

DAFTAR ISI

	Hal
HALAMAN JUDUL	i
HALAMAN PERNYATAAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
INTISARI	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL	xv
DAFTAR RUMUS	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Ruang Lingkup.....	6
1.3 Rumusan Masalah.....	6
1.4 Tujuan dan Manfaat Penelitian.....	6
1.5 Sistematika Penulisan	7
BAB II TINJAUAN PUSTAKA	8
2.1 Studi Literatur	8
2.2 Data Mining	15
2.3 Teknik Data Mining.....	16
2.4 Tahapan Data Mining	17
2.5 Algoritma Klasifikasi Data Mining	19

2.6 Algoritma <i>Decision Tree</i>	21
2.7 Algoritma <i>Naive Bayes</i>	25
2.8 Evaluasi dan Validasi Metode Klasifikasi Data Mining	28
2.9 Normalisasi	30
2.10 Matriks Korelasi	31
2.11 Diskritisasi Variabel	32
2.12 Rapidminer.....	32
BAB III METODE PENELITIAN.....	35
3.1 Desain Penelitian	35
3.2 Spesifikasi Hardware dan Software.....	37
3.3 Dataset.....	38
3.4 Normalisasi Data.....	39
3.5 Korelasi Matriks	40
3.6 Diskritisasi Data.....	42
3.7 Modelingg	45
3.8 Evaluation	45
3.9 Deployment.....	45
BAB IV ANALISA DAN PEMBAHASAN	46
4.1 Diskritisasi Data.....	46
4.2 Decision Tree C4.5	54
4.3 Gaussian Naïve Bayes	74
4.4 Pengujian Model	81
4.5 Analisa Komparasi.....	105
BAB V KESIMPULAN DAN SARAN	111
5.1 Kesimpulan	111

5.2 Saran.....	111
DAFTAR PUSTAKA	112

DAFTAR GAMBAR

	Hal
Gambar 1.1 Peringkat Ancaman Kejahatan Siber di Dunia (Top Ten).....	1
Gambar 1.2 Ancaman dan serangan pada ICS.....	3
Gambar 2.1 Tahapan Data Mining	19
Gambar 3.1 Kerangka Kerja	35
Gambar 3.2 Diagram Proses Pengolahan Data	36
Gambar 3.3 Proses Normalisasi Data	40
Gambar 3.4 Proses Proses Korelasi Matrik	41
Gambar 3.5 Hasil Proses Proses Korelasi Matrik.....	41
Gambar 3.6 Proses Diskritisasi Secara Umum	43
Gambar 3.7 Pemodelan Diskritisasi Data	44
Gambar 3.8 Hasil Proses Diskritisasi Data.....	45
Gambar 4.1 Pohon Keputusan Node 1	65
Gambar 4.2 Pohon Keputusan Node 2	67
Gambar 4.3 Pohon Keputusan Node 3	71
Gambar 4.4 Pohon Keputusan Akhir.....	72
Gambar 4.5 Pemodelan Klasifikasi Decision Tree (Rapidminer).....	73
Gambar 4.6 Hasil Pemodelan Klasifikasi Decision Tree (Rapidminer)	73
Gambar 4.7 Pemodelan Klasifikasi Naïve Bayes (Rapidminer)	80
Gambar 4.8 Hasil Distribusi Klasifikasi Naïve Bayes (Rapidminer)	80
Gambar 4.9 Struktur Proses Pengujian dengan 10-Folds Cross Validation	82
Gambar 4.10 Struktur Decision Tree Kriteria Gini Index.....	82
Gambar 4.11 Confusion Matrix C4.5 Tanpa Diskritisasi	82
Gambar 4.12 Pengukuran AUC Pesimistic C4.5 Tanpa Diskritisasi	84
Gambar 4.13 Proses C4.5 Diskritisasi 3 Bin No Define Boundaries.....	84
Gambar 4.14 Confusion Matrix C4.5 Diskritisasi 3 Bin No Define Boundaries	85
Gambar 4.15 Pengukuran AUC C4.5 Diskritisasi 3 Bin No Define Boundaries	86
Gambar 4.16 Proses C4.5 Diskritisasi 3 Bin Define Boundaries.....	87
Gambar 4.17 Confusion Matrix C4.5 Diskritisasi 3 Bin Define Boundaries	87
Gambar 4.18 Pengukuran AUC C4.5 Diskritisasi 3 Bin Define Boundaries	88

Gambar 4.19 Proses C4.5 Diskritisasi 5 Bin No Define Boundaries.....	89
Gambar 4.20 Confusion Matrix C4.5 Diskritisasi 5 Bin No Define Boundaries	89
Gambar 4.21 Pengukuran AUC C4.5 Diskritisasi 5 Bin No Define Boundaries	90
Gambar 4.22 Proses C4.5 Diskritisasi 5 Bin Define Boundaries.....	91
Gambar 4.23 Confusion Matrix C4.5 Diskritisasi 5 Bin Define Boundaries	91
Gambar 4.24 Pengukuran AUC C4.5 Diskritisasi 5 Bin Define Boundaries	92
Gambar 4.25 Diagram Akurasi, Precision dan Recall C4.5 Gini Index	93
Gambar 4.26 Diagram AUC C4.5 Gini Index.....	94
Gambar 4.27 Struktur Naïve Bayes.....	94
Gambar 4.28 Confusion Naïve Bayes Tanpa Diskritisasi.....	94
Gambar 4.29 Pengukuran AUC Naïve Bayes Tanpa Diskritisasi.....	96
Gambar 4.30 Proses Naïve Bayes Diskritisasi 3 Bin No Define Boundaries....	96
Gambar 4.31 Confusion Matrix Naïve Bayes Diskritisasi 3 Bin No Define Boundaries.....	97
Gambar 4.32 Pengukuran AUC Naïve Bayes Diskritisasi 3 Bin No Define Boundaries.....	98
Gambar 4.33 Proses Naïve Bayes Diskritisasi 3 Bin Define Boundaries.....	98
Gambar 4.34 Confusion Matrix Naïve Bayes Diskritisasi 3 Bin Define Boundaries.....	99
Gambar 4.35 Pengukuran AUC Naïve Bayes Diskritisasi 3 Bin Define Boundaries.....	100
Gambar 4.36 Proses Naïve Bayes Diskritisasi 5 Bin No Define Boundaries..	100
Gambar 4.37 Confusion Matrix Naïve Bayes Diskritisasi 5 Bin No Define Boundaries.....	101
Gambar 4.38 Pengukuran AUC Naïve Bayes Diskritisasi 5 Bin No Define Boundaries.....	102
Gambar 4.39 Proses Naïve Bayes Diskritisasi 5 Bin Define Boundaries.....	102
Gambar 4.40 Confusion Matrix Naïve Bayes Diskritisasi 5 Bin Define Boundaries.....	103
Gambar 4.41 Pengukuran AUC Naïve Bayes Diskritisasi 5 Bin Define Boundaries.....	104
Gambar 4.42 Diagram Akurasi, Precision dan Recall Naïve Bayes Gini Index	105
Gambar 4.43 Diagram AUC Naïve Bayes	105
Gambar 4.44 T-Test C4.5 dan NB Tanpa Diksritisasi	106

Gambar 4.45 Hasil T-Test C4.5 dan NB Tanpa Diksritisasi.....	106
Gambar 4.46 T-Test C4.5 dan NB Dengan Diksritisasi	106
Gambar 4.47 Hasil T-Test C4.5 dan NB Dengan Diksritisasi 3 Bin Tanpa Parameter Batasan -1 dan 1	107
Gambar 4.48 Hasil T-Test C4.5 dan NB Dengan Diksritisasi 3 Bin Input Parameter Batasan -1 dan 1	107
Gambar 4.49 Hasil T-Test C4.5 dan NB Dengan Diksritisasi 5 Bin Tanpa Parameter Batasan -1 dan 1	107
Gambar 4.50 Hasil T-Test C4.5 dan NB Dengan Diksritisasi 5 Bin Input Parameter Batasan -1 dan 1	107
Gambar 4.51 Diagram Kenaikan Tingkat Akurasi Metode C4.5 dan Naïve Bayes	109
Gambar 4.52 Diagram Tingkat Kesalahan klasifikasi Metode C4.5 dan Naïve Bayes	110

DAFTAR TABEL

	Hal
Tabel 2.1 Hasil pengujian penelitian sebelumnya	12
Tabel 2.2 Sate Of The Art.....	13
Tabel 2.3 Model Confusion Matrix	29
Tabel 2.4 Akurasi AUC	30
Tabel 3.1 Tipe Data Attribute Dataset Malware.....	38
Tabel 3.2 Contoh Isi Dataset Malware	39
Tabel 3.3 Hasil Normalisasi Data	40
Tabel 3.4 Dataset Hasil Korelasi Matrik	42
Tabel 4.1 Hasil Diskritisasi 3 Bin dan 5 Bin Tanpa Parameter Batasan (Manual)	48
Tabel 4.2 Hasil Diskritisasi 3 Bin dan 5 Bin Tanpa Parameter Batasan (Rapidminer).....	49
Tabel 4.3 Hasil Diskritisasi 3 Bin dan 5 Bin Tanpa Parameter Batasan (Manual)	52
Tabel 4.4 Hasil Diskritisasi 3 Bin dan 5 Bin Tanpa Parameter Batasan (Rapidminer).....	53
Tabel 4.5 Dataset Malware Sampel 50 Record.....	55
Tabel 4.6 Pemisahan Nilai Atribut Millisecond.....	57
Tabel 4.7 Nilai Gini Atribut Millisecond	60
Tabel 4.8 Kandidat Split Atribut Pada Node 1	63
Tabel 4.9 Kandidat Split Atribut Pada Node 2	65
Tabel 4.10 Kandidat Split Atribut Pada Node 3.1	68
Tabel 4.11 Kandidat Split Atribut Pada Node 3.2	69
Tabel 4.12 Mean dan Standar Deviasi.....	76
Tabel 4.13 Pengujian Model C4.5 Gini Index dan Gaussian Naïve Bayes	84
Tabel 4.14 Confusion Matrix C4.5 Tanpa Diskritisasi	83
Tabel 4.15 Confusion Matrix C4.5 Diskritisasi 3 Bin No Define Boundaries ...	85
Tabel 4.16 Confusion Matrix C4.5 Diskritisasi 3 Bin Define Boundaries	87
Tabel 4.17 Confusion Matrix C4.5 Diskritisasi 5 Bin No Define Boundaries ...	92
Tabel 4.18 Confusion Matrix C4.5 Diskritisasi 5 Bin Define Boundaries	89
Tabel 4.19 Hasil Pengujian Algoritma C4.5 Kriteria Gini Index	93
Tabel 4.20 Confusion Matrix Naïve Bayes Tanpa Diskritisasi	95

Tabel 4.21 Confusion Matrix Naïve Bayes Diskritisasi 3 Bin No Define Boundaries.....	97
Tabel 4.22 Confusion Matrix Naïve Bayes Diskritisasi 3 Bin Define Boundaries	99
Tabel 4.23 Confusion Matrix Naïve Bayes Diskritisasi 5 Bin No Define Boundaries.....	101
Tabel 4.24 Confusion Matrix Naïve Bayes Diskritisasi 5 Bin Define Boundaries	103
Tabel 4.25 Hasil Pengujian Algoritma Naïve Bayes	104
Tabel 4.26 Hasil Pengujian T-Test C4.5 dan Naïve Bayes.....	108
Tabel 4.27 Tingkat Kesalahan Klasifikasi C4.5 dan Naïve Bayes	109

DAFTAR RUMUS

	Hal
Rumus 2.1 Information Gain	22
Rumus 2.2 Entropy	22
Rumus 2.3 Gain Ratio	23
Rumus 2.4 Split Info	23
Rumus 2.5 Gini Index	23
Rumus 2.6 Gini Index Subset.....	23
Rumus 2.7 Teorema Bayes.....	26
Rumus 2.8 Probabilitas Prior	26
Rumus 2.9 Gaussian Naïve Bayes.....	27
Rumus 2.10 Bernouli Naïve Bayes	27
Rumus 2.11 Multinomial Naïve Bayes	27
Rumus 2.12 Penghitungan Frekuensi Relatif pada multinomial	28
Rumus 2.13 Precision.....	29
Rumus 2.14 Recall.....	29
Rumus 2.15 Accuracy	29
Rumus 2.16 AUC	30
Rumus 2.17 Z-Score.....	30
Rumus 2.18 Koefisien Korelasi Pearson.....	31