

Pengembangan Sistem Penerjemahan Otomatis Bahasa Batak ke Bahasa Indonesia untuk Konten Edukasi Digital

Fricles Ariwisanto Sianturi,¹⁾ Ismail Marzuki Sianturi

¹⁾Universitas Tjut Nyak Dhien, Sumatera Utara, Indonesia

¹⁾Universitas Audi Indonesia, Sumatera Utara, Indonesia

Email Corresponding: sianturifricles@utnd.ac.id*

Kata Kunci

Sistem penerjemahan otomatis,
Bahasa batak,
Bahasa indonesia,
Edukasi digital,
Neural machine translation,
Pelestarian bahasa daerah.

ABSTRAK

Pengabdian ini bertujuan untuk mengembangkan sistem penerjemahan otomatis Bahasa Batak ke Bahasa Indonesia guna mendukung konten edukasi digital serta memperkuat pelestarian bahasa daerah. Permasalahan yang diangkat adalah terbatasnya ketersediaan perangkat komputasional untuk Bahasa Batak dan kompleksitas struktur morfologi serta konteks linguistik yang memengaruhi akurasi terjemahan ke dalam Bahasa Indonesia. Pengabdian ini menggunakan pendekatan hibrida yang menggabungkan metode berbasis aturan (rule-based) dan Neural Machine Translation (NMT). Tahapan Pengabdian meliputi pengumpulan dan anotasi korpus paralel bilingual, proses prapemrosesan (tokenisasi, normalisasi, dan stemming), pelatihan model, evaluasi sistem menggunakan metrik BLEU dan tingkat akurasi, serta uji kegunaan pada platform pembelajaran digital. Hasil Pengabdian menunjukkan bahwa sistem yang dikembangkan memperoleh skor BLEU sebesar 78,4 dan tingkat akurasi terjemahan sebesar 85% pada pengujian terkontrol. Integrasi aturan morfologi terbukti meningkatkan kualitas terjemahan, khususnya pada struktur imbuhan dan ekspresi kontekstual. Uji coba pada modul pembelajaran digital menunjukkan peningkatan aksesibilitas dan pemahaman materi bagi peserta didik dan pendidik yang menggunakan konten berbahasa Batak. Simpulannya, sistem penerjemahan otomatis yang dikembangkan mampu secara efektif mengonversi konten Bahasa Batak ke Bahasa Indonesia sehingga mendukung pendidikan digital yang inklusif dan berkontribusi pada keberlanjutan bahasa daerah. Pendekatan hibrida yang diterapkan juga memiliki potensi untuk dikembangkan pada sistem penerjemahan bahasa daerah lainnya

Keywords

Automatic translation system,
Batak language,
Indonesian language,
Digital education,
Neural machine translation,
Language preservation.

ABSTRACT

This study aims to develop an automatic translation system from Batak language to Indonesian to support digital educational content and promote regional language preservation. The research addresses the limited availability of computational linguistic tools for Batak and the challenges of translating morphological and contextual structures into standard Indonesian. The study employs a hybrid approach combining rule-based linguistic modeling and neural machine translation (NMT). The development stages include corpus collection and annotation, preprocessing (tokenization, normalization, and stemming), model training using a bilingual parallel corpus, system evaluation using BLEU score and accuracy metrics, and usability testing in digital learning platforms. The results indicate that the proposed system achieved a BLEU score of 78.4 and translation accuracy of 85% in controlled testing scenarios. The integration of morphological rules significantly improved translation performance for affixation and contextual expressions. User testing in digital educational modules demonstrated increased accessibility and comprehension for students and educators utilizing Batak-language materials. In conclusion, the developed automatic translation system effectively facilitates the conversion of Batak language content into Indonesian, contributing to inclusive digital education and regional language sustainability. The hybrid translation approach proves reliable and adaptable for further development in other local language translation systems..

Informasi Artikel

Diterima Redaksi: 14 April 2025

Revisi Terakhir: 20 Juli 2025

Diterbitkan: 30 Agustus 2025

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



I. PENDAHULUAN

Transformasi digital dalam bidang pendidikan telah mendorong pemanfaatan teknologi berbasis kecerdasan buatan untuk meningkatkan aksesibilitas dan kualitas pembelajaran. Di Indonesia, yang memiliki ratusan bahasa daerah, integrasi teknologi bahasa menjadi isu strategis, terutama dalam mendukung inklusivitas pendidikan dan pelestarian bahasa lokal. Bahasa Batak sebagai salah satu bahasa daerah yang memiliki kekayaan struktur morfologi dan variasi dialek menghadapi tantangan serius dalam adaptasi ke dalam ekosistem pembelajaran digital, khususnya dalam bentuk sistem penerjemahan otomatis yang akurat dan kontekstual (Aritonang & Sitohang, 2024; Baharuddin & Nasution, 2023).

Secara global, perkembangan sistem penerjemahan otomatis telah mengalami kemajuan signifikan melalui pendekatan Statistical Machine Translation (SMT) dan kemudian Neural Machine Translation (NMT). Model NMT berbasis encoder–decoder dan attention mechanism terbukti meningkatkan kualitas terjemahan pada bahasa dengan sumber daya besar (high-resource languages). Namun, pada bahasa dengan sumber daya terbatas (low-resource languages), performa model neural murni cenderung menurun akibat keterbatasan korpus paralel dan variasi morfologis yang kompleks. Oleh karena itu, pendekatan hibrida yang menggabungkan metode berbasis aturan linguistik (rule-based) dengan model neural menjadi alternatif yang banyak direkomendasikan dalam literatur terkini (Ginting & Pratama, 2025).

Dalam konteks nasional, berbagai Pengabdian dan pengabdian sebelumnya telah mengembangkan sistem penerjemahan untuk beberapa bahasa daerah, seperti Jawa dan Sunda, baik berbasis kamus digital maupun model statistik sederhana. Beberapa program pengabdian juga berfokus pada digitalisasi kamus dan dokumentasi bahasa daerah sebagai bagian dari pelestarian budaya (Hasanah & Sitorus, 2022; Koehn, 2021). Namun, sebagian besar inisiatif tersebut belum secara spesifik mengintegrasikan sistem penerjemahan otomatis ke dalam platform edukasi digital, serta belum mengoptimalkan analisis morfologi lokal dalam arsitektur model berbasis neural. Selain itu, Bahasa Batak masih relatif minim dukungan korpus terstandar yang siap digunakan untuk pelatihan model penerjemahan otomatis.

Berdasarkan tinjauan tersebut, terdapat kesenjangan (gap analysis) yang jelas, yaitu: (1) keterbatasan sistem penerjemahan otomatis Bahasa Batak yang teruji secara komputasional, (2) belum optimalnya integrasi pendekatan hibrida dalam menangani struktur morfologi Bahasa Batak, dan (3) belum adanya implementasi sistem yang secara khusus dirancang untuk mendukung konten edukasi digital. Kesenjangan ini menunjukkan perlunya pengembangan sistem yang tidak hanya berfokus pada aspek teknis penerjemahan, tetapi juga pada kebermanfaatannya dalam lingkungan pembelajaran digital (Lumbantobing, 2024).

Dengan demikian, kontribusi baru dari Pengabdian ini terletak pada pengembangan sistem penerjemahan otomatis Bahasa Batak ke Bahasa Indonesia berbasis pendekatan hibrida yang mengintegrasikan analisis morfologi lokal dan model Neural Machine Translation, serta implementasinya secara langsung dalam konteks konten edukasi digital. Pendekatan ini diharapkan mampu meningkatkan akurasi terjemahan, memperluas akses pembelajaran, dan mendukung pelestarian bahasa daerah secara berkelanjutan melalui teknologi.

II. MASALAH

Berdasarkan latar belakang dan tujuan Pengabdian yang berjudul “Pengembangan Sistem Penerjemahan Otomatis Bahasa Batak ke Bahasa Indonesia untuk Konten Edukasi Digital”, maka rumusan masalah dalam Pengabdian ini disusun sebagai berikut:

- a. Bagaimana merancang dan mengembangkan sistem penerjemahan otomatis Bahasa Batak ke Bahasa Indonesia yang mampu bekerja secara akurat dan kontekstual untuk mendukung konten edukasi digital?
- b. Bagaimana mengintegrasikan pendekatan berbasis aturan (rule-based) dan Neural Machine Translation (NMT) dalam menangani karakteristik morfologi dan variasi struktur Bahasa Batak
- c. Bagaimana membangun dan mengelola korpus paralel bilingual Bahasa Batak–Bahasa Indonesia yang memadai untuk pelatihan dan evaluasi model penerjemahan?

- d. Sejauh mana tingkat akurasi dan kualitas terjemahan yang dihasilkan sistem, berdasarkan pengukuran menggunakan metrik evaluasi seperti BLEU score dan uji pemahaman pengguna?
 - e. Bagaimana efektivitas implementasi sistem penerjemahan otomatis dalam meningkatkan aksesibilitas dan pemahaman konten edukasi digital bagi pengguna?
- Rumusan masalah ini menjadi dasar dalam perancangan metodologi, pengembangan sistem, serta evaluasi hasil Pengabdian guna memastikan bahwa sistem yang dikembangkan benar-benar relevan dengan kebutuhan pendidikan digital dan pelestarian bahasa daerah (Nasution, 2023).

III. METODE

Pengabdian ini menggunakan pendekatan *Research and Development* (R&D) dengan fokus pada perancangan, pengembangan, dan evaluasi sistem penerjemahan otomatis Bahasa Batak ke Bahasa Indonesia untuk konten edukasi digital. Metodologi disusun secara sistematis agar dapat direplikasi serta merujuk pada praktik standar dalam pengembangan sistem Neural Machine Translation (NMT) dan pendekatan rule-based dalam pemrosesan bahasa alami (Pangaribuan, 2022)

3.1. Desain Pengabdian

Desain Pengabdian terdiri atas lima tahapan utama:

1. Analisis Kebutuhan (Needs Assessment)

Dilakukan melalui studi literatur dan observasi pada platform pembelajaran digital untuk mengidentifikasi kebutuhan penerjemahan Bahasa Batak–Bahasa Indonesia. Analisis ini juga mencakup identifikasi karakteristik linguistik Bahasa Batak, terutama aspek morfologi dan variasi dialek.

2. Pengumpulan dan Penyusunan Korpus

- a. Mengumpulkan korpus paralel bilingual dari:
- b. Buku pelajaran muatan lokal
- c. Cerita rakyat Batak
- d. Dokumen edukasi daerah
- e. Wawancara dan transkripsi lisan
- f. Data dianotasi secara manual oleh ahli bahasa.
- g. Korpus dibagi menjadi data pelatihan (80%), validasi (10%), dan pengujian (10%).

3. Prapemrosesan Data (Preprocessing)

Tahapan ini meliputi:

- a. Tokenisasi
- b. Normalisasi ejaan
- c. Case folding
- d. Stemming dan analisis morfologi berbasis aturan

Prosedur prapemrosesan merujuk pada standar pemrosesan teks dalam Pengabdian Natural Language Processing (NLP) sebelumnya, khususnya pada model encoder–decoder berbasis attention mechanism yang umum digunakan dalam NMT (Pane & Putro, 2024; Purba & Siregar, 2025).

4. Pengembangan Model Penerjemahan

Sistem dikembangkan menggunakan pendekatan hibrida, yaitu:

- a. Pendekatan Rule-Based untuk menangani struktur imbuhan, duplikasi, dan pola morfologi khas Bahasa Batak.
- b. Neural Machine Translation (NMT) berbasis arsitektur encoder–decoder dengan mekanisme perhatian (attention mechanism).

Arsitektur model mengacu pada model NMT standar berbasis sequence-to-sequence yang banyak digunakan dalam Pengabdian terdahulu. Proses pelatihan dilakukan menggunakan dataset paralel yang telah diproses.

5. Evaluasi dan Pengujian Sistem

Evaluasi dilakukan melalui dua pendekatan:

1. Evaluasi Kuantitatif

- a. Menggunakan metrik BLEU (Bilingual Evaluation Understudy).
- b. Menghitung tingkat akurasi terjemahan berbasis kesesuaian semantik dan struktur kalimat.

2. Evaluasi Kualitatif

- a. Uji keterbacaan oleh ahli bahasa.
- b. Uji kegunaan (usability testing) pada pengguna (guru dan siswa) dalam platform pembelajaran digital.

3.2. Bahan dan Instrumen Pengabdian

Beberapa bahan dan instrumen yang digunakan sebagai penunjang data Pengabdian meliputi:

Tabel 1. Bahan dan Instrumen Pengabdian

No	Bahan/Instrumen	Fungsi
1	Korpus paralel Bahasa Batak–Bahasa Indonesia	Data pelatihan dan evaluasi model
2	Dokumen edukasi digital	Sumber data implementasi sistem
3	Perangkat lunak NLP	Proses tokenisasi dan pelatihan model
4	Modul pembelajaran digital	Uji implementasi sistem
5	Kuesioner uji kegunaan	Evaluasi persepsi pengguna
6	Validator ahli bahasa	Validasi kualitas terjemahan

3.3. Teknik Analisis Data

- a. Analisis Kuantitatif:
 Menggunakan skor BLEU dan persentase akurasi untuk mengukur performa sistem.
- b. Analisis Kualitatif:
 Menggunakan metode deskriptif untuk mengevaluasi kesesuaian konteks, struktur morfologi, dan tingkat pemahaman pengguna.

3.4. Alur Pengembangan Sistem

Secara ringkas, alur metode Pengabdian adalah sebagai berikut:



Gambar 1. Alur Metode Pelaksanaan (Sari & Wijaya, 2021)

3.5. Validitas dan Keandalan

Untuk memastikan validitas:

- a. Data dianotasi oleh ahli bahasa.
- b. Evaluasi dilakukan menggunakan standar metrik internasional (BLEU).
- c. Pengujian dilakukan pada data yang tidak termasuk dalam dataset pelatihan.

Keandalan sistem diuji melalui pengulangan eksperimen dan perbandingan hasil pada beberapa subset data uji.

Metode yang disusun ini memberikan kerangka kerja yang jelas, sistematis, dan terukur dalam pengembangan sistem penerjemahan otomatis Bahasa Batak ke Bahasa Indonesia, serta memastikan bahwa hasil Pengabdian dapat dipertanggungjawabkan secara ilmiah dan aplikatif dalam konteks edukasi digita(Symbolon & Pardede, 2023).

IV. HASIL DAN PEMBAHASAN

4.1 Hasil Pengembangan Korpus dan Model

Pengabdian ini berhasil menyusun korpus paralel Bahasa Batak–Bahasa Indonesia sebanyak 18.500 pasangan kalimat, yang diperoleh dari buku muatan lokal, cerita rakyat, dan modul pembelajaran digital. Distribusi data terdiri dari 14.800 data pelatihan (80%), 1.850 data validasi (10%), dan 1.850 data pengujian (10%).

Model yang dikembangkan terdiri atas dua konfigurasi utama:

- Model NMT Murni (Baseline)
- Model Hibrida (Rule-Based + NMT)

Pelatihan dilakukan selama 40 epoch dengan parameter pembelajaran terstandar (learning rate 0,001; batch size 32).

4.2 Evaluasi Kuantitatif

Evaluasi dilakukan menggunakan metrik BLEU dan akurasi semantik. Hasil pengujian ditampilkan pada Tabel 1.

Tabel 2. Perbandingan Performa Model

Model	BLEU Score	Akurasi (%)	Waktu Proses (detik/100 kalimat)
NMT Murni	70,2	79%	2,4
Hibrida (Rule + NMT)	78,4	85%	2,8

4.3 Analisis Kesalahan (Error Analysis)

Analisis tambahan dilakukan terhadap 300 sampel kalimat uji untuk mengidentifikasi jenis kesalahan.

Tabel 3. Distribusi Jenis Kesalahan

Jenis Kesalahan	NMT Murni	Model Hibrida
Kesalahan Morfologi	34%	15%
Kesalahan Konteks	27%	19%
Kesalahan Leksikal	21%	18%
Struktur Kalimat	18%	12%

Terlihat bahwa integrasi aturan morfologi secara signifikan menurunkan kesalahan morfologi hampir 50% dibandingkan model baseline. Hal ini menunjukkan bahwa karakteristik Bahasa Batak yang kaya imbuhan lebih efektif ditangani dengan pendekatan berbasis aturan yang dikombinasikan dengan pembelajaran neural(Sitorus, 2024; Tarihoran, 2022).

4.4 Hasil Uji Implementasi pada Konten Edukasi Digital

Sistem diuji pada 5 modul pembelajaran digital tingkat SMP. Uji kegunaan melibatkan 30 responden (guru dan siswa).

Tabel 4. Hasil Uji Kegunaan (Usability Testing)

Aspek Penilaian	Skor Rata-rata (Skala 1–5)
Kemudahan Penggunaan	4,3
Kejelasan Terjemahan	4,4
Kecepatan Sistem	4,1
Manfaat dalam Pembelajaran	4,5

Hasil menunjukkan tingkat penerimaan yang tinggi terhadap sistem, terutama dalam aspek peningkatan pemahaman materi.

4.5 Skema Arsitektur Sistem

Berikut skema alur kerja sistem yang dikembangkan:



Gambar 2. Skema Alur Kerja Sistem

Skema ini menunjukkan integrasi antara komponen linguistik berbasis aturan dan arsitektur neural secara berurutan.

4.6 Pembahasan dan Perbandingan dengan Pengabdian Sebelumnya

Jika dibandingkan dengan pengabdian sebelumnya (Purba & Siregar, 2025; Simbolon & Pardede, 2023; Tarihoran, 2022) yang berfokus pada digitalisasi kamus Bahasa Batak atau sistem terjemahan berbasis kamus statis, Pengabdian ini menunjukkan beberapa keunggulan:

a. Pendekatan Adaptif dan Kontekstual

Sistem sebelumnya cenderung menghasilkan terjemahan literal, sedangkan sistem ini mampu mempertimbangkan konteks kalimat melalui mekanisme attention pada NMT.

b. Integrasi Morfologi Lokal

Berbeda dari model neural murni pada Pengabdian terdahulu, integrasi aturan morfologi secara signifikan meningkatkan akurasi struktur imbuhan.

c. Implementasi Langsung pada Platform Edukasi

Sebagian pengabdian sebelumnya berhenti pada tahap prototipe atau dokumentasi bahasa. Pengabdian ini menguji sistem secara langsung dalam modul pembelajaran digital.

4.7 Analisis Tambahan (Eksperimen Lanjutan)



Gambar 2. Tingkat Akurasi Pengabdian dengan Sebelumnya

Sebagai analisis tambahan, dilakukan eksperimen dengan pengurangan ukuran korpus menjadi 50%. Hasilnya menunjukkan penurunan skor BLEU menjadi 72,1. Hal ini menegaskan bahwa ukuran dan kualitas korpus sangat memengaruhi performa sistem.

Selain itu, pengujian pada variasi dialek menunjukkan bahwa model masih mengalami penurunan akurasi sebesar 5–7% pada dialek yang kurang terwakili dalam data pelatihan. Temuan ini menjadi rekomendasi untuk pengayaan korpus pada Pengabdian selanjutnya.

4.8 Sintesis Menuju Simpulan

Data kuantitatif, analisis kesalahan, serta hasil uji implementasi menunjukkan hubungan yang konsisten:

- Integrasi aturan morfologi → menurunkan kesalahan struktural
- Pendekatan hibrida → meningkatkan skor BLEU dan akurasi
- Implementasi edukatif → meningkatkan pemahaman pengguna

Dengan demikian, hasil Pengabdian secara logis mengarah pada simpulan bahwa sistem penerjemahan otomatis berbasis pendekatan hibrida efektif dan relevan untuk mendukung konten edukasi digital serta berkontribusi terhadap pelestarian Bahasa Batak melalui teknologi.

Hasil performansi, skor BLEU 78,4 menunjukkan peningkatan yang cukup signifikan dibandingkan rata-rata sistem berbasis kamus yang umumnya berada di bawah skor 65 dalam studi sebelumnya.

V. KESIMPULAN

Pengabdian ini bertujuan untuk mengembangkan sistem penerjemahan otomatis Bahasa Batak ke Bahasa Indonesia yang akurat, adaptif, dan terintegrasi dengan konten edukasi digital. Berdasarkan hasil pengembangan, pengujian, dan analisis yang telah dilakukan, dapat disimpulkan bahwa sistem yang dibangun dengan pendekatan hibrida (rule-based dan Neural Machine Translation) berhasil memenuhi tujuan Pengabdian.

Secara kuantitatif, sistem memperoleh skor BLEU sebesar 78,4 dan tingkat akurasi 85%, yang menunjukkan peningkatan signifikan dibandingkan model NMT murni. Analisis kesalahan membuktikan bahwa integrasi aturan morfologi efektif menurunkan kesalahan struktur imbuhan dan konteks kalimat secara substansial. Eksperimen tambahan terkait variasi ukuran korpus dan dialek juga menunjukkan bahwa kualitas dan representasi data berpengaruh langsung terhadap performa model, sehingga memperkuat validitas pendekatan yang digunakan.

Secara kualitatif, hasil uji implementasi pada modul pembelajaran digital menunjukkan tingkat penerimaan yang tinggi dari guru dan siswa, dengan skor rata-rata di atas 4 (skala 1–5) pada aspek kejelasan terjemahan dan manfaat pembelajaran. Temuan ini menegaskan bahwa sistem tidak hanya layak secara teknis, tetapi juga aplikatif dalam mendukung aksesibilitas dan pemahaman materi edukasi.

Dengan demikian, klaim bahwa sistem penerjemahan otomatis ini efektif untuk mendukung konten edukasi digital dan berkontribusi pada pelestarian Bahasa Batak didukung oleh data eksperimen, evaluasi kuantitatif, serta pengujian implementatif yang memadai. Pengabdian ini memberikan kontribusi penting dalam pengembangan teknologi bahasa daerah berbasis kecerdasan buatan dan membuka peluang pengembangan lebih lanjut untuk bahasa daerah lainnya di Indonesia.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih yang sebesar-besarnya kepada seluruh pihak yang telah memberikan dukungan dalam pelaksanaan Pengabdian dan pengembangan sistem penerjemahan otomatis Bahasa Batak ke Bahasa Indonesia untuk konten edukasi digital ini.

Ucapan terima kasih disampaikan kepada para ahli bahasa dan validator yang telah membantu dalam proses anotasi korpus serta validasi hasil terjemahan. Apresiasi juga diberikan kepada para guru dan siswa yang berpartisipasi dalam uji coba sistem pada modul pembelajaran digital, sehingga Pengabdian ini memperoleh data evaluasi yang komprehensif.

Penulis juga mengucapkan terima kasih kepada institusi dan tim pengembang yang telah menyediakan dukungan teknis, fasilitas Pengabdian, serta sumber daya komputasi yang diperlukan selama proses pengembangan dan pengujian sistem.

Semoga hasil Pengabdian ini dapat memberikan manfaat bagi pengembangan teknologi bahasa daerah, mendukung pendidikan digital yang inklusif, serta berkontribusi pada pelestarian Bahasa Batak secara berkelanjutan.

DAFTAR PUSTAKA

- Aritonang, M., & Sitohang, R. (2024). Pemanfaatan transformer-based model dalam penerjemahan dialek Batak Toba ke Bahasa Indonesia. *Jurnal Komputasi dan Informatika Indonesia*, 9(2), 112–125.
- Baharuddin, A., & Nasution, S. H. (2023). Tantangan morfologi dalam mesin penerjemah otomatis bahasa daerah: Studi kasus Bahasa Batak Mandailing. *Linguistik Komputasional*, 6(1), 45–58.
- Ginting, F., & Pratama, E. (2025). Digitalizing ethnic languages: Development of Batak Karo translation tools for educational resources. *International Journal of Digital Humanities*, 14(1), 89–104.
- Hasanah, U., & Sitorus, P. (2022). Strategi pengembangan konten edukasi digital berbasis kearifan lokal melalui teknologi bahasa. *Jurnal Teknologi Pendidikan Nusantara*, 5(3), 201–215.
- Koehn, P. (2021). *Neural machine translation*. Cambridge University Press.
- Lumbantobing, J. (2024). Analisis akurasi Rule-Based Machine Translation pada penerjemahan teks pendidikan Bahasa Batak. *Informatika Mulawarman: Jurnal Ilmiah Ilmu Komputer*, 19(1), 33–42.
- Nasution, A. Z. (2023). Corpus building for low-resource languages: The case of Batak Angkola. *Proceedings of the Southeast Asian Language Processing Conference*, 112–119.
- Pane, S. F., & Putro, M. D. (2024). Optimasi algoritma Deep Learning untuk penerjemahan otomatis bahasa daerah di Sumatera Utara. *Jurnal Sistem Informasi dan Teknologi*, 12(2), 150–165.
- Pangaribuan, T. (2022). *Struktur sintaksis Bahasa Batak: Fondasi bagi pemodelan mesin*. Penerbit Aksara Budaya.
- Purba, D., & Siregar, R. (2025). Efektivitas media pembelajaran digital berbasis bahasa ibu terhadap pemahaman konsep siswa di daerah terpencil. *Jurnal Inovasi Pembelajaran Digital*, 3(1), 12–28.

-
- Sari, R. P., & Wijaya, A. (2021). Neural Machine Translation (NMT) untuk bahasa sumber daya rendah (low-resource language) di Indonesia. *Jurnal Sains dan Informatika*, 7(2), 178–187.
- Simbolon, N., & Pardede, H. (2023). Evaluasi metode Bilingual Evaluation Understudy (BLEU) pada sistem penerjemah Bahasa Batak-Indonesia. *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi*, 9(1), 22–30.
- Sitorus, L. (2024). *Membangun korpus paralel Batak-Indonesia untuk aplikasi edukasi seluler*. Gramedia Pustaka Utama.
- Tarihoran, E. (2022). Peran penerjemahan otomatis dalam pelestarian bahasa daerah di era digital. *Jurnal Dokumentasi dan Informasi*, 43(2), 311–325.