

# Analisis Sentimen Pengguna Aplikasi MyPertamina Menggunakan Algoritma *Bidirectional Long Short Term Memory*

Angga Saputra<sup>1</sup>, Rito Cipta Sigitta Hariyono<sup>2</sup>, Nurul Mega Saraswati<sup>3</sup>

<sup>1,3</sup>Fakultas Sains dan Teknologi, <sup>2</sup>Fakultas Ilmu Komputer

<sup>1,3</sup>Universitas Peradaban, <sup>2</sup>Universitas Bhamada Slawi

<sup>1,3</sup>Brebes, <sup>2</sup>Tegal; Indonesia

e-mail: <sup>1</sup>anggaegae@gmail.com, <sup>2</sup>rintocipta13@gmail.com, <sup>3</sup>nurul.mega.s@gmail.com

Diajukan: 11 Juli 2023; Direvisi: 22 September 2023; Diterima: 25 September 2023

## Abstrak

Aplikasi mobile telah menjadi sarana penting bagi perusahaan untuk berinteraksi dengan pengguna dan mengumpulkan umpan balik dari mereka. Salah satu aspek penting dari analisis data aplikasi pengguna adalah analisis sentimen, yang dapat memberikan wawasan pengguna tentang pandangan dan penggunaan aplikasi. Aplikasi MyPertamina yang merupakan layanan e-money juga telah terdaftar dan diawasi oleh Bank Indonesia. MyPertamina berfungsi sebagai sistem pembayaran non tunai. Pada aplikasi MyPertamina dapat memberikan ulasan aplikasi melalui kolom komentar atau memberikan rating dari 1 hingga 5 pada google playstore, namun seringkali pengguna memberikan rating yang tidak sesuai dengan ulasannya sehingga dapat mempengaruhi pelanggan jika ulasannya kurang bagus. Agar mempermudah pengelolaan data ulasan, penelitian ini menggunakan analisis sentiment. Analisis sentimen merupakan suatu metode yang digunakan untuk memahami, mengekstrak, dan mengolah secara otomatis data opini dan tekstual untuk mendapatkan sentimen yang terkandung dalam suatu opini. Penggunaan algoritma *Bidirectional Long Short Term Memory (BiLSTM)* dan *Long Short Term Memory (LSTM)* untuk analisis sentimen dari umpan balik pengguna pada aplikasi myPertamina. Penelitian ini menggunakan dataset yang berisi ulasan pengguna dari aplikasi myPertamina yang dikumpulkan dari platform Google PlayStore. Dataset ini secara otomatis dianotasi menggunakan library transformers untuk mengidentifikasi sentimen dari ulasan sebagai positif atau negatif. Selanjutnya, algoritma BiLSTM dan LSTM diterapkan untuk melakukan analisis sentimen pada dataset yang dianotasi. Hasil eksperimen menunjukkan bahwa kedua algoritma BiLSTM dan LSTM mampu menghasilkan sentimen yang cukup baik pada dataset ulasan pengguna myPertamina. BiLSTM menunjukkan kinerja yang lebih baik dibandingkan dengan LSTM, dengan akurasi 90% dan 86,25% untuk model LSTM.

**Kata kunci:** Analisis sentimen, MyPertamina, *Bidirectional Long Short Term Memory (BiLSTM)*, *Long Short Term Memory (LSTM)*.

## Abstract

Mobile apps have become an important means for companies to interact with users and gather feedback from them. One important aspect of user app data analysis is sentiment analysis, which can provide user insights into the views and usage of the app. MyPertamina application which is an e-money service has also been registered and supervised by Bank Indonesia. MyPertamina functions as a non-cash payment system. The MyPertamina application can provide application reviews through the comments column or provide a rating from 1 to 5 on Google Playstore, but often users provide ratings that do not match the reviews so that they can influence customers if the reviews are not good. In order to facilitate the management of review data, this research uses sentiment analysis. Sentiment analysis is a method used to understand, extract, and automatically process opinion and textual data to get the sentiment contained in an opinion. The use of *Bidirectional Long Short Term Memory (BiLSTM)* and *Long Short Term Memory (LSTM)* algorithms for sentiment analysis of user feedback on the myPertamina application. This research uses a dataset containing user reviews of the myPertamina application collected from the Google PlayStore platform. The dataset is automatically annotated using library transformers to identify the sentiment of the reviews as positive or negative. Next, BiLSTM and LSTM algorithms are applied to perform sentiment analysis on the annotated dataset. The experimental results show that both BiLSTM and LSTM algorithms

are able to generate fairly good sentiment on the myPertamina user review dataset. BiLSTM shows better performance compared to LSTM, with 90% accuracy and 86.25% for the LSTM model.

**Keywords:** Sentiment analysis, MyPertamina application, Bidirectional Long Short Term Memory (BiLSTM), Long Short Term Memory (LSTM).

## 1. Pendahuluan

Pertamina merupakan perusahaan BUMN yang mengelola minyak dan gas bumi terus berupaya meningkatkan pelayanan dengan adopsi teknologi, Salah satunya adalah melalui digitalisasi SPBU dan pembayaran non-tunai menggunakan aplikasi MyPertamina. Digitalisasi SPBU bertujuan untuk memperbaiki pengelolaan data penyaluran, stok, penjualan, dan pembayaran BBM di SPBU. MyPertamina juga memberikan keuntungan seperti efisiensi pembayaran, *point reward*, serta berbagai *event* dan promo menarik [1]. Aplikasi MyPertamina mulai diterapkan pada 1 September 2022 [2] yang bisa didapatkan melalui *App Markets* seperti *Google Playstore* dan *AppStore*. Pengguna aplikasi dapat memberikan distribusi berupa ulasan atau rating bintang dari 1 hingga 5 pada *google playstore*, namun seringkali pengguna memberikan *rating* yang tidak sesuai dengan ulasannya sehingga rating tersebut tidak cukup menggambarkan kualitas aplikasi [3]. Ulasan berbentuk kalimat lebih menggambarkan tanggapan pengguna terhadap aplikasi. Hal ini dapat memungkinkan bahwa ulasan-ulasan tersebut mempengaruhi pengunjung atau pengguna yang akan mengunduh aplikasi MyPertamina bahwa aplikasi tersebut berkualitas atau tidak [4].

Agar mempermudah pengelolaan data ulasan, penelitian ini menggunakan metode analisis sentimen. Analisis sentimen merupakan suatu metode yang digunakan untuk memahami, mengekstrak, dan mengolah secara otomatis data opini dan tekstual untuk mendapatkan sentimen yang terkandung dalam suatu opini [5]. Penelitian sebelumnya telah melakukan studi mengenai analisis sentimen dengan variasi topik yang mencakup ulasan film, twitter, aplikasi di *Google Play Store*, dan *App Store*. Dalam penelitian tersebut, beberapa peneliti telah menggunakan model *machine learning* dan *deep learning*. Sebagai contoh, dalam penelitian mengenai analisis sentimen aplikasi *e-learning Quipper* selama pandemi COVID-19, peneliti menggunakan metode SVM dan *dataset* yang terdiri dari 1000 ulasan *Google Play Store*. Evaluasi kinerja algoritma SVM dengan kernel linier menunjukkan hasil terbaik pada skenario ketiga, di mana data training dan data testing memiliki perbandingan 60:40. Dalam skenario ini, akurasi yang dicapai adalah sebesar 90,25%. [6]. Penelitian selanjutnya mengenai sentimen analisis terhadap aplikasi pada *Google Playstore* menggunakan algoritma *Naïve Bayes* dan Algoritma *Genetika*. Objek penelitian ini adalah 4 aplikasi yang berasal dari *google playstore* yaitu Gojek, Tokopedia, Shopee dan Ruang Guru. Penelitian ini menunjukkan hasil akhir yang sangat baik dimana nilai akurasi lebih dari 90% [7]. Berikutnya penelitian mengenai analisis sentimen ulasan film IMDb menggunakan LSTM hasil dari penelitian ini mendapatkan tingkat akurasi yaitu 89,9% secara keseluruhan penelitian ini cukup baik karena menggunakan jumlah *dataset* 50.000 ulasan namun yang menjadi permasalahan pada penelitian ini *dataset* yang dibuat masih menggunakan proses pelabelan secara otomatis dengan menggunakan rating dimana rating  $\leq 4$  akan dilabeli sebagai kelas negatif dan  $\geq 7$  akan dilabeli positif, *dataset* yang dibuat dengan tidak menggunakan metode analisis teks namun masih menggunakan rating sehingga hasil akhir sulit untuk diketahui perbandingan nilai akurasi dari rating dan metode analisis sentimen dengan data teks [8]. Penelitian-penelitian sebelumnya menggunakan algoritma LSTM yaitu pada analisis sentimen *review* film menggunakan *Adam* dan *RMSprop Optimizer*. Data yang digunakan dalam penelitian ini adalah data sekunder, yaitu data yang sudah tersedia sebelum peneliti mulai melakukan penelitian. Data yang diambil adalah kumpulan dokumen *review* film melalui *dataset* *IMDB Largest Review* yang didapat dari *web IEEE Dataport*. Data yang digunakan adalah 14.000 data *review* film dalam bahasa Inggris, yang terdiri dari 7.000 *review* dengan label positif dan 7.000 *review* dengan label negatif dalam format *.csv*. Optimasi berupa *Adam* dan *RMSprop* yang mampu menghasilkan akurasi 77,11% untuk optimasi *Adam* dan 80,07% untuk *RMSprop* [9]. Penelitian selanjutnya mengenai analisis sentimen terhadap layanan *grab indonesia* menggunakan *Bidirectional Long Short Term Memory* (Bi-LSTM) dengan menggunakan 5000 *dataset* ulasan pada *google playstore* dan menghasilkan tingkat akurasi sebesar 91% [10]. Penelitian mengenai analisis sentimen ulasan aplikasi MyPertamina juga telah dilakukan dengan menggunakan metode *Random Forest* dan menghasilkan akurasi sebesar 90% [11].

Metode LSTM memiliki kelemahan yaitu hanya menggunakan satu *input* yaitu informasi masa lalu untuk mempelajari informasi hanya dari satu arah yaitu ke depan (*forward*), untuk mengatasi masalah ini, dikenalkan sebuah varian dari LSTM yaitu *Bidirectional Long Short-Term Memory* (Bi-LSTM) yang memiliki dua *input* yaitu *forward* dan *backward*, dengan adanya dua *input* tersebut, sistem dapat

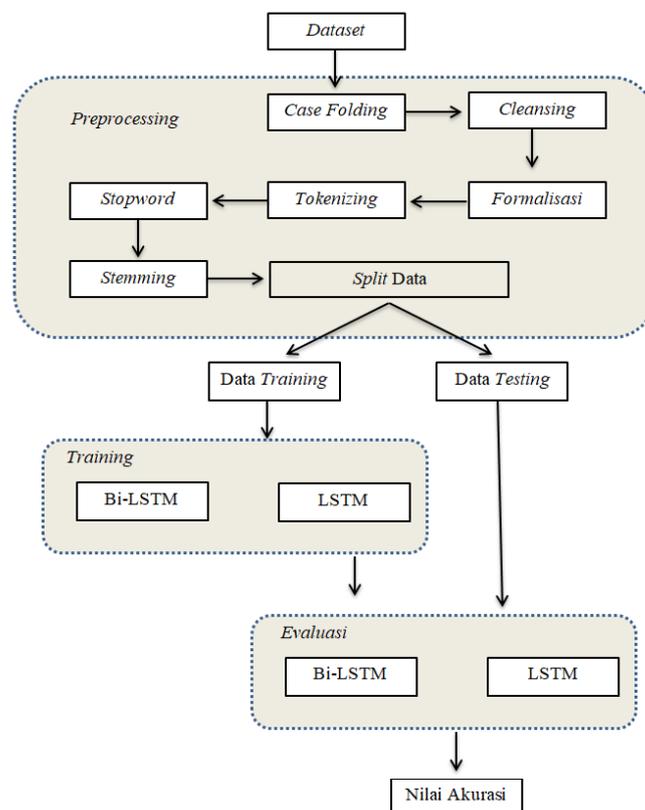
---

*Analisis Sentimen Pengguna Aplikasi MyPertamina Menggunakan Algoritma Bidirectional Long Short Term Memory (Angga Saputra)*

mempelajari informasi dari masa lalu maupun masa depan [12]. Berdasarkan hasil penelitian sebelumnya penelitian ini bertujuan untuk melanjutkan penelitian sebelumnya, melakukan eksperimen model menggunakan Bi-LSTM menggunakan data berbeda yaitu data ulasan pengguna aplikasi MyPertamina yang ada di kolom komentar *Google Playstore* kemudian membandingkan akurasi yang dihasilkan dengan model yang dibangun menggunakan model LSTM untuk mengetahui tingkat akurasi kedua metode tersebut dalam melakukan analisis sentimen ulasan aplikasi MyPertamina. Sehingga penulis memilih judul “Analisis Sentimen Pengguna Aplikasi MyPertamina Menggunakan Algoritma *Bidirectional Long Short Term Memory*”.

## 2. Metode Penelitian

Bagian ini akan membahas tahapan-tahapan yang dilakukan peneliti dalam melakukan penelitian analisis sentimen seperti, pengumpulan data ulasan yang akan digunakan sebagai *dataset* menggunakan teknik *web scraping*, *preprocessing*, *split data*, *training* dan evaluasi LSTM dan Bi-LSTM. Tahapan-tahapan yang dilakukan pada penelitian ini dapat dilihat pada Gambar 1 dibawah ini. Berdasarkan gambar 1, setelah *dataset* berhasil dibuat selanjutnya lakukan proses *preprocessing* yang meliputi *casefolding*, *cleansing*, *formalisasi*, *tokenizing*, *stopword*, *stemming*. Setelah *preprocessing* berhasil dilakukan data akan dibagi menjadi 2 yaitu data *training* dan data *testing*. Setelah data dibagi proses pertama dilakukan proses training model LSTM dan Bi-LSTM kemudian dilakukan evaluasi model menggunakan data *testing* sampai menemukan hasil akhir akurasi masing-masing model.



Gambar 1. Metode Penelitian

### 2.1. Sumber data

Data yang digunakan dalam penelitian ini adalah data yang berasal dari kolom ulasan aplikasi MyPertamina yang tersedia di Google Play Store. Google Play Store adalah aplikasi resmi milik Google yang dapat digunakan oleh perangkat yang menggunakan sistem operasi Android ataupun Web. Penelitian ini menggunakan 6000 data yang diambil dari kolom ulasan aplikasi MyPertamina menggunakan *web scraping*. Data yang diambil merupakan data terbaru yang bersifat *time series*. Proses *scraping* menggunakan *library scraping* milik google dengan *parameter NEWEST* sehingga data yang diambil

---

adalah 6000 data terbaru. Selain mengambil data ulasan peneliti juga mengambil beberapa data seperti author dan atribut lainnya.

## 2.2. Pelabelan Data

Setelah data dikumpulkan melalui proses *web scraping*, data kemudian diberi label sesuai dengan sentimen yang terkandung dalam setiap ulasannya. Dalam penelitian ini, label yang digunakan adalah positif dan negatif, dalam proses pelabelan dilakukan secara otomatis menggunakan *library python* yang disebut dengan Transformers. Proses Pelabelan data dilakukan menggunakan library Transformers dimana sebanyak 6000 data dilakukan pelabelan secara otomatis. Label yang digunakan adalah label positif dan negatif. Dimana label negatif lebih dominan mencapai 4169 data berlabel negatif dan 1831 data berlabel positif.

## 2.3. Preprocessing

Preprocessing atau praproses adalah tahap pertama dari teks mining. Setelah keputusan praproses diambil dan diimplementasikan, fitur (yaitu, variabel) diekstrak dari teks dengan menggunakan teks mining kosakata tertutup, penambahan teks kosakata terbuka, atau keduanya. Kemudian, fitur tersebut digunakan sebagai unit analisis [13]. Data ulasan aplikasi MyPertamina diambil melalui teknik web scraping dari situs Google Play Store dan kemudian disimpan dalam format CSV. Setelah itu, seluruh data diberi label sentimen agar dapat digunakan dalam analisis selanjutnya. Pada tahap *preprocessing*, data akan diproses agar dapat lebih mudah dianalisis dan dimengerti. Berbagai teknik seperti *casefolding*, *cleansing*, *formalisasi*, *tokenizing*, penghapusan *stopwords* dan *stemming* akan diterapkan agar hasil analisis lebih akurat dan efisien.

## 2.4. Data Splitting

Data *Splitting* atau pembagian data merupakan metode yang sering digunakan dalam pembelajaran mesin untuk membagi data menjadi set latih, uji, atau validasi. Pendekatan ini memungkinkan untuk menemukan *parameter hiper-model* dan juga mengestimasi kinerja generalisasi [14]. Dalam penelitian ini data yang sudah dilakukan *preprocessing* kemudian data dibagi sebagai data *testing* dan data *training*. Proses membagi dataset dilakukan secara otomatis menggunakan *python regular expression* dengan menggunakan *library sklearn*.

## 2.5. Modelling

*Modelling* merupakan tahapan dalam membuat model yang sesuai dengan masalah yang akan dipecahkan, dengan memperhatikan aspek performa dan juga kompleksitas pemrosesan data. Sehingga, dapat menghasilkan model yang efektif dan efisien dalam menyelesaikan masalah yang dihadapi.

### a. Word Embedding

Layer pertama pada arsitektur jaringan digunakan sebagai input, dan salah satu teknik yang sering digunakan pada layer pertama adalah *embedding*. *Embedding* adalah teknik vektorisasi teks yang mengubah kata-kata dalam kosakata menjadi vektor bilangan *real kontinu*, sehingga kata-kata dapat diolah dengan algoritma pembelajaran mesin.

### b. Bi-LSTM

Pada penelitian ini menggunakan *Bidirectional LSTM*. *Bidirectional Long Short Term Memory* atau Bi-LSTM adalah model pemrosesan berurutan yang terdiri dari dua LSTM : satu mengambil input dalam arah maju (*forward direction*), dan yang lainnya di arah belakang (*backward direction*). LSTM adalah salah satu tipe dari RNN yang dapat membuat jaringan mampu mempertahankan ketergantungannya dalam jangka panjang antara data pada waktu tertentu dengan beberapa data pada waktu sebelum-sebelumnya. Komponen utama dari lapisan LSTM adalah suatu unit yang disebut blok memori. Gerbang-gerbang pada lstm berfungsi untuk membuat, membaca, dan mengulang kembali informasi pada *cell state*. *Cell state* pada LSTM adalah komponen yang membawa informasi dari suatu unit ke unit lainnya. Dalam menentukan bobot awal LSTM menggunakan nilai acak sedangkan pada penelitian ini menggunakan *default* dari *library keras* yang secara *default* inialisasi bobot awal LSTM pada *library Keras* menggunakan distribusi normal dengan *mean* (rerata) nol dan standar deviasi satu (yaitu *Gaussian distribution with mean 0 and standard deviation 1*). Proses ini sering disebut sebagai "*glorot uniform initialization*" atau "*Xavier initialization*". Hal ini bertujuan untuk menghasilkan bobot awal yang acak namun stabil untuk membantu dalam proses pelatihan jaringan yang lebih baik.

## 2.6. Evaluasi

Evaluasi merupakan proses untuk mengevaluasi kinerja suatu sistem atau model yang telah dibuat. Evaluasi dapat dilakukan dengan menggunakan berbagai metode seperti analisis *Confusion Matrix*, *regresi linear*, indeks kesalahan, dan lain-lain. Hasil dari evaluasi ini digunakan untuk mengevaluasi keakuratan, sensitivitas, dan performa klasifikasi dari sistem atau model tersebut. Evaluasi ini juga digunakan untuk mengevaluasi kemampuan sistem atau model dalam menangani kesalahan dan menemukan sumber kesalahan.

## 2.7. Hasil

Hasil akhir dari penelitian ini adalah mengetahui nilai akurasi dari masing-masing model yang telah dibangun menggunakan model Bi-LSTM dan LSTM. Hasil akhir akurasi didapatkan melalui proses perhitungan menggunakan Rumus *confusion matrix*. Sehingga setelah didapatkan nilai akurasi dari masing-masing model maka akan didapat perbandingan akurasi untuk menentukan model mana yang paling baik.

## 3. Hasil dan Pembahasan

### 3.1. Dataset

Data penelitian ini diperoleh dari kolom ulasan aplikasi MyPertamina di *Google Play*. *Google Play* adalah *platform* resmi Google yang dapat diakses melalui perangkat Android atau web. Pengguna dapat mengunduh aplikasi MyPertamina dan memberikan ulasan. Ulasan tersebut digunakan sebagai sumber data dengan metode *Web Scraping*, yaitu pengambilan dan penyimpanan data dari web. Penelitian ini menggunakan 6000 data terbaru dengan fitur ulasan, *username*, dan *rating*. Data disimpan dalam format *.csv*. Tahap pengolahan data awal meliputi pengelompokan, penentuan kelas target, dan pelabelan sentimen positif dan negatif secara otomatis menggunakan *library Transformers*. *Transformers* adalah *library Python* yang menyediakan berbagai model NLP *state-of-the-art* dan fungsi utilitas untuk memproses teks sekuensial, mengubah teks menjadi representasi numerik, melatih dan mengevaluasi model, serta melakukan *fine-tuning* dengan data yang ada atau baru contoh data yang diambil dapat di lihat pada Tabel 1 dibawah.

Tabel 1. Sampel Hasil Pengambilan Data

<i>username</i>	<i>content</i>	<i>score</i>
Amin Sahidin	Tidak bagus lama lemot mau masukin no yg lwat email kmbali trus ngk akan masuk ujung2	1
Pak Aseplagi	Pelayanan cukup baik dan operatornya ramah	2
Katiman	Mohon untuk dapat mengganti nomor telephone, minimal 90hari sekali	3
Yatna Su	Aplks mudah dipahami, cuma sulit mendapatkan rekomendasi	4
Muhammad Zacky	Jelek kode gk masuk,pake wa Kan bisa juga verifikasinya ribet banget pake sms	5

### 3.2. Preprocessing

Data ulasan aplikasi MyPertamina diambil melalui *web scraping* dari *Google Play Store* dan disimpan dalam format CSV. Seluruh data diberi label sentimen untuk analisis selanjutnya. Namun, karena data teks yang tidak terstruktur, *preprocessing* diperlukan untuk meningkatkan kualitas data. Proses *preprocessing* melibatkan *casefolding*, *cleansing*, formalisasi, *tokenizing*, penghapusan *stopwords*, dan *stemming* agar analisis lebih akurat dan efisien. *Dataset* hasil *preprocessing* dapat dilihat pada Tabel 2.

Tabel 2. Sampel Data Hasil *Preprocessing*

<i>Text</i>	<i>Casefolding</i>	<i>Cleansing</i>	<i>Formalisasi</i>	<i>Tokenizing</i>	<i>Stopword</i>	<i>Stemming</i>
Tidak bagus lama lemot mau masukin no yg lwat email kmbali trus ngk akan masuk ujung2	tidak bagus lama lemot mau masukin no yg lwat email kmbali trus ngk akan masuk ujung2	tidak bagus lama lemot mau masukin lwat email kmbali trus ngk akan masuk ujung	tidak bagus lama lemot mau memasukkan lewat email kembali terus enggak akan masuk ujung	['tidak', 'bagus', 'lama', 'lemot', 'mau', 'memasukkan', 'lewat', 'email', 'kembali', 'terus', 'enggak', 'akan', 'masuk', 'ujung']	['bagus', 'lemot', 'memasukka n', 'lewat', 'email', 'masuk', 'ujung']	['bagus', 'lot', 'masuk', 'lewat', 'email', 'masuk', 'ujung']

### 3.3. Data Splitting

Proses pembagian data *training* dan testing menggunakan *library sklearn* dengan fungsi *train\_test\_split* untuk membagi *dataset* menjadi *subset train* dan *subset test*. *Subset train* digunakan untuk melatih model, sementara *subset test* digunakan untuk menguji kinerja model pada data yang belum dikenal sebelumnya. Dengan menggunakan *train\_test\_split*, data dapat dibagi secara acak menjadi dua *subset* yang *independen*. Setelah *preprocessing*, dari 6.000 ulasan tersisa 4.657. Terdapat ketidakseimbangan data di mana kelas negatif lebih dominan. Untuk mengatasi hal ini, digunakan *library sklearn* untuk melakukan *resampling* data agar kelas menjadi lebih seimbang. Modul *sklearn utils resample* digunakan untuk melakukan *resampling* pada *dataset*, mengatasi ketidakseimbangan kelas, dan memperluas *dataset*. Setelah *resampling*, jumlah data positif dan negatif menjadi 3.863, sehingga total data setelah *resampling* menjadi 7.726 yang kemudian data dibagi 40% untuk testing dan 60% untuk *training*.

### 3.4. Training dan Evaluasi

Setelah data dibagi menjadi dua *subset* kemudian dilakukan proses *training* model, dimana setiap pelatihan akan dilakukan tahap evaluasi. Proses *training* model menguji beberapa *parameter* seperti jumlah *neuron* yang digunakan dan fungsi aktivasi. Hasil evaluasi model menggunakan *parameter neuron* dapat dilihat pada Tabel 3.

Tabel 3. Hasil pengujian menggunakan parameter neuron

Jumlah Neuron	Akurasi Evaluasi
100	0.8961
200	0.8951
300	0.9000
400	0.8861

Hasil uji coba dengan berbagai jumlah neuron menunjukkan bahwa jumlah *neuron* 300 menghasilkan akurasi tertinggi diantara jumlah *neuron* lainnya, yaitu sebesar 0.9000.

Tabel 4. Hasil pengujian fungsi aktivasi

Fungsi Aktivasi	Akurasi Evaluasi
<i>Sigmoid</i>	0.9000
<i>Softmax</i>	0.4972

Berdasarkan hasil pengujian yang tercatat pada Tabel 4 fungsi aktivasi *sigmoid* memiliki akurasi tertinggi sebesar 0.9000 dibandingkan dengan fungsi aktivasi *softmax* yang hanya mencapai 0.4972. Model kemudian akan diuji dengan data *testing* yang terdiri dari 3.091 data, di mana 1554 ulasan termasuk dalam kelas negatif dan 1537 ulasan termasuk dalam kelas positif. Setelah model menghasilkan prediksinya di masing-masing kelas, akan dilakukan perhitungan akurasi, presisi, dan *recall* untuk mengetahui tingkat kepercayaan model, dengan melakukan hal ini, diharapkan dapat dipahami sejauh mana model dapat diandalkan dalam memprediksi kelas yang benar. Hasil Uji model dapat dilihat pada Tabel 5.

Tabel 5. Hasil Evaluasi Model Menggunakan Data Testing

Model	Akurasi
Bi-LSTM	0.9000
LSTM	0.8625

Untuk mengetahui hasil akurasi, presisi dan *recall* model maka dilakukan proses evaluasi model menggunakan metode perhitungan *confusion matrix*. Dari hasil pengujian model Bi-LSTM didapat bahwa model berhasil memprediksi dengan tepat ulasan berlabel negatif sebanyak 1318 kali, namun salah memprediksi sebanyak 236 kali. Begitu juga dengan ulasan berlabel positif, model berhasil memprediksi sebanyak 1464 data dengan benar dan salah memprediksi sebanyak 73 data.

Tabel 6. Hasil *Confusion Matrix* Bi-LSTM

Nilai Prediksi	Nilai Input	
	POSITIF	NEGATIF
POSITIF	1464	236
NEGATIF	73	1318

Berdasarkan Tabel 6 maka dapat melakukan pengukuran untuk menentukan hasil dari model Bi-LSTM.

1.  $Accuracy = \frac{1464+1318}{1464+1318+73+236} \times 100\% = 90.00\%$
2.  $Precision = \frac{1464}{1464+236} \times 100\% = 86.1176\%$
3.  $Recall = \frac{1464}{1464+73} \times 100\% = 95.25\%$

Dari hasil pengujian model LSTM didapat bahwa model berhasil memprediksi dengan tepat ulasan berlabel negatif sebanyak 1249 kali, namun salah memprediksi sebanyak 305 kali. Begitu juga dengan ulasan berlabel positif, model berhasil memprediksi sebanyak 1417 data dengan benar, namun salah memprediksi sebanyak 120 data.

Tabel 7. Hasil *Confusion Matrix* LSTM

Nilai Prediksi	Nilai Input	
	POSITIF	NEGATIF
POSITIF	1417	305
NEGATIF	120	1249

Berdasarkan Tabel 7 maka dapat melakukan pengukuran untuk menentukan hasil dari model LSTM.

1.  $Accuracy = \frac{1417+1249}{1417+1249+305+120} \times 100\% = 86.25\%$
2.  $Precision = \frac{1417}{1417+305} \times 100\% = 82.28\%$
3.  $Recall = \frac{1417}{1417+120} \times 100\% = 92.19\%$

#### 4. Kesimpulan

Penelitian mengenai analisis sentimen ulasan aplikasi MyPertamina di *Google Play Store* telah berhasil dilakukan. Penelitian menggunakan metode LSTM dan Bi-LSTM mampu mendapatkan tingkat akurasi model dan evaluasi yang cukup tinggi hal ini karena melakukan eksperimen penambahan, perubahan pada *parameter* jumlah *neuron* yang digunakan dan fungsi aktivasi pada *layer output*. Hasil maksimal didapatkan setelah menerapkan jumlah *neuron* 300 dan menggunakan fungsi aktivasi *sigmoid*. Model yang dikembangkan telah diuji coba dengan menggunakan *dataset* ulasan dari aplikasi MyPertamina, dan menghasilkan tingkat akurasi yang cukup baik. Dalam uji coba tersebut, model Bi-LSTM mampu mencapai tingkat akurasi sebesar 90% dan model LSTM mencapai 86,25%.

#### Daftar Pustaka

- [1] R. A. Softina, F. M. Amin, and N. Wahyudi, “Analisis Faktor yang Mempengaruhi Innovation Resistance dan Intention to Use Terhadap Penerapan Pembayaran Non Tunai,” *J. Sist. Inf. Bisnis*, vol. 12, no. 1, pp. 26–35, 2022, doi: 10.21456/vol12iss1pp26-35.
- [2] A. Lutfi, “EFEKVITAS PENGGUNAAN APLIKASI MY PERTAMINA DI ERA KENAIKAN BBM BERSUBSIDI,” vol. 1, no. 2, 2022.
- [3] G. S. Heryadi, “Strategi Promosi Loyalty Program My Pertamina yang Dilakukan Oleh PT Pertamina (PERSERO) Tahun 2018,” *J. Visi Komun.*, vol. 17, no. 01, 2018.
- [4] A. Faadilah, “Analisis Sentimen Pada Ulasan Aplikasi Tokopedia di Google Play Store Menggunakan Metode Long Short Term Memory,” pp. 1–46, 2020.
- [5] E. M. Sipayung, H. Maharani, and I. Zefanya, “Perancangan Sistem Analisis Sentimen Komentar Pelanggan Menggunakan Metode Naive Bayes Classifier,” *J. Sist. Inf.*, vol. 2, no. 10, pp. 2355–4614, 2022.
- [6] N. F. Ahsinil Amal Annur, Aang Alim Murtopo, “Analisis Sentimen Aplikasi E-Learning Selama Pandemi Covid-19 Dengan Menggunakan Metode Support Vector Machine Dan ...,” *IJIR*, vol. 3, pp. 9–17, 2022.
- [7] A. Rahman, E. Utami, and S. Sudarmawan, “Sentimen Analisis Terhadap Aplikasi pada Google Playstore Menggunakan Algoritma Naïve Bayes dan Algoritma Genetika,” *J. Komtika (Komputasi dan Inform.)*, vol. 5, no. 1, pp. 60–71, 2021, doi: 10.31603/komtika.v5i1.5188.
- [8] S. M. Qaisar, “Sentiment Analysis of IMDb Movie Reviews Using Long Short-Term Memory,” *2020 2nd Int. Conf. Comput. Inf. Sci. ICCIS 2020*, pp. 12–15, 2020, doi:

- 
- 10.1109/ICCIS49240.2020.9257657.
- [9] L. G. Astuti, P. S. Informatika, and U. Udayana, "Implementasi LSTM pada Analisis Sentimen Review Film Menggunakan Adam dan RMSprop Optimizer," *J. Elektron. Ilmu Komput. Udayana*, vol. 10, no. 4, pp. 351–362, 2022.
- [10] D. R. Alghifari, M. Edi, and L. Firmansyah, "Implementasi Bidirectional LSTM untuk Analisis Sentimen Terhadap Layanan Grab Indonesia Bidirectional LSTM Implementation for Sentiment Analysis Against Grab Indonesia Services," *J. Manaj. Inform.*, vol. 12, pp. 89–99, 2022.
- [11] C. G. Indrayanto, D. E. Ratnawati, and B. Rahayudi, "Analisis Sentimen Data Ulasan Pengguna Aplikasi MyPertamina di Indonesia pada Google Play Store menggunakan Metode Random Forest," vol. 7, no. 3, 2023.
- [12] M. Ilmiah, *Implementasi Metode Bidirectional Long Short-Term Memory (Bi-LSTM) untuk Prediksi Kasus Positif Covid-19 di Indonesia*. 2022.
- [13] L. Hickman, S. Thapa, L. Tay, M. Cao, and P. Srinivasan, "Text Preprocessing for Text Mining in Organizational Research: Review and Recommendations," *Organ. Res. Methods*, vol. 25, no. 1, 2022, doi: 10.1177/1094428120971683.
- [14] D. E. Birba, "A Comparative study of data splitting algorithms for machine learning model selection," 2020.