

DAFTAR ISI

Hlm.

HALAMAN PERSEMBAHAN.....	ii
ABSTRAK	iii
ABSTRACT.....	iv
DAFTAR ISI.....	vii
DAFTAR GAMBAR	xi
DAFTAR TABEL	i
DAFTAR LAMPIRAN.....	ii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah Penelitian	6
1.3 Tujuan Penelitian	6
1.4 Batasan Masalah Penelitian.....	6
1.5 Kerangka Pemikiran Penelitian.....	7
1.6 Sistematika Penulisan.....	9
BAB II TUJUAN PUSTAKA.....	10
2.1 <i>State of the Art</i>	10
2.2 Dasar Teori.....	14
2.2.1. Pariwisata	14
2.2.2. Tempat Wisata di Indonesia.....	15
2.2.3. <i>Natural Language Processing (NLP)</i>	15
2.2.4 BERT.....	17
2.2.5 <i>Bidirectional Encoder Representations from Transformers Score</i>	18
3. <i>Cross-Industry Standard Process for Data Mining (Crisp-DM)</i>	21
BAB III METODOLOGI PENELITIAN.....	24

3.1 <i>Business Understanding</i> / Pemahaman Bisnis	25
3.1.1 <i>Determine Business Objective</i>	25
3.1.2 <i>Assess Situation</i>	26
3.1.3 <i>Determine Data Mining Goals</i>	26
3.1.4 <i>Project Plan</i>	26
3.2 <i>Data Understanding</i> / Pemahaman Data	26
3.2.1 <i>Data Collection</i>	26
3.2.2 <i>Data Description</i>	27
3.2.3 <i>Initial Data Exploration</i>	28
3.3 <i>Data Preparation</i>	28
3.3.1 <i>Load Dataset Awal</i>	28
3.3.2 <i>Hapus Nilai Kosong</i>	28
3.3.3 <i>Normalisasi dan Pembersihan Teks Ulasan</i>	28
3.3.4 <i>Hapus Teks Kosong Hasil Cleaning</i>	29
3.3.5 <i>Konversi Rating Menjadi Label Sentimen</i>	29
3.3.6 <i>Seleksi Kolom dan Simpan Dataset Bersih</i>	29
3.4 <i>Modeling</i>	29
3.4.1 <i>Select Modelling</i>	29
3.4.2 <i>Build Model</i>	30
3.4.3 <i>Generate Test Design</i>	30
3.4.4 <i>EDA</i>	30
3.5 <i>Evaluation</i>	31
3.5.1 <i>Evaluate Result</i>	31
3.5.2 <i>Determine Best Model</i>	31
3.5.3 <i>Determine Next Step</i>	32
BAB IV	33
HASIL DAN PEMBAHASAN.....	33

4.1 Hasil <i>Business Understanding</i>	33
4.1.1 <i>Determine Business Objective</i>	34
4.1.2 <i>Assess Situation</i>	34
4.2 Hasil <i>Data Understanding</i>	36
4.2.1 <i>Data Collection</i>	36
4.2.2 <i>Data Description</i>	36
4.2.3 EDA.....	37
a. Tripadvisor	37
b. Google Maps	41
4.3 Hasil <i>Data Preparation</i>	45
4.3.1 <i>Load Dataset Awal</i>	45
4.3.2 Hapus Nilai Kosong	46
4.3.3 Normalisasi dan Pembersihan Teks Ulasan	47
4.3.4 Hapus Teks Kosong Hasil <i>Cleaning</i>	49
4.3.5 Konversi <i>Rating</i> Menjadi Label Sentimen.....	51
4.3.6 Seleksi Kolom dan Simpan Dataset Bersih.....	52
4.4 Hasil <i>Modelling</i>	53
4.4.1 Konfigurasi Skenario dan <i>Logging</i> Awal.....	53
4.4.2 Pemuatan dan Sampling Dataset	53
4.4.3 Tokenisasi Menggunakan BERT <i>Tokenizer</i>	53
4.4.4 Pembagian <i>Dataset: Train, Validation, Test</i>	54
4.4.5 Konfigurasi <i>Training Arguments</i>	54
4.4.6 Inisialisasi dan Pelatihan Model.....	55
4.4.7 Penyimpanan Model dan <i>Tokenizer</i>	55
4.4.8 Evaluasi Model pada Data Pengujian	55
4.4.9 Pembentukan Dataframe Hasil Prediksi.....	56
4.4.10 Perhitungan Akurasi dan Laporan Klasifikasi	56

4.4.11 Visualisasi Confusion Matrix.....	56
4.5 Hasil Evaluation.....	57
4.5.1 Skenario 70:15:15	57
a. Skenario 1: LR 2e-5, Batch Size 32, Epoch 3	57
b. Skenario 2: LR 3e-5, Batch Size 32, Epoch 5	58
c. Skenario 3: LR 5e-5, Batch Size 32, Epoch 3	60
d. Skenario 4: LR 3e-5, Batch Size 32, Epoch 10	61
4.5.2 Skenario 70:20:10	63
a. Skenario 1: LR 2e-5, Batch Size 16, Epoch 3	63
b. Skenario 2: LR 3e-5, Batch Size 16, Epoch 5	64
c. Skenario 3: LR 3e-5, Batch Size 16, Epoch 3	66
4.5.3 Skenario 80:10:10	67
a. Skenario 1: LR 2e-5, Batch Size 16, Epoch 3	67
b. Skenario 2: LR 1e-5, Batch Size 16, Epoch 3	68
c. Skenario 3: LR 2e-5, Batch Size 16, Epoch 3	70
4.7 Pembahasan.....	72
4.7.1 Model Analisis Sentimen IndoBERT.....	72
4.7.2 Akurasi Model.....	72
BAB V SIMPULAN DAN SARAN	74
5.1 Simpulan	74
5.2 Saran.....	74
DAFTAR PUSTAKA	76
LAMPIRAN	81

DAFTAR GAMBAR

	Hlm.
Gambar 1. 1 Kerangka Pemikiran Peneliti	8
Gambar 2.1 Metode CRISP-DM [26].	22
Gambar 3. 1 Tahapan Crisp-DM.....	24
Gambar 4. 1 Distribusi Peringkat Perjalanan	37
Gambar 4. 2 Frekuensi Kata.....	38
Gambar 4. 3 Distribusi Berdasarkan Tahun	39
Gambar 4. 4 Distribusi Ulasan Berdasarkan Provinsi.....	40
Gambar 4. 5 <i>Word Cloud</i>	40
Gambar 4. 6 Distribusi Peringkat Pengguna	41
Gambar 4. 7 Frekuensi Kata.....	42
Gambar 4. 8 15 Besar Distribusi Ulasan Berdasarkan Provinsi.....	43
Gambar 4. 9 Distribusi Ulasan Berdasarkan Tahun.....	44
Gambar 4. 10 <i>Word Cloud</i>	45
Gambar 4. 11 Konfigurasi Skenario dan <i>Logging</i> Awal.....	53
Gambar 4. 12 Pemuatan dan Sampling Dataset	53
Gambar 4. 13 Tokenisasi Menggunakan BERT.....	54
Gambar 4. 14 Pembagian Dataset	54
Gambar 4. 15 Konfigurasi <i>Training Arguments</i>	54
Gambar 4. 16 Inisialisasi dan Pelatihan Model.....	55
Gambar 4.17 Penyimpanan Model.....	55
Gambar 4.18 Evaluasi Model.....	55
Gambar 4.19 Pembentukan Data <i>frame</i>	56
Gambar 4.20 Perhitungan Akurasi dan Klasifikasi.....	56
Gambar 4.21 <i>Visualisasi Confusion Matrix</i>	56
Gambar 4.22 Skenario 1 (70:15:15).....	57
Gambar 4.23 <i>Confusion Matrix</i> 1.....	58
Gambar 4.24 Skenario 2 (70:15:15).....	58
Gambar 4.25 <i>Confusion Matrix</i> 2.....	59
Gambar 4.26 Skenario 3 (70:15:15).....	60
Gambar 4.27 <i>Confusion Matrix</i> 3.....	61

Gambar 4.28 Sekenario 1 (70:15:15).....	61
Gambar 4.29 <i>Confusion Matrix</i> 4.....	62
Gambar 4.30 Sekenario 1 (70:20:10).....	63
Gambar 4.31 <i>Confusion Matrix</i> 5.....	64
Gambar 4.32 Sekenario 2 (70:20:10).....	64
Gambar 4.33 <i>Confusion Matrix</i> 6.....	65
Gambar 4.34 Sekenario 3 (70:20:10).....	66
Gambar 4.35 <i>Confusion Matrix</i> 7.....	67
Gambar 4.36 Sekenario 1 (80:10:10).....	67
Gambar 4.37 <i>Confusion Matrix</i> 8.....	68
Gambar 4.38 Sekenario 2 (80:10:10).....	69
Gambar 4.39 <i>Confusion Matrik</i> 9	69
Gambar 4.40 Sekenario 3 (80:10:10).....	70
Gambar 4.41 <i>Confusion Matrix</i> 10	71



DAFTAR TABEL

	Hlm.
Tabel 1.1 Jumlah Perjalanan Wisatawan Mancanegara Tahun 2018-2023	1
Tabel 1.2 Jumlah Perjalanan Wisatawan Nusantara Tahun 2019-2023.....	2
Tabel 2.1 <i>State Of The Art</i>	10
Tabel 3.1 Contoh Tempat Wisata di Indonesia.....	27
Tabel 3.2 <i>Test Design</i>	30
Tabel 4. 1 <i>Load Dataset</i> Awal	45
Tabel 4. 2 Normalisasi dan Pembersihan Teks	48
Tabel 4. 3 Hapus Teks Hasil <i>Cleaning</i>	50
Tabel 4. 4 Konversi Ranting Menjadi Label Sentimen.....	51
Tabel 4. 5 Seleksi Kolom dan Simpan Dataset Bersih.....	52
Tabel 4.6 Skenario dan Nilai Akurasinya	71



DAFTAR LAMPIRAN

	Hlm.
Lampiran 1 Data Wisata Di Indonesia	81
Lampiran 2 Code Model	100
Lampiran 3 Code Contoh untuk Skenario.....	102

