

KOMBINASI *INTENT CLASSIFICATION* DAN *NAMED ENTITY RECOGNITION* PADA DATA BERBAHASA INDONESIA DENGAN METODE DUAL *INTENT AND ENTITY TRANSFORMER*

Zahra Asma Annisa¹, Rizal Setya Perdana², Putra Pandu Adikara³

^{1,2,3}Universitas Brawijaya, Malang

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

*Penulis Korespondensi

(Naskah masuk: 07 November 2023, diterima untuk diterbitkan: 30 Oktober 2024)

Abstrak

Pelayanan pelanggan atau *customer service* adalah sebuah bentuk upaya pemenuhan keinginan dan kebutuhan pelanggan yang disertai dengan ketepatan penyampaian sesuai standar perusahaan demi memenuhi harapan pelanggan. Pada beberapa kasus seperti layanan perbankan, diperlukan layanan pelanggan yang dapat diakses setiap saat dengan ketepatan dan daya tanggap tinggi. Atas kebutuhan pelayanan dengan kualitas tinggi ini, perusahaan dapat mengaplikasikan konsep pelayanan prima. Salah satu penerapan konsep kecerdasan buatan demi pelaksanaan pelayanan prima adalah penggunaan *chatbot*, yang memerlukan metode yang tepat bagi proses klasifikasi intensi pengguna maupun *Named Entity Recognition* (NER). Salah satu kekurangan dari pelaksanaan klasifikasi intensi dan NER secara terpisah terletak pada representasi numerik yang digunakan dalam tiap model. Meski menggunakan data latih dan arsitektur model yang serupa, model dapat menghasilkan representasi numerik yang berbeda dalam tahap fiturisasi, sehingga berpotensi mengurangi tingkat generalisasi model. Untuk mengatasi masalah tersebut, klasifikasi intensi dan NER dapat digabungkan dengan menggunakan mekanisme *multi-task learning* dalam bentuk model Dual Intent and Entity Transformer (DIET). Penelitian dilakukan dengan memanfaatkan data sekunder dari Helpdesk TIK UB, merancang model DIET menggunakan pustaka PyTorch dan Transformers, lalu mengevaluasi model DIET menggunakan *f1-score*. Kombinasi *hyperparameter* terbaik yang didapatkan adalah *warm-up step* sebesar 70, *early stopping patience* sebesar 15, *weight decay* sebesar 0,01, bobot *loss* NER sebesar 0,6, dan bobot *loss* klasifikasi intensi berupa 0,4. Kombinasi *hyperparameter* yang telah diperoleh menghasilkan kapabilitas yang berbeda apabila terdapat perubahan dalam data yang digunakan, karena kapabilitas model DIET baik dalam melakukan klasifikasi intensi maupun NER sangat bergantung terhadap data.

Kata kunci: *klasifikasi intensi, named entity recognition, dual intent and entity transformer, IndoBERT, pelayanan pelanggan*

COMBINING *INTENT CLASSIFICATION* AND *NAMED ENTITY RECOGNITION* ON *INDONESIAN DATA USING DUAL INTENT AND ENTITY TRANSFORMER*

Abstract

*Customer service is a way to fulfill the wants and needs of customers accompanied by the accuracy of delivery according to company standards in order to meet customer expectations. In some cases such as banking services, customer service is needed that can be accessed at any time with high accuracy and responsiveness. For this high-quality service requirement, companies can implement the concept of excellent service. One application of artificial intelligence for service excellence is the use of a chatbot, which requires an appropriate method for the classification of user intent and Named Entity Recognition (NER). One of the drawbacks of performing intent classification and NER separately lies in the different numerical representations used in each model. Despite using similar training data and model architecture, the models may produce different numerical representations in the featurization stage, potentially reducing the generalization ability of the model. To overcome this problem, intent and NER classification can be combined using a multi-task learning mechanism in the form of a Dual Intent and Entity Transformer (DIET) model. The research was conducted by utilizing secondary data from Helpdesk TIK UB, designing DIET models using PyTorch and Transformers libraries, then evaluating DIET models using *f1-score*. The best hyperparameter combination obtained is a warm-up step of 70, early stopping patience of 15, weight decay of 0.01, NER loss weight of 0.6, and intent classification loss weight of 0.4. The combination of hyperparameters that have been obtained produce different capabilities if there are changes in the data that is*

used because the capabilities of the DIET model in both intention and NER classification are highly dependent on the data.

Keywords: *intent classification, named entity recognition, dual intent and entity transformer, IndoBERT, customer service*

1. PENDAHULUAN

Pelayanan pelanggan atau *customer service* adalah sebuah bentuk upaya pemenuhan keinginan dan kebutuhan pelanggan yang disertai dengan ketepatan penyampaian sesuai standar perusahaan demi memenuhi harapan pelanggan (Tjiptono, 2005). Dengan kualitas pelayanan yang baik, pelanggan akan merasa seluruh kebutuhannya atas produk yang terkait telah terpenuhi oleh perusahaan. Oleh karena itu, dapat disimpulkan bahwa pelayanan pelanggan merupakan hal yang sangat berpengaruh terhadap kepuasan pelanggan juga nama baik perusahaan (Kotler & Keller, 2009). Salah satu tolak ukur kualitas pelayanan pelanggan yaitu daya tanggap suatu kebijakan untuk membantu pelanggan dengan memberikan pelayanan yang cepat juga tepat.

Terdapat beberapa bentuk pelayanan pelanggan yang dilaksanakan secara daring salah satu di antaranya yaitu layanan pesan tertulis secara langsung (*live chat*) melalui situs perusahaan. Waktu tersedianya layanan pelanggan bergantung pada jenis layanan yang diberikan serta tingkat kepentingan masalah yang berpotensi dialami oleh pelanggan. Pada beberapa kasus seperti layanan perbankan, diperlukan layanan pelanggan yang dapat diakses setiap saat dengan ketepatan dan daya tanggap tinggi. Atas kebutuhan pelayanan dengan kualitas tinggi ini, perusahaan dapat mengaplikasikan konsep pelayanan prima.

Salah satu penerapan konsep kecerdasan buatan demi pelaksanaan pelayanan prima adalah penggunaan *chatbot*. *Chatbot* merupakan sebuah sistem perangkat lunak yang dapat berinteraksi dengan manusia menggunakan bahasa alami layaknya manusia pada umumnya. Sistem ini bergantung pada berbagai teknik pemrosesan bahasa alami baik untuk melakukan interpretasi data teks masukan maupun memberikan respon teks (Ruindungan & Jacobus, 2021).

Dalam pengembangan *chatbot*, diperlukan metode yang tepat untuk melaksanakan *intent classification* atau klasifikasi intensi dan *Named Entity Recognition* (NER) agar *chatbot* dapat menyediakan respon yang sesuai. Namun, dalam beberapa penelitian yang telah dilakukan, ditemukan bahwa pembangunan model berskala besar dengan data yang telah dilatih sebelumnya ternyata tidak ideal untuk membangun *chatbot* akibat banyaknya biaya serta waktu yang dibutuhkan.

Dalam pengembangan *chatbot*, klasifikasi intensi diperlukan agar *chatbot* dapat menentukan inti dari sekumpulan data teks masukan. *Named Entity Recognition* (NER) diperlukan untuk menemukan

"entitas bernama" dalam teks secara otomatis lalu mengklasifikasikannya ke dalam kategori yang telah ditentukan sebelumnya seperti orang, lokasi dan organisasi (Jauregi Unanue, Zare Borzeshi & Piccardi, 2017). Kedua bentuk pemrosesan bahasa alami ini dapat dilakukan dengan menggunakan arsitektur model yang serupa, seperti metode linear dengan *pre-trained* BERT untuk mendapatkan konteks data teks dalam bentuk numerik yang telah dilakukan oleh Pandey, Daw & Pudi (2022) dan Li et al. (2019).

Salah satu kekurangan dari pelaksanaan klasifikasi intensi dan NER secara terpisah terletak pada representasi numerik yang digunakan dalam tiap model. Meski menggunakan data latih dan arsitektur model yang serupa, model dapat menghasilkan representasi numerik yang berbeda dalam tahap fiturisasi, sehingga berpotensi mengurangi tingkat generalisasi model. Untuk mengatasi masalah tersebut, klasifikasi intensi dan NER dapat digabungkan dengan menggunakan mekanisme *multi-task learning*. Dalam mekanisme ini, pelatihan dua atau lebih jenis algoritma dilakukan secara paralel (Caruana, 1997).

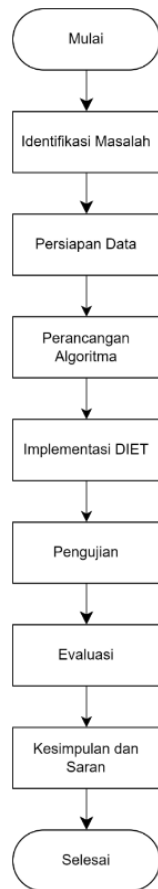
Salah satu penerapan konsep *multi-task learning* dalam pengembangan *chatbot* adalah Dual Intent & Entity Transformer (DIET). DIET merupakan sebuah arsitektur *multi-task transformer* yang dapat menangani klasifikasi intensi dan NER secara bersamaan. Berbeda dengan arsitektur model bahasa lainnya, DIET merupakan arsitektur modular yang fleksibel dan lebih cepat dilatih (Bunk et al., 2020).

Berdasarkan permasalahan yang telah diuraikan, penelitian ini akan memanfaatkan konsep arsitektur model DIET untuk mengeksekusi klasifikasi intensi dan NER secara paralel terhadap data berbahasa Indonesia. Penelitian ini juga memanfaatkan *pre-trained* BERT dalam proses fiturisasi. Dengan dilaksanakannya penelitian ini, diharapkan penerapan model DIET sebagai salah satu bentuk *multi-task learning* dapat membantu mewujudkan pelayanan prima yang ideal menggunakan Bahasa Indonesia.

2. METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder berupa data daftar pertanyaan yang sering dipertanyakan kepada Helpdesk TIK UB. Data diperoleh melalui *database* Helpdesk TIK UB dengan melakukan kueri terhadap daftar topik yang dan entitas yang sering digunakan dalam pesan pengguna. Daftar kosakata yang digunakan dalam proses tokenisasi dan fiturisasi terhadap data ini

terdapat dalam modul IndoBERT. Diagram alir penelitian dapat dilihat pada Gambar 1.



Gambar 1 Diagram Alir Metodologi Penelitian

3. PERANCANGAN

3.1 Identifikasi Masalah

Permasalahan yang diangkat pada penelitian ini adalah penerapan metode DIET dalam melaksanakan klasifikasi intensi dan NER secara bersamaan terhadap data berbahasa Indonesia. Dengan mempertimbangkan hasil penelitian Bunk et al. (2020), penelitian ini akan menggunakan model BERT dengan variasi IndoBERT. Algoritma DIET yang dirancang akan berfokus dalam arsitektur DIET yang telah dikenalkan oleh Bunk et al. (2020). IndoBERT merupakan sebuah variasi model pre-trained BERT yang dilatih menggunakan kosakata Bahasa Indonesia dengan lebih dari 220 juta kata dari berbagai sumber resmi seperti Wikipedia Indonesia. Model ini termasuk model yang bersifat monolingual atau berfokus dalam satu bahasa (Koto et al., 2020).

Setelah arsitektur telah terbentuk, dilakukan pengujian model DIET dengan menggunakan kombinasi *hyperparameter* dengan metode sekuensial. Dalam metode pengujian ini, nilai hasil pengujian terbaik pada kombinasi parameter pertama digunakan dalam pengujian kombinasi parameter kedua, dan seterusnya hingga parameter terakhir untuk mempersingkat waktu pengujian.

3.2 Persiapan Data

Persiapan data menggunakan *dataset* daftar pertanyaan yang sering dipertanyakan di Helpdesk TIK UB mencakup beberapa tahap.

3.2.1 Perolehan Data

Pertama, *dataset* diperoleh dari *database* aplikasi sistem tiket daring Helpdesk TIK Universitas Brawijaya. Konteks dari data yang digunakan adalah data kumpulan pertanyaan yang sering dipertanyakan oleh sivitas akademika Universitas Brawijaya yang memerlukan bantuan perihal layanan IT di Universitas Brawijaya. Data teks ini disajikan dalam Bahasa Indonesia. *Dataset* lalu ditinjau kembali untuk memilih kueri yang dapat digunakan dalam pembangunan *chatbot* sesuai dengan intensi yang digunakan, seperti bertanya mengenai cara pembuatan surel Universitas Brawijaya. Pelabelan intensi dilakukan dengan menandai intensi dari tiap baris data latih sesuai dengan kueri yang digunakan untuk mengambil data latih, dengan contoh data latih beserta label intensinya yang terpapar pada tabel 1.

Tabel 1 Contoh Data Latih

Data Teks	Intensi
</p>Saya tidak bisa login SIAM dan BAIS</p>	gagal_login

3.2.2 Pembersihan dan Pelabelan Data

Pembersihan data dilakukan dengan memperbaiki data mentah menjadi data yang dapat diproses oleh *chatbot* dengan mengubah seluruh huruf dalam data latih menjadi huruf kecil serta menghilangkan *tag* HTML dan karakter non-ASCII. Pelabelan entitas dilakukan secara manual untuk menandai intensi dan entitas secara singkat namun deskriptif dengan menggunakan kurung siku untuk mengurung entitas yang terdapat dalam data latih dan kurung Penambahan label pada data sangat membantu dalam mengevaluasi kinerja sistem. Tabel 2 merupakan contoh pelabelan data intensi maupun entitas data latih. Selanjutnya, data yang sudah siap digunakan diformat dalam fail YAML. Seluruh label entitas dan intensi yang digunakan dalam penelitian ini dijelaskan dalam tabel 3 dan 4 dengan civitas akademika yang menghubungi Helpdesk TIK UB sebagai identitas pengirim pesan.

Tabel 2 Contoh Hasil Pembersihan dan Pelabelan Data

Data Asli		Hasil Persiapan Data
Data Teks	Intensi	
</p>Saya tidak bisa login SIAM dan BAIS</p>	gagal_login	- intent: gagal_login examples: - saya tidak bisa login [siam](aplikasi) dan bais

Tabel 3 Deskripsi Label Entitas

Label Entitas	Deskripsi
aplikasi	Judul aplikasi
nim	Nomor induk mahasiswa dan pegawai
email	E-mail pengirim pesan

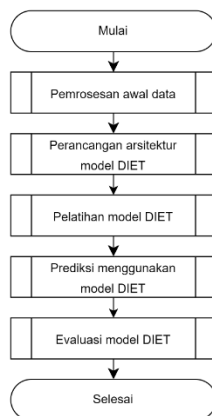
Label Entitas	Deskripsi
orang	Nama pengirim pesan
organisasi	Organisasi dari pengirim pesan
waktu	Waktu ujian sertifikasi IT
domain	Domain situs
nomor	Nomor telepon pengirim pesan
OS	Sistem operasi dari perangkat milik pengirim pesan
gelombang	Gelombang ujian sertifikasi IT yang diikuti oleh pengirim pesan

Tabel 4 Deskripsi Label Intensi

Label Intensi	Deskripsi
gagal_login	Pengirim pesan gagal melakukan proses <i>login</i> ke dalam suatu aplikasi atau domain
masalah_vpn	Pengirim pesan memiliki masalah terkait VPN
mengajukan_email	Pengirim pesan berniat mengajukan alamat email baru
nilai_it_tidak_keluar	Pengirim pesan mempertanyakan nilai hasil sertifikasi IT yang belum tertera di SIAM UB
permintaan_lisensi_office	Pengirim pesan mengajukan akses Microsoft Office 365
mengajukan_webhosting	Pengirim pesan mengajukan webhosting
ucapan_terima_kasih	Pengirim pesan hanya menyatakan terima kasih

3.3 Perancangan Algoritma

Diagram alir perancangan algoritma untuk melaksanakan klasifikasi intensi dan NER secara bersamaan menggunakan metode DIET dapat dilihat pada Gambar 2. Model DIET dibangun dengan memanfaatkan pustaka PyTorch dan Transformers untuk mendefinisikan blok *transformer* dalam model DIET.



Gambar 2 Diagram Alir Perancangan Algoritme

Tahap pemrosesan awal data mencakup tahap tokenisasi data dan ekstraksi label intensi maupun entitas. Perancangan arsitektur model DIET dilakukan dengan memanfaatkan PyTorch dan Transformers dalam membangun arsitektur model DIET. Dalam model DIET yang dibangun, digunakan lapisan linear serta BERT untuk melakukan tahap NER dan klasifikasi intensi.

3.4 Perancangan Pengujian

Pengujian dilakukan untuk menemukan pengaruh perubahan serta nilai *hyperparameter* yang sesuai untuk model DIET yang telah disusun secara sekuensial. Nilai *hyperparameter* terbaik ditentukan oleh nilai *F1-Score* dari masing-masing kombinasi *hyperparameter*. Nilai *hyperparameter* yang diuji tertera pada Tabel 5.

Tabel 5 Nilai Hyperparameter-Tuning

Hyperparameter	Nilai		
Warm-up step	65	70	75
Early stopping patience	5	10	15
Weight decay	0,1	0,01	0,001
Rasio bobot loss NER dan klasifikasi intensi	0,1 & 0,9	...	0,9 & 0,1

4. PENGUJIAN DAN ANALISIS

4.1 Pengujian terhadap Warm-up Step

Nilai *warm-up step* dipilih berdasarkan 10% dari total *steps* dalam pelatihan model menggunakan *hyperparameter* rasio data latih sebesar 0,95, *epoch* maksimum sebesar 100, *batch size* sebesar 8, *weight decay* sebesar 0,01, *early stopping patience* sebesar 15, *early stopping threshold* sebesar 0,0001, bobot *loss* NER sebesar 0,1 dan bobot *loss* klasifikasi intensi sebesar 0,9.

Tabel 6 Hasil Pengujian terhadap Warm-up Step

Warm-up Step	F1-Score NER	F1-Score Klasifikasi intensi
65	0,9814	0,9387
70	0,9847	0,9476
75	0,9824	0,9473

Berdasarkan hasil pengujian, didapatkan bahwa kombinasi *hyperparameter* senilai 70 dan 75 menghasilkan *f1-score* dengan selisih yang sangat kecil. Jumlah *warm-up step* bergantung pada jumlah data yang digunakan dalam proses pelatihan dan berpengaruh terhadap waktu yang diperlukan dalam melatih model. Dengan nilai *f1-score* tertinggi, dipilih nilai *hyperparameter warm-up step* sebesar 70.

4.2 Pengujian terhadap Early Stopping Patience

Berdasarkan hasil pengujian pada Tabel 7, didapatkan bahwa model dengan nilai *early stopping patience* sebesar 5 hampir tidak dapat melakukan NER dan klasifikasi intensi. Model dengan nilai *early stopping patience* sebesar 10 meraih konvergensi lebih cepat dari model dengan nilai *early stopping patience* sebesar 15, tetapi memiliki nilai *f1-score* yang lebih rendah.

Tabel 7 Hasil Pengujian terhadap Early Stopping Patience

Early Stopping Patience	F1-Score NER	F1-Score Klasifikasi intensi
5	0,9809	0,0987
10	0,9836	0,9441
15	0,9847	0,9476

Hasil pengujian terhadap *early stopping patience* pada Tabel 7 menunjukkan bahwa makin bertambahnya nilai *early stopping*, makin baik pula kapabilitas model dalam melakukan NER dan klasifikasi intensi. Pengujian lebih lanjut lalu dilakukan menggunakan nilai *early stopping patience* sebesar 18 dan 20. Nilai *f1-score* model menggunakan kedua nilai tersebut ditampilkan pada Tabel 8 yang menunjukkan bahwa kapabilitas model paling optimal masih terletak pada kombinasi dengan nilai *early stopping patience* sebesar 15. Model dengan nilai *early stopping patience* sebesar 20 menghasilkan model yang bersifat *underfitting*. Berdasarkan nilai *f1-score* yang tertinggi, dipilih *hyperparameter early stopping patience* sebesar 15.

Tabel 8 Hasil Pengujian terhadap *Early Stopping Patience*

<i>Early Stopping Patience</i>	F1-Score NER	F1-Score Klasifikasi intensi
5	0,9809	0,0987
10	0,9836	0,9441
15	0,9847	0,9476
18	0,9809	0,9119
20	0,9809	0,0000

4.3 Pengujian terhadap *Weight Decay*

Hasil pengujian terhadap *weight decay* terdapat pada Tabel 9. Dengan nilai *f1-score* tertinggi, dipilih nilai *hyperparameter weight decay* sebesar 0,01. Nilai ini sama dengan nilai *weight decay* yang digunakan dalam penelitian yang dilakukan oleh Izsak, Berchansky dan Levy (2021) yang berfokus dalam pelatihan BERT menggunakan biaya yang terbatas.

Tabel 9 Hasil Pengujian terhadap *Warm-up Step*

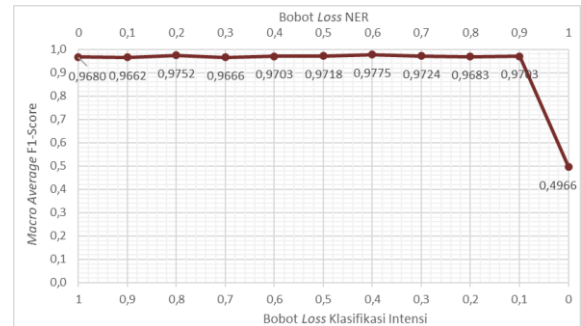
<i>Weight Decay</i>	F1-Score NER	F1-Score Klasifikasi intensi
0,1	0,9896	0,9231
0,01	0,9847	0,9476
0,001	0,9814	0,8710

4.4 Pengujian terhadap Bobot *Loss* NER dan Klasifikasi Intensi

Pengujian terhadap bobot *loss* NER dan klasifikasi intensi untuk menghitung *loss* total model dilakukan dengan menggunakan 11 kombinasi nilai dalam jangkauan 0 hingga 1. Naiknya bobot *loss* NER berbanding lurus dengan nilai *f1-score* NER model DIET, begitu pula sebaliknya. Model dengan kombinasi bobot *loss* NER sebesar 0 dan bobot *loss* klasifikasi intensi sebesar 1 hanya bergantung pada *loss* klasifikasi intensi untuk menghitung *loss* total, tetapi masih mendapatkan nilai *f1-score* yang tinggi bagi proses NER. Hal ini dikarenakan mayoritas dari data latih yang digunakan bukan merupakan bagian dari daftar entitas bernama dan model dengan kombinasi nilai bobot tersebut hanya dapat menghasilkan prediksi *subword* yang bukan merupakan entitas.

Dalam hasil pengujian penyetalan bobot *loss* NER dan klasifikasi intensi model DIET, kombinasi

terbaik dipilih melalui perbandingan *macro average* dari *f1-score* masing-masing kombinasi. Nilai *macro average* dipilih karena terdapat kombinasi bobot *loss* yang hanya bergantung pada *loss* NER, begitu pula pada *loss* klasifikasi intensi. Perbandingan nilai *macro average f1-score* tiap kombinasi dapat dilihat pada Gambar 3. Berdasarkan hasil pengujian, didapatkan kombinasi dengan nilai *macro average f1-score* terbaik di antara seluruh kombinasi bobot *loss* yang diuji, yaitu nilai bobot *loss* NER sebesar 0,6 dan bobot *loss* klasifikasi intensi sebesar 0,4.

Gambar 3 Grafik *Macro Average F1-Score* Pengujian Bobot *Loss* NER dan Klasifikasi intensi

4.5 Hasil Kombinasi *Hyperparameter* Terbaik

Berdasarkan hasil pengujian yang telah dilakukan, didapatkan kombinasi nilai *hyperparameter* terbaik dengan nilai *f1-score* NER sebesar 0,9330 dan nilai *f1-score* klasifikasi intensi sebesar 0,9621 berupa *warm-up step* sebesar 70, *early stopping patience* sebesar 15, *weight decay* sebesar 0,01, bobot *loss* NER sebesar 0,6 serta bobot *loss* klasifikasi intensi sebesar 0,4. Kombinasi *hyperparameter* ini menghasilkan model yang melalui proses pelatihan sebanyak 17 *epoch* dengan menggunakan *learning rate* dinamik yang berubah pada tiap *epoch* hingga mencapai nilai 4,19E-05.

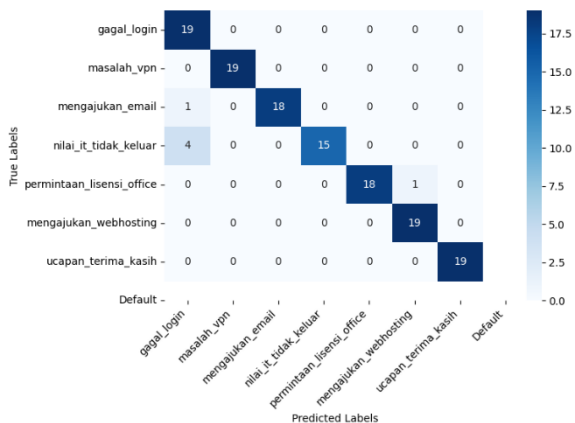
Model DIET dengan *dataset* Helpdesk TIK UB menggunakan model *pre-trained* BERT dalam variasi yang berbeda dengan penelitian sebelumnya, yaitu IndoBERT. Dengan jumlah parameter dan kamus kata yang berbeda, hal ini dapat menyebabkan *hyperparameter* optimal yang berbeda pula dengan penelitian sebelumnya. Model ini juga tidak menggunakan *learning rate* konstan. *Learning rate* diubah dalam tiap *epoch* proses pelatihan hingga model menemui titik stabil. Dengan ukuran data yang lebih kecil, perlakuan *warm-up* dalam pelatihan model juga lebih besar dari penelitian sebelumnya, yaitu 0,00004.

Model DIET dapat digunakan untuk pelaksanaan klasifikasi intensi dan NER secara sekaligus, terbukti dengan *confusion matrix* yang dihasilkan oleh model DIET pada Gambar 4 dan Gambar 5. Model dapat melakukan klasifikasi intensi dengan baik, dengan kesalahan yang dapat disebabkan oleh adanya konteks data dari suatu intensi yang sama dengan intensi lainnya seperti pengenalan diri. Berbeda dengan *confusion matrix*

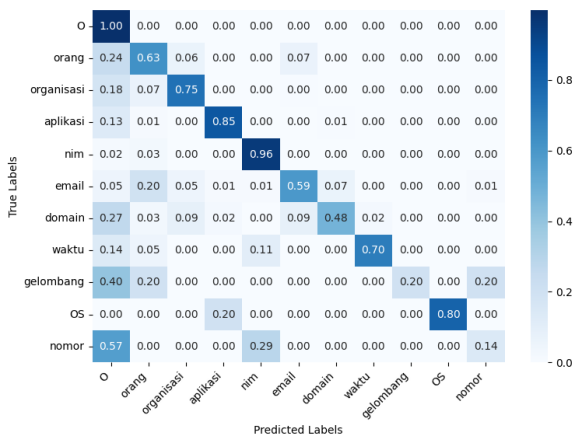
klasifikasi intensi, *confusion matrix* NER menunjukkan kapabilitas prediksi yang lebih tersebar dan bias terhadap kata non-entitas.

Pada Gambar 5, terdapat tingkat kesalahan yang cukup tinggi bagi label “gelombang” dan “nomor”. Bagian berlabel “gelombang” sering diklasifikasikan sebagai “O”, sedangkan bagian berlabel “nomor” sering diklasifikasikan sebagai “O” dan “nim”.

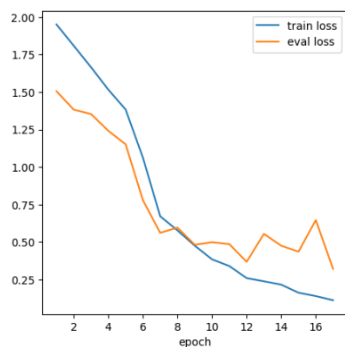
Nilai *f1-score* yang tinggi selama pengujian untuk NER didukung oleh struktur data latih maupun uji yang didominasi oleh data teks non-entitas. Oleh karena itu, dapat disimpulkan bahwa model DIET sangat bergantung pada data yang digunakan.



Gambar 4 Confusion Matrix Klasifikasi menggunakan Model DIET



Gambar 5 Confusion Matrix NER menggunakan Model DIET



Gambar 6 Grafik Perubahan Training Loss dan Evaluation Loss Model DIET

Melalui grafik perubahan nilai *training loss* dan *evaluation loss* dari hasil kombinasi *hyperparameter* terbaik pada Gambar 6, dapat disimpulkan bahwa kapabilitas model terhadap data latih makin meningkat. Masih terdapat kemungkinan bahwa model belum stabil karena terdapat peningkatan *evaluation loss* pada beberapa *epoch* akhir.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Melalui penelitian ini, dapat disimpulkan bahwa model DIET menggunakan data berbahasa Indonesia dapat dibangun menggunakan pustaka PyTorch dan Transformers. Pustaka PyTorch digunakan untuk membangun arsitektur DIET secara garis besar, dan pustaka Transformers digunakan untuk mendefinisikan blok *transformer* yang diperlukan dalam model DIET. Model DIET dalam penelitian ini menggunakan IndoBERT dalam perhitungan representasi numerik data teks yang juga difasilitasi oleh Hugging Face.

Berdasarkan tahap pengujian, kapabilitas model terus berubah seiring pengubahan *hyperparameter* model DIET secara sekensial. Kombinasi *hyperparameter* terbaik dengan nilai *f1-score* NER sebesar 0,9330 dan nilai *f1-score* klasifikasi intensi sebesar 0,9621 adalah *warm-up step* sebesar 70, *early stopping patience* sebesar 15, *weight decay* sebesar 0,01, bobot *loss* NER sebesar 0,6, dan bobot *loss* klasifikasi intensi berupa 0,4. Berdasarkan observasi terhadap kapabilitas model, model bersifat bias terhadap *subword* non-entitas karena data latih didominasi oleh kata non-entitas. Kombinasi *hyperparameter* yang telah diperoleh dapat menghasilkan kapabilitas yang berbeda apabila terdapat perubahan dalam data yang digunakan, karena kapabilitas model DIET baik dalam melakukan klasifikasi intensi maupun NER sangat bergantung terhadap data.

5.2 Saran

Berdasarkan hasil penelitian, terdapat beberapa saran perbaikan pengembangan model DIET yang dapat diterapkan dalam penelitian selanjutnya. Dengan *dataset* yang masih terbatas, disarankan untuk menambah data latih, terutama data entitas bernama. Dalam data yang baru, bisa ditambahkan label jenis entitas yang lebih banyak atau pemilihan data latih yang lebih selektif agar perbedaan jumlah antara kata non-entitas dan entitas bernama tidak terlalu besar.

Untuk penelitian selanjutnya, disarankan penggunaan metode lain seperti *grid search* untuk mencari *hyperparameter* yang bersifat *global optimum*. Pengaplikasian model DIET juga tidak terbatas pada penggunaan lapisan linear dan IndoBERT dalam melakukan klasifikasi intensi maupun NER. Dalam penelitian selanjutnya, bisa diaplikasikan metode klasifikasi lain seperti LSTM

dan model *pre-trained* lain seperti Language-Agnostic BERT Sentence Embedding (LaBSE) untuk melaksanakan tokenisasi dan fiturisasi

DAFTAR PUSTAKA

- BUNK, T., VARSHNEYA, D., VLASOV, V. DAN NICHOL, A., 2020. DIET: Lightweight Language Understanding for Dialogue Systems. [online] <https://doi.org/10.48550/arxiv.2004.09936>.
- CARUANA, R., 1997. Multitask Learning. *Machine Learning*, [online] 28(1), pp.41–75. <https://doi.org/10.1023/A:1007379606734>.
- IZSAK, P., BERCHANSKY, M. DAN LEVY, O., 2021. How to Train BERT with an Academic Budget. [online] Tersedia di: <<http://arxiv.org/abs/2104.07705>> [Diakses 18 Juni 2023].
- JAUREGI UNANUE, I., ZARE BORZESHI, E. DAN PICCARDI, M., 2017. Recurrent neural networks with specialized word embeddings for health-domain named-entity recognition. *Journal of Biomedical Informatics*, [online] 76, pp.102–109. <https://doi.org/10.1016/j.jbi.2017.11.007>.
- KOTLER, P. DAN KELLER, K.L., 2009. *Manajemen Pemasaran*. Jakarta: Erlangga.
- KOTO, F., RAHIMI, A., LAU, J.H. DAN BALDWIN, T., 2020. IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP. [online] Tersedia di: <<https://arxiv.org/abs/2011.00677>> [Diakses 25 Juni 2023].
- LI, X., BING, L., ZHANG, W. DAN LAM, W., 2019. Exploiting BERT for End-to-End Aspect-based Sentiment Analysis. [online] Tersedia di: <<https://arxiv.org/abs/1910.00883>> [Diakses 19 Juni 2023].
- PANDEY, A., DAW, S. DAN PUDI, V., 2022. Multilinguals at SemEval-2022 Task 11: Transformer Based Architecture for Complex NER. [online] Tersedia di: <<https://arxiv.org/abs/2204.02173>> [Diakses 19 Juni 2023].
- RUINDUNGAN, D.G.S. DAN JACOBUS, A., 2021. Chatbot Development for an Interactive Academic Information Services using the Rasa Open Source Framework. *Jurnal Teknik Elektro dan Komputer*, [online] 10(1), pp.61–68. <https://doi.org/https://doi.org/10.35793/jtek.10.1.2021.31150>.
- TJIPTONO, F., 2005. *Pemasaran Jasa*. Yogyakarta: Andi.

Halaman ini sengaja dikosongkan.