## DAFTAR ISI

**LEMBAR PENGESAHAN SKRIPSI ..... iii**
**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI ..... iv**
**HALAMAN PERNYATAAN ORISINALITAS ..... v**
**LEMBAR HASIL FINAL DETEKSI PLAGIASI SKRIPSI ..... vi**
**ABSTRAK ..... vii**
***ABSTRACT* ..... viii**
**KATA PENGANTAR ..... ix**
**DAFTAR ISI ..... x**
**DAFTAR GAMBAR ..... xii**
**DAFTAR TABEL ..... xiii**
**BAB I PENDAHULUAN ..... 1**
1.1 Latar Belakang ..... 1
1.2 Rumusan Masalah ..... 6
1.3 Tujuan Penelitian ..... 6
1.4 Batasan Masalah ..... 7
1.5 Manfaat Penelitian ..... 8
**BAB II TINJAUAN PUSTAKA ..... 9**
2.1 Penelitian Terdahulu ..... 9
2.2 Analisis Sentimen ..... 12
2.3 SMOTE-Tomek ..... 14
2.4 Prapemrosesan Teks ..... 15
*2.4.1 Case Folding* ..... 15
2.4.2 *Cleaning* ..... 16
2.4.3 *Tokenization* ..... 17
*2.4.4 Stopword Removal* ..... 17
*2.4.5 Stemming* ..... 17
2.5 *Manual Annotation dan Annotation Guideline* ..... 18
2.6 Ekstraksi Fitur ..... 19
2.7 *Machine Learning* ..... 20
*2.7.1 Random Forest* ..... 21

*2.7.2 Support Vector Machine (SVM)* ..... 24
*2.7.3 K-Nearest Neighbor (KNN)* ..... 26
*2.7.4 Naïve Bayes* ..... 29
2.8 *Term Frequency – Inverse Document Frequency* (TF-IDF) ..... 31
2.9 Evaluasi Model ..... 34
**BAB III METODOLOGI PENELITIAN ..... 39**
3.1 Pengumpulan Data ..... 40
3.2 Pelabelan Data ..... 42
3.3 Preprocessing ..... 45
3.4 Pembagian Dataset ..... 46
3.5 Ekstraksi Fitur TF-IDF ..... 48
3.6 Implementasi Model ..... 49
3.7 Evaluasi Model dan Perbandingan Model ..... 50
3.8 Implementasi Website ..... 52
**BAB IV HASIL DAN PEMBAHASAN ..... 54**
4.1 Hasil *Fleiss' Kappa* ..... 54
4.2 Distribusi Dataset dan TF-IDF ..... 55
4.3 Perbandingan Sebelum dan Sesudah SMOTE-Tomek ..... 57
4.4 Hasil Implementasi *Naïve Bayes* ..... 62
4.5 Hasil Implementasi *SVM* ..... 68
4.6 Hasil Implementasi *Random Forest* ..... 74
4.7 Hasil Implementasi *KNN* ..... 79
4.8 Evaluasi dan Perbandingan Model ..... 84
4.9 Implementasi Website ..... 86
**BAB V KESIMPULAN DAN SARAN ..... 92**
5.1 Kesimpulan ..... 92
5.2 Saran ..... 93
**DAFTAR PUSTAKA ..... 95**
**LAMPIRAN ..... 99**
**BIODATA PENULIS ..... 105**