

DAFTAR ISI

| | |
|--|------|
| HALAMAN JUDUL..... | i |
| HALAMAN JUDUL..... | ii |
| HALAMAN PENGESAHAN..... | iii |
| PERNYATAAN BEBAS PLAGIASI..... | iv |
| HALAMAN PERSEMBAHAN | v |
| PRAKATA..... | vi |
| DAFTAR ISI | viii |
| DAFTAR TABEL..... | xi |
| DAFTAR GAMBAR | xii |
| DAFTAR LAMPIRAN | xiii |
| INTISARI..... | xiv |
| <i>ABSTRACT</i> | xv |
| BAB I PENDAHULUAN..... | 1 |
| 1.1 Latar Belakang..... | 1 |
| 1.2 Rumusan Masalah | 3 |
| 1.3 Batasan Masalah | 3 |
| 1.4 Tujuan Penelitian | 4 |
| 1.5 Manfaat Penelitian..... | 4 |
| 1.6 Tinjauan Pustaka..... | 4 |
| 1.7 Metodologi Penelitian | 6 |
| 1.8 Sistematika Penulisan | 6 |
| BAB II LANDASAN TEORI | 8 |
| 2.1 Data Mining..... | 8 |
| 2.2 Machine Learning..... | 9 |
| 2.3 Analisis Klasifikasi..... | 10 |
| 2.4 Evaluasi Performa Klasifikasi | 11 |
| 2.5 Jarak <i>Euclidean</i> | 15 |
| 2.6 Matriks..... | 16 |
| 2.6.1 Jenis Matriks..... | 17 |
| 2.6.2 Penjumlahan Matriks..... | 18 |
| 2.6.3 Perkalian Matriks dengan Bilangan Real | 19 |

| | | |
|--|---|----|
| 2.6.4 | Perkalian Antar Matriks..... | 19 |
| 2.6.5 | Transpos Matriks | 20 |
| 2.6.6 | Invers Matriks..... | 21 |
| 2.6.7 | Trace Matriks..... | 22 |
| 2.7 | Decision Tree | 23 |
| 2.7.1 | Pembentukan Decision Tree | 24 |
| 2.7.2 | Attribute Selection Measures | 25 |
| 2.8 | Cross Validation..... | 26 |
| 2.9 | Bootstrap Sampling | 28 |
| 2.10 | Feature Scaling..... | 29 |
| BAB III PENANGANAN IMBALANCED DATA MENGGUNAKAN SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE AND BOOSTING (SMOTEBOOST) | | 31 |
| 3.1 | Data Tidak Seimbang | 31 |
| 3.2 | Penanganan Data Tidak Seimbang | 32 |
| 3.3 | Synthetic Minority Oversampling Technique (SMOTE) | 33 |
| 3.3.1 | Algoritma SMOTE | 33 |
| 3.3.2 | Algoritma <i>K Nearest Neighbor</i> (KNN) dalam SMOTE | 34 |
| 3.3.3 | Pembentukan Observasi Sintetis | 34 |
| 3.3.4 | SMOTE-NC (<i>Nominal and Continuous</i>)..... | 35 |
| 3.4 | Ensemble Learning | 36 |
| 3.4.1 | Pembentukan Ensemble Learning | 37 |
| 3.5 | AdaBoost (<i>Adaptive Boosting</i>) | 42 |
| 3.6 | Gradient Boosting | 46 |
| 3.7 | LightGBM (<i>Light Gradient Boosting Machine</i>) | 54 |
| 3.8 | Synthetic Minority Oversampling Technique and Boosting (SMOTEBoost) | 57 |
| 3.8.1 | Algoritma SMOTEBoost..... | 59 |
| 3.8.2 | Kelebihan dan Kekurangan SMOTEBoost | 60 |
| 3.9 | Flowchart Metode SMOTEBoost | 63 |
| BAB IV STUDI KASUS | | 64 |
| 4.1 | Deskripsi Data | 64 |
| 4.2 | Data Eksplorasi..... | 65 |
| 4.3 | Data Pre-processing | 69 |
| 4.4 | Analisis Klasifikasi menggunakan Decision Tree | 70 |

| | | |
|----------------------|--|----|
| 4.5 | Analisis Klasifikasi dengan Penanganan <i>Imbalanced Data</i> menggunakan <i>Ensemble Learning</i> | 74 |
| 4.5.1 | Analisis Klasifikasi menggunakan AdaBoost | 74 |
| 4.5.2 | Analisis Klasifikasi menggunakan <i>Gradient Boosting</i> | 76 |
| 4.5.3 | Analisis Klasifikasi menggunakan <i>LightGBM</i> | 77 |
| 4.6 | Analisis Klasifikasi dengan Penanganan <i>Imbalanced Data</i> menggunakan SMOTEBoost..... | 79 |
| 4.6.1 | Klasifikasi SMOTEBoost (menggunakan AdaBoost)..... | 83 |
| 4.6.2 | Klasifikasi SMOTEBoost (menggunakan <i>Gradient Boosting</i>)..... | 85 |
| 4.6.3 | Klasifikasi SMOTEBoost (menggunakan <i>LightGBM</i>)..... | 86 |
| 4.7 | Perbandingan Performa Klasifikasi | 88 |
| BAB V PENUTUP..... | | 92 |
| 5.1 | Kesimpulan..... | 92 |
| 5.2 | Saran | 93 |
| DAFTAR PUSTAKA | | 94 |
| LAMPIRAN..... | | 98 |