

ABSTRAK

Tradisi Kerapan Sapi Madura merupakan warisan budaya yang masih dilestarikan hingga saat ini, meskipun menghadapi beragam persepsi masyarakat seiring dengan perubahan sosial dan meningkatnya kesadaran terhadap isu hak hewan. Untuk memahami dinamika persepsi publik secara lebih mendalam, penelitian ini bertujuan untuk mengidentifikasi dan mengklasifikasikan sentimen masyarakat terhadap tradisi Kerapan Sapi dengan memanfaatkan pendekatan analisis sentimen berbasis data digital. Metode yang diterapkan dalam penelitian ini merupakan *Long Short-Term Memory* (LSTM), sebuah algoritma deep learning yang memiliki keunggulan dalam memproses data teks berurutan. LSTM digunakan untuk mengklasifikasikan sentimen menjadi tiga kategori, yaitu positif, negatif, dan netral. Data dikumpulkan dari media sosial TikTok untuk mencerminkan opini publik secara luas. Hasil penelitian menunjukkan bahwa model LSTM setelah penerapan *Random Oversampling* memperoleh akurasi sebesar 78,52%. Nilai precision untuk masing-masing kelas, yaitu negatif, netral, dan positif, berturut-turut sebesar 0,88; 0,71; dan 0,78. Sementara itu, nilai recall untuk ketiga kelas tersebut adalah 0,82; 0,76; dan 0,78. Adapun nilai F1-score yang dihasilkan pada kelas negatif, netral, dan positif masing-masing sebesar 0,85; 0,73; dan 0,78. Berdasarkan hasil evaluasi tersebut, model LSTM dengan *Random Oversampling* mampu mengklasifikasikan sentimen pada data komentar terkait tradisi Kerapan Sapi Madura dengan performa yang baik.

Kata Kunci: Kerapan Sapi Madura, Analisis Sentimen, Media Sosial, *Long Short-Term Memory* (LSTM), Budaya Lokal

ABSTRACT

The Madura Kerapan Sapi tradition is a cultural heritage that is still preserved to this day, despite facing various public perceptions along with social changes and increasing awareness of animal rights issues. To understand the dynamics of public perception more deeply, this study aims to identify and classify public sentiment towards the Kerapan Sapi tradition by utilizing a digital data-based sentiment analysis approach. The method applied in this study is Long Short-Term Memory (LSTM), a deep learning algorithm that has advantages in processing sequential text data. LSTM is used to classify sentiment into three categories: positive, negative, and neutral. Data was collected from the social media platform TikTok to broadly reflect public opinion. The research results show that the LSTM model, after the application of Random Oversampling, achieved an accuracy of 78.52%. The precision values for each class, namely negative, neutral, and positive, are 0.88; 0.71; and 0.78, respectively. Meanwhile, the recall values for these three classes are 0.82; 0.76; and 0.78. The F1-score values for the negative, neutral, and positive classes are 0.85; 0.73; and 0.78, respectively. Based on these evaluation results, the LSTM model with Random Oversampling is able to classify sentiments in comments related to the Madura Bull Racing tradition with good performance.

Keywords: Madura Bull Racing, Sentiment Analysis, Social Media, Long Short-Term Memory (LSTM), Local Culture

KATA PENGANTAR

Penulis memanjatkan puji syukur ke hadirat Allah SWT atas segala rahmat dan karunia-Nya, sehingga skripsi yang berjudul “Analisis Sentimen Publik terhadap Tradisi Kerapan Sapi Madura Menggunakan Metode *Long Short-Term Memory* (LSTM)” dapat disusun dengan baik. Skripsi ini dibuat sebagai salah satu prasyarat untuk menyelesaikan pendidikan di program gelar sarjana dalam bidang Sistem Informasi.

Penelitian ini dilatarbelakangi oleh berkembangnya opini publik di media sosial mengenai tradisi Kerapan Sapi Madura. Tradisi ini memunculkan berbagai pandangan, baik dari segi budaya maupun etika. Untuk memahami beragam persepsi tersebut secara lebih menyeluruh, penulis menggunakan metode analisis sentimen berlandaskan deep learning, terutama *Long Short-Term Memory* (LSTM), yang dikenal efektif dalam mengolah dan memahami pola dalam data teks.

Dalam proses pembuatan skripsi ini, penulis mengakui bahwa banyak tantangan yang harus dihadapi, baik terkait aspek teknis, pengumpulan sumber, maupun pengolahan informasi. Namun, berkat bantuan dari banyak pihak, baik dosen pembimbing, rekan-rekan mahasiswa, maupun keluarga, Skripsi ini akhirnya berhasil diselesaikan dengan baik. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih atas semua dukungan, saran, dan dorongan yang telah diterima.

Penulis berkeinginan agar karya tulis ini bisa memberikan kontribusi dalam dunia akademis, terutama dalam sektor analisis sentimen serta konservasi budaya lokal yang didukung oleh teknologi. Penulis juga memahami bahwa studi ini memiliki batasan-batasan, sehingga umpan balik dan masukan yang konstruktif sangat dinantikan untuk pengembangan penelitian di waktu yang akan datang.

Bangkalan, April 2026

Akhmad Ramdani Syam
Universitas Trunojoyo Madura

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN ORISINALITAS.....	iii
ABSTRAK.....	iv
<i>ABSTRACT</i>	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah.....	4
1.2.1 Permasalahan.....	4
1.2.2 Metode Usulan	4
1.2.3 Pertanyaan Penelitian	4
1.3 Tujuan dan Manfaat.....	4
1.3.1 Tujuan Penelitian	4
1.3.2 Manfaat Penelitian	5
1.3 Batasan Masalah.....	5
1.5 Sistemika Penulisan Proposal Skripsi	5
BAB II KAJIAN PUSTAKA	7
2.1 Analisis Sentimen.....	7
2.2 Media Sosial TikTok	8
2.3 Tradisi Kerapan Sapi Madura.....	9
2.4 <i>Text Mining</i> dan Pra-Pemrosesan Data Teks.....	10
2.5 <i>Long Short-Term Memory</i> (LSTM).....	12
2.5.1 Arsitektur LSTM.....	13
2.5.2 Penerapan LSTM dalam Analisis Sentimen	18
2.6 Evaluasi Kinerja Model.....	19
2.7 Penelitian Terkait	21

BAB III METODE PENELITIAN.....	26
3.1 Alur Penelitian.....	26
3.2 Metode Pengumpulan Data	26
3.2.1 Studi Litetatur	27
3.2.2 Studi Lapangan.....	27
3.3 Implementasi	27
3.3.1 Pengumpulan Data	27
3.3.2 <i>Preprocessing</i>	28
3.3.3 Labelling	34
3.3.4 Perhitungan LSTM.....	34
3.3.5 Pelatihan LSTM	37
3.3.6 Pengujian LSTM	39
BAB IV HASIL DAN PEMBAHASAN	40
4.1 Pengumpulan dan Pengolahan Data	40
4.1.1 Sumber dan Teknik Pengumpulan Data.....	40
4.1.2 Deskripsi Dataset	40
4.1.3 Tahapan <i>Preprocessing</i> Data	41
4.1.4 Hasil <i>Preprocessing</i> Data	46
4.1.5 Pelabelan Data (<i>Labelling</i>).....	47
4.1.6 Analisis Distribusi Data	48
4.2 Pembahasan dan Analisis Hasil Penelitian.....	49
4.2.1 Implementasi Model LSTM.....	49
4.2.2 Pelatihan Model (<i>Training</i>) Sebelum Random Oversampling	52
4.2.3 Pelatihan Model (<i>Training</i>) Setelah Random Oversampling.....	53
4.2.4 Hasil Pengujian Model Sebelum Random Oversampling.....	55
4.2.4 Hasil Pengujian Model Setelah Random Oversampling.....	57
4.2.5 Analisis Performa Model Sebelum Random Oversampling	58
4.2.6 Analisis Performa Model Setelah Random Oversampling	59
4.2.7 Analisis Sentimen Publik	61
4.2.8 Analisis Kata/Frasa Dominan	61
4.2.9 Pembahasan Hasil Penelitian	63
4.2.10 Implikasi Penelitian.....	64
BAB V PENUTUP.....	65
5.1 Kesimpulan.....	65

5.2	Saran.....	65
DAFTAR PUSTAKA		67
LAMPIRAN.....		71

DAFTAR GAMBAR

Gambar 2. 1 Arsitektur LSTM	13
Gambar 2. 2 Forget Gate	14
Gambar 2. 3 Input Gate	15
Gambar 2. 4 Cell State	16
Gambar 2. 5 Output Gate	17
Gambar 3. 1 Alur Penelitian.....	26
Gambar 4. 1 Souce Code Cleaning	42
Gambar 4. 2 Source Code Case Folding	43
Gambar 4. 3 Source Code Tokenizing	43
Gambar 4. 4 Source Code Normalisasi Kata	44
Gambar 4. 5 Source Code Stopword Removal	45
Gambar 4. 6 Source Code Stemming	45
Gambar 4. 7 Wordcloud Sebelum dan Sesudah Preprocessing.....	47
Gambar 4. 8 Diagram Pelabelan Data.....	48
Gambar 4. 9 Jumlah Data Training dan Testing Sebelum Random Oversampling	52
Gambar 4. 10 Grafik Accuraccy Sebelum Random Oversampling	53
Gambar 4. 11 Jumlah Data Training dan Testing Setelah Random Oversampling	53
Gambar 4. 12 Grafik Accuraccy Setelah Random Oversampling	54
Gambar 4. 13 Confusion Matrix Sebelum Random Oversampling	56
Gambar 4. 14 Confusion Matrix Setelah Random Oversampling	57
Gambar 4. 15 Perbandingan Analisis Sentimen Annual vs Predict	59
Gambar 4. 16 Perbandingan Analisis Sentimen Annual vs Predict Setelah Random Oversampling	60
Gambar 4. 17 Frekuensi Kata Sebelum dan Sesudah Preprocessing	62
Gambar 4. 18 Wordcloud Positif	62

DAFTAR TABEL

Tabel 2.1 Penelitian Terkait	21
Tabel 3. 1 <i>Dataset</i>	28
Tabel 3. 2 Hasil Dataset Setelah <i>Cleaning</i>	28
Tabel 3. 3 Hasil Dataset Setelah <i>Case Folding</i>	29
Tabel 3. 4 Hasil Dataset Setelah <i>Tokenizing</i>	30
Tabel 3. 5 <i>Hasil Dataset Setelah Normalisasi Kata</i>	31
Tabel 3. 6 Hasil Dataset Setelah <i>Stopword Removal</i>	32
Tabel 3. 7 Hasil Dataset Setelah <i>Stemming</i>	33
Tabel 3. 8 Hasil Pelabelan Data	34
Tabel 3. 9 Hasil <i>Word Embedding</i>	35
Tabel 3. 10 <i>Bobot dan bias</i>	35
Tabel 3. 11 Hasil Perhitungan LSTM	36
Tabel 3. 12 Parameter LSTM	37
Tabel 3. 13 Skenario Pengujian LSTM	38
Tabel 3. 14 Skenario Pengujian LSTM Setelah Random Oversampling	38
Tabel 4. 1 Contoh Dataset	41
Tabel 4. 2 <i>Preprocessing</i> Data	46
Tabel 4. 3 Distribusi Sentimen	47
Tabel 4. 4 Hasil Uji Coba Model LSTM	49
Tabel 4. 5 Hasil Skenario Pengujian LSTM Setelah Random Oversampling	51
Tabel 4. 6 Hasil Evaluasi Model Sebelum Random Oversampling	56
Tabel 4. 7 Hasil Evaluasi Model Setelah Random Oversampling	58

DAFTAR LAMPIRAN

Lampiran 1. Source Code Preprocessing Cleaning.....	71
Lampiran 2. Source Code Preprocessing Normalisasi Kata	71
Lampiran 3. Source Code Preprocessing Case Folding	71
Lampiran 4. Source Code Preprocessing Tokenizing.....	71
Lampiran 5. Source Code Preprocessing Stopword Removal dan Stemming	71
Lampiran 6. Source Code Pelabelan Lexicon	71
Lampiran 7. Source Code Model Training	71
Lampiran 8. Source Code Training dan Testing	71