

TUGAS AKHIR

**PERFORMA MODEL NAMED ENTITY
RECOGNITION MENGGUNAKAN METODE
BLSTM–HMM PADA DOKUMEN TWITTER
BERBAHASA INDONESIA**



**ADI WIJAYA
19102099**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO
2024**

TUGAS AKHIR

**PERFORMA MODEL NAMED ENTITY
RECOGNITION MENGGUNAKAN METODE
BLSTM–HMM PADA DOKUMEN TWITTER
BERBAHASA INDONESIA**

Disusun Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Komputer



**ADI WIJAYA
19102099**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO
2024**

HALAMAN PERSETUJUAN PEMBIMBING

**PERFORMA MODEL NAMED ENTITY
RECOGNITION MENGGUNAKAN METODE
BLSTM-HMM PADA DOKUMEN TWITTER
BERBAHASA INDONESIA**

Dipersiapkan dan Disusun oleh

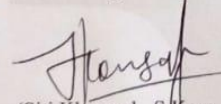
ADI WIJAYA
19102099

Fakultas Informatika

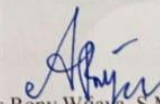
Institut Teknologi Telkom Purwokerto

Pada Tanggal: 24 Januari 2024

Pembimbing Utama,


(Siti Khomsah, S.Kom., M.Cs)
NIDN. 051708101

Pembimbing Pendamping,


(Andreas Rony Wijaya, S.Mat., M.Sc)
NIDN. 0630050602

HALAMAN PENGESAHAN

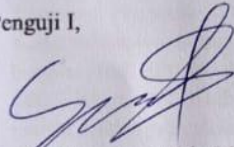
**PERFORMA MODEL NAMED ENTITY
RECOGNITION MENGGUNAKAN METODE
BLSTM-HMM PADA DOKUMEN TWITTER
BERBAHASA INDONESIA**

Dipersiapkan dan Disusun oleh

Adi Wijaya
19102099

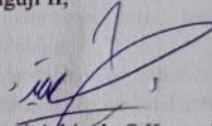
Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tugas Akhir
Pada hari, Rabu 24 Januari 2024

Penguji I,



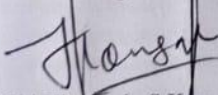
Sena Wijayanto, S.Pd., M.T
NIDN. 0613109201

Penguji II,



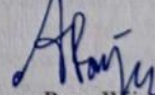
Ummi Azzahrah, S.Kom., M.Kom
NIDN. 0621129001

Pembimbing Utama,

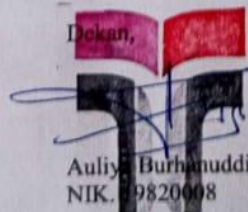


Siti Khomsah, S.Kom., M.Cs
NIDN. 0517108101

Pembimbing Pendamping,



Andreas Rony Wijaya, S.Mat., M.Sc
NIDN. 0630059602



Auliy Burhanuddin, S.Si., M.Kom
NIK. 9820008

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertandatangan di bawah ini,

Nama mahasiswa : Adi Wijaya
NIM : 19102099
Program Studi : Teknik Informatika

Menyatakan bahwa Tugas Akhir dengan judul berikut:
**PERFORMA MODEL NAMED ENTITY RECOGNITION
MENGUNAKAN METODE BLSTM-HMM PADA DOKUMEN
TWITTER BERBAHASA INDONESIA**

Dosen Pembimbing Utama : Siti Khomsah, S.Kom., M.Cs
Dosen Pembimbing Pendamping : Andreas Rony Wijaya, S.Mat., M.Sc

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Institut Teknologi Telkom Purwokerto maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan, dan penelitian Saya Sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing.
3. Dalam Karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab Saya, bukan tanggung jawab Institut Teknologi Telkom Purwokerto.
5. Pernyataan ini Saya buat dengan sesungguhnya, apabila dikemudian hariterdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka Saya bersedia menerima Sanksi Akademik dengan pencabutan gelar yang sudah diperoleh serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Purwokerto, 9 Januari 2024,
Yang Menyatakan,



(Adi Wijaya)

KATA PENGANTAR

Pertama penulis panjatkan puji dan syukur kepada kehadiran Allah Subhanahu Wa Ta'ala atas segala rahmat serta hidayah-Nya sehingga penulis bisa mengerjakan tugas akhir yang berjudul **“PERFORMA MODEL NAMED ENTITY RECOGNITION MENGGUNAKAN METODE BLSTM–HMM PADA DOKUMEN TWITTER BERBAHASA INDONESIA”** hingga selesai tepat waktu. Penulis dalam kesempatan ini mengucapkan terima kasih kepada:

1. Dr. Tenia Wahyuningrum, S.Kom., M.T. Selaku Rektor Institut Teknologi Telkom Purwokerto.
2. Auliya Burhanuddin, S.Si., M.Kom. Selaku Dekan dari Fakultas Informatika, dan dosen wali penulis.
3. Amalia Beladonna Arifa, S.Pd., M.Cs. Selaku Kepala Program Studi S1 Informatika Institut Teknologi Telkom Purwokerto.
4. Siti Khomsah, S.Kom., M.Cs. Selaku dosen pembimbing utama yang memberikan bimbingan serta arahan hingga tugas akhir penulis ini dapat selesai dengan baik dan tepat waktu.
5. Andreas Rony Wijaya, S.Mat., M.Sc. Selaku dosen pembimbing pendamping yang juga memberikan bimbingan serta arahan hingga tugas akhir penulis ini dapat selesai dengan baik dan tepat waktu.
6. Orang tua dan keluarga penulis yang mendukung, mendoakan, serta memberikan banyak bantuan untuk penulis selama pengerjaan tugas akhir.
7. Teman-teman dari IT Telkom Purwokerto yang tidak bisa penulis sebutkan satu per satu yang sudah membantu serta mendukung penulis.

Purwokerto, 9 Januari 2024



Adi Wijaya

DAFTAR ISI

TUGAS AKHIR.....	i
HALAMAN PERSETUJUAN PEMBIMBING	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR.....	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vi
DAFTAR TABEL.....	viii
DAFTAR GAMBAR	ix
ABSTRAK	x
ABSTRACT.....	xi
BAB I.....	1
PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Pertanyaan Penelitian.....	4
1.4. Tujuan Penelitian	4
1.5. Batasan Masalah	4
1.6. Manfaat Penelitian	5
BAB II.....	6
TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	6
2.1. Tinjauan Pustaka	6
2.2. Landasan Teori.....	8
2.2.1. Natural Language Processing	8
2.2.2. Information Extraction	9
2.2.3. Named Entity Recognition	9
2.2.4. Bidirectional Long Short-Term Memory.....	11
2.2.5. Hidden Markov Model	14
2.2.6. Confusion Matrix.....	16
BAB III	19
METODOLOGI PENELITIAN.....	19

3.1. Subyek dan Obyek Penelitian	19
3.2. Alat dan Bahan Penelitian.....	19
3.2.1. Perangkat Keras	19
3.2.2. Perangkat Lunak	19
3.2.3. Data Penelitian.....	19
3.3. Diagram Alir Penelitian	20
3.3.1. Studi Literatur	21
3.3.2. Pengumpulan Dataset	21
3.3.3. Pembangunan Model Named Entity Recognition	22
3.3.4. Pengujian Model NER.....	28
3.3.5. Analisis Hasil.....	29
3.3.6. Kesimpulan	29
BAB IV	30
HASIL DAN PEMBAHASAN.....	30
4.1. Analisis Awal Data	30
4.2. Pembahasan Hasil Implementasi	37
4.2.1. Preprocessing.....	37
4.2.2. Ekstraksi Fitur FastText.....	41
4.2.3. Model BLSTM	48
4.2.4. Model HMM.....	49
4.3. Pengujian Model NER	50
4.4. Analisis Hasil	52
BAB V	54
KESIMPULAN DAN SARAN.....	54
5.1. Kesimpulan	54
5.2. Saran	54
DAFTAR PUSTAKA	55
LAMPIRAN.....	58

DAFTAR TABEL

Tabel 2.1 Kajian Penelitian <i>NER</i> Sebelumnya	7
Tabel 3.1 Preview Data Train	21
Tabel 3.2 Preview Data Test	22
Tabel 3.3 Contoh <i>Embedding</i> Layer Tabel	25
Tabel 3.4 Contoh <i>Average Pooling</i>	26
Tabel 4.1 Entity pada data train awal	30
Tabel 4.2 Entity pada data test awal	31
Tabel 4.3 <i>POS tag</i> pada data train awal	32
Tabel 4.4 <i>POS tag</i> pada data test awal	33
Tabel 4.5 <i>POS tag</i> yang ditentukan	35
Tabel 4.6 Label entity yang ditentukan	36
Tabel 4.7 Entity data train setelah preprocessing	38
Tabel 4.8 Entity data test setelah preprocessing	39
Tabel 4.9 <i>POS tag</i> data train setelah preprocessing	39
Tabel 4.10 <i>POS tag</i> data test setelah preprocessing	40
Tabel 4.11 <i>build_input_vector</i> pada data train	42
Tabel 4.12 <i>build_input_vector</i> pada data test	43
Tabel 4.13 <i>build_output_vector</i> pada data train	45
Tabel 4.14 <i>build_output-vektor</i> pada data test	46
Tabel 4.15 Dimensi Data	47
Tabel 4.16 Hasil Pengujian	50
Tabel 4.17 Perbandingan Hasil	52

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi <i>BLSTM</i>	12
Gambar 3.1 Diagram alir penelitian.....	20
Gambar 3.2 Desain sistem model <i>NER</i>	23
Gambar 3.3 Arsitektur <i>BLSTM-HMM</i>	27
Gambar 4.1 Ringkasan Model <i>BLSTM</i>	49
Gambar 4.2 Ringkasan Model <i>HMM</i>	50
Gambar 4.3 Hasil <i>F1 score</i>	52