

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

<sup>1)</sup>Bogy Dharma Sandi, <sup>2)</sup>Susan Dian Purnmasari, <sup>3)</sup>Yesi Novaria Kunang, <sup>4)</sup>Ilman Zuhri Yadi, <sup>5)</sup>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](mailto:bogylhkenantan@gmail.com).

### ABSTRAK

Bahasa daerah merupakan salah satu kebudayaan yang ada di Indonesia dengan ragam budaya dari berbagai macam daerah di seluruh Nusantara. Bahasa Besemah, merupakan salah satu bahasa daerah yang dituturkan oleh suku Besemah. Bahasa Besemah digunakan sehari-hari oleh penduduk Masyarakat 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 yang ada di provinsi Sumatera Selatan, merupakan salah satu penduduk asli bahasa Besemah. Untuk meningkatkan pelestarian bahasa Besemah, maka penelitian ini perlu menggunakan media sistem penerjemah bahasa Besemah ke Indonesia dan sebaliknya. Adanya mesin terjemah bahasa daerah Besemah yang mengikuti perkembangan dunia teknologi ini akan memudahkan untuk mengetahui terjemahan dari bahasa Besemah ke Indonesia dan sebaliknya. Oleh karena itu, maka peneliti menggunakan algoritma *Encoder-Decoder*. Algoritma *encoder-decoder* merupakan salah satu arsitektur *machine learning* yang bisa digunakan untuk mesin penerjemah. Algoritma ini terbagi menjadi dua bagian, *Encoder* berfungsi untuk mempelajari dokumen *input* lalu mengolahnya menjadi terjemahan untuk tiap kata bahasa Besemah. *Decoder* berfungsi untuk mengambil representasi data *input* yang dihasilkan *encoder* dan mengolahnya menjadi hasil terjemah ke bahasa Besemah ke Indonesia dan sebaliknya.

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

### ABSTRACT

Regional languages are one of the cultures in Indonesia with a variety of cultures from various regions throughout the archipelago. Besemah language is one of the regional languages spoken by the Besemah tribe. Besemah language is used daily by the people of Pagar Alam City and Lahat Regency. Although there are differences in sound and vocabulary between villages, Besemah language still has the same parent form. The people of Pagar Alam, which covers the Gunung area in South Sumatra province, are one of the native inhabitants of the Besemah language. To improve the preservation of the Besemah language, this study needs to use a Besemah language translation system media to Indonesian and vice versa. The existence of a Besemah regional language translation machine that follows the development of the world of technology will make it easier to find out the translation from Besemah to Indonesian and vice versa. Therefore, the researcher uses the Encoder-Decoder algorithm. The encoder-decoder algorithm is one of the machine learning architectures that can be used for machine translation. This algorithm is divided into two parts, the Encoder functions to study input documents and then process them into translations for each word of the Besemah language. The decoder functions to take the representation of input data produced by the encoder and process it into translation results from Besemah to Indonesian and vice versa.

**Keyword:** Besemah Language Translator, Encoder-Decoder Algorithm.

### ENDAHULUAN

Besemah adalah nama suku yang mendiami daerah lembah Gunung Dempo di Kota Pagar Alam, Sumatera Selatan. Wilayah itu merupakan bagian dari Pulau Sumatera Selatan yang luas. Daerah yang didiami oleh suku Besemah itu kemudian disebut juga dengan tanah Besemah. Mereka menyebut dirinya sebagai jeme (orang) Besemah. (Bloom & Reenen, 2013)

Bahasa Besemah, sebagai salah satu bahasa daerah, menjadi bahasa sehari-hari bagi masyarakat di kota Pagar Alam dalam interaksi dengan penduduk desa. Saat terlibat dalam pertemuan resmi, warga Kota Pagar Alam beralih ke penggunaan bahasa Indonesia. Pemeliharaan dan penggunaan bahasa daerah ini mencerminkan rasa cinta dan upaya untuk melestarikan serta mempertahankan keberlanjutan bahasa

tersebut. Kekayaan kearifan lokal dalam bahasa menjadi bagian integral dari identitas setiap daerah. Dalam bahasa daerah, juga terdapat variasi bentuk kata yang menjadi ciri khas yang digunakan oleh masyarakat..(Wardiah, 2021)

Mesin penerjemah merupakan perangkat lunak penerjemah bahasa yang proses penerjemahannya dilakukan oleh mesin (komputer) secara otomatis dan hasilnya dapat langsung disajikan kepada pengguna. Opsi penggunaan mesin penerjemah lebih umum dipilih karena cepat dan praktis dalam melakukan penerjemahan teks, berbanding terbalik dengan penerjemahan oleh manusia yang memerlukan waktu lama. Meskipun penerjemah manusia yang berkompeten akan menghasilkan penerjemahan yang lebih baik dan dalam beberapa keadaan tak tergantikan oleh penerjemah otomatis, 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.(Wismoyo, 2018)

Sebelumnya ada beberapa mesin penerjemah yang sudah digunakan oleh orang lain seperti *stemming*, *Algoritme Edit Distance*, *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. 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.(AKBAR, 2021).

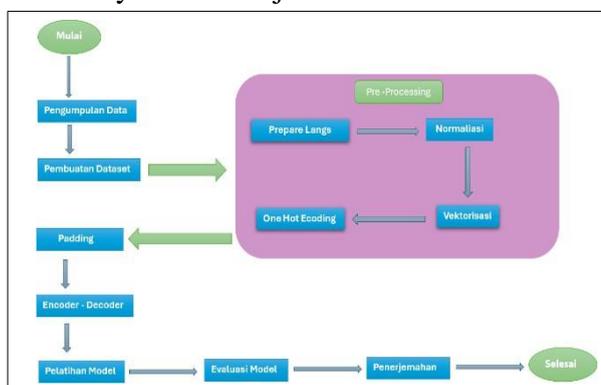
Oleh karena itu untuk menyelesaikan permasalahan yang ada maka diperlukan algoritma *Encoder-Decoder* ,merupakan algoritma yang dapat membantu pengembangan mesin penerjemah Bahasa daerah.Pada Bahasa daerah Besemah, Untuk itu diperlukan teknologi penerjemahan otomatis dari Bahasa Besemah ke Indonesia agar Bahasa Besemah semakin dikenal dan tidak ketinggalan dengan budaya asing.Selanjutnya, tujuan dari penelitian ini adalah merancang dan membangun sebuah mesin penerjemah sebagai solusi untuk memperluas pemahaman dan penggunaan bahasa Besemah. Maka dari itu penulis mengangkat penelitian dengan judul "**Mesin Penerjemah Bahasa Besemah Berbasis *Machine Learning* dengan Algoritma Model *Encoder-Decoder***".

### Algoritma Encoder-Decoder

Encoder-Decoder adalah salah satu arsitektur dalam *Machine Learning* yang umumnya digunakan untuk pembelajaran *unsupervised* yaitu data yang tidak memiliki label. Bagian ini akan dibagi menjadi dua bagian utama. Bagian pertama adalah *encoder*, bagian ini berfungsi untuk mengolah data masukan menjadi sebuah *state* atau vektor representasi yang mewakili informasi-informasi dari data masukan tersebut. Bagian kedua adalah *decoder*, bagian ini bertugas untuk mengambil *state* atau vektor representasi yang dihasilkan *encoder*, lalu mengolahnya menjadi keluaran yang diinginkan.(AKBAR, 2021)

Pada bagian *encoder* dan *decoder* dapat terdiri dari beberapa *sub-layer*. Tergantung kebutuhan layer *neural machine translation* yang biasa digunakan. Arsitektur *encoder decoder* telah banyak dimanfaatkan untuk beberapa tugas, diantaranya adalah untuk kompresi gambar, denoising gambar, deteksi anomali, dimensionality reduction, machine translation hingga automatic text summarization.(AKBAR, 2021)

Selanjutnya kelebihan dari Algoritma *encoder-decoder* dapat menghasilkan terjemahan yang lebih akurat karena sistem memproses seluruh input sebelum menerjemahkan. Hal ini memungkinkan *decoder* untuk menggunakan informasi dari seluruh kalimat sumber saat menghasilkan kata berikutnya dalam terjemahan.



### Gambar 1. Tahap Penelitian

#### a. Pengumpulan Data

Tahap pertama peneliti melakukan pengumpulan data, dan ditahap mengumpulkan data ini dibagi menjadi 3 tim untuk mengumpulkan data tersebut, dan pengumpulan data tersebut di mulai dari bulan Oktober 2023 sampai dengan Maret 2024, dan data tersebut didapatkan dari Kamus Bahasa Besemah-Indonesia-Inggris yang dibuat oleh Dr.Sutiono Mahdi, Drs., M.Hum. yang nantinya di kumpulkan dalam bentuk *file Excel*. dan Dataset yang berhasil dikumpulkan dari kamus tersebut berjumlah 5104 kata perkalimat dalam bahasa Besemah ke Indonesia.

#### b. Pembuatan Dataset

Dalam tahap ini, setelah terkumpulnya Dataset yang didapat dari Kamus Besemah-Indonesia-Inggris maka langkah selanjutnya adalah pembuatan Dataset untuk menghasilkan penerjemah yang akan di uji coba melalui model Encoder-Decoder pada Machine Learning, yang dimana Dataset yang sudah terkumpul 5104 dataset berbentuk file Excel akan di jadikan kedalam bentuk file txt.

#### c. Pre-Processing

*Pre-Processing* terdiri dari beberapa tahapannya yaitu, *Prepare Langs*, *Normalisasi*, *Vektorisasi*, *One Hot Encoding* Pada langkah *Prepare Langs* yang dimana fungsinya untuk menyiapkan kosakata dari bahasa masukan dan keluaran, Pada langkah *Normalisasi* berfungsi sebagai langkah-langkah untuk menyederhanakan teks ke bentuk standar, Pada langkah *Vektorisasi* berfungsi proses mengubah data non-numerik menjadi *format numerik* yang dapat digunakan oleh algoritma pembelajaran mesin, dan Pada langkah *One Hot Encoding* adalah teknik representasi data yang digunakan untuk mengkonversi kategori atau label

### 1. Prepare Langs

Dalam konteks Prepare Langs ini dijelaskan, Prepare Langs atau *Prepare Languages* akan merujuk pada fungsi yang bertujuan untuk menyiapkan kosakata (*vocabulary*) dari bahasa masukan (input language) dan bahasa keluaran (output language).

### 2. Normalisasi

Proses ini berupa langkah-langkah untuk menyederhanakan teks ke bentuk standar. Normalisasi melibatkan beberapa tindakan seperti mengonversi huruf menjadi huruf kecil dan menghapus spasi berlebihan.

### 3. Vektorisasi

Vektorisasi adalah proses mengubah data non-numerik menjadi format numerik yang dapat digunakan oleh algoritma pembelajaran mesin. Dalam konteks pemrosesan teks, ini biasanya melibatkan mengubah teks menjadi vektor.

### 4. One Hot Encoding

*One-hot encoding* adalah metode representasi data yang digunakan untuk mengubah data kategorikal, seperti kata-kata dalam teks, menjadi format numerik yang dapat diproses oleh model pembelajaran mesin

#### d. Encoder-Decoder

Encoder-Decoder pada tahapan ini berfungsi pada pengatur permodelan algoritma dalam penerjemahan yang disesuaikan dalam urutan input dan output bahasa yang akan diterjemahkan. Maka model ini dapat fokus pada tahapan atau bagian untuk mempelajari alignment atau penyelarasan urutan input dan output.

#### e. Pelatihan Model

Pelatihan model yang telah dibentuk sebelumnya akan diberikan dataset yang telah dibuat untuk melatih model. Tujuannya agar model algoritma dapat mempelajari tahapan hubungan antara

kalimat bahasa Besemah dan bahasa Indonesia

#### f. Evaluasi Model

Tahapan pada evaluasi model algoritma ini untuk menunjukkan hasil kegiatan yang dilakukan, dengan melakukan evaluasi hasil dari algoritma model encoder-decoder dengan pengujian dan perbaikan yang cukup baik

#### g. Penerjemahan

Pada tahapan ini, jika proses hasil evaluasi telah selesai maka pada penerjemahan ini akan menghasilkan proses penerjemahan bahasa. Pada proses penerjemahan bahasa Besemah ke bahasa Indonesia dan sebaliknya

## HASIL DAN PEMBAHASAN

### a. Pengumpulan Data

Pada penelitian yang dilakukan, data yang dikumpulkan dan akan digunakan bersumber dari data hasil KAMUS BESEMAH-INDONESIA-INGGRIS. Berdasarkan hasil data yang didapatkan berjumlah 5104 data yang terdiri dari 1729 untuk dataset kalimat dan 3375 untuk dataset kata. 5104 data yang dikumpulkan tersebut disimpan di dalam file excel.

### b. Pembuatan Dataset

Setelah data terkumpul, data tersebut akan dipecah menjadi 2 bagian yaitu Dataset yang berbentuk file Excel dan Dataset yang berbentuk file txt dan data bahasa Besemah

#### Contoh kalimat Bahasa Besemah

Dengah tunggulah ghuma kami kah ke kebun  
naniu keciq uue uuaq ngikul  
maruq ngisun ayam becuq

Sulha ngambiq ayiq di kambing

Die ngudut siung cap jambu di kebun

Rani nymabal caluq

Bani nampit celane

**Penerjemahan dari Kalimat Besemah ke Indonesia**

Kamu jagalah rumah kami mau ke kebun  
 Kamu anak-anak tidak usah ikut  
 Maruk memelihara ayam hutan  
 Sulha menimbah air di sumur  
 Dia merokok keretek cap jambu di kebun  
 Rani membuat sambal terasi  
 Bani menambal celana

Setelah dijadikan file Excel maka Dataset dijadikan ke bentuk file Txt agar dapat di proses didalam codingan mesin penerjemahan Bahasa Besemah ke Indonesia

**Dataset yang sudah dijadikan file Txt**

Bogy polisi. Bogy pelisi.  
 Bogy Tentra. Bogy Tentra.  
 Bapaknya Ketua RT. Bapaqe riye.  
 Dewi Penyanyi. Dewi penyanyi.  
 Sapi Binatang. Sapi Binatang.  
 Mujair Ikan. Mujair Ikan.  
 Hape Telepon. Hape telepun.  
 Ibu yesi guru kami. Ibu Yesi guru kami.

**c. Preprocessing**

Pada tahap ini, dataset yang telah dibentuk akan mengalami pra-pemrosesan terlebih dahulu. Pra-pemrosesan ini mencakup empat tahap yaitu *Prepare Langs*, Normalisasi, Vektorisasi, dan *One Hot Encoding*.

**1. Prepare Langs**

Dalam konteks kode yang sebelumnya dijelaskan, *PrepareLangs* akan merujuk pada fungsi yang bertujuan untuk menyiapkan kosakata (*vocabulary*) dari bahasa masukan (*input language*) dan bahasa keluaran (*output language*). Fungsi ini mengumpulkan semua kata unik dari setiap kalimat dalam bahasa masukan dan keluaran.

Berikut codingan pada bagian *Prepare Langs*

```
"""Lang class, used to store the vocabulary of each language"""
```

```
class Lang:
    def __init__(self, language):
```

```
        self.language_name = language
        self.word_to_index = {"SOS":SOS_token,
                              "EOS":EOS_token , "<UNK>":UNK_token}
        self.word_to_count = {}
        self.index_to_word = {SOS_token: "SOS",
                              EOS_token: "EOS" , UNK_token: "<UNK>"}
        self.vocab_size = 3
        self.cutoff_point = -1
```

**2. Normalisasi**

Proses ini mencakup langkah-langkah untuk menyederhanakan teks menjadi bentuk standar. Normalisasi melibatkan tindakan seperti mengubah semua huruf menjadi huruf kecil dan menghapus spasi yang berlebihan. Berikut adalah penjelasan langkah demi langkah untuk setiap bagian dalam fungsi koding yang ditunjukkan pada gambar di bawah ini :

**Codingan Mengganti Pola Khusus**

```
s = re.sub (r" ##AT##-##AT## ", r" ", s)
```

**Penjelasan**

Langkah ini menggantikan pola teks spesifik " ##AT##-##AT## " dengan satu spasi. Ini berguna jika teks masukan memiliki pola ini dan ingin menggantinya dengan spasi.

**Kode Codingan Mengubah Huruf menjadi Huruf Kecil**

```
s = uniToAscii(s.lower().strip())
```

**Penjelasan**

`s.lower()` mengubah semua huruf menjadi huruf kecil.

`s.strip()` menghapus spasi kosong di awal dan akhir string

**Kode Codingan Memisahkan Tanda Baca dari Kata**

```
s = re.sub (r " ([.!?])", r " \1", s)
```

**Penjelasan**

Langkah ini menambahkan satu spasi sebelum setiap tanda baca seperti titik (.), tanda seru (!), dan tanda tanya (?). Misalnya, "Hello!" akan menjadi "Hello"

**Codingan Menghapus Karakter Non- Huruf**

```
s = re.sub ( r " [^a-zA-Z.!?]+", r " ", s)
```

### Penjelasan

z, A-Z), titik (.), tanda seru (!), atau tanda tanya (?) dengan spasi. Karakter non huruf seperti angka, tanda baca lain, dan karakter khusus lainnya akan dihapus.

#### d. Vektorisasi

Vektorisasi adalah proses mengubah data non-numerik menjadi format data non-numerik menjadi format numerik yang dapat digunakan oleh algoritma pembelajaran mesin. Dalam konteks pemrosesan teks, ini biasanya melibatkan mengubah teks menjadi vektor. Vektorisasi memungkinkan model untuk bekerja dengan data teks dan mengekstrak informasi yang diperlukan untuk tugas pembelajaran.

#### e. One Hot Encoding

*One-hot encoding* adalah metode representasi data yang digunakan untuk mengubah data kategorikal, seperti kata-kata dalam teks, menjadi format numerik yang dapat diproses oleh model pembelajaran mesin. Dalam konteks pemrosesan bahasa alami, one-hot encoding sering digunakan untuk mengubah kata atau token menjadi vektor numerik yang merepresentasikan kata tersebut dalam ruang vektor.

Berikut codingan pada proses *One Hot Encoding*

```
def indexesFromSentence ( lang, sentence ):
    indexes = [ ]
    for word in sentence.split ( ' ' ):
        try:
            indexes.append ( lang.word_to_index [word]
        )
        except:
            indexes.append(lang.word_to_index["<UNK>"]
        )
    return indexes
```

```
def tensorFromSentence (lang, sentence):
    indexes = indexesFromSentence(lang,
    sentence)
    indexes.append(EOS_token)
    result = torch.LongTensor(indexes).view(-1)
    if use_cuda:
        return result.cuda()
    else:
        return result
```

#### f. Padding

*Padding* adalah proses dalam pemrosesan data yang digunakan untuk memastikan bahwa semua item dalam batch memiliki ukuran yang seragam. Ini penting dalam pembelajaran mesin, terutama dalam pemrosesan bahasa dan jaringan saraf, karena banyak model memerlukan input dengan ukuran tetap.

Berikut codingan pada proses *Padding*

```
"""seperates data into batches of size
batch_size"""
def batchify(data, input_lang, output_lang,
batch_size, shuffle_data=True):
    if shuffle_data == True:
        shuffle ( data )
    number_of_batches = len (data ) // batch_size
    batches = list(range(number_of_batches))
    longest_elements=list(range(number_of_batche
s))
```

#### g. Encoder-Decoder

Dalam penelitian ini algoritma *Encoder-Decoder* menerapkan untuk pembuatan mesin penerjemah dari bahasa Besemah ke bahasa Indonesia. Algoritma ini terdiri dari dua komponen utama ialah *encoder* dan *decoder*. *Encoder* bertanggung jawab untuk mengubah input menjadi teks dalam bahasa Besemah menjadi representasi vektor yang lebih abstrak. Sementara *decoder* menerima representasi vektor dan menghasilkan teks

dalam bahasa Indonesia  
yang merupakan hasil terjemahan.  
**Berikut Codingan Pada Proses Encoder-Decoder**

```
def train_batch(input_batch,target_batch,
encoder,decoder,encoder_optimizer,
decoder_optimizer, loss_criterion):
encoder_optimizer.zero_grad()
decoder_optimizer.zero_grad()
loss = 0
enc_h_hidden,enc_c_hidden=encoder.create_in
it_hiddens(input_batch.shape[1])

enc_hiddens,enc_outputs =encoder(input_batch,
enc_h_hidden,enc_c_hidden)decoder_input=
Variable(torch.LongTensor(1,input_batch.shape
[1])).
```

```
fill_(output_lang.word_to_index.get("SOS"))
```

#### h. Pelatihan Model

Dalam tahap ini, akan dilakukan pelatihan model *Encoder-Decoder*. Setelah model dibuat sebelumnya, akan dilakukan pelatihan agar model tersebut dapat melakukan penerjemahan bahasa Indonesia dan bahasa Besemah. Pelatihan yang dilakukan menggunakan 100 *epoch*

#### Hasil dari Pelatihan Model

Reading lines...  
Read 5104 sentence pairs  
Trimmed to 4615 sentence pairs  
Counting words...  
Train pairs: 4615

Test pairs: 0  
Counted Words -> Trimmed Vocabulary Sizes  
(w/ EOS and SOS tags):  
ind, 3388 -> 3391  
bs, 4092 -> 4095

Train Pairs #  
4615  
Current mem usage:  
0.0  
Current mem usage:  
0.0  
Encoder and Decoder Created  
Current mem usage:  
0.0  
Cuda being used  
Number of epochs: 100

#### i. Penerjemahan

Pada proses hasil penerjemahan ini dapat menghasilkan proses pengubahan teks dari bahasa Besemah ke bahasa Indonesia dan sebaliknya. penerjemahan dari bahasa inti agar pesan dan makna dapat dipahami dengan baik oleh pembaca yang menerjemahkan dari bahasa tersebut.

#### Hasil dari Penerjemahan

##### Bahasa Besemah

Dengah tunggulah ghuma kami kah ke kebun  
Kamu keciq dide ndaq ngikut  
Maruq ngingun ayam beeruge  
Sulha ngambiq ayiq di kambang  
Die ngudut siung cap jambu di kebun  
Rani nymabal caluq  
Bani nampit celane

##### Bahasa Indonesia

Kamu jagalah rumah kami mau ke kebun  
Kamu anak-anak tidak usah ikut  
Maruk memelihara ayam hutan  
Sulha menimbah air di sumur  
Dia merokok keretek cap jambu di kebun  
Rani membuat sambal terasi  
Bani menambal celana

## 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] Abidin, Z. (2017). Penerapan Neural Machine Translation untuk Eksperimen Penerjemahan secara Otomatis pada Bahasa Lampung - Indonesia. *Prosiding Seminar Nasional Metode Kuantitatif*, 978, 53-68.
- [2] Afria, R., Sanjaya, D., & Tiara, M. (2020). Leksikostatistik dan Grotokronologi Bahasa Melayu Palembang, Basemah Lahat, Basemah Pagaralam, dan Kayu Agung: Kajian Linguistik Historis Komparatif. *Madah*, 11(1), 27-42. <https://doi.org/10.31503/madah.v11i1.223>
- [3] AKBAR, G. (2021). *Automatic Text Summarization Berita Bahasindonesia Menggunakan Metode Attentional Encoder Decoder*. <https://dspace.uui.ac.id/handle/123456789/37691><https://dspace.uui.ac.id/bitstream/handle/123456789/37691/14523091.pdf?sequence=1>
- [4] Amalia Yunia Rahmawati. (2020). BIAS GENDER DALAM MESIN PENERJEMAHAN (GOOGLE TRANSLATE) DARI BAHASA INDONESIA KE BAHASA INGGRIS. July, 1-23.
- [5] Aristyanto, M. Y., & Kurniawan, R. (2021). Pengembangan Metode Neural Machine Translation Berdasarkan Hyperparameter Neural Network. *Seminar Nasional Official Statistics, 2021*(1), 935-946. <https://doi.org/10.34123/semnasoffstat.v2021i1.789>
- [6] Bloom, N., & Reenen, J. Van. (2013). 濟無 No Title No Title No Title. *NBER Working Papers*, 89. <http://www.nber.org/papers/w16019>
- [7] Darwis, R., Sujaini, H., & Nyoto, R. D. (2019). Peningkatan Mesin Penerjemah Statistik dengan Menambah Kuantitas Korpus Monolingual (Studi Kasus : Bahasa Indonesia - Sunda). *Jurnal Sistem Dan Teknologi Informasi (JUSTIN)*, 7(1), 27. <https://doi.org/10.26418/justin.v7i1.27254>
- [8] Fauziyah, Y., Ilyas, R., & Kasyidi, F.

- (2022). Mesin Penterjemah Bahasa Indonesia-Bahasa Sunda Menggunakan Recurrent Neural Networks. *Jurnal Teknoinfo*, 16(2), 313.  
<https://doi.org/10.33365/jti.v16i2.1930>
- [9] HENUARIANTO, I. (2023). MESIN PENERJEMAH BAHASA KOMERING KE BAHASA INDONESIA MENGGUNAKAN ALGORITMA RNN EMBEDDING [PhD Thesis]. UNIVERSITAS BINA DARMA.
- [10] Irfan, M. (2022). Named Entity Recognition Untuk Data Review Tempat Wisata Dengan Metode "Bidirectional Encoder Representations from Transformers." *Universitas Islam Indonesia*.
- [11] Musdholifah, A. (2022). EVALUASI PENGARUH MEKANISME ATTENTION TERHADAP MODEL MESIN PENERJEMAH BAHASA JAWA-INDONESIA BERBASIS NEURAL MACHINE TRANSLATION SHERINE DEVI SUTOMO,
- [12] Aina Musdholifah, S.Kom., M.Kom., Ph.D. Nasution, A. H., Rizki, Y., Nasution, S., & Muhammad, R. (2019). Mesin Penerjemah Interaktif Dengan Animasi 3D Berbasis Augmented Reality. *It Journal Research and Development*, 4(1), 28-39.  
<https://doi.org/10.25299/itjrd.2019.vol>
- [13] Ronja, R., Sujaini, H., & Nyoto, R. D. (2020). Perbandingan Nilai Akurasi Algoritma Smoothing pada Mesin Penerjemah Statistik Bahasa Indonesia Bahasa Melayu Sambas dengan Language Model Toolkit IRSTLM. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 6(3), 291.  
<https://doi.org/10.26418/jp.v6i3.42471>
- [14] Razsiah, F. R. F. (2023). Aplikasi Penerjemah Bahasa Bangka-Indonesia-Inggris Berbasis Website Dengan Neural Machine Translation (Nmt). <http://repository.polman-babel.ac.id/id/eprint/710/1/Laporan-Proyek-Akhir-Fadel-Razsiah-TRPL.pdf>
- [15] Wardiah, D. (2021). Pronomina Bahasa Besemah Dialek Tanjung Periuk Kecamatan Gumay Talang Kabupaten Lahat Sebagai Pemertahanan Bahasa Daerah Dalam Menjaga Kearifan Lokal. *Prosiding Seminar Nasional Pendidikan Program Pascasarjana Universitas PGRI Palembang*, 22-30.
- [16] Wismoyo, P. A. (2018). Mesin penerjemah bahasa inggris - indonesia berbasis jaringan saraf tiruan dengan mekanisme attention menggunakan arsitektur transformer skripsi. *Undip E-Journal System*.
- [17]

## SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini:

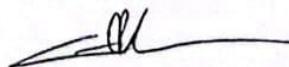
Nama : Bogy Dharma Sandi  
Nim : 201410145  
Program Studi : Sistem Informasi  
Fakultas : Sains Teknologi  
No. WA : 081279345430  
Nama Pembimbing : Susan Dian Purnamasari, M.Kom  
Judul Artikel : Mesin Penerjemahan Bahasa Besemah Berbasis Machine Learning Dengan Algoritma Model Encoder-Decoder

Menyatakan memang benar belum mendapatkan *Letter of Acceptance* (LoA) dan masih tahap *submit/menunggu proses review* dari pihak penerbit jurnal. Mengingat pendaftaran wisuda sedang berlangsung, untuk itu saya mohon dapat diizinkan mendaftar wisuda walaupun belum mendapatkan LoA, dengan konsekuensi tidak mendapatkan Transkrip Akademik saya. Saya secara sadar tidak akan menuntut Transkrip Akademik saya sebelum saya mendapatkan LoA dan mengumpulkan ke Pusat Pelayanan Mahasiswa (PPM).

Demikian surat pernyataan ini saya buat dengan sebenarnya untuk dipergunakan sebagaimana mestinya. Terima kasih.

Mengetahui,  
Ketua Program Studi

Palembang, 9 September 2024  
Hormat saya,



Nita Rosa Damayanti, M.Kom., Ph.D



Bogy Dharma Sandi

Lampiran:  
Bukti submit artikel

Rabit : Jurnal Teknologi dan Sistem Informasi Univrab ▾ Tasks 0 English View Site bogydharma sandi

RABIT  
Jurnal Teknologi dan Sistem Informasi Univrab

Submission Library View Metadata

MESTIN PENERJEMAHAN BAHASA BESEMAH BERBASIS MACHINE LEARNING DENGAN ALGORITMA MODEL ENCODER-DECODER

Bogy Dharma Sandi

Submission Review Copyediting Production

Submission Files Search

19621-1 bogydharma sandi, Jurnal Bogy Revisi FIX.docx Article Text

Download All Files

Pre-Review Discussions Add discussion

Name	From	Last Reply	Replies	Closed
No Items				