

# Pemrosesan Bahasa Alami Untuk Analisis Sentimen Dengan Menggunakan Bidirectional GRU

Mohamad Maulana Firdaus Ramadhan  
Sekolah Teknik Elektro dan  
Informatika  
Institut Teknologi Bandung  
Bandung, Indonesia  
18222140@mahasiswa.itb.ac.id

Rizqi Andhika Pratama  
Sekolah Teknik Elektro dan  
Informatika  
Institut Teknologi Bandung  
Bandung, Indonesia  
18222118@mahasiswa.itb.ac.id

Rizqika Mulia Pratama  
Sekolah Teknik Elektro dan  
Informatika  
Institut Teknologi Bandung  
Bandung, Indonesia  
13522126@mahasiswa.itb.ac.id

Ikhwan Al Hakim  
Sekolah Teknik Elektro dan  
Informatika  
Institut Teknologi Bandung  
Bandung, Indonesia  
13522147@mahasiswa.itb.ac.id

**Abstract**— Sosial media telah menjadi bagian penting dalam keseharian manusia saat ini, banyak aplikasi yang digunakan untuk membagikan konten berupa opini dan pengalaman pribadi, salah satunya adalah Twitter. Penulis mempunyai ide untuk mencari solusi terkait permasalahan sehari-hari yang terjadi di masyarakat dengan memanfaatkan fenomena tersebut. Pada penelitian ini, penulis melakukan analisis sentimen dalam pemrosesan bahasa alami untuk mengetahui opini dan sikap masyarakat di Twitter terhadap isu-isu penting menggunakan model Bidirectional Gated Recurrent Unit (GRU). Metode ini digunakan untuk menggali pemahaman yang lebih dalam terhadap sentimen dalam teks, dengan memanfaatkan arah maju dan mundur dalam GRU untuk menangkap konteks kata-kata secara menyeluruh. Kemudian, kata-kata berupa teks yang ditangkap tersebut diproses melalui tahap data preprocessing untuk meningkatkan kualitas data sebelum dilakukan modelling. Model yang dilatih adalah jaringan saraf tiruan (neural network) dengan arsitektur BiGRU. Evaluasi model ditinjau dari metrics Classification Report. Pada dataset yang dipakai, BiGRU memberikan hasil yang baik dengan akurasi sebesar 0.70, precision rata-rata 0.72, dan recall rata-rata 0.71.

**Keywords**—Sosial media, BiGRU, Twitter, jaringan saraf tiruan, analisis sentimen

## I. PENDAHULUAN

### A. Latar Belakang

Media sosial saat ini menjadi pilihan baru bagi orang-orang untuk menjalin hubungan dan berkomunikasi. Karena itu, angka pengguna media sosial terus meningkat dengan cepat tiap tahunnya. Twitter adalah salah satu media sosial yang mengalami peningkatan pengguna terbesar. Menurut Stacy Jo Dixon, Twitter memiliki sekitar 368 juta pengguna aktif bulanan di seluruh dunia pada Desember 2022, dan diperkirakan akan meningkat menjadi sekitar 497 juta pada tahun 2025. Isi dari postingan Twitter, yang disebut “tweet”, telah banyak digunakan oleh para peneliti, pemerintah, atau industri untuk mendapatkan pengetahuan yang membantu mereka menangani masalah-masalah sehari-hari. Berbagai sentimen dari masyarakat dapat dilihat dari tweet mereka.

Teknologi pemrosesan bahasa alami (*Natural Language Processing*, NLP) telah berkembang pesat di era digital saat ini dan memberikan dampak yang luas pada berbagai bidang kehidupan manusia. Salah satu bidang yang mendapat manfaat dari teknologi NLP adalah analisis sentimen. Analisis sentiment adalah proses mengidentifikasi dan mengklasifikasikan opini, emosi, dan sikap yang terkandung dalam teks. Analisis sentimen dapat digunakan untuk menganalisis teks yang dihasilkan oleh pengguna internet di berbagai platform, seperti media sosial, ulasan daring, blog, dan lain-lain.

Analisis sentimen memiliki nilai penting dalam memahami pandangan dan perasaan masyarakat terhadap berbagai hal, seperti produk, layanan, isu, atau entitas tertentu. Dengan mengetahui sentiment masyarakat, organisasi dan pemerintah dapat meningkatkan kualitas dan kinerja mereka, menangani keluhan dan masalah yang ada, serta menciptakan hubungan yang baik dengan masyarakat. Salah satu tantangan dalam analisis sentimen adalah mengatasi masalah konteks, ironi, dan sarkasme yang dapat mempengaruhi makna sebuah kalimat.

Salah satu metode untuk mengatasi masalah tersebut adalah dengan menggunakan bidirectional GRU. Bidirectional GRU adalah jenis *recurrent neural network* (RNN) yang dapat mengingat informasi jangka panjang dan jangka pendek dalam proses pembelajaran. Hal ini dimungkinkan karena *bidirectional GRU* membaca teks dari dua arah, yaitu dari awal ke akhir dan dari akhir ke awal. Dengan demikian, bidirectional GRU dapat memanfaatkan informasi konteks dari kedua arah dan menghasilkan representasi vektor yang lebih kaya untuk setiap kata.

### B. Rumusan Masalah

1. Bagaimana cara mengimplementasikan teknologi NLP untuk mengekstrak dan menginterpretasikan opini, emosi, dan sikap masyarakat terhadap isu-isu penting yang mempengaruhi masyarakat?

2. Bagaimana cara menerapkan *Bidirectional GRU* untuk analisis sentimen pada teks Twitter?
3. Bagaimana performa *Bidirectional GRU* dalam analisis sentiment pada teks Twitter?

#### C. Tujuan

1. Mengimplementasikan teknologi NLP metode *Bidirectional GRU* untuk mengekstrak dan menginterpretasikan opini, emosi, dan sikap masyarakat terhadap isu-isu penting yang mempengaruhi masyarakat
2. Mengimplementasikan *Bidirectional GRU* untuk analisis sentiment pada tingkat aspek atau entitas pada teks Twitter berbahasa Indonesia.

#### D. Manfaat

1. Mengetahui sentimen masyarakat terhadap isu-isu penting yang mempengaruhi kesejahteraan masyarakat.
2. Menggunakan sentimen masyarakat sebagai tolak ukur kepuasan masyarakat terhadap kinerja pemerintah atau organisasi.
3. Memanfaatkan informasi sentimen masyarakat sebagai dasar pengambilan keputusan yang lebih baik terutama untuk perbaikan dalam berbagai bidang, seperti kesehatan, pendidikan, bisnis, dan layanan masyarakat.

## II. PEMBAHASAN

### A. Landasan Teori

#### 1) Twitter

Twitter adalah platform media sosial yang memungkinkan pengguna untuk berbagi pesan pendek yang disebut "tweets" dalam format teks, gambar, atau video. Twitter adalah alat komunikasi yang populer dan memiliki dampak yang signifikan pada berbagai aspek sosial, politik, dan budaya. Twitter sering digunakan sebagai platform untuk menyebarkan dan memengaruhi opini publik tentang berbagai topik, termasuk politik, berita, dan isu-isu sosial. Dalam konteks ini, Twitter dapat memainkan peran penting dalam membentuk pandangan masyarakat.

#### 2) Emosi

Emosi adalah reaksi psikologis dan fisik terhadap situasi, perasaan, atau stimulus tertentu yang melibatkan perubahan mood, ekspresi wajah, perubahan fisiologis, dan pengalaman subjektif. Penelitian tentang emosi melibatkan berbagai bidang, termasuk psikologi, ilmu neurologi, ilmu kognitif, dan ilmu sosial. Ekman (1992) mengemukakan bahwa Teori Emosi Dasar adalah pendekatan yang mengidentifikasi sejumlah emosi dasar yang mendasari spektrum emosi manusia, seperti kebahagiaan, sedih, marah, takut, dan terkejut.

#### 3) Analisis Sentimen

Analisis sentimen adalah proses komputasional untuk mengevaluasi dan mengidentifikasi perasaan atau pendapat yang terkandung dalam teks, seperti ulasan produk, komentar sosial media, atau artikel berita, dengan tujuan mengklasifikasikan perasaan tersebut menjadi kategori sentimen utama seperti positif, negatif, atau netral. Melalui pendekatan berbasis aturan, statistik, atau machine learning, analisis sentimen memungkinkan penggunaan komputer

untuk secara otomatis memahami dan mengukur respons atau reaksi terhadap berbagai subjek atau entitas.

#### 4) Recurrent Neural Network

Recurrent Neural Network (RNN) adalah jenis arsitektur jaringan saraf yang digunakan untuk memproses data urutan, seperti teks, waktu, atau data deret waktu. RNN memiliki kapabilitas untuk mengingat informasi dari masa lalu dan menggunakannya dalam pemrosesan yang lebih lanjut. Hal ini membuatnya cocok untuk berbagai aplikasi yang melibatkan data urutan. RNN memiliki "memori" internal yang memungkinkannya mengingat informasi dari langkah-langkah sebelumnya dalam urutan data. Hal ini dicapai melalui sambungan siklik dalam arsitektur, yang memungkinkan informasi kembali ke dirinya sendiri.

#### 5) Bidirectional GRU

Bidirectional GRU (Gated Recurrent Unit) adalah jenis model jaringan saraf rekuren (RNN) yang digunakan dalam pemrosesan urutan data, seperti teks, audio, atau waktu. Model ini terdiri dari dua lapisan GRU yang saling terhubung, satu untuk mengolah urutan data maju dan satu lagi untuk urutan data mundur. Model ini juga memungkinkan informasi untuk mengalir maju (dari masa lalu ke masa depan) dan mundur (dari masa depan ke masa lalu) dalam jaringan, sehingga memungkinkannya untuk memahami konteks dari kedua arah.

#### 6) Word Embedding

Word embedding adalah teknik dalam pemrosesan bahasa alami (Natural Language Processing, NLP) yang digunakan untuk merepresentasikan kata-kata dalam bentuk vektor numerik dalam ruang dimensi rendah. Teknik ini berguna untuk mengubah kata-kata menjadi representasi vektor yang memiliki makna semantik, sehingga mesin dapat lebih baik memahami hubungan antara kata-kata dalam teks.

#### 7) Classification Report

Classification Report merupakan salah satu metrik evaluasi performa dari model machine learning berbasis klasifikasi. Metrik ini memberikan informasi yang komprehensif mengenai precision, recall, F1 score, dan support untuk setiap kelas yang ada dalam dataset.

Precision (presisi) mengukur sejauh mana model dapat mengidentifikasi dengan benar anggota kelas tertentu. Presisi dihitung dengan membagi jumlah true positive dengan jumlah true positive dan false positive. Semakin tinggi nilai presisi, semakin sedikit false positive yang dilakukan oleh model.

Recall mengukur sejauh mana model dapat mengenali secara lengkap anggota kelas tertentu. Recall dihitung dengan membagi jumlah true positive dengan jumlah true positive dan false negative. Semakin tinggi nilai recall, semakin sedikit false negative yang dilakukan oleh model.

F1 Score adalah rata-rata harmonik terbobot dari presisi dan recall. F1 Score menggabungkan kedua metrik tersebut menjadi satu nilai yang mencerminkan performa keseluruhan model. Semakin mendekati nilai F1 Score ke angka 1.0, semakin baik performa yang diharapkan dari model.

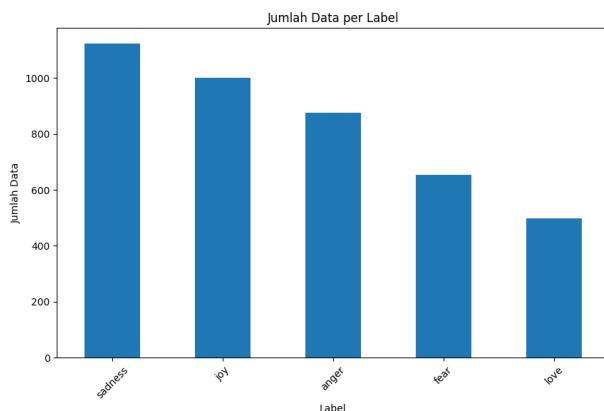
Support adalah jumlah kemunculan aktual dari setiap kelas dalam dataset. Meskipun support tidak berubah antara model, metrik ini berguna dalam proses evaluasi performa untuk memberikan konteks tentang seberapa sering kelas tersebut muncul dalam dataset.

## B. Metode Penelitian

### 1) Pengumpulan Data

Data yang digunakan dalam penelitian ini berisi dua kolom yaitu cuitan dari beberapa pengguna media sosial twitter dan kolom label yang berisi emosi yang ditunjukkan dari cuitan tersebut. Emosi tersebut diantaranya joy, anger, fear, love dan sadness. Dataset ini berisi 5153 baris dan 2 kolom.

Data ini kemudian dibagi menjadi 2, data pertama terdiri dari 1000 baris terakhir tidak memiliki label yang akan diprediksi nantinya dan diberi nama data\_predict, dan data lainnya akan dipakai sebagai data pelatihan dan validasi untuk membangun sebuah model prediksi yang diberi nama data\_train berjumlah 4153. Berikut ini merupakan distribusi dari kolom label dari dataset.



Gambar 1. Persebaran label

Dapat dilihat bahwa label memiliki jumlah yang berbeda dimana "sadness" memiliki 1123 sampel, "joy" sebanyak 1002, "anger" sebanyak 874, "fear" sebanyak 654, dan "love" sebanyak 498.

Selain itu, Teks dalam dataset ini telah diproses sebelumnya menggunakan beberapa kriteria. Nama pengguna (@) digantikan dengan istilah [USERNAME]. URL telah digantikan dengan istilah [URL]. Angka-angka pribadi, seperti nomor telepon dan faktur, telah digantikan dengan istilah [SENSITIVE-NO]. Berikut merupakan salah satu contoh dari data yang ada.

love,"[USERNAME] Centil, rempong abis, gatau malu, cerewet parah, temen usil, koki terbaik sepanjang masa, yg paling sayang sm aku, yang bela belain nganter kesana kemari, dan 1001++ hal lain yg gabisa diungkapkan dengan kata kata wkwk"

### 2) Data Preprocessing

Dalam dunia analisis data, terdapat sebuah prinsip yang dikenal dengan istilah "Garbage In, Garbage Out" yang berarti bahwa jika data yang dimasukkan bermasalah atau tidak berkualitas, maka hasil analisis yang dihasilkan juga akan bermasalah. Oleh karena itu, tahap data preprocessing merupakan langkah yang sangat penting dalam proses analisis data. Data preprocessing adalah proses penyesuaian data mentah menjadi data yang siap digunakan oleh model serta meningkatkan kualitas data tersebut.. Dalam penelitian kali ini digunakan beberapa tahapan data preprocessing diantaranya :

#### a) Menghilangkan data null

Data yang mengandung nilai kosong atau null tidak dapat digunakan dalam proses pelatihan model. Oleh

karena itu, langkah pertama adalah mengidentifikasi dan menghapus data yang memiliki nilai null.

#### b) Menghilangkan data duplikat

Data yang memiliki duplikat tidak memberikan variasi tambahan yang berguna dalam pembelajaran model. Oleh karena itu, data duplikat akan diidentifikasi dan dihapus untuk meningkatkan variasi data yang digunakan dalam proses pembelajaran model.

#### c) Case Folding

Case folding merupakan sebuah tahap untuk mengubah semua huruf yang ada menjadi huruf kecil. Case folding membantu menghindari masalah sensitivitas huruf besar-kecil saat mencari kata kunci dalam dokumen atau saat membandingkan teks. Dengan mengubah semua huruf menjadi huruf kecil, kita memastikan bahwa pencarian dan perbandingan teks dapat dilakukan dengan benar tanpa memperhatikan perbedaan huruf besar-kecil.

#### d) Menghapus simbol simbol dan karakter lain

Tahap ini melibatkan penghapusan semua karakter yang tidak relevan atau tidak diinginkan dalam teks. Ini termasuk tanda baca, angka numerik, karakter khusus, format [URL] dan semacamnya, dan mengubah emoji menjadi representasi teks. Hal ini membantu menghilangkan noise atau informasi yang tidak relevan dalam teks, sehingga mendorong pemrosesan data yang lebih fokus dan akurat.

#### e) Konversi Kata

Konversi kata merupakan proses mengubah kata yang tidak baku menjadi kata baku. Tahapan ini salah satunya dilakukan dengan menggunakan bantuan kamus singkatan. Jika terdapat kata tidak baku yang terdeteksi maka akan diubah sesuai dengan kamus singkatan tersebut.

Kedua, tahap konversi slangword dilakukan dengan menggunakan pustaka indoNLP. Pustaka ini dapat digunakan untuk melakukan proses normalisasi kata-kata slang dan mereduksinya ke bentuk dasar atau kata baku.

#### f) Menghapus Stopword

Stopword adalah kata-kata yang dianggap tidak penting atau kontribusi informasi yang rendah dalam analisis teks. Oleh karena itu, tahap penghapusan stopwords dilakukan untuk menghilangkan kata-kata tersebut dari teks. Penghapusan stopwords dilakukan dengan menggunakan dua pustaka, yaitu NLTK dan indoNLP.

#### g) Tokenisasi

Tokenisasi dimaknai sebagai mengubah teks menjadi representasi numerik. Ini karena setelah tokenisasi, setiap token dalam teks diberikan kode numerik yang unik sesuai dengan kamus kata-kata yang digunakan dalam pemrosesan. Representasi numerik ini memungkinkan komputer untuk melakukan berbagai tugas analisis teks, seperti pemodelan bahasa, analisis sentimen, atau klasifikasi teks.

#### h) Padding

Padding merupakan penambahan nilai (0) ke dalam sebuah teks agar memiliki panjang yang sama untuk seluruh teks.

### i) Word Embedding

Word embedding adalah teknik dalam pemrosesan bahasa alami (Natural Language Processing, NLP) yang digunakan untuk merepresentasikan kata-kata atau token dalam teks ke dalam bentuk vektor numerik. Representasi ini memungkinkan komputer untuk memahami dan memproses makna kata-kata dalam konteks matematis. Dalam penelitian kali ini word embedding yang digunakan yaitu fasttext yang didapat dari laman

### j) RandomOverSampler

RandomOverSampler merupakan teknik untuk pengelolaan ketidakseimbangan kelas dalam dataset. Ini adalah pendekatan yang digunakan untuk mengatasi masalah di mana jumlah sampel dalam kelas minoritas sangat sedikit dibandingkan dengan kelas mayoritas.

### 3) Modelling

Model utama yang digunakan dalam penelitian ini adalah jaringan saraf tiruan (neural network) dengan arsitektur BiGRU atau Bidirectional GRU. Model ini dikembangkan untuk tugas klasifikasi teks, yang bertujuan untuk mengklasifikasikan teks ke dalam kategori yang sesuai. Arsitektur model ini memiliki beberapa komponen inti yang bekerja sama untuk mencapai tujuan ini.

Pertama, kita memiliki lapisan Embedding yang bertanggung jawab untuk mengubah kata-kata dalam teks menjadi representasi vektor numerik. Representasi ini didasarkan pada matriks embedding yang telah dipersiapkan sebelumnya, yaitu matriks embedding FastText. Lapisan ini memungkinkan model untuk memahami makna kata dalam konteks tugas klasifikasi.

Kemudian, kita memiliki lapisan Bidirectional GRU, yang merupakan lapisan rekuren. Lapisan ini memiliki kemampuan untuk memahami hubungan kontekstual dalam teks, tidak hanya dari kiri ke kanan, tetapi juga sebaliknya. Ini membantu model dalam menangkap struktur dan pola yang lebih kompleks dalam teks.

Untuk menghindari overfitting, kita menggunakan lapisan Dropout setelah lapisan rekuren. Dropout secara acak mematikan beberapa unit selama pelatihan, sehingga mengurangi risiko overfitting.

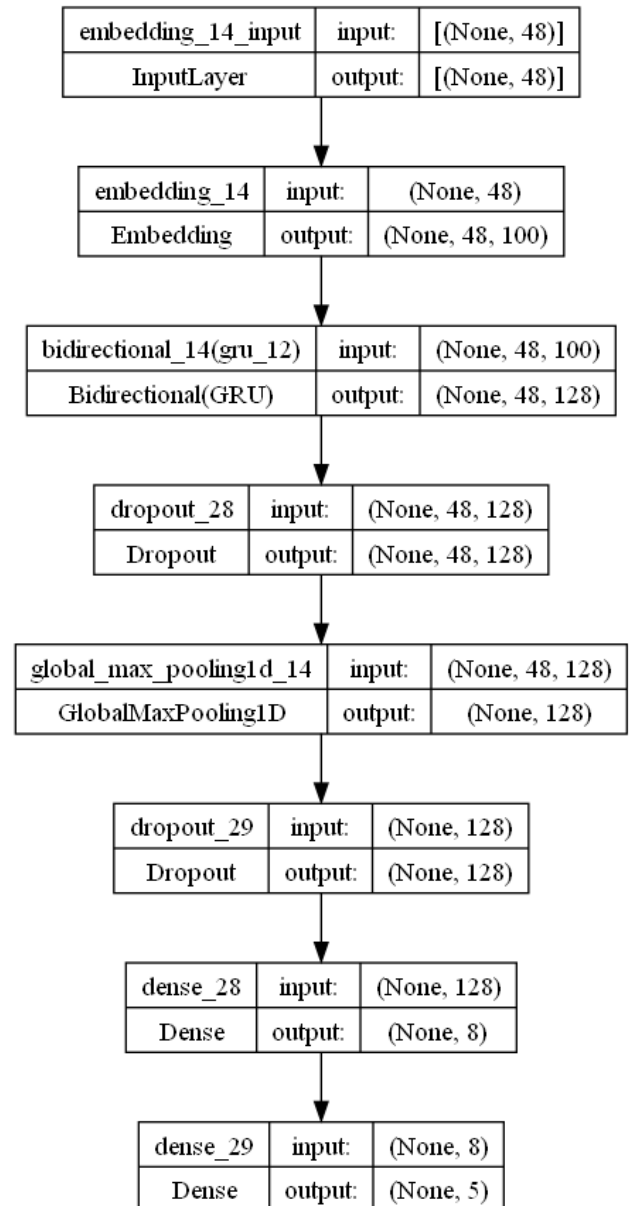
Selanjutnya, ada lapisan GlobalMaxPooling1D yang bertujuan untuk mengambil fitur terpenting dari urutan kata. Lapisan ini berguna dalam mengurangi dimensi data dan fokus pada fitur-fitur yang paling relevan.

Selanjutnya, kita memiliki dua lapisan Dense yang merupakan lapisan jaringan saraf biasa dengan aktivasi ReLU. Lapisan ini berfungsi untuk memproses fitur-fitur yang telah diekstrak sebelumnya.

Lapisan terakhir adalah lapisan Dense dengan aktivasi softmax. Ini adalah lapisan output yang menghasilkan distribusi probabilitas untuk setiap kategori klasifikasi. Model ini mengoptimalkan fungsi kerugian binary cross-entropy menggunakan optimizer Adam dengan laju pembelajaran 0.001.

Selama pelatihan, model dipantau menggunakan beberapa metrik, termasuk fungsi kerugian (loss) dan akurasi, pada data pelatihan dan data validasi. Kita juga menggunakan tiga callback, yaitu ModelCheckpoint untuk menyimpan model terbaik, ReduceLROnPlateau untuk mengurangi laju pembelajaran jika perlu, dan EarlyStopping untuk menghentikan pelatihan jika tidak ada peningkatan yang

signifikan dalam data validasi. Berikut merupakan visualisasi model yang dibuat.



Gambar 2. Visualisasi Model

### C. Hasil dan Pembahasan

Dalam tahap pra-pemrosesan data, sejumlah langkah kritis telah diambil untuk mempersiapkan data teks secara optimal sebelum digunakan dalam pelatihan model. Langkah pertama adalah menghilangkan data duplikat dan data kosong. Langkah kedua adalah menghilangkan karakter non-alfabet dan mengubah semua huruf menjadi huruf kecil. Hal ini dilakukan untuk memastikan konsistensi dalam representasi teks, sehingga model dapat dengan mudah memahami dan memprosesnya. Selanjutnya, dilakukan penghapusan kata-kata yang dianggap tidak relevan dalam konteks analisis, serta menggantinya dengan kata-kata yang memiliki makna yang lebih signifikan. Tindakan ini membantu meningkatkan kualitas data dengan fokus pada informasi penting, sehingga memungkinkan model untuk memahami dan mempelajari pola yang lebih bermakna. Hasil akhir dari langkah-langkah ini adalah data yang lebih bersih, lebih terstruktur, dan siap digunakan dalam pelatihan model

klasifikasi emosi. Dengan cara ini, kita dapat memaksimalkan potensi model untuk memahami dan mengklasifikasikan teks dengan akurat ke dalam berbagai kelas emosi. Data yang semula berbentuk seperti ini :

“[USERNAME] Centil, rempong abis, gatau malu, cerewet parah, temen usil, koki terbaik sepanjang masa, yg paling sayang sm aku, yang bela belain nganter kesana kemari, dan 1001++ hal lain yg gabisa diungkapkan dengan kata kata wkwk”

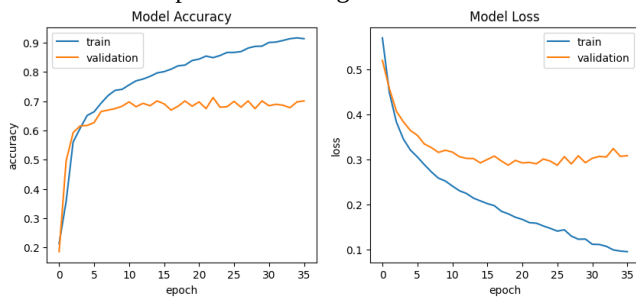
menjadi :

“centil repot habis malu cerewet parah teman usil koki terbaik sayang bela membela kesana kemari”

Selanjutnya, dalam tahap tokenisasi, setiap kata atau token dalam teks diberikan kode numerik unik sesuai dengan kamus kata-kata yang digunakan dalam pemrosesan. Representasi numerik ini memungkinkan komputer untuk melakukan berbagai tugas analisis teks. Terakhir, word embedding, khususnya FastText, digunakan untuk merepresentasikan kata-kata sebagai vektor numerik, memungkinkan model memahami makna kata dalam konteks matematis. Semua langkah ini bertujuan mempersiapkan data yang lebih bersih, lebih terstruktur, dan siap digunakan dalam pelatihan model klasifikasi emosi, sehingga memaksimalkan potensi model untuk memahami dan mengklasifikasikan teks dengan akurat ke dalam berbagai kelas emosi.

Data pelatihan telah dibagi menjadi dua bagian: satu untuk pelatihan model dan satu lagi untuk validasi. Setelah penerapan oversampling, data pelatihan telah diperluas menjadi 4590 baris dengan setiap kelas memiliki 918 sampel. Data validasi terdiri dari 608 baris.

Model klasifikasi telah dilatih dengan 50 epoch dan batch size sebesar 64. Early stopping dengan patience sebesar 10 digunakan untuk mencegah overfitting, yang berarti pelatihan akan berhenti jika tidak ada peningkatan yang signifikan dalam performa model selama 10 epoch berturut-turut. Hasil pelatihan model dapat dilihat sebagai berikut.



Gambar 3. Grafik hasil pelatihan model

Dapat dilihat, terjadi peningkatan accuracy dan penurunan loss seiring dengan bertambahnya epoch. Namun ketika mencapai epoch ke 26 model tidak mengalami perkembangan yang baik. Meskipun accuracy dan loss model mengalami peningkatan namun tidak untuk validasinya. Yang mengindikasikan model mengalami overfitting, yaitu ketika model "terlalu" mempelajari data pelatihan dan kehilangan kemampuan untuk melakukan generalisasi yang baik pada data yang tidak pernah dilihat sebelumnya. Maka dari itu model terbaik diambil dari epoch tersebut sesuai dengan penggunaan EarlyStopping. Selain itu digunakan pula metrik classification report untuk mengukur performa model. Didapatkan hasil sebagai berikut.

Tabel 1. Classification metrics

| Label     | Precision | Recall | F1-Score | Support |
|-----------|-----------|--------|----------|---------|
| Anger     | 0.65      | 0.85   | 0.74     | 134     |
| Fear      | 0.84      | 0.72   | 0.77     | 86      |
| Joy       | 0.74      | 0.67   | 0.70     | 157     |
| Love      | 0.76      | 0.72   | 0.74     | 75      |
| Sadness   | 0.62      | 0.58   | 0.60     | 156     |
| Macro avg | 0.72      | 0.71   | 0.71     | 608     |

Tabel 1 menunjukkan hasil evaluasi model klasifikasi untuk masing-masing kelas emosi, yaitu Anger, Fear, Joy, Love, dan Sadness, beserta nilai rata-rata (macro avg) dari metrik precision, recall, dan F1-score, serta jumlah sampel (Support) untuk setiap kelas. Mari kita bahas lebih mendalam tentang hasil evaluasi ini:

1. **Anger:** Model memiliki tingkat presisi (precision) sebesar 0.65, yang mengindikasikan bahwa dari semua teks yang diklasifikasikan sebagai "Anger," sekitar 65% di antaranya adalah benar-benar "Anger." Recall sebesar 0.85 menunjukkan bahwa model mampu mengidentifikasi sebagian besar teks yang sebenarnya adalah "Anger," yaitu sekitar 85%. F1-score yang seimbang sebesar 0.74 menggambarkan harmonisasi antara precision dan recall untuk kelas ini.
2. **Fear:** Untuk kelas "Fear," model memiliki tingkat presisi yang sangat baik (0.84), yang mengindikasikan bahwa ketika model mengklasifikasikan teks sebagai "Fear," sekitar 84% dari teks tersebut benar-benar mencerminkan emosi tersebut. Recall sebesar 0.72 menunjukkan bahwa model mampu mengidentifikasi sebagian besar teks yang sebenarnya adalah "Fear." Ini menunjukkan kemampuan model untuk mengenali teks-teks yang mengandung emosi "Fear" dengan sangat baik. F1-score yang tinggi (0.77) mencerminkan keseimbangan yang baik antara precision dan recall, menunjukkan bahwa model memiliki performa yang sangat baik untuk kelas ini.
3. **Joy:** Dalam mengklasifikasikan teks sebagai "Joy," model memiliki tingkat presisi yang baik (0.74), yang berarti ketika model mengatakan suatu teks adalah "Joy," sekitar 74% dari teks tersebut benar-benar mengandung emosi "Joy." Recall sebesar 0.67 menunjukkan bahwa model berhasil mengidentifikasi sebagian besar teks yang sebenarnya adalah "Joy." F1-score yang seimbang (0.70) mencerminkan kemampuan model untuk mengenali teks dengan emosi positif ini dengan baik.
4. **Love:** Model juga memiliki tingkat presisi yang baik (0.76) dalam mengklasifikasikan teks sebagai "Love." Recall yang memadai (0.72) menunjukkan bahwa model mampu mengidentifikasi sebagian besar teks yang mengandung emosi "Love." Ini menghasilkan F1-score yang solid (0.74), yang menunjukkan kemampuan model dalam mengidentifikasi teks dengan konteks cinta atau kasih sayang.

5. **Sadness:** Kelas "Sadness" memiliki tingkat presisi yang cukup baik (0.62) dan recall yang layak (0.58). Presisi yang cukup tinggi menunjukkan bahwa ketika model mengklasifikasikan teks sebagai "Sadness," sebagian besar teks tersebut memang mencerminkan emosi sedih. Meskipun recall lebih rendah, F1-score yang mencerminkan keseimbangan antara precision dan recall (0.60) menunjukkan bahwa model memiliki kinerja yang cukup baik dalam mengklasifikasikan teks yang mencerminkan emosi sedih.
6. **Rata-Rata (Macro Avg):** Rata-rata dari seluruh metrik, yaitu precision, recall, dan F1-score, adalah sekitar 0.72. Ini menunjukkan bahwa model memiliki kinerja yang konsisten dan seimbang di seluruh kelas emosi. Support menunjukkan jumlah sampel yang digunakan dalam evaluasi model untuk masing-masing kelas.

Secara keseluruhan, hasil evaluasi ini mengindikasikan bahwa model klasifikasi telah berhasil dalam mengklasifikasikan teks-teks ke dalam kelas emosi yang berbeda dengan baik dan seimbang. Akurasi model sebesar 0.70 juga memvalidasi kemampuan model dalam mengenali emosi dalam teks.

### III. PENUTUP

#### A. Kesimpulan

Hasil pengujian pada model yang telah dibuat memperlihatkan bahwa BiGRU dapat memberikan hasil yang baik dengan akurasi sebesar 0.70, precision rata-rata 0.72, recall rata-rata 0.71, dan support yang bervariasi tergantung pada kelas emosi.

### DAFTAR PUSTAKA

- [1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955. (*references*)
- [2] R. Hans, "Proses Pengolahan Data Teks dengan Analysis Sentiment, Kenali Tahapannya Yuk," *dqlab.id*. <https://dqlab.id/proses-pengolahan-data-teks-dengan-analysis-sentiment> (accessed Sep. 04, 2023).
- [3] U. B. Komaragiri, "Sentiment Analysis with NLP," *Analytics Vidhya*, Sep. 26, 2022. <https://www.analyticsvidhya.com/blog/2022/09/sentiment-analysis-with-nlp/>
- [4] "Apa itu Analisis Sentimen? - Penjelasan tentang Analisis Sentimen - AWS," *Amazon Web Services, Inc.* <https://aws.amazon.com/id/what-is/sentiment-analysis/>
- [5] S. Dixon, "Twitter: number of users worldwide 2019-2020 | Statista," *Statista*, Dec. 14, 2022. <https://www.statista.com/statistics/303681/twitter-users-worldwide/>
- [6] Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. <https://arxiv.org/abs/1412.3555/>
- [7] Graves, A., Wayne, G., & Danihelka, I. (2014). Neural Turing Machines. <https://arxiv.org/abs/1410.5401/>
- [8] Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3-4), 169-200. <https://www.paulekman.com/wp-content/uploads/2013/07/An-Argument-For-Basic-Emotions.pdf>
- [9] Bojanowski, P., Grave, E., Joulin, A., Mikolov, T. (2017). Enriching Word Vectors with Subword Information. <https://arxiv.org/abs/1607.04606/>
- [10] Boyd, D., & Ellison, N. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210-230. <https://onlinelibrary.wiley.com/doi/full/10.1111/j.1083-6101.2007.00393.x/>
- [11] "What is Case Folding | IGI Global," *www.igi-global.com*. <https://www.igi-global.com/dictionary/case-folding/36602#:~:text=The%20process%20of%20converting%20all> (accessed Sep. 07, 2023).
- [12] A. Pai, "What is Tokenization | Tokenization In NLP," *Analytics Vidhya*, May 25, 2020. <https://www.analyticsvidhya.com/blog/2020/05/what-is-tokenization-nlp/>
- [13] Caner, "Padding for NLP," *Medium*, Apr. 03, 2020. <https://medium.com/@canerkilinc/padding-for-nlp-7dd8598c916a>
- [14] I. Zulfa and E. Winarko, "Sentimen Analisis Tweet Berbahasa Indonesia Dengan Deep Belief Network," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 11, no. 2, pp. 187–198, Jul. 2017, Accessed: Sep. 07, 2023. [Online]. Available: <https://jurnal.ugm.ac.id/ijccs/article/view/24716/16691>
- [15] N. D. Putranti and E. Winarko, "Analisis Sentimen Twitter untuk Teks Berbahasa Indonesia dengan Maximum Entropy dan Support Vector Machine," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 8, no. 1, pp. 91–100, Jan. 2014, Accessed: Sep. 07, 2023. [Online]. Available: <https://jurnal.ugm.ac.id/ijccs/article/view/3499/3025>
- [16] "Papers with Code - BiGRU Explained," *paperswithcode.com*. <https://paperswithcode.com/method/bigr>
- [17] S. Kohli, "Understanding a Classification Report For Your Machine Learning Model," *Medium*, Nov. 18, 2019. <https://medium.com/@kohlishivam5522/understanding-a-classification-report-for-your-machine-learning-model-88815e2ce397>
- [18] A. Kharwal, "Classification Report in Machine Learning," *Data Science | Machine Learning | Python | C++ | Coding | Programming | JavaScript*, Jul. 07, 2021. <https://thecleverprogrammer.com/2021/07/07/classification-report-in-machine-learning/>

### ANALISIS TAMBAHAN

kamus singkatan dan word embedding fasttext diperoleh dari laman <https://github.com/meisaputri21/Indonesian-Twitter-Emotion-Dataset>