

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
PERNYATAAN BEBAS PLAGIASI	iii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN.....	xii
INTISARI.....	xiii
ABSTRACT.....	xiv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Batasan Masalah.....	3
1.3. Tujuan dan Manfaat Penelitian.....	4
1.4. Tinjauan Pustaka	4
1.5. Metode Penulisan	8
1.6. Sistematika Penulisan.....	8
BAB II LANDASAN TEORI.....	10
2.1. Skala Pengukuran Data	10
2.2. Probabilitas.....	12
2.2.1. Sifat-sifat Probabilitas.....	12
2.2.1. Probabilitas Bersyarat	15
2.3. Variabel Random.....	16
2.3.1. Variabel Random Diskrit	16
2.3.2. Variabel Random Kontinu	17
2.3.3. Ekspektasi	18
2.4. Matriks.....	18
2.4.1. Pengertian Matriks	19
2.4.2. Jenis-jenis Matriks	19
2.4.3. Operasi Matriks.....	22

2.4.4.	Transpos Matriks.....	24
2.4.5.	Determinan Matriks	25
2.4.6.	Invers Matriks	26
2.5.	Vektor.....	26
2.6.	<i>Data Mining</i>	27
2.6.1.	Fungsi Data Mining	28
2.7.	<i>Machine Learning</i>	30
2.8.	Klasifikasi.....	31
2.8.1.	Langkah Pembentukan Model Klasifikasi	31
2.8.2.	Komponen Dasar Proses Klasifikasi.....	33
2.8.3.	Jenis Teknik Klasifikasi	34
2.9.	Pohon Keputusan (<i>Decision Tree</i>)	34
2.9.1.	Jenis Algoritma <i>Decision Tree</i>	36
2.10.	Evaluasi Performa Model Klasifikasi	36
BAB III PENERAPAN METODE ENSEMBLE BOOTSTRAP AGGREGATING (BAGGING) PADA ALGORITMA POHON KEPUTUSAN C5.0		40
3.1.	Pohon Keputusan C5.0.....	40
3.1.1.	Keunggulan Algoritma Pohon Keputusan C5.0.....	41
3.1.2.	Algoritma Pohon Keputusan C5.0	41
3.1.3.	Entropi.....	42
3.1.4.	<i>Information Gain</i>	43
3.1.5.	<i>Gain Ratio</i>	47
3.1.6.	<i>Splitting Point</i> Atribut Numerik.....	48
3.2.	Metode <i>Ensemble</i>	51
3.2.1.	Jenis – jenis Metode <i>Ensemble</i>	52
3.3.	<i>Bootstrap Aggregating (Bagging)</i>	53
3.3.1.	Tahapan <i>Bagging</i>	55
3.3.2.	Algoritma <i>Bagging</i>	59
3.4.	Algoritma <i>Bagging</i> untuk Klasifikasi pada Pohon Keputusan C5.0	61
3.5.	Diagram Alur Penelitian.....	62
BAB IV STUDI KASUS		64
4.1.	Deskripsi Data	64
4.2.	Pra-pemrosesan Data	66
4.3.	Partisi Data	67

4.4.	Pembentukan Model Pohon Keputusan C5.0.....	68
4.5.	Klasifikasi Risiko Penyakit Jantung dengan Pohon Keputusan C5.0	73
4.6.	<i>Bagging</i> pada <i>Decision Tree</i> C5.0.....	74
4.6.1.	<i>Resampling Bootstrap</i> pada <i>Data Training</i>	75
4.6.2.	<i>Aggregating</i>	77
4.7.	Perbandingan Performa Klasifikasi.....	78
BAB V PENUTUP.....		80
5.1.	Kesimpulan.....	80
5.2.	Saran.....	81
DAFTAR PUSTAKA		82
LAMPIRAN.....		86
APENDIKS.....		94

DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i>	37
Tabel 3.1 <i>Training dataset</i> dari <i>database</i> pelanggan <i>Allelectronics</i>	45
Tabel 3.2 Sampel data atribut numerik	49
Tabel 3.3 Sampel data atribut numerik setelah diurutkan.....	49
Tabel 3.4 Contoh pengambilan keputusan <i>averaging</i>	58
Tabel 3.5 Contoh pengambilan keputusan <i>majority voting</i>	59
Tabel 4.1 Deskripsi data <i>Heart Disease</i>	64
Tabel 4.2 Ilustrasi <i>resampled data</i>	68
Tabel 4.3 <i>Crosstab</i> atribut Jenis Kelamin	69
Tabel 4.4 <i>Crosstab</i> atribut Gula Darah >120 mg/dl (<i>fb</i> s)	70
Tabel 4.5 <i>Crosstab</i> atribut Latihan Induksi Angina (<i>exng</i>).....	71
Tabel 4.6 Hasil klasifikasi metode pohon keputusan C5.0.....	73
Tabel 4.7 Perbandingan akurasi data <i>training</i> dan data <i>testing</i>	75
Tabel 4.8 Data <i>bootstrap</i> pertama pada data <i>Heart Disease</i>	76
Tabel 4.9 <i>Majority vote bagging</i> pada data <i>Heart Disease</i>	77
Tabel 4.10 Hasil <i>bagging</i> pada pohon keputusan C5.0.....	78
Tabel 4.11 Perbandingan performa pohon keputusan C5.0 dan C5.0 + <i>Bagging</i> . 78	