisis Layanan Pelanggan PT PLN Berdasarkan Media Sosial Twitter Dengan Menggunakan Metode Naïve Bayes Classifier,” *OKTAL : Jurnal Ilmu Komputer dan Sains*, vol. 1, no. 6, pp. 573–582, Jun. 2022, [Online]. Available: <https://journal.mediapublikasi.id/index.php/oktal>
12. I. G. M. D. W. Wiartha, I. M. A. D. Suarjaya, and N. K. D. Rusjyanthi, “Analisis Sentimen Masyarakat terhadap Brand Lokal di E-Commerce dengan Pendekatan Deep Learning Menggunakan Metode XLNet,” *JITTER-Jurnal Ilmiah Teknologi dan Komputer*, vol. 4, no. 2, pp. 1–12, Aug. 2023.
13. K. S. Witanto, N. A. S. ER, A. E. Karyawatia, I. G. A. G. Kadyanana, I. K. G. Suhartanaa, and L. G. Astuti, “Implementasi LSTM pada Analisis Sentimen Review Film Menggunakan Adam dan RMSprop Optimizer,” *Jurnal Elektronik Ilmu Komputer Udayana*, vol. 10, no. 4, pp. 351–362, May 2022.
14. H. B. Tambunan and T. W. D. Hapsari, “Analisis Opini Pengguna Aplikasi New PLN Mobile Menggunakan Text Mining,” *PETIR*, vol. 15, no. 1, pp. 121–134, Dec. 2021, doi: 10.33322/petir.v15i1.1352.
15. L. Wiranda and M. Sadikin, “PENERAPAN LONG SHORT TERM MEMORY PADA DATA TIME SERIES UNTUK MEMPREDIKSI PENJUALAN PRODUK PT. METISKA FARMA,” *Jurnal Nasional Pendidikan Teknik Informatika*, vol. 8, no. 3, pp. 184–196, Dec. 2019.