JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9465
Hal 11-18
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Optimasi Hyperparameter Optuna Pada Model mT5 Untuk Penerjemahan Angkola-Indonesia

Awal Ridho Harahap, Hanafi*

Fakultas Ilmu Komputer, PJJ Informatika, Universitas Amikom Yogyakarta, Yogyakarta, Indonesia
Email: [1] ridhoharahap@students.amikom.ac.id, [2,*]hanafi@amikom.ac.id
Email Penulis Korespondensi: hanafi@amikom.ac.id
Submitted **03-01-2026;** Accepted **12-01-2026;** Published **19-02-2026**

Abstrak

Penelitian ini bertujuan untuk mengatasi tantangan pelestarian Bahasa Angkola di era digital akibat ketiadaan korpus data digital yang memadai , melalui pengembangan sistem penerjemahan mesin otomatis Angkola ke Indonesia yang akurat dan efisien. Metode yang diajukan berfokus pada pendekatan fine-tuning model Multilingual Text-to-Text Transfer Transformer (mT5-base) menggunakan korpus data teks Angkola-Indonesia. Dataset awal yang terdiri dari pasangan kalimat Angkola-Indonesia dibersihkan dan menghasilkan 28.775 pasangan kalimat yang digunakan untuk pelatihan. Data kemudian dibagi menjadi 70% data latih (20.142 baris), 15% data validasi (4.316 baris), dan 15% data uji (4.317 baris). Optimasi performa model dilakukan secara cerdas menggunakan Optuna Hyperparameter Tuning untuk menemukan kombinasi hyperparameter terbaik. Fungsi objektif Optuna dirancang untuk memaksimalkan skor gabungan yang didasarkan pada metrik BLEU dan chrF dari hasil evaluasi validasi. Proses optimasi menghasilkan Trial terbaik (Trial 50) dengan hyperparameter kunci: learning rate = 0.0004316 dan num beams 4. Model terbaik yang diperoleh dari proses fine-tuning kemudian dievaluasi pada set data Uji (Test) yang terpisah. Evaluasi akhir pada data uji menggunakan metrik standar penerjemahan menghasilkan kinerja yang sangat baik, dengan skor BLEU = 73.84 dan chrF = 83.34. Secara keseluruhan, penelitian ini berhasil mengimplementasikan optimasi hyperparameter menggunakan Optuna untuk model mT5, menghasilkan model penerjemahan Angkola-Indonesia yang menunjukkan akurasi tinggi dan kinerja yang lebih efisien. Hasil ini memberikan kontribusi nyata bagi pelestarian Bahasa Angkola dengan menyediakan alat terjemahan yang modern dan akurat.

Kata Kunci: Penerjemahan Mesin; Bahasa Angkola; mT5; Optuna; Hyperparameter Tuning; BLEU Score; chrF Score

Abstract

This research aims to address the challenges of preserving the Angkola language in the digital era, which are exacerbated by the lack of an adequate digital data corpus, by developing an accurate and efficient automatic Angkola-to-Indonesian machine translation system. The proposed method focuses on a fine-tuning approach for the Multilingual Text-to-Text Transfer Transformer (mT5-base) model using an Angkola-Indonesian text data corpus.The initial dataset, consisting of Angkola-Indonesian sentence pairs, was cleaned, resulting in 28,775 sentence pairs used for training. The data was subsequently split into 70% training data (20,142 lines), 15% validation data (4,316 lines), and 15% test data (4,317 lines). Intelligent model performance optimization was conducted using Optuna Hyperparameter Tuning to find the best hyperparameter combination. Optuna's objective function was designed to maximize a composite score based on the BLEU and chrF metrics from the validation evaluation results. The optimization process yielded the best Trial (Trial 50) with key hyperparameters: learning rate = 0.0004316 and num beams = 4. The best model obtained from the fine-tuning process was then evaluated on a separate Test dataset. The final evaluation on the test data using standard translation metrics demonstrated excellent performance, achieving a BLEU score of 73.84 and a chrF score of 83.34. Overall, this research successfully implemented hyperparameter optimization using Optuna for the mT5 model, resulting in an Angkola-to-Indonesian translation model that exhibits high accuracy and more efficient performance. These results provide a tangible contribution to the preservation of the Angkola language by offering a modern and accurate translation tool.

Keywords: Machine Translation; Angkola Language; mT5; Optuna; Hyperparameter Tuning; BLEU Score; chrF Score

1. PENDAHULUAN

Kemajuan pesat dalam bidang kecerdasan buatan (Artificial Intelligence) dan pemrosesan bahasa alami (Natural Language Processing/NLP) telah merevolusi industri penerjemahan. Adopsi Neural Machine Translation (NMT) berbasis Transformer telah terbukti meningkatkan akurasi dan efisiensi dalam menangkap nuansa semantik antarbahasa secara signifikan[1][2]. Namun, dominasi teknologi ini masih terpusat pada bahasa-bahasa global dengan sumber daya melimpah (high-resource languages). Tantangan praktis dan kesenjangan teknologi masih sangat terasa pada pengembangan sistem terjemahan yang inklusif bagi bahasa daerah dengan sumber daya terbatas (low-resource languages), yang sering kali terabaikan dalam diskursus teknologi global[1].

Batak Angkola, sebagai entitas bahasa dan budaya yang unik di Tapanuli Selatan, menghadapi ancaman digitalisasi yang serius. Berbeda dengan kerabat etnisnya (Toba dan Mandailing), Bahasa Batak Angkola memiliki karakteristik morfologi yang sangat kaya dan kompleks. Studi terdahulu mendokumentasikan adanya dua belas pola perubahan fonem serta keragaman afiks (prefiks, sufiks, konfiks) yang rumit dalam komunikasi sehari-hari[6]. Namun, kompleksitas linguistik ini justru berbenturan dengan kendala utama: ketiadaan korpus digital yang memadai dan terstruktur. Kelangkaan data teks paralel ini menjadi hambatan terbesar dalam melatih model AI yang akurat, menjadikan bahasa ini rentan punah di era digital karena minimnya representasi teknolog[3][4][5].

Penelitian terkait komputasi Bahasa Angkola sejauh ini masih sangat terbatas. Upaya yang dilakukan oleh Hrp dkk., misalnya, baru berfokus pada pendekatan berbasis aturan (*rule-based*) seperti algoritma *stemming* dengan akurasi tinggi[7]. Namun, belum ada penelitian yang melangkah lebih jauh ke tahap penerjemahan mesin otomatis (machine translation) berbasis Deep Learning untuk bahasa ini. Pendekatan berbasis aturan tradisional sering kali gagal menangkap

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9465
Hal 11-18
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

konteks semantik yang luas, sementara pendekatan Deep Learning standar sering kali performanya buruk pada data yang sedikit (data scarcity). Di sinilah letak kesenjangan riset yang krusial: bagaimana membangun sistem penerjemahan yang handal untuk bahasa dengan morfologi kompleks namun minim data seperti Angkola?.

Untuk menjembatani kesenjangan tersebut, pendekatan Transfer Learning menggunakan model pre-trained multibahasa menjadi solusi potensial. Model mT5 (Multilingual T5), yang dilatih pada 101 bahasa, telah menunjukkan kemampuan generalisasi yang kuat pada kasus low-resource dan dialek tertentu, seperti pada bahasa Jerman Swiss dan Hinglish, di mana pendekatan kurikulum ganda mampu meningkatkan skor BLEU secara drastis[8][9][10][11]. Meskipun demikian, penerapan mT5 secara langsung sering kali tidak optimal tanpa penyesuaian parameter yang presisi, terutama pada arsitektur encoder-decoder[12][13]. Kebanyakan penelitian hanya menggunakan parameter default, padahal setiap bahasa low-resource memiliki karakteristik unik yang membutuhkan konfigurasi hyperparameter spesifik.

Berdasarkan kondisi tersebut, penelitian ini mengisi kekosongan riset dengan mengusulkan integrasi model mT5 yang dioptimasi menggunakan Optuna. Optuna dipilih karena fleksibilitas dan efisiensinya dalam mencari konfigurasi hyperparameter terbaik di ruang pencarian yang kompleks, yang sering kali sulit dilakukan secara manual[14]. Penelitian ini tidak hanya bertujuan memperkaya khazanah leksikon digital Angkola-Indonesia, tetapi secara spesifik membuktikan bahwa kombinasi Transfer Learning (mT5) dengan Automated Hyperparameter Optimization (Optuna) dapat mengatasi masalah kelangkaan data dan kompleksitas morfologi pada bahasa daerah. Ini merupakan langkah pionir dalam pelestarian Bahasa Batak Angkola melalui teknologi mutakhir.

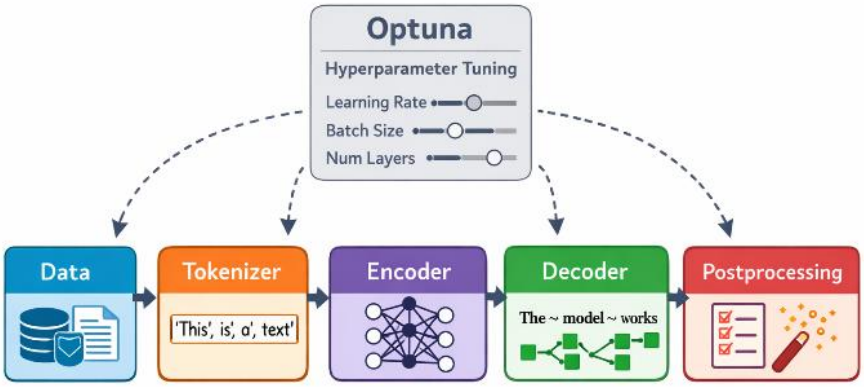
2. METODOLOGI PENELITIAN

2.1 Pendekatan Penelitian

Penelitian ini adalah penelitian eksperimental kuantitatif yang bertujuan mengembangkan dan mengevaluasi sistem penerjemahan Bahasa Angkola → Bahasa Indonesia. Pendekatan eksperimen meliputi: konstruksi korpus paralel, praproses teks, fine-tuning model pretrained mT5, optimasi hyperparameter otomatis menggunakan Optuna, dan evaluasi kuantitatif (BLEU, CHRF) serta analisis kualitatif hasil terjemahan[9].

2.2 Alur Kerja Penelitian (Pipeline)

Secara ringkas penelitian mengikuti alur: Pengumpulan korpus paralel Angkola–Indonesia, Prapemrosesan (cleaning, normalisasi, filtering), Tokenisasi subword (SentencePiece mT5 tokenizer), Fine-tuning mT5 (encoder–decoder) - baseline, Optimasi hyperparameter dengan Optuna (TPE), Evaluasi (validation/test) menggunakan BLEU & CHRF, Analisis kuantitatif & kualitatif. Gambar 1 ini mengilustrasikan alur kerja (workflow) dari sebuah model Machine Learning khususnya model pemrosesan bahasa alami atau NLP (seperti arsitektur Transformer) yang sedang dioptimalkan menggunakan Optuna.



Gambar 1. Alur Kerja Penelitian

2.3 Dataset

Sumber dataset diperoleh dari kamus Angkola–Indonesia Edisi II, kurasi manual, dan penyelarasan pasangan kalimat (sentence alignment). Dengan format pasangan paralel (source_target) dalam file CSV, tiap baris berisi satu pasangan kalimat. Tabel 1 berikut menjelaskan tentang sumber dan pembagian dataset untuk proyek Machine Learning, kemungkinan besar untuk tugas penerjemahan bahasa (Angkola ke Indonesia).

Tabel 1. Distribusi dataset

Subset	Jumlah Kalimat
Training	22.516
Validation	4.824
Testing	4.824
Total	**32.164**

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9465
Hal 11-18
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

2.4 Prapemrosesan Teks

Langkah-langkah prapemrosesan yang diterapkan:
a. Normalisasi Unicode & case - mengganti karakter non-standar, konsisten pada lowercasing bila perlu.
b. Pembersihan karakter - membuang simbol non-leksikal, markup, dan noise.
c. Penghapusan duplikasi - menghapus pasangan identik.
d. Filter panjang kalimat - mengeliminasi kalimat dengan token di bawah 2 atau melebihi max_length.
e. Penanganan token khusus - menjaga ekspresi numerik, tanda baca penting, singkatan lokal sesuai aturan linguistik Angkola.

Prapemrosesan tersebut meningkatkan kualitas data input sehingga stabilitas pelatihan model bertambah [15].

2.5 Tokenisasi Dan Representasi

Bahasa Angkola memiliki morfologi yang produktif (banyak afiks). Tokenisasi subword mengurangi OOV dan menangkap morfem secara lebih efisien. Pendekatan ini direkomendasikan untuk bahasa bermorfologi kaya[16]. Tokenizer: SentencePiece (unigram or BPE) seperti yang digunakan pada pretraining mT5. Parameter: vocab_size (mis. 32k), model_type ('unigram' atau 'bpe'), max_length (mis. 128/256), pad_token, bos_token, eos_token. Proses: kalimat sumber dan target masing-masing di-encode menjadi ID token:

$$X = (x_1, x_2, \dots, x_n) \xrightarrow{Tokenizer} (t_1, t_2, \dots, t_m) \quad (1)$$

dengan padding/truncation:

$$\hat{X} = \begin{cases} X \| \langle pad \rangle, & \text{jika } m < L \\ X_1 : L, & \text{jika } m > L \end{cases} \quad (2)$$

Referensi implementasi: Kudo & Richardson (SentencePiece)[16]
Dimana :
X : satu kalimat input (Bahasa Angkola)
xi : token/kata ke-i sebelum tokenisasi
n : jumlah token awal
Tokenizer : fungsi tokenisasi (SentencePiece milik mT5)
tj : token subword hasil tokenisasi
m : jumlah token setelah tokenisasi (biasanya $m \neq n$)
Kalimat bahasa alami direpresentasikan sebagai urutan token awal, kemudian dipetakan oleh tokenizer menjadi urutan token subword yang digunakan model.

2.6 mT5

mT5 adalah versi multilingual T5 (text-to-text) yang dilatih pada Common Crawl multibahasa. Ia memakai encoder–decoder Transformer standar: encoder menghasilkan representasi H dari input, decoder memproduksi token target autoregresif[9]. Mekanisme attenion dasar adalah:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{dk}}\right)V \quad (3)$$

Decoder menghasilkan probabilitas token berikutnya:

$$P(Yt \mid y < t, X) = softmax\,(Woht + b) \quad (4)$$

Loss yang dioptimalkan adalah cross-entropy:

$$L = -\sum_{t=1}^{T} log\, P(Yt \mid Y < t, X) \quad (5)$$

Detail arsitektur dan alasan memilih mT5

Konfigurasi Pelatihan Baseline pada Tabel 2, sebagai standar pembanding, penelitian ini menggunakan model google/mt5-base yang dilatih selama 6 epoch dengan parameter optimizer AdamW, learning rate 5e-5, dan batch size 4 tanpa menggunakan label smoothing.

Tabel 2. Konfigurasi Baseline yang Dipakai untuk Perbandingan

Parameter	Nilai (baseline)
Model checkpoint	google/mt5-base
Optimizer	AdamW
Learning rate	5e-5
Batch size	4
Epochs	6

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9465
Hal 11-18
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Max length	128
Scheduler	Linear warmup
Label smoothing	0.0

Model mT5 dibangun di atas arsitektur encoder-decoder standar Transformer, namun keunggulan utamanya terletak pada pendekatan revolusionernya terhadap pemrosesan multibahasa. Fondasi kemampuan ini berasal dari pelatihan berskala besar menggunakan korpus mC4, sebuah varian multibahasa dari dataset C4. Dengan mengumpulkan teks dari beragam bahasa, mT5 mampu mendalami representasi linguistik yang kaya dan lintas budaya. Proses pelatihannya memanfaatkan tujuan pemodelan bahasa tersembunyi (masked language modeling), di mana sebagian token masukan ditutup dan model dilatih untuk memprediksi token-token tersebut secara akurat.

Secara operasional, mT5 menyeragamkan seluruh tugas pemrosesan bahasa melalui format "Teks-ke-Teks" yang universal. Dalam paradigma ini, tugas-tugas yang berbeda seperti penerjemahan mesin, peringkasan, penjawaban pertanyaan, hingga klasifikasi teks, semuanya dirumuskan sebagai masalah input teks dan output teks. Sebagai ilustrasi dalam kasus penerjemahan, sistem menerima masukan dengan format instruksi spesifik seperti "terjemahkan bahasa Inggris ke bahasa Indonesia: [teks Inggris]" dan kemudian menghasilkan keluaran berupa teks sasaran yang sesuai.

Keunggulan arsitektur ini memfasilitasi kemampuan transfer pembelajaran lintas bahasa (cross-lingual transfer learning). Latihan pada kumpulan data yang luas memungkinkan model untuk mentransfer pemahaman antarbahasa, di mana wawasan yang diperoleh dari satu bahasa dapat mendongkrak kinerja pada bahasa lain. Fitur ini sangat krusial, terutama dalam meningkatkan kualitas pemrosesan untuk bahasa-bahasa serumpun atau bahasa yang memiliki sumber daya data terbatas (low-resource languages).

Berkat fleksibilitas tersebut, mT5 memiliki spektrum penerapan yang sangat luas di berbagai sektor teknik bahasa. Implementasinya mencakup penerjemahan mesin otomatis, peringkasan dokumen multibahasa, serta sistem tanya-jawab (question answering) yang mampu merespons pertanyaan lintas bahasa. Lebih jauh lagi, model ini mampu melakukan klasifikasi teks—seperti penentuan kategori tema atau analisis sentimen—tanpa memandang bahasa asalnya, serta mendukung pembuatan konten kreatif multibahasa untuk kebutuhan pemasaran dan media sosial.

2.7 Optimasi Hyperparameter (Optuna)

Optuna adalah framework HPO (hyperparameter optimization) yang menggunakan pendekatan define-by-run dan TPE (Tree-structured Parzen Estimator) untuk Bayesian optimization; cocok untuk ruang parameter yang heterogen dan mahal dievaluasi[17]. Tabel 3 mendefinisikan "Ruang Pencarian" (Search Space) yang komprehensif untuk proses optimasi hyperparameter menggunakan framework Optuna dengan pendekatan Tree-structured Parzen Estimator (TPE). Dalam skema ini, tujuh parameter kunci dievaluasi secara dinamis untuk menemukan konfigurasi terbaik, mencakup learning rate yang dieksplorasi dengan distribusi log-uniform (1e-6 hingga 5e-4), variasi batch size (2, 4, 8), dan jumlah epoch (4, 5, 6, 8), serta parameter regularisasi seperti dropout dan label smoothing dengan rentang hingga 0.3 dan 0.2. Selain itu, tabel ini juga menyertakan opsi untuk warmup ratio dan tipe scheduler (linear atau cosine), menegaskan bahwa strategi optimasi ini dirancang untuk menangani ruang parameter yang heterogen guna memaksimalkan performa model di luar konfigurasi standar.

Tabel 3. Ruang Pencarian

Hyperparameter	**Rentang / Pilihan**
Learning rate	[1e-6, 5e-4] (log-uniform)
Batch size	{2,4,8}
Epochs	{4,5,6,8}
Dropout	[0.0, 0.3]
Label smoothing	[0.0, 0.2]
Warmup ratio	[0.0, 0.2]
Scheduler type	{linear, cosine}
Hyperparameter	**Rentang / Pilihan**

Objective function (mengoptimalkan nilai gabungan BLEU/CHRF di validation set):

$$f(\theta) = \alpha \cdot BLEU_{val}(\theta) + \beta \cdot CHRF_{val}(\theta) \tag{6}$$

dengan biasanya $\alpha = 1.0$ dan $\beta = 0.1$ (bisa disesuaikan). Optimasi mencari:

$$\theta * = \arg\max_{\theta \in \Theta} f(\theta) \tag{7}$$

Referensi review HPO terkini dan praktik[18]

2.8 Evaluasi

Untuk mengevaluasi kualitas terjemahan mesin secara kuantitatif, penelitian ini menggunakan metrik standar Bilingual Evaluation Understudy (BLEU). Metrik ini mengukur tingkat kesesuaian antara teks hasil prediksi model (candidate) dengan terjemahan acuan manusia (reference) berdasarkan presisi n-gram. Namun, sekadar menghitung presisi dapat menjadi bias jika model menghasilkan kalimat yang sangat pendek. Oleh karena itu, algoritma ini menerapkan faktor

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9465
Hal 11-18
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Brevity Penalty (BP) untuk menghukum hasil terjemahan yang panjangnya kurang dari referensi, memastikan keseimbangan antara akurasi kata dan kelengkapan kalimat. Perhitungan skor BLEU [19] dan penentuan nilai penalti BP [20] didefinisikan secara matematis sebagai berikut:

$$BLEU = BP \cdot \exp\left(\sum_{n=1}^{N} Wn \log Pn\right) \tag{8}$$

dengan

$$BP = \begin{cases} 1, & c > r \\ e^{\left(1-\frac{r}{c}\right)}, & c \leq r \end{cases} \tag{9}$$

dimana Pn adalah modified precision untuk n-gram
chrF : character n-gram F-score, lebih sensitif pada pilihan morfem dan variasi morfologi. Rumus (F-score):

$$CHRF = \frac{(1+\beta 2) \cdot Pchar \cdot Rchar}{\beta 2 \cdot Pchar + Rchar} \tag{10}$$

3. HASIL DAN PEMBAHASAN

3.1 Gambaran Umum Eksperimen

Bagian ini memaparkan hasil eksperimen penerjemahan otomatis Bahasa Angkola–Bahasa Indonesia yang dikembangkan menggunakan model mT5 melalui pendekatan fine-tuning serta optimasi hyperparameter berbasis Optuna. Rangkaian eksperimen dilaksanakan secara sistematis yang diawali dengan pelatihan model baseline, dilanjutkan dengan proses optimasi menggunakan Optuna, dan diakhiri dengan pelatihan ulang model menggunakan konfigurasi parameter terbaik yang diperoleh. Kinerja model kemudian dievaluasi secara komprehensif pada data validasi dan data uji, mencakup pengukuran kuantitatif menggunakan metrik BLEU dan CHRF serta analisis kualitatif melalui pemeriksaan sampel hasil terjemahan.

3.2 Hasil Prapemrosesan dan Distribusi Dataset

Dari volume awal sebanyak 32.164 data mentah, proses pembersihan ini menghasilkan korpus final bersih sejumlah 28.775 pasangan kalimat. Tabel 4 merinci struktur akhir dataset yang digunakan dalam eksperimen setelah melalui tahapan prapemrosesan yang ketat, meliputi normalisasi teks, eliminasi duplikasi, dan penghapusan baris kosong. Dataset tersebut kemudian didistribusikan menggunakan skema rasio 70:15:15 menjadi tiga subset terpisah: data latih (training) sebanyak 20.142 kalimat, serta data validasi dan data uji masing-masing sebesar 4.316 dan 4.317 kalimat, sebuah strategi pembagian yang dirancang untuk menjamin objektivitas evaluasi dan mencegah terjadinya kebocoran data (data leakage).

Tabel 4. Distribusi Dataset Setelah Prapemrosesan

Subset	Jumlah Kalimat
Training	20.142
Validation	4.316
Testing	4.317
Total	**28.775**

3.3 Hasil Pelatihan Model Baseline mT5

Nilai BLEU yang masih rendah mengindikasikan bahwa konfigurasi default mT5 belum mampu memanfaatkan potensi penuh model dalam menerjemahkan Bahasa Angkola yang memiliki kompleksitas morfologi tinggi. Tabel 5 merangkum kinerja awal model pada fase validasi yang dilatih menggunakan konfigurasi standar tanpa intervensi optimasi hyperparameter. Berdasarkan hasil evaluasi, model baseline ini mencatatkan skor BLEU sebesar ≠ 20,16 dan skor CHRF pada kisaran ≠ 39–41, dengan capaian validation loss di angka ≠ 4,45. Meskipun tren penurunan loss selama pelatihan mengindikasikan adanya proses pembelajaran, capaian skor metrik evaluasi yang relatif rendah menegaskan bahwa penggunaan parameter default belum cukup optimal untuk menangkap kompleksitas penerjemahan, sehingga diperlukan strategi optimasi lebih lanjut untuk mendongkrak akurasi model.

Tabel 5. Hasil Evaluasi Model Baseline (Validasi)

Metrik	Nilai
BLEU	±20,16
CHRF	±39–41
Validation Loss	±4,45

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9465
Hal 11-18
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

3.4 Hasil Optimasi Hyperparameter Menggunakan Optuna

Dalam proses pencarian yang melibatkan 50 trial ini, algoritma berhasil menemukan kombinasi parameter terbaik yang berbeda signifikan dari baseline, di antaranya learning rate sebesar 0.00043, peningkatan batch size menjadi 16, serta penerapan regularisasi yang lebih presisi seperti label smoothing di angka 0.1254 dan dropout sebesar 0.0856. Tabel 6 merangkum konfigurasi hyperparameter optimal yang berhasil diidentifikasi setelah melalui serangkaian eksperimen intensif menggunakan framework Optuna dengan metode Tree-structured Parzen Estimator (TPE). Konfigurasi terpilih ini, yang juga mencakup penggunaan scheduler tipe Linear dan weight decay 0.0139, selanjutnya dijadikan acuan utama untuk pelatihan ulang model guna mencapai kinerja penerjemahan yang maksimal.

Tabel 6. Hyperparameter Terbaik Hasil Optuna

Hyperparameter	Nilai Terbaik
Learning Rate	0.000431687
Batch Size	16
Epoch	6
Label Smoothing	0.1254
Warmup Ratio	0.0346
Weight Decay	0.0139
Scheduler	Linear
Dropout	0.0856
Attention Dropout	0.1592
Num Beams	4

3.5 Evaluasi Model Teroptimasi

Hasil evaluasi kinerja model final yang telah dioptimasi menggunakan parameter terbaik pada set data validasi. Tabel 7 merangkum hasil pengujian menunjukkan lonjakan performa yang sangat signifikan, di mana model berhasil mencapai skor BLEU sebesar 70,58 dan CHRF 81,67, disertai dengan penurunan validation loss yang tajam menjadi 2,65. Peningkatan skor BLEU yang melampaui 50 poin dibandingkan baseline ini secara empiris membuktikan bahwa intervensi optimasi hyperparameter berperan krusial dalam menghasilkan stabilitas model dan kualitas penerjemahan yang jauh lebih superior.

Tabel 7. Hasil Evaluasi Model Teroptimasi (Validasi)

Metrik	Nilai
BLEU	70,58
CHRF	81,67
Validation Loss	2,65

3.6 Hasil Evaluasi pada Data Uji (Test Set)

Pada tahap pengujian ini, model menunjukkan kemampuan generalisasi yang impresif dengan mencatatkan skor BLEU sebesar 73,84 dan CHRF 83,34, serta tingkat validation loss terendah di angka 2,21. Fakta bahwa skor evaluasi pada data uji ini justru melampaui hasil pada fase validasi mengonfirmasi secara empiris bahwa model tidak mengalami overfitting dan mampu menerjemahkan kalimat baru dengan tingkat akurasi yang konsisten serta reliabel. Tabel 8 menyajikan puncak dari rangkaian eksperimen, yaitu evaluasi performa model final pada dataset uji (testing set) yang sepenuhnya terisolasi dari proses pelatihan maupun optimasi.

Tabel 8. Hasil Evaluasi Model pada Data Uji

Metrik	Nilai
BLEU	73,84
CHRF	83,34
Validation Loss	2,21

3.7 Analisis Perbandingan Baseline vs Model Teroptimasi

Sementara konfigurasi baseline dengan parameter standar hanya mampu menghasilkan skor BLEU sekitar 20,16 dan CHRF pada kisaran 40, penerapan optimasi otomatis berhasil mendongkrak akurasi penerjemahan secara drastis hingga mencapai skor BLEU 73,84 dan CHRF 83,34. Kesenjangan skor yang signifikan ini secara empiris memvalidasi sensitivitas tinggi arsitektur mT5 terhadap konfigurasi hyperparameter, sekaligus menegaskan superioritas metode optimasi otomatis dalam mengeksplorasi potensi maksimal model dibandingkan dengan pengaturan manual. Tabel 9 menyajikan perbandingan komprehensif antara kinerja model baseline dan model mT5 yang telah dioptimasi menggunakan Optuna, menyoroti disparitas performa yang sangat tajam di antara kedua pendekatan tersebut.

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9465
Hal 11-18
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Tabel 9. Perbandingan Kinerja Model

Model	BLEU	CHRF
mT5 Baseline	~20,16	~40
mT5 + Optuna	73,84	83,34

3.8 Analisis Kualitatif Hasil Terjemahan

Berdasarkan sampel yang dievaluasi, seperti penerjemahan ungkapan kultural "Horas ma di hita sude..." yang dialihbahasakan secara tepat menjadi "Salam untuk kita semua...", model terbukti berhasil mempertahankan struktur kalimat yang alami sekaligus menangkap makna semantik dan kontekstual secara akurat. Meskipun kinerja umum menunjukkan hasil yang positif, evaluasi ini juga mengidentifikasi adanya residu kesalahan leksikal minor, terutama pada kosakata idiomatik, yang diatribusikan pada keterbatasan variasi data latih yang tersedia. Tabel 10 menyajikan analisis kualitatif terhadap kapabilitas penerjemahan model melalui komparasi langsung antara kalimat sumber Bahasa Angkola dan luaran model dalam Bahasa Indonesia.

Tabel 10. Contoh Hasil Terjemahan

Bahasa Angkola	Terjemahan Model
Horas ma di hita sude…	Salam untuk kita semua…
Piga jam do tarida ho tu jolo?	Jam berapa kamu akan berangkat?

3.9 Pembahasan

Hasil penelitian menunjukkan bahwa penggunaan model mT5 yang dikombinasikan dengan optimasi hyperparameter menggunakan Optuna memberikan peningkatan performa yang sangat signifikan dalam tugas penerjemahan Bahasa Angkola–Bahasa Indonesia. Tingginya skor CHRF mengindikasikan bahwa model berhasil menangkap variasi morfologi dan afiksasi Bahasa Angkola secara lebih baik. Hal ini sejalan dengan karakteristik mT5 sebagai model multilingual yang mampu melakukan cross-lingual transfer learning, khususnya pada bahasa dengan sumber daya terbatas. Selain itu, penggunaan label smoothing, dropout, dan scheduler yang optimal terbukti meningkatkan stabilitas pelatihan dan mencegah overfitting, sebagaimana tercermin dari rendahnya nilai test loss. Namun demikian, keterbatasan penelitian ini terletak pada ukuran dan domain dataset yang masih terbatas pada kamus. Pengayaan korpus berbasis percakapan atau teks naratif berpotensi meningkatkan kualitas terjemahan lebih lanjut.

4. KESIMPULAN

Berdasarkan hasil eksperimen dan pembahasan, dapat disimpulkan bahwa pengembangan sistem penerjemahan Bahasa Angkola ke Bahasa Indonesia menggunakan model multilingual Text-to-Text Transfer Transformer (mT5) dengan optimasi hyperparameter berbasis Optuna memberikan peningkatan kinerja yang sangat signifikan. Dataset paralel hasil prapemrosesan berjumlah 28.775 pasangan kalimat terbukti mampu mendukung proses pelatihan, validasi, dan pengujian model secara objektif. Model mT5 dengan konfigurasi standar hanya menghasilkan kualitas terjemahan yang terbatas, yang ditunjukkan oleh nilai BLEU dan CHRF yang relatif rendah. Namun, setelah dilakukan optimasi hyperparameter menggunakan Optuna, performa model meningkat secara drastis dengan nilai BLEU mencapai 73,84 dan CHRF sebesar 83,34 pada data uji. Peningkatan ini menunjukkan bahwa pemilihan hyperparameter yang optimal berperan penting dalam memaksimalkan kemampuan model mT5, khususnya dalam menangani bahasa dengan sumber daya terbatas seperti Bahasa Angkola. Selain peningkatan kuantitatif, hasil terjemahan model teroptimasi menunjukkan struktur kalimat yang lebih alami dan kesesuaian semantik yang lebih baik. Dengan demikian, pendekatan mT5 yang dioptimasi menggunakan Optuna terbukti efektif dan berpotensi diterapkan untuk pengembangan sistem penerjemahan bahasa daerah lainnya.

REFERENCES

[1] G. P. M. Virgilio, F. Saavedra Hoyos, and C. B. Bao Ratzemberg, “The impact of artificial intelligence on unemployment: a review,” *Int. J. Soc. Econ.*, no. January, pp. 25553–25579, 2024, doi: 10.1108/IJSE-05-2023-0338.
[2] I. Bakhov, “Artificial intelligence tools for automating philological text research Herramientas de inteligencia artificial para automatizar la investigación filológica de textos,” 2025, doi: 10.62486/latia2025293.
[3] S. Cahyawijaya *et al.*, “NusaWrites: Constructing High-Quality Corpora for Underrepresented and Extremely Low-Resource Languages,” 2023.
[4] D. Arbian, A. Prasetya, D. Dwi, and F. Almu, “Multilingual Parallel Corpus for Indonesian Low-Resource Languages,” vol. 9, no. September, 2025.
[5] N. Nadra, R. Marnita, and K. A. Amini, “Assimilation of the Batak Angkola Language in Pintu Padang, North Sumatra, Indonesia,” *J. Arbitrer*, vol. 11, no. 1, pp. 29–38, 2024, doi: 10.25077/ar.11.1.29-38.2024.
[6] R. M. Khofifah Aisah Amini, Nadra Nadra, “Affix Form and Morphonemic Process of Batak Angkola Language Bentuk Afiks Dan Proses Morfofonemik Bahasa Batak Angkola,” *Pendidijan, Bahasa, Dan Sastra*, vol. 11, no. 1, pp. 30–38, 2023.
[7] N. H. Hrp, M. Fikry, and Y. Yusra, “Algoritma Stemming Teks Bahasa Batak Angkola Berbasis Aturan Tata Bahasa,” *J. Comput. Syst. Informatics*, vol. 4, no. 3, pp. 642–648, 2023, doi: 10.47065/josyc.v4i3.3458.
[8] V. Agarwal, S. B. Pooja Rao, and D. B. Jayagopi, “Hinglish to English Machine Translation using Multilingual Transformers,” *Int. Conf. Recent Adv. Nat. Lang. Process. RANLP*, vol. 2021-Septe, pp. 16–21, 2021, doi: 10.26615/issn.2603-2821.2021_003.

[9] L. Xue, N. Constant, A. Roberts, and M. Kale, "mT5 : A Massively Multilingual Pre-trained Text-to-Text Transformer," pp. 483–498, 2021.
[10] M. Kale and A. Siddhant, "nmT5 - Is parallel data still relevant for pre-training massively multilingual language models ?," pp. 683–691, 2021.
[11] M. Kresic and N. Abbas, "ScienceDirect Normalizing Swiss German Dialects with the Power of Large Language Models," *Procedia Comput. Sci.*, vol. 244, pp. 287–295, 2024, doi: 10.1016/j.procs.2024.10.202.
[12] Z. Liu, Y. Lin, and M. Sun, *Representation Learning for Natural Language Processing, Second Edition*. 2023. doi: 10.1007/978-981-99-1600-9.
[13] B. Zhang *et al.*, "Encoder-Decoder Gemma: Improving the Quality-Efficiency Trade-Off via Adaptation," 2025, [Online]. Available: http://arxiv.org/abs/2504.06225
[14] N. Alinda and S. Defit, "The Use of Hyperparameter Tuning in Model Classification : A Scientific Work Area Identification," vol. 8, no. December, pp. 2181–2188, 2024.
[15] B. Haddow, R. Bawden, A. Valerio, M. Barone, and A. Birch, "Survey of Low-Resource Machine Translation," no. August 2021, 2022.
[16] J. Richardson, "SentencePiece : A simple and language independent subword tokenizer and detokenizer for Neural Text Processing," pp. 66–71, 2018.
[17] T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, and P. Networks, "O ptuna : A N ext - generation H yperparameter O ptimization F ramework," pp. 1–10, 2019.
[18] M. Azam, K. Raiaan, S. Sakib, and N. Mohammad, "A systematic review of hyperparameter optimization techniques in Convolutional Neural Networks," *Decis. Anal. J.*, vol. 11, no. September 2023, p. 100470, 2024, doi: 10.1016/j.dajour.2024.100470.
[19] K. Papineni, S. Roukos, T. Ward, and W. Zhu, "B LEU : a Method for Automatic Evaluation of Machine Translation," no. July, pp. 311–318, 2002.
[20] M. Popovi, "CHR F : character n -gram F-score for automatic MT evaluation," no. September, pp. 392–395, 2015.