

Implementasi *Bidirectional LSTM* untuk Analisis Sentimen Terhadap Layanan Grab Indonesia

Bidirectional LSTM Implementation for Sentiment Analysis Against Grab Indonesia Services

Dloifur Rohman Alghifari¹, Mohammad Edi^{2*}, Lutfi Firmansyah³

Magister Teknik Informatika, Universitas AMIKOM Yogyakarta, Indonesia

*E-mail: mohammad.edi100@students.amikom.ac.id

Abstrak

Grab Indonesia adalah salah satu perusahaan ojek online terkemuka di Indonesia dan memiliki jumlah pelanggan cukup besar di Indonesia. Tingkat kepuasan pelanggan yang bervariasi terhadap layanan yang diberikan, sehingga pasti ada saran dan keluhan dari para pelanggan. Analisis sentiment dapat dijadikan salah satu solusi untuk mengetahui tingkat kepuasan pelayan guna meningkatkan sistem dan pelayan. Penelitian ini bertujuan untuk mengetahui tingkat kepuasan pengguna grab Indonesia melalui aplikasi grab yang ada di playstore. Model pendekatan yang dapat digunakan salah satunya yaitu LSTM. LSTM merupakan pengembangan algoritma RNN untuk mengatasi masalah vanishing gradien. LSTM memiliki kelemahan hanya menjalankan hanya dapat menangkap informasi dari satu arah. Bidirectional LSTM (BiLSTM) merupakan metode LSTM yang telah dikembangkan, dimana BiLSTM dapat menangkap informasi dari dua arah. Metode BiLSTM ini semakin banyak data maka akan meningkat performa algoritmanya. Hasil pengujian menunjukkan BiLSTM lebih baik handal dari LSTM dalam untuk kasus analisi sentiment terhadap layanan grab indonesia. BiLSTM menghasilkan akurasi terbaik sebesar 91% dan training loss sebesar 28%. Saran untuk penelitian yang berikutnya dapat menghasilkan representasi kata yang lebih banyak dan variatif dengan cara mempertimbangkan pada kombinasi word embedding-nya.

Kata kunci: Analisis Sentimen; Grab Indonesia; LSTM, Bidirectional LSTM.

Abstract

Grab Indonesia is one of the leading online motorcycle taxi companies in Indonesia and has a large number of customers in Indonesia. The level of customer satisfaction varies with the services provided, so there must be suggestions and complaints from customers. Sentiment analysis can be used as a solution to determine the level of service satisfaction in order to improve the system and service. This study aims to determine the level of satisfaction of Grab Indonesia users through the Grab application in the Playstore. One of the approaches that can be used is LSTM. LSTM is an RNN algorithm development to solve the vanishing gradient problem. LSTM has the disadvantage of only running can only capture information from one direction. Bidirectional LSTM (BiLSTM) is an LSTM method that has been developed, where BiLSTM can capture information from two directions. In this BiLSTM method, the more data, the better the algorithm's performance. The test results show that BiLSTM is more reliable than LSTM in the case of sentiment analysis on the Indonesian Grab service. BiLSTM produces the best accuracy of 91% and training loss of 28%. Suggestions for future research can produce more and varied word representations by considering the word embedding combinations.

Keywords: Sentiment Analysis; Grab Indonesia; LSTM, Bidirectional LSTM.

Naskah diterima 31 Jul. 2022; direvisi 23 Sept. 2022; dipublikasikan 1 Okt. 2022.

JAMIKA is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.



I. PENDAHULUAN

Transportasi merupakan salah satu kebutuhan pokok masyarakat Indonesia. Bagi kebanyakan orang, transportasi merupakan hal penting sebagai menunjang aktivitas mereka sehari-hari. Akan tetapi transportasi pribadi hanya dimiliki hanya beberapa orang saja. Oleh karena itu, Sebagian masyarakat banyak yang menggunakan transportasi umum seperti ojek atau angkot.

Di Indonesia, perusahaan yang bergerak dibidang transportasi *online* banyak bermunculan, mulai dari perusahaan besar seperti Gojek dan Grab. Perusahaan terus bekerja untuk layanan terbaik bagi masyarakat seperti menawarkan berbagai diskon untuk layanan yang ada. Seiring berkembangnya setiap layanan, minat masyarakat terhadap transportasi *online* meningkat. Perusahaan transportasi *online* juga kerap dipuji. Masyarakat menyampaikan tanggapan terhadap transportasi *online* bisa melalui media sosial, *playstore* tempat mengunduh aplikasi transaksi *online* tersebut. Tanggapan masyarakat dapat berupa pujian, saran, maupun kritik [1]. Grab Indonesia sebagai salah satu perusahaan transportasi *online* di Indonesia yang memiliki jumlah

pelanggan mencakup hampir seluruh wilayah di Indonesia [2]. Setiap pelanggan memiliki tingkat kepuasan berbeda terhadap layanan yang diberikan oleh Grab Indonesia, sehingga selalu ada saran dan keluhan.

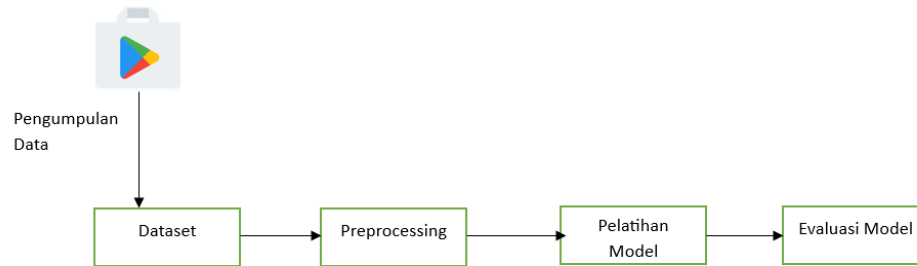
Analisis sentimen merupakan salah satu solusi yang dapat digunakan untuk mengetahui tingkat kepuasan pengguna. Dari data yang tidak terstruktur dapat ditarik sebuah kesimpulan dengan menggunakan analisis sentimen [3]. *Deep Learning* merupakan salah satu metode yang dapat digunakan dalam pengimplementasian analisis sentimen. Cara kerja *deep learning* dengan mengekstraksi data menggunakan *neural network*, kemudian model melakukan pembelajaran melalui nilai error. Salah satu metode *deep learning* untuk klasifikasi teks yaitu *Long Short-Term Memory (LSTM)* [4]. LSTM merupakan metode hasil perkembangan dari metode *Recurrent Neural Network (RNN)*. LSTM memiliki kelemahan hanya memproses kata dalam satu arah saja. Bidirectional *Long Short-Term Memory (BiLSTM)* merupakan LSTM yang telah dikembangkan yang dapat mengatasi kelemahan pada metode LSTM, dimana BiLSTM dapat memproses kata dalam 2 arah.

Analisis sentimen merupakan klasifikasi teks yang bertujuan untuk mengklasifikasikan teks (dokumen) yang berisi opini sebagai opini yang positif, negatif, atau netral. Adapun penelitian sebelumnya yang pernah dilakukan terkait klasifikasi sentimen yaitu penelitian yg dilakukan oleh Krisnasari [5], dengan menggunakan 1.034.329 tweet dari Twitter antara November 2021 dan Maret 2022. Data Tweet dikategorikan dengan menggunakan model klasifikasi metode *Bidirectional Long Short-Term Memory (BiLSTM)* dan menghasilkan 45,98% sentimen negatif, 28,04% sentimen positif, dan 25,98% netral. Penelitian lainnya oleh R. Watrionthos [6], algoritma Naïve Bayes digunakan untuk mengevaluasi persepsi dan penerimaan publik terhadap model pembelajaran campuran pada data tweet dari Twitter. Hasil penelitian menunjukkan polarisasi sentimen positif 44,51% dan negatif 45,80%. Dalam penelitian yang dilakukan Wahyunita [7] tentang transportasi *online*, analisis sentimen dilakukan pada tweet berbahasa Indonesia dengan metode pembobotan *Hybrid TF-IDF* dan metode algoritma *K-Nearest Neighbor (KNN)* sebagai metode klasifikasi sentimen. Dengan menggunakan pengujian *cross validation* mendapatkan hasil terbaik dengan nilai akurasi 70% untuk $k=5$, presisi kelas sentimen positif 68%, negatif 75%, recall kelas sentimen positif 82%, negatif 59%, *F-measure* kelas sentimen positif 74% dan negatif 65%. Penelitian lainnya oleh Wahyudi [8], menggunakan *Support Vector Machine (SVM)* untuk menganalisis ulasan atau rating pengguna aplikasi Grab di Google Play Store. Hasil dari analisis *Support Vector Machine* menunjukkan akurasi sebesar 85,54% hasil ulasan positif adalah "ovo", dan ulasan negative adalah "driver". Dalam penelitian yang dilakukan oleh Rahmatullah [9], data diperoleh dari komentar di twitter dan menggunakan algoritma *Naïve Bayes*, *Support Vector Machine* dan *K-Nearest Neighbor* sebagai perbandingan akurasi yang dihasilkan. Ketiga algoritma tersebut menggunakan salah satu fitur seleksi *Particle Swarm Optimization*. Hasil akurasi yang didapat membuktikan dengan fitur seleksi *Particle Swarm Optimization* lebih tinggi dan akurasi yang paling tinggi oleh Algoritma *Support Vector Machine (SVM-PSO)* dengan Akurasi 98 % dan AUC 0.988. Pada penelitian yang telah dilakukan Afrianto [10], menggunakan opini publik terhadap saham. Model dibangun menggunakan metode *Bidirectional Long Short-Term Memory (BiLSTM)* dengan menambahkan opini publik dalam proses training dan testing. Penelitian lainnya Hermanto [11] mengimplemntasikan metode LSTM, LSTM-CNN dan CNN-LSTM untuk klasifikasi sentimen positif dan negatif pada judul berita berbahasa Indonesia. Dataset berdasarkan judul artikel berbahasa Indonesia yang tersedia di website Detik Finance. Secara berurutan metode LSTM, LSTM-CNN dan CNN-LSTM menghasilkan akurasi sebesar, 62%, 65% dan 74% setelah melalui pengujian. Penelitian Murty [12] mencapai akurasi 85% ketika melakukan penelitian mengenai *sentiment analysis* menggunakan algoritma LSTM.

Berdasarkan permasalahan yang telah diuraikan dan beberapa penelitian yang telah dipaparkan, maka penelitian ini mengusulkan metode *bidirectional LSTM* untuk melakukan analisis sentimen pada layanan grab Indonesia, dengan menggunakan data komentar pengguna aplikasi di *playstore*. Dalam penelitian ini juga dilakukan perbandingan dengan beberapa metode *machine learning* sebagai pembanding dengan metode yang diusulkan. Metode *machine learning* sebagai pembanding, yaitu *Multinomial Naïve Bayes (Multinomial NB)*, *Logistic Regression* dan *Linear Support Vector Classifier (Linear SVC)*. Kami menggunakan pembobotan kata dan ekstraksi *feature* menggunakan *Bag of Word (BOW)*. Penjelasan hasil penelitian ini dibagi menjadi beberapa bagian: Metodologi Penelitian, Hasil dan Pembahasan, dan Kesimpulan.

II. METODE PENELITIAN

Penelitian ini dilakukan dengan menjalankan beberapa tahap yaitu pengumpulan dataset, *preprocessing*, pelatihan model dan evaluasi hasil. Metode penelitian secara jelas dapat dilihat pada Gambar 1.



Gambar 1. Metode Penelitian

Dataset

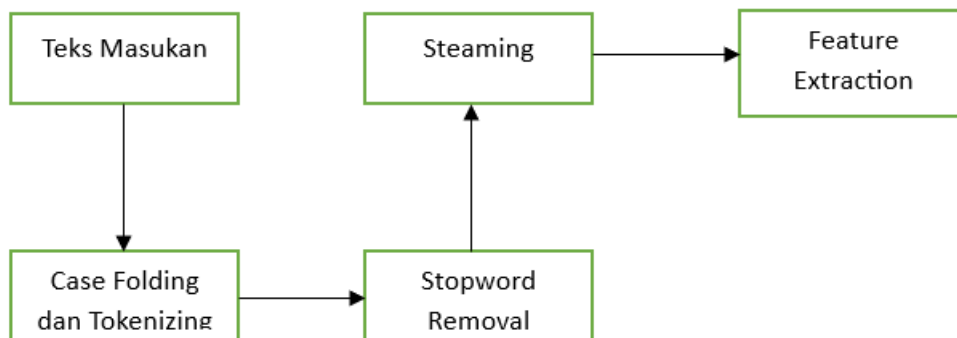
Dataset yang digunakan diambil dari hasil *scrapping* kolom komentar pada aplikasi grab yang ada playstore antara tanggal 04 juli sampai 23 juli 2022. Dataset yang akan digunakan pada penelitaian berjumlah 5000 data berbahasa indonesia. Dataset akan dikelompokkan menjadi 3 label, yaitu negatif, netral dan positif. 1615 data berlabel negatif, 272 data netral dan 3113 data berlabel positif. Contoh dataset yang digunakan pada penelitian ini dapat dilihat pada Tabel 1.

TABEL 1
SAMPLE DATASET

Index	Content	Label
5	Setiap hari cuma di kasih 1 orderan apa ga gilaaaaa	1
98	Terimakasih kasih	1
1697	Saya kecewa pada pihak grab karna penawaran grab modal tidak merata....	-1
4787	Orderan nya muhon di sams ratakan	0

Preprocessing

Pada tahap preprosesing 2 tahapan, yaitu tranformasi data dan *text-preprocessing*. Pada tranformasi data dilakukan tranformasi label menjadi numerik, yaitu *negative* menjadi (-1), netral menjadi 0 dan positif menjadi 1. Pada tahap *text-preprocessing* dilakukan untuk mengubah teks menjadi data numerik [13], dengan cara melakukan ekstraksi menggunakan metode BOW (*Bag of Word*). Tahapan pada *text-preprocessing* dapat dilihat pada Gambar 2.



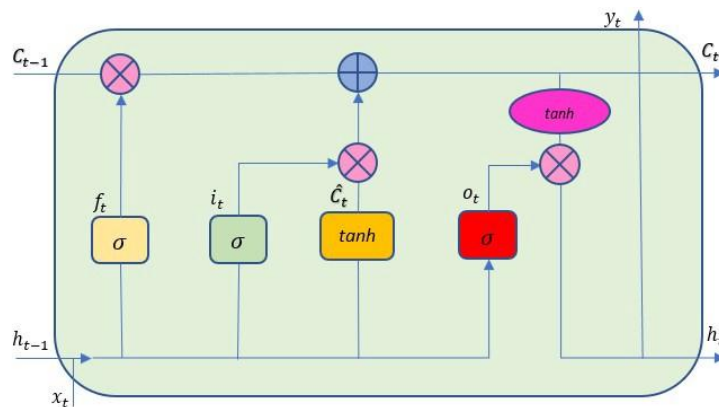
Gambar 2. Tahapan *Text-Preprocessing*

Pada Gambar 2 dapat dilihat teks masukan yang digunakan adalah teks dari kolom komentar pada aplikasi grab di *playstore*. Proses selanjutnya dilakukan *case folding* yaitu konversi huruf menjadi huruf kecil semua serta perbersihan kata seperti penghapusan *username*, angka, url, emoji, spasi yang berlebih dan karakter yang diulang lebih dari dua, dan selanjutnya memisahkan kalimat menjadi per-kata yang dapat disebut *tokenizing*. Kemudian dilakukan *stopword removal*, yaitu penghapusan kata berdasarkan *stopword list* yang digunakan. Proses *stemming*, yaitu pencarian kata dasar pada tiap kata yang telah dipisahkan sebelumnya.

Kemudian dilakukan proses pembobotan kata atau *feature extraction* menggunakan metode BOW (*Bag of Word*). Sehingga setiap kata akan memiliki sebuah nilai yang nantinya kan digunakan untuk perhitungan dalam model pembelajaran.

Bidirectional LSTM

Algoritma *Deep Learning* menghasilkan performa lebih baik dan waktu komputasi yang lebih cepat dari algoritma *machine learning*. Salah satu algoritma *Deep Learning* yang dapat digunakan untuk klasifikasi kata adalah LSTM. LSTM atau *Long Short-Term Memory* merupakan metode RNN yang telah dikembangkan dengan menambahkan interaksi tambahan pada tiap modulnya [14]. LSTM menggunakan blok *memory-cell* yang terdiri dari *input gate*, *forget gate* dan *output gate* untuk mengganti lapisan RNN agar dapat mengatasi masalah *vanishing gradient* pada RNN [15], [16]. Pada LSTM informasi dapat di simpan dalam waktu panjang karena dapat mempelajari *long-term dependency* [17]. Arsitektur LSTM sendiri dapat dilihat pada Gambar 3.



Gambar 3. Arsitektur LSTM

Gerbang pertama pada LSTM disebut *forget gate*, dimana terdapat proses memilah informasi yang ada pada *cell state* dengan menggunakan persamaan 1. informasi akan dibuang dari *cell state* jika *forget gate* bernilai 0, sebaliknya informasi akan disimpan *cell state* jika *forget gate* bernilai 1.

$$f_t = \sigma(W_f \cdot [h_{t-1}, X_t] + b_f) \tag{1}$$

Gerbang kedua yaitu *input gate*, disini informasi akan melewati 2 lapisan yaitu *sigmoid* sesuai dengan persamaan 2 dan *tanh* perhitungannya dengan menggunakan persamaan 3. *Cell state* dihasilkan dari nilai *output* kedua lapisan yang telah digabungkan.

$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \tag{2}$$

$$C_t = \tanh(W_c \cdot [h_{t-1}, X_t] + b_c) \tag{3}$$

Kemudian dilakukan proses perbaruan nilai *cell state* dengan menggunakan persamaan 4. Selanjutnya akan melewati *output gate*, yang mana didalamnya nilai *output cell state* akan ditentukan. pada lapisan *sigmoid* digunakan untuk memilih *output* berdasarkan *cell state* yang ada dengan menggunakan persamaan 5, dan selanjutnya akan diteruskan ke lapisan *tanh* menggunakan persamaan 6.

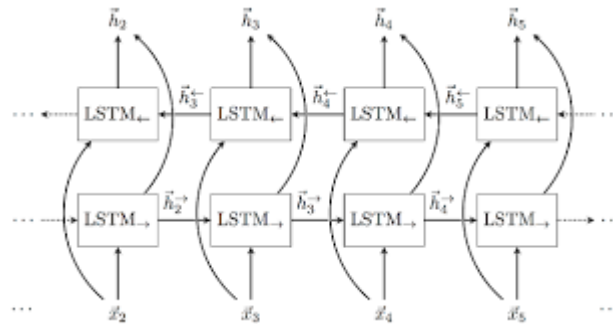
$$C_t = f_t \cdot C_{t-1} + i_t \cdot C_t \tag{4}$$

$$O_t = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o) \tag{5}$$

$$h_t = O_t \cdot \tanh(C_t) \tag{6}$$

LSTM memiliki kelemahan hanya memproses kata satu satu arah saja, yaitu dari awal sampai akhir [18]. Kami dalam penelitian ini mempertimbangkan untuk menggunakan *bidirectional LSTM* agar dapat membaca kata dari 2 arah, yaitu awal sampai akhir dan akhir sampai awal [19]. Sesuai arsitektur *bidirectional*

LSTM pada gambar 4, *bidirectional LSTM* memiliki 2 lapisan LSTM yang terpisah satu untuk maju dan satu untuk mundur sesuai dengan persamaan 7.



Gambar 4. Arsitektur Bidirectional LSTM

$$h_t^{BiLSTM} = h_t^{forward} \oplus h_t^{backward} \quad (7)$$

Evaluasi Hasil

Evaluasi hasil digunakan untuk menghitung kinerja model yang sudah dibuat, dimana dalam penelitian ini evaluasi akan menggunakan *confusion matrix* dengan cara menghitung rasio prediksi benar dan salah. *Confusion matrix* lebih jelasnya dapat dilihat pada Gambar 5 [20].

		Predicted Class	
		Positive	Negative
Actual Class	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

Gambar 5. *Confusion Matrix*

Confusion matrix nantinya akan menghasilkan beberapa model evaluasi diantaranya, yaitu *accuracy* didapat nilai kedekatan hasil prediksi dengan hasil sebenarnya dapat dihitung dengan persamaan 8. *Precision* didapat berdasarkan nilai dari rasio prediksi benar positif terhadap data bernilai positif pada keseluruhan data, perhitungannya dapat menggunakan persamaan 9. *Recall* didapat berdasarkan hasil rasio benar positif terhadap data yang bernilai benar positif pada keseluruhan data, perhitungannya menggunakan persamaan 10. *F1 score* yaitu nilai *precision* dan *recall* yang telah digabungkan [21] dengan menggunakan persamaan 11 [22].

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (8)$$

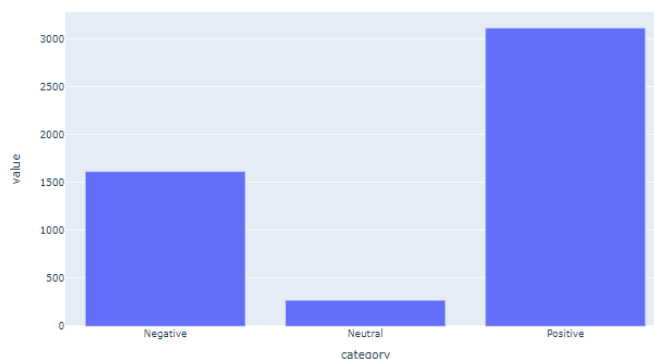
$$Precision = \frac{TP}{TP+FP} \quad (9)$$

$$Recall = \frac{TP}{TP+FN} \quad (10)$$

$$F1\ score = \frac{2 \times Precision \times Recall}{Precision+recall} \quad (11)$$

III. HASIL DAN PEMBAHASAN

Pada pembahasan diatas, dijelaskan bahwa dataset yang digunakan diambil dari hasil *scrapping* kolom komentar pada aplikasi grab yang ada *playstore* antara tanggal 04 juli sampai 23 juli 2022. Dataset yang akan digunakan pada penelitaian berjumlah 5000 data berbahasa indonesia. Dataset akan dikelompokkan menjadi 3 label, yaitu negatif, netral, dan positif. Label *negative* diambil dari komentar yang mempunyai rating 1 sampai 2, label netral komentar yang mempunyai rating 3 dan label positif diambil dari komentar rating 4 sampai 5. Didapatkan 1615 data berlabel *negative*, 272 data netral, dan 3113 data berlabel positif. Gambar 6 menampilkan pengelompokan dataset. Contoh dataset yang digunakan pada penelitian ini dapat dilihat pada Tabel 1.



Gambar 6. Grafik Pengelompokan Dataset

Sebelum dataset dilakukan pemodelan dan pengujian, dataset akan dilakukan preprocessing, diantaranya *case folding* dan *tokenizing*, *stopword removal*, *steaming*, dan *feature extraction*. *Case folding* mengubah semua karakter menjadi *lower case*, selanjutnya dilakukan proses *tokenizing* yang merupakan memecah kalimat menjadi per kata. Contoh hasil *preprocessing*, *Case folding* dan *tokenizing* dapat dilihat pada Tabel 2.

TABEL 2
CONTOH HASIL CASE FOLDING DAN TOKENIZING

Index	Content	Hasil Case Folding dan Tokenizing
5	Setiap hari cuma di kasih 1 orderan apa ga gilaaaaa	['Setiap', 'hari', 'cuma', 'di', 'kasih', 'orderan', 'apa', 'ga', 'gilaaaaa']
23	Sangat baik dan sangat membantu saya dalam memenuhi kebutuhan terimakasih Grab.	['Sangat', 'baik', 'dan', 'sangat', 'membantu', 'saya', 'dalam', 'memenuhi', 'kebutuhan', 'terimakasih', 'Grab']
35	Oke untuk nambah pendapatan, tapi tolong untuk pengembang diminalisir order fiktif saya sudah 3xmendapat order fiktif	['Oke', 'untuk', 'nambah', 'pendapatan', 'tapi', 'tolong', 'untuk', 'pengembang', 'diminalisir', 'order', 'fiktif', 'saya', 'sudah', 'order', 'fiktif']
325	Alhamdulillah dengan ada y grab uang dapur saya bisa kebantu sukses terus untuk grab	['Alhamdulillah', 'dengan', 'ada', 'y', 'grab', 'uang', 'dapur', 'saya', 'bisa', 'kebantu', 'sukses', 'terus', 'untuk', 'grab']
452	Bagus banget apk nya	['Bagus', 'banget', 'apk', 'nya']

Preprocessing selanjutnya, yaitu *stopword removal* proses ini menghapus kata-kata yang tidak penting. Contoh hasil proses *stopword removal* dapat dilihat pada Tabel 3.

TABEL 3
CONTOH HASIL STOPWORD REMOVAL

Index	Content	Hasil Stopword Removal
5	Setiap hari cuma di kasih 1 orderan apa ga gilaaaaa	['Setiap', 'hari', 'cuma', 'kasih', 'orderan', 'gilaaaaa']
23	Sangat baik dan sangat membantu saya dalam memenuhi kebutuhan terimakasih Grab.	['baik', 'membantu', 'saya', 'memenuhi', 'kebutuhan', 'terimakasih', 'Grab']
35	Oke untuk nambah pendapatan, tapi tolong untuk pengembang diminalisir order fiktif saya sudah 3xmendapat order fiktif	['Oke', 'nambah', 'pendapatan', 'tolong', 'pengembang', 'diminalisir', 'order', 'fiktif', 'saya', 'order', 'fiktif']
325	Alhamdulillah dengan ada y grab uang dapur saya bisa kebantu sukses terus untuk grab	['Alhamdulillah', 'grab', 'uang', 'dapur', 'saya', 'kebantu', 'sukses', 'terus', 'grab']
452	Bagus banget apk nya	['Bagus', 'banget', 'apk']

Setelah *stopword removal*, tahap selanjutnya dalam *preprocessing* adalah *steaming*. Tahap ini merupakan tahapan untuk membuang imbuhan dalam kata sehingga yang ditampilkan adalah kata dasar. Contoh hasil *steaming* dapat dilihat pada Tabel 4. Tahap selanjutnya adalah *feature extraction*. Dalam proses ini peneliti menggunakan *library* dari python, yaitu *Bag of words* (BOW).

TABEL 4
 CONTOH HASIL STEAMING

Index	Content	Hasil Steaming
5	Setiap hari cuma di kasih 1 orderan apa ga gilaaaaa	['tiap', 'hari', 'cuma', 'kasih', 'order', 'gilaaaaa']
23	Sangat baik dan sangat membantu saya dalam memenuhi kebutuhan terimakasih Grab.	['baik', 'bantu', 'saya', 'penuh', 'butuh', 'terimakasih', 'Grab']
35	Oke untuk nambah pendapatan, tapi tolong untuk pengembang diminalisir order fiktif saya sudah 3xmendapat order fiktif	['Oke', 'tambah', 'pendapatan', 'tolong', 'pengembang', 'minalisir', 'order', 'fiktif', 'saya', 'order', 'fiktif']
325	Alhamdulillah dengan ada y grab uang dapur saya bisa kebantu sukses terus untuk grab	['Alhamdulillah', 'grab', 'uang', 'dapur', 'saya', 'bantu', 'sukses', 'terus', 'grab']
452	Bagus banget apk nya	['Bagus', 'banget', 'apk']

Pada penelitian ini seluruh experiment pengujian dilakukan pada komputer dengan RAM 16 GB, VGA Nvidia GTX dan processor core i5. Dalam pengaplikasiannya menggunakan aplikasi jupyter notebook dengan python versi 3.9.7. Setelah dataset melalui tahap *preprocessing*, selanjutnya dataset dibagi menjadi menjadi dua, yaitu 80% sebagai data *training* dan 20% sebagai data *testing*. Selajutnya data *training* akan dibagi menjadi 2 sebagai data validasi sebanyak 25% selama proses pelatihan model. Dalam penelitian ini juga dilakukan perbandingan dengan beberapa metode *machine learning* sebagai pembanding dengan metode yang diusulkan. metode *machine learning* sebagai pembanding, yaitu *Multinomial Naïve Bayes* (*Multinomial NB*), *Logistic Regression* dan *Linear Support Vector Classifier* (*Linear SVC*). Kami menggunakan pembobotan kata dan ekstraksi *feature* menggunakan *Bag of Word* (BOW). Berdasarkan Tabel 5 dapat dilihat bahwa metode *Multinomial NB* menghasilkan akurasi sebesar 0.63, metode *Logistic Regression* menghasilkan nilai akurasi sebesar 0.67 dan *Linear SVC* menghasilkan akurasi 0.65. Metode *Logistic Regresion* menghasilkan nilai akursi yang lebih baik dari metode lainnya.

TABEL 5
 PERBANDINGAN ALGORITMA LAINNYA

Algoritma	Accuracy
Multinomial NB	0.63
Logistic Regression	0.67
Linear SVC	0.65

Selanjutnya melakukan pengujian terhadap algoritma yang diusulkan yaitu *bidirectional LSTM* (BiLSTM). Kami juga akan melakukan perbandingan dengan metode LSTM. Dalam metode LSTM terdapat nilai parameter yang dapat mempengaruhi nilai akuarasi algoritma. Parameter LSTM dan BiLSTM pada penelitian ini telah ditentukan diawal, nilai parameter yang digunakan untuk pengujian dapat dilihat pada Tabel 6. Agar dapat menghindari kondisi *overfitting* kami memakai nilai dropout sebesar 0.4 dan jumlah epoch yang digunakan berjumlah 25.

TABEL 6
 PARAMETER LSTM DAN BiLSTM

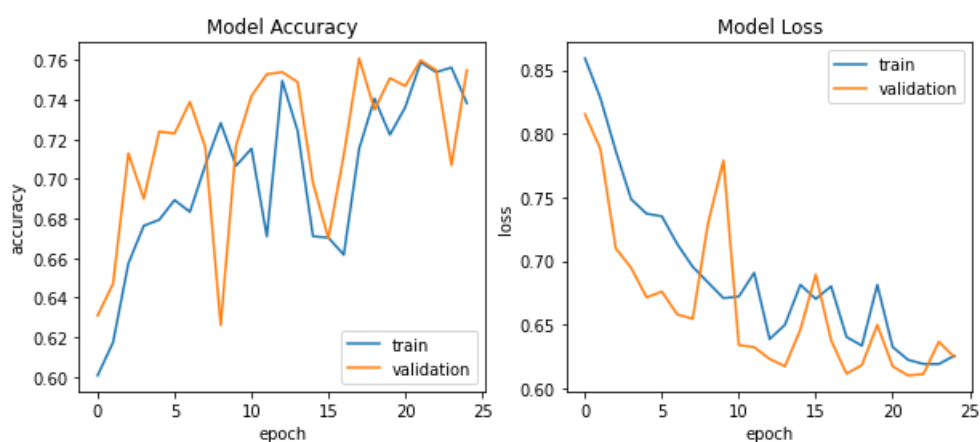
Nama Parameter	Nilai Parameter
Embedding Size	100
Activation	Sigmoid
Optimizer	Stochastic Gradient Descent (sgd)
Learning Rate	0.1
Batch Size	64

Berdasarkan hasil pengujian antara LSTM dan BiLSTM pada Tabel 7 dapat dilihat hasil nilai metrik evaluasi menunjukkan bahwa BiLSTM memiliki kinerja lebih handal. BiLSTM juga lebih baik jika dibandingkan algoritma lainnya pada Tabel 7. Dapat dilihat juga pada grafik hasil pelatihan pada Gambar 7 dan Gambar 8,

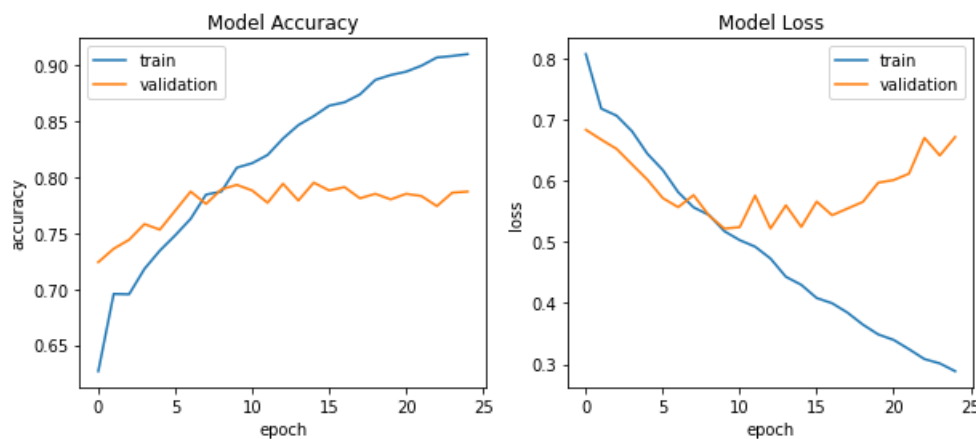
menunjukkan BiLSTM akurasi pelatihan stabil terus naik setiap *epoch*-nya dan *training loss* yang lebih stabil dari LSTM. Sedangkan pada LSTM akurasi pelatihan dan training loss cenderung naik turun pada setiap *epoch*-nya. Akurasi terbaik yang dihasilkan metode BiLSTM, yaitu akurasi pengujian sebesar 0.9097 atau 91% dan training loss sebesar 0.2884 atau 29%. Berdasarkan pengujian yang telah dilakukan BiLSTM memiliki kinerja yang handal dari LSTM karena memiliki kelebihan dapat memproses kata 2 arah. BiLSTM juga memiliki kelemahan, yaitu membutuhkan data yang banyak serta waktu dan biaya komputasi yang lebih tinggi dari metode lain pada penelitian ini.

TABEL 7
HASIL PENGUJIAN LSTM DAN BiLSTM

Algoritma	Akurasi	Training Loss	Precision	Recall
LSTM	0.7563	0.4822	0.5701	0.9053
BiLSTM	0.9097	0.2884	0.7916	0.9330



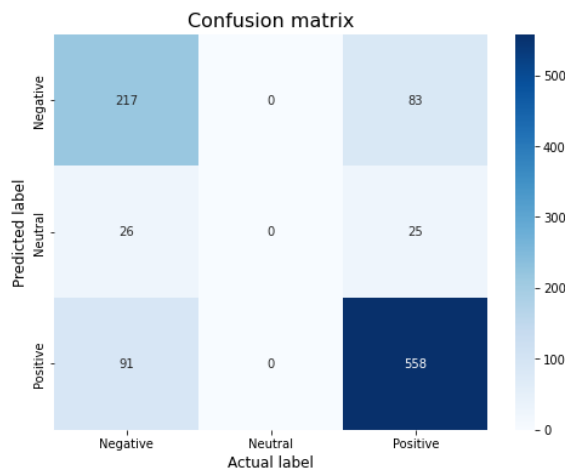
Gambar 7. Grafik Hasil Pelatihan LSTM



Gambar 8. Grafik Hasil Pelatihan BiLSTM

Pada Gambar 7 dan 8 dapat dilihat juga nilai akurasi pada data testing dari algoritma BiLSTM lebih baik dari LSTM dengan nilai akurasi sebesar 78% berbanding dengan 76% pada LSTM. Kemudian dilakukan evaluasi hasil menggunakan metode *confusion matrix*, hasilnya dapat dilihat pada Gambar 9. Dengan menggunakan persamaan 9, 10, dan 11 maka akan didapatkan nilai *precision* sebesar 0.69, nilai *recall* sebesar 0.82 dan *F1-score* sebesar 0.75. Dengan menggunakan model pelatihan yang terbaik, maka dilakukan pengujian dengan menggunakan kalimat yang akan di inputkan, kemudian sistem akan mengklasifikasi kalimat tersebut termasuk sentimen positif, negatif atau netral. Contoh kalimat dalam ujicoba adalah ‘aplikasi bagus dan bermanfaat’, ‘mengecewakan sekali’, ‘pelayanan grabnya sangat payah dan parah’. Dari tiga kali

pengujian, kalimat aplikasi bagus dan bermanfaat menghasilkan sentimen positif tetapi saat dimasukkan kalimat mengecewakan sekali malah diklasifikasikan sebagai sentimen positif. Hasil kesalahan ini didapat karena dataset yang digunakan kurang banyak. Metode BiLSTM ini semakin banyak data maka akan meningkat performa algoritmanya. Selain itu faktor lainnya yang dapat mempengaruhi yaitu pada saat proses *preprocessing* dilakukan, serta perlu dilakukan *POS-Tagging* dan *Negation Handling* untuk meningkatkan ekstraksi sentimennya. Hasil pengujian berdasarkan kalimat masukan dapat dilihat pada Gambar 10.



Gambar 9. Confusion Matrix BiLSTM

```

1 predict_class(['aplikasi bagus dan bermanfaat'])
1/1 [=====] - 0s 19ms/step
The predicted sentiment is Positive

1 predict_class(['mengecewakan sekali'])
1/1 [=====] - 0s 19ms/step
The predicted sentiment is Positive

1 predict_class(['pelayanan grabnya sangat payah dan parah'])
1/1 [=====] - 0s 28ms/step
The predicted sentiment is Negative
    
```

Gambar 10. Pengujian Model Terhadap Kalimat Masukan

IV. KESIMPULAN

Berdasarkan hasil pengujian yang telah dilakukan, algoritma *Bidirectional LSTM* (BiLSTM) telah berhasil diimplementasi untuk kasus analisis sentimen terhadap pelayanan grab Indonesia dan terbukti menghasilkan akurasi yang lebih baik dari algoritma LSTM biasa yang menghasilkan akurasi 76%, sedangkan BiLSTM menghasilkan akurasi terbaik mencapai 91%. BiLSTM juga menghasilkan akurasi yang lebih baik dari algoritma *Multinomial Naïve Bayes*, *Logistic Regression* dan *Support Vektor Classifier* dalam kasus analisis sentiment pada pelayanan grab Indonesia dengan akurasi secara berturut-turut sebesar 63%, 67% dan 65%. Hasil tersebut didapatkan karena pada BiLSTM informasi dibaca secara 2 arah sekaligus. Tetapi BiLSTM juga memiliki kelemahan, yaitu membutuhkan dataset yang besar agar dapat menghindari *overfitting*, dan juga waktu dan biaya komputasi yang dibutuhkan juga cukup tinggi. Saran untuk penelitian yang berikutnya dapat menghasilkan representasi kata yang lebih banyak dan variatif dengan cara mempertimbangkan pada kombinasi *word embedding*-nya.

DAFTAR PUSTAKA

- [1] D. Puspitasari, D. Subhi, B. Novia, "Pengembangan Analisis Sentimen Pada Twitter Untuk Layanan Ojek Online Menggunakan Rule Based," SEMINAR INFORMATIKA APLIKATIF POLINEMA (SIAP), p. 2020.
- [2] S. Mandasari, B. H. Hayadi, and R. Gunawan, "Nomor 2," *Volume*, vol. 5, pp. 118–126, 2022, [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jsk/index>
- [3] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "ANALISIS SENTIMEN APLIKASI RUANG GURU DI TWITTER MENGGUNAKAN ALGORITMA KLASIFIKASI," *Jurnal Teknoinfo*, vol. 14, no. 2, p. 115, Jul. 2020, doi: 10.33365/jti.v14i2.679.
- [4] R. Naquitasia, D. Hatta Fudholi, and L. Iswari, "ANALISIS SENTIMEN BERBASIS ASPEK PADA WISATA HALAL DENGAN METODE DEEP LEARNING," 2022. [Online]. Available: <https://ejournal.teknokrat.ac.id/index.php/teknoinfo/index>
- [5] A. K. Ni Komang, "Classification of Public Figures Sentiment on Twitter using Big Data Technology," *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, vol. 6, no. 1, pp. 157–169, Jul. 2022, doi: 10.31289/jite.v6i1.7329.
- [6] R. Watrianthos, M. Giatman, W. Simatupang, R. Syafriyetti, and N. K. Daulay, "Analisis Sentimen Pembelajaran Campuran Menggunakan Twitter Data," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 1, p. 166, Jan. 2022, doi: 10.30865/mib.v6i1.3383.
- [7] S. Wahyunita, Y. Azhar, and N. Hayatin, "Analisa Sentimen Tweet Berbahasa Indonesia dengan Menggunakan Metode Pembobotan Hybrid TF-IDF pada Topik Transportasi Online," *REPOSITOR*, vol. 2, no. 2, pp. 185–192, 2020, [Online]. Available: <http://apps.twitter.com/>.
- [8] R. Wahyudi *et al.*, "Analisis Sentimen pada review Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine," *JURNAL INFORMATIKA*, vol. 8, no. 2, 2021, [Online]. Available: <http://ejournal.bsi.ac.id/ejurnal/index.php/ji>
- [9] B. Rahmatullah, ; Pungkas Budiyo, ; Suwanda, and A. Saputra, "SENTIMEN ANALISIS TRANSPORTASI ONLINE MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE, NAIVE BAYES DAN KNN," *Jurnal Ilmu Komputer JIK*, [Online]. Available: <https://t.co/ZYbdHbWj5>
- [10] N. Afrianto, D. H. Fudholi, and S. Rani, "Prediksi Harga Saham Menggunakan BiLSTM dengan Faktor Sentimen Publik," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 1, pp. 41–46, Feb. 2022, doi: 10.29207/resti.v6i1.3676.
- [11] D. Tri Hermanto, A. Setyanto, and E. T. Luthfi, "Algoritma LSTM-CNN untuk Sentimen Klasifikasi dengan Word2vec pada Media Online LSTM-CNN Algorithm for Sentiment Clasification with Word2vec On Online Media".
- [12] G. S. N Murthy, S. Rao Allu, B. Andhavarapu, M. Bagadi, and M. Belusonti, "Text based Sentiment Analysis using LSTM; Text based Sentiment Analysis using LSTM." [Online]. Available: www.ijert.org
- [13] M. Tri Anjasmos and dan Fitri Marisa, "Seminar Nasional Hasil Riset Prefix-RTR ANALISIS SENTIMEN APLIKASI GO-JEK MENGGUNAKAN METODE SVM DAN NBC (STUDI KASUS: KOMENTAR PADA PLAY STORE)," 2020.
- [14] K. K. Rekayasa and M. D. Hilmawan, "Terbit online pada laman web jurnal: <http://journal.itelkom-pwt.ac.id/index.php/dinda> Journal of Dinda Deteksi Sarkasme Pada Judul Berita Berbahasa Inggris Menggunakan Algoritme Bidirectional LSTM," *Data Institut Teknologi Telkom Purwokerto*, vol. 2, no. 1, pp. 46–51, 2022, [Online]. Available: <https://www.huffingtonpo>
- [15] K. Setyo Nugroho, I. Akbar, and A. Nizar Suksmawati, "Seminar Nasional Hasil Riset Prefix-RTR DETEKSI DEPRESI DAN KECEMASAN PENGGUNA TWITTER MENGGUNAKAN BIDIRECTIONAL LSTM," 2021.
- [16] F. Shahid, A. Zameer, and M. Muneeb, "Predictions for COVID-19 with Deep Learning models of LSTM, GRU and Bi-LSTM," *Chaos, Solitons and Fractals*, vol. 140, Nov. 2020, doi: 10.1016/j.chaos.2020.110212.
- [17] X. H. Le, H. V. Ho, G. Lee, and S. Jung, "Application of Long Short-Term Memory (LSTM) neural network for flood forecasting," *Water (Switzerland)*, vol. 11, no. 7, 2019, doi: 10.3390/w11071387.
- [18] H. Elfaiik and E. H. Nfaoui, "Deep Bidirectional LSTM Network Learning-Based Sentiment Analysis for Arabic Text," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 395–412, Jan. 2021, doi: 10.1515/jisys-2020-0021.

-
- [19] S. Ahmed *et al.*, “Att-bil-sl: Attention-based bi-lstm and sequential lstm for describing video in the textual formation,” *Applied Sciences (Switzerland)*, vol. 12, no. 1, Jan. 2022, doi: 10.3390/app12010317.
- [20] I. Markoulidakis, G. Kopsiaftis, I. Rallis, and I. Georgoulas, “Multi-Class Confusion Matrix Reduction method and its application on Net Promoter Score classification problem,” in *ACM International Conference Proceeding Series*, Jun. 2021, pp. 412–419. doi: 10.1145/3453892.3461323.
- [21] M. Y. Aldean, M. David Hilmawan, R. Indriyati, J. Lasama, and A. Junaidi, *Analisa Relevansi Tweet terhadap Hashtag dengan Metode Logistic Regression*.
- [22] W. Gouda, M. Almurafteh, M. Humayun, and N. Z. Jhanjhi, “Detection of COVID-19 Based on Chest X-rays Using *Deep Learning*,” *Healthcare (Switzerland)*, vol. 10, no. 2, Feb. 2022, doi: 10.3390/healthcare10020343.