

Analisis Sentimen Sosial Media dengan Metode Bidirectional Gated Recurrent Unit

Fadly Faturrohman¹, Dewi Rosmala²

^{1,2} Program Studi Informatika Institut Teknologi Nasional Bandung, Indonesia

Email : faturrohman309@itenas.ac.id

Received DD MM YYYY | Revised DD MM YYYY | Accepted DD MM YYYY

ABSTRAK

Sentimen analisis banyak diterapkan untuk dapat menganalisa opini dari suatu kalimat berbagai metode telah dicoba untuk melakukan klasifikasi di dalam sebuah teks. Salah satu metode yang populer digunakan adalah dengan deep learning, klasifikasi sentimen masuk kedalam ranah Natural Language Processing maka metode yang umum digunakan adalah metode recurrent. Dalam penelitian ini diusulkan metode Bidirectional GRU (BiGRU) untuk ekstraksi fitur teks dengan parameter yang diukur adalah perubahan nilai dari Learning Rate, Batch Size, dan Epoch. Dalam penerapan metodenya dibantu dengan word embedding GloVe dalam menghasilkan model terbaik, yang akhirnya digunakan untuk melakukan klasifikasi polaritas pada data sentimen dari sosial media Twitter. Hasil yang didapatkan pada model terbaik menghasilkan akurasi sebesar 97.7% dengan nilai precision 97.8%, recall 97.5% dan nilai loss 1.3%

Kata kunci: Sentimen Analisis, BiGRU, Word Embedding, GloVe, Natural Language Processing

ABSTRACT

Sentiment analysis is widely applied to be able to analyze the opinion of a sentence, various methods have been tried to classify in a text. One of the popular methods used is deep learning, sentiment classification is included in the realm of Natural Language Processing, so the method commonly used is the iterative method. In this study, the Bidirectional GRU (BiGRU) method is proposed for text feature extraction with the measured parameters being the change in the value of Learning Rate, Batch Size, and Epoch. The application of the method is assisted by word embedding GloVe in producing the best model, which is finally used to classify polarity on sentiment data from Twitter social media. The results obtained in the best model are 97.7% accuracy with 97.8% precision value, 97.5% recall and 1.3% loss value..

Keywords: Sentiment Analysis, BiGRU, Word Embedding, GloVe, Natural Language Processing

1.

PENDAHULUAN

Analisis sentimen dikenal juga sebagai penambangan opini merupakan salah satu bidang studi yang berkaitan dengan menganalisa pendapat, sentimen, evaluasi, penilaian, sikap, dan emosi orang terhadap entitas yang mewakili ruang masalah yang besar (Liu, 2012). Indikator sentimen juga termasuk dari kata sentimen atau disebut juga kata opini. Menurut Bing Liu Indikator sentimen ini adalah kata-kata yang biasa digunakan untuk mengekspresikan sentimen positif atau negatif (Liu, 2012). Seperti contohnya, "bagus" dan "luar biasa" adalah kata sentimen positif, "buruk" dan "mengerikan" merupakan kata dengan sentimen negatif.

Perkembangan bisnis serta organisasi selalu mencari tahu tentang opini konsumen mengenai produk serta layanan yang disediakan. Konsumen juga ingin mengetahui pendapat pengguna produk lain yang sudah pernah membeli atau memakainya, atau pendapat orang lain tentang kandidat politik sebelum mengambil memilihnya. Sebelumnya, untuk mengumpulkannya sebuah entitas perlu melakukan sebuah survei, jajak pendapat, serta diskusi grup guna memperoleh opini publik. Namun dengan pertumbuhan yang cepat dari media sosial individu dan organisasi semakin banyak menggunakan konten di media ini untuk pengambilan keputusan (Liu, 2012). Pendekatan untuk sentimen analisis yang diusulkan adalah dengan metode deep learning, pembelajaran mendalam yang khususnya digunakan untuk memproses teks adalah arsitektur Recurrent Neural Network (RNN). RNN merupakan jenis jaringan saraf tiruan di mana koneksi antar unit membentuk siklus dua arah (recurrent), dan melakukan tugas yang sama dari masing-masing unit untuk setiap elemen dalam urutan. Teknik RNN sangat optimal untuk masalah data sekuensial terutama untuk menangkap informasi dalam satuan waktu didalam loop (Zulqarnain et al., 2020). Walaupun RNN telah sangat efektif dalam tugas data sekuensial dan pengenalan suara terdapat kelemahan RNN yaitu *gradien vanishing* yang terjadi saat suatu fitur dilewatkan ke dalam sel secara berulang melalui banyak langkah waktu, nilainya cenderung membesar (exploding) atau menghilang (vanishing), sangat sulit untuk membuat RNN dapat mengolah data dalam langkah waktu yang lama. Di sisi lain, pengembangan untuk mengatasi masalah di tersebut telah diusulkan, seperti Long Short-Term Memory (LSTM) dan Gated Recurrent Unit (GRU). Arsitektur GRU dan LSTM populer dipakai karena memiliki kemampuan untuk menjaga memori dari aktivasi sebelumnya daripada mengganti seluruh nilai di dalam prosesnya seperti yang RNN lakukan (Lien Minh et al., 2018). Pada penelitian ini diusulkan arsitektur GRU karena berhasil menghindari masalah RNN juga masalah *overfitting*, serta menghemat waktu pelatihan (Zulqarnain et al., 2020).

Pada penelitian Zulqarnain dan kawan-kawan yang meneliti tentang klasifikasi sentimen dengan menggunakan dataset IMDB dan menggunakan varian model RNN yaitu LSTM dengan menerapkan perubahan alur yang disebut Bidirectional (BiLSTM) dan metode GRU dengan penambahan pada state/lapisannya yang disebut Two-State Gated Recurrent Unit (TGRU), mendapatkan hasil accuracy untuk BiLSTM yaitu 84.29 % dan untuk arsitektur TGRU yaitu 83.95 %.

Dari usulan penelitian sebelumnya yang menerapkan perkembangan arsitektur Bidirectional ternyata menghasilkan akurasi yang cukup baik pada kasus sentimen analisis maka dari itu pada penelitian ini diusulkan suatu sistem klasifikasi analisis sentimen untuk mendeteksi polaritas suatu sentimen berbahasa Indonesia dengan memakai arsitektur usulan yaitu Bidirectional Gated Recurrent Unit (BiGRU) yang mengubah dan mengamati beberapa penerapan parameter dan penelitian ini diukur dari performa model pada tingkat accuracy, loss, precision, recall, dan F1-Score yang dihasilkan dari hasil model hasil pelatihan.

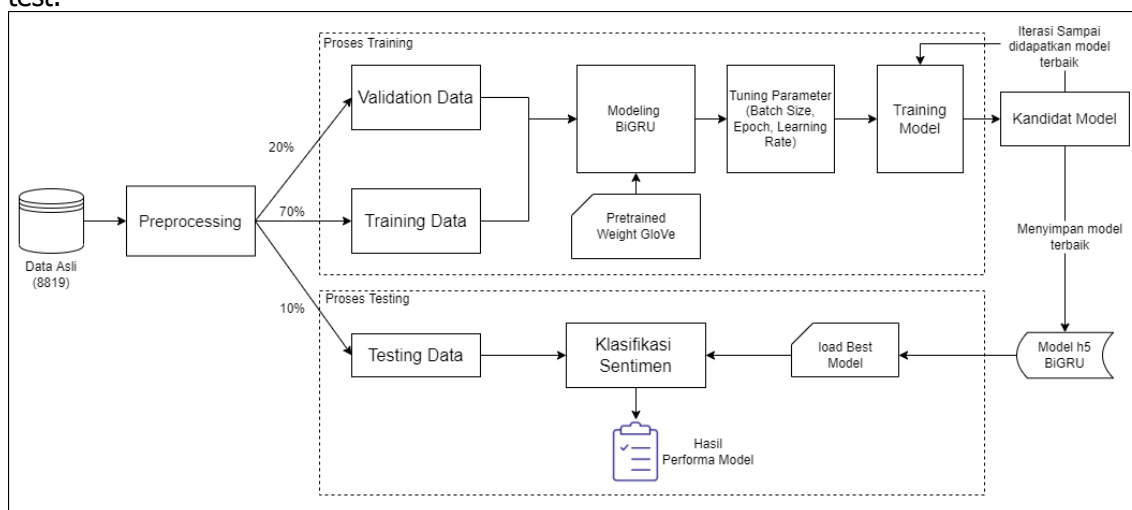
2.

METODE PENELITIAN

Pada tahapan ini dijelaskan metode yang terdiri dari pengumpulan data, perancangan, proses pra pemrosesan teks, proses pelatihan untuk mendapatkan model terbaik, dan proses pengujian model untuk melakukan evaluasi pada kinerja model. Penerapan arsitektur BiGRU dengan menggunakan dua variasi untuk menghasilkan model terbaik yang digunakan untuk klasifikasi polaritas sentimen analisis.

2.1. Perancangan

Tahap pada penelitian ini dibagi menjadi beberapa langkah seperti terlihat pada gambar 1 menunjukkan proses penerapan metode dari BiGRU dimulai dari proses input data dan sampai dengan langkah terakhir yang menghasilkan output performa model. Proses terbagi menjadi proses preprocessing, pemodelan, training model atau proses pelatihan dan proses testing yang menunjukkan hasil klasifikasi dari performa model menggunakan dataset yang sudah di split dengan rasio 70 % untuk data training, 20 % untuk data validasi, dan 10 % untuk data test.



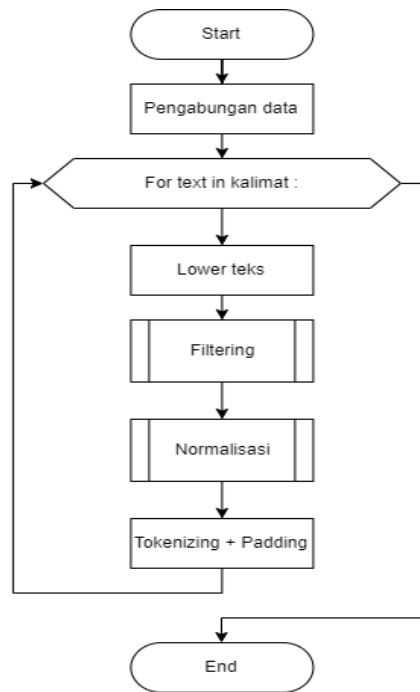
Gambar 1. Blok Diagram Perancangan Sistem

2.2. Data yang digunakan

Pengumpulan data didapatkan dari situs *github* yang mengambil data dari beberapa penelitian terdahulu yang datanya sudah dipublikasi di situs *github*, yang penelitiannya juga mengenai sentimen analisis yaitu pilkada DKI Jakarta (Rossi et al., 2017), Penggunaan layanan Telekomunikasi (Rofiqoh et al., 2017), *Cyberbullying* (Luqyana et al., 2018), Opini tayangan televisi (Nurjanah et al., 2017), Opini Film (Antinasari et al., 2017) Data yang digunakan berupa kumpulan teks dari sosial media *Twitter* yang telah diberi dianotasi atau label yang diambil dari sekitar tahun 2017 sampai 2018.

2.3. Preprocessing Data

Beberapa tahapan *preprocessing* yang digunakan untuk membentuk data bersih dari noise pada serangkaian teks. proses ditunjukkan seperti pada gambar 2. Dan dibagi menjadi 3 tahapan yaitu filtering, normalisasi, dan tokenisasi.



Gambar 2 Flowchart Preprocessing

Filterisasi berisi kumpulan proses pembersihan noise yang menggunakan library regular expression (regex) yang bekerja dengan melakukan match case dari karakter yang telah diset sebagai noise untuk dihilangkan dari teks. Prosesnya terbagi menjadi beberapa perintah yaitu mengubah kalimat menjadi huruf kecil (lowercase), pembersihan dari simbol, angka, dan tanda baca (puctuational removal). URL Removal menghapus URL atau hyperlinks. Proses lainnya seperti penghapusan username, hashtag, baris baru (new line), spasi berlebih dan simbol yang diulang.

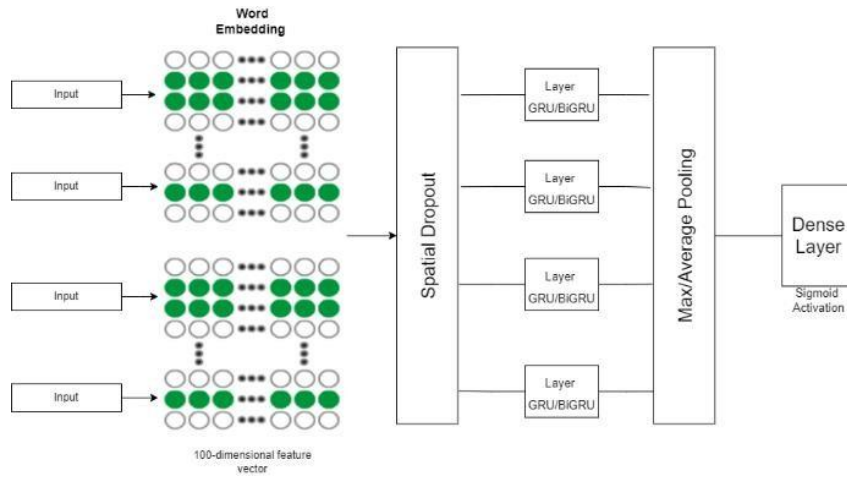
Pada teks di sosial media sering ditemui noise berupa kata yang tidak formal seperti "horeeeeeee" untuk mengekspresikan kebahagiaan namun kata tersebut tidak termasuk kedalam kata baku. Maka karena itu, dilakukan normalisasi dengan menghapus huruf yang berlebihan. Ciri kedua yang terkait dengan kata nonstandar adalah singkatan. contoh, biasanya orang menulis "g", "gk" atau "tdk" yang berarti kata "tidak". Selain itu, sering menggunakan kata untuk menyatakan "tidak" seperti "ngak" atau "ga". Juga, biasanya menggabungkan huruf dan angka, seperti "hati2", dan Anda harus "hati-hati" ketika harus mengulang "se7" (setuju).

Proses tokenisasi menggunakan fungsi dari keras tokenize yang akan bekerja dengan memisahkan serangkaian kalimat menjadi token kata terpisah dalam bentuk array agar bisa diproses dengan numerik di dalam sistem one-hot encoding dengan library keras yaitu Tokenize untuk mendapat data numerik dari suatu kata, dan dengan penerapan padding yang ditujukan untuk membuat semua data memiliki panjang yang sama

2.4 Perancangan Model

Pembuatan model dari arsitektur BiGRU yang digunakan untuk melakukan pelatihan dibagi ditambah dengan beberapa lapisan seperti lapisan word embedding, lapisan pooling juga dense layer. Word embedding yang dipakai matriks fitur pre-trained GLOVE yang berasal dari

korpus kata situs TEMPO (Kurniawan, 2019) dengan panjang dimensinya vektornya 100, susunan arsitektur dengan susunan terdapat pada gambar 3.

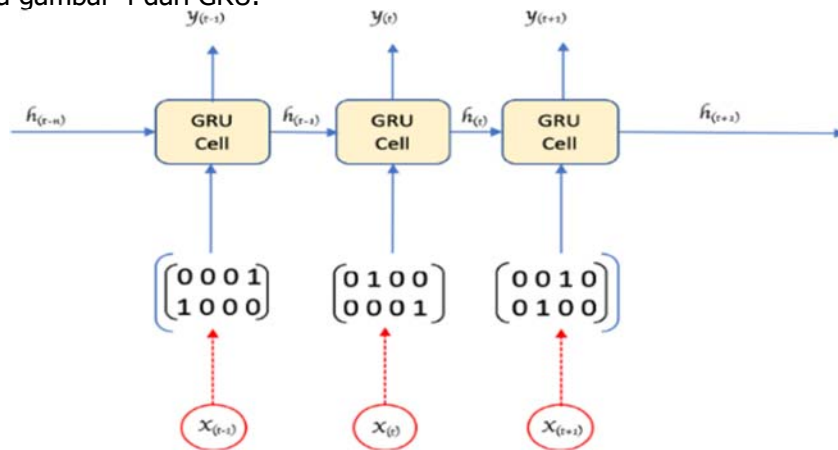


Gambar 3 Arsitektur Model

Model yang diusulkan memiliki lapisan yang disusun dari beberapa layer berbeda yang dijadikan satu dan dalam urutan pertama layer embedding yang menggunakan pre-trained Glove dengan dimensi vektor 100 untuk membuat matriks fitur dari data kontekstual.

Selanjutnya masuk ke layer spasial dropout 1D yang merupakan variasi lain dari layer dropout, namun layer ini menjatuhkan (dropping) nilai dari seluruh peta fitur 1D daripada elemen individual secara acak.

Setelah itu masuk kedalam lapisan GRU atau BiGRU yang merupakan unit-unit cell recurrent seperti pada gambar 4 dari GRU.



Gambar 4 Proses input layer GRU

Pada usulan penelitian ini model BiGRU pada tahapannya sedikit berbeda dengan GRU biasanya, BiGRU memiliki satu layer lebih banyak dari GRU dikarenakan proses ekstraksi fitur yang berjalan secara *forward* dan *backward* melalui dua arah yaitu layer pertama forward yang melakukan ekstraksi kata depan sampai akhir dan backward dari kata belakang sampai awal (Pasaribu et al., 2020). Lalu Dilakukan penggabungan dua fitur nilai dan operasi fungsi aktivasi ReLU. Arsitektur yang digunakan memakai formulasi hitungan yang terbagi menjadi 4 tahapan yaitu (Suyanto et al., 2019):

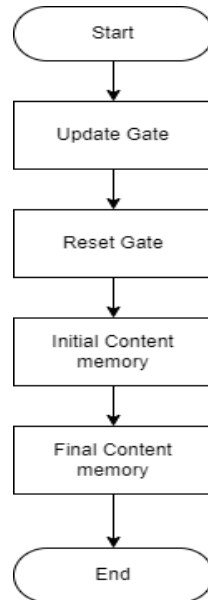
$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (1)$$

$$r_t = \sigma(w_r \cdot [h_{t-1}, x_t]) \quad (2)$$

$$\tilde{h}_t = \tanh(w \cdot [r_t * h_{t-1}, x_t]) \quad (3)$$

$$h_t = (1 - z_t) * z_t + h_{t-1} + h_t \quad (4)$$

Dalam arsitektur GRU digunakan gating system yang dapat mengatur aliran data seperti update gate (z) dan reset gate (r) urutannya proses formula disusun seperti pada gambar 5. Kedua gerbang digunakan untuk menentukan apakah informasi bisa dipakai lagi atau harus dibuang.



Gambar 5. Alur Arsitektur GRU

Setelah diproses di lapisan GRU diterapkan metode pooling dengan dua penerapan yaitu average pooling dan max pooling. Terakhir dense layer dengan aktivasi sigmoid untuk menekan nilai dari tiap peta fitur menjadi antara 0 sampai 1 untuk dapat menggunakan uji binary classification.

2.5 Pelatihan Model

Pelatihan model dilakukan dengan dalam beberapa skenario pengujian dengan parameter pengujian yang berbeda, konfigurasi pada hyperparameter yang diuji adalah Batch Size dari 32, 64, 128, Learning Rate dari 0.1, 0.001, 0.0001, dan epoch 25 dan 50. Dan parameter pembantu yang sudah ditetapkan nilainya yaitu nilai dropout pada angka 0.2 dan konfigurasi jumlah unit pada model BiGRU 32, dan menggunakan optimizer adam yang optimal untuk pelatihan gradient descent (Kingma & Ba, 2015).

2.6 Pengujian Model

Proses pengujian model pada penelitian ini dilakukan setelah mendapatkan model hasil pelatihan yang optimal atau terbaik di setiap skenario dengan mengukur performa model yang dibangun dengan arsitektur GRU. Pengujian pada model ini dilakukan dengan menerapkan menghitung nilai accuracy, precision, recall, dan f1-score sebagai tingkat kedekatan nilai prediksi dengan nilai aktual juga melihat dari nilai loss sebuah model untuk mengetahui tingkat salah dari modelnya. Performa model dapat dihitung dengan formula berikut (Pardede & Ibrahim, 2020) :

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

$$\text{F-Measure} = 2 \times \frac{TP + TN}{TP + FP + FN + TN} \quad (7)$$

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

Keterangan :

TP = True Positive

FP = False Positive

FN = False Negative

TN = True Negative

3.

HASIL DAN PEMBAHASAN

Pada implementasi penerapan arsitektur dengan menggunakan metode BiGRU dengan melihat nilai akurasi dan loss dari tiap parameter yang diamati dalam mencapai keadaan optimal, juga melihat performa dari model menggunakan confusion matrix. Pelatihan pertama yang dilakukan yaitu pelatihan dengan perbedaan learning rate seperti pada tabel 1.

Tabel 1 Hasil Pelatihan Pada Learning Rate Berbeda

Batch Size	Epoch	Learning rate	Acc	Loss
16	50	0.1	0.62	0.861
		0.01	0.843	0.351
		0.001	0.949	0.132

Pelatihan yang menggunakan parameter bantuan yaitu batch size 16 dan epoch 50 didapatkan hasil optimum pada parameter learning rate 0.001 di model BiGRU mendapatkan akurasi 94.9% dan nilai loss 13.2%. Parameter kedua yang diuji ada epoch pada angka 25 dan 50 dengan hasil yang didapatkan seperti pada tabel 2.

Tabel 2 Hasil Pelatihan Pada Batch Size Berbeda

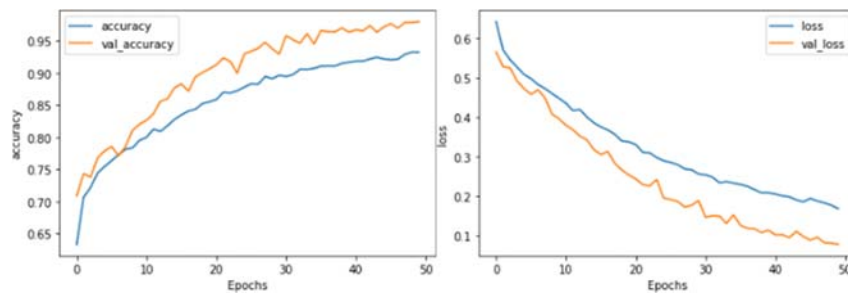
Learning Rate	Batch size	Epoch	Acc	Loss
16	16	25	0.897	0.246
		50	0.949	0.132

dengan parameter pembantu yaitu batch size 16 dan menggunakan parameter learning rate terbaik dari yaitu 0.001. didapatkan hasil optimal yaitu pada model BiGRU epoch 50 dengan akurasi mencapai 94,9 %. Skenario terakhir adalah perubahan nilai pada parameter batch size, dengan menggunakan parameter pembantu terbaik yaitu learning rate 0.001 dan epoch 50, dengan hasil yang didapatkan seperti pada tabel 3.

Learning Rate	Epoch	Batch Size	Acc	Loss
16	50	16	0.949	0.132
		32	0.935	0.170
		64	0.900	0.243
		128	0.838	0.362

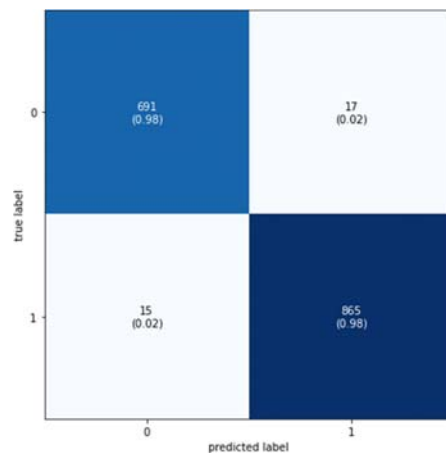
Dengan hasil pada model BiGRU terdapat pada angka batch size 16 dengan nilai akurasi tertinggi yaitu 94.9 %, dengan perubahan semakin besar nilai batch size tingkat akurasi dan loss semakin buruk juga.

Setelah didapatkan nilai optimum dari tiap parameter, dilakukan test dengan model terbaik dari pengukuran performa menggunakan grafik nilai akurasi dan loss juga confusion matrix. Untuk grafik model terbaik yang didapatkan hasil seperti pada gambar 6, yang menunjukkan model BiGRU membutuhkan epoch yang lebih banyak untuk mencapai nilai optimumnya.



Gambar 6. Grafik Accuracy dan Loss Model BiGRU

Pada penelitian nilai optimum dari tiap model ditentukan dari parameter terbaik yang telah diuji dan didapatkan dari sana tiap model ditentukan performanya dengan confusion matrix pada gambar 8



Gambar 7. Confusion Matrix model GRU

Dari confusion matrix yang terbagi menjadi dua label yaitu negatif (0) dan positif (1) didapatkan hasil false positive dan false negatif yang kecil untuk setiap labelnya. Dari confusion matrix bisa dilakukan perhitungan performa dengan hasil pada tabel 4.

Tabel 3. Hasil Performa model

Acc	Precision	Recall	F1-Score
97,7%	97,8%	97,5%	97,6%

Performa model dengan accuracy yang diperoleh yaitu 97,7% dengan precision pada 97.8%, recall pada 97.5% dan F1-Score yang mendapatkan nilai 97,6%. model BiGRU juga mendapatkan nilai loss 1,329% yang berarti selisihnya sekitar 0.499%. Untuk hasil test dengan menggunakan data test didapatkan untuk model BiGRU adalah nilai akurasi 80% yang diuji dengan data test.

4.

KESIMPULAN

Perancangan sistem analisis sentimen dengan arsitektur GRU menggunakan model BiGRU telah mampu melakukan klasifikasi sentimen berbahasa Indonesia pada dataset dari sosial media twitter, dalam dua kelas yang berbeda yaitu positif dan negatif. Kedua model yang dibangun berhasil memberikan nilai performa optimum dari parameter terbaik yang didapatkan dan menghasilkan model terbaik dari proses pelatihan. Pengujian sistem model pada skenarionya memperoleh nilai Accuracy tertinggi menggunakan parameter terbaik pada learning rate 0,001, batch size 16, dan epoch 50. Pada nilai accuracy terbaik yang didapatkan 97.6 % lalu nilai performa yang didapatkan precision 97,8%, Recall 97,5%, dan F-Measure 97.6% terakhir untuk nilai loss yang didapatkan adalah 1,329% pada model BiGRU. Dari hasil yang didapatkan dapat disimpulkan model BiGRU yang dibangun sudah cukup optimal untuk melakukan klasifikasi sentimen dengan akurasi yang baik.

DAFTAR PUSTAKA

- Antinasari, P., Perdana, R. S., & Fauzi, M. A. (2017). Analisis Sentimen Tentang Opini Film Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Naive Bayes Dengan Perbaikan Kata Tidak Baku. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 1(12), 1718–1724. <http://j-ptiik.ub.ac.id>
- Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–15.
- Kurniawan, K. (2019). *KaWAT: A Word Analogy Task Dataset for Indonesian*. 2013–2015. <http://arxiv.org/abs/1906.09912>
- Lien Minh, D., Sadeghi-Niaraki, A., Huy, H. D., Min, K., & Moon, H. (2018). Deep learning approach for short-term stock trends prediction based on two-stream gated recurrent unit network. *IEEE Access*, 6, 55392–55404. <https://doi.org/10.1109/ACCESS.2018.2868970>
- Liu, B. (2012). Sentiment Analysis and Opinion Mining. In *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions* (Issue May). Morgan & Claypool Publishers. <https://doi.org/10.1017/CBO9781139084789>
- Luqyana, W. A., Cholissodin, I., & Perdana, R. S. (2018). Analisis Sentimen Cyberbullying Pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*,

2(11), 4704–4713.

- Nurjanah, W. E., Perdana, R. S., & Fauzi, M. A. (2017). Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 1(12), 1750–1757.
- Pardede, J., & Ibrahim, R. G. (2020). Implementasi Long Short-Term Memory untuk Identifikasi Berita Hoax Berbahasa Inggris pada Media Sosial. *Journal of Computer Science and Informatics Engineering (J-Cosine)*, 4(2), 179–187. <https://doi.org/10.29303/jcosine.v4i2.361>
- Pasaribu, D. J. M., Kusriani, K., & Sudarmawan, S. (2020). Peningkatan Akurasi Klasifikasi Sentimen Ulasan Makanan Amazon dengan Bidirectional LSTM dan Bert Embedding. *Inspiration: Jurnal Teknologi Informasi Dan Komunikasi*, 10(1), 9–20. <https://doi.org/10.35585/inspir.v10i1.2568>
- Rofiqoh, U., Perdana, R. S., & Fauzi, M. A. (2017). Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexion Based Feature. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 1(12), 1725–1732. <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/628>
- Rossi, A., Lestari, T., Setya Perdana, R., & Fauzi, M. A. (2017). Analisis Sentimen Tentang Opini Pilkada DKI 2017 Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Naïve Bayes dan Pembobotan Emoji. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 1(12), 1718–1724. <http://j-ptiik.ub.ac.id>
- Suyanto, Ramadhani, K. N., & Mandala, S. (2019). *Deep Learning Modernisasi Machine Learning Untuk Big Data*. Penerbit Informatika.
- Zulqarnain, M., Abd, S., Ghazali, R., Mohd, N., Aamir, M., & Mazwin, Y. (2020). An Improved Deep Learning Approach based on Variant Two-State Gated Recurrent Unit and Word Embeddings for Sentiment Classification. *International Journal of Advanced Computer Science and Applications*, 11(1), 594–603. <https://doi.org/10.14569/ijacsa.2020.0110174>