

KLASIFIKASI INTENSI DENGAN METODE MODIFIED EASY DATA AUGMENTATION (EDA) DAN SEQUENCE-BASED MODELS PADA CHATBOT BAHASA INDONESIA

Vincent Revel Aldisa¹, Rizal Setya Perdana², Putra Pandu Adikara³

^{1,2,3}Universitas Brawijaya

Email: ¹vincentrevel@student.ub.ac.id, ²rizalespe@ub.ac.id, ³adikara.putra @ub.ac.id

*Penulis Korespondensi

(Naskah masuk: dd mmm yyyy, diterima untuk diterbitkan: dd mmm yyyy)

Abstrak

Helpdesk TIK Universitas Brawijaya (UB) adalah layanan *helpdesk online* yang dirancang untuk mendukung mahasiswa, dosen, dan staf UB dalam hal teknologi informasi. Layanan ini menghadapi kendala karena staf hanya merespons selama jam kerja, sementara permasalahan IT *civitas academica* UB sering terjadi di luar waktu tersebut. Sebagai solusi, diusulkan penerapan *chatbot* yang dikembangkan berdasarkan *dataset* Helpdesk TIK UB. Namun, riset sebelumnya mengalami masalah *overfitting* akibat keterbatasan *dataset* yang berkualitas, terdiri dari 600 data dengan 7 kelas intensi. Untuk mengatasi masalah ini, penelitian ini menggunakan metode Easy Data Augmentation (EDA) yang dimodifikasi, termasuk empat operasi dasar: *synonym replacement*, *random insert*, *random swap*, dan *random deletion*. Hasilnya, *dataset* diperluas menjadi 5.112 baris data. Dengan menggunakan tiga *sequence-based models* (LSTM, RNN, GRU), penelitian menunjukkan peningkatan signifikan dalam kinerja model setelah augmentasi data. Hal ini ditandai dengan penurunan *loss* pengujian dan peningkatan akurasi pada setiap model, mengurangi perbedaan kinerja antara data pelatihan dan pengujian. Penelitian ini juga mengeksplorasi *hyperparameter tuning* dengan 486 kombinasi berbeda, mengidentifikasi konfigurasi terbaik untuk klasifikasi intensi pertanyaan pada setiap model. GRU menunjukkan performa terbaik pada tahap pengujian dengan *loss* pengujian terendah yaitu 0,0195 dan akurasi 0,9960.

Kata kunci: Easy Data Augmentation (EDA), klasifikasi intensi, model berbasis urutan, LSTM, RNN, GRU

INTENT CLASSIFICATION USING MODIFIED EASY DATA AUGMENTATION (EDA) METHOD AND SEQUENCE-BASED MODELS ON INDONESIAN CHATBOTS

Abstract

The IT Helpdesk at Universitas Brawijaya (UB) is an online service designed to support students, faculty, and staff with information technology issues. The service faces challenges as staff can only respond during working hours, while IT problems often occur outside these times within the UB academic community. To address this, the implementation of a chatbot developed using the UB IT Helpdesk dataset is proposed. However, previous research encountered overfitting issues due to the limited quality of the dataset, which consisted of 600 data entries across 7 intent classes. To overcome this, the current study applies a modified Easy Data Augmentation (EDA) method, involving four basic operations: *synonym replacement*, *random insertion*, *random swap*, and *random deletion*. As a result, the dataset expanded to 5112 data entries. Using three *sequence-based models* (LSTM, RNN, GRU), the study demonstrates a significant improvement in model performance after data augmentation. This is marked by a decrease in test loss and an increase in accuracy for each model, reducing the performance gap between training and testing data. The study also explores *hyperparameter tuning* with 486 different combinations, identifying the best configuration for intent classification in each model. GRU showed the best performance at the testing stage with the lowest test loss 0.0195 and highest accuracy of 0.9960.

Keywords: Easy Data Augmentation (EDA), intention classification, sequence-based model, LSTM, RNN, GRU

1. PENDAHULUAN

Di era digital saat ini, pelayanan publik berbasis Teknologi Informasi dan Komunikasi (TIK) menjadi penting, termasuk di Universitas Brawijaya (UB)

yang mengimplementasi teknologi, seperti pada Helpdesk TIK untuk mendukung *civitas academica* dalam mengatasi permasalahan TIK. Namun, layanan ini memiliki keterbatasan karena staf hanya merespon

dalam jam kerja. Sebagai solusi, diusulkan penggunaan *chatbot* yang dapat beroperasi 24 jam. Riset sebelumnya menunjukkan masalah *overfitting* dalam pengembangan *chatbot* akibat *dataset* terbatas (Al Farisi et al., 2023). *Overfitting*, yaitu ketika model terlalu spesifik pada data pelatihan, mengurangi keakuratan pada data baru (Piratla, 2023). Hal ini yang akan menjadi fokus penelitian, terutama dengan *dataset* kecil.

Solusi untuk masalah *dataset* terbatas dan *overfitting* adalah penambahan *dataset*, baik secara manual atau menggunakan metode Data Augmentation. Data Augmentation memungkinkan perluasan *dataset* dengan cara melakukan manipulasi pada data yang ada dengan waktu yang singkat. Penelitian ini akan menggunakan teknik *modified Easy Data Augmentation* (EDA) untuk mengatasi masalah *dataset* terbatas dan *overfitting*. EDA memodifikasi teks dengan cara seperti mengganti sinonim dan penghapusan, sangat cocok untuk *dataset* QA Helpdesk TIK UB karena mempertahankan integritas semantik.

Dalam konteks penerapan Easy Data Augmentation (EDA), berbagai penelitian telah memberikan wawasan berharga. Penelitian oleh Natasya & Girsang (2022) menggunakan model LSTM dan Bi-LSTM untuk analisis sentimen berbasis aspek dalam Bahasa Indonesia, dengan memodifikasi EDA dan menggabungkannya dengan backtranslation. Dengan *dataset* terbatas, penelitian ini berhasil meningkatkan akurasi menjadi 88,9% dan F1-score 84,3% melalui pendekatan ini.

Kartika et al. (2023) mengembangkan model berbasis IndoBERT dengan data augmentasi untuk identifikasi paraphrase dalam Bahasa Indonesia. Mereka menggunakan *dataset Quora Question Pairs* yang telah diterjemahkan ke Bahasa Indonesia dan dilakukan proses augmentasi, yang meningkatkan jumlah data menjadi lebih dari satu juta pasang pertanyaan. Hasilnya, model klasifikasi ANN yang mereka kembangkan mencapai akurasi 90,1% dan skor F1 86,7%.

Gennaro et al. (2020) fokus pada klasifikasi intensi dalam *dataset Question Answering* (QA) menggunakan LSTM. *Dataset* mereka berisi ribuan pertanyaan dalam bahasa Inggris, dan penelitian ini berhasil mencapai akurasi 91,20% dengan model LSTM. Uji coba pada prototype responder mengindikasikan kemampuan model dalam merespons pertanyaan sesuai intensi, meskipun masih terdapat keterbatasan dalam jumlah knowledge yang dimiliki oleh model.

Penelitian-penelitian ini memberikan dasar yang kuat untuk penggunaan EDA dalam peningkatan kinerja model pembelajaran mesin, khususnya dalam konteks Bahasa Indonesia dan tugas-tugas berbasis teks. Berdasarkan tinjauan literatur yang telah dibahas, penelitian ini bertujuan untuk menerapkan teknik *modified Easy Data Augmentation* (EDA) dalam pengembangan *chatbot* untuk Helpdesk TIK

UB. Fokus utama adalah untuk mengatasi masalah *overfitting* yang terjadi dalam pengembangan model berbasis pembelajaran mesin dengan *dataset* terbatas pada penelitian-penelitian yang telah dilakukan sebelumnya. Penelitian ini akan menggunakan *dataset* yang diperluas melalui proses augmentasi data EDA, yang diharapkan dapat meningkatkan kinerja dan keakuratan model dalam klasifikasi intensi pertanyaan.

Lebih lanjut, penelitian ini akan menguji penggunaan *sequence-based models* (LSTM, RNN, dan GRU) dalam memproses dan mengklasifikasikan pertanyaan yang diajukan ke Helpdesk TIK UB. Model ini akan dievaluasi berdasarkan akurasi dan kemampuannya dalam mengurangi kesenjangan kinerja antara data latihan dan pengujian. Diharapkan bahwa keluaran atau hasil dari penelitian ini, yang terdiri dari *dataset* yang dihasilkan dari augmentasi data dan model klasifikasi intensi, akan menjadi komponen dari perancangan *chatbot* yang baik. *Chatbot* yang baik dapat diandalkan untuk menjawab pertanyaan, seperti pada studi kasus dari Helpdesk TIK UB. *Chatbot* yang baik juga dapat membantu staf UPT TIK UB menangani pertanyaan dari *civitas academica* UB dan juga dapat menjawab pertanyaan pengguna di luar jam kerja.

2. LANDASAN KEPUSTAKAAN

2.1 Helpdesk TIK UB

Salah satu layanan yang disediakan oleh Unit Pelaksana Teknis Teknologi Informasi dan Komunikasi (UPT TIK UB) adalah Helpdesk TIK UB. Melalui layanan ini, seluruh *civitas akademica* UB dapat mengajukan pertanyaan atau melaporkan masalah yang mereka miliki terkait layanan teknologi informasi yang ada di Universitas Brawijaya. Dengan memasukkan data identitas dan memilih topik, *civitas academica* UB dapat membuka tiket baru. Setelah itu, mereka dapat menjelaskan masalah yang dihadapi. Staf UPT TIK UB kemudian akan merespon tiket tersebut dan membantu masalah pengguna (Helpdesk TIK, n.d.).

2.2 Easy Data Augmentation (EDA)

Easy Data Augmentation (EDA) merupakan pendekatan inovatif dalam pengolahan *dataset* teks yang bertujuan untuk memperbaiki kuantitas dan kualitas data pelatihan dengan metode yang sederhana dan efisien (Wei & Zou, 2019). EDA berfokus pada empat operasi utama: *Synonym replacement* (SR), yang menggantikan kata-kata dengan sinonimnya, seperti mengubah kata "bahagia" menjadi "senang" atau "gembira"; *Random insert* (RI), yang menyisipkan kata-kata tambahan secara acak, contohnya, kata "terkadang" dapat disisipkan ke dalam kalimat "Saya suka membaca buku" menjadi "Saya terkadang suka membaca buku".

Selanjutnya ada *Random swap* (RS), yang mengubah urutan kata dalam teks, seperti menukar kata "pizza" dan "makan" dalam kalimat "Dia suka makan pizza di restoran favoritnya" menjadi "Dia makan suka pizza di restoran favoritnya"; dan *Random deletion* (RD), yang menghapus kata-kata secara acak, seperti menghilangkan kata "besar" dari kalimat "Dia sedang mempersiapkan presentasi untuk pertemuan besar besok," sehingga menjadi "Dia sedang mempersiapkan presentasi untuk pertemuan besok."

2.3 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) adalah jenis arsitektur yang dikembangkan dari RNN, cocok untuk *deep learning*, klasifikasi, dan prediksi berdasarkan data *time series* (Khumaidi et al., 2020). LSTM memiliki 4 *gate*: *forget gate* (untuk menghapus informasi tidak relevan), *input gate* (menambahkan informasi yang terpilih), *output gate* (menghasilkan informasi aktual), dan *candidate memory*. Perhitungan dari setiap *gate* yang terdapat pada LSTM dilakukan dengan Persamaan 1-6.

$$f_t = \sigma(x_t \cdot W_f + h_{t-1} \cdot U_f + b_f) \quad (1)$$

$$i_t = \sigma(x_t \cdot W_i + h_{t-1} \cdot U_i + b_i) \quad (2)$$

$$g_t = \tanh(x_t \cdot W_g + h_{t-1} \cdot U_g + b_g) \quad (3)$$

$$c_t = f_t \cdot c_{(t-1)} + i_t \cdot g_t \quad (4)$$

$$o_t = \sigma(x_t \cdot W_o + h_{t-1} \cdot U_o + b_o) \quad (5)$$

$$h_t = o_t \cdot \tanh(c_t) \quad (6)$$

2.4 Recurrent Neural Networks (RNN)

Recurrent Neural Networks (RNN) adalah jaringan saraf tiruan yang digunakan untuk mengenali pola dalam data berurutan seperti teks, suara, atau urutan waktu. RNN memiliki kemampuan unik untuk mengirimkan informasi *output* kembali ke *input* jaringan, memungkinkannya untuk mempertimbangkan informasi sebelumnya dalam membuat prediksi atau keputusan (Rumelhart et al., 1986). Perhitungan dari setiap *gate* yang terdapat pada LSTM dilakukan dengan Persamaan 7 dan 8.

$$h_t = \tanh(W_h \cdot x_t + U_h \cdot h_{t-1} + b_h) \quad (7)$$

$$y_t = W_{fc1} \cdot h_t + b_{fc1} \quad (8)$$

2.5 Gated Recurrent Unit (GRU)

Gated Recurrent Unit (GRU) merupakan variasi dari Recurrent Neural Networks (RNN) yang dikembangkan oleh Cho et al. (2014) untuk mengatasi masalah *vanishing gradient* yang sering terjadi pada RNN konvensional. GRU memanfaatkan dua *gate*, yaitu *update gate* dan *reset gate*, untuk

mengontrol aliran informasi. *Update gate* membantu model menentukan sejauh mana informasi masa lalu harus dipertahankan, sementara *reset gate* menentukan sejauh mana informasi tersebut harus diabaikan. Struktur GRU memungkinkan penangkapan dependensi jangka panjang dalam data. Perhitungan dari setiap *gate* yang terdapat pada LSTM dilakukan dengan Persamaan 9-12.

$$z_t = \sigma(W_z \cdot x_t + U_z \cdot h_{t-1} + b_z) \quad (9)$$

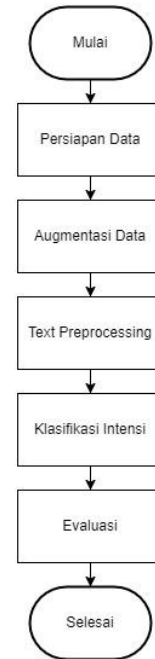
$$r_t = \sigma(W_r \cdot x_t + U_r \cdot h_{t-1} + b_r) \quad (10)$$

$$\tilde{h}_t = \tanh(W_n \cdot x_t + U_n (r_t \cdot h_{t-1}) + b_n) \quad (11)$$

$$h_t = z_t \cdot h_{t-1} + (1 - z_t) \cdot \tilde{h}_t \quad (12)$$

3. METODOLOGI

Penelitian ini menerapkan *modified Easy Data Augmentation* (EDA) pada *dataset* awal lalu dilakukan evaluasi dengan klasifikasi intensi menggunakan *sequence-based models* (LSTM, RNN, dan GRU). Strategi penelitian dapat dilihat pada Gambar 1.



Gambar 1 Diagram strategi penelitian

3.1 Persiapan Data

Penelitian menggunakan data dari *database Helpdesk TIK UB* yang berisi riwayat percakapan pengguna dengan staf Helpdesk TIK UB. Data ini terdiri dari 91,977 percakapan dari Maret 2020 hingga Desember 2022 dalam format HTML. Setelah dianalisis, terdapat 7 intensi yang dijadikan kelas target untuk penelitian *chatbot*. Data yang digunakan hanya berfokus pada isi kolom "body" karena relevansi dengan penelitian, menghasilkan 707 data. Selama proses ini, dilakukan pembersihan format

HTML agar tidak mengganggu proses augmentasi dan klasifikasi intensi.

3.2 Augmentasi Data

Pada tahap ini, dirancang *modified Easy Data Augmentation* (EDA) yang menggabungkan *POS Tagging* dan *similarity measurement* dalam operasi EDA. Terdapat empat operasi EDA, yaitu *Synonym replacement* (SR), *Random insert* (RI), *Random swap* (RS), dan *Random deletion* (RD). EDA yang dimodifikasi menerapkan *POS Tagging* menggunakan *NLTK Tagger Conditional Random Field* (CRF) untuk memastikan bahwa operasi tersebut tidak mengubah makna kalimat, sehingga data ditambahkan sambil menjaga integritas kalimat.

3.3 Text Preprocessing

Pada langkah ini, data dibersihkan untuk persiapan pelatihan model klasifikasi intensi. Proses pembersihan mencakup penghapusan tanda baca, mengkonversi huruf menjadi *lowercase*, menghilangkan *stopwords*, dan tindakan lainnya sesuai dengan kebutuhan penelitian. Setelah pembersihan, *dataset* akan di-tokenisasi dan di-vektorisasi sebelum digunakan dalam pelatihan model.

3.4 Klasifikasi Intensi

Pada penelitian ini klasifikasi intensi akan dilakukan dengan menggunakan *sequence-based models* (LSTM, RNN, dan GRU), dengan tahap inisialisasi arsitektur serta pelatihan dan pengujian masing-masing model. Pelatihan dilakukan sejumlah iterasi yang ditentukan sekaligus mengevaluasi kinerja model pada setiap iterasi, sehingga kemudian menghasilkan bobot akhir yang optimal dari masing-masing model.

3.3 Evaluasi

Pada tahap ini, dilakukan evaluasi kinerja dari algoritma klasifikasi intensi *sequence-based models* (LSTM, RNN, dan GRU) yang sudah dilakukan pada tahap sebelumnya. Evaluasi melibatkan metrik-metrik seperti akurasi, presisi, *recall*, dan *F1-score*. Selain itu, efektivitas penerapan teknik augmentasi *modified Easy Data Augmentation* (EDA) juga diukur untuk mengatasi *overfitting* akibat keterbatasan *dataset*. *Overfitting* adalah fenomena di mana model cenderung "menghafal" data pelatihan daripada "menggeneralisasi" dari data tersebut, yang dapat mengakibatkan performa model yang buruk pada data yang belum pernah dilihat sebelumnya.

4. HASIL DAN ANALISIS

4.1 Pengujian Augmentasi Data

Pada pengujian pertama ini melibatkan perbandingan kinerja tiga *sequence-based models* (LSTM, RNN, dan GRU) sebelum dan sesudah

implementasi EDA pada *dataset* awal Helpdesk TIK, menggunakan *hyperparameter* optimizer Adam, hidden size 64, batch size 32, learning rate 0,001 dan weight decay 0,00017. Data sebelum EDA berisi 707 baris dengan 7 label intensi, sedangkan setelah EDA, jumlah data meningkat menjadi 5.112 baris dengan lebih banyak variasi dalam tiap label intensi. Hasil tiap operasi lebih rinci dapat dilihat pada Tabel 1.

Tabel 1. Hasil tiap operasi dari implementasi EDA

Operasi EDA	Nilai
<i>Synonym replacement</i>	2.984
<i>Random insert</i>	707
<i>Random swap</i>	707
<i>Random deletion</i>	707
Total	5.112

Pelatihan dilakukan pada tiga model sebelum dan sesudah EDA selama 20 *epoch*. Hasil pelatihan dibandingkan pada Tabel 2-7.

Tabel 2. Perbandingan hasil pelatihan LSTM

Dataset	Acc	Loss	F1
Sebelum EDA	0,9890	0,0951	0,9891
Sesudah EDA	0,9985	0,0090	0,9985

Tabel 3. Perbandingan hasil pengujian LSTM

Dataset	Acc	Loss	F1
Sebelum EDA	0,8732	0,4636	0,8318
Sesudah EDA	0,9941	0,0204	0,9941

Tabel 4. Perbandingan hasil pelatihan RNN

Dataset	Acc	Loss	F1
Sebelum EDA	0,9874	0,1038	0,9875
Sesudah EDA	0,9985	0,0077	0,9985

Tabel 5. Perbandingan hasil pengujian RNN

Dataset	Acc	Loss	F1
Sebelum EDA	0,8732	0,4890	0,8318
Sesudah EDA	0,9980	0,0187	0,9980

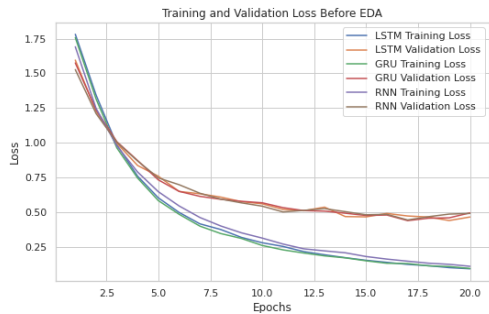
Tabel 6. Perbandingan hasil pelatihan GRU

Dataset	Acc	Loss	F1
Sebelum EDA	0,9921	0,0916	0,9920
Sesudah EDA	0,9989	0,0073	0,9989

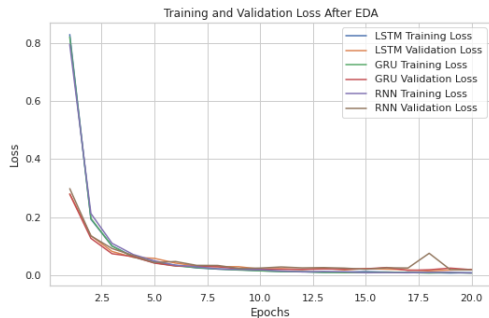
Tabel 7. Perbandingan hasil pengujian GRU

Dataset	Acc	Loss	F1
Sebelum EDA	0,8592	0,4854	0,8214
Sesudah EDA	0,9941	0,0192	0,9941

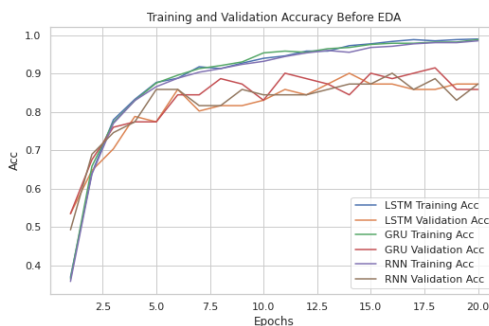
Hasil perbandingan antara model LSTM, RNN, dan GRU sebelum dan sesudah penerapan *modified* EDA pada *dataset* Helpdesk TIK menunjukkan peningkatan yang signifikan dalam kinerja model. Setelah EDA, ketiga model mengalami peningkatan akurasi yang signifikan, dengan nilai akurasi yang lebih dari 0,9985. Selain itu *loss* model juga menurun drastis setelah EDA, menjadi kurang dari 0,01.



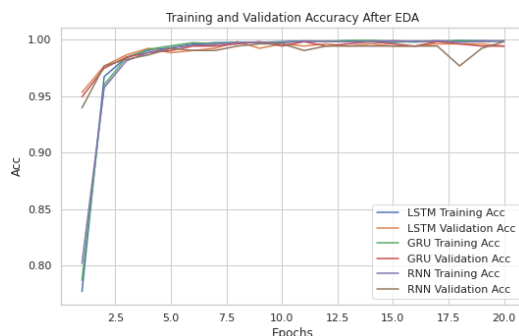
Gambar 2 Grafik *loss sequence-based models* sebelum penerapan EDA



Gambar 3 Grafik *loss sequence-based models* sesudah penerapan EDA



Gambar 4 Grafik akurasi *sequence-based models* sebelum penerapan EDA



Gambar 5 Grafik akurasi *sequence-based models* sesudah penerapan EDA

Grafik untuk *sequence-based models* sebelum penerapan (Gambar 2 dan 4), menunjukkan adanya potensi *overfitting*, dengan gap yang besar antara dan akurasi pengujian yang berfluktuasi. Hal ini sejalan dengan tabel perbandingan hasil yang menunjukkan akurasi yang lebih rendah dan *loss* yang lebih tinggi sebelum EDA. Namun, setelah EDA diterapkan (Gambar 3 dan 5), terlihat bahwa *loss* pelatihan dan pengujian menurun secara stabil dengan gap yang kecil antara keduanya, dan akurasi mendekati sempurna dengan *loss* yang jauh lebih rendah. Hal ini mencerminkan peningkatan akurasi dan skor F1, serta penurunan *loss* dalam tabel pelatihan dan pengujian.

Tabel 8. *Classification report* LSTM

<i>Dataset</i>	Prec (B)	Prec (A)	Rec (B)	Rec (A)	F1 (B)	F1 (A)
buat_reset_web	1,00	1,00	0,82	0,96	0,90	0,98
hosting						
gagal_login	0,79	0,99	0,92	1,00	0,85	0,99
masalah_vpn	1,00	1,00	0,91	1,00	0,95	1,00
mengajukan_e	0,82	0,97	0,90	1,00	0,86	0,99
mail						
nilai_it_tidak_k	0,79	1,00	0,85	1,00	0,81	1,00
eluar						
permintaan_lise	0,88	1,00	0,88	1,00	0,88	1,00
nsi_office						
ucapan_terima_kasih	1,00	1,00	0,83	1,00	0,91	1,00

** B = sebelum EDA, A = sesudah EDA

Tabel 9. *Classification report* GRU

<i>Dataset</i>	Prec (B)	Prec (A)	Rec (B)	Rec (A)	F1 (B)	F1 (A)
buat_reset_web	1,00	1,00	0,91	0,97	0,95	0,99
hosting						
gagal_login	0,73	1,00	0,92	1,00	0,81	1,00
masalah_vpn	1,00	1,00	0,91	1,00	0,96	1,00
mengajukan_e	0,90	0,97	0,90	0,99	0,90	0,98
mail						
nilai_it_tidak_k	0,73	1,00	0,85	1,00	0,79	1,00
eluar						
permintaan_lise	0,86	0,98	0,75	1,00	0,80	0,99
nsi_office						
ucapan_terima_kasih	1,00	1,00	0,67	1,00	0,80	1,00

** B = sebelum EDA, A = sesudah EDA

Tabel 10. *Classification report* RNN

<i>Dataset</i>	Prec (B)	Prec (A)	Rec (B)	Rec (A)	F1 (B)	F1 (A)
buat_reset_web	1,00	1,00	0,91	0,99	0,95	0,99
hosting						
gagal_login	0,83	1,00	0,83	1,00	0,83	1,00
masalah_vpn	0,91	1,00	0,91	1,00	0,91	1,00
mengajukan_e	0,90	0,99	0,90	1,00	0,90	0,99
mail						
nilai_it_tidak_k	0,79	1,00	0,85	1,00	0,81	1,00
eluar						
permintaan_lise	0,78	1,00	0,88	1,00	0,82	1,00
nsi_office						
ucapan_terima_kasih	1,00	1,00	0,83	1,00	0,91	1,00

** B = sebelum EDA, A = sesudah EDA

Selanjutnya dilakukan pembuatan classification report untuk masing-masing model sebelum dan sesudah EDA untuk mengevaluasi akurasi, presisi, *recall*, dan *F1-score* model dalam kasus klasifikasi intensi yang tidak seimbang, yang dapat dilihat pada Tabel 8-10. Sebelum EDA, model LSTM memiliki kinerja yang efektif dalam memprediksi kelas 'gagal_login' dengan tingkat *recall* yang tinggi, tetapi mengalami kesalahan prediksi pada kelas 'ucapan_terima_kasih' dan 'buat_reset_webhosting'. Setelah EDA, terjadi peningkatan signifikan dalam kinerja model LSTM, dengan skor yang mendekati sempurna di semua kelas, meskipun masih terdapat beberapa kesalahan prediksi pada kelas 'buat_reset_webhosting', menunjukkan perlunya penyesuaian lebih lanjut. Begitu juga dialami untuk model GRU dan model RNN, dengan skor hampir sempurna di semua kelas. mengindikasikan bahwa EDA meningkatkan pemahaman model terhadap variasi data.

Tabel 11. Hasil prediksi intensi *sequence-based models*

Teks	Aktual	LSTM	GRU	RNN
Assalamualaikum wr. wb., atau ataupun mengganggu waktunya. Perkenalkan Matheus Albert perwakilan Laboratorium komputer Teknik Industri sekarang baru berganti nama menjadi Laboratorium Optimasi, Sistem Informasi, Pendukung Keputusan. Berikut lampirkan formulir pengajuan hosting beserta dokumen diperlukan pengajuan hosting. kasih.	0	3	3	3
Assalamualaikum Wr. Wb. Saya Rayhan Kumala NIM 175060500111013, lupa password SIAM saya. Apakah dibantu? Saya mencoba pilihan lupa password, lupa pertanyaan Hint question juga. beserta sebelumnya.	1	1	1	1
sayaa cobaa terus menggunakan vpn bisaa dulu sempat hanyaa sebentar sajaa,mohon bantuannyaa	2	2	2	2
dibuatkan email pribadi nama fadilla mengisi form pengajuan email pribadi https : 3tcUPj9 .	3	3	5	3
Saya melihat nilai hasil sertifikasi IT kerjakan. Saya lupa mengambil screenshot nilai akhir ujiannya. Bagaimana melihat hasilnya??	4	4	4	4
info, departemen	5	5	5	5

matematika, bagaimana caranya membaca lisensi office lab komputer kami. Ada 76 pc 16 laptop,				
terimakasih atas layanannya mba, udah jadi siap mba	6	6	6	6

Pada Tabel 11 dapat dilihat cuplikan beberapa kelas intensi dari hasil prediksi ketiga *sequence-based models*. Dapat dilihat terdapat kesalahan prediksi pada semua model di kalimat yang sama, yakni pada baris pertama pada tabel. Hal ini kemungkinan besar dikarenakan *noise* dan beberapa kata yang ambigu pada data yang tidak diproses pada tahap *preprocessing*. *Noise* disini seperti pengucapan salam, pengenalan nama, organisasi, asal, dan lainnya yang tidak relevan dengan intensi pertanyaan dan seharusnya dihilangkan. Terdapat beberapa kata ambigu pada *dataset*, seperti kesalahan ketik pada baris ke-5 pada Tabel 11, frase seperti "mohon bantuannya" dan "Apakah dibantu?", penggunaan bahasa informal dan bahasa asing, dan banyak lainnya.

4.2 Pengujian Hyperparameter Tuning

Pada pengujian kedua, yang merupakan proses *hyperparameter tuning*, lima *hyperparameter* diuji sesuai dengan skenario pengujian yang telah direncanakan. Evaluasi hasil pengujian dilakukan berdasarkan nilai *loss* pada data uji. Pemilihan faktor *loss* data uji sebagai parameter penentu disebabkan oleh fokus penelitian pada pengurangan *overfitting*, yang terjadi ketika model memiliki performa yang baik pada data pelatihan tetapi performa menurun drastis pada data yang belum pernah dilihat sebelumnya. Hal ini sesuai dengan fokus pada *loss* data uji, karena *loss* ini mengukur sejauh mana model dapat beradaptasi dengan data yang tidak pernah digunakan dalam pelatihan, yang penting untuk memastikan generalisasi yang baik.

Pengujian dilakukan pada *dataset* yang telah mengalami *modified* EDA dengan 20 *epoch*, dan tuning *hyperparameter* dilakukan pada tiga model yang telah dirancang sebelumnya menggunakan PyTorch dengan bantuan library *itertools* untuk menghasilkan semua kombinasi *hyperparameter* yang mungkin. Setelah menguji 486 kombinasi *hyperparameter*, ditemukan kombinasi terbaik berdasarkan nilai *loss* terendah untuk setiap model. Hasil kombinasi *hyperparameter* terbaik terperinci dapat dilihat pada Tabel 12, dan metrik performa model secara rinci dapat ditemukan di Tabel 13 hingga Tabel 4.15.

Tabel 12. Hasil kombinasi *best performance hyperparameter tuning*

Model	Val loss	BS	LR	Opt	WD	HS
LSTM	0,02105	8	1e-4	RM	1e-4	32
				Spro		

RNN	0,01994	128	1e-4	P RM Spro p	1e-4	32
GRU	0,01951	32	1e-4	RM Spro p	1e-4	32

Tabel 13. *Performance metrics* LSTM berdasarkan kombinasi *best hyperparameter*

<i>Dataset</i>	<i>Acc</i>	<i>Loss</i>	<i>F1</i>
Pelatihan	0,9978	0,0145	0,9978
Pengujian	0,9941	0,0210	0,9941

Tabel 14. *Performance metrics* RNN berdasarkan kombinasi *best hyperparameter*

<i>Dataset</i>	<i>Acc</i>	<i>Loss</i>	<i>F1</i>
Pelatihan	0,9989	0,0124	0,9989
Pengujian	0,9960	0,0199	0,9960

Tabel 15. *Performance metrics* GRU berdasarkan kombinasi *best hyperparameter*

<i>Dataset</i>	<i>Acc</i>	<i>Loss</i>	<i>F1</i>
Pelatihan	0,9980	0,0131	0,9980
Pengujian	0,9960	0,0195	0,9960

Hasil pengujian *hyperparameter tuning* pada model LSTM, RNN, dan GRU mengungkapkan beberapa *insight* penting terkait dengan konfigurasi *hyperparameter*. Model LSTM, dengan *loss* pengujian sebesar 0,02105, menunjukkan efektivitasnya dalam memahami pola data dengan menggunakan *batch size* yang relatif kecil. Di sisi lain, model RNN mendapatkan *loss* pelatihan terendah di antara ketiganya, yaitu 0,0124, dengan menggunakan *batch size* yang jauh lebih besar, yakni 128. Hal ini menunjukkan kemampuan RNN dalam memproses lebih banyak data dalam satu iterasi, yang meningkatkan efisiensi pembelajaran dengan konfigurasi *hyperparameter* yang serupa dengan LSTM. Sementara itu, model GRU, dengan *loss* pengujian terendah di 0,0195, beroperasi secara efektif pada *batch size* yang berukuran sedang, yakni 32, mengindikasikan kemampuannya dalam mengelola *batch* dengan efisiensi tinggi.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil implementasi, pengujian, dan analisis terhadap penerapan metode *modified* Easy Data Augmentation (EDA) pada *dataset* Helpdesk TIK dengan menggunakan model LSTM, RNN, dan GRU untuk klasifikasi intensi pada *chatbot* bahasa Indonesia, dapat disimpulkan beberapa hal, antara lain:

1. Penerapan *modified* Easy Data Augmentation (EDA) menghasilkan peningkatan signifikan dalam performa model LSTM, RNN, dan GRU pada data yang belum dikenal. Hal ini terlihat dari penurunan signifikan dalam *loss* pengujian dan peningkatan akurasi pengujian. Grafik *loss* dan akurasi menunjukkan bahwa gap antara

performa pada data pelatihan dan pengujian berkurang setelah EDA, mengindikasikan pengurangan *overfitting*. Meskipun terdapat beberapa kesalahan prediksi pada kelas 'buat_reset_webhosting', peningkatan keseluruhan memvalidasi efektivitas EDA dalam meningkatkan akurasi dan keandalan model.

2. Berdasarkan hasil *hyperparameter tuning* yang dilakukan sebanyak 486 kombinasi *hyperparameter*, didapatkan kombinasi *hyperparameter* terbaik tiap *sequence-based models* (LSTM, RNN, dan GRU) pada *dataset* yang telah diaugmentasi. Hasil evaluasi terbaik dari ketiga *sequence-based models* dapat dilihat pada **Error! Reference source not found.** hingga **Error! Reference source not found.** RNN mendapatkan *loss* pelatihan terendah (0,0124) dan akurasi pelatihan tertinggi (0,9989) di antara ketiganya. Untuk performa pada pengujian, model GRU lebih tinggi jika dilihat dari *loss* pengujian (0,0195), terdapat perbedaan 0,0004 dengan model RNN dengan akurasi kedua model yang sama di 0,9960.

Berdasarkan penelitian ini, terdapat beberapa saran dari penulis yang dapat dijadikan pertimbangan untuk penelitian mendatang yang terkait dengan topik dari penelitian ini, antara lain:

1. Penggunaan model klasifikasi lain seperti penggunaan *Large Language Models* (GPT-3 atau model serupa yang lebih besar) dalam meningkatkan pemahaman intensi pertanyaan.
2. Eksplorasi metode *preprocessing* lain yang dapat memberikan masukan lebih baik bagi model.
3. Eksplorasi strategi modifikasi untuk teknik augmentasi data Easy Data Augmentation (EDA), gabungkan dengan metode augmentasi data lainnya untuk mendapatkan hasil yang lebih variatif.

6. DAFTAR PUSTAKA

- AL FARISI, F. A., PERDANA, R. S., & ADIKARA, P. P. (2023). Klasifikasi Intensi dengan Metode Long Short-Term Memory pada Chatbot Bahasa Indonesia. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 7(13). <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/12759>
- CHO, K., VAN MERRIENBOER, B., GULCEHRE, C., BAHDANAU, D., BOUGARES, F., SCHWENK, H., & BENGIO, Y. (2014). *Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation*.
- GENNARO, G. DI, BUONANNO, A., GIROLAMO, A. DI, OSPEDALE, A., & PALMIERI, F. A. N. (2020). Intent Classification in Question-

- Answering Using LSTM Architectures. *CoRR*, [abs/2001.09330](https://arxiv.org/abs/2001.09330).
<https://arxiv.org/abs/2001.09330>
- HELPDESK TIK. (N.D.). *Layanan Helpdesk Teknologi Informasi UB*. Retrieved October 15, 2023, from <https://helpdesk-tik.ub.ac.id/>
- KARTIKA, B. V., ALFREDO, M. J., & KUSUMA, G. P. (2023). Fine-Tuned IndoBERT Based Model and Data Augmentation for Indonesian Language Paraphrase Identification. *Revue d'Intelligence Artificielle*, 37(3), 733–743.
<https://doi.org/10.18280/ria.370322>
- KHUMAIDI, A., RAAFI'UDIN, R., & SOLIHIN, I. P. (2020). Pengujian Algoritma Long Short Term Memory untuk Prediksi Kualitas Udara dan Suhu Kota Bandung. *Jurnal Telematika*, 15(1), 13–18.
- NATASYA, ., & GIRSANG, A. S. (2022). Modified EDA and Backtranslation Augmentation in Deep Learning Models for Indonesian Aspect-Based Sentiment Analysis. *Emerging Science Journal*, 7(1), 256–272.
<https://doi.org/10.28991/ESJ-2023-07-01-018>
- PIRATLA, V. (2023). *Robustness, Evaluation and Adaptation of Machine Learning Models in the Wild*.
- RUDER, S. (2016). *An overview of gradient descent optimization algorithms*.
- RUMELHART, D. E., HINTON, G. E., & WILLIAMS, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536.
<https://doi.org/10.1038/323533a0>
- SOYDANER, D. (2020). A Comparison of Optimization Algorithms for Deep Learning. *International Journal of Pattern Recognition and Artificial Intelligence*, 34(13), 2052013.
<https://doi.org/10.1142/S0218001420520138>
- WEI, J., & ZOU, K. (2019). *EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks*.