

## DAFTAR ISI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI .....	ii
HALAMAN PENGESAHAN .....	iii
KATA PENGANTAR .....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI .....	v
ABSTRAK .....	vi
ABSTRACT .....	vii
DAFTAR ISI .....	viii
DAFTAR TABEL .....	xi
DAFTAR GAMBAR .....	xiii
BAB I PENDAHULUAN .....	1
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah .....	5
1.3. Tujuan dan Manfaat Penelitian.....	5
1.4. Ruang Lingkup .....	5
1.5. Sistematika Penulisan.....	6
BAB II TINJAUAN PUSTAKA .....	8
2.1. Studi Pustaka .....	8
2.2. Analisis Sentimen.....	13
2.3. Aspect-Based Sentiment Analysis.....	14
2.4. Ulasan Shopee .....	15
2.5. Identifikasi Aspek dan Pembentukan Final Dataset.....	16
2.5.1. Reduksi Dimensi .....	17
2.5.2. Hierarchical Clustering.....	20
2.5.3. Anotasi Manual dan Kontrol Kualitas.....	24
2.6. Representasi Teks.....	27
2.7. Term Frequency-Inverse Document Frequency.....	28
2.8. Sentence Bidirectional Encoder Representations from Transformers.....	30
2.9. Logistic Regression (LR) .....	36
2.10. Grid Search dan K-Fold Cross-Validation untuk Optimasi Hyperparameter.....	38
2.11. Metriks Evaluasi.....	39
2.12. McNemar Test.....	40

BAB III	METODOLOGI PENELITIAN.....	43
3.1.	Pengumpulan Data Eksploratori.....	44
3.2.	Preprocessing Data Eksploratori .....	45
3.2.1.	Pembersihan dan Standarisasi Teks .....	45
3.2.2.	Normalisasi Slang dengan Aho-Corasick Automaton.....	46
3.2.3.	Stopword Removal .....	47
3.2.4.	Stemming dan Seleksi Akhir Token.....	48
3.3.	Identifikasi Aspek Melalui Hierarchical Clustering dan Manual Review.....	49
3.3.1.	Representasi Fitur dengan TF-IDF .....	49
3.3.2.	Reduksi Dimensi dengan Truncated SVD dan PCA .....	50
3.3.3.	Hierarchical Clustering dengan Ward Linkage .....	56
3.3.4.	Penentuan Jumlah Cluster Optimal .....	58
3.3.5.	Ekstraksi Keyword per Cluster dan Identifikasi Aspek .....	60
3.4.	Pengumpulan Data Terfokus .....	61
3.5.	Anotasi Manual dan Kontrol Kualitas.....	62
3.5.1.	Protokol Anotasi Manual.....	62
3.5.2.	Kontrol Kualitas dengan Fleiss Kappa dan Cohen's Kappa .....	63
3.5.3.	Adjudikasi dan Penetapan Label Akhir.....	66
3.6.	Preprocessing Final Dataset .....	66
3.7.	Pengembangan Model .....	67
3.7.1.	Pembagian Data Train dan Test .....	67
3.7.2.	Term Frequency-Inverse Document Frequency .....	68
3.7.2.	Sentence Bidirectional Encoder Representations from Transformers .....	70
3.7.3.	Logistic Regression (LR) .....	77
3.7.4.	Grid Search dan K-Fold Cross-Validation untuk Optimasi Hyperparameter .	78
3.8.	Evaluasi Model.....	80
3.8.1.	Evaluasi Per-Aspek .....	82
3.8.2.	Evaluasi Menyeluruh.....	82
3.9.	McNemar Test untuk Analisis Komparatif .....	83
BAB IV	HASIL DAN PEMBAHASAN.....	84
4.1.	Hasil Fase Identifikasi Aspek.....	84
4.2.	Hasil Pembentukan Final Dataset.....	87
4.3.	Perbandingan Performa Model dengan Hyperparameter Terbaik Per-Aspek.....	89

4.3.1. Hasil Cross-Validation dengan Hyperparameter Terbaik Per-Aspek .....	91
4.3.2. Hasil Held-Out Test Set dengan Hyperparameter Terbaik Per-Aspek.....	92
4.3.3. Hasil McNemar Test dengan Hyperparameter Terbaik Per-Aspek.....	93
4.4. Perbandingan Performa Model dengan Hyperparameter Universal.....	95
4.4.1. Hasil Cross-Validation dengan Hyperparameter Universal .....	95
4.4.2. Hasil Held-Out Test Set dengan Hyperparameter Universal .....	96
4.4.3. Hasil McNemar Test dengan Hyperparameter Universal .....	98
4.5. Evaluasi Model Combined .....	99
4.5.1. Hasil Cross-Validation Model Combined .....	99
4.5.2. Hasil Held-Out Test Set Model Combined .....	100
4.5.3. Hasil McNemar Test Model Combined .....	100
4.6. Sintesis dan Pembahasan Komprehensif.....	101
4.6.1. Karakteristik Aspek dan Kesesuaian Representasi .....	101
4.6.2. Trade-off antara Spesialisasi dan Generalisasi.....	102
4.6.3. Konsistensi antara Cross-Validation dan Test Set .....	102
4.6.4. Interpretasi Hasil McNemar Test .....	103
4.6.5. Implikasi Praktis untuk Sistem ABSA pada E-commerce Indonesia.....	104
4.7. Ringkasan Temuan Utama .....	105
BAB V KESIMPULAN DAN SARAN.....	106
5.1. Kesimpulan.....	106
5.2. Saran .....	106
DAFTAR PUSTAKA.....	107
LAMPIRAN 1. TABEL KATA KUNCI TERATAS DI SETIAP KLASTER.....	115
LAMPIRAN 2. TEMPAT PENGUMPULAN HASIL LABELING .....	116
LAMPIRAN 3. TABEL SAMPEL PERBEDAAN HASIL ANOTASI .....	117
LAMPIRAN 4. TABEL SAMPEL PERBANDINGAN PREDIKSI MODEL.....	118
LAMPIRAN 5. PSEUDO CODE PROJECT .....	119