

# Machine Translation Task :

## Translate Informal to Formal Sentence Indonesian Language with Transform Model

Dimiyadi Suhartono  
Fakultas Teknologi Maju dan  
Multidisiplin  
Universitas Airlangga  
Surabaya, Indonesia  
dimiyadi.suhartono-2021@ftmm.unair.a  
c.id

Davano Al Raffi Abdul Jabbar  
Fakultas Teknologi Maju dan  
Multidisiplin  
Universitas Airlangga  
Surabaya, Indonesia  
davano.al.raffi-2021@ftmm.unair.ac.id

Aditya Ananda  
Fakultas Teknologi Maju dan  
Multidisiplin  
Universitas Airlangga  
Surabaya, Indonesia  
aditya.ananda-2021@ftmm.unair.ac.id

Stevanus Sembiring  
Fakultas Teknologi Maju dan  
Multidisiplin  
Universitas Airlangga  
Surabaya, Indonesia  
stevanus.sembiring-2021@ftmm.unair.a  
c.id

M Zidane Ramadhan  
Fakultas Teknologi Maju dan  
Multidisiplin  
Universitas Airlangga  
Surabaya, Indonesia  
m.zidane.ramadhan-2021@ftmm.unair.  
ac.id

**Abstrak** — Mesin Terjemahan merupakan sebuah aplikasi yang penting dalam Pencarian informasi Lintas Bahasa. Terjemahan menggunakan kamus adalah dapat diandalkan meskipun kosakatanya terbatas. Dalam penggunaan sehari-hari, bahasa Indonesia penuh dengan kesalahan pada formalitas kalimat, yaitu penyimpangan dari standar dalam hal kosa kata, ejaan, dan urutan kata. Di samping itu, Model NLP Indonesia yang tersedia saat ini biasanya dikembangkan dengan mempertimbangkan standar bahasa Indonesia. Penelitian ini akan membahas perpindahan gaya kalimat dari bahasa Indonesia informal ke formal sebagai masalah dari terjemahan mesin dengan sumber daya yang tersedia. Penelitian ini berfungsi untuk menerjemahkan kalimat informal menjadi kalimat formal. Penelitian ini menggunakan berbagai model dengan pendekatan *machine learning* serta *deep learning* dalam proses penerjemahan dan membandingkan satu sama lain untuk melihat model yang lebih baik. Evaluasi BLEU menentukan diterima atau tidaknya kalimat hasil prediksi dengan mengukur kedekatan konteks kalimat formal dari dataset. Kesimpulan yang didapatkan ialah *Pre-Trained Language Model* menggunakan MarianMT menempati peringkat satu dalam perbandingan model yang terbaik dengan skor BLEU yaitu 41,6366. Pada model tersebut hasil yang dikeluarkan memiliki kelebihan dapat mengenali pola kalimat dan memiliki informasi dasar dari *transform* model. Akan tetapi, hasil tersebut memberikan penilaian terhadap model *training* masih belum cukup baik. Sehingga model tersebut terkadang membuat kata-kata yang di luar *labels* dan mengganti tanda baca yang kurang sesuai. Penelitian ini perlu peningkatan dalam hal menambah jumlah *dataset* yang digunakan, memperbaiki *preprocessing* pada data latih, dan memperbaiki model dengan melakukan perubahan arsitektur model, penyesuaian *hyperparameter*, atau peningkatan kapasitas model.

**Kata Kunci** — BLEU, Deep Learning, Machine Learning, Machine Translation, NLP

### I. PENDAHULUAN

Penerjemahan mesin merupakan salah satu aplikasi dari bidang *natural language processing* (NLP), di mana proses

ini memanfaatkan teknik matematika dan algoritma untuk menerjemahkan kata atau dokumen dari satu bahasa ke bahasa lain [1]. Tujuan terjemahan mesin adalah menyampaikan secara akurat makna isi bahasa sumber dalam bahasa target. Proses ini penting untuk menghilangkan hambatan bahasa dan memfasilitasi komunikasi antar komunitas linguistik yang berbeda.[6] Konsep ini pertama kali dikemukakan oleh Warren Weaver pada tahun 1947, hanya satu tahun setelah komputer pertama, integrator numerik elektronik dan komputer, dikembangkan. Sejak saat itu, Mesin ini dianggap sebagai salah satu tugas paling menantang di bidang pemrosesan bahasa alami (NLP) [2]. Terjemahannya sering kali menggunakan kamus sebagai sumbernya. Kamus lebih dipercaya karena dibuat oleh ahlinya siapa yang bisa dipercaya dengan hasilnya. Terjemahan menggunakan kamus sangat bergantung pada jumlah kata di dalam kamus [3].

Mesin penerjemah sampai saat ini telah berkembang dengan pesat, tetapi masih terdapat sejumlah permasalahan yang perlu diatasi, terutama pada bahasa Indonesia. Salah satu permasalahan yang dihadapi yaitu kompleksitas bahasa Indonesia yang memiliki banyak variasi dan dialek. Selain itu, bahasa Indonesia memiliki struktur kalimat yang kompleks dan banyak kata yang memiliki makna ganda. Hal ini membuat mesin penerjemah sulit untuk memahami konteks dan makna yang tepat. Meskipun mesin penerjemah jaringan saraf tiruan telah mencapai kesuksesan yang besar dalam beberapa tahun terakhir, masih diperlukan penelitian lebih lanjut untuk meningkatkan akurasi dan performa mesin penerjemah bahasa Indonesia. Beberapa penelitian telah dilakukan untuk meningkatkan akurasi mesin penerjemah bahasa Indonesia, seperti menggunakan mekanisme *attention* dari Bahdanau dan Luong pada Neural Machine Translation Bahasa Indonesia ke Bahasa Melayu Ketapang dengan Arsitektur Recurrent Neural Network [4]. Selain itu, terdapat penelitian yang melakukan pendekatan statistical machine translation (SMT) pada Bahasa Lampung Dialek Api ke Bahasa Indonesia [5].

Dalam konteks machine translation, khususnya translasi gaya dari bahasa informal ke formal dalam bahasa Indonesia, terdapat tantangan kritis yang harus dihadapi yaitu kurangnya data pelatihan yang memadai untuk melatih model yang mampu melakukan style transfer dengan tingkat akurasi dan konsistensi optimal [6]. Saat ini, dataset yang tersedia seperti dataset STIF parallel masih belum mencukupi kebutuhan pelatihan model secara menyeluruh. Keterbatasan inilah yang menjadi penghalang utama dalam upaya penciptaan model yang mampu menghasilkan terjemahan yang tepat dan sesuai konteks dari bahasa informal ke formal. Maka dari itu, diperlukan langkah konkrit untuk mengatasi masalah ini agar mendapatkan hasil terjemahan yang optimal

Untuk mengatasi kendala yang ada, terdapat usulan metode yang didasarkan pada penelitian sebelumnya dengan menggunakan pendekatan model pretrained MarianMT. Model ini mampu menangani kalimat yang panjang dan menangkap ketergantungan jarak jauh antar kata, yang memungkinkan terjemahan yang lebih koheren

## II. TINJAUAN PUSTAKA

### A. Natural Language Preprocessing

Natural Language Processing (NLP) adalah bidang interdisipliner yang menggabungkan ilmu komputer, kecerdasan buatan, dan linguistik untuk memungkinkan komputer memahami, menafsirkan, dan memanipulasi bahasa manusia [7]. NLP melibatkan pemrosesan kumpulan data bahasa alami, seperti teks atau ucapan, menggunakan pendekatan berbasis aturan atau pembelajaran mesin. Beberapa teknik NLP yang umumnya digunakan seperti analisis sentimen, klasifikasi teks, chatbot, terjemahan bahasa, dan asisten pintar. NLP juga digunakan di berbagai industri untuk mengotomatisasi tugas, mendapatkan wawasan dari data yang tidak terstruktur, dan meningkatkan interaksi pelanggan.

NLP melibatkan konsep linguistik komputasi (CL), yang berfokus pada deskripsi komputasi bahasa sebagai sebuah sistem. Sementara linguistik komputasi (CL) lebih fokus ke analisis linguistik, NLP lebih berorientasi pada penerapan praktis pemrosesan bahasa. Secara keseluruhan, NLP telah diterapkan dalam berbagai teknik dan aplikasi, seperti chatbots yang menggunakan model *deep learning*, sistem pengenalan suara, dan asisten virtual yang semakin memahami konteks dan kebutuhan pengguna. Dengan terus berkembangnya teknologi ini, NLP menjadi pilar utama dalam menciptakan solusi cerdas yang dapat beradaptasi dengan bahasa manusia.

### B. Bahasa Resmi dan tidak resmi

Bahasa resmi adalah bahasa yang digunakan dalam situasi formal, seperti pidato, rapat, surat resmi, dan sebagainya. Bahasa resmi memiliki ciri-ciri seperti menggunakan ejaan baku, menghindari singkatan, menggunakan kata baku, dan mengikuti kaidah tata bahasa yang baku [7]. Bahasa tidak resmi adalah bahasa yang digunakan dalam situasi informal, seperti percakapan sehari-hari, media sosial, pesan singkat, dan sebagainya. Bahasa tidak resmi memiliki ciri-ciri seperti menggunakan ejaan tidak baku, menggunakan singkatan, menggunakan kata tidak baku, dan mengabaikan kaidah tata bahasa yang baku [8]. Bahasa tidak resmi seringkali lebih mudah dipahami dan lebih mengekspresikan perasaan daripada

bahasa resmi. Namun, bahasa tidak resmi juga dapat menimbulkan kesalahpahaman dan kurangnya rasa hormat jika digunakan dalam situasi yang tidak tepat. Oleh karena itu, penggunaan bahasa resmi dan tidak resmi harus disesuaikan dengan konteks dan tujuan komunikasi.

### C. Mesin Penerjemah

Mesin penerjemah merupakan alat penerjemah otomatis yang mampu mengubah teks dari satu bahasa ke bahasa lain tanpa melibatkan manusia. Terjemahan mesin modern tidak hanya menerjemahkan kata demi kata, tetapi juga menganalisis semua elemen teks dan bagaimana kata-kata memengaruhi satu sama lain untuk mengomunikasikan makna keseluruhan dari teks bahasa asli ke dalam bahasa target [9]. Dalam konteks pembuatan mesin penerjemah dari bahasa tidak resmi ke resmi, terdapat berbagai penelitian yang bertujuan untuk mengembangkan aplikasi mesin penerjemah yang mampu mentransformasi teks tidak resmi ke dalam teks resmi. Misalnya, penelitian ini merancang sebuah sistem untuk menerjemahkan teks dari bahasa Indonesia ke dalam teks bahasa Banjar, dengan harapan dapat berkontribusi dalam menjaga kelestarian Bahasa Banjar di Indonesia [9]. Selain itu, strategi dan teknik pengembangan mesin penerjemah berbasis kecerdasan buatan juga telah diterapkan untuk mengatasi masalah terjemahan pragmatis, terkait-konvensi, linguistik, dan khusus teks [10]. Dengan demikian, pengembangan mesin penerjemah dari bahasa tidak resmi ke resmi terus menjadi fokus penelitian yang penting dalam mendukung komunikasi lintas bahasa.

### D. Dictionary Based Translation

Salah satu gagasan dasar adalah dengan sekadar mengubah barisan pada dasar tingkat kata dengan bantuan kamus. Sebagai garis dasar, Penelitian ini mengembangkan sistem penerjemah berbasis kamus sederhana. Model ini menggunakan tataran kata bahasa Indonesia formal-informal kamus yang telah digunakan dalam banyak karya. Sistem ini cukup terjemahkan kata informal ke dalam bentuk formalnya jika itu muncul di kamus [11].

Metode penerjemahan ini didasarkan pada entri kamus bahasa. Padanan kata tersebut digunakan untuk mengembangkan ayat terjemahannya. Generasi pertama dari terjemahan mesin adalah seluruhnya didasarkan pada mesin yang dapat dibaca atau elektronik kamus. Sampai batas tertentu, metode ini masih berlaku membantu dalam penerjemahan frasa tetapi tidak kalimat. Sebagian besar pendekatan penerjemahan dikembangkan kemudian kurang lebih menggunakan kamus bilingual dengan aturan tata bahasa.

### E. Neural Machine Translation

*Neural Machine Translation* (NMT) adalah jenis model penerjemahan mesin yang memanfaatkan jaringan saraf untuk mempelajari pemetaan dari kalimat sumber ke kalimat target [6]. Tidak seperti model penerjemahan mesin tradisional, model NMT tidak bergantung pada aturan yang dibuat dengan tangan atau kamus yang telah ditentukan sebelumnya. Sebaliknya, model ini mempelajari proses penerjemahan dari sejumlah besar data paralel. Hal ini memungkinkan model NMT untuk menangkap pola dan nuansa yang kompleks dalam penerjemahan bahasa.

Model NMT terdiri dari dua komponen utama: encoder dan decoder. Encoder mengambil kalimat sumber sebagai input dan menghasilkan representasi vektor dengan panjang tetap dari kalimat tersebut, menangkap informasi semantik dan kontekstual [12]. Decoder mengambil representasi vektor ini sebagai masukan dan menghasilkan kalimat target kata per kata, menggunakan informasi yang dikodekan untuk memandu proses pembuatan [12]. Dalam encoder dan decoder digunakan arsitektur RNN dengan layer LSTM. LSTM merupakan salah satu jenis arsitektur jaringan saraf tiruan (neural network) yang digunakan untuk memproses data berurutan, seperti teks, suara, dan video. LSTM memiliki struktur yang kompleks dengan tiga gerbang (*gate*) yaitu *forget gate*, *input gate*, dan *output gate*, yang memungkinkan jaringan untuk mengingat informasi penting dan mengabaikan informasi yang tidak relevan [13]. Selama pelatihan, model disajikan dengan pasangan kalimat sumber dan target, dan parameter model disesuaikan untuk meminimalkan perbedaan antara kalimat target yang diprediksi dan yang sebenarnya. Hal ini dicapai melalui backpropagation, di mana kesalahan disebarkan kembali melalui jaringan, dan bobotnya diperbarui untuk meningkatkan akurasi terjemahan model.

Model NMT menawarkan beberapa keunggulan dibandingkan model penerjemahan mesin tradisional. Pertama, model ini mampu menangani kalimat yang panjang dan menangkap ketergantungan jarak jauh antar kata, yang memungkinkan terjemahan yang lebih koheren. Kedua, model NMT dapat belajar menerjemahkan ekspresi idiomatik dan bahasa non-harfiah lainnya, sehingga menghasilkan terjemahan yang lebih akurat dan sesuai dengan konteks. Ketiga, model NMT memiliki kemampuan untuk menghasilkan terjemahan yang lebih lancar dan terdengar alami, sehingga meningkatkan kualitas hasil terjemahan secara keseluruhan.

Terlepas dari kelebihanannya, model NMT juga memiliki beberapa keterbatasan. Pertama, model ini membutuhkan data paralel dalam jumlah besar untuk dilatih secara efektif, yang dapat menjadi tantangan bagi bahasa dengan sumber daya yang terbatas. Kedua, model ini dapat menjadi mahal secara komputasi untuk dilatih dan dijalankan, sehingga membutuhkan sumber daya komputasi yang signifikan. Terakhir, model NMT mungkin mengalami kesulitan dengan kata-kata yang langka atau di luar kosa kata, karena model ini sangat bergantung pada kosa kata yang ada dalam data pelatihan, yang berpotensi menyebabkan ketidakakuratan dalam penerjemahan untuk istilah-istilah yang kurang umum.

#### F. Pretrained Language Modeling

*Pretrained language model* (PLM) merupakan model bahasa yang telah dilatih pada dataset teks yang besar (model yang menggunakan arsitektur Transformer) [14]. Model ini dapat digunakan untuk berbagai tugas NLP, termasuk *machine translation*. Pada *machine translation*, PLM dapat digunakan untuk menerjemahkan teks dari satu bahasa ke bahasa lain. Cara kerjanya adalah dengan mempelajari pola-pola bahasa yang umum terjadi di kedua bahasa tersebut[1]. Pada kasus bahasa Indonesia tidak resmi ke resmi, PLM dapat digunakan untuk mempelajari

pola-pola perbedaan antara bahasa Indonesia informal dan formal. Dengan mempelajari pola-pola ini, PLM dapat menghasilkan terjemahan yang lebih akurat dan sesuai dengan konteks.

Salah satu *pretrained language model* yaitu MarianMT. Marian adalah alat *open-source* untuk pelatihan dan pelayanan terjemahan mesin *neural*, yang sebagian besar dikembangkan di University of Edinburgh, Adam Mickiewicz University di Poznań, dan Microsoft [15]. MarianMT menggunakan teknologi pembelajaran mesin dan kecerdasan buatan untuk memahami dan menerjemahkan teks dalam berbagai konteks dan ragam bahasa.

salah satu model dari MarianMT adalah Helsinki-NLP/opus-mt-id-en, di mana model ini dirancang untuk menerjemahkan bahasa Indonesia ke dalam bahasa Inggris. Model ini merupakan bagian dari proyek OPUS yang mencakup berbagai model penerjemahan mesin berbasis transformer.

Untuk menerjemahkan teks dari bahasa tidak resmi ke resmi, dapat dipakai model MarianMT. Model ini belajar untuk memahami konteks dan nuansa bahasa dari data pelatihan OPUS dan kemudian menerapkan pemahaman tersebut saat menerjemahkan teks baru (informal ke formal dalam dataset STIF Indonesia). Oleh karena itu, model ini dapat digunakan untuk menerjemahkan teks dari bahasa informal ke formal dengan mempertahankan makna asli teks sambil mengubah ragam bahasanya.

### III. DATA DAN METODOLOGI

Dataset yang diperoleh berasal dari platform *Hugging Face* yang telah disediakan oleh akun haryoaw (Haryo Akbarianto Wibowo) dengan judul ‘stif-Indonesia’ (kepanjangan: *Semi-Supervised Low-Resource Style Transfer of Indonesian Informal to Formal Language*).

Dataset ini terdiri dari 1.920 baris yang terdiri dari dua pasang kolom yakni kata informal dan formalnya dalam bahasa Indonesia (*training set*). Sedangkan untuk *testing set* terdapat 363 baris dan 214 baris *dev set* (untuk pengembangan dan pemilihan model).

Pada proses pembangunan model *pre-trained* dengan MarianMT digunakan *training set* untuk pelatihan model *machine translation* atau mesin penerjemah serta *dev set* yang berguna sebagai data validasi untuk mengetahui performa model selama proses pelatihan. Selanjutnya kedua jenis data tersebut akan diproses dalam pembangunan model *fine tuning* dengan MarianMT. Sedangkan *test set* digunakan untuk evaluasi model.

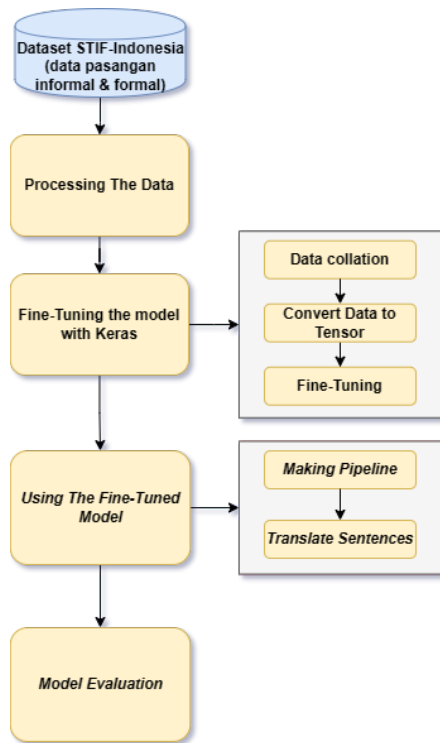


Fig 1. Flowchart Machines Translation with MarianMT

Metodologi dalam penelitian ini ditunjukkan pada gambar di atas, di mana langkah pertama yang dilakukan yakni pemrosesan pada data berupa tokenisasi teks informal dan formal menggunakan *Marian Indonesian to English model* sehingga dihasilkan ID token. Setelah data telah diproses menjadi numerik selanjutnya dilakukan *fine tuning*, proses yang pertama yaitu melakukan *data collation* atau kolektor data untuk mengatasi *padding* sehingga menghasilkan *batch* atau pengelompokan yang dinamis. Setelah selesai melakukan kolektor data, selanjutnya dilakukan pengubahan format menjadi bentuk tensor yang akan digunakan dalam proses pelatihan. Proses *fine tuning* dilakukan dengan pembuatan *optimizer* serta *schedule*, di mana *optimizer* digunakan untuk penyesuaian parameter model berdasarkan kesalahan yang dibuat sedangkan *schedule* merupakan sebuah alat untuk mengatur laju pembelajaran yakni ukuran *step* atau langkah yang diambil oleh *optimizer* dalam mengoptimalkan parameter model (*schedule* dapat menyesuaikan laju *training* berdasarkan *epoch*). Kemudian dilakukan kompilasi atau *compile* untuk menghubungkan antara *optimizer* dengan model, proses penyesuaian sebelum *fine tuning* selesai dan *fine tuning* dijalankan menggunakan dataset pelatihan *STIF-Indonesia*.

Proses selanjutnya yaitu menggunakan model *fine tuning* atau dalam platform transformers kerap disebut *inference* (menghasilkan teks *output* dari teks *input*), proses *inference* pada model *fine tuning* digunakan *pipeline*. *Pipeline* merupakan alat yang dapat menyederhanakan proses memunculkan output berupa hasil terjemahan secara otomatis tanpa harus melakukan pemrosesan pada teks atau data yang akan diketahui terjemahannya. Proses yang terakhir yakni evaluasi model dengan melakukan prediksi atau menerjemahkan kalimat-kalimat informal menjadi bentuk formal pada data validasi dan pengujian, setelah didapatkan hasil terjemahan maka selanjutnya dihitung skor BLEU untuk tiap kalimat dan menghitung rata-rata skor BLEU sebagai ukuran kebaikan akhir model.

#### IV. HASIL DAN PEMBAHASAN

Untuk verifikasi model yang terbaik untuk digunakan dalam *machine translation informal to formal* dalam korpus yang sama, maka model Dictionary-based, Neural Machine Translation, dan Pretrained dibandingkan menurut rata-rata BLEU Score.

##### A. Perbandingan Hasil Model

Gambaran dari hasil terdiri dari skor BLEU dari semua model yang diuji, seperti yang ditunjukkan pada Tabel 1, dan sampel terjemahan Informal to Formal untuk setiap model ditunjukkan pada Tabel 2.

TABLE I. AVERAGE BLEU SCORE

Model	Rata-Rata BLEU
Dictionary Based	41,465
NMT GRU	22,486
NMT LSTM	34,2
Pre-trained Model	41,636633

Pada tabel I, telah disajikan hasil nilai BLEU dari Dictionary Based, NMT GRU, NMT LSTM, dan PLM MarianMT. Dapat dilihat bahwa Pre-trained Model yang telah di *fine tuning* menghasilkan performa terbaik dengan nilai BLEU 41,636633, dilanjutkan dengan pendekatan *Dictionary Base* yang menghasilkan nilai BLEU 41,465. *Pre-trained Model* dan *Dictionary Based* memiliki nilai BLEU yang sama-sama tinggi dengan selisih hanya 0,171, tetapi *Pre-trained Model* lebih unggul dalam hal kualitas terjemahan. Kemudian terdapat dua model yang menggunakan pendekatan NMT, dimana model pertama menggunakan arsitek GRU yang menghasilkan prediksi dengan nilai BLEU 22,486. Sedangkan model kedua menggunakan arsitektur LSTM yang menghasilkan nilai BLEU sebesar 34,2. Perbandingan hasil dari kedua model NMT berbeda signifikan, dimana NMT dengan arsitektur GRU sederhana menghasilkan prediksi yang paling rendah diantara 3 model lainnya.

TABLE II. SAMPLE OF OMISSION ERROR

Model	Contoh 1	Contoh 2
Teks Formal Aktual	admin, tolong dibantu periksa resi xxxnumberxxx, sudah xxxnumberxxx hari tidak ada pergerakan.	maaf tapi saya masih bermasalah lagi .
Dictionary Based	admin, tolong dibantu periksa resi xxxnumberxxx, sudah xxxnumberxxx hari tidak ada pergerakan.	maaf tapi saya <b>masuk</b> bermasalah lagi.
NMT GRU	<b>min</b> tolong dibantu cek resi xxxnumberxxx , sudah xxxnumberxxx hari <b>nih ada masalah</b> ?	maaf tapi saya masih bermasalah ?

NMT LSTM	admin tolong di bantu periksa resi sudah hari tidak ada ada	tapi saya masih bermasalah lagi
Pre-trained Model	admin tolong dibantu <b>cek</b> resi xxxnumberxxx , sudah xxxnumberxxx hari tidak ada pergerakan .	maaf , tapi saya masih bermasalah lagi .

Pada tabel II, telah disajikan perbandingan dua contoh teks formal aktual dengan hasil prediksi dari masing-masing model. Dapat dilihat model NMT GRU menghasilkan prediksi teks formal yang kurang sesuai dengan teks formal aktual, dimana pada hasil prediksi contoh kalimat 1 terdapat banyak kata yang masih belum diterjemahkan ke bentuk formal seperti “min” seharusnya menjadi “admin”, ada kata yang tidak sesuai dengan konteks “ni ada masalah”, pada contoh kalimat 2 terdapat perbedaan tanda baca yang seharusnya “.” tetapi dalam prediksi “?”. Selanjutnya, hasil prediksi model NMT LSTM yang kehilangan beberapa kata pada contoh kalimat 1 dengan kata “xxxnumberxxx” dan contoh kalimat 2 dengan kata “maaf”, terdapat terjemahan yang kurang sesuai, dan terdapat perubahan tanda baca yang tidak sesuai dengan teks formal aktual pada kedua contoh kalimat. Kemudian untuk *Pre-trained Model* terdapat kekurangan seperti kata yang belum formal dan terdapat penambahan tanda baca yang kurang sesuai dengan teks formal aktual.

Pemilihan *Pre-trained Model* sebagai model utama didasari oleh rata-rata BLEU score, dimana model ini memiliki nilai yang paling tinggi ketimbang model *Dictionary Based*, NMT berbasis GRU, dan NMT berbasis LSTM. Selanjutnya akan dibahas dengan lebih detail mengenai performa dari *Pre-trained Model*.

### B. Performa dari Pretrained Model

Pada bagian di atas, ditunjukkan bahwa performa model pre-trained jauh lebih baik daripada yang lain. Sekarang, akan dijelaskan performa model ini secara detail mulai dari *history training*, *validation loss*, skor BLEU serta analisis kesalahan terjemahan.

Untuk Train Loss dan Validation loss pada model pre-trained terlihat pada Fig.2 bahwa pada epoch ke 20, nilai *loss* dari *train* dan *validation* tetap.

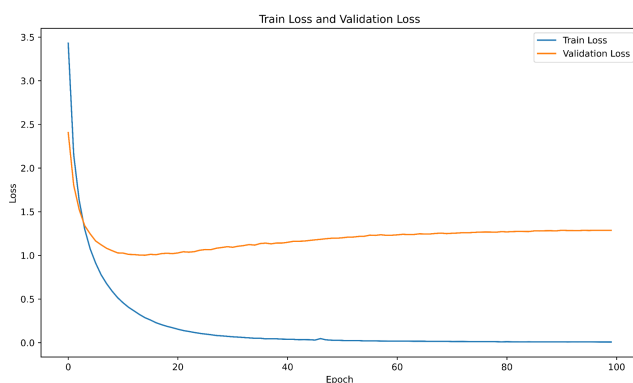


Fig 2. Training dan Validasi Loss

Nilai BLEU pada model pretrained untuk terjemahan informal ke formal memiliki rata-rata 41,636633, untuk nilai minimum BLEU adalah 4,9323516 dan untuk nilai maksimum BLEU adalah 100. Rentang BLEU paling banyak berkisar di antara 37,93235157 sampai 48,93235157 sebanyak 73 data. Pada nilai BLEU yang paling rendah berkisar antara 4,932351569 sampai 15,93235157 sebanyak 48 data. Pada nilai BLEU yang paling tinggi berkisar antara 92,93235157 sampai 103,9323516 sebanyak 9 data.

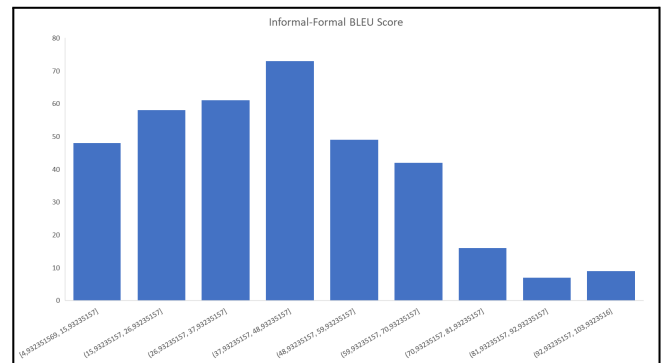


Fig 3. Informal-Formal BLEU Score

Setelah dilakukan analisis terhadap hasil terjemahan pada data pengujian, ditemukan tiga kesalahan meliputi: *omission*, *punctuation*, dan leksikal.

TABLE III. SAMPLE OF OMISSION ERROR

Hasil Terjemahan Model	Label
mengapa dari kemarin saya daftar paket yellow tidak bisa terdaftar ? mohon dibantu .	ini mengapa dari kemarin saya daftar paket yellow tidak bisa terdaftar ? mohon <b>dibantu</b> .
mau daftar lagi juga tidak bisa .	mau daftar ulang lagi juga tidak bisa .

Dalam label yang benar, kalimat pertama dimulai dengan kata “ini” tetapi dalam prediksi model, kata “ini” tidak ada. Begitupun dalam kalimat kedua, kata “ulang” dalam hasil prediksi model tidak ada padahal dalam label yang benar kata tersebut ada. Oleh karena itu, model telah menghilangkan kata tersebut dari terjemahan yang merupakan kesalahan *Omission*. Dalam konteks machine translation, *Omission* dalam bahasa Indonesia berarti penghilangan atau kelalaian. Kesalahan ini merujuk pada situasi di mana beberapa bagian dari teks sumber atau target tidak diterjemahkan oleh model.

TABLE IV. SAMPLE OF PUNCTUATION ERROR

Hasil Terjemahan Model	Label
tolong admin . paket saya belum terkirim dengan resi jp xxxphonexxx .	tolong admin , paket saya belum terkirim dengan resi jp xxxphonexxx .
admin , mengapa indosat jelek terus - sinyalnya hari ini ?	admin , mengapa indosat jelek terus sinyalnya hari ini ?

Dalam label yang benar, setelah kalimat “admin” seharusnya diikuti tanda koma (“,”) namun hasil prediksi model adalah kata titik (“.”). Sedangkan pada kalimat kedua tidak memiliki tanda hubung (“-”), tetapi dalam prediksi

model tanda hubung ditambahkan setelah kata “terus”. Oleh karena itu, model telah menggunakan tanda baca yang tidak perlu di mana kesalahan tersebut termasuk dalam kesalahan *punctuation* (tanda baca).

TABLE V. SAMPLE OF LEXICAL ERROR

Hasil Terjemahan Model	Label
bagaimana ini ? seharian tidak bisa , pulsa terpotong melulu .	bagaimana ini ? seharian tidak bisa , pulsa terpotong terus .
tidak tegapi , takut pengemudinya kena masalah .	tidak berani , takut pengemudinya kena masalah .

Dalam label yang benar, kata “terus” pada kalimat pertama digunakan tetapi dalam prediksi model kata “melulu” digunakan sebagai pengganti. Meskipun “melulu” dan “terus” memiliki makna yang serupa dalam bahasa Indonesia, namun penggunaan kata “melulu” mungkin kurang tepat dalam konteks ini dan terdengar kurang formal. Sedangkan hal yang sama dengan kalimat kedua kata “berani” digunakan, tetapi dalam prediksi model, kata “tegapi” digunakan sebagai pengganti. Kata “tegapi” yang diterjemahkan tidak memiliki arti dalam bahasa Indonesia. Oleh karena itu, model telah menerjemahkan kata yang kurang tepat sehingga termasuk dalam kesalahan leksikal.

## V. Kesimpulan dan Saran

Dalam penelitian yang telah dilakukan, dapat ditarik kesimpulan bahwa dengan menggunakan pendekatan kamus (*dictionary*) serta *deep learning* dapat menghasilkan model prediksi untuk menerjemahkan bentuk kalimat Indonesia informal menjadi kalimat formal. Dari keempat model yang dibangun, didapatkan model yang terbaik adalah model Pre-Trained dengan menggunakan MarianMT. Model ini memiliki kelebihan yaitu cenderung mudah dalam proses *inferences* karena menggunakan metode *pipeline*, kemudian karena model ini merupakan hasil *fine-tuning* dari Transformer yang sudah dilatih dengan banyak data atau korpus sehingga proses pelatihan dan pengenalan pola data pasangan bahasa Indonesia informal formal tidak dimulai dari nol karena memiliki informasi dari model Transformer (MarianMT) sebelumnya. Model ini menghasilkan rata-rata skor BLEU yang lebih besar dibandingkan model lain yaitu sebesar 41,6366. Model ini memiliki kelemahan, seperti tanda baca sering dihilangkan atau diganti dengan yang tidak sesuai dengan *labels*, terdapat kata yang tidak diterjemahkan dari terjemahan *labels* lalu terdapat beberapa prediksi atau terjemahan yang masih menghasilkan kata dengan nuansa bahasa Indonesia informal.

Saran atau rekomendasi yang dapat dilakukan untuk pengembangan selanjutnya adalah Pengumpulan dataset yang lebih besar dengan konteks kata yang lebih luas, Membuat strategi *fine-tuning* untuk Model *Machine Translation* yang lebih kuat dengan menggunakan beberapa teknik regularisasi. Serta melakukan kombinasi model dan tahapan alur menerjemahkan yang lebih kompleks agar hasil yang didapatkan bisa lebih bagus.

## REFERENCES

- [1] D. W. Otter, J. R. Medina, and J. K. Kalita, “A Survey of the Usages of Deep Learning for Natural Language Processing,” *IEEE Trans. NEURAL Netw. Learn. Syst.*, vol. 32, no. 2, 2021.
- [2] H. Wang, “Progress in Machine Translation,” 2022.
- [3] M. Maryamah, A. Z. Arifin, R. Sarno, and A. M. Hasan, “Adapting Google Translate using Dictionary and Word Embedding for Arabic-Indonesian Cross-lingual Information Retrieval,” in *2020 IEEE International Conference on Internet of Things and Intelligence System (IoT&IS)*, Jan. 2021, pp. 205–209. doi: 10.1109/IoT&IS50849.2021.9359708.
- [4] W. Gunawan, H. Sujaini, and T. Tursina, “Analisis Perbandingan Nilai Akurasi Mekanisme Attention Bahdanau dan Luong pada Neural Machine Translation Bahasa Indonesia ke Bahasa Melayu Ketapang dengan Arsitektur Recurrent Neural Network,” *J. Edukasi Dan Penelit. Inform. JEPIN*, vol. 7, no. 3, p. 488, Dec. 2021, doi: 10.26418/jp.v7i3.50287.
- [5] P. Permata and Z. Abidin, “Statistical Machine Translation Pada Bahasa Lampung Dialek Api Ke Bahasa Indonesia,” *J. MEDIA Inform. BUDIDARMA*, vol. 4, no. 3, p. 519, Jul. 2020, doi: 10.30865/mib.v4i3.2116.
- [6] B. R. Irawan and R. Adi, “Improving Indonesian Informal to Formal Style Transfer via Pre-Training Unlabelled Augmented Data,” in *2023 6th International Conference of Computer and Informatics Engineering (IC2IE)*, Lombok, Indonesia: IEEE, Sep. 2023, pp. 25–29. doi: 10.1109/IC2IE60547.2023.10331336.
- [7] S. L. Lo, E. Cambria, R. Chiong, and D. Cornforth, “Multilingual sentiment analysis: from formal to informal and scarce resource languages,” *Artif. Intell. Rev.*, vol. 48, no. 4, pp. 499–527, Dec. 2017, doi: 10.1007/s10462-016-9508-4.
- [8] “Penggunaan Bahasa Indonesia dan Bahasa Gaul di Kalangan Remaja | Jurnal Multidisiplin Dehasen (MUDE).” Accessed: Dec. 28, 2023. [Online]. Available: <https://jurnal.unived.ac.id/index.php/mude/article/view/2339>
- [9] “The Journal of Universitas Negeri Surabaya,” Accessed: Dec. 13, 2023. [Online]. Available: <https://ejournal.unesa.ac.id>
- [10] W. Untara and T. Setiawan, “PROBLEMA MESIN PENERJEMAH BERBASIS AI DALAM PROSES PENERJEMAHAN BUKU INGGRIS-INDONESIA DAN SOLUSINYA,” *Adab. J. Bhs. Dan Sastra*, vol. 4, no. 1, p. 92, Jun. 2020, doi: 10.14421/ajbs.2020.04105.
- [11] H. A. Wibowo *et al.*, “Semi-Supervised Low-Resource Style Transfer of Indonesian Informal to Formal Language with Iterative Forward-Translation,” arXiv, Dec. 22, 2020. Accessed: Dec. 27, 2023. [Online]. Available: <http://arxiv.org/abs/2011.03286>
- [12] Y. Wu, Y. Wang, and S. Liu, “A Dataset for Low-Resource Stylized Sequence-to-Sequence Generation,” *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 05, pp. 9290–9297, Apr. 2020, doi: 10.1609/aaai.v34i05.6468.
- [13] N. Rezaputra and Y. D. Prabowo, “Alih Bentuk Kalimat Non-Formal Menjadi Kalimat Formal Menggunakan Pendekatan Machine Translation,” vol. 8, no. 1, 2022.
- [14] J. Li, T. Tang, W. X. Zhao, J.-Y. Nie, and J.-R. Wen, “Pretrained Language Models for Text Generation: A Survey,” arXiv, May 13, 2022. Accessed: Dec. 25, 2023. [Online]. Available: <http://arxiv.org/abs/2201.05273>
- [15] M. Junczys-Dowmunt *et al.*, “Marian: Fast Neural Machine Translation in C++,” arXiv, Apr. 04, 2018. Accessed: Dec. 25, 2023. [Online]. Available: <http://arxiv.org/abs/1804.00344>