UNIVERSITAS
GADJAH MADA
ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN RUMAH SAKIT MENGGUNAKAN ROBERTA
Alifia Intan Andrean Nunyai, Prof. Dra. Sri Hartati, M.Sc., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR ISI

HALAMAN PERSETUJUAN LAPORAN TESIS ii
HALAMAN PENGESAHAN iii
PERNYATAAN BEBAS PLAGIASI iv
PRAKATA v
DAFTAR ISI vii
DAFTAR GAMBAR x
DAFTAR TABEL xiii
INTISARI xiv
ABSTRACT xv
BAB I PENDAHULUAN 1
1.1 Latar Belakang Masalah 1
1.2 Rumusan Masalah 4
1.3 Batasan Penelitian 4
1.4 Tujuan Penelitian 5
1.5 Manfaat Penelitian 5
1.6 Kontribusi Penelitian 5
BAB II TINJAUAN PUSTAKA 7
2.1 Perkembangan Model Transformer dalam Analisis Sentimen 7
2.2 Penelitian Analisis Sentimen Berbasis Transformer 9
2.3 Analisis Sentimen pada Ulasan Rumah Sakit 10
BAB III METODOLOGI PENELITIAN 15
3.1 Alur Penelitian 15
3.2 Akuisisi Data 15
3.2.1 Sumber dan Deskripsi Dataset 16
3.2.2 Karakteristik Dataset 16
3.2.3 Pembagian Dataset 20
3.3 Rancangan Metode 21
3.3.1 *Exploratory Data Analysis (EDA)* 23
3.3.2 Data *Preprocessing* 23

UNIVERSITAS
GADJAH MADA
ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN RUMAH SAKIT MENGGUNAKAN ROBERTA
Alifia Intan Andrean Nunyai, Prof. Dra. Sri Hartati, M.Sc., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

3.3.2 *Experiment Setup* 24
3.3.4 Dataset Splitting 26
3.3.5 *DataLoader* 27
3.3.6 *RoBERTa Fine-Tuning* 30
3.3.7 *Training and Validation* 34
3.4 Rancangan Pengujian dan Evaluasi 35
3.4.1 Protokol Evaluasi 35
a. *Accuracy* 35
b. *Macro F1-score* 36
c. *Weighted F1-score* 37
d. Classification Report 37
e. Confusion Matrix 38
3.5 *Result Analysis and Comparison* 39
BAB IV IMPLEMENTASI 40
4.1 Lingkungan Implementasi 40
4.3 Implementasi *Exploratory Data Analysis (EDA)* 43
4.4 *Implementasi Preprocessing Data* 44
4.5 Implementasi *Experiment Setup* 45
4.6 Implementasi *Text Tokenization* 47
4.7 Implementasi *Dataset Splitting* 50
4.8 Implementasi *DataLoader Construction* 52
4.9 Implementasi *Class Weight* 53
4.10 Implementasi *RoBERTa Fine-Tuning* 55
4.11 Implementasi *Training and Validation* 59
4.12 Implementasi Protokol Evaluasi 62
4.12.1 Implementasi Evaluasi Model 62
4.12.2 Implementasi Pengujian Model 62
4.12.3 Implementasi *Result Comparison* 64
BAB V HASIL DAN PEMBAHASAN 66
5.1 Hasil *Exploratory Data Analysis (EDA)* 66
5.1.1 Distribusi Data pada Setiap Aspek 66
5.1.2 Distribusi Sentimen pada Setiap Aspek 67
5.2 Hasil *Data Preprocessing* 69
5.2.1 Hasil Pemeriksaan Struktur Dataset 69

UNIVERSITAS
GADJAH MADA
ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN RUMAH SAKIT MENGGUNAKAN ROBERTA
Alifia Intan Andrean Nunyai, Prof. Dra. Sri Hartati, M.Sc., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

5.2.2 Hasil *Label Encoding* 70
5.3 Hasil *Experiment Setup* 71
5.3.1 Hasil *Text Tokenization* 72
5.4 Hasil *Dataset* Splitting 73
5.5 Hasil *DataLoader Construction* 75
5.6 Hasil *Class Weight* 76
5.7 Hasil *Training RoBERTa Fine-Tuning* 78
5.7.1 Hasil *Fine-Tuning* Aspek SDM 78
5.7.2 Hasil *Fine-Tuning* Aspek Fasilitas 79
5.7.4 Ringkasan Hasil *Fine-Tuning* 81
5.8 Hasil Evaluasi Model 82
5.8.1 Hasil Evaluasi Aspek SDM 83
5.8.2 Hasil Evaluasi Aspek Fasilitas 86
5.8.3 Hasil Evaluasi Aspek Pelayanan 88
5.9. Hasil Accuracy 91
5.10 Hasil Macro F1-score 92
5.11 Hasil Weighted F1-score 94
5.12 Perbandingan Performa Model dengan Penelitian Sebelumnya 96
BAB VI KESIMPULAN DAN SARAN 101
6.1 Kesimpulan 101
6.2 Saran 102
DAFTAR PUSTAKA 104

UNIVERSITAS
GADJAH MADA
ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN RUMAH SAKIT MENGGUNAKAN ROBERTA
Alifia Intan Andrean Nunyai, Prof. Dra. Sri Hartati, M.Sc., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian 15
Gambar 3.2 Total Data pada Setiap Aspek 17
Gambar 3.3 Jumlah Data untuk Sentimen Positif 18
Gambar 3.4 Jumlah Data untuk Sentimen Negatif 18
Gambar 3.5 Jumlah Data untuk Sentimen Netral 19
Gambar 3.6 Jumlah Data untuk Sentimen None 19
Gambar 3.7 Alur Metode Penelitian 21
Gambar 3.8 Pembagian Dataset 27
Gambar 3.9 Arsitektur RoBERTa 31

Gambar 4.1 Implementasi pembacaan dataset 41
Gambar 4.2. Hasil informasi struktur *dataset* 42
Gambar 4.3 menghitung jumlah data pada setiap kategori sentimen 43
Gambar 4.4 Kode program visualisasi distribusi sentimen 44
Gambar 4.5 Kode program konversi label sentimen ke bentuk numerik 45
Gambar 4.6 Kode program penentuan perangkat komputasi 45
Gambar 4.7 Kode program pemilihan tokenizer RoBERTa 46
Gambar 4.8 Kode program pembagian data latih dan data uji. 46
Gambar 4.9 Kode program konfigurasi parameter pelatihan 46
Gambar 4. 10 Kode program penentuan jumlah label dan aspek 47
Gambar 4.11 Kode program tokenisasi teks ulasan 48
Gambar 4.12 Kode program pengambilan hasil tokenisasi 48
Gambar 4.13 Kode program pembuatan kelas dataset 49
Gambar 4.14 Kode program tokenisasi pada kelas dataset 49
Gambar 4. 15 Kode program pembuatan dataset hasil tokenisasi 50
Gambar 4.16 Kode program pembagian dataset secara stratified 50
Gambar 4. 17 Kode program pembuatan objek dataset 51
Gambar 4. 18 Kode program pembuatan DataLoader 51

UNIVERSITAS
GADJAH MADA
ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN RUMAH SAKIT MENGGUNAKAN ROBERTA
Alifia Intan Andrean Nunyai, Prof. Dra. Sri Hartati, M.Sc., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Gambar 4.19 Kode program perhitungan bobot kelas dan penyimpanan informasi hasil pembagian dataset 52
Gambar 4.20 Kode program pembuatan objek dataset 52
Gambar 4.21 Kode program pembuatan DataLoader 53
Gambar 4. 22 Kode program perhitungan class weight 54
Gambar 4.23 Kode program perhitungan performa keseluruhan model 54
Gambar 4.24 Kode program pemanggilan fungsi make_class_weights 54
Gambar 4. 25 Kode program penerapan class weight pada fungsi loss 55
Gambar 4. 26 Kode program pemuatan tokenizer dan model RoBERTa 56
Gambar 4.27 Kode program arsitektur RobertaClassifier 56
Gambar 4.28 Kode program konfigurasi loss, optimizer, dan scheduler 57
Gambar 4.29 Kode program proses fine-tuning model RoBERTa 58
Gambar 4. 30 Kode program fungsi validasi model 59
Gambar 4.31 Kode program Output Logits 59
Gambar 4.32 Kode Program Prediksi Kelas Sentimen Berdasarkan Nilai *Logit* Tertinggi 60
Gambar 4. 33 Kode program proses training dan validation 60
Gambar 4. 34 Kode Penyimpanan Riwayat Hasil *Training* dan *Validation* 61
Gambar 4.35 Kode Penyimpanan Model Terbaik Berdasarkan Akurasi Validasi 61
Gambar 4.36 Kode Program Evaluasi Model Menggunakan Data Pengujian 62
Gambar 4.37 Kode Program Pengujian Seluruh Aspek 63
Gambar 4.38 Kode Program Visualisasi Hasil Pengujian 63
Gambar 4.39 Kode Program Visualisasi *Confusion Matrix* 63
Gambar 4.40 Kode Program Rekapitulasi Hasil Pengujian 64
Gambar 4.41 Kode Program Penyusunan Tabel Perbandingan 64
Gambar 4.42 Kode Program Visualisasi Perbandingan Model 65

Gambar 5.1 Jumlah Data pada Setiap Aspek 67
Gambar 5.2 Distribusi Sentimen pada Setiap Aspek 68
Gambar 5.3 Kurva loss, accuracy, dan Macro F1-score aspek SDM 78
Gambar 5.4 Kurva loss, accuracy, dan Macro F1-score aspek Fasilitas 79

UNIVERSITAS
GADJAH MADA
ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN RUMAH SAKIT MENGGUNAKAN ROBERTA
Alifia Intan Andrean Nunyai, Prof. Dra. Sri Hartati, M.Sc., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

Gambar 5. 5 Kurva loss, accuracy, dan Macro F1-score aspek Pelayanan..... 80
Gambar 5. 6 *Confusion Matrix* Aspek SDM..... 85
Gambar 5. 7 Confusion Matrix Aspek Fasilitas..... 87
Gambar 5. 8 Confusion Matrix Aspek Pelayanan..... 90
Gambar 5. 9 Accuracy Model pada Setiap Aspek 91
Gambar 5. 10 *Macro F1-score* Model pada Setiap Aspek..... 93
Gambar 5. 11 *Weighted F1-score* Model pada Setiap Aspek 95
Gambar 5. 12 Diagram Batang Perbandingan *Macro F1-score* 97

UNIVERSITAS
GADJAH MADA
ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN RUMAH SAKIT MENGGUNAKAN ROBERTA
Alifia Intan Andrean Nunyai, Prof. Dra. Sri Hartati, M.Sc., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR TABEL

Tabel 3.1 Contoh Struktur Dataset 16
Tabel 3.2 Jumlah Data pada Setiap Aspek Berdasarkan Kategori Sentimen 20
Tabel 3.3 Rincian Dataset Split 21
Tabel 3.4 Skema Label Encoding 23
Tabel 3.5 Hasil Tokenization 26
Tabel 3.6 Struktur Data pada DataLoader 28

Tabel 4.1 Lingkungan Implementasi 41
Tabel 4. 2 Lima Data Pertama Dataset 42

Tabel 5.1 yang menampilkan lima data pertama 70
Tabel 5.2 Hasil Label Encoding 71
Tabel 5.3 Hasil Pembagian Dataset 73
Tabel 5 4 Distribusi Kelas Hasil Dataset Splitting Aspek SDM 74
Tabel 5. 5 Distribusi Kelas Hasil Dataset Splitting Aspek Fasilitas 74
Tabel 5. 6 Distribusi Kelas Hasil Dataset Splitting Aspek Pelayanan 74
Tabel 5. 7 Distribusi Data Latih dan Nilai Class Weight 76
Tabel 5. 8 Perbandingan Hasil Fine-Tuning dengan dan tanpa Class Weight 77
Tabel 5. 9 Ringkasan hasil *fine-tuning* 81
Tabel 5.10 Ringkasan Hasil Evaluasi Model 82
Tabel 5. 11 *Classification Report* Aspek SDM 83
Tabel 5. 12 Hasil Evaluasi Model RoBERTa pada Aspek Fasilitas 86
Tabel 5. 13 Classification Report Aspek Pelayanan 89
Tabel 5.14 Perbandingan Macro F1-score dengan Penelitian Sebelumnya 98