

Analisis Sentimen dan Emosi Masyarakat terhadap Prestasi Timnas di AFC U23 Asian Cup Qatar 2024™ pada Sosial Media YouTube Menggunakan BERT dan Algoritma Sequence Network GRU

Ariq Pradipa Santoso 2006527052, Fadel Akbar Putra 2006577492, Kemas Rafly Omar Thoriq 2006577422
Dept. of Electrical Engineering, Computer Engineering Study Program
Faculty of Engineering, Universitas Indonesia
Kampus Baru UI, Depok 16424

Abstract—Penelitian ini menganalisis sentimen dan emosi masyarakat terhadap prestasi Tim Nasional Indonesia di AFC U23 Asian Cup Qatar 2024™ melalui komentar YouTube menggunakan metode Natural Language Processing (NLP). Data dikumpulkan menggunakan library googleapiclient dan diproses melalui tahap preprocessing yang meliputi labelling, cleaning, tokenizing, dan word embedding. Kami membandingkan performa dua algoritma populer, yaitu BERT (Bidirectional Encoder Representations from Transformers) dan GRU (Gated Recurrent Unit), masing-masing dengan dua optimizer, ADAM dan SGD. Evaluasi performa dilakukan menggunakan metrik ROC-AUC dan F1-score. Hasil eksperimen menunjukkan bahwa model Transformer BERT memiliki akurasi tertinggi (0.726) dan nilai ROC-AUC tertinggi (0.842), mengungguli model GRU. Namun, model GRU dengan optimizer SGD memiliki Hamming Loss terendah (0.213), menunjukkan lebih sedikit kesalahan klasifikasi. Penelitian ini menekankan pentingnya pemilihan model dan tuning hyperparameter untuk meningkatkan performa analisis sentimen dan emosi dalam konteks komentar YouTube.

Index Terms—Analisis Sentimen, Natural Language Processing (NLP), BERT, GRU, ADAM, SGD

I. PENDAHULUAN

Prestasi Tim Nasional (Timnas) dalam ajang AFC U23 Asian Cup Qatar 2024™ menjadi sorotan utama masyarakat Indonesia. Masyarakat secara aktif mengekspresikan perasaan dan opini mereka melalui media sosial, khususnya YouTube. Komentar-komentar yang muncul mencerminkan sentimen dan emosi masyarakat, mulai dari pujian hingga kritik, serta harapan dan kekecewaan. Analisis terhadap sentimen dan emosi ini penting untuk memahami reaksi publik secara mendalam dan menyeluruh.

Analisis sentimen dan emosi adalah cabang dari pemrosesan bahasa alami (Natural Language Processing, NLP) yang bertujuan untuk mengidentifikasi dan mengkategorikan opini yang diekspresikan dalam teks. Dalam konteks ini, analisis dilakukan terhadap komentar-komentar YouTube yang berkaitan dengan performa Timnas di AFC U23 Asian Cup Qatar 2024™. Algoritma yang digunakan

adalah Sequence Network BILSTM (Bidirectional Long Short-Term Memory) dan GRU (Gated Recurrent Unit).

Penelitian ini akan membahas dan membandingkan performa dua algoritma NLP yang populer untuk analisis sentimen dan emosi: BERT (Bidirectional Encoder Representations from Transformers) dan GRU (Gated Recurrent Unit).

BERT adalah model Transformer yang dilatih dengan menggunakan sejumlah besar data teks dan mampu memahami konteks kata dalam sebuah kalimat. Kemampuan ini memungkinkan BERT untuk melakukan berbagai tugas NLP dengan tingkat akurasi yang tinggi, termasuk analisis sentimen dan emosi.

GRU adalah jenis algoritma recurrent neural network yang terkenal dengan kemampuannya untuk memproses data sekuensial. GRU memiliki arsitektur yang lebih sederhana dan efisien dibandingkan LSTM, karena menggunakan lebih sedikit parameter, yang dapat membuatnya lebih cepat dalam proses training dan inferensi.

Penelitian ini akan membahas dan membandingkan performa kedua algoritma ini dalam menganalisis sentimen dan emosi masyarakat terhadap prestasi Timnas di AFC U23 Asian Cup Qatar 2024™.

A. Permasalahan utama yang akan dibahas dalam penelitian ini adalah:

1. Bagaimana performa BERT dan algoritma Sequence Network GRU dalam menganalisis sentimen dan emosi dari komentar YouTube terkait prestasi Timnas di AFC U23 Asian Cup Qatar 2024™?
2. Apa saja faktor-faktor yang mempengaruhi keakuratan dan efektivitas kedua algoritma tersebut dalam analisis sentimen dan emosi?
3. Apakah ada pendekatan atau modifikasi baru yang dapat meningkatkan performa analisis sentimen dan emosi dalam konteks ini?

B. Penelitian ini bertujuan untuk:

1. Mengkomparasi performa BERT dan algoritma Sequence Network GRU dalam menganalisis sentimen dan emosi dari komentar YouTube terkait prestasi Timnas di AFC U23 Asian Cup Qatar 2024™.

2. Mengevaluasi keakuratan dan efektivitas kedua algoritma tersebut dalam berbagai kasus sentimen dan emosi.

C. Penelitian ini memiliki beberapa batasan sebagai berikut:

1. Data yang digunakan terbatas pada komentar YouTube yang terkait dengan prestasi Timnas di AFC U23 Asian Cup Qatar 2024™.
2. Analisis hanya difokuskan pada sentimen (positif, negatif, netral) dan beberapa emosi dasar (misalnya, kebahagiaan, kekecewaan).
3. Algoritma yang dikomparasi terbatas pada BERT dan GRU, dengan kemungkinan modifikasi minor untuk peningkatan performa.

II. STUDI LITERATUR

Ilmania et al. (2023) dalam karya mereka "Aspect Detection and Sentiment Classification Using Deep Neural Network for Indonesian Aspect-Based Sentiment Analysis" menekankan pentingnya analisis sentimen berbasis aspek untuk menangani kalimat yang mengandung lebih dari satu opini. Penelitian ini dilakukan pada dataset berbahasa Indonesia dengan dua langkah proses: deteksi aspek dan klasifikasi sentimen. Pada deteksi aspek, mereka membandingkan dua model jaringan saraf dalam-dalam dengan vektor masukan dan topologi yang berbeda: vektor embedding kata yang diproses menggunakan Gated Recurrent Unit (GRU) dan vektor bag-of-words yang diproses menggunakan lapisan fully-connected. Untuk klasifikasi sentimen, mereka membandingkan dua pendekatan jaringan saraf dalam-dalam: pendekatan pertama menggunakan embedding kata, leksikon sentimen, dan POS tags sebagai vektor masukan dengan topologi berbasis bi-GRU, sementara pendekatan kedua menggunakan matriks aspek untuk mereskalakan vektor embedding kata sebagai vektor masukan dengan topologi berbasis Convolutional Neural Network (CNN). Hasil penelitian menunjukkan bahwa model mereka memberikan kinerja terbaik pada 4 dari 7 aspek dibandingkan dengan kerangka kerja baseline yang menggunakan model berbeda untuk setiap aspek. Penelitian ini relevan dalam konteks analisis sentimen dan emosi terhadap prestasi Timnas di AFC U23 Asian Cup Qatar 2024™ pada media sosial YouTube, karena menyediakan kerangka kerja yang kuat untuk menangani data berbahasa Indonesia dan menawarkan pendekatan berbasis aspek yang dapat digunakan untuk analisis sentimen yang lebih mendalam dan terfokus [4].

Pradhan (2023) dalam artikel "Extracting Sentiments from YouTube Comments" menyoroti pentingnya analisis sentimen pada komentar YouTube. YouTube sebagai platform media sosial yang paling banyak digunakan, memungkinkan pengguna untuk mengekspresikan opini mereka melalui komentar. Analisis sentimen pada komentar YouTube dapat memberikan wawasan tentang opini publik atau perilaku terkait video tertentu. Penelitian ini menggunakan teknik Pemrosesan Bahasa Alami (NLP) untuk mengklasifikasikan sentimen komentar menjadi positif, negatif, atau netral. Hasil

analisis ini dapat membantu pengguna mendapatkan laporan tentang video YouTube mereka. Penelitian ini penting dalam konteks analisis sentimen dan emosi terhadap prestasi Timnas di AFC U23 Asian Cup Qatar 2024™ karena menawarkan metode untuk memahami bagaimana masyarakat mengekspresikan opini mereka melalui komentar di YouTube. Dengan menggunakan teknik NLP, kita dapat menganalisis sentimen secara otomatis dan efisien [5].

Muhammad et al. (2023) dalam karya mereka "Sentiment Analysis of Positive and Negative of YouTube Comments Using Naïve Bayes – Support Vector Machine (NBSVM) Classifier" mengeksplorasi penggunaan kombinasi metode Naïve Bayes dan Support Vector Machine (SVM) untuk klasifikasi sentimen komentar YouTube. Analisis sentimen pada komentar video YouTube merupakan proses untuk memahami, mengekstrak, dan memproses data tekstual secara otomatis guna mendapatkan informasi sentimen yang terkandung dalam satu kalimat komentar video YouTube. Kombinasi Naïve Bayes dan SVM menunjukkan kinerja yang lebih baik dengan nilai presisi 91%, recall 83%, dan f1-score 87%. Studi ini relevan dalam konteks analisis sentimen dan emosi terhadap prestasi Timnas di AFC U23 Asian Cup Qatar 2024™ karena menunjukkan bahwa kombinasi metode klasifikasi yang berbeda dapat meningkatkan akurasi dan kinerja dalam analisis sentimen. Dengan memanfaatkan kombinasi Naïve Bayes dan SVM, analisis sentimen terhadap komentar YouTube dapat dilakukan dengan lebih akurat dan memberikan hasil yang lebih kuat [6].

Ketiga literatur ini memberikan landasan yang kuat untuk analisis sentimen dan emosi masyarakat terhadap prestasi Timnas di AFC U23 Asian Cup Qatar 2024™ pada media sosial YouTube. Literatur pertama memberikan pendekatan berbasis aspek yang mendalam dan sesuai untuk data berbahasa Indonesia, sementara literatur kedua dan ketiga menawarkan metode yang efisien dan akurat untuk analisis sentimen pada komentar YouTube. Dengan menggabungkan pendekatan dan metode dari ketiga literatur ini, laporan dapat memberikan analisis yang komprehensif dan mendalam tentang sentimen dan emosi masyarakat terhadap prestasi Timnas di AFC U23 Asian Cup Qatar 2024™.

III. DASAR TEORI

A. Sentimen Analisis

Analisis sentimen adalah subbidang dari pemrosesan bahasa alami (Natural Language Processing, NLP) yang berfokus pada ekstraksi informasi subjektif dan sentimen dari data teks. Tujuan utama analisis sentimen adalah untuk menentukan nada emosional atau polaritas yang terkait dengan teks tertentu, yang dapat berupa sentimen positif, negatif, atau netral. Teknik ini digunakan untuk mengidentifikasi opini, sikap, atau perasaan penulis terhadap suatu subjek yang diekspresikan melalui teks.

Berbagai teknik klasifikasi digunakan dalam analisis sentimen, antara lain:

1. Metode Berbasis Leksikon: Menggunakan kamus sentimen yang berisi kata-kata dan frasa dengan label

polaritas (positif, negatif, atau netral). Analisis dilakukan dengan mencocokkan kata-kata dalam teks dengan kamus tersebut.

2. Algoritma Pembelajaran Mesin: Melibatkan pelatihan model pembelajaran mesin menggunakan dataset yang telah diberi label sentimen. Algoritma seperti Naive Bayes, Support Vector Machine (SVM), dan Random Forest sering digunakan untuk klasifikasi sentimen.
3. Model Pembelajaran Mendalam (Deep Learning): Model deep learning seperti Convolutional Neural Networks (CNN) dan Recurrent Neural Networks (RNN), termasuk Long Short-Term Memory (LSTM) dan Gated Recurrent Unit (GRU), terbukti efektif dalam menangani data teks yang kompleks dan bervariasi.
4. Pendekatan Hibrida: Menggabungkan metode berbasis leksikon dan algoritma pembelajaran mesin untuk meningkatkan akurasi dan efektivitas analisis sentimen.

Analisis sentimen memiliki berbagai aplikasi di berbagai industri, seperti:

1. Pemantauan Reputasi Merek: Perusahaan memantau bagaimana merek mereka dibicarakan di media sosial dan platform online lainnya, membantu mereka memahami persepsi publik terhadap produk atau layanan mereka.
2. Penilaian Kepuasan Pelanggan: Digunakan untuk menilai kepuasan pelanggan dari ulasan dan feedback yang diberikan, membantu perusahaan meningkatkan kualitas produk dan layanan.
3. Prediksi Tren Pasar: Dengan menganalisis sentimen publik terhadap topik tertentu, perusahaan dapat memprediksi tren pasar dan membuat keputusan bisnis yang lebih informasional.
4. Deteksi Propaganda Media Sosial: Digunakan untuk mendeteksi propaganda atau informasi yang salah di media sosial, penting untuk menjaga integritas informasi dan mencegah penyebaran berita palsu.
5. Analisis Wacana Politik: Membantu dalam memahami opini publik terhadap kandidat atau kebijakan tertentu, penting dalam strategi kampanye dan pengambilan keputusan politik.

Namun, analisis sentimen menghadapi beberapa tantangan:

- Pemrosesan Data Teks yang Tidak Terstruktur: Teks dari media sosial dan platform online sering kali tidak terstruktur dan berisik, membuat analisis menjadi sulit.
- Ambiguitas: Kata-kata atau frasa dalam teks bisa memiliki makna yang berbeda tergantung pada konteksnya, sehingga sulit untuk menentukan sentimen yang sebenarnya.
- Sarkasme: Identifikasi sarkasme merupakan tantangan besar karena sarkasme sering kali bertentangan dengan makna literal dari kata-kata yang digunakan.

- Bahasa Sehari-hari: Penggunaan slang, idiom, dan bahasa informal dalam teks membuat analisis sentimen menjadi lebih kompleks.

B. BERT (Bidirectional Encoder Representations from Transformers)

BERT (Bidirectional Encoder Representations from Transformers) adalah model bahasa yang revolusioner yang dikembangkan oleh Google AI. Model ini dilatih dengan menggunakan sejumlah besar data teks dan mampu memahami konteks kata dalam sebuah kalimat. Kemampuan ini memungkinkan BERT untuk melakukan berbagai tugas pemrosesan bahasa alami (NLP) dengan tingkat akurasi yang tinggi.

BERT Transformer menggunakan arsitektur jaringan saraf yang disebut "Transformer" yang dikembangkan oleh Google Research pada tahun 2017. Transformer memungkinkan BERT untuk memproses kata-kata dalam sebuah kalimat secara bersamaan, sehingga model dapat memahami konteks dan hubungan antar kata.

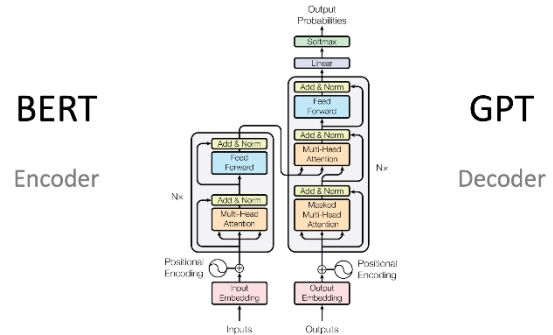


Fig. 2. Diagram GRU yang sudah disederhanakan [7]

Gambar di atas menunjukkan arsitektur BERT Transformer. Model ini terdiri dari dua encoder utama:

- Encoder Encoder: Encoder pertama memproses kata-kata dalam sebuah kalimat dari kiri ke kanan.
- Encoder Decoder: Encoder kedua memproses kata-kata dalam sebuah kalimat dari kanan ke kiri.

Kedua encoder ini menghasilkan representasi vektor untuk setiap kata dalam kalimat. Representasi vektor ini kemudian digunakan untuk melakukan berbagai tugas NLP. BERT Transformer memiliki beberapa manfaat utama, antara lain:

- Lebih akurat: BERT Transformer lebih akurat dalam melakukan berbagai tugas NLP dibandingkan model sebelumnya.
- Lebih serbaguna: BERT Transformer dapat digunakan untuk berbagai tugas NLP, seperti menjawab pertanyaan, ringkasan teks, terjemahan mesin, dan klasifikasi teks.
- Lebih mudah digunakan: BERT Transformer lebih mudah digunakan dibandingkan model sebelumnya, karena model ini dilatih dengan menggunakan sejumlah besar data teks dan tidak memerlukan banyak penyesuaian.

C. GRU (Gated Recurrent Circuit)

GRU, atau Gated Recurrent Unit, adalah arsitektur jaringan

saraf tiruan (JST) berulang yang dirancang untuk memproses urutan data, seperti teks. GRU menawarkan keseimbangan antara kompleksitas dan performa, membuatnya menjadi pilihan populer untuk berbagai tugas pemrosesan bahasa alami (NLP), termasuk analisis sentimen.

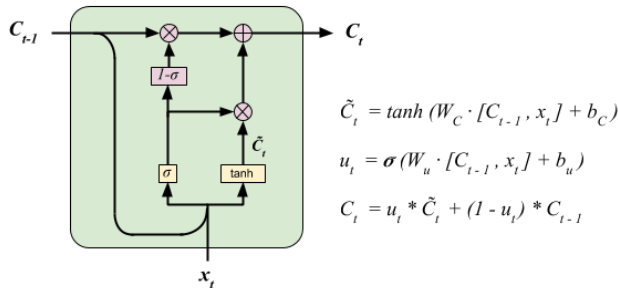


Fig. 3. Diagram GRU yang sudah disederhanakan [3]

GRU terdiri dari dua gerbang utama:

1. Gerbang pembaruan (update gate): Mengontrol seberapa banyak informasi dari keadaan tersembunyi sebelumnya ($h_{(t-1)}$) yang akan diintegrasikan ke dalam keadaan tersembunyi saat ini (h_t).
2. Gerbang reset (reset gate): Mengontrol seberapa banyak informasi dari keadaan tersembunyi sebelumnya yang akan dihapus.

Kedua gerbang ini bekerja sama untuk memungkinkan GRU belajar dari informasi yang relevan dari urutan data, sekaligus melupakan informasi yang tidak relevan.

IV. EKSPERIMEN

A. Metode Eksperimen

1) Data

Data dalam proyek ini dikumpulkan dari video-video YouTube terkait sorotan Timnas U23 di AFC U23 Asian Cup Qatar 2024™. Hasil yang diambil adalah seluruh komentar dari video-video terkait. Total data yang berhasil dikumpulkan adalah sekitar 3000 data.

2) Preprocessing

Preprocessing data diperlukan untuk memastikan bahwa data yang akan diproses merupakan data yang bersih dan mudah dipahami. Proses ini melalui beberapa tahap, dimulai dengan labeling untuk menentukan apakah sebuah komentar termasuk positif, negatif, atau netral. Selanjutnya adalah tahap cleaning.

Pada tahap ini, dilakukan pembersihan data dari elemen-elemen yang tidak diperlukan atau dapat mengganggu proses analisis. Cleaning mencakup penghapusan emoji, tanda baca, karakter khusus, dan angka yang mungkin tidak relevan. Selain itu, semua teks diubah menjadi huruf kecil untuk konsistensi, stopwords dihapus untuk memfokuskan analisis pada kata-kata yang lebih bermakna, dan dilakukan stemming untuk mengubah kata ke bentuk dasarnya.

Dilanjutkan dengan proses Tokenizing, proses ini memecah teks menjadi unit-unit yang lebih kecil, biasanya kata atau frasa, yang disebut token. Tokenizing membantu dalam mempersiapkan data teks

untuk analisis lebih lanjut dengan algoritma machine learning dan deep learning, karena algoritma tersebut biasanya membutuhkan input dalam bentuk token.

Terakhir adalah Word Embedding, pada tahap ini, token yang telah dihasilkan dari proses tokenizing diubah menjadi representasi numerik yang dapat digunakan oleh model machine learning. Word embedding, seperti Word2Vec atau GloVe, adalah teknik yang digunakan untuk merepresentasikan kata-kata dalam vektor angka yang mengandung informasi tentang makna dan hubungan antar kata. Ini memungkinkan model untuk memahami konteks dan makna kata dalam teks.

3) Skenario

Penelitian ini menggunakan dua jenis algoritma sequence network, yaitu BERT dan GRU (Gated Recurrent Unit), dengan dua optimizer yang berbeda, yaitu ADAM dan SGD. Berikut adalah skenario eksperimen yang akan dijalankan:

a) GRU dengan Optimizer ADAM

Model GRU akan dilatih menggunakan optimizer ADAM untuk memanfaatkan kecepatan dan stabilitas dalam pembaruan bobot, mirip dengan eksperimen pada BILSTM.

b) GRU dengan Optimizer SGD

Model GRU akan dilatih menggunakan optimizer SGD untuk melihat bagaimana kinerja algoritma ini dengan pendekatan optimizer yang berbeda.

c) Transformer BERT

Model Transformer BERT akan digunakan untuk perbandingan dengan model GRU dan untuk menyelidiki kemampuannya dalam memahami konteks kalimat secara lebih baik. BERT dikenal dengan performa yang luar biasa dalam berbagai tugas NLP, dan menarik untuk melihat bagaimana performa model ini dalam konteks penelitian ini.

Setiap skenario akan dievaluasi menggunakan dua metrik utama, yaitu ROC-AUC dan F1-score untuk melakukan dan menentukan evaluasi hasil setiap skenarionya.

4) Hyperparameter Tuning

Penelitian ini kita melakukan hyperparameter tuning dimana hyperparamter tuning ini akan membuat dan menghasilkan hasil yang kita inginkan.

Pada model GRU dengan optimizer Adam, nilai learning rate diubah dari 0.01 menjadi 0.001. Hal ini diharapkan dapat meningkatkan akurasi dan stabilitas model, serta mengurangi overfitting dan mempercepat konvergensi model.

Pada model SGD, nilai learning rate diubah dari 0.1 menjadi 0.01. Penyesuaian ini diharapkan dapat meningkatkan akurasi model, mengurangi overfitting, dan mempercepat konvergensi model.

B. Alat Eksperimen

Dalam penelitian ini, berbagai alat dan library digunakan untuk mendukung proses scrapping data, serta pembangunan dan evaluasi model machine learning.

1) Scrapping

Untuk mengumpulkan data komentar dari YouTube, digunakan library `googleapiclient`. Library ini memungkinkan akses ke API YouTube, yang digunakan untuk mengambil komentar secara programatis dan efisien. Penggunaan `googleapiclient` digabungkan menggunakan script yang dibuat di python

2) Machine Learning

Untuk membangun dan melatih model machine learning, digunakan beberapa library utama

- TensorFlow: Library open-source yang digunakan untuk membangun dan melatih model deep learning. TensorFlow menyediakan fleksibilitas tinggi dan performa yang baik untuk berbagai jenis model neural network.
- Keras: Sebuah API tingkat tinggi untuk TensorFlow yang memudahkan definisi dan pelatihan model neural network dengan sintaks yang sederhana dan intuitif.
- Gensim: Library yang digunakan untuk pemrosesan teks dan pemodelan topik, termasuk word embedding. Gensim menyediakan berbagai alat untuk membuat representasi vektor dari teks yang digunakan dalam analisis sentimen.

Dengan memanfaatkan `googleapiclient` untuk scrapping dan library seperti TensorFlow, Keras, dan Gensim untuk machine learning, penelitian ini dapat melakukan proses pengumpulan, pengolahan, dan analisis data secara efisien dan akurat.

C. Hasil Eksperimen

Table di bawah menunjukkan hasil eksperimen untuk tugas pemrosesan bahasa alami (NLP). Tabel ini membandingkan akurasi dan kerugian hamming dari tiga model yang berbeda: GRU-SGD, GRU-Adam, dan Transformer.

TABLE I

TABLE HASIL AKURASI, ROC-AUC DAN HAMMING LOSS

Model	Metrics		
	Accuracy	ROC-AUC	Hamming Loss
GRU-SGD	0.684	0.803	0.213
GRU-Adam	0.648	0.792	0.222
Transformer	0.726	0.842	0.274

Berikut dibawah ini adalah grafik daripada ROC GRU-SGD yang didapatkan. Grafik ini menunjukkan hasil ROC curve class untuk setiap sentimen para pemain termasuk positif, negatif dan netral.

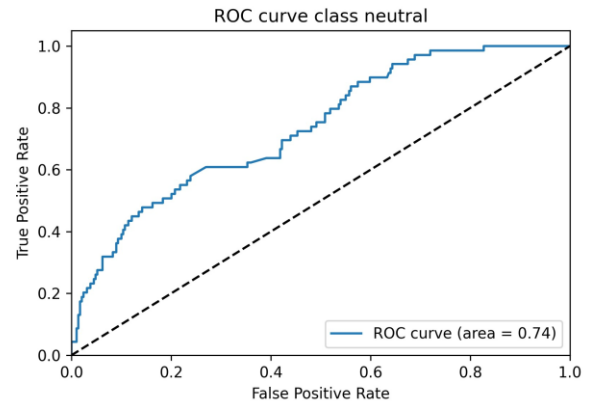


Fig. 4. Grafik ROC GRU-SGD, kelas netral

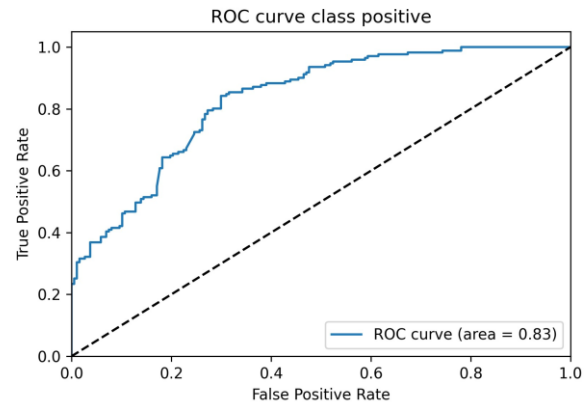


Fig. 5. Grafik ROC GRU- SGD, kelas positif

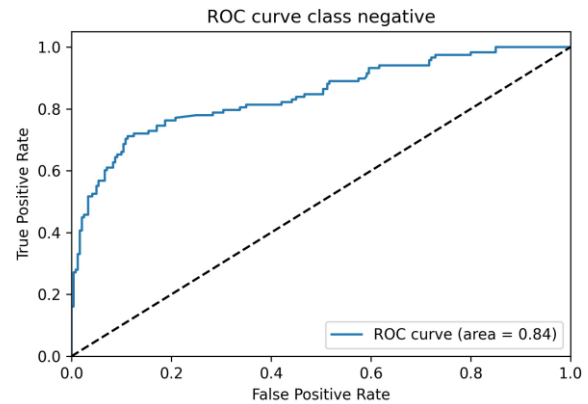


Fig. 6. Grafik ROC GRU- SGD, kelas negatif

Berikut dibawah ini adalah grafik daripada ROC GRU-Adam

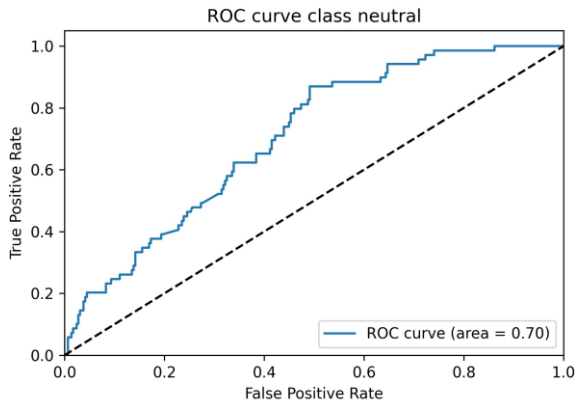


Fig. 7. Grafik ROC GRU-Adam, kelas netral

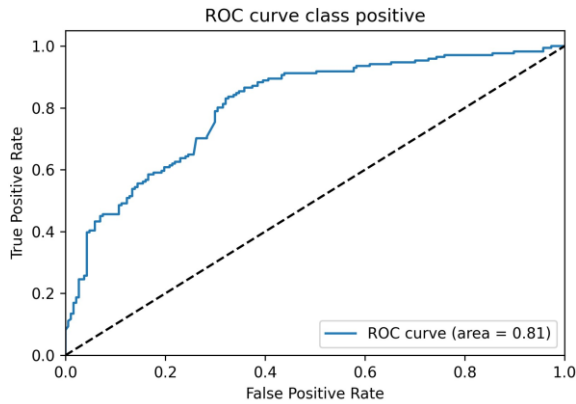


Fig. 8. Grafik ROC GRU-Adam, kelas positif

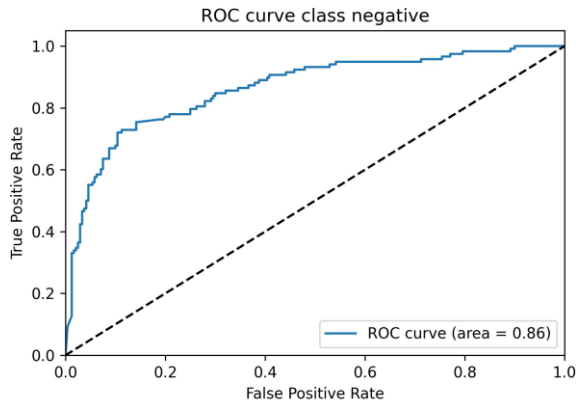


Fig. 8. Grafik ROC GRU-Adam, kelas negatif

Dibawah ini merupakan hasil sentimen masyarakat terhadap para pemain. Hasil distribusi sentimen masyarakat terhadap para pemain menunjukkan bahwa kelas netral memiliki frekuensi yang paling sedikit, sehingga nilai ROC-nya juga paling kecil.

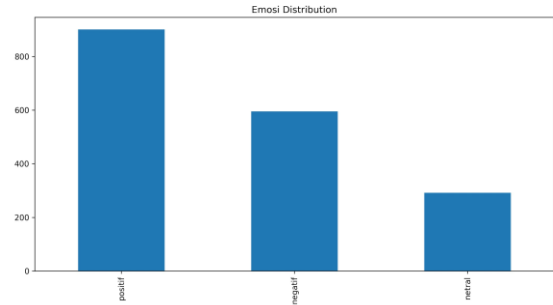


Fig. 10. Distribusi Hasil Sentimen

V. ANALISIS EKSPERIMEN

Tabel hasil akurasi, ROC-AUC, dan Hamming Loss menunjukkan perbandingan kinerja tiga model berbeda, yaitu GRU dengan optimizer SGD, GRU dengan optimizer Adam, dan Transformer. Metrik yang digunakan untuk evaluasi adalah akurasi, ROC-AUC, dan Hamming Loss, yang memberikan gambaran menyeluruh tentang performa masing-masing model dalam tugas pemrosesan bahasa alami (NLP) untuk analisis sentimen dan emosi masyarakat terhadap prestasi Timnas di AFC U23 Asian Cup Qatar 2024™ pada platform YouTube.

Dari hasil eksperimen, model Transformer menunjukkan performa terbaik dengan akurasi tertinggi sebesar 0.726 dan nilai ROC-AUC sebesar 0.842, mengindikasikan kemampuan superior dalam memahami konteks kalimat dan melakukan klasifikasi yang akurat dibandingkan dengan model GRU. Namun, Transformer memiliki nilai Hamming Loss tertinggi sebesar 0.274, menunjukkan bahwa meskipun mampu memberikan prediksi yang akurat, model ini juga melakukan beberapa kesalahan dalam klasifikasi label.

Model GRU dengan optimizer SGD memiliki performa yang baik dengan akurasi 0.684 dan nilai ROC-AUC 0.803, serta Hamming Loss 0.213, yang merupakan nilai Hamming Loss terendah di antara ketiga model. Hal ini menunjukkan bahwa meskipun GRU-SGD tidak seakurat Transformer dalam metrik akurasi dan ROC-AUC, model ini lebih baik dalam meminimalkan kesalahan klasifikasi.

Model GRU dengan optimizer Adam memiliki performa yang paling rendah dengan akurasi 0.648 dan nilai ROC-AUC 0.792. Hamming Loss model ini berada di tengah-tengah dengan nilai 0.222. Meskipun optimisasi dengan Adam dikenal karena stabilitas dan kecepatan konvergensi, dalam eksperimen ini, GRU-Adam tidak mampu mengungguli GRU-SGD maupun Transformer dalam hal akurasi dan kemampuan membedakan antara kelas positif dan negatif.

VI. KESIMPULAN

Kesimpulan dari penelitian ini menunjukkan bahwa model Transformer memiliki kinerja terbaik dalam tugas analisis sentimen dan emosi masyarakat terhadap prestasi Timnas di AFC U23 Asian Cup Qatar 2024™ pada platform YouTube, dengan akurasi tertinggi (0.726) dan nilai ROC-AUC tertinggi (0.842). Hal ini mengindikasikan bahwa Transformer lebih unggul dalam memahami konteks kalimat dan memberikan klasifikasi yang akurat dibandingkan dengan model GRU. Meskipun demikian, model GRU dengan optimizer SGD

memiliki nilai Hamming Loss terendah (0.213), menunjukkan bahwa model ini membuat lebih sedikit kesalahan klasifikasi. Model GRU dengan optimizer Adam, meskipun dioptimalkan untuk stabilitas dan kecepatan konvergensi, menunjukkan kinerja yang paling rendah di antara ketiga model dalam hal akurasi (0.648) dan ROC-AUC (0.792). Penelitian ini juga menekankan pentingnya hyperparameter tuning, seperti penyesuaian learning rate, dalam meningkatkan kinerja model. Secara keseluruhan, pilihan model harus disesuaikan dengan prioritas metrik evaluasi, di mana Transformer cocok untuk aplikasi yang memerlukan akurasi tinggi, sementara GRU-SGD dapat dipertimbangkan untuk meminimalkan kesalahan klasifikasi.

REFERENCES

- [1] Puteri, Dian. (2023). Implementasi Long Short Term Memory (LSTM) dan Bidirectional Long Short Term Memory (BiLSTM) Dalam Prediksi Harga Saham Syariah. *Euler : Jurnal Ilmiah Matematika, Sains dan Teknologi*. 11. 35-43. 10.34312/euler.v11i1.19791.
- [2] Zaytar, Mohamed Akram, and Chaker El Amrani. "Sequence to sequence weather forecasting with long short-term memory recurrent neural networks." *International Journal of Computer Applications* 143.11 (2016): 7-11.
- [3] Prijono, Benny. (2018). "Pengenalan Long Short Term Memory (LSTM) dan Gated Recurrent Unit (GRU) – RNN Bagian 2." *IndoML.com*, 13 April. Diakses pada 28 Mei 2024. [<https://indoml.com/2018/04/13/pengenalan-long-short-term-memory-lstm-dan-gated-recurrent-unit-gru-rnn-bagian-2/>].
- [4] A. Ilmania, Abdurrahman, S. Cahyawijaya and A. Purwarianti, "Aspect Detection and Sentiment Classification Using Deep Neural Network for Indonesian Aspect-Based Sentiment Analysis," 2018 International Conference on Asian Language Processing (IALP), Bandung, Indonesia, 2018, pp. 62-67, doi: 10.1109/IALP.2018.8629181.
- [5] R. Pradhan, "Extracting Sentiments from YouTube Comments," 2021 Sixth International Conference on Image Information Processing (ICIIP), Shimla, India, 2021, pp. 1-4, doi: 10.1109/ICIIP53038.2021.9702561.
- [6] A. N. Muhammad, S. Bukhori and P. Pandunata, "Sentiment Analysis of Positive and Negative of YouTube Comments Using Naïve Bayes – Support Vector Machine (NBSVM) Classifier," 2019 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE), Jember, Indonesia, 2019, pp. 199-205, doi: 10.1109/ICOMITEE.2019.8920923.
- [7] Alissa, Sarah & Wald, Mike. (2023). Text Simplification Using Transformer and BERT. *Computers, Materials & Continua*. 75. 3479-3495. 10.32604/cmc.2023.033647.