

ABSTRAK

Pada tahun 2015 kurang lebih 17 juta orang meninggal dunia yang disebabkan oleh penyakit jantung dan rusaknya pembuluh darah, serta menurut data *World Health Organization* (WHO) 8.7 juta orang meninggal yang diakibatkan oleh penyakit jantung. Di seluruh dunia, penyakit jantung menjadi penyebab utama kematian termasuk Indonesia. Jadi sangat penting agar penyakit jantung dideteksi secara dini, akan tetapi diagnosa sering kali dibuat dengan berfokus pada naluri dan pengalaman dokter, bukan bersumber pada data yang dicakup dalam database. Hal ini menimbulkan dampak yang tidak diinginkan seperti kesalahan dalam mendiagnosis penyakit dan biaya pengobatan yg tidak perlu pada pasien, apalagi biaya pengobatan penyakit jantung sangat mahal. Dalam penelitian ini, peneliti mengusulkan metode MKNN (*Modifed K-Nearest Neighbor*) dengan seleksi fitur *Pearson Correlation* dan normalisasi data *MinMax* untuk untuk identifikasi penyakit jantung. Dataset yang digunakan pada penelitian ini adalah dataset *Heart Disease* yang diambil dari website *Kaggle Repository* yang terdiri dari 1025 data, 13 fitur dan *binary class* dengan persentase pembagian data 90% *train* dan 10% *test*. Hasil penelitian ini menunjukkan akurasi tertinggi adalah 97% dengan 11 fitur dan k tetangga terdekat adalah 1.

Kata Kunci : *Data mining, Modified K-Nearest Neighbor, Penyakit Jantung, Pearson Correlation*

DAFTAR ISI

HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN ORISINALITAS.....	iii
KATA PENGANTAR	iv
ABSTRAK.....	v
DAFTAR ISI.....	vi
DAFTAR GAMBAR	viii
DAFTAR TABEL.....	ix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah.....	3
1.2.1 Permasalahan.....	3
1.2.2 Metode Usulan	3
1.2.3 Pertanyaan Peneliti.....	3
1.3 Tujuan dan Manfaat.....	3
1.3.1 Tujuan Peneliti	3
1.3.2 Manfaat Penelitian	4
1.4 Batasan – Batasan.....	4
1.5 Sistematika Laporan Skripsi.....	4
BAB II KAJIAN PUSTAKA	6
2.1. Klasifikasi.....	6
2.2. Penyakit Jantung.....	6
2.3. Normalisasi <i>Min-Max</i>	7
2.4. <i>Pearson Correlation</i>	7
2.5. MKNN (<i>Modified K-Nearest Neighbor</i>)	8
2.6. Evaluasi	10
2.6.1 <i>K-fold Cross Valiation</i>	10
2.6.2 <i>Confusion Matrix</i>	11
2.7. Penelitian Terkait	12
BAB III METODE USULAN.....	17
3.1 Metodologi Penelitian	17
3.2 Arsitektur Sistem.....	18

3.3	<i>Flowchart</i> Sistem	19
3.4	<i>Dataset</i>	21
BAB IV HASIL DAN PEMBAHASAN		23
4.1	Lingkungan Uji Coba	23
4.2	Implementasi <i>User Interface</i>	23
4.2.1.	Tampilan Awal Sistem.....	23
4.2.2.	Tampilan Menu <i>Dataset</i>	24
4.2.3.	Tampilan Menu <i>Projects</i>	25
4.2.4.	Tampilan Input Data Baru.....	30
4.3	Implementasi Metode	32
4.3.1.	Skenario Pengujian Pertama	32
4.3.2.	Skenario Pengujian Kedua	37
4.3.3.	Skenario Pengujian Ketiga.....	39
4.4	Evaluasi	40
BAB V PENUTUP.....		43
5.1.	Kesimpulan.....	43
5.2.	Saran.....	43
DAFTAR PUSTAKA		44
LAMPIRAN.....		47

DAFTAR GAMBAR

Gambar 3. 1 Arsitektur Sistem.....	18
Gambar 3. 2 <i>Flowchat</i> Sistem.....	19
Gambar 4. 1 Tampilan Awal Sistem	23
Gambar 4. 2 Tampilan Menu Dataset	24
Gambar 4. 3 Tampilan Menu <i>Projects</i> 1	25
Gambar 4. 4 Tampilan Menu <i>Projects</i> 2	25
Gambar 4. 5 Tampilan Menu <i>Projects</i> 3.....	26
Gambar 4. 6 Tampilan Menu <i>Projects</i> 4.....	26
Gambar 4. 7 Tampilan Menu <i>Projects</i> 5.....	27
Gambar 4. 8 Tampilan Menu <i>Projects</i> 6.....	28
Gambar 4. 9 Tampilan Menu <i>Projects</i> 7	28
Gambar 4. 10 Tampilan Menu <i>Projects</i> 8.....	29
Gambar 4. 11 Tampilan Menu <i>Projects</i> 9.....	29
Gambar 4. 12 Tampilan Menu <i>Projects</i> 10.....	30
Gambar 4. 13 Tampilan <i>Input</i> data baru 1	30
Gambar 4. 14 Tampilan <i>Input</i> data baru 2	31
Gambar 4. 15 Grafik hasil skenario pertama	40
Gambar 4. 16 Grafik hasil skenario kedua.....	41
Gambar 4. 17 Grafik hasil skenario ketiga dengan <i>k-fold</i>	42
Gambar 4. 18 Grafik hasil skenario ketiga tanpa <i>k-fold</i>	42

DAFTAR TABEL

Tabel 2. 1 <i>Confusion Matrix</i>	11
Tabel 3. 1 Keterangan <i>Dataset</i>	21
Tabel 4. 1 Spesifikasi Perangkat	23
Tabel 4. 2 Hasil Normalisasi <i>Min-Max</i>	32
Tabel 4. 3 Korelasi fitur <i>Pearson Correlation</i>	32
Tabel 4. 4 Hasil Skenario Pengujian Pertama	33
Tabel 4. 5 Hasil Skenario Pengujian Kedua.....	37
Tabel 4. 6 Hasil Skenario Pengujian ketiga <i>k-fold</i>	39
Tabel 4. 7 Hasil Skenario Pengujian ketiga tanpa <i>k-fold</i>	40