

Daftar Isi

PERSEMBAHAN	iii
KATA PENGANTAR	v
DAFTAR ISI	vii
DAFTAR GAMBAR	xi
DAFTAR TABEL	xv
DAFTAR ALGORITMA	xxi
DAFTAR CONTOH	xxiii
BAB I PENDAHULUAN	1
1.1 Pengertian Data Mining dan Bukan Data Mining	2
1.2 Pekerjaan dalam Data Mining	5
1.3 Proses Data Mining	7
1.4 Set Data	7
1.5 Konsep Klasifikasi	10
1.6 Konsep Cluster	13
BAB II PEMROSESAN AWAL	17
2.1 Pembuangan Outlier	18
2.2 Normalisasi Data	18
2.3 Data yang Salah	22
2.4 Pemilihan Fitur Berbasis Statistik	24
2.4.1 Pengujian Hipotesis dengan T-Test	24
2.4.2 Fisher's Discriminant Ratio	28
2.5 Pengukuran Kelas yang Dapat Dipisahkan	30
2.5.1 Thornton's Separable Index	30
2.5.2 Direct Class Separability Measure	34
BAB III FITUR	41
3.1 Tipe Fitur	41
3.2 Konsep Kedekatan	42

3.3	Ukuran Kedekatan Nilai Kontinyu	43
3.3.1	Ukuran Ketidakmiripan	43
3.3.2	Ukuran Kemiripan	46
3.4	Ukuran Kedekatan Nilai Diskret	49
3.4.1	Ukuran Ketidakmiripan	49
3.4.2	Ukuran Kemiripan	50
3.5	Ukuran Kedekatan Vektor dengan Fitur Campuran	52
3.6	Ukuran Ketidakmiripan Vektor dengan Himpunan	54
3.7	Ukuran Ketidakmiripan Himpunan dengan Himpunan	54
BAB IV	KLASIFIKASI BERBASIS DECISION TREE	57
4.1	Konsep Decision Tree	57
4.2	Algoritma ID3	59
4.3	Algoritma C4.5	65
4.4	Implementasi Decision Tree di MATLAB	80
4.5	Isu Terkait Decision Tree	83
BAB V	KLASIFIKASI BERBASIS ARTIFICIAL NEURAL NETWORK	85
5.1	Konsep Artificial Neural Network	85
5.2	ANN Perceptron	88
5.2.1	Konsep ANN Perceptron	88
5.2.2	Implementasi ANN Perceptron di MATLAB	91
5.3	ANN Error Backpropagation	95
5.3.1	Konsep Multilayer Perceptron	95
5.3.2	Cara Pelatihan MLP Backpropagation	97
5.3.3	Implementasi MLP Backpropagation di MATLAB	101
5.4	ANN Learning Vector Quantization	112
5.4.1	Konsep Learning Vector Quantization	112
5.4.2	Cara Pelatihan Learning Vector Quantization	112
5.4.3	Implementasi ANN LVQ di MATLAB	114
5.5	Isu Terkait ANN	121
BAB VI	KLASIFIKASI SUPPORT VECTOR MACHINE.....	123
6.1	Konsep SVM	123
6.1.1	SVM Linear	124
6.1.2	Hyperplane SVM	125
6.1.3	SVM Nonlinear	128
6.2	Implementasi SVM di MATLAB	130

6.3	SVM Multikelas	139
6.4	Isu Terkait SVM	147
BAB VII	KLASIFIKASI BERBASIS NEAREST NEIGHBOR.....	149
7.1	Konsep Nearest Neighbor	149
7.2	K-Nearest Neighbor	150
7.3	Template Reduction K-Nearest Neighbor	156
7.4	Support Vector K-Nearest Neighbor Classifier	161
7.5	K-Support Vector Nearest Neighbor	166
7.6	Weight K-Support Vector Nearest Neighbor	178
7.6.1	Pelatihan Model WK-SVNN	180
7.6.2	Prediksi Menggunakan Model WK-SVNN	181
BAB VIII	ANALISIS CLUSTER BERBASIS PARTISI	189
8.1	K-Means	189
8.2	K-Harmonic Means	206
8.3	K-Modes	211
8.4	Fuzzy C-Means	218
BAB IX	ANALISIS CLUSTER BERBASIS HIERARKI	245
9.1	Konsep Pengelompokan Berbasis Hierarki	245
9.2	Metode-Metode Terkait AHC	246
9.3	Implementasi AHC di MATLAB	247
BAB X	EVALUASI SISTEM KLASIFIKASI	257
10.1	Pengukuran Kinerja Klasifikator	257
10.1.1	Akurasi	257
10.1.2	Sensitivitas dan Spesifisitas	259
10.1.3	Precision dan Recall	261
10.2	Metode Evaluasi Klasifikator	263
10.2.1	Hold-Out	263
10.2.2	Random Subsampling	264
10.2.3	K-Fold Cross Validation	264
10.2.4	Bootstrap	264
10.2.5	Teknik Partisi Set Data di MATLAB	265
10.3	Metode Pembandingan Klasifikator	273
10.3.1	Perkiraan Interval Kepercayaan Akurasi	273
10.3.2	Pembandingan Kinerja Dua Model	274
10.3.3	Pembandingan Kinerja Dua Klasifikator	276

BAB XI VALIDITAS CLUSTER	277
11.1 Validitas Internal	278
11.1.1 Davies-Bouldin Index	279
11.1.2 Silhouette Index	283
11.1.3 Dunn Index	289
11.1.4 Validitas Fuzzy Clustering	293
11.2 Validitas Eksternal	304
11.2.1 Validitas Berorientasi Klasifikasi	304
11.2.2 Validitas Berorientasi Kemiripan	307
DAFTAR PUSTAKA	309
LAMPIRAN A	313
LAMPIRAN B	321
LAMPIRAN C	331
GLOSARIUM	354
INDEKS	356
TENTANG PENULIS	359