

**ANALISIS SENTIMEN PADA PORTAL BERITA  
ONLINE INTERNASIONAL TENTANG PARIWISATA  
BALI MENGGUNAKAN METODE LSTM**



**OLEH  
NGAKAN MADE KRISNA SEDANA  
NIM 1915101025**

**JURUSAN TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KEJURUAN  
UNIVERSITAS PENDIDIKAN GANESHA  
SINGARAJA**

**2024**

**ANALISIS SENTIMEN PADA PORTAL BERITA  
ONLINE INTERNASIONAL TENTANG PARIWISATA  
BALI MENGGUNAKAN METODE LSTM**

**SKRIPSI**



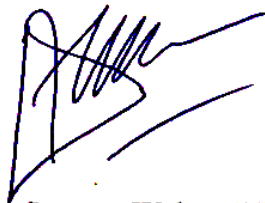
**JURUSAN TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KEJURUAN  
UNIVERSITAS PENDIDIKAN GANESHA  
SINGARAJA**

**2024**

**DIAJUKAN UNTUK MELENGKAPI TUGAS  
DAN MEMENUHI SYARAT-SYARAT UNTUK  
MENCAPAI GELAR SARJANA**

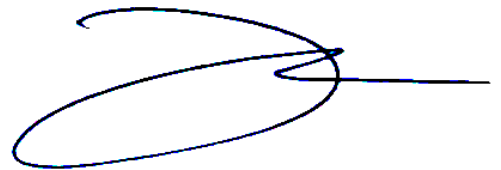
**Menyetujui**

**Pembimbing I**



I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs.  
NIP. 198910262019031004

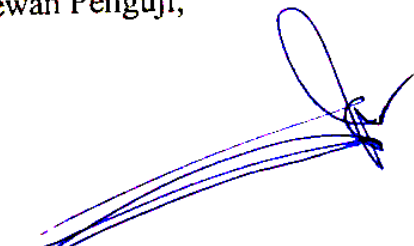
**Pembimbing II**



I Ketut Resika Arthana, S.T., M.Kom.  
NIP. 198412012012121002


**Skripsi oleh Ngakan Made Krisna Sedana ini**  
telah di pertahankan di depan dewan penguji  
pada tanggal 29 Januari 2024

Dewan Penguji,




Agus Aan Jiwa Permana, S.Kom., M.Cs.  
NIP. 19870804 201504 1 001

( Ketua )



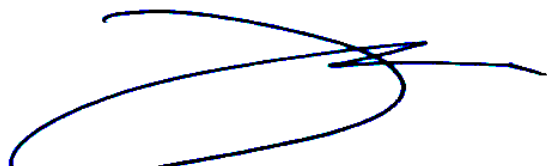
Dr. Ni Ketut Kertiasih, S.Si., M.Pd.  
NIP. 19701118 199703 2 001

( Anggota )



I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs.  
NIP. 19891026 201903 1 004

( Anggota )



I Ketut Resika Arthana, S.T., M.Kom.  
NIP. 19841201 201212 1 002

( Anggota )

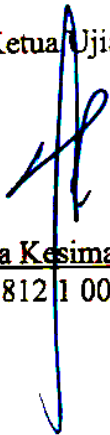
Diterima oleh Panitia Ujian Fakultas Teknik dan Kejuruan  
Universitas Pendidikan Ganesha  
Guna memenuhi syarat-syarat untuk mencapai gelar sarjana

Pada:

Hari : Senin  
Tanggal : 29 Januari 2024

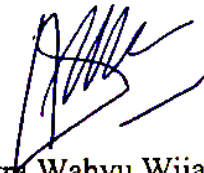
**Mengetahui,**

Ketua Ujian,



Made Windu Antara Kesiman, S.T., M.Sc., Ph.D.  
NIP. 19821111 200812 1 001

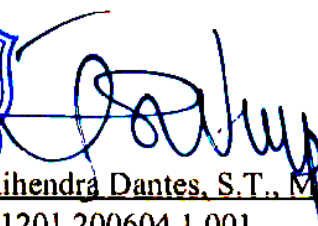
Sekretaris Ujian,



I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs.  
NIP. 19891026 201903 1 004

**Mengesahkan,**

**Dekan Fakultas Teknik dan Kejuruan**



Dr. Kadek Rihendra Dantes, S.T., M.T.  
NIP. 19791201 200604 1 001

## Pernyataan

Dengan ini saya menyatakan bahwa karya tulis yang berjudul “**Analisis Sentimen pada Portal Berita *Online* Internasional tentang Pariwisata Bali Menggunakan Metode LSTM**” beserta keseluruhan isinya adalah benar-benar hasil karya sendiri dan tidak melakukan penjiplakan dan pengutipan dengan cara yang tidak sesuai dengan etika yang berlaku pada buku pendoman skripsi . Jika dikemudian hari ditemukan adanya pelanggaran saya siap menanggung resiko atau sanksi yang di jatuhkan kepada saya.

Singaraja, 29 Januari 2024

Yang membuat pernyataan,



Ngakan Made Krisna Sedana

NIM. 1915101025

## PRAKATA

Puji syukur peneliti panjatkan kehadapan Ida Sang Hyang Widhi Wasa/Tuhan Yang Maha Esa karena berkat rahmat-Nya lah sehingga memungkinkan peneliti menyelesaikan skripsi dengan judul "Analisis Sentimen pada Portal Berita Online Internasional tentang Pariwisata Bali Menggunakan Metode LSTM" sebagai salah satu syarat untuk memperoleh gelar sarjana Ilmu Komputer di Universitas Pendidikan Ganesha. Dalam menyelesaikan skripsi ini, peneliti banyak mendapatkan motivasi dan bantuan baik berupa moral maupun material dari berbagai pihak. Untuk itu dalam kesempatan ini peneliti mengucapkan rasa terima kasih kepada :

1. Bapak Prof. Dr. I Wayan Lasmawan, M.Pd., selaku Rektor Universitas Pendidikan Ganesha beserta stafnya yang telah memberikan motivasi dan fasilitas sehingga peneliti dapat menyelesaikan studi dan melakukan penelitian sesuai dengan rencana yang dicanangkan.
2. Bapak Dr. Kadek Rihendra Dantes, S.T., M.T., selaku Dekan Fakultas Teknik dan Kejuruan atas motivasi dan fasilitas yang diberikan sehingga peneliti dapat menyelesaikan studi sesuai dengan rencana.
3. Bapak Putu Hendra Suputra, S.Kom., M.Cs., selaku Ketua Jurusan Teknik Informatika atas motivasi dan fasilitas yang diberikan sehingga peneliti dapat menyelesaikan studi ini dengan tepat waktu.
4. Bapak I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs., selaku Koorprodi Ilmu Komputer dan selaku pembimbing I yang telah memberikan banyak motivasi, bimbingan, arahan, petunjuk skripsi, dengan penuh kesabaran kepada peneliti dalam menyelesaikan skripsi ini sehingga peneliti dapat menyelesaikan skripsi dengan tepat waktu.
5. Bapak I Ketut Resika Arthana, S.T., M.Kom., selaku pembimbing II yang telah banyak memberikan motivasi, semangat, bimbingan, arahan dengan penuh kesabaran kepada peneliti dalam menyelesaikan skripsi ini sehingga peneliti dapat menyelesaikan skripsi dengan tepat waktu.

6. Seluruh staf dosen di lingkungan Jurusan Teknik Informatika/ Program Studi Ilmu Komputer yang telah memberikan ilmu, pengalaman, motivasi, serta semangat selama menjalani studi di Universitas Pendidikan Ganesha
7. Bapak I Putu Edy Suardiyana Putra, PH.D., selaku CEO Bit House yang telah banyak memberikan ide kepada peneliti dalam menyelesaikan skripsi ini sehingga peneliti dapat menyelesaikan skripsi dengan tepat waktu.
8. Ungkapan terimakasih dan penghargaan yang sangat spesial peneliti haturkan kepada kedua orang tua tercinta peneliti (Bpk. Ngakan Nyoman Oka dan Ibu Desak Putu Griastini) serta kedua saudara peneliti, yang selalu memberikan doa restu, nasihat, dukungan, semangat, bimbingan dan motivasi serta cinta kasih kepada peneliti, sehingga skripsi ini dapat selesai dengan tepat waktu.
9. Untuk seluruh teman – teman prodi Ilmu Komputer Angkatan 2019 yang telah memberikan berbagai pengalaman dan kebersamaan selama peneliti menempuh pendidikan di Universitas Pendidikan Ganesha.

Semoga Ida Sang Hyang Widhi Wasa/Tuhan Yang Maha Esa senantiasa memberikan karunianya dan membantu peneliti dalam menyelesaikan skripsi ini. Peneliti menyadari sepenuhnya bahwa apa yang disajikan dalam skripsi ini masih jauh dari kata sempurna mengingat keterbatasan peneliti. Peneliti mengharapkan segala masukan, kritik dan saran yang bersifat membangun dari berbagai pihak. Peneliti berharap apa yang dibuat oleh peneliti dapat bermanfaat dan berharga bagi kita semua, khususnya untuk mengembangkan dunia Pendidikan

Singaraja, 29 Januari 2024

Peneliti



## DAFTAR ISI

PRAKATA.....	vii
ABSTRAK.....	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL.....	xiv
BAB I.....	1
1.1. Latar belakang.....	1
1.2. Rumusan Masalah.....	5
1.3. Tujuan.....	5
1.4. Batasan Masalah.....	6
1.5. Manfaat.....	6
BAB II.....	8
2.1. <i>Web Scrapping</i> .....	8
2.2. <i>Preprocessing Text</i> .....	8
2.2.1. Cleaning.....	9
2.2.2. Case Folding.....	9
2.2.3. Stopword Removal.....	9
2.2.4. Stemming.....	9
2.3. <i>LSTM (Long Short-Term Memory)</i> .....	10
2.3.1. <i>Forget Gate</i> .....	11
2.3.2. <i>Input Gate</i> .....	11
2.3.3. <i>Output Gate</i> .....	11
2.4. Analisis Sentimen.....	12
2.5. <i>Confusion Matrix</i> .....	13
2.6. <i>LDA (Latent Dirichlet Allocation)</i> .....	15
2.7. Penelitian Terkait.....	16
BAB III.....	20

3.1.	Alur Penelitian.....	20
3.2.	<i>Scraping Data Website</i> .....	21
3.3.	Pelabelan Data.....	22
3.4.	<i>Preprocessing Data</i> .....	23
3.4.1.	<i>Cleaning</i> .....	24
3.4.2.	<i>Case Folding</i> .....	25
3.4.3.	<i>Stopwords Removal</i> .....	26
3.4.4.	<i>Stemming</i> .....	27
3.5.	Pembuatan Model LSTM .....	28
3.5.1.	Ekstraksi Fitur .....	29
3.5.2.	<i>Splitting Data</i> .....	32
3.5.3.	Pelatihan Model .....	32
3.5.4.	Pengujian Model .....	34
3.6.	<i>Confusion Matrix</i> .....	34
3.7.	Pemodelan Topik LDA .....	34
BAB IV	.....	36
4.1.	Pengumpulan Data .....	36
4.2.	Pembuatan Dataset dan Pelabelan .....	37
4.3.	<i>Preprocessing Data</i> .....	40
4.3.1.	<i>Cleaning</i> .....	40
4.3.2.	<i>Case Folding</i> .....	41
4.3.3.	<i>Stopwords Removal</i> .....	42
4.3.4.	<i>Stemming</i> .....	43
4.4.	Membangun Model .....	45
4.4.1.	Ekstraksi Fitur .....	45
4.4.2.	Pembagian dataset.....	47
4.4.3.	Pembuatan Model.....	48
4.4.4.	Pelatihan Model .....	50
4.4.5.	Pengujian Model .....	52
4.4.6.	Evaluasi <i>Confusion Matrix</i> .....	54
4.5.	Analisis Topic Modeling LDA.....	59
4.6.	Analisis Hasil dan Pembahasan Penelitian.....	68

BAB V.....	70
5.1 Kesimpulan.....	70
5.2 Saran.....	72
DAFTAR PUSTAKA.....	73
LAMPIRAN.....	77
Riwayat Hidup.....	86



## DAFTAR GAMBAR

Gambar 1. 1 Persentase Sumber Membaca Berita di Inggris.....	2
Gambar 1. 2 Persentase Sumber Membaca Berita di Australia .....	3
Gambar 2. 1 Arsitektur LSTM .....	12
Gambar 2. 2 <i>Coincidence Matrix</i> .....	13
Gambar 2. 3 Representasi Model LDA .....	15
Gambar 3. 1 Alur Penelitian.....	20
Gambar 3. 2 Alur Proses <i>Scraping</i> .....	21
Gambar 3. 3 Contoh Data <i>Input</i> .....	29
Gambar 3. 4 Contoh <i>Stack</i> Model LSTM .....	33
Gambar 4. 1 Halaman Isi Berita.....	36
Gambar 4. 2 Persentase Distribusi Jumlah Label Sentimen .....	39
Gambar 4. 3 Proses <i>Training Testing</i> Model .....	51
Gambar 4. 4 Grafik <i>Training Testing</i> Model .....	51
Gambar 4. 5 <i>Confusion Matrix</i> .....	55
Gambar 4. 6 Grafik Hasil <i>Coherence Score</i> dan Jumlah Topik untuk Berita Sentiment Positif .....	60
Gambar 4. 7 Grafik Hasil <i>Coherence Score</i> dan Jumlah Topik untuk Berita Sentiment Netral.....	61
Gambar 4. 8 Grafik Hasil <i>Coherence Score</i> dan Jumlah Topik untuk Berita Sentiment Negatif.....	62
Gambar 4. 9 Hasil Pemodelan LDA Berita Sentimen Positif .....	63
Gambar 4. 10 Hasil Pemodelan LDA Berita Sentimen Netral.....	64
Gambar 4. 11 Hasil Pemodelan LDA Berita Sentimen Negatif.....	64
Gambar 4. 12 Hasil Visualisasi LDA dengan <i>Intertopic Distance Map</i> pada Berita Positif .....	66
Gambar 4. 13 Hasil Visualisasi LDA dengan <i>Intertopic Distance Map</i> pada Berita Netral.....	67
Gambar 4. 14 Hasil Visualisasi LDA dengan <i>Intertopic Distance Map</i> pada Berita Negatif.....	67

## DAFTAR TABEL

Tabel 3. 1 Contoh Data <i>Scraping</i> .....	21
Tabel 3. 2 Contoh Label.....	23
Tabel 3. 3 Contoh <i>Cleaning</i> .....	24
Tabel 3. 4 Contoh <i>Case Folding</i> .....	25
Tabel 3. 5 Contoh <i>Stopwords</i> .....	26
Tabel 3. 6 Contoh <i>Stemming</i> .....	27
Tabel 3. 7 Contoh Tokenisasi Data.....	30
Tabel 3. 8 Contoh Hasil <i>Padding</i> Data .....	31
Tabel 4. 1 Hasil Data Berita dari <i>Scraping</i> .....	37
Tabel 4. 2 Hasil Pelabelan.....	38
Tabel 4. 3 Distribusi Jumlah Label Sentimen .....	38
Tabel 4. 4 Hasil <i>Cleaning</i> .....	40
Tabel 4. 5 Hasil <i>Case Folding</i> .....	41
Tabel 4. 6 Hasil <i>Stopwords Removal</i> .....	42
Tabel 4. 7 Hasil <i>Stemming</i> .....	44
Tabel 4. 8 Hasil Tokenisasi.....	45
Tabel 4. 9 Hasil <i>Padding</i> .....	46
Tabel 4. 10 Pembagian Dataset.....	48
Tabel 4. 11 Akurasi Setiap Data Pelatihan dan Uji.....	52
Tabel 4. 12 Hasil Pengujian Data Baru .....	53
Tabel 4. 13 <i>Confusion Matrix</i> .....	55
Tabel 4. 14 Distribusi Kata Berita Sentimen Positif.....	64
Tabel 4. 15 Distribusi Kata Berita Sentimen Netral .....	64
Tabel 4. 16 Distribusi Kata Berita Sentimen Negatif .....	65