

DAFTAR ISI

| | |
|--|------|
| HALAMAN PENGESAHAN..... | ii |
| PERNYATAAN BEBAS PLAGIASI..... | iii |
| PRAKATA..... | iv |
| ARTI LAMBANG DAN SINGKATAN..... | vi |
| ABSTRACT..... | vii |
| INTISARI..... | viii |
| DAFTAR ISI..... | ix |
| DAFTAR GAMBAR..... | xi |
| DAFTAR TABEL..... | xiii |
| BAB I PENDAHULUAN..... | 1 |
| 1.1 Latar Belakang..... | 1 |
| 1.2 Perumusan masalah..... | 5 |
| 1.3 Keaslian penelitian..... | 5 |
| 1.4 Tujuan Penelitian..... | 8 |
| 1.5 Manfaat Penelitian..... | 8 |
| BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI..... | 9 |
| 2.1 Tinjauan Pustaka..... | 9 |
| 2.2 Landasan Teori..... | 12 |
| 2.2.1 <i>Short Message Service (SMS)</i> | 12 |
| 2.2.2 Penggunaan SMS dalam layanan pemerintahan..... | 13 |
| 2.2.3 <i>Text Mining</i> | 13 |
| 2.2.4 <i>Preprocessing</i> | 14 |
| 2.2.5 <i>Feature extraction</i> | 14 |
| 2.2.6 <i>Synthetic Minority Oversampling Technique (SMOTE)</i> | 15 |
| 2.2.7 <i>Support Vector Machine (SVM)</i> | 17 |
| 2.2.8 <i>Naïve Bayes (NB)</i> | 18 |
| 2.2.9 <i>K-Nearest Neighbor (KNN)</i> | 19 |
| 2.2.10 <i>Logistic Regression (LR)</i> | 20 |
| 2.2.11 <i>Decision Tree (DT)</i> | 21 |
| 2.2.12 <i>Random Forest (RF)</i> | 22 |
| BAB III METODOLOGI..... | 24 |
| 3.1 Alat dan Bahan Penelitian..... | 24 |
| 3.1.1 Alat Penelitian..... | 24 |
| 3.1.2 Bahan Penelitian..... | 24 |
| 3.2 Jalannya Penelitian..... | 25 |
| 3.2.1 Studi Literatur..... | 26 |

| | | |
|----------------------------------|--|----|
| 3.2.2 | Pengumpulan Dataset..... | 26 |
| 3.2.3 | Perancangan, Pengujian dan Evaluasi Sistem..... | 27 |
| 3.2.4 | Dokumentasi dan Penulisan Tesis..... | 28 |
| 3.3 | Perancangan Sistem..... | 28 |
| 3.3.1 | Input <i>Dataset</i> | 29 |
| 3.3.2 | <i>Preprocessing</i> | 29 |
| 3.3.3 | Pemisahan <i>Train Data</i> dan <i>Test Data</i> | 30 |
| 3.3.4 | <i>Feature Extraction</i> | 31 |
| 3.3.5 | SMOTE..... | 32 |
| 3.3.6 | Klasifikasi..... | 33 |
| 3.3.7 | Evaluasi Kinerja Metode Klasifikasi..... | 37 |
| 3.4 | Cara Analisis..... | 39 |
| BAB IV HASIL DAN PEMBAHASAN..... | | 40 |
| 4.1 | Hasil penginputan dataset..... | 40 |
| 4.2 | Hasil <i>preprocessing</i> | 40 |
| 4.3 | Hasil Pemisahan <i>Train Data</i> dan <i>Test Data</i> | 41 |
| 4.4 | Hasil SMOTE..... | 42 |
| 4.5 | Hasil Evaluasi Kinerja Metode Klasifikasi..... | 42 |
| 4.5.1 | <i>Support Vector Machine (SVM)</i> | 43 |
| 4.5.2 | <i>Naïve Bayes (NB)</i> | 46 |
| 4.5.3 | <i>K-Nearest Neighbor (KNN)</i> | 51 |
| 4.5.4 | <i>Logistic Regression (LR)</i> | 55 |
| 4.5.5 | <i>Decision Tree (DT)</i> | 58 |
| 4.5.6 | <i>Random Forest (RF)</i> | 61 |
| BAB V KESIMPULAN DAN SARAN..... | | 66 |
| 5.1 | Kesimpulan..... | 66 |
| 5.2 | Saran..... | 67 |
| DAFTAR PUSTAKA..... | | 69 |
| LAMPIRAN..... | | 1 |