

## DAFTAR ISI

LEMBAR PERNYATAAN ORISINALITAS .....	ii
LEMBAR PENGESAHAN .....	iii
KATA PENGANTAR .....	iv
ABSTRAK .....	v
ABSTRACT .....	vi
Daftar Isi.....	vii
Daftar Gambar.....	x
Daftar Tabel .....	xii
Daftar Simbol.....	xiv
Daftar Istilah.....	xv
Bab I    Pendahuluan .....	1
I.1    Latar Belakang.....	1
I.2    Perumusan Masalah.....	5
I.3    Tujuan Penelitian.....	6
I.4    Batasan Penelitian .....	6
I.5    Manfaat Penelitian.....	6
Bab II    Tinjauan Pustaka .....	7
II.1    Penelitian Terdahulu.....	7
II.2    Online Travel Agency (OTA) .....	12
II.2.1    Traveloka.....	12
II.2.2    Tiket.com .....	12
II.3    Twitter .....	13
II.4    Text mining .....	13

II.4.1	<i>Natural Language Processing</i> .....	13
II.4.2	<i>Data Preprocessing</i> .....	14
II.4.3	<i>N-gram</i> .....	15
II.4.4	<i>Naive Bayes Classifier</i> .....	15
II.4.5	<i>Term Frequency-Inverse Document Frequency (TF-IDF)</i> .....	17
II.4.6	<i>Data splitting</i> .....	18
II.4.7	<i>Random oversampling</i> .....	19
II.4.8	<i>K-Fold cross validation</i> .....	19
II.4.9	<i>Confusion matrix</i> .....	20
Bab III	<b>METODOLOGI PENELITIAN</b> .....	21
III.1	<b>Kerangka Berpikir</b> .....	21
III.2	<b>Sistematika Penyelesaian Masalah</b> .....	22
III.3	<b>Pengumpulan Data</b> .....	24
III.4	<b>Pengolahan Data atau Pengembangan Produk / Artifak</b> .....	25
III.5	<b>Metode Evaluasi</b> .....	26
III.6	<b>Alasan Pemilihan Metode</b> .....	27
Bab IV	<b>Analisa pengolahan data</b> .....	28
IV.1	<b>Data selection</b> .....	28
IV.1.1	<b>Pengumpulan Data</b> .....	28
IV.2	<b>Data Preprocessing</b> .....	29
IV.2.1	<b>Labeling</b> .....	29
IV.2.2	<b>Cleaning</b> .....	31
IV.2.3	<b>Case Folding</b> .....	32
IV.2.4	<b>Stemming</b> .....	32
IV.2.5	<b>Tokenizing</b> .....	32
IV.2.6	<b>Stopwords</b> .....	33

IV.2.7	<i>Spell Correction</i> .....	33
IV.3	<i>Transformation</i> .....	34
IV.3.1	TF-IDF .....	34
IV.4	Data mining .....	36
IV.4.1	Data Splitting .....	36
IV.4.2	<i>Naïve bayes</i> .....	36
IV.4.3	<i>Random Oversampling</i> .....	40
IV.5	Evaluasi .....	40
IV.5.1	K-Fold Cross Validation .....	40
IV.5.2	<i>Confusion Matrix</i> .....	41
Bab V	implementasi dan evaluasi.....	42
V.1	Analisis Sentimen.....	42
V.2	Implementasi Algoritma.....	44
V.3	<i>Random oversampling</i> .....	45
V.4	Hasil Akurasi .....	46
V.5	Hasil Validasi .....	50
V.6	Visualisasi .....	50
V.7	N-gram.....	55
Bab VI	kesimpulan dan saran.....	57
VI.1	Kesimpulan.....	57
VI.2	Saran.....	57
VI.2.1	Saran untuk perusahaan .....	57
VI.2.2	Saran untuk masyarakat .....	58
VI.2.3	Saran untuk penelitian selanjutnya .....	58
Daftar Pustaka	.....	59