

“PENGARUH SELEKSI FITUR TERHADAP PERFORMANSI MODEL KLASIFIKASI NAÏVE BAYES PADA DIAGNOSA ARRHYTHMIA”

Hanin Rana Delilah
(14.04.1.1.1.00118)

Pembimbing I : Mula'ab, S.Si., M.Kom.

Pembimbing II : Rosida Vivin Nahari, S.Kom., M.T.

ABSTRAK

Arrhythmia merupakan salah satu kelainan pada jantung yang ditandai dengan ritme jantung yang tidak normal, menjadi lebih cepat atau lebih lambat. Untuk mendiagnosa *arrhythmia* perlu dilakukan prosedur *Electrocardiograf*. Hasil prosedur *Electrocardiograf* ini berupa *Electrodiogram* atau yang biasa di kenal sebagai ECG. ECG inilah yang kemudian dijadikan rujukan untuk mendiagnosa *arrhythmia*. Data ECG tersebut memiliki ukuran data yang besar sehingga dibutuhkan proses analisa yang tepat. Klasifikasi merupakan salah satu pendekatan yang dapat dilakukan untuk memahami data ECG tersebut. Salah satu algoritma klasifikasi yang dapat menangani data dalam jumlah besar adalah *Naïve Bayes*. Banyaknya fitur yang dimiliki data ECG menjadi hambatan tersendiri dalam proses klasifikasi. Seleksi fitur dapat mengurangi fitur yang kurang relevan pada dataset. Penelitian ini memiliki tujuan untuk mengetahui pengaruh seleksi fitur terhadap performansi model klasifikasi *Naive Bayes* pada diagnosa *arrhythmia*. Hasil yang didapat dari penelitian ini adalah nilai akurasi, presisi, *recall*, dan *f-measure* mengalami peningkatan setelah menerapkan seleksi fitur. Hal ini dibuktikan dengan nilai akurasi yang meningkat sebesar 4%, presisi meningkat sebesar 0.13, nilai *recall* mengalami kenaikan sebesar 0.13, nilai *f-measure* mengalami kenaikan sebesar 0.14, dan waktu komputasi yang dibutuhkan 0.03 detik lebih cepat. Performa tertinggi didapatkan pada klasifikasi dengan 80 fitur. Dihasilkan akurasi sebesar 93 %, presisi dan *recall* sebesar 0.45, *f-measure* sebesar 0.42, dan waktu komputasi selama 0.10 detik.

Kata Kunci: *arrhythmia*, *electrocardiogram* (ECG), klasifikasi, *Naïve Bayes*, seleksi fitur

DAFTAR ISI

HALAMAN PENGESAHAN	v
HALAMAN PERNYATAAN ORISINALITAS.....	vii
HALAMAN MOTTO.....	ix
HALAMAN PERSEMBAHAN	xi
ABSTRAK.....	xiii
KATA PENGANTAR	xv
DAFTAR ISI	xvii
DAFTAR GAMBAR.....	xix
DAFTAR TABEL	xxi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah.....	2
1.3 Tujuan dan Manfaat.....	3
1.4 Batasan Masalah.....	3
1.5 Metodologi Penelitian.....	3
1.6 Sistematika Penulisan	4
BAB II STUDI PUSTAKA	5
2.1 Tinjauan Pustaka	5
2.2 Electrocardiogram	7
2.3 Arrhythmia Dataset	8
2.4 Model Umum Analisa Data.....	9
2.5 Seleksi Fitur.....	9
2.6 Information Gain	10
2.7 Klasifikasi.....	11
2.8 Naïve Bayes	12
2.9 Cross Validation	14
2.10 Evaluasi Klasifikasi	14
BAB III PERANCANGAN SISTEM.....	19
3.1 Input Dataset	20
3.2 Seleksi Fitur.....	20
3.3 Klasifikasi.....	20
3.4 Pengujian dan Evaluasi.....	21
BAB IV IMPLEMENTASI SISTEM	23
4.1 Lingkungan Uji Coba	23
4.2 Pembuatan Program	23
4.2.1 Pembuatan GUI	23

4.2.2	Memuat Dataset	26
4.2.3	Seleksi Fitur	26
4.2.4	Model Klasifikasi.....	27
4.2.5	Evaluasi	28
4.3	Evaluasi dan Hasil Analisa Sistem	28
BAB V PENUTUP.....		47
5.1	Kesimpulan.....	47
5.2	Saran	47
DAFTAR PUSTAKA		49
DAFTAR RIWAYAT HIDUP.....		51

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Gelombang ECG	7
Gambar 2.2 Ilustrasi Kurva ROC	17
Gambar 3.1 Rancangan Sistem	19
Gambar 4.1 Rancangan GUI Sistem	23
Gambar 4.2 Tab Loading Dataset	24
Gambar 4.3 Tab Seleksi Fitur	25
Gambar 4.4 Tab Klasifikasi	26
Gambar 4.5 Uji Coba Memuat Dataset	29
Gambar 4.6 Uji Coba Seleksi Fitur	30
Gambar 4.7 Uji Coba Klasifikasi	30
Gambar 4.8 Uji Coba Kurva ROC	31
Gambar 4.9 Kurva ROC Skenario 1	33
Gambar 4.10 Perbandingan Akurasi Percobaan Skenario 2	34
Gambar 4.11 Perbandingan Presisi, Recall, dan F-measure Percobaan Skenario 2	35
Gambar 4.12 Perbandingan Waktu Komputasi Percobaan Skenario 2	36
Gambar 4.13 Kurva ROC Skenario 2 (n=40)	37
Gambar 4.14 Kurva ROC Skenario 2 (n=80)	38
Gambar 4.15 Kurva ROC Skenario 2 (n=120)	39
Gambar 4.16 Kurva ROC Skenario 2 (n=160)	40
Gambar 4.17 Kurva ROC Skenario 2 (n=200)	41
Gambar 4.18 Kurva ROC Skenario 2 (n=240)	42
Gambar 4.19 Perbandingan Akurasi Skenario 1 dan 2	43
Gambar 4.20 Perbandingan Waktu Komputasi Skenario 1 dan 2	44
Gambar 4.21 Perbandingan Presisi, Recall, dan F-measure Skenario 1 dan Skenario 2	45

DAFTAR TABEL

Tabel 2.1 Rangkuman Penelitian Sebelumnya.....	6
Tabel 2.2 Distribusi Kelas pada Dataset	8
Tabel 2.3 Model Confusion Matrix	15
Tabel 4.1 Detail Spesifikasi Perangkat	23
Tabel 4.2 Hasil Percobaan Skenario 1	32
Tabel 4.3 Hasil Percobaan Skenario 2	33
Tabel 4.3 Perbandingan Hasil Percobaan.....	42