

**PENERAPAN ALGORITMA LONG SHORT-TERM MEMORY
(LSTM) DALAM PENGENALAN NAMA, LOKASI, DAN
ORGANISASI PADA BERITA PEMILU 2024**

SKRIPSI

Oleh:

Ozzi Ardhiyanto

202010225030



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2024**

**LEMBAR PERSETUJUAN PEMBIMBING
(DIGUNAKAN UNTUK TUGAS AKHIR)**

Judul Tugas akhir : PENERAPAN ALGORITMA LONG SHORT-
TERM MEMORY (LSTM) DALAM
PENGENALAN NAMA, LOKASI, DAN
ORGANISASI PADA BERITA PEMILU 2024

Nama Mahasiswa : OZZI ARHIYANTO

Nomor : 202010225030

Pokok
Mahasiswa

Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian : 28, Juni 2024

Skripsi

Jakarta, 05/07/2024

MENYETUJUI,

Pembimbing I



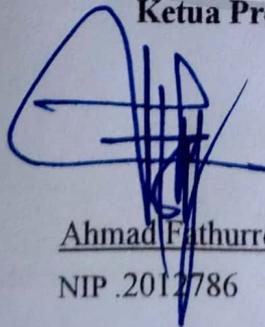
Dr. Mujiono, S.T., M.T
NIDN.0406127002

Pembimbing II



Fata Nidaul Khasanah, S. Kom., M.Eng
NIDN.0327059202

Ketua Program Studi



Ahmad Fathurrozi, S.E., M.M.S.I
NIP.2012786

**Program Studi Informatika
Fakulta Ilmu Komputer Universitas Bhayangkara Jakarta Raya
2024**

LEMBAR PENGESAHAN

Judul Skripsi : PENERAPAN ALGORITMA LONG SHORT-TERM
MEMORY (LSTM) DALAM PENGENALAN
NAMA, LOKASI, DAN ORGANISASI PADA
BERITA PEMILU 2024

Nama Mahasiswa : OZZI ARDHIYANTO

Nomor Pokok Mahasiswa : 202010225030

Program Studi/Fakultas : Informatika/Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 28, Juni 2024

Jakarta, 05 Juli 2024

MENGESAHKAN,

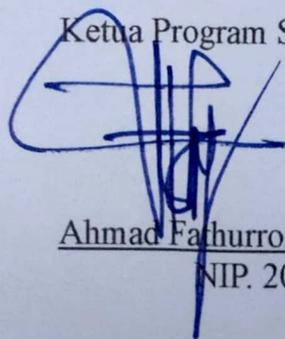
Ketua Tim Penguji : Wowon Priatna, ST, M
NIDN. 0312079403

Penguji I : Nurfyah, ST., M.Kom
NIDN. 0312079403

Penguji II : Dr. Mujiono, S.T., M.T
NIDN.0406127002

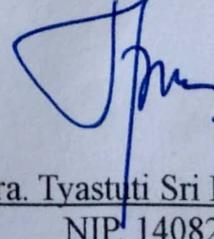
MENGETAHUI,

Ketua Program Studi Informatika



Ahmad Fathurrozi, S.E., M.M.S.I
NIP. 2012486

Dekan Fakultas Ilmu Komputer



Dr. Dra. Tyastuti Sri Lestari, M.M.
NIP. 1408206



LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : OZZI ARDHIYANTO
NPM : 202010225030
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : PENERAPAN ALGORITMA LONG SHORT-TERM
MEMORY (LSTM) DALAM PENGENALAN NAMA,
LOKASI, DAN ORGANISASI PADA BERITA PEMILU
2024

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Jakarta, 05 Juli 2024

Penulis



OZZI ARDHIYANTO

ABSTRAK

Ozzi Ardhiyanto, 202010225030. Pada era informasi digital saat ini, pengenalan entitas seperti nama, lokasi, dan organisasi dalam konteks berita pemilu memiliki peran yang sangat penting dalam memahami dan menganalisis dinamika politik serta pandangan masyarakat. Penelitian ini bertujuan untuk meningkatkan pengenalan entitas tersebut dalam teks berita pemilu 2024 dengan menggunakan algoritma Long Short-Term Memory (LSTM). Metode yang diusulkan melibatkan penggunaan LSTM untuk mengidentifikasi dan mengekstraksi entitas nama, lokasi, dan organisasi secara otomatis dari teks berita pemilu. Melalui serangkaian eksperimen, pendekatan ini dievaluasi dan dibandingkan dengan metode tradisional. Selain itu, evaluasi model menggunakan *confusion matrix* memberikan gambaran yang lebih rinci tentang kinerja klasifikasi entitas oleh model LSTM. Penerapan model LSTM dapat diterapkan dalam pembuatan sistem dalam menentukan entitas orang, lokasi, organisasi dan yang bukan termasuk dari ketiga entitas tersebut pada berita pemilu 2024 dengan menghasilkan nilai *average precision* 0.9280, nilai *average recall* 0.9489, nilai *average F1 score* 0.9381, dan nilai *average accuracy* 0.9489. Maka dapat dikatakan bahwa model memiliki kinerja yang cukup baik dalam melakukan prediksi entitas dalam berita pemilu 2024.

Kata Kunci: Berita Pemilu 2024, Pengenalan Entitas, LSTM, Analisis Berita, Pemrosesan Teks

ABSTRACT

Ozzi Ardhiyanto 202010225030. *In the current digital information era, recognizing entities such as names, locations, and organizations in the context of election news plays a crucial role in understanding and analyzing political dynamics and public opinions. This research aims to enhance the recognition of these entities in the 2024 election news texts using Long Short-Term Memory (LSTM) algorithms. The proposed method involves using LSTM to automatically identify and extract name, location, and organization entities from election news texts. Through a series of experiments, this approach is evaluated and compared with traditional methods. Additionally, the model evaluation using a confusion matrix provides a detailed overview of the entity classification performance by the LSTM model. The application of the LSTM model can be utilized in developing systems to determine entities such as persons, locations, organizations, and those that do not belong to these three categories in the 2024 election news, achieving an average precision of 0.9280, an average recall of 0.9489, an average F1 score of 0.9381, and an average accuracy of 0.9489. Thus, it can be concluded that the model performs quite well in predicting entities in the 2024 election news.*

Keyword: 2024 Election News, Entity Recognition, LSTM, News Analysis, Text Processing

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : OZZI ARDHIYANTO
NPM : 202010225030
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Esklusif (*Non-Exclusive Royalty-Free Right*)**, atas karya ilmiah saya yang berjudul :

PENERAPAN ALGORITMA LONG SHORT-TERM MEMORY (LSTM)
DALAM PENGENALAN NAMA, LOKASI, DAN ORGANISASI PADA
BERITA PEMILU 2024

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 05 Juli 2024
Yang Menyatakan


Ozzi Ardhiyanto

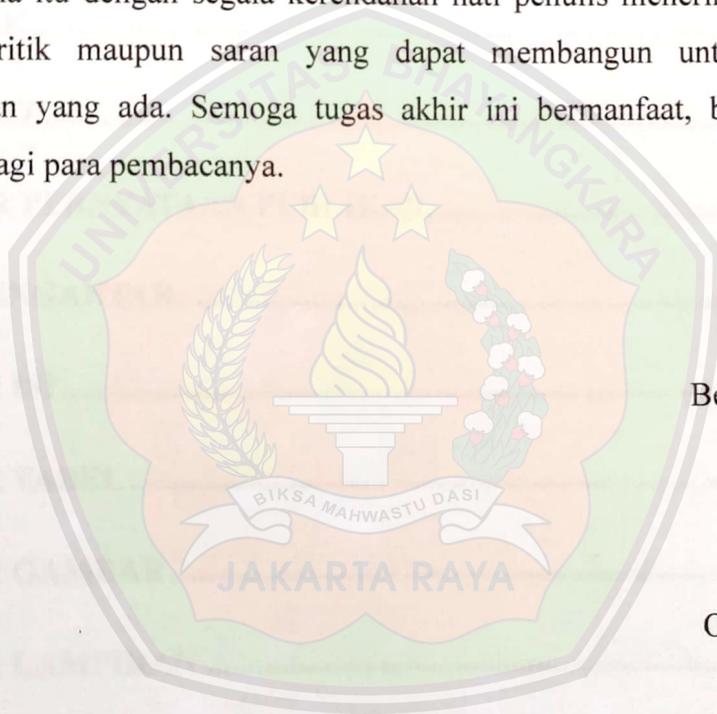
KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kepada Allah SWT, yang telah melimpahkan rahmat dan hidayahnya sehingga penulis dapat menyelesaikan tugas akhir yang berjudul “Pengenalan Entitas Dalam Berita Pemilu 2024 Menggunakan Algoritma LSTM” dengan baik dan lancar sebagai salah satu persyaratan yang harus dipenuhi dalam menyelesaikan jenjang strata satu di Jurusan Informatika, Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya. Penyusunan dan penulisan tugas akhir ini tidak akan terselesaikan tanpa adanya dukungan, bantuan dan bimbingan dari banyak pihak. Oleh sebab itu, penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Bapak Irjen Polisi (Purn) Prof. Dr. Drs. Bambang Karsono, S.H., M.M., Ph.D., D.Crim (HC) selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Ibu Dr. Dra. Tyastuti Sri Lestari, M.M selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Bapak Ahmad Fathurrozi, S.E., M.M.S.I. selaku Kepala Program Studi Informatika Universitas Bhayangkara Jakarta Raya.
4. Bapak Dr. Mujiono, M.T selaku Dosen Pembimbing satu, yang telah memberikan bimbingan, arahan dan masukan yang sangat berharga selama proses penulisan tugas akhir ini.
5. Ibu Fata Nidaul Khasanah, S.Kom., M.Eng selaku Dosen Pembimbing dua, yang telah memberikan bimbingan, arahan dan masukan yang sangat berharga selama proses penulisan tugas akhir ini.
6. Bapak Mugiarto, S.Kom., M.Kom. selaku Dosen Pembimbing Akademik, yang telah membantu, memberikan arahan dan inspirasi dalam perjalanan pendidikan penulis.
7. Kedua Orang Tua yang selalu mendukung, mendoakan dan memberikan semangat kepada penulis agar dapat menyelesaikan perkuliahan dan tugas akhir ini dengan lancar.

8. Rekan-rekan Program Studi Informatika Angkatan 2020, Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya yang menjadi tempat bertukar ilmu dan bertukar informasi akademik maupun non akademik serta berjuang bersama untuk menyelesaikan kuliah dan menyanggah gelar sarjana.

Penulis menyadari bahwa masih banyak kekurangan pada tugas akhir ini, oleh karena itu dengan segala kerendahan hati penulis menerima masukan baik berupa kritik maupun saran yang dapat membangun untuk memperbaiki kekurangan yang ada. Semoga tugas akhir ini bermanfaat, baik bagi penulis maupun bagi para pembacanya.



Bekasi, Juni 2024

Ozzi Ardhiyanto

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iii
ABSTRAK	iv
ABSTRACT	v
LEMBAR PERNYATAAN PUBLIKAS.....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah.....	3
1.3 Rumusan Masalah.....	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Batasan Masalah	4

1.7	Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA		6
2.1	Tinjauan Pustaka	6
2.2	Pemilu	8
2.3	Python	8
2.4	Knowledge Discovery in Databases (KDD)	9
2.5	Named Entity Recognition(NER)	10
2.6	Long Short Term Memory	11
2.7	Bidirectional LSTM (Bi-LSTM).....	15
2.8	K-fold Cross Validation	17
2.9	Evaluasi Model	19
BAB III METODOLOGI PENELITIAN		21
3.1	Kerangka Penelitian	21
3.3	Metode Pengumpulan Data	22
3.2.1	Studi Kepustakaan.....	22
3.2.2	Pengembangan Data Kualitatif	22
3.4	Data Penelitian	22
3.5	Metode Analisis	23
3.4.1	Preprocessing Data.....	23
3.4.2	Data Cleaning.....	23

3.4.3	Tagging.....	24
3.4.4	Pembuatan Model LSTM dan Bi-LSTM	24
3.4.5	Pelatihan dan Optimasi Model.....	24
3.4.6	Evaluasi Model	25
3.4.7	Analisis Kuantitatif	25
3.6	Rancangan Penelitian.....	25
BAB IV HASIL DAN PEMBAHASAN.....		27
4.1	Persiapan Data	27
4.1.1	Pengumpulan Data	27
4.2	Preprocessing Data.....	27
4.2.1	Data Cleaning.....	27
4.2.3	Inisialisasi Proyek Python.....	29
4.2.4	Import paket-paket yang diperlukan	29
4.2.5	Check Data.....	31
4.2.6	Tokenize Data	34
4.2.7	Pemilahan Data	36
4.2.8	Transformasi Data Teks menjadi Matriks Sequence.....	38
4.3	Modeling	40
4.4	Evaluasi Model	43
4.4.1	Evaluasi kinerja model melalui prediksi data validasi.....	47
BAB V PENUTUP.....		50

5.1 Kesimpulan	50
5.2 Saran	51
DAFTAR PUSTAKA.....	52
LAMPIRAN.....	55



DAFTAR TABEL

Tabel 1. 1 Contoh Dataset	2
Tabel 2. 1 Tinjauan Pustaka	6
Tabel 2. 2 Tabel contoh pembagian cross validation K=5	18
Tabel 3. 1 Jadwal Penelitian.....	25
Tabel 4. 1 Dataset Awal.....	28
Tabel 4. 2 Dataset yang Telah diolah	29
Tabel 4. 3 penjelasan paket modul yang digunakan.....	30
Tabel 4. 4 parameter word count.....	32
Tabel 4. 5 hasil perhitungan word count	32
Tabel 4. 6 menampilkan mismatch row	33
Tabel 4. 7 hasil dari pemisahan setiap kata	34
Tabel 4. 8 parameter tokenized data.....	35
Tabel 4. 9 Parameter pemilahan data	37
Tabel 4. 10 data test.....	37
Tabel 4. 11 data train.....	38
Tabel 4. 12 parameter tokenisasi, padding dan encoding.....	39
Tabel 4. 13 data yang telah dilakukan tokenisasi, padding dan encoding.....	39
Tabel 4. 14 Parameter model LSTM.....	42
Tabel 4. 15 Perbandingan data train dan data test.....	46
Tabel 4. 16 Tabel perbandingan hasil sebenarnya dan hasil prediksi.....	48

DAFTAR GAMBAR

Gambar 2. 1 LSTM Gates	11
Gambar 2. 2 Forget Gate.....	12
Gambar 2. 3 Input Gate.....	13
Gambar 2. 4 Output Gate	14
Gambar 2. 5 Bi-LSTM Layer.....	16
Gambar 2. 6 Confusion Matrix	19
Gambar 3. 1 Alur Penelitian.....	21
Gambar 4. 1 Paket dan Modul	30
Gambar 4. 2 Paket dan modul untuk membuat model dan evaluasi	30
Gambar 4. 3 menghitung jumlah karakter pada kolom word dan tag	32
Gambar 4. 4 Miss Match Row	33
Gambar 4. 5 Pemisahan kalimat menjadi per kata.....	34
Gambar 4. 6 Tokenisasi data	35
Gambar 4. 7 Tokenize Data.....	36
Gambar 4. 8 Split data dan Convert menjadi list python	36
Gambar 4. 9 Tokenizer, Flattened dan Encode	39
Gambar 4. 10 Inisialisasi Variabel dan Library.....	40
Gambar 4. 11 LSTM model	40
Gambar 4. 12 Matriks evaluasi	41
Gambar 4. 13 Arsitektur model.....	43
Gambar 4. 15 Membandingkan data asli dan prediksi	47

DAFTAR LAMPIRAN

Lampiran 1 : SOURCE CODE NER LSTM.....	56
--	----

