



**MESIN PENERJEMAH BAHASA PALEMBANG KE BAHASA INDONESIA
MENGUNAKAN NO LANGUAGE LEFT BEHIND**

LAPORAN PENELITIAN

**MUHAMMAD FINALDO
201410129**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS TEKNOLOGI
UNIVERSITAS BINA DARMA
2024**



**MESIN PENERJEMAH BAHASA PALEMBANG KE BAHASA INDONESIA
MENGUNAKAN NO LANGUAGE LEFT BEHIND**

LAPORAN PENELITIAN

MUHAMMAD FINALDO

201410129

**Laporan penelitian ini diajukan sebagai syarat memperoleh gelar
Sarjana Komputer**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS TEKNOLOGI
UNIVERSITAS BINA DARMA
2024**

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

**MESIN PENERJEMAH BAHASA PALEMBANG KE BAHASA INDONESIA
MENGUNAKAN NO LANGUAGE LEFT BEHIND**

**MUHAMMAD FINALDO
201410129**

**Telah diterima sebagai salah satu syarat untuk memperoleh gelar
Sarjana Komputer pada program Studi Sistem Informasi**

**Palembang, 14 Maret 2025
Fakultas Sains dan Teknologi
Universitas Bina Darma**

Dosen Pembimbing



Dr. Yesi Novaria Kunang, S.T., M.Kom

Dekan



Universitas Bina Darma
Fakultas Sains Teknologi

Dr. Tata Sutabri., MMSI., MKM.

HALAMAN PERSETUJUAN

HALAMAN PERSETUJUAN

Laporan penelitian ini berjudul "MESIN PENERJEMAH BAHASA PALEMBANG KE BAHASA INDONESIA MENGGUNAKAN NO LANGUAGE LEFT BEHIND" yang disusun oleh "Muhammad Finaldo", telah dipertahankan di depan komisi penguji

Komisi Penguji

Ketua : Dr.Yesi Novaria Kunang, S.T.,M.Kom. (.....)

Anggota : Susan Dian Purnamasari, M.Kom. (.....)

Anggota : Nia Oktaviani, M.Kom. (.....)

Mengetahui,
Program Studi Sistem Informasi
Universitas Bina Darma
Ketua,

Universitas Bina Darma
Fakultas Sains Teknologi

Dr.Ari Muzakir,S.kom.,M.Cs
NIDN. 0223128701

SURAT PERNYATAAN

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini :

Nama : Muhammad Finaldo

Nim : 201410129

Dengan ini menyatakan bahwa :

1. Karya tulis saya, yang merupakan hasil riset, adalah karya yang sepenuhnya orisinal dan belum pernah diajukan untuk memperoleh gelar akademik (Sarjana) di Universitas Bina Darma maupun perguruan tinggi lainnya.
2. Karya tulis ini merupakan hasil pemikiran, gagasan, rumusan dan penelitian saya, dengan bimbingan dari tim pembimbing.
3. Dalam karya tulis ini, tidak ada karya atau pendapat yang berasal dari orang lain yang telah dipublikasikan, kecuali bila dinyatakan secara jelas dengan dikutip tertulis yang mencantumkan nama pengarang dan disertakan dalam daftar rujukan.
4. Saya bersedia untuk tugas skripsi saya, yang akan diperiksa keasliannya menggunakan perangkat plagiarism checker dan diunggah ke internet, sehingga dapat diakses secara daring.
5. Saya menulis surat pernyataan ini dengan sungguh-sungguh. Jika nanti bukti saya melakukan penyimpangan atau ketidak benaran dalam pernyataan ini, saya siap menerima sanksi sesuai dengan peraturan dan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat agar dapat dipergunakan sebagaimana mestinya

Palembang, 8 Maret 2025

Yang membuat pernyataan



Muhammad Finaldo

201410129

HALAMAN MOTTO DAN PERSEMBAHAN

MOTTO :

“Allah tidak akan membebani mereka melainkan sesuai dengan kadar kesanggupannya”

(Q.S Al-Baqarah :286)

“Terlambat Bukan Berarti Gagal,Cepat Bukan Berarti Hebat. Terlambat bukan menjadi alasan untuk menyerah, setiap orang memiliki proses yang berbeda.

PERCAYA PROSES itu yang paling penting, Karena Allah telah mempersiapkan Hal Baik dibalik Proses yang kamu anggap rumit”

PERSEMBAHAN :

Laporan penelitian ini dipersembahkan sebagai ungkapan syukur kepada Allah SWT. yang telah memberikan rahmat-Nya kepada penulis dan kepada kedua orang tua dan kakak-kakak serta sahabat-sahabat yang selalu membantu, mendukung dan mendoakan penulis selama proses menempuh pendidikan sarjana.

ABSTRACT

Indonesia is a country with a vast diversity of languages, one of which is the Palembang language, popular in South Sumatra. This language has great potential to be developed into a machine translator to facilitate communication between Palembang language speakers and Indonesian language speakers. This will be very beneficial for tourists visiting areas in South Sumatra, ensuring they do not face difficulties in communication. The importance of preserving the Palembang language cannot be overlooked, as it is one of Indonesia's cultural treasures. With rapidly and sustainably advancing technology, the demand for technology to accelerate progress and development continues to grow. One such technology is artificial intelligence (AI). Machine learning, as a part of AI, can be utilized to create an effective machine translator. To translate the Palembang language into Indonesian, researchers use the No Language Left Behind (NLLB) algorithm. This algorithm works by analyzing input documents in the Palembang language and generating translations for each word. The translations are then arranged into accurate and easily understandable text in Indonesian.

Keywords: Palembang Language, No Language Left Behind, Machine Translation.

ABSTRAK

Indonesia merupakan negara yang memiliki keberagaman bahasa yang sangat banyak, salah satunya adalah Bahasa Palembang yang populer di Sumatera Selatan. Bahasa ini memiliki potensi besar untuk dikembangkan menjadi mesin penerjemah, guna memudahkan komunikasi antara penutur Bahasa Palembang dengan penutur Bahasa Indonesia. Ini akan sangat bermanfaat bagi wisatawan yang berkunjung ke daerah-daerah di Sumatera Selatan, sehingga mereka tidak akan mengalami kesulitan dalam berkomunikasi. Pentingnya pelestarian Bahasa Palembang tidak dapat diabaikan, karena merupakan salah satu kekayaan budaya Indonesia. Dengan teknologi yang berkembang pesat dan berkelanjutan, kebutuhan akan teknologi untuk mempercepat kemajuan dan pembangunan semakin meningkat. Salah satu teknologi yang dapat digunakan adalah kecerdasan buatan (AI). *Machine learning* adalah sebagai bagian dari AI, dapat digunakan untuk membuat mesin penerjemah yang efektif. Untuk menerjemahkan Bahasa Palembang ke Bahasa Indonesia, peneliti menggunakan algoritma No Language Left Behind (NLLB). Algoritma ini bekerja dengan cara mempelajari dokumen input dalam Bahasa Palembang dan menghasilkan terjemahan untuk setiap kata. Kemudian, hasil terjemahan tersebut disusun menjadi teks dalam Bahasa Indonesia yang akurat dan mudah dipahami.

Kata Kunci: Bahasa Palembang, *No Language Left Behind*, Mesin Penerjemah.

KATA PENGANTAR

Puji syukur kita panjatkan atas kehadiran Allah SWT, yang telah memberikan keimanan, kesempatan, kesehatan, kekuatan serta berkah yang luar biasa atas rahmat-Nya sehingga penulis mampu menyelesaikan penulisan laporan penelitian dengan judul mengenai “MESIN PENERJEMAH BAHASA PALEMBANG KE BAHASA INDONESIA MENGGUNAKAN NO LANGUAGE LEFT BEHIND”. Laporan Penelitian ini dituliskan sebagai wujud untuk memenuhi syarat untuk memperoleh gelar sarjana (S1) dari Program Studi Sistem Informasi, Fakultas Sains Teknologi Universitas Bina Darma Palembang.

Pada kesempatan ini, dengan segala kerendahan hati penulis mengucapkan terimakasih kepada semua pihak atas bantuan, bimbingan saran serta arahan yang telah diberikan dalam penulisan laporan penelitian ini, antara lain:

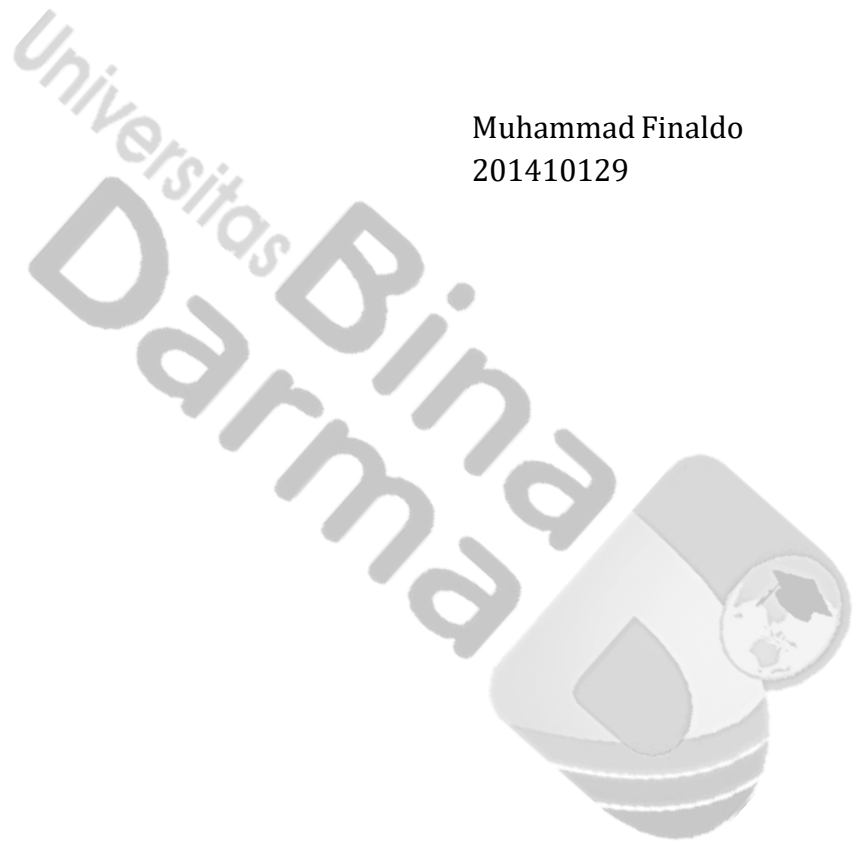
1. Kedua orang tua yang sangat saya cintai, yaitu Mama dan Papa. Terimakasih banyak yang sebesar besarnya atas semua pengorbanan, keringat dan kasih sayang yang selalu mereka berikan untuk saya. Saya juga sangat menghargai segala motivasi, dukungan yang tiada henti, serta doa-doa yang selalu dipanjatkan oleh mama dan papa.
2. Prof. Dr. Sunda Ariana, M.Pd., M.M. selaku Rektor Universitas Bina Darma Palembang.
3. Dr. Tata Sutabri, S.Kom., MMSI., MKM. selaku Dekan Fakultas Sains Teknologi.
4. Nita Rosa Damayanti, M. Kom., Ph.D. selaku Ketua Program Studi Sistem Informasi.
5. Dengan hormat, saya ingin mengucapkan terima kasih kepada Dr. Yesi Novaria Kunang, S. T. , M. Kom. , selaku pembimbing saya dan Ketua ISRG (Intelligent System Research Group). Beliau telah memberikan bimbingan dan arahan yang berarti sejak awal hingga akhir dalam proses menyusun laporan penelitian ini. Saya sangat menghargai semua ilmu yang telah diberikan.

6. Seluruh dosen pengajar program studi Sistem Informasi yang telah memberikan ilmu, pengetahuan, dan wawasan selama saya menempuh pendidikan di Universitas Bina Darma.
7. kedua kakak saya Kak Tia dan Kak Nanda. yang selalu menjadi sumber inspirasi dan motivasi dalam setiap langkahku. Terima kasih atas doa, dukungan, dan semangat yang tiada henti. Nasihat dan keberanianmu telah membantuku menghadapi setiap tantangan, dan kehadiranmu selalu menjadi penguat di saat aku merasa lelah dan ragu.
8. Abangku, Tito Nugraha. yang selalu menjadi panutan dan inspirasi. Terima kasih atas segala dukungan, semangat, dan nasihat yang tak pernah putus.
9. Penulis ingin menyampaikan rasa terima kasih yang tulus kepada seluruh sahabat-sahabat Kosang, yang telah setia mendampingi dan memberikan semangat yang selalu ada di setiap suka dan duka. Dukungan, candaan, dan semangatmu yang tak pernah pudar. Kebersamaan kita selama ini menjadi kekuatan besar bagiku untuk tetap bertahan dan menyelesaikan laporan penelitian ini.
10. Dan yang terakhir, kepada diri saya sendiri Muhammad Finaldo saya ingin mengucapkan terima kasih kepada diri saya sendiri. Terima kasih telah berusaha keras, tidak menyerah di tengah jalan, dan terus berjuang meski banyak rintangan yang harus dihadapi. Keberanian untuk memulai, kesabaran untuk bertahan, dan keteguhan untuk menyelesaikan adalah hal-hal yang tidak mudah. Saya bersyukur atas setiap detik waktu yang dihabiskan, setiap usaha yang dilakukan, dan setiap langkah kecil yang diambil menuju penyelesaian karya ini. Terima kasih telah percaya pada diri sendiri dan terus berusaha memberikan yang terbaik.

Demikianlah kata pengantar yang dapat diucapkan jika ada kesalahan penulisan saya mohon maaf. Sekian Terima Kasih.

Palembang, Maret 2025

Muhammad Finaldo
201410129



DAFTAR ISI

<u>HALAMAN PENGESAHAN</u>	iii
<u>HALAMAN PERSETUJUAN</u>	iv
<u>SURAT PERNYATAAN</u>	v
<u>HALAMAN MOTTO DAN PERSEMBAHAN</u>	vi
<u>ABSTRACT</u>	vii
<u>ABSTRAK</u>	viii
<u>KATA PENGANTAR</u>	ix
<u>DAFTAR ISI</u>	xii
<u>DAFTAR GAMBAR</u>	xiv
<u>DAFTAR TABEL</u>	xv
<u>BAB 1 PENDAHULUAN</u>	1
1.1 <u>Latar Belakang</u>	1
1.2 <u>Rumusan Masalah</u>	2
1.3 <u>Tujuan Penelitian</u>	2
1.4 <u>Batasan Masalah</u>	2
1.5 <u>Manfaat Penelitian</u>	3
<u>BAB II TINJAUAN PUSTAKA</u>	4
2.1. <u>Landasan Teori</u>	4
2.1.1 <u>Mesin Penerjemah</u>	4
2.1.2 <u>Neural Machine Translation (NMT)</u>	4
2.1.3 <u>Machine Learning (ML)</u>	4
2.1.4 <u>No Language Left Behind (NLLB)</u>	5
2.2 <u>Penelitian Terdahulu</u>	6
<u>BAB III METODOLOGI PENELITIAN</u>	9
3.1 <u>Pengumpulan Data</u>	10
3.2 <u>Pre – Processing</u>	10

3.2.2	<u>PrepareLanguages</u>	10
3.2.2	<u>Normalisasi</u>	10
3.2.3	<u>Tokenisasi</u>	10
3.3	<u>Pelatihan Model</u>	11
3.4	<u>Push Hugging Face</u>	11
3.5	<u>Evaluasi Model</u>	11
3.6	<u>Terjemahan</u>	11
<u>BAB IV HASIL DAN PEMBAHASAN</u>		12
4.1	<u>Pengumpulan Data</u>	12
4.2	<u>Preprocessing</u>	13
4.2.1	<u>PrepareLanguages</u>	13
4.2.2	<u>Tokenisasi</u>	16
4.2.3	<u>Normalisasi</u>	21
4.3	<u>Pelatihan Model</u>	21
3.4	<u>Push Hugging Face</u>	22
4.5	<u>Evaluasi Model</u>	22
4.6	<u>Terjemahan</u>	24
<u>BAB V PENUTUP</u>		25
5.1	<u>Kesimpulan</u>	25
5.2	<u>Saran</u>	25
<u>DAFTAR PUSTAKA</u>		26
<u>LAMPIRAN</u>		27

DAFTAR GAMBAR

Gambar 3. 1 Flowchart.....	9
Gambar 4. 1 Tampilan Dataset.....	12
Gambar 4. 2 Tampilan File JSON.....	13
Gambar 4. 3 Kode PrepareLanguages.....	13
Gambar 4. 4 Kode Membaca File JSON.....	14
Gambar 4. 5 Kode Memproses Dalam Data JSON.....	15
Gambar 4. 6 Kode Konversi JSON ke DataFrame	16
Gambar 4. 7 Kode mengimpor tokenisasi.....	16
Gambar 4. 8 Kode Menginisialisasi Tokenizer	17
Gambar 4. 9 Kode Mendefinisikan tokenizer.....	17
Gambar 4. 10 Kode Sampling Dataset.....	18
Gambar 4. 11 Kode Mengidentifikasi token.....	19
Gambar 4. 12 Kode Menambahkan Bahasa Baru.....	19
Gambar 4. 13 Kode Mendukung Bahasa Baru	19
Gambar 4. 14 Kode Ukuran Menyesuaikan Embedding	20
Gambar 4. 15 Kode Penyesuaian dan Inisialisasi Embedding Token Baru	20
Gambar 4. 16 Kode Normalisasi	21
Gambar 4. 17 Kode Pelatihan Model.....	21
Gambar 4. 18 Kode Push Huggingface.....	22
Gambar 4. 19 Kode Evaluasi Model.....	22
Gambar 4. 20 Output Evaluasi Model	23
Gambar 4. 21 Terjemahan	24

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu.....	7
---	---

