UNIVERSITAS
GADJAH MADA
ANALISIS KLASIFIKASI SENTIMEN MULTIKELAS PADA DATA TWITTER (X) MENGGUNAKAN CONVOLUTIONAL NEURAL
NETWORK BERBASIS WORD EMBEDDING FASTTEXT DAN GLOVE (Studi Kasus: Busana Muslim Merek Elzatta dan
Dauky)
Fitri Hartanti, Drs. Danardono, MPH., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR ISI

HALAMAN JUDUL i
HALAMAN PENGESAHAN ii
HALAMAN PERNYATAAN iii
HALAMAN PERSEMBAHAN iv
HALAMAN MOTTO v
PRAKATA vi
DAFTAR ISI ix
DAFTAR TABEL xiii
DAFTAR GAMBAR xiv
INTISARI xvi
ABSTRACT xvii
I PENDAHULUAN 1
1.1. Latar Belakang 1
1.2. Rumusan Masalah 3
1.3. Pembatasan Masalah 4
1.4. Tujuan Penelitian 5
1.5. Manfaat Penelitian 6
1.6. Tinjauan Pustaka 7
1.7. Metodologi Penelitian 9
1.8. Sistematika Penulisan 11
II LANDASAN TEORI 13
2.1. Ukuran Pemusatan 13
2.1.1. Rata-rata (Mean) 13
2.1.2. Median 13
2.1.3. Modus 14
2.2. Ukuran Penyebaran 14
2.2.1. Range (Jangkauan) 14
2.2.2. Variansi 15
2.2.3. Standar Deviasi 15
2.2.4. Interquartile Range (IQR) 15
2.3. Probabilitas 15
2.3.1. Ruang Sampel dan Kejadian 16
2.3.2. Aturan Dasar Probabilitas 16

UNIVERSITAS GADJAH MADA
ANALISIS KLASIFIKASI SENTIMEN MULTIKELAS PADA DATA TWITTER (X) MENGGUNAKAN CONVOLUTIONAL NEURAL
NETWORK BERBASIS WORD EMBEDDING FASTTEXT DAN GLOVE (Studi Kasus: Busana Muslim Merek Elzatta dan
Dauky)
Fitri Hartanti, Drs. Danardono, MPH., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

2.3.3. Kejadian Saling Lepas dan Independen . . . 17
2.3.4. Probabilitas Bersyarat dan Teorema Perkalian . . . 17
2.3.5. Hukum Probabilitas Total . . . 18
2.3.6. Teorema Bayes . . . 18
2.4. Variabel Random . . . 18
2.4.1. Variabel Random Diskrit . . . 19
2.4.2. Variabel Random Kontinu . . . 20
2.5. Matriks . . . 20
2.5.1. Penjumlahan dan Pengurangan Matriks . . . 21
2.5.2. Perkalian Skalar . . . 21
2.5.3. Perkalian Matriks . . . 22
2.5.4. Transpos Matriks . . . 22
2.6. Operasi Hadamard . . . 22
2.7. *Machine Learning* . . . 23
2.7.1. Jenis-Jenis *Machine Learning* . . . 24
2.7.2. Data Latih, Validasi, dan Uji . . . 25
2.7.3. *Overfitting* dan *Underfitting* . . . 27
2.7.4. *Cross Validation* dan *Hyperparameter Tuning* . . . 29
2.8. Jaringan Saraf Tiruan (*Artificial Neural Network*) . . . 30
2.8.1. *Perceptron* . . . 33
2.8.2. Fungsi Aktivasi . . . 33
2.8.3. *Multilayer Perceptron* (MLP) . . . 37
2.8.4. *Backpropagation* . . . 38
2.9. *Natural Language Processing* (NLP) . . . 40
2.9.1. *Text Preprocessing* . . . 40
2.9.2. Representasi Teks Tradisional . . . 42
2.9.3. ***Web Scraping* untuk Pengumpulan Data Ulasan** . . . 43
III METODOLOGI PENELITIAN . . . 44
3.1. Analisis Sentimen . . . 44
3.2. Penanganan Data Tidak Seimbang . . . 46
3.2.1. *Random Over-Sampling* (ROS) . . . 47
3.3. *Convolutional Neural Network* (CNN) . . . 48
3.3.1. Lapisan CNN . . . 49
3.3.2. Contoh Implementasi CNN pada Analisis Sentimen Ulasan Produk . . . 51
3.4. *Long Short-Term Memory* (LSTM) . . . 53

UNIVERSITAS GADJAH MADA
ANALISIS KLASIFIKASI SENTIMEN MULTIKELAS PADA DATA TWITTER (X) MENGGUNAKAN CONVOLUTIONAL NEURAL
NETWORK BERBASIS WORD EMBEDDING FASTTEXT DAN GLOVE (Studi Kasus: Busana Muslim Merek Elzatta dan
Dauky)
Fitri Hartanti, Drs. Danardono, MPH., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

3.4.1. Struktur *Cell* . . . 55
3.4.2. Contoh Implementasi LSTM pada Analisis Sentimen . . . 57
3.5. *Bidirectional Long Short-Term Memory* (BiLSTM) . . . 59
3.5.1. Mekanisme Membaca Dua Arah . . . 60
3.5.2. Contoh Implementasi BiLSTM pada Analisis Sentimen . . . 61
3.6. Kombinasi CNN–LSTM . . . 62
3.7. Kombinasi CNN–BiLSTM . . . 64
3.8. *Feature Extraction* . . . 66
3.8.1. Tokenisasi . . . 67
3.8.2. *Padding Sequences* . . . 67
3.8.3. *Word Embedding* . . . 68
3.9. Optimasi Adam . . . 71
3.10. Evaluasi Model . . . 72
3.10.1. *Confusion Matrix* . . . 73
3.10.2. Metrik Evaluasi . . . 74
IV STUDI KASUS . . . 77
4.1. Spesifikasi Perangkat Keras . . . 77
4.2. Deskripsi Data . . . 78
4.3. Pelabelan Kelas Sentimen . . . 81
4.4. Eksplorasi dan Visualisasi Data Asli . . . 82
4.4.1. Pemeriksaan Data Hilang dan Duplikasi . . . 82
4.4.2. Visualisasi Perbandingan Karakteristik Sentimen Ulasan . . . 83
4.4.3. *Text Preprocessing* . . . 85
4.4.4. Analisis Gambaran Opini pada *Tweet* . . . 89
4.4.5. Pembagian Data Latih dan Data Validasi . . . 93
4.4.6. Penanganan Kelas Tidak Seimbang . . . 94
4.4.7. *Feature Extraction* . . . 95
4.5. Pembentukan Model Klasifikasi Sentimen . . . 101
4.5.1. Konfigurasi dan Parameter Model . . . 102
4.5.2. Arsitektur CNN–LSTM dan CNN–BiLSTM . . . 104
4.5.3. Evaluasi Model pada Data Tidak Seimbang . . . 105
4.5.4. Evaluasi Model pada Data Seimbang . . . 107
4.6. Implementasi Model Terbaik pada Data Uji . . . 108
V PENUTUP . . . 112
5.1. Kesimpulan . . . 112
5.2. Saran . . . 113

UNIVERSITAS
GADJAH MADA
ANALISIS KLASIFIKASI SENTIMEN MULTIKELAS PADA DATA TWITTER (X) MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK BERBASIS WORD EMBEDDING FASTTEXT DAN GLOVE (Studi Kasus: Busana Muslim Merek Elzatta dan Dauky)
Fitri Hartanti, Drs. Danardono, MPH., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR PUSTAKA 114
A Lampiran Syntax *Web Scraping* 122
B Lampiran Syntax Eksplorasi dan Visualisasi Data 124
C Lampiran Syntax Tokenisasi 129
D Lampiran Syntax *Padding* 135
E Lampiran Syntax *Embedding* 138
F Lampiran Syntax *Random Over-Sampling* 143
G Lampiran Syntax CNN 146
H Lampiran Syntax LSTM 150
I Lampiran Syntax BiLSTM 154
J Lampiran Syntax CNN-LSTM 158
K Lampiran Syntax CNN-BiLSTM 162
L Lampiran Syntax Implementasi Model Terbaik 167