



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

LAPORAN PENELITIAN

BOGY DHARMA SANDI

201410145

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS SAINS TEKNOLOGI

UNIVERSITAS BINA DARMA

2024



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

BOGY DHARMA SANDI

201410145

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

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS SAINS TEKNOLOGI

UNIVERSITAS BINA DARMA

2024

HALAMAN PENGESAHAN

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

BOGY DHARMA SANDI

201410145

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

Palembang, 19 Agustus 2024
Fakultas Sains Teknologi
Universitas Bina Darma
Dekan

Pembimbing



Susan Dian Purnamasari, M.Kom



Universitas Bina Darma
Fakultas Sains Teknologi

Dr. Tata Sutabri, S.Kom., MMSI., MKM

HALAMAN PERSETUJUAN

Skripsi Berjudul "MESIN PENERJEMAHAN BAHASA BESEMAH BERBASIS MACHINE LEARNING DENGAN ALGORITMA MODEL ENCODER-DECODER Oleh "Bogy Dharma Sandi",telah dipertahankan di depan komisi penguji pada hari Senin tanggal 19 Agustus 2024

Komisi Penguji

1. Ketua : Susan Dian Purnamasari,M.Kom.


(.....)

2. Anggota : Fatmasari,S.Kom.,M.M.,M.Kom.


(.....)

3. Anggota : Muhamad Ariandi,S.Kom.,M.Kom.


(.....)

Mengetahui,
Program Studi Sistem Informasi
Fakultas Sains Teknologi
Universitas Bina Darma
Ketua,




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

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini :

Nama : BOGY DHARMA SANDI

Nim : 201410145

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, 05 Agustus 2024

Yang membuat pernyataan




BOGY DHARMA SANDI
201410145

MOTTO DAN PERSEMBAHAN

Motto

"Hidup bukan tentang dipukul,lalu kau balas,tapi tentang bagaimana kau mampu berdiri tegak walau banyak pukulan yang kau terima. Berani hidup,harus berani mati. Takut mati jangan hidup,Takut hidup mati saja. Karena Hidup itu hanya sekali maka hidupla yang penuh arti"

(Bogy LH Kenantan)

Persembahan

Laporan penelitian ini dipersembahkan untuk :

1. Seorang perempuan cantik yang memiliki mahkota dari surga,dan yang mempertaruhkan nyawa untuk melahirkan aku ke dunia,kemudian membesarkanku dengan penuh kasih sayangnya dan perempuan itu ialah ibuku,Mama panggilannya.Terimakasih banyak untuk mama ku,perempuan tangguh yang selalu bekerja keras dan tak pernah kenal lelah untuk mewujudkan anak-anaknya menjadi sarjana,maaf jika anakmu ini,belum bisa membalas kebaikanmu Ma..
Suatu saat nanti anakmu ini pasti akan membuatmu bangga dengan apa yang sudah sejauh ini dia jalani ma..
2. Seorang Laki-laki yang kuat dan gagah berani dengan segala jerih payahnya untuk memberikan yang terbaik untuk keluarganya dan mengajarkan arti kehidupan dan banyak hal lainnya.Laki-laki itu ialah Ayahku atau aku sering menyebutnya Papa.sejak SD Kelas IV saya menjalani hidup tanpa adanya sosok dia,akan tetapi saya selalu percaya dan yakin, karena robot buatannya ini sangat kuat dan tahan disegala kondisi apapun,dengan didewasakannya saya dengan keadaan, sehingga saya banyak belajar tentang kehidupan yang penuh lika-liku ini. Terimakasih Pa..

"Aku dan Papa

Adalah pasangan yang gagal,aku gagal membahagikan Papa sampai tua,dan Papa gagal menemani aku beranjak dewasa.

Akan tetapi aku percaya Papa selalu ada didalam diriku walaupun kita terpisah jarak dan waktu dan tidak bisa ditempu dengan apapun itu..

3. Ketiga aku ucapkan juga untuk saudara dan saudari ku yang memberikan do'a dan semangat untuk adik kecil,adik bungsu satunya ini,do'a kan saja adik kecilmu ini bisa membuat bangga nama baik keluarga kita dan juga bisa mengangkat derajat keluarga kita.Aamiin.
4. Keempat terimakasih juga aku ucapkan untuk teman-teman seperjuanganku yang telah mengabdikan waktunya dan memperbolehkan saya untuk makan minum tidur dikosannya.semoga nanti kita bisa menjadi orang yang sukses dan membuat bangga keluarga kita masing-masing.jika sudah sukses jangan lupakan pertemanan kita yang sudah kita jalani sampai saat ini.
5. Terimakasih banyak untuk diriku sendiri. Bogy Dharma Sandi.LH
Sudah bertahan sampai sejauh ini dengan segala rintangan yang dihadapi jangan pernah menyerah karena perjalanan yang sesungguhnya baru saja akan kita mulai..
Tetap tegar dan semangat banyak rintangan dan cobaan didepan sana yang sudah menunggu kita.

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, Teknologi berkembang pesat dan berkelanjutan seiring perkembangan zaman. Teknologi dibutuhkan oleh semua orang untuk mempercepat kemajuan dan pembangunan. Salah satu contoh teknologi yang dapat digunakan adalah kecerdasan buatan (AI). Machine learning adalah cabang dari AI yang dapat dimanfaatkan untuk membuat mesin penerjemah. 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 Penerjemahan, Algoritma Encoder-Decoder

KATA PENGANTAR

Puji syukur kita panjatkan atas kehadiran Allah SWT. Dimana atas berkat rahmat serta karunianya penulis dapat menyelesaikan Laporan Penelitian dengan baik, yang berjudul “Mesin Penerjemahan Bahasa Besemah Berbasis *Machine Learning* Dengan Algoritma Model *Encoder-Decoder*”. Laporan Penelitian ini disusun sebagai salah satu syarat untuk memperoleh gelar sarjana (S1) dari Program Studi Sistem Informasi, Fakultas Sains Teknologi Universitas Bina Darma Palembang.

Dalam proses penyusunan Laporan Penelitian ini, banyak sekali tahapan-tahapan yang harus diselesaikan satu-persatu. Semua itu tidak lepas dari adanya bantuan, bimbingan dan dukungan oleh berbagai pihak. Rasa terima kasih penulis ucapkan dengan sepenuh hati kepada :

1. Ucapan terima kasih saya haturkan kepada kedua orangtua saya Ayah saya Almarhum AIPTU Lukman Hakim dan Ibu saya Hernawati atas semua Do'a, dukungan, serta kasih sayangnya yang telah mengantarkan saya ke jenjang pendidikan tinggi ini.
2. Ibu Susan Dian Purnamasari, M. Kom. Selaku pembimbing saya yang telah membimbing saya dari awal hingga selesai dalam membuat dan menyusun laporan penelitian. Saya ucapkan terima kasih banyak.
3. Dr. Yesi Novaria Kunang, S. T., M. Kom Selaku dosen Intelligent System Riset Group terima kasih telah memberi arahan, bimbingan dan pengetahuan dalam laporan penelitian, saya ucapkan terima kasih banyak atas semua ilmu yang diberi.
4. Kepada Saudara-saudariku dan Keluarga besarku.
5. Prof. Dr. Sunda Ariana, M.Pd., M.M. selaku Rektor Universitas Bina DarmaPalembang.
6. Dr. Tata Sutabri, S.Kom., MMSI., MKM. Selaku Dekan Fakultas

SainsTeknologi.

7. Nita Rosa Damayanti, M. Kom., Ph.D. selaku Ketua Program Studi Sistem Informasi.
8. Seluruh Dosen pengajar program studi Sistem Informasi yang telah memberikan ilmu, pengetahuan dan wawasan selama saya menempuh pendidikan di Universitas Bina Darma.
9. Kepada teman-teman yang telah membantu dan saling support terkhususnya Teman-teman HIMSIF
10. Seluruh tim Universitas Bina Darma yang telah kebersamai penulis dalam menyelesaikan tugas akhir Laporan Penelitian ini.

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

Palembang 05 Agustus 2024

Bogy Dharma Sandi

DAFTAR ISI

HALAMAN PENGESAHAN.....	i
HALAMAN PERSETUJUAN.....	ii
SURAT PERNYATAAN.....	iii
MOTTO DAN PERSEMBAHAN.....	iv
ABSTRAK.....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xiii
DAFTAR LAMPIRAN.....	xiv
BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian.....	4
1.4 Batasan Masalah.....	4
1.5 Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Landasan Teori.....	5
2.1.1 Mesin Penerjemah.....	5
2.1.2 Machine Learning.....	6
2.1.3 Text Mining.....	8
2.1.4 Algoritma Encoder-Decoder.....	10
2.2 Penelitian Terdahulu.....	12
BAB III METODE PENELITIAN.....	19
3.1 <i>Flowchart</i>	19
3.2 Pengumpulan Data.....	19
3.3 Pembuatan Dataset.....	20
3.4 <i>Pre-Processing</i>	20
3.4.1 Prepare Langs.....	20
3.4.2 Normalisasi.....	23
3.4.3 <i>Vektorisasi</i>	25
3.4.4 One Hot Encoding.....	28
3.5 Padding.....	30

3.6 Encoder-Decoder	33
3.7 Pelatihan Model.....	35
3.8 Evaluasi Model.....	36
3.9 Penerjemahan.....	38
BAB IV	39
HASIL DAN PEMBAHASAN	39
4.1 Pengumpulan Data	39
4.2 Pembuatan Dataset	39
4.3 Preprocessing.....	40
4.3.1 Prepare Langs.....	40
4.3.2 Normalisasi	42
4.3.3 Vektorisasi.....	44
4.3.4 One Hot Encoding	44
4.4 Padding	45
4.6 Pelatihan Model.....	49
4.7 Evaluasi Model.....	50
4.8 Penerjemahan	52
BAB V KESIMPULAN	56
5.1 Kesimpulan	56
5.2 Saran	56
DAFTAR PUSTAKA	57

DAFTAR GAMBAR

Gambar 3. 1 Flowchart Progres Mesin Penerjemahan	19
Gambar 3. 2 Codingan Proses Kelas Lang	21
Gambar 3. 3 Proses Codingan Menghitung Frekuensi Kata dalam Kalimat.....	21
Gambar 3. 4 Proses Codingan Menghitung Frekuensi Kata	22
Gambar 3. 5 Proses Codingan Membuat Titik Batas	22
Gambar 3. 6 Proses Codingan Menambahkan Kalimat	22
Gambar 3. 7 Proses Codingan Menambahkan Kata.....	23
Gambar 3. 8 Codingan Fungsi Normalisasi	23
Gambar 3. 9 Variabel Pada Proses Codingan Normalisasi	24
Gambar 3. 10 Codingan Proses Filter Pair	24
Gambar 3. 11 Codingan awal Proses Vektorisasi.....	25
Gambar 3. 12 Codingan Proses Membaca Pasangan Kalimat.....	26
Gambar 3. 13 Codingan Trim Pasangan Kalimat	26
Gambar 3. 14 Menghitung Kata dan Membuat Cutoff Kosakata.....	26
Gambar 3. 15 Codingan Menambahkan Kalimat ke Kosakata dan Mengacak Pasangan	27
Gambar 3. 16 Codingan Data Menjadi Set Pelatihan dan Pengujian	27
Gambar 3. 17 Codingan Mencetak Informasi dan Mengembalikan Hasil	27
Gambar 3. 18 Proses Codingan Indexes From Sentence.....	28
Gambar 3. 19 Proses Codingan Tensor From Sentence	29
Gambar 3. 20 Proses Codingan Tensors From Pair	29
Gambar 3. 21 Proses Codingan Sentence From Tensor	30
Gambar 3. 22 Codingan Proses Batchify	31
Gambar 3. 23 Proses Codingan Pad Batch	32
Gambar 3. 24 Proses Method Encoder	33
Gambar 3. 25 Proses Codingan Forward Method Encoder.....	34
Gambar 3. 26 Proses Codingan Create_init_hiddens Method	35
Gambar 3. 27 Pelatihan Model Penerjemahan	36
Gambar 3. 28 Evaluasi Model.....	37
Gambar 4. 1 Kamus Sumber Dataset yang didapat	39
Gambar 4. 2 Dataset dalam format file excel.....	40

Gambar 4. 3 Dataset dalam format file txt.....	40
Gambar 4. 4 Codingan Proses PREPARE LANGS.....	41
Gambar 4. 5 Lanjutan Codingan Proses Pembuatan PREPARE LANGS.....	42
Gambar 4. 6 Codingan Proses Normalisasi.....	42
Gambar 4. 7 Proses Codingan Mengganti Pola Khusus Pada Normalisasi.....	43
Gambar 4. 8 Proses Codingan Mengubah Huruf Menjadi Huruf Kecil Pada Normalisasi.....	43
Gambar 4. 9 Proses Codingan Memisahkan Tanda Baca Pada Normalisasi.....	43
Gambar 4. 10 Proses Codingan Menghapus Karakter Non-Huruf.....	44
Gambar 4. 11 Codingan Proses One Hot Encoding.....	45
Gambar 4. 12 Lanjutan Codingan Proses One Hot Encoding.....	45
Gambar 4. 13 Codingan Proses Padding.....	46
Gambar 4. 14 Lanjutan Codingan Proses Padding.....	46
Gambar 4. 15 Proses Codingan Model Encoder.....	47
Gambar 4. 16 Lanjuttan Proses Codingan Model Encoder.....	47
Gambar 4. 17 Proses Codingan Model Decoder.....	48
Gambar 4. 18 Lanjutan Prosesn Codingan Model Decoder.....	48
Gambar 4. 19 Hasil Pelatihan Model Encoder-Decoder.....	49
Gambar 4. 20 Hasil Evaluasi Model Encoder-Decoder.....	50

DAFTAR TABEL

Table 2. 1 Penelitian Sebelumnya18



DAFTAR LAMPIRAN

LAMPIRAN 1: SURAT KETERANGAN LULUS SEMINAR HASIL	1
LAMPIRAN 2: SURAT KETERANGAN LULUS SEMINAR HASIL	1
LAMPIRAN 3 : SK PEMBIMBING	2
LAMPIRAN 4 : SK PEMBIMBING	2
LAMPIRAN 5 : LEMBAR ACC PENGAJUAN JUDUL	3
LAMPIRAN 6 : LEMBAR ACC PENGAJUAN JUDUL	3
LAMPIRAN 7 : LEMBAR KONSULTASI PEMBIMBING SEMINAR HASIL	4
LAMPIRAN 8 : LEMBAR KONSULTASI PEMBIMBING SEMINAR HASIL	4
LAMPIRAN 9 : LEMBAR PERBAIKAN SEMHAS	6
LAMPIRAN 10 : LEMBAR PERBAIKAN SEMHAS	6
LAMPIRAN 11 : SURAT KETERANGAN LULUS SEMINAR PROPOSAL	7
LAMPIRAN 12 : SURAT KETERANGAN LULUS SEMINAR PROPOSAL	7
LAMPIRAN 13 : LEMBAR KONSULTASI PEMBIMBING PROPOSAL PENELITIAN	8
LAMPIRAN 14 : LEMBAR KONSULTASI PEMBIMBING PROPOSAL PENELITIAN	8
LAMPIRAN 15 : FORM VERIFIKASI PENJILIDAN	9
LAMPIRAN 16 : FORM VERIFIKASI PENJILIDAN	9
LAMPIRAN 17 : HASIL TURNITIN	10
LAMPIRAN 18 : HASIL TURNITIN	10
LAMPIRAN 19 : SURAT PERNYATAAN BELUM ADA LOA JURNAL	16
LAMPIRAN 20 : SURAT PERNYATAAN BELUM ADA LOA JURNAL	16
LAMPIRAN 21 : BUKTI SUBMIT JURNAL	17
LAMPIRAN 22 : LEMBAR PERBAIKAN SEMINAR PROPOSAL	18