

## DAFTAR ISI

LEMBAR PENGESAHAN .....	ii
LEMBAR PERNYATAAN .....	iii
ABSTRAK .....	iv
<i>ABSTRACT</i> .....	v
KATA PENGANTAR .....	vi
DAFTAR ISI .....	viii
DAFTAR GAMBAR .....	xi
DAFTAR TABEL .....	xiii
DAFTAR LAMPIRAN .....	xiv
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Identifikasi Dan Rumusan Masalah .....	3
1.2.1 Identifikasi Masalah .....	3
1.2.2 Rumusan Masalah .....	3
1.3 Tujuan Penelitian .....	4
1.4 Pembatasan Masalah .....	4
1.5 Kegunaan Penelitian .....	4
1.5.1 Kegunaan Pengembangan Ilmu .....	4
1.5.2 Kegunaan Operasional .....	5
1.6 Sistematika Penulisan .....	5
BAB II KAJIAN PUSTAKA, KERANGKA PEMIKIRAN DAN HIPOTESIS ....	6
2.1 Kajian Pustaka .....	6
2.1.1 <i>Data Mining</i> .....	6
2.1.2 Metode Klasifikasi .....	7
2.1.3 <i>Decision Tree</i> .....	8
2.1.4 Bahasa Pemrograman <i>Python</i> .....	9

2.1.4.1	<i>Scikit-Learn</i> .....	10
2.1.4.2	<i>Pandas</i> .....	11
2.1.5	Algoritma CART .....	12
2.1.6	Penyakit Jantung .....	13
2.1.7	COVID-19 .....	14
2.1.8	Penelitian Terdahulu .....	15
2.2	Kerangka Pemikiran .....	17
2.3	Hipotesis .....	18
BAB III METODE PENELITIAN.....		19
3.1	Metode Penelitian.....	19
3.1.1	Pengumpulan Data Penelitian.....	20
3.1.1.1	Metode Pengumpulan Data.....	20
3.1.1.2	Jenis Sumber Data .....	22
3.1.1.3	Sumber Data .....	23
3.1.2	Pembersihan Data ( <i>Data Cleansing</i> ) .....	23
3.1.3	Pembagian Data .....	24
3.1.4	Pembuatan Model Klasifikasi <i>Decision Tree</i> .....	25
3.1.5	Evaluasi Model .....	37
3.1.5.1	<i>Confusion Matrix</i> .....	37
3.1.5.2	Metrik Evaluasi.....	39
BAB IV ANALISIS DAN PEMBAHASAN.....		50
4.1.	Pengumpulan Data .....	50
4.2	Variabel Bebas dan Variabel Target .....	54
4.3.	Pembersihan Data ( <i>Data Cleansing</i> ).....	55
4.3.1.	Menangani Atribut Data Tidak Lengkap.....	55
4.3.2.	Menangani Data Duplikat.....	57
4.3.3.	Menangani Nilai Data Tidak Valid.....	59
4.3.4.	Normalisasi Data.....	64
4.4.	Pembagian Data.....	73

4.4.1.	Menentukan Rasio Pembagian.....	73
4.4.2.	Pengacakan Data.....	73
4.4.3.	Pembagian Data .....	76
4.4.4.	Pengecekan Keseimbangan Kelas .....	77
4.5.	Pembuatan Model Klasifikasi <i>Decision Tree</i> .....	80
4.5.1.	Spesifikasi Perangkat.....	80
4.5.1.1.	Perangkat keras .....	80
4.5.1.2.	Perangkat lunak:.....	80
4.5.2.	Pembuatan Model .....	80
4.5.2.1.	Memuat data ke dalam program.....	81
4.5.2.2.	Pemisahan data variabel bebas dan variabel target .....	83
4.5.2.3.	Pembuatan model klasifikasi <i>Decision Tree</i> .....	86
4.5.2.4.	Atribut Paling Berpengaruh .....	89
4.6.	Evaluasi dan Analisis Model.....	91
4.6.1.	<i>Confusion Matrix</i> .....	91
4.6.2.	Metrik Evaluasi.....	92
4.6.2.1.	Akurasi ( <i>Accuracy</i> ) .....	92
4.6.2.2.	Presisi ( <i>Precision</i> ) .....	93
4.6.2.3.	<i>Recall (Sensitivity)</i> .....	94
4.6.2.4.	Spesifisitas ( <i>Specificity</i> ) .....	95
4.6.2.5.	F1-Score .....	95
4.6.2.6.	<i>Matthews Correlation Coefficient (MCC)</i> .....	96
4.6.2.7.	<i>Cohen's Kappa</i> .....	97
4.6.2.8.	<i>Balanced Accuracy</i> .....	99
4.6.2.9.	AUC ROC .....	100
BAB V KESIMPULAN DAN SARAN.....		103
5.1	Kesimpulan.....	103
5.2	Saran .....	104
DAFTAR PUSTAKA .....		105

## DAFTAR GAMBAR

Gambar 2. 1 Contoh model Decision Tree.....	9
Gambar 3.1 Metode Penelitian.....	19
Gambar 3. 2 Kode pengacakan data.....	23
Gambar 3. 3 Flowchart algoritma CART.....	26
Gambar 3.4 pustaka Python .....	28
Gambar 3. 5 perintah instal pandas .....	29
Gambar 3. 6 perintah instal scikit-learn .....	29
Gambar 3. 7 perintah instal matplotlib.....	30
Gambar 3. 8 kode memuat data training dan data testing .....	31
Gambar 3. 9 kode memisahkan variabel bebas dan target .....	31
Gambar 3. 10 kode mengatur entropy .....	33
Gambar 3. 11 kode membuat dan melatih model .....	33
Gambar 3. 12 kode membuat visualisasi model.....	35
Gambar 3. 13 kode menampilkan visualisasi model.....	36
Gambar 3. 14 skema Confusion Matrix .....	38
Gambar 4. 1 Data terkumpul.....	54
Gambar 4. 2 Informasi Data yang Kosong.....	56
Gambar 4. 3 Data setelah pembersihan data .....	57
Gambar 4. 4 Fitur Cek Duplikat.....	58
Gambar 4. 5 Mengatur Pengecekan Data Duplikat.....	58
Gambar 4. 6 Hasil pengecekan data duplikat.....	59
Gambar 4. 7 Hasil validasi data usia.....	60
Gambar 4. 8 Hasil validasi data jenis kelamin .....	60
Gambar 4. 9 Hasil validasi data cp .....	60
Gambar 4. 10 Hasil validasi data trestbps .....	61
Gambar 4. 11 Hasil validasi data chol .....	61
Gambar 4. 12 Hasil validasi data fbs .....	61
Gambar 4. 13 Hasil validasi data restecg .....	62
Gambar 4. 14 Hasil validasi data thalach.....	62

Gambar 4. 15 Hasil validasi data exang.....	62
Gambar 4. 16 Hasil validasi data oldpeak.....	63
Gambar 4. 17 Hasil validasi data slope.....	63
Gambar 4. 18 Hasil validasi data ca.....	63
Gambar 4. 19 Hasil validasi data thal.....	64
Gambar 4. 20 Data “jenis kelamin” setelah konversi.....	65
Gambar 4. 21 Data “cp” setelah konversi.....	66
Gambar 4. 22 Data “fbs” konversi.....	67
Gambar 4. 23 Data "restecg" hasil konversi.....	68
Gambar 4. 24 Data "exang" konversi.....	69
Gambar 4. 25 Data "slope" hasil konversi.....	70
Gambar 4. 26 Data "thal" hasil konversi.....	71
Gambar 4. 27 Data "target" konversi.....	72
Gambar 4. 28 Data sebelum pengacakan.....	74
Gambar 4. 29 Data setelah pengacakan.....	75
Gambar 4. 30 Pembagian data training.....	76
Gambar 4. 31 Pembagian data testing.....	77
Gambar 4. 32 hasil formula kelas 1 data training.....	78
Gambar 4. 33 hasil formula kelas 0 data training.....	78
Gambar 4. 34 hasil formula kelas 1 data testing.....	79
Gambar 4. 35 hasil formula kelas 0 data testing.....	79
Gambar 4. 36 hasil memuat data training.....	81
Gambar 4. 37 hasil memuat data testing.....	82
Gambar 4. 38 hasil pemisahan variabel bebas data training.....	84
Gambar 4. 39 hasil pemisahan variabel target data training.....	84
Gambar 4. 40 hasil pemisahan variabel bebas data testing.....	85
Gambar 4. 41 hasil pemisahan variabel target data testing.....	86
Gambar 4. 42 model Klasifikasi Decision Tree.....	88
Gambar 4. 43 hasil atribut paling berpengaruh.....	90
Gambar 4. 44 hasil Confusion Matrix.....	92
Gambar 4. 45 hasil AUC ROC.....	101

## DAFTAR TABEL

Tabel 2. 1 Penelitian terdahulu.....	17
--------------------------------------	----