DAFTAR ISI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI..... **Error! Bookmark not defined.**

HALAMAN PENGESAHAN ..... **Error! Bookmark not defined.**

KATA PENGANTAR iv

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI **Error! Bookmark not defined.**

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 3

1.3 Tujuan 4

1.4 Manfaat 4

1.5 Ruang Lingkup 4

1.6 Sistematika Penulisan 4

BAB II LANDASAN TEORI 6

2.1 Studi Literatur 6

2.2 *Natural Language Processing* (NLP) 8

2.3 Analisis Sentimen 9

2.4 *Part-of-Speech* (POS) Tagging 11

2.5 Star Rating 12

2.6 *Rule Based Method* 13

2.7 *CountVectorizer* 14

2.8 *Term Frequency-Inverse Document Frequency* (TF-IDF) 15

2.9 Naïve Bayes Classifier 20

2.10 Confusion Matrix 26

2.11 Stratified K-Fold Cross Validation 28

BAB III METODE PENELITIAN 34

3.1 Studi Literatur 35

3.2 *Scraping Data* 35
3.3 *Preprocessing* 36
3.3.1 *Data Preparation* 36
3.3.2 Menghapus Karakter Tidak Penting 36
3.3.3 *Case Folding* 37
3.3.4 Tokenisasi 37
3.3.5 *Filtering* (*Slang Word Replacement* dan *Stop Word Removal*) 38
3.3.6 *Stemming* 38
3.4 *Labelling* dan Ekstraksi Fitur 39
3.4.1 POS Tagging 39
3.4.2 *Rule Based Method* 40
3.4.3 Star Rating *Labelling* 41
3.4.4 *CountVectorizer* 42
3.4.5 TF-IDF 42
3.5 Naïve Bayes Classifier 43
3.6 Stratified K-Fold Cross Validation 43
3.7 Evaluasi Model 44
3.8 Analisis Hasil 44
BAB IV HASIL DAN PEMBAHASAN 46
4.1 *Scraping Data* 46
4.2 *Preprocessing* 47
4.2.1 *Data Preparation* 47
4.2.2 Menghapus Karakter Tidak Penting dan *Case Folding* 48
4.2.3 Tokenisasi 49
4.2.4 *Filtering* 50
4.2.5 *Stemming* 51
4.3 *Labelling* dan Ekstraksi Fitur 52
4.3.1 Implementasi Metode POS Tagging, *Rule Based Method*, dan *CountVectorizer* 52
4.3.2 Implementasi Metode Star Rating dan TF-IDF 57
4.4 Naïve Bayes Classifier 61
4.4.1 Implementasi Naïve Bayes Classifier pada Skenario 1 61
4.4.2 Implementasi Naïve Bayes Classifier pada Skenario 2 64

4.5 Stratified K-Fold Cross Validation 67
4.5.1 Skenario 1 67
4.5.2 Skenario 2 68
4.6 Evaluasi Model 69
4.6.1 Evaluasi Model Metode POS Tagging dan *Rule Based Method* 69
4.6.2 Evaluasi Model Metode Star Rating dan TF-IDF 72
4.7 Analisis Hasil 75
BAB V PENUTUP 77
5.1 Kesimpulan 77
5.2 Saran 77
DAFTAR PUSTAKA 79
LAMPIRAN 1. *Source Code Preprocessing* 83
LAMPIRAN 2. *Source Code* Metode POS Tagging, *Rule Based*, dan Naïve Bayes 87
LAMPIRAN 3. *Source Code* Metode Star Rating, TF-IDF, dan Naïve Bayes 90
LAMPIRAN 4. Dataset Setelah *Data Preparation* 91