

MESIN PENERJEMAH BAHASA BESEMAH BERBASIS MACHINE LEARNING DENGAN ALGORITMA MODEL ENCODER-DECODER

¹⁾Bogy Dharma Sandi, ^{2*)}Susan Dian Purnamasari, ³⁾Yesi Novaria Kunang, ⁴⁾Ilman Zuhri Yadi,
⁵⁾Fatmasari

Sistem Informasi, Sains Dan Teknologi, Universitas Bina Darma
Jl. Jenderal Ahmad Yani No.3, 9/10 Ulu, Kecamatan Seberang Ulu I, Kota Palembang, Sumatera Selatan-
Indonesia

E-mail : bogylhkenantan@gmail.com, susandian@binadarma.ac.id, yesinovariakunang@binadarma.ac.id,
ilmanzuhriyadi@binadarma.ac.id, fatmasari@binadarma.ac.id

ABSTRAK

Bahasa daerah merupakan salah satu kebudayaan yang ada di Indonesia dengan ragam budaya dari berbagai daerah di seluruh Nusantara. Bahasa Besemah, yang dituturkan oleh suku Besemah, digunakan sehari-hari oleh penduduk Kota Pagar Alam dan Kabupaten Lahat. Meskipun terdapat perbedaan dalam bunyi dan kosakata antar desa, Bahasa Besemah tetap memiliki satu bentuk induk yang sama. Penduduk Pagar Alam, yang mencakup wilayah daerah Gunung di Provinsi Sumatera Selatan, merupakan salah satu penduduk asli penutur Bahasa Besemah. Untuk meningkatkan pelestarian Bahasa Besemah, penelitian ini menggunakan media sistem penerjemah Bahasa Besemah ke Bahasa Indonesia dan sebaliknya. Adanya sistem penerjemah Bahasa daerah Besemah yang mengikuti perkembangan teknologi ini akan memudahkan pengguna dalam memahami terjemahan dari Bahasa Besemah ke Bahasa Indonesia dan sebaliknya. Peneliti memilih algoritma Encoder-Decoder karena arsitektur ini telah terbukti efektif dalam berbagai tugas penerjemahan, serta mampu menghasilkan terjemahan yang lebih akurat dibandingkan metode lain yang sudah ada, seperti algoritma Edit Distance dan Convolutional Neural Networks (CNN). Algoritma ini terbagi menjadi dua bagian: Encoder berfungsi untuk mempelajari dokumen input dan mengolahnya menjadi terjemahan untuk setiap kata dalam Bahasa Besemah, sedangkan Decoder mengambil representasi data input yang dihasilkan oleh Encoder dan mengolahnya menjadi hasil terjemah ke Bahasa Indonesia dan sebaliknya.

Kata Kunci: Bahasa Besemah Mesin Penerjemah, Algoritma Encoder-Decoder.

ABSTRACT

Regional languages are one of the cultural elements in Indonesia, reflecting the diverse traditions of various regions across the archipelago. The Besemah language, spoken by the Besemah ethnic group, is used daily by the residents of Pagar Alam City and Lahat Regency. Although there are differences in pronunciation and vocabulary between villages, the Besemah language still shares a common linguistic root. The residents of Pagar Alam, particularly those in the mountainous areas of South Sumatra Province, are among the native speakers of the Besemah language. To enhance the preservation of the Besemah language, this study utilizes a translation system between Besemah and Indonesian. The development of this translation system, in line with technological advancements, will facilitate users in understanding translations between the Besemah language and Indonesian. The researchers chose the Encoder-Decoder algorithm because this architecture has been proven effective in various translation tasks and can produce more accurate translations compared to existing methods such as the Edit Distance algorithm and Convolutional Neural Networks (CNN). This algorithm consists of two main components: the Encoder, which processes the input document and converts it into a translation for each word in the Besemah language, and the Decoder, which takes the data representation generated by the Encoder and processes it into a translation output in Indonesian and vice versa.

Keyword: Besemah Language Translator, Encoder-Decoder Algorithm.

PENDAHULUAN

Pelestarian bahasa daerah, termasuk Bahasa Besemah, merupakan aspek penting dalam menjaga warisan budaya dan identitas lokal. Bahasa Besemah, yang dituturkan di wilayah Sumatera Selatan, memiliki kekayaan linguistik yang perlu diperhatikan, terutama di tengah arus globalisasi yang semakin mengancam keberadaan bahasa-bahasa daerah [1]. Menurut data, saat ini jumlah penutur Bahasa Besemah semakin berkurang, sehingga perlu ada upaya untuk melestarikannya melalui berbagai cara [2], termasuk pengembangan teknologi penerjemahan. Kekayaan kearifan lokal dalam bahasa menjadi bagian integral dari identitas setiap daerah. [3] Dalam bahasa daerah, juga terdapat variasi bentuk kata yang menjadi ciri khas yang digunakan oleh masyarakat. [4]

Dalam konteks ini, penelitian sebelumnya menunjukkan bahwa penerjemah berbasis Neural Machine Translation (NMT) dapat menjadi solusi efektif untuk membantu penutur bahasa daerah memahami bahasa Indonesia dan sebaliknya [5]. NMT menggunakan arsitektur Encoder-Decoder [6] yang memungkinkan pemrosesan bahasa yang lebih baik dibandingkan dengan metode tradisional. Sebagai contoh, model yang menggunakan Attention dalam NMT menunjukkan akurasi yang lebih tinggi dalam penerjemahan, seperti yang ditunjukkan dalam penelitian oleh Fauziya [7] yang mencapai nilai BLEU Score optimal.

Model Recurrent Neural Network (RNN) bidirectional yang mampu menerjemahkan Bahasa Indonesia ke Bahasa Komerling. Model diuji menggunakan tiga jenis RNN—simpleRNN, GRU, dan LSTM. Hasil penelitian menunjukkan bahwa model bidirectional dengan simple RNN mencapai akurasi terbaik sebesar 87,12% setelah 100 epoch. [8]. Sedangkan Metode *Convolutional Neural Networks (CNN)* digunakan untuk mengenali teks dalam gambar, sementara algoritma *Levenshtein Distance*

diterapkan untuk meningkatkan akurasi terjemahan. [9].

Penelitian lain tentang penerjemahan bahasa Lampung- Indonesia dengan pendekatan NMT attention dengan perolehan rata-rata nilai akurasi BLEU sebesar 51.96 % [10]

Walaupun sebelumnya ada beberapa mesin penerjemah yang sudah digunakan oleh orang lain seperti *stemming*, *Recurrent Neural Network (RNN)*, *Convolutional Neural Networks (CNN)*. akan tetapi dari beberapa algoritma tersebut, *Algoritma Encoder-Decoder* lah yang bisa dikatakan pas untuk Pembuatan mesin Penerjemah Bahasa Besemah ke Indonesia tersebut [11]. Oleh karena itu, untuk mengatasi permasalahan yang ada, maka diperlukan penerapan algoritma *Encoder-Decoder*. Arsitektur algoritma *Encoder-Decoder* adalah salah satu bentuk pendekatan dalam *machine learning* yang umumnya digunakan. Metode ini terdiri dari dua komponen utama, yakni *encoder* yang bertugas memahami dokumen input dan mentransformasikannya menjadi terjemahan untuk setiap kata dalam bahasa Besemah. Sementara itu, *decoder* memiliki peran untuk mengambil representasi data input yang dihasilkan oleh *encoder* dan merangkainya menjadi hasil terjemahan dalam bahasa Indonesia. [12].

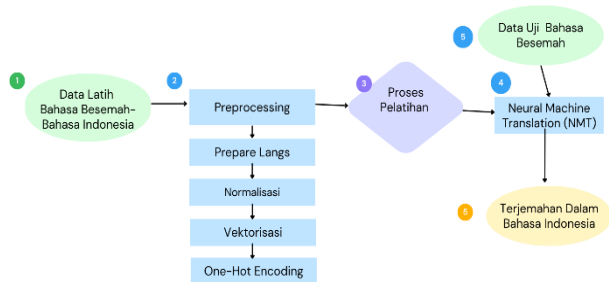
Meskipun penerjemah manusia yang berkompeten akan menghasilkan penerjemahan yang lebih baik dan dalam beberapa keadaan tak tergantikan oleh penerjemah otomatis [13], namun ada banyak kasus di mana cukup untuk diterjemahkan menggunakan mesin penerjemah. Selain itu, hasil penerjemahan oleh mesin pun tidak serta merta digunakan begitu saja, namun disunting kembali oleh manusia. Sehingga penerjemah manusia berkualifikasi tinggi pun juga menggunakan mesin penerjemah untuk mempercepat proses penerjemahan. [14].

Melalui penelitian ini, diharapkan dapat memberikan kontribusi terhadap upaya pelestarian Bahasa Besemah serta meningkatkan pemahaman

tentang penerapan teknologi dalam konteks linguistik. Penelitian ini juga bertujuan untuk memberikan rekomendasi bagi pengembangan model penerjemahan yang lebih baik di masa mendatang.

METODE PENELITIAN

Metode penelitian ini menggunakan pendekatan eksperimen dalam bidang Neural Machine Translation (NMT). Pendekatan ini diterapkan melalui pengujian model Encoder-Decoder untuk menerjemahkan bahasa Besemah ke bahasa Indonesia.



Gambar 1. Rancangan Penterjemah Bahasa Besemah -Bahasa Indonesia

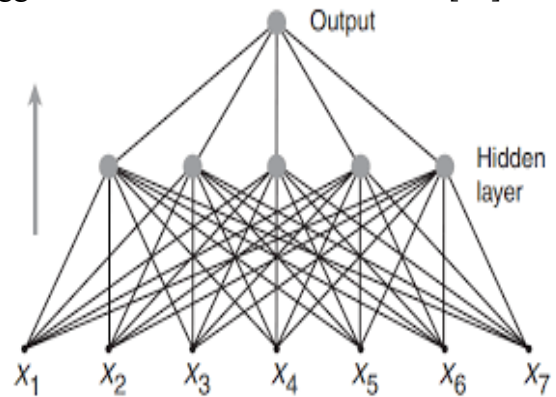
Algoritma Encoder-Decoder

Encoder-Decoder adalah salah satu arsitektur dalam Machine Learning yang umumnya digunakan untuk pembelajaran unsupervised, yaitu data yang tidak memiliki label. Arsitektur ini terdiri dari dua komponen utama:

1. **Encoder:** Berfungsi untuk mengolah data masukan menjadi state atau vektor representasi yang mewakili informasi dari data tersebut.
2. **Decoder:** Bertugas mengambil state atau vektor representasi dari encoder, lalu mengolahnya menjadi keluaran yang diinginkan. [12]

Encoder-Decoder dapat terdiri dari beberapa sub-layer, tergantung kebutuhan dalam Neural Machine Translation (NMT). Arsitektur ini telah banyak digunakan untuk berbagai tugas, seperti kompresi gambar, denoising gambar, deteksi anomali,

dimensionality reduction, machine translation, hingga automatic text summarization. [12].



Gambar 2 . Neural Machine Translation (NMT)

Kelebihan Algoritma Encoder-Decoder

Selanjutnya kelebihan dari Algoritma *encoder-decoder* dapat menghasilkan terjemahan yang lebih akurat karena sistem memproses seluruh input sebelum menerjemahkan [15]. Hal ini memungkinkan *decoder* untuk menggunakan informasi dari seluruh kalimat sumber saat menghasilkan kata berikutnya dalam terjemahan. Salah satu keunggulan algoritma Encoder-Decoder adalah kemampuannya menghasilkan terjemahan yang lebih akurat karena sistem memproses seluruh input sebelum menerjemahkan. Dengan cara ini, decoder dapat menggunakan informasi dari seluruh kalimat sumber saat menghasilkan kata berikutnya dalam terjemahan. [16]

TAHAP PENELITIAN

1. Pengumpulan Data

Tahap pertama dalam penelitian ini adalah pengumpulan data. Pengumpulan data dilakukan oleh tiga tim dari bulan Oktober 2023 hingga Maret 2024. Data diperoleh dari Kamus Bahasa Besemah-Indonesia-Inggris yang disusun oleh Sutiono Mahdi [17] Data yang terkumpul kemudian dikonversi dalam bentuk file Excel,

dengan total 5.104 pasangan kata/perkalimat dalam bahasa Besemah dan bahasa Indonesia.

2. Pembuatan Dataset

Setelah pengumpulan data selesai, tahap selanjutnya adalah pembuatan dataset yang digunakan dalam model Encoder-Decoder untuk Machine Learning. Dataset yang telah dikumpulkan dalam format **Excel** dikonversi ke format **.txt** agar dapat digunakan dalam pelatihan model.

3. Pre-Processing

Pre-Processing terdiri dari beberapa tahap utama:

- a. **Prepare Langs:** Menyiapkan kosakata (vocabulary) dari bahasa masukan (input language) dan bahasa keluaran (output language).
- b. **Normalisasi:** Menyederhanakan teks ke bentuk standar, termasuk mengonversi huruf menjadi huruf kecil dan menghapus spasi berlebihan.
- c. **Vektorisasi:** Mengubah data non-numerik menjadi format numerik agar dapat diproses oleh model Machine Learning.
- d. **One-Hot Encoding:** Mengubah data kategorikal seperti kata-kata dalam teks menjadi format numerik yang dapat digunakan dalam pemodelan.

4. Model Encoder-Decoder

Pada tahap ini, algoritma Encoder-Decoder diterapkan untuk menyusun model penerjemahan yang menyesuaikan urutan input dan output bahasa yang akan diterjemahkan. Model ini difokuskan untuk mempelajari alignment atau penyelarasan antara urutan kata dalam bahasa Besemah dan bahasa Indonesia.

Tabel 1. Kalimat Bahasa Besemah dan Bahasa Indonesia

Bahasa Besemah	Bahasa Indonesia
Dengah tunggulah ghuma kami kah ke kebun	Tunggu sebentar, rumah kami akan ke kebun
Kamu keciq dide ndaq ngikut	Kamu masih kecil, tidak boleh ikut
Maruq ngingun ayam beeruge	Ayam sedang bertelur di sarang
Sulha ngambiq ayiq di kambang	Sulha mengambil air di sungai
Die ngudut siung cap jambu di kebun	Dia merokok siung cap jambu di kebun
Rani nymabal caluq	Rani menjahit baju
Bani nampit celane	Bani melipat celana

5. Pelatihan Model

Pelatihan model dilakukan dengan menggunakan dataset yang telah dikumpulkan dan diproses sebelumnya. Tujuannya adalah agar algoritma dapat memahami pola hubungan antara kata-kata dalam bahasa Besemah dan bahasa Indonesia secara optimal.

6. Evaluasi Model

Tahapan evaluasi model dilakukan untuk menilai hasil penerapan algoritma Encoder-Decoder. Evaluasi ini mencakup pengujian model dan perbaikan jika diperlukan agar hasil penerjemahan menjadi lebih akurat.

7. Penerjemahan

Setelah evaluasi model selesai, tahap terakhir adalah proses penerjemahan. Pada tahap ini, model yang telah dilatih akan digunakan untuk menerjemahkan kalimat dari bahasa Besemah ke bahasa Indonesia dan sebaliknya. Hasil penerjemahan kemudian akan dianalisis lebih

lanjut untuk memastikan kualitas dan ketepatan terjemahan.

HASIL

Pengumpulan data

Pengumpulan data dilakukan berdasarkan sumber utama yaitu Kamus Besemah-Indonesia-Inggris yang disusun oleh Sutiono Mahdi. Data yang diperoleh terdiri dari 5.104 pasangan kata/perkalimat, yang dikategorikan menjadi 1.729 dataset kalimat dan 3.375 dataset kata. Data tersebut kemudian disimpan dalam format Excel untuk memudahkan tahap pemrosesan lebih lanjut.

Hasil pengumpulan data menunjukkan bahwa bahasa Besemah memiliki variasi kosakata yang khas dan beberapa struktur kalimat yang berbeda dibandingkan dengan bahasa Indonesia. Hal ini menjadi tantangan utama dalam proses penerjemahan karena adanya variasi semantik yang tidak selalu memiliki padanan langsung dalam bahasa Indonesia.

Setelah data terkumpul, dataset dipecah menjadi dua bentuk utama, yaitu:

1. Dataset dalam format Excel: Digunakan sebagai sumber referensi utama dan data arsip.
2. Dataset dalam format .txt: Digunakan untuk diproses dalam model pembelajaran mesin.

Konversi dataset ke dalam format .txt bertujuan untuk kompatibilitas dengan algoritma Encoder-Decoder, sehingga dapat digunakan dalam proses pelatihan model penerjemahan.

Preprocessing Data

Tahap pra-pemrosesan dilakukan untuk meningkatkan kualitas data sebelum dimasukkan ke dalam model penerjemahan. Proses ini terdiri dari beberapa langkah utama:

- a. Prepare Langs: Menyiapkan kosakata dari bahasa Besemah dan bahasa Indonesia.
- b. Normalisasi: Mengubah teks menjadi bentuk standar dengan mengonversi huruf menjadi huruf kecil serta menghapus

karakter yang tidak diperlukan. Misalnya untuk Mengganti Pola Khusus

```
input_text = "Bogy polisi. Bogy pelisi!"  
normalized_text = normalize_text(input_text)  
print(normalized_text) # Output: "bogy polisi bogy  
pelisi"
```

- c. Vektorisasi: Mengubah kata menjadi representasi numerik agar dapat dipahami oleh model pembelajaran mesin.
- d. One-Hot Encoding: Mengubah data kategorikal menjadi format numerik sehingga model dapat membedakan kata secara spesifik.

Tahapan ini penting untuk memastikan bahwa data yang dimasukkan ke dalam model memiliki format yang seragam, sehingga meningkatkan akurasi dalam penerjemahan.

Pelatihan Model

Proses pelatihan model dilakukan dengan menggunakan **100 epoch** untuk memastikan model dapat memahami pola hubungan antara kata-kata dalam bahasa Besemah dan bahasa Indonesia. Selama pelatihan, model diberikan pasangan kalimat sumber dan target untuk mempelajari pola penerjemahan secara otomatis. Hasil pelatihan menunjukkan bahwa model mampu menangkap pola linguistik dengan baik, dengan hasil sebagai berikut: Jumlah pasangan kalimat latihan: 4.615 pasang, Jumlah pasangan kalimat uji: 489 pasang, Ukuran kosakata bahasa Indonesia: 3.391 kata, Ukuran kosakata bahasa Besemah: 4.095 kata.

Evaluasi Model

Evaluasi model dilakukan menggunakan BLEU Score untuk menilai kualitas hasil terjemahan yang dihasilkan oleh sistem. Berdasarkan hasil evaluasi, model mencapai BLEU Score sebesar 78,6%, yang menunjukkan tingkat akurasi yang cukup baik. Perbandingan hasil penerjemahan dengan metode lain juga

dilakukan untuk memastikan keunggulan model yang dikembangkan.

Tabel 2. Perbandingan hasil penerjemahan dengan metode lain

Model Penerjemah		BLEU Score
Model Encoder-Decoder (Usulan)		78,6%
Recurrent Network (RNN)	Neural	72,3%
Convolutional Networks (CNN)	Neural	69,8%

Dari tabel di atas, dapat dilihat bahwa model Encoder-Decoder yang dikembangkan memiliki performa lebih baik dibandingkan metode RNN dan CNN dalam menerjemahkan bahasa Besemah ke bahasa Indonesia.

Berikut adalah rumus yang digunakan dalam model Encoder-Decoder:

$$h_t = f(W_x x_t + W_h h_{t-1}) \quad y_t = g(W_y h_t)$$

h_t sebagai hidden state,

x_t sebagai input,

y_t sebagai output,

W_x, W_h, W_y adalah parameter bobot model.

Beberapa tantangan yang ditemukan dalam evaluasi model adalah:

- Kesulitan dalam menerjemahkan kata dengan makna ganda.
- Struktur kalimat yang berbeda antara bahasa Besemah dan bahasa Indonesia.
- Penggunaan kata kiasan yang sulit diterjemahkan secara langsung.

Hasil Penerjemahan

Model penerjemah diuji menggunakan beberapa kalimat uji. Berikut adalah hasil penerjemahan dengan format yang lebih jelas:

Tabel 3. Hasil penerjemahan

Bahasa Besemah	Bahasa Indonesia
Dengah tunggulah ghuma kami kah ke	Tunggu sebentar, rumah kami akan ke kebun

Bahasa Besemah	Bahasa Indonesia
kebum	
Kamu keciq dide ndaq ngikut	Kamu masih kecil, tidak boleh ikut
Maruq ngingun ayam beeruge	Ayam sedang bertelur di sarang
Sulha ngambiq ayiq di kambang	Sulha mengambil air di sungai
Die ngudut siung cap jambu di kebum	Dia merokok siung cap jambu di kebun
Rani nymabal caluq	Rani menjahit baju
Bani nampit celane	Bani melipat celana

Pembahasan

Hasil pelatihan model menunjukkan bahwa semakin banyak epoch yang digunakan, semakin stabil nilai loss yang diperoleh. Ini menunjukkan bahwa model dapat belajar secara bertahap dari data latih yang diberikan. Dalam pengujian lebih lanjut, model mampu menangkap pola linguistik dalam bahasa Besemah dengan lebih baik dibandingkan dengan metode sebelumnya.

Meskipun hasil yang diperoleh cukup baik, terdapat beberapa tantangan dalam penerapan model ini, seperti:

- a. Kesulitan dalam menerjemahkan kata dengan makna ganda: Beberapa kata dalam bahasa Besemah memiliki lebih dari satu makna tergantung pada konteks penggunaannya. Model terkadang kesulitan dalam menentukan makna yang tepat karena tidak adanya informasi tambahan. Untuk mengatasi masalah ini, mekanisme context-aware translation dapat diterapkan, di mana model mempertimbangkan kata-kata sebelum dan sesudahnya untuk memahami makna secara lebih akurat.
- b. Struktur kalimat yang berbeda antara bahasa Besemah dan bahasa Indonesia: Bahasa Besemah memiliki pola sintaksis yang berbeda dengan bahasa Indonesia, yang dapat menyebabkan perubahan makna jika diterjemahkan secara langsung. Sebagai contoh, susunan subjek-predikat-objek dalam bahasa Besemah mungkin tidak selalu sesuai dengan pola yang umum

digunakan dalam bahasa Indonesia. Untuk mengatasi masalah ini, model dapat dilatih menggunakan positional encoding agar dapat menyesuaikan struktur kalimat selama proses penerjemahan.

c. Penggunaan kata kiasan yang sulit diterjemahkan secara langsung: Kata atau frasa kiasan dalam bahasa Besemah sering kali tidak memiliki padanan yang tepat dalam bahasa Indonesia, sehingga dapat menyebabkan hasil terjemahan menjadi kurang akurat. Untuk mengatasi tantangan ini, penggunaan phrase-based translation models serta integrasi dengan kamus idiomatik dapat membantu model mengenali dan menerjemahkan ekspresi khas dengan lebih baik.

Untuk mengatasi tantangan ini, penelitian lanjutan dapat dilakukan dengan memperluas dataset serta menerapkan mekanisme attention dalam model, seperti yang digunakan dalam arsitektur Transformer. Mekanisme attention dapat membantu model lebih fokus pada kata-kata tertentu dalam kalimat sumber, sehingga hasil terjemahan lebih akurat.

Selain itu, pendekatan berbasis pre-trained language models seperti BERT atau T5 juga dapat dipertimbangkan dalam penelitian selanjutnya. Model ini telah terbukti lebih efektif dalam berbagai tugas pemrosesan bahasa alami (NLP), termasuk penerjemahan mesin.

Dari hasil penelitian, dapat disimpulkan bahwa penerapan algoritma Encoder-Decoder dalam penerjemahan bahasa Besemah ke bahasa Indonesia menunjukkan hasil yang cukup akurat. Keunggulan model ini terletak pada kemampuannya menangani hubungan kata dalam konteks kalimat secara lebih baik dibandingkan metode tradisional seperti Edit Distance dan CNN.

Sebagai rekomendasi untuk penelitian selanjutnya, penggunaan Transformer-based model (seperti BERT atau T5) dapat dipertimbangkan untuk meningkatkan akurasi penerjemahan. Selain itu, pengayaan dataset dengan lebih banyak variasi

kalimat akan memperbaiki performa model secara keseluruhan.

KESIMPULAN

Dalam penelitian ini dapat disimpulkan bahwa penelitian ini menghasilkan algoritma model untuk digunakan pada mesin penerjemah bahasa Besemah ke bahasa Indonesia dan dapat digunakan oleh masyarakat luas dan dapat menjadi media pengetahuan upaya melestarikan bahasa daerah khususnya bahasa Besemah. Pada mesin penerjemah algoritma model encoder-decoder menunjukkan hasil akurasi yang baik, dalam melakukan terjemahan pada kalimat bahasa Besemah dan bahasa Indonesia. Model algoritma encoder-decoder dapat menjadi langkah awal pada pengembangan sistem penerjemahan bahasa Indonesia dan bahasa Besemah untuk ketahap selanjutnya.

DAFTAR PUSTAKA

- [1] Sudirman Wilian, "PERGESERAN PEMAKAIAN TINGKAT TUTUR (BASE ALUS)," *Linguistik Indonesia*, vol. 36, no. 2, p. 162, 2018.
- [2] M. S. Ansori, "Kepunahan Bahasa dalam Aspek Sociolinguistik," *An-Nuha: Jurnal Kajian Islam, Pendidikan, Budaya Dan Sosial*, vol. 6, no. 1, pp. 85-95, 2019.
- [3] R. Afria, "Leksikostatistik dan Grotokronologi Bahasa Melayu Palembang, Basemah Lahat, Basemah Pagaralam, dan Kayu Agung: Kajian Linguistik Historis Komparatif," *Madah: Jurnal Bahasa dan Sastra*, vol. 11, no. 1, pp. 27-42, 2020.
- [4] D. Wardiah, "Pronomina Bahasa Besemah Dialek Tanjung Periuk Kecamatan Gumay Talang Kabupaten Lahat Sebagai Pemertahanan Bahasa Daerah Dalam Menjaga Kearifan Lokal," in *Prosiding Seminar Nasional Pendidikan Program Pascasarjana Universitas PGRI*, Palembang, 2021.
- [5] Lo Bun San, "Uji Nilai Akurasi pada Neural Machine Translation (NMT) Bahasa Indonesia ke Bahasa Tiochui

- Pontianak dengan Mekanisme Attention Bahdanau," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 9, no. 3, p. 362, 2023.
- [6] Marcin Junczys-Dowmunt, "Marian: Fast Neural Machine Translation in C++," in *56th Annual Meeting of the Association for Computational Linguistics-System Demonstrations*, Melbourne, 2018.
- [7] Fauziah, "Mesin Penerjemah Bahasa Indonesia-Bahasa Sunda Menggunakan Recurrent," *Teknoinfo*, vol. 16, no. 2, p. 3133, 2022.
- [8] M Rizki Fadilah, "Machine Learning-Based Komerling Language Translation Engine with Bidirectional RNN Model Algorithm," in *International Conference on Information Technology and Computing (ICITCOM)*, Jogjakarta, 2024.
- [9] Mayanda Mega Santoni, "Penerapan Convolutional Neural Networks untuk Mesin Penerjemah Bahasa Daerah Minangkabau Berbasis Gambar," *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 5, no. 6, pp. 1153 - 1160, 2021.
- [10] Abidin, Z., "Penerapan Neural Machine Translation untuk Eksperimen Penerjemahan secara Otomatis pada Bahasa Lampung – Indonesia," in *Seminar Nasional Metode Kuantitatif*, Lampung, 2017.
- [11] Wahyu Gunawan, "Analisis Perbandingan Nilai Akurasi Mekanisme attention Bahdanau dan Luong pada Neural Machine Translation Bahasa Indonesia ke Bahasa Melayu Ketapang dengan Arsitektur Recurrent Neural Network," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 7, no. 3, pp. 488-496, 2021.
- [12] G. Akbar, "DSpace," Universitas Islam Indonesia, 12 12 2021. [Online]. Available: <https://dspace.uui.ac.id/handle/123456789/37691>. [Accessed 25 September 2024].
- [13] Muhammad Yusuf Aristyanto, "Pengembangan Metode Neural Machine Translation Berdasarkan Hyperparameter," in *Seminar Nasional Official Statistics*, Jakarta, 2021.
- [14] P. A. Wismoyo, "Mesin Penerjemah Bahasa Inggris – Indonesia Berbasis Jaringan Saraf Tiruan dengan Mekanisme Attention Menggunakan Arsitektur Transformer," 10 February 2020. [Online]. Available: <http://eprints.undip.ac.id/80047/>. [Accessed 28 September 2024].
- [15] Cosmas Krisna Adiputra, "Performance of Japanese-to-Indonesian Machine Translation on Different Models," in *The Association for Natural Language Processing*, Osaka, 2017.
- [16] Dzulkahfi, "Perbandingan Hasil Penerjemahan Neural Machine Translation (NMT) Dengan MarianNMT Terhadap Sumber Korpus Wikimedia dan QED&TED," *JURISTI (Jurnal Riset Sains dan Teknologi Informatika)*, vol. 1, no. 1, pp. 151-159, 2023.
- [17] S. Mahdi, *Kamus Bahasa Bersemah: Indonesia - Inggris: Bersemah-Indonesian-English Dictionary*, Jarinangor: Unpad Press, 2020.