



**MODEL MESIN PENERJEMAH BAHASA BESEMAH BERBASIS MACHINE
LEARNING DENGAN ALGORITMA MODEL RECURRENT NEURAL
NETWORK (RNN)**

LAPORAN PENELITIAN

MUHAMAD ANDIKA

191410165

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

2025



**MODEL MESIN PENERJEMAH BAHASA BESEMAH BERBASIS MACHINE
LEARNING DENGAN ALGORITMA MODEL RECURRENT NEURAL
NETWORK (RNN)**

MUHAMAD ANDIKA

191410165

**Laporan Penelitian ini diajukan sebagai syarat untuk memperoleh gelar
Sarjana Komputer**

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS SAINS TEKNOLOGI

UNIVERSITAS BINA DARMA

PALEMBANG

2025

HALAMAN PENGESAHAN

**MODEL MESIN PENERJEMAH BAHASA BESEMAH BERBASIS MACHINE
LEARNING DENGAN ALGORITMA MODEL RECURRENT NEURAL
NETWORK (RNN)**

MUHAMAD ANDIKA
191410165

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

Dosen Pembimbing

Palembang, 28 February 2024
Program Studi Sistem Informasi
Dekan Fakultas,


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


Dr. Tata Sutabri, MMSI, MKM.

HALAMAN PERSETUJUAN

Penelitian berjudul "MODEL MESIN PENERJEMAH BAHASA BESEMAH BERBASIS MACHINE LEARNING DENGAN ALGORITMA MODEL RECURRENT NEURAL NETWORK (RNN)" oleh MUHAMAD ANDIKA telah dipertahankan didepan komisi penguji pada hari Rabu tanggal 28 February 2024.

Komisi Penguji

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

(.....)

2. Anggota : Assoc. Prof. Leon A. Abdillah, S.Kom., M.M., M.TA, CDM,)

(.....)

3. Anggota : Irman Effendy, M.Kom.

(.....)

Mengetahui,

Program Studi Sistem Informasi

Universitas Bina Darma

Ketua,

Universitas Bina Darma
Fakultas Sains Teknologi

Dr. Ari Muzakir, S.Kom., M.Cs.

SURAT PERNYATAAN

Saya yang bertanda tangan bawah ini :

Nama : MUHAMAD ANDIKA

NIM : 191410165

Dengan ini menyatakan bahwa :

1. Karya tulis saya (Riset) adalah asli dan belum pernah diajukan untuk mendapat gelar akademik (Sarjana) di Universitas Bina Darma atau perguruan tinggi lainnya ;
2. Karya tulis ini murni gagasan, rumusan dan penelitian saya dengan arahan tim pembimbing ;
3. Di dalam karya tulis ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dikutip dengan mencantumkan nama pengarang dan memasukkan ke dalam daftar rujukan ;
4. Saya bersedia tugas skripsi, di cek keasliannya menggunakan plagiarism checker serta di unggah ke internet, sehingga dapat diakses secara daring;
5. Surat pernyataan ini saya tulis dengan sungguh-sungguh dan apabila terbukti melakukan penyimpanan atau ketidakbenaran dalam pernyataan ini maka saya bersedia menerima sanksi dengan peraturan dan perundang- undangan yang berlaku ;

Demikian surat pernyataan ini saya buat agar dapat dipergunakan sebagaimana mestinya.

Palembang, February 2024

Yang membuat pernyataan,



MUHAMAD ANDIKA
191410165

ABSTRAK

Indonesia terdiri atas berbagai suku dengan Bahasa daerahnya masing-masing salah satunya ialah suku Besemah di provinsi Sumatera Selatan dengan budaya bahasanya yaitu Bahasa Besemah. Hingga saat ini Bahasa Besemah masih digunakan oleh masyarakat suku Besemah namun seiring berjalannya waktu jumlah penutur Bahasa Besemah ini semakin berkurang belum lagi kebanyakan masyarakat luas tidak mengetahui apa itu Bahasa Besemah. Mesin penerjemah adalah sebuah alat yang dapat melakukan peralihan bahasa satu ke bahasa lain. Penelitian ini bertujuan untuk mengumpulkan dataset berupa kalimat maupun kata dari Bahasa Besemah kemudian membuat mesin penerjemah Bahasa Besemah ke Bahasa Indonesia dan sebaliknya. Dalam penelitian ini, diterapkan teknologi Neural Machine Translation (NMT) dengan pendekatan Recurrent Neural Network (RNN). Hasil untuk *val_accuracy* besemah-indonesia adalah 0,8469 dan untuk Indonesia-besemah mendapatkan nilai *val_accuracy* 0.8492, pada uji coba penerjemahan yang dilakukan menggunakan model RNN, 100 epoch, 64 batch size dan 0.2 validation split.

Kata Kunci: Neural Machine Translation, Bahasa Besemah, Recurrent Neural Network.

KATA PENGANTAR

Penulis ingin menyampaikan ungkapan terima kasih dan rasa syukur kepada Allah SWT, yang dengan rahmat dan karunia-Nya, penulis berhasil menyelesaikan laporan penelitian ini. Harapannya, laporan penelitian ini dapat memberikan manfaat bagi semua pembaca. Proses pembuatan laporan ini juga merupakan salah satu persyaratan untuk memperoleh gelar Sarjana Komputer (S.Kom.) dari Fakultas Sains Teknologi, Program Studi Informasi di Universitas Bina Darma.

Penulis juga ingin menyampaikan apresiasi yang tak terhingga kepada semua pihak dan teman-teman yang telah membantu dalam penyusunan laporan penelitian ini. Terima kasih yang tulus penulis sampaikan kepada:

1. Bu Rektor Universitas Bina Darma Prof. Dr. Sunda Ariana, M.Pd., M.M.
2. Bu Dr. Yesi Novaria Kunang, S.T., M.Kom. dan bapak Ilman Zuhri Yadi, M.M., M.Kom. selaku dosen pembimbing yang senantiasa membantu, menasehati dan membimbing saya selama penelitian.
3. Pak Assoc. Prof. Leon A. Abdillah, S.Kom., M.M., MTA, CDM dan pak Irman, Effendy, M.Kom. selaku dosen penguji yang telah membimbing dan menasehati saya.
4. Kedua orang tua serta adik saya dan orang tercinta yang memberikan doa dan dukungan selama berkuliah dan menyelesaikan laporan akhir penelitian ini.
5. Seluruh tim perpustakaan Universitas Bina Darma yang telah membantu penulis dalam menyelesaikan tugas akhir Laporan Penelitian ini.
6. Seluruh rekan – rekan yang terlibat dalam penelitian ini.

Penulis menutup kata pengantar dengan mengungkapkan permohonan maaf yang tulus apabila terdapat kesalahan atau kekeliruan. Sekian.

Palembang February 2024

Muhamad Andika



DAFTAR ISI

HALAMAN PENGESAHAN.....	ii
HALAMAN PERSETUJUAN	iii
SURAT PERNYATAAN	iv
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
DAFTAR LAMPIRAN	xii
BAB I	xii
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah	3
1.5 Manfaat Penelitian.....	4
BAB II.....	5
2.1 Mesin Penerjemah	5
2.2 Neural Machine Translation	5
2.3 Natural Language Toolkit (NLTK)	6
2.4 Reccurent Neural Network.....	6
2.5 Peneltian Sebelumnya.....	8
BAB III	12
3.1 Alat dan Bahan	12
3.2 Flowchart	14
3.3 Pengumpulan data.....	15
3.4 Dataset.....	15
3.5 Pra-Pemrosesan.....	15
3.5.1 Penghapusan Special Characters.....	15
3.5.2 Normalisasi.....	15
3.5.3 Tokenisasi	15
3.5.4 Padding.....	16

3.6	Model RNN	16
3.7	Pelatihan Model	16
3.8	Evaluasi Model	16
BAB IV		17
4.1	Pengumpulan Data	17
4.2	Dataset.....	17
4.3	Pra-Pemrosesan.....	18
4.3.1	Remove Special Characters	19
4.3.2	Normalisasi.....	19
4.3.3	Tokenisasi	20
4.3.4	Padding.....	21
4.4	Model RNN.....	21
4.5	Pelatihan Model.....	23
4.6	Evaluasi Model	24
BAB V		26
5.1	Kesimpulan.....	26
5.2	Saran	26
DAFTAR PUSTAKA		27
LAMPIRAN		29

DAFTAR GAMBAR

Gambar 3. 1 Alur penelitian	14
Gambar 4. 1. Data hasil scan kamus berbentuk file excel.....	17
Gambar 4. 2. Dataset bahasa Indonesia berbentuk txt.....	18
Gambar 4. 3. Dataset bahasa Besemah berbentuk txt.....	18
Gambar 4. 4. Sebelum proses remove special characters.....	19
Gambar 4. 5. Setelah proses remove special characters.	19
Gambar 4. 6. Sebelum proses normalisasi	20
Gambar 4. 7. Setelah proses normalisasi	20
Gambar 4. 8. Hasil tokenisasi.....	21
Gambar 4. 9. Hasil padding	21
Gambar 4. 10. Kode untuk membuat model RNN bahasa indonesia- besemah	22
Gambar 4. 11. Kode untuk membuat model RNN bahasa indonesia- besemah	22
Gambar 4. 12. Hasil pelatihan model bahasa indonesia-besemah	23
Gambar 4. 13. Hasil pelatihan model bahasa besemah-indonesia	24
Gambar 4. 14. Hasil uji coba penerjemahan indonesia-besemah.....	24
Gambar 4. 15. Hasil uji coba penerjemahan besemah -indonesia	24

DAFTAR TABEL

Tabel 2.1 Penelitian terdahulu.....	10
--	-----------



DAFTAR LAMPIRAN

Lampiran 1 SK Pembimbing.....	29
Lampiran 2 Permohonan Judul.....	30
Lampiran 3 Lembar Konsultasi Semhas	31
Lampiran 4 Lembar Konsultasi Sempro	32
Lampiran 5 Surat Keterangan Lulus.....	33
Lampiran 6 Hasil Turnitin	34
Lampiran 7 Formulir Perbaikan Seminar Hasil.....	38
Lampiran 8 Formulir Kelayakan Jilid Skripsi	39

