

**KLASIFIKASI BERITA MENGGUNAKAN METODE
*LONG SHORT-TERM MEMORY (LSTM)***

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika



Oleh :

Elwina
NIM : 09021182025018

**Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2024**

LEMBAR PENGESAHAN SKRIPSI

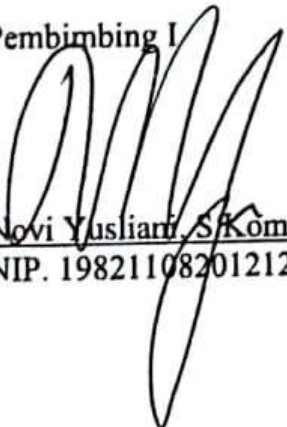
KLASIFIKASI BERITA MENGGUNAKAN METODE *LONG SHORT-TERM MEMORY (LSTM)*

Oleh :


Elwina
NIM : 09021182025018

Palembang, 13 Januari 2024

Pembimbing I


Novi Yusliani, S.Kom., M.T.
NIP. 198211082012122001

Pembimbing II


Hadipurnawan Satria, Ph.D.
NIP. 198004182020121001

Mengetahui,
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom
NIP. 197812222006042003

TANDA LULUS UJIAN KOMPREHENSIF

Pada hari Selasa tanggal 9 Januari 2024 telah dilaksanakan ujian Komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Elwina
NIM : 09021182025018
Judul : Klasifikasi Berita Menggunakan Metode *Long Short-Term Memory*
(LSTM)

dan dinyatakan LULUS.

1. Ketua Penguji

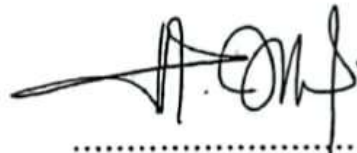
Desty Rodiah, M.T.
NIP. 198912212020122011



.....

2. Penguji

Annisa Darmawahyuni, M.Kom.
NIP. 199006302023212044



.....

3. Pembimbing 1

Novi Yusliani, M.T.
NIP. 198211082012122001



.....

4. Pembimbing 2

Hadipurnawan Satria, Ph.D.
NIP. 198004182020121001



.....

Mengetahui,
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.
NIP. 197812222006042003

HALAMAN PERNYATAAN BEBAS PLAGIAT

Yang bertanda tangan dibawah ini :

Nama : Elwina
NIM : 09021182025018
Program Studi : Teknik Informatika Reguler
Judul : Klasifikasi Berita Menggunakan Metode *Long Short-Term Memory (LSTM)*

Hasil Pengecekan *Software iThenticate/Turnitin* : 16%

Menyatakan bahwa laporan skripsi saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam laporan skripsi ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan dari siapapun

Palembang, 3 Januari 2024

Penulis,



Elwina

NIM. 09021182025018

MOTTO DAN PERSEMBAHAN

“Tidak Ada Hasil Yang Menghianati Proses”

Kupersembahkan karya tulis ini kepada :

- Allah SWT
- Orang Tua dan Keluarga Besarku
- Sahabat Seperjuanganku
- Universitas Sriwijaya

ABSTRACT

In today's digital information age, the abundance of news available on the internet creates the need for an automated system that can classify news very accurately based on its category. This research presents an approach that uses Long Short-Term Memory (LSTM) to classify news and Term Frequency-Inverse Document Frequency (TF-IDF) as word weighting. The LSTM method is used to process news text and extract features that represent relevant news content. The data used is multiclass and taken from Kaggle with a total data of 32,259 news titles which are then divided into 70% training data and 30% test data. After testing the test data, the LSTM classification performance results are obtained with an accuracy value of 87%, precision 88%, recall 87%, and f-score 88%.

Keywords: Classification, Long Short-Term Memory, Term Frequency-Inverse Document Frequency, multiclass.

ABSTRAK

Di era informasi digital saat ini, banyaknya berita yang tersedia di internet menciptakan kebutuhan akan sistem otomatis yang dapat mengklasifikasikan berita dengan sangat akurat berdasarkan kategorinya. Penelitian ini menyajikan pendekatan yang menggunakan *Long Short-Term Memory* (LSTM) untuk mengklasifikasikan berita dan *Term Frequency-Inverse Document Frequency* (TF-IDF) sebagai pembobotan kata. Metode LSTM digunakan untuk mengolah teks berita dan mengekstrak fitur-fitur yang mewakili konten berita yang relevan. Data yang digunakan bersifat *multiclass* dan diambil dari Kaggle dengan jumlah data yaitu 32.259 judul berita yang kemudian dibagi menjadi data latih 70% dan data uji 30%. Setelah dilakukan pengujian pada data uji diperoleh hasil kinerja klasifikasi LSTM dengan nilai *accuracy* 87%, *precision* 88%, *recall* 87%, dan *f-score* 88%.

Kata Kunci : Klasifikasi, *Long Short-Term Memory*, *Term Frequency-Inverse Document Frequency*, *multiclass*.

KATA PENGANTAR

Puji dan syukur penulis ucapkan kepada Allah SWT atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan Skripsi dengan judul **“Klasifikasi Berita Menggunakan Metode *Long Short-Term Memory (LSTM)*”**. Penulisan Skripsi ini ditujukan untuk melengkapi salah satu syarat dalam menyelesaikan pendidikan dan mendapatkan gelar Sarjana Komputer di Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Dalam menyelesaikan skripsi ini banyak pihak yang telah memberikan bantuan dan dukungan baik secara langsung maupun tidak langsung. Untuk itu penulis mengucapkan terima kasih setulus-tulusnya kepada :

1. Kedua Orang Tua saya yaitu Johari Syahrial dan Tri Koria Susie, saudara kembarku yaitu Elwansyah. Terima kasih untuk semua do'a yang telah dipanjatkan dan terima kasih juga untuk semua dukungan dan bantuan yang telah diberikan.
2. Bapak Jaidan Jauhari, S.Pd., M.T. (alm) selaku Dekan Fakultas Ilmu Komputer
3. Ibu Alvi Syahrini Utami, M.Kom. selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Samsuryadi, M.Kom., Ph.D. selaku Dosen Pembimbing Akademik yang telah memberikan bimbingan, arahan, dan motivasi dalam proses perkuliahan.
5. Ibu Novi Yusliani, M.T. dan Bapak Hadipurnawan Satria, Ph.D. selaku Dosen Pembimbing Skripsi yang telah memberikan arahan, masukan, motivasi, serta semangat kepada penulis dalam menyelesaikan Skripsi ini.
6. Ibu Annisa Darmawahyuni, M.Kom. selaku Dosen Penguji yang telah memberikan masukan dan pengetahuan pada penulis dalam penyelesaian Skripsi ini.
7. Seluruh Dosen dan Staff Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

8. Sahabat-sahabatku Hania, Rizka, Mifta, Nadya, dan Lastri yang telah memberikan dukungan, semangat dan kasih sayang selama perkuliahan. Serta Teman-teman Teknik Informatika Reguler B 2020.

Penulis menyadari bahwa masih banyak keterbatasan dan kekurangan yang ada dalam penulisan Skripsi ini. Maka dari itu, segala kritik dan saran sangat penulis butuhkan agar dapat menghasilkan karya tulis yang lebih baik. Sekian yang dapat penulis sampaikan, semoga Skripsi ini dapat berguna dan bermanfaat bagi banyak orang.

Palembang, 29 Januari 2024



Elwina

DAFTAR ISI

	Halaman
LEMBAR PENGESAHAN PROPOSAL SKRIPSI.....	ii
TANDA LULUS UJIAN KOMPREHENSIF.....	iii
HALAMAN PERNYATAAN BEBAS PLAGIAT	iv
MOTTO DAN PERSEMBAHAN.....	v
ABSTRACT	vi
ABSTRAK	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xiv
DAFTAR GAMBAR.....	xvii
BAB I PENDAHULUAN	I-1
1.1 Pendahuluan.....	I-1
1.2 Latar Belakang	I-1
1.3 Rumusan Masalah.....	1-3
1.4 Tujuan Penelitian	I-4
1.5 Manfaat Penelitian	I-4
1.6 Batasan Masalah	I-5
1.7 Sistematika Penulisan	I-5
1.8 Kesimpulan	I-7
BAB II KAJIAN LITERATUR.....	II-1
2.1 Pendahuluan.....	II-1
2.2 Landasan Teori.....	II-1

2.2.1	Berita.....	II-1
2.2.2	Klasifikasi	II-2
2.2.3	Pra-Pengolahan	II-2
2.2.4	LSTM.....	II-5
2.2.5	Pembobotan Kata TF-IDF	II-9
2.2.6	<i>Confusion Matrix</i>	II-10
2.3	Penelitian Lain yang Relevan	II-12
2.4	Kesimpulan	II-14
BAB III METODOLOGI PENELITIAN		III-1
3.1	Pendahuluan.....	III-1
3.2	Pengumpulan Data	III-1
3.2.1	Jenis Data	III-1
3.2.2	Sumber Data.....	III-2
3.2.3	Metode Pengumpulan Data.....	III-2
3.3	Tahapan Penelitian	III-2
3.3.1	Kerangka Kerja	III-3
3.3.2	Kriteria Pengujian	III-7
3.3.3	Format Data Pengujian.....	III-7
3.3.4	Alat yang digunakan dalam Pelaksanaan Penelitian	III-8
3.3.5	Pengujian Penelitian.....	III-8
3.3.6	Analisis Hasil Pengujian dan Membuat Kesimpulan.....	III-9
3.4	Metode Pengembangan Perangkat Lunak.....	III-9
3.4.1	Fase Insepsi	III-10
3.4.2	Fase Elaborasi	III-10
3.4.3	Fase Kontruksi	III-10

3.4.4	Fase Transisi.....	III-11
3.5	Manajemen Proyek Penelitian	III-11
3.6	Kesimpulan	III-16
BAB IV PENGEMBANGAN PERANGKAT LUNAK		IV-1
4.1	Pendahuluan.....	IV-1
4.2	Fase Insepsi.....	IV-1
4.2.1	Pemodelan Bisnis	IV-1
4.2.2	Kebutuhan Sistem	IV-2
4.2.3	Analisis dan Desain.....	IV-4
4.2.3.1	Analisis Kebutuhan Perangkat Lunak	IV-4
4.2.3.2	Analisis Data	IV-5
4.2.3.3	Analisis Pra-Pengolahan Teks	IV-5
4.2.3.4	Analisis Proses Klasifikasi	IV-13
4.2.3.5	Analisis Hasil Klasifikasi	IV-14
4.2.4	Implementasi.....	IV-14
4.3	Fase Elaborasi	IV-19
4.3.1	Pemodelan Bisnis	IV-19
4.3.2	Perancangan Data.....	IV-19
4.3.3	Perancangan Tampilan Antarmuka	IV-20
4.3.4	Kebutuhan Sistem	IV-20
4.3.5	Diagram Aktivitas	IV-21
4.3.6	Diagram Urutan.....	IV-24
4.4	Fase Konstruksi.....	IV-27
4.4.1	Kebutuhan Sistem	IV-27
4.4.2	Diagram Kelas.....	IV-27
4.4.3	Implementasi.....	IV-28
4.4.3.1	Implementasi Kelas.....	IV-28
4.4.3.2	Implementasi <i>Interface</i>	IV-29

4.5	Fase Transisi.....	IV-30
4.5.1	Pemodelan Bisnis.....	IV-30
4.5.2	Rencana Pengujian.....	IV-30
4.5.3	Implementasi.....	IV-32
4.5.3.1	Pengujian <i>Use Case Input Data</i>	IV-32
4.5.3.2	Pengujian <i>Use Case</i> Melakukan Proses Klasifikasi Menggunakan Metode LSTM.....	IV-33
4.5.3.3	Pengujian <i>Use Case</i> Melakukan Pengujian dengan <i>Input</i> Kalimat dari Pengguna.....	IV-34
4.6	Kesimpulan	IV-35
BAB V HASIL DAN PENELITIAN.....		V-1
5.1	Pendahuluan.....	V-1
5.2	Data Hasil Penelitian.....	V-1
5.2.1	Konfigurasi Percobaan.....	V-1
5.2.2	Data Hasil Konfigurasi.....	V-2
5.2.3	Data Konfigurasi Hasil Pengujian Klasifikasi Berita	V-9
5.3	Analisis Hasil Penelitian	V-14
5.3.1	Analisis Hasil Klasifikasi Menggunakan LSTM	V-14
5.3.2	Analisis Hasil Pengujian Kalimat Masukan.....	V-15
5.4	Kesimpulan	V-16
BAB VI KESIMPULAN DAN SARAN.....		VI-1
6.1	Pendahuluan.....	VI-1
6.2	Kesimpulan	VI-1
6.3	Saran.....	VI-1
DAFTAR PUSTAKA		xix
LAMPIRAN		xxi

DAFTAR TABEL

Tabel II-1. Model <i>Confusion Matrix</i>	II-11
Tabel III-1. Rancangan <i>Confusion Matrix</i>	III-7
Tabel III-2. Rancangan Tabel Hasil Analisis Klasifikasi.....	III-9
Tabel III-3. Perencanaan Aktivitas Penelitian dalam bentuk WBS.....	III-12
Tabel IV-1. Tabel Fungsional.....	IV-3
Tabel IV-2. Kebutuhan Non-Fungsional.....	IV-4
Tabel IV-3. Contoh Data Judul Berita	IV-6
Tabel IV-4. Hasil <i>Cleaning</i>	IV-7
Tabel IV-5. Hasil <i>Case Folding</i>	IV-8
Tabel IV-6. Hasil <i>Tokenization</i>	IV-9
Tabel IV-7. Hasil Normalisasi	IV-10
Tabel IV-8. Hasil <i>Stopword Removal</i>	IV-11
Tabel IV-9. Hasil <i>Stemming</i>	IV-12
Tabel IV-10. Hasil <i>Training</i> pada model	IV-13
Tabel IV-11. Definisi <i>Actor</i>	IV-15
Tabel IV-12. Definisi <i>Use Case</i>	IV-16
Tabel IV-13. Skenario Melakukan <i>Input Data</i>	IV-17
Tabel IV-14. Skenario Melakukan Pra-Pengolahan Data	IV-17
Tabel IV-15. Skenario Melakukan Proses Klasifikasi Menggunakan Metode LSTM	IV-18

Tabel IV-16. Skenario Melakukan Pengujian dengan <i>Input</i> Kalimat dari Pengguna	IV-18
Tabel IV-17. Implementasi Kelas	IV-28
Tabel IV-18. Rencana Pengujian <i>Input</i> Data	IV-31
Tabel IV-19. Rencana Pengujian Melakukan Proses Klasifikasi Menggunakan Metode LSTM	IV-31
Tabel IV-20. Rencana Pengujian Melakukan Pengujian dengan <i>input</i> kalimat dari pengguna.....	IV-32
Tabel IV-21. Pengujian <i>Use Case Input</i> Data	IV-32
Tabel IV-22. Pengujian <i>Use Case</i> Klasifikasi Menggunakan Metode LSTM	IV-33
Tabel IV-23. Pengujian <i>Use Case</i> Melakukan Pengujian dengan <i>Input</i> Kalimat dari Pengguna	IV-34
Tabel V-1. Data Hasil Evaluasi <i>Validation</i> Klasifikasi LSTM.....	V-8
Tabel V-2. Data Hasil Pengujian Kinerja Klasifikasi LSTM Dengan Data Uji.....	V-10
Tabel V-3. Data Hasil Pengujian Prediksi Klasifikasi Berita.....	V-11

DAFTAR GAMBAR

Gambar II-1. Diagram proses klasifikasi teks	II-2
Gambar II-2. Contoh Proses pada <i>Cleaning</i>	II-4
Gambar II-3. Contoh Proses pada <i>Case Folding</i>	II-4
Gambar II-4. Contoh Proses pada <i>Tokenization</i>	II-4
Gambar II-5. Contoh Proses Normalisasi	II-5
Gambar II-6. Contoh Proses pada <i>Stopword Removal</i>	II-5
Gambar II-7. Contoh Proses <i>Stemming</i>	II-6
Gambar II-8. Arsitektur <i>Long Shost-Term Memory</i>	II-7
Gambar III-1. Tahapan Penelitian	III-2
Gambar III-2. Kerangka Sistem	III-4
Gambar III-3. Diagram Tahapan Pra Pengolahan	III-5
Gambar IV-1. Diagram <i>Use Case</i>	IV-15
Gambar IV-2. Rancangan Antarmuka Perangkat Lunak.....	IV-20
Gambar IV-3. Diagram Aktivitas <i>Input Data</i>	IV-21
Gambar IV-4. Diagram Aktivitas Pra-Pengolahan Data	IV-22
Gambar IV-5. Diagram Aktivitas Klasifikasi Menggunakan Metode LSTM.....	IV-23
Gambar IV-6. Diagram Aktivitas Melakukan Pengujian dengan <i>Input</i> Kalimat dari Pengguna	IV-24
Gambar IV-7. Diagram Urutan <i>Input Data</i>	IV-25
Gambar IV-8. Diagram Urutan Pra-Pengolahan Data	IV-25

Gambar IV-9. Diagram Urutan Klasifikasi Menggunakan Metode LSTM.....	IV-26
Gambar IV-10. Diagram Urutan Pengujian <i>Input</i> Kalimat dari Pengguna.....	IV-26
Gambar IV-11. Diagram Kelas	IV-28
Gambar IV-12. <i>Interface</i> Perangkat Lunak.....	IV-30
Gambar V-1. <i>Confusion Matrix</i> Hasil Validasi Klasifikasi LSTM <i>Fold</i> 1.....	V-3
Gambar V-2. <i>Confusion Matrix</i> Hasil Validasi Klasifikasi LSTM <i>Fold</i> 2.....	V-3
Gambar V-3. <i>Confusion Matrix</i> Hasil Validasi Klasifikasi LSTM <i>Fold</i> 3.....	V-4
Gambar V-4. <i>Confusion Matrix</i> Hasil Validasi Klasifikasi LSTM <i>Fold</i> 4.....	V-4
Gambar V-5. <i>Confusion Matrix</i> Hasil Validasi Klasifikasi LSTM <i>Fold</i> 5.....	V-5
Gambar V-6. <i>Confusion Matrix</i> Hasil Validasi Klasifikasi LSTM <i>Fold</i> 6.....	V-5
Gambar V-7. <i>Confusion Matrix</i> Hasil Validasi Klasifikasi LSTM <i>Fold</i> 7.....	V-6
Gambar V-8. <i>Confusion Matrix</i> Hasil Validasi Klasifikasi LSTM <i>Fold</i> 8.....	V-6
Gambar V-9. <i>Confusion Matrix</i> Hasil Validasi Klasifikasi LSTM <i>Fold</i> 9.....	V-7
Gambar V-10. <i>Confusion Matrix</i> Hasil Validasi Klasifikasi LSTM <i>Fold</i> 10.....	V-7
Gambar V-11. Grafik Hasil Evaluasi <i>Validation</i> Klasifikasi LSTM	V-9
Gambar V-12. <i>Confusion Matrix</i> Hasil Pengujian Klasifikasi LSTM menggunakan data uji.....	V-10
Gambar V-13. Hasil Model <i>Accuracy</i>	V-10

BAB I

PENDAHULUAN

1.1 Pendahuluan

Pada bab pendahuluan membahas latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, sistematika penulisan dan kesimpulan. Bab ini juga merangkum secara umum dari kegiatan yang dilakukan dalam penelitian. Pendahuluan dimulai dengan pembahasan mengenai Klasifikasi serta penelitian yang relevan dengan Long Short-Term Memory.

1.2 Latar Belakang

Berita adalah laporan tentang suatu peristiwa, opini, kecenderungan, situasi, kondisi interpretasi yang penting, menarik, masih baru dan perlu disampaikan secepatnya kepada khalayak, kata "berita" mengacu pada informasi yang disampaikan melalui surat kabar, radio, televisi, internet, dan media lainnya (Lijuan et al., 2015). Seiring berkembangnya berita di Indonesia, banyak portal berita yang beralih dari media cetak ke media elektronik yang bisa diakses secara online menggunakan internet, ratusan berita ditulis setiap harinya di berbagai portal berita Indonesia berbasis internet (Rizaldy & Santoso, 2017). Oleh karena itu berita-berita tersebut perlu diklasifikasikan untuk memudahkan pembaca dalam menemukan berita yang diinginkan. Klasifikasi secara manual membutuhkan waktu yang lama dan bisa membuat pembaca bosan, sedangkan dengan sistem yang dibantu

algoritma *Long Short-Term Memory* (LSTM) dapat membantu mengklasifikasi berita secara otomatis menggunakan *machine learning*.

Salah satu metode atau pendekatan dalam bidang pembelajaran mesin (*machine learning*) yang sering digunakan untuk mengklasifikasi adalah *deep learning*. *Deep learning* merupakan sub bidang dari *machine learning* yang menggunakan algoritma dan model jaringan syaraf tiruan (*neural networks*) yang terdiri dari beberapa lapisan (*layer*) untuk memahami dan memproses data yang kompleks. Pada dasarnya, *deep learning* merupakan jaringan syaraf yang memiliki tiga atau lebih lapisan (*Artificial Neural Network*) ANN. *Deep learning* mampu mempelajari fitur-fitur yang relevan secara otomatis dari data mentah, berkinerja dengan baik saat menggunakan data yang besar, serta memiliki kemajuan arsitektur jaringan *neural* dan komputasi yang lebih kuat dibandingkan dengan algoritma *machine learning* lainnya.

Terdapat banyak macam metode *deep learning* untuk mengklasifikasi teks, salah satunya adalah metode *Long Short-Term Memory* (LSTM). Metode ini dapat memproses data secara sekuensial seperti teks, suara, dan video (Nurrohmat & SN, 2019). Pada penelitian yang dilakukan oleh (Ivan & Azhar, 2022) yang berjudul Deteksi Depresi Pengguna Twitter Indonesia Menggunakan *Long Short-Term Memory-Recurrent Neural Network* (LSTM-RNN), LSTM menghasilkan akurasi yang tinggi jika dibandingkan dengan penggunaan metode lainnya dalam proses klasifikasi. Penelitian lainnya yang dilakukan oleh (Zhou & Xu, 2016) yang berjudul *Amazon Food Review Classification using Deep Learning and*

Recommender System, akurasi paling tinggi yaitu pada LSTM dengan ukuran *embedding* 100, rasio regulasi 0.001, *learning rate* 0.001 dan probabilitas *dropout* 0.9 sebesar 0.89 %.

Long Short-Term Memory (LSTM) adalah sebuah jenis arsitektur jaringan syaraf tiruan (*neural network*) yang khusus dirancang untuk mengatasi masalah dependensi jarak jauh dalam data berurutan. LSTM merupakan bentuk khusus dari *Recurrent Neural Networks* (RNN) yang memiliki kemampuan khusus dalam mengingat dan memproses informasi jangka panjang dalam data berurutan. LSTM mampu mengatasi masalah dependensi jarak jauh dengan lebih baik daripada RNN tradisional (Hochreiter dan Schmidhuber, 1997). Oleh karena itu, LSTM telah berhasil digunakan dalam berbagai tugas pemrosesan bahasa alami seperti analisis sentimen, penerjemahan mesin, dan lainnya, serta dalam tugas-tugas yang melibatkan data berurutan seperti prediksi harga saham atau prediksi cuaca.

Berdasarkan penjelasan pada paragraf-paragraf sebelumnya dapat disimpulkan bahwa penelitian ini bertujuan untuk mengembangkan sistem klasifikasi berita menggunakan metode *Long Short-Term Memory* (LSTM).

1.3 Rumusan Masalah

Berdasarkan penjelasan latar belakang sebelumnya, rumusan masalah dari penelitian ini adalah bagaimana mengklasifikasikan teks dengan menggunakan metode *Long Short—Term Memory*. Terdapat beberapa *research question* (RQ) dalam penelitian ini yaitu sebagai berikut :

1. Bagaimana mengembangkan sistem klasifikasi berita menggunakan metode *Long Short-Term Memory (LSTM)*?
2. Bagaimana mengukur kinerja model LSTM dalam mengklasifikasi berita pada teks berdasarkan matrik evaluasi *Accuracy, Precision, Recall*, dan *F-1 Score*?

1.4 Tujuan Penelitian

1. Menghasilkan sistem klasifikasi berita pada teks yang sudah dikembangkan menggunakan metode *Long Short-Term Memory (LSTM)*.
2. Mengetahui hasil kinerja metode *Long Short-Term Memory (LSTM)* untuk klasifikasi berita pada teks berdasarkan matrik evaluasi *Accuracy, Precision, Recall*, dan *F-1 Score*.

1.5 Manfaat Penelitian

1. Membantu dalam mengorganisasi dan mengelompokkan berita berdasarkan kategori atau topik tertentu, sehingga dapat dengan mudah mencari informasi yang diinginkan tanpa perlu menyaring banyak berita secara manual.
2. Hasil penelitian dapat digunakan sebagai referensi penelitian berikutnya, peneliti berikutnya dapat memperluas pengetahuan mengenai penggunaan LSTM dalam klasifikasi teks dan berkontribusi pada

pengembangan metode klasifikasi yang lebih canggih dan efisien.

1.6 Batasan Masalah

1. Data hanya terdiri dari 5 kelas/kategori yaitu *finance*, *travel*, *health*, *food*, dan *sport*.
2. Data yang digunakan berupa file berformat .csv.
3. Bahasa yang digunakan adalah bahasa Indonesia.

1.7 Sistematika Penulisan

Penelitian ini ditulis dalam beberapa subbagian yang menjelaskan secara detail, jelas, dan tepat mengenai hal yang didapatkan peneliti ketika melakukan penelitian. Penjelasan – penjelasan tersebut disusun dalam masing – masing bab pada penelitian. Adapun sistematika penulisan pada penelitian ini adalah sebagai berikut :

BAB I. PENDAHULUAN

Bab ini menguraikan landasan dari penelitian , seperti latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah serta sistematika penulisan.

BAB II. KAJIAN LITERATUR

Bab ini membahas mengenai dasar – dasar teori yang yang berkaitan dengan penelitian, seperti klasifikasi teks, pra-pengolahan teks, dan algoritma *Long Short-Term Memory*. Serta menguraikan penelitian –

penelitian terdahulu yang relevan terhadap penelitian ini.

BAB III. METODOLOGI PENELITIAN

Bab ini mendeskripsikan mengenai data yang digunakan dalam penelitian dan tahapan yang akan dilaksanakan dalam penelitian ini. Pada akhir bab, akan diuraikan perancangan manajemen proyek untuk pelaksanaan penelitian.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Pada bab ini membahas tahapan pengembangan perangkat lunak, diawali dengan analisis kebutuhan, perancangan dan konstruksi perangkat lunak, kemudian diakhiri dengan pengujian untuk memastikan apakah sistem yang dikembangkan sesuai dengan rancangan dan kebutuhan penelitian.

BAB V. HASIL DAN ANALISIS PENELITIAN

Bab ini berisi hasil pengujian berdasarkan perencanaan yang telah dilakukan. Tabel untuk hasil pengujian yang dibuat akan digunakan sebagai acuan untuk menyusun kesimpulan untuk bab selanjutnya.

BAB VI. KESIMPULAN DAN SARAN

Pada bab ini menyimpulkan hasil penelitian yang didapatkan secara rinci serta memberikan saran-saran terkait penelitian yang dapat dilakukan selanjutnya.

1.8 Kesimpulan

Berdasarkan penjelasan diatas, dapat disimpulkan bahwa masalah yang perlu diselesaikan adalah mengetahui akurasi paling baik pada klasifikasi menggunakan algoritma *Long Short-Term Memory (LSTM)*.

DAFTAR PUSTAKA

- Andriani, A. 2012. Penerapan Algoritma C4. 5 Pada Program Klasifikasi Mahasiswa Dropout. In Seminar Nasional Matematika (pp. 139-147).
- Bunkhumpornpat, C., Sinapiromsaran, K., & Lursinsap, C. 2009. Safe-level-smote: Safe-level-synthetic minority over-sampling technique for handling the class imbalanced problem. In *Advances in Knowledge Discovery and Data Mining: 13th Pacific-Asia Conference, PAKDD 2009 Bangkok, Thailand, April 27-30, 2009 Proceedings* 13 (pp. 475-482). Springer Berlin Heidelberg.
- Chandra, D. N., Indrawan, G., & Sukaraja, I. N. 2016. Klasifikasi Berita Lokal Radar Malang Menggunakan Metode Naïve Bayes Dengan Fitur N-Gram. *Jurnal Ilmiah Teknologi Informasi Asia*, 10(1), 11-19.
- Chevalier, G. 2018. Schematic of the Long-Short Term Memory cell, a component of recurrent neural networks.
- Darujati, C., & Gumelar, A. B. 2012. Pemanfaatan teknik supervised untuk klasifikasi teks bahasa indonesia. *Jurnal Bandung Text Mining*, 16(1), 5-1.
- Fauziah, D. A., Maududie, A., & Nuritha, I. 2018. Klasifikasi Berita Politik Menggunakan Algoritma K-nearest Neighbor. *Berkala Sainstek*, 6(2), 106-114.
- Feldman, R., & Sanger, J. 2007. *The text mining handbook: advanced approaches in analyzing unstructured data*. Cambridge university press.
- Handayani, F., & Pribadi, F. S. 2015. Implementasi algoritma naive bayes classifier dalam pengklasifikasian teks otomatis pengaduan dan pelaporan masyarakat melalui layanan call center 110. *Jurnal Teknik Elektro*, 7(1), 19-24.
- Hochreiter, S., & Schmidhuber, J. 1997. Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- Irham, L. G., Adiwijaya, A., & Wisesty, U. N. 2019. Klasifikasi Berita Bahasa Indonesia Menggunakan Mutual Information dan Support Vector Machine. *Jurnal Media Informatika Budidarma*, 3(4), 284-292.
- Ivan Dwi Nugraha and Azhar, Y. 2022. Deteksi Depresi Pengguna Twitter Indonesia Menggunakan LSTM-RNN. *Jurnal Nasional Pendidikan Teknik Informatika : JANAPATI*. 11, 3 (Dec. 2022), 320–329. DOI: <https://doi.org/10.23887/janapati.v11i3.50674>.
- Jones, K.S., 1973. Indexing term weighting, *Inf. Storage Retr*, 9, Pp. 619-633.

- Khan, A., Baharudin, B., Lee, L. H., & Khan, K. 2010. A review of machine learning algorithms for text-documents classification. *Journal of advances in information technology*, 1(1), 4-20.
- Kurniawan, B., Fauzi, M. A., & Widodo, A. W. 2017. Klasifikasi Berita Twitter Menggunakan Metode Improved Naïve Bayes. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 1(10), 1193-1200.
- Li-Juan, Z., Feng, Z., Qing-Qing, P., Xin, Y., & Zheng-Tao, Y. 2015. A classification method of Vietnamese news events based on maximum entropy model. In 2015 34th Chinese Control Conference (CCC) (pp. 3981-3986). IEEE.
- M. A. Nurrohmat And A. Sn. 2019. Sentiment Analysis Of Novel Review Using Long Short-Term Memory Method. *Ijccs (Indonesian J. Comput. Cybern. Syst., Vol. 13, No. 3, P. 209, Doi: 10.22146/Ijccs.41236*.
- Pratama, E. D. 2022. Implementasi Model Long-Short Term Memory (LSTM) pada Klasifikasi Teks Data SMS Spam Berbahasa Indonesia. *The Journal on Machine Learning and Computational Intelligence (JMLCI)*, 1(2).
- Rizaldy, A., & Santoso, H. A. 2017. Performance improvement of Support Vector Machine (SVM) With information gain on categorization of Indonesian news documents. 2017 International Seminar on Application for Technology of Information and Communication (iSemantic).
- Rofiqi, A. Y. 2017. Clustering Berita Olahraga Berbahasa Indonesia Menggunakan Metode K-Medoid Bersyarat. *Jurnal Simantec*, 6(1).
- Saadah, M. N., Atmagi, R. W., Rahayu, D. S., & Arifin, A. Z. (2013). Sistem temu kembali dokumen teks dengan pembobotan tf-idf dan LCS. *JUTI J. Ilm. Teknol. Inf*, 11(1), 19.
- Samodra, J., Sumpeno, S., & Hariadi, M. 2009. Klasifikasi dokumen teks berbahasa Indonesia dengan menggunakan Naive Bayes. In Seminar nasional electrical, informatics, and it's education.
- Wahyudi, M. D. R. 2019. Penerapan Algoritma Cosine Similarity pada Text Mining Terjemah Al-Qur'an Berdasarkan Keterkaitan Topik. *Semesta Teknika*, 22(1), 41-50.
- Zhou, Z., & Xu, L. 2009. Amazon Food review classification using deep learning and recommender system. In Stanford University Stanford.