

BAB IV

KLASIFIKASI DATA UNTUK DETEKSI ANOMALI

4.1 PERSIAPAN DATA EKSPERIMEN

4.1.1 Profil Data Eksperimen

Dataset yang digunakan pada penelitian ini adalah salah satu *dataset* CICIDS2017 yaitu *Thursday-WorkingHours-Morning-WebAttack.pcap_ISCX* yang merupakan subset dari CICIDS2017 yang dirilis pada tahun 2017. Data terdiri dari data normal dan data anomali trafik *web attack*, yaitu *web attack brute force*, *sql injection* dan *cross site scripting* (xss). Jumlah *record* data pada *dataset* adalah 170.366 *record* data dengan komposisi data normal maupun data intrusi yang berbeda-beda dan dapat dilihat pada tabel 4.1. Untuk fitur-fitur yang terdapat pada *dataset* ini telah disajikan pada tabel 2.2.

Tabel 4.1 Profil Data Penelitian

| Jenis Data | Jumlah |
|--------------------------|----------------|
| Benign (Normal) | 168.186 |
| WebAttack-Brute Force | 1.507 |
| WebAttack-XSS | 652 |
| Web Attack-Sql Injection | 21 |
| Total | 170.366 |

4.1.2 Data Preprocessing

Pada tahapan ini data dilakukan tahap data *cleaning* atau melakukan persiapan data sebelum diolah melalui proses data *mining* yang akan dilakukan dengan menggunakan alat bantu WEKA. Data *cleaning* dapat dilakukan dengan menghapus data yang tidak sesuai atau relevan seperti duplikasi data. Pada *dataset*

4.1.4 Split Data

Untuk keperluan eksperimen dan pengujian dilakukan pemisahan *dataset*. *Dataset* dipisahkan menjadi data *training* dan data *testing*. Pada penelitian ini dilakukan pemisahan data dengan perbandingan 60 : 40 atau 60% data *training* dan 40% data *testing*, perbandingan 70 : 30 atau 70% data *training* dan 30% data *testing*, serta perbandingan 80 : 20 atau 80% data *training* dan 20% data *testing*. Setelah dilakukan tahap split data pada setiap pembagian data (*dataset*) akan memiliki jumlah *record* data dan komposisi data yang berbeda-beda, sehingga profil *dataset* setelah dilakukan split data disajikan sebagai berikut :

Tabel 4.2 Profil Data Pada Split Data 60 : 40

| Jenis Data | 60% | 40% |
|-------------------------|---------|--------|
| Benign (Normal) | 100.912 | 67.274 |
| WebAttack-Brute Force | 904 | 603 |
| Web Attack-XSS | 391 | 261 |
| WebAttack-Sql Injection | 13 | 8 |
| Total | 102.220 | 68.146 |

Tabel 4.3 Profil Data Pada Split Data 70 : 30

| Jenis Data | 70% | 30% |
|-------------------------|---------|--------|
| Benign (Normal) | 117.730 | 50.456 |
| WebAttack-Brute Force | 1.055 | 452 |
| Web Attack-XSS | 456 | 196 |
| WebAttack-Sql Injection | 15 | 6 |
| Total | 119.256 | 51.110 |

Tabel 4.4 Profil Data Pada Split Data 80 : 20

| Jenis Data | 80% | 20% |
|-------------------------|---------|--------|
| Benign (Normal) | 134.549 | 33.637 |
| WebAttack-Brute Force | 1.206 | 301 |
| Web Attack-XSS | 522 | 130 |
| WebAttack-Sql Injection | 17 | 4 |
| Total | 136.294 | 34.072 |

4.1.5 Feature Selection

Menurut Vriza dkk.[31] untuk mengurangi dimensi atribut yang kurang relevan maka perlu teknik seleksi fitur untuk menghapus fitur tersebut, dengan penggunaannya memungkinkan didapatkan hasil yang lebih baik dari sisi akurasi daripada tanpa menggunakannya. Teknik seleksi fitur secara luas telah digunakan hampir di banyak bidang seperti kategorisasi teks, analisis *genomic*, deteksi intrusi, *bio informatic* dan lain-lain [32]. *Information gain* merupakan salah satu metode filter dalam seleksi fitur yang dapat bekerja dengan baik terhadap data berdimensi tinggi dan cocok digunakan terhadap *dataset* yang datanya bertipe *numeric*. [33] Pada penelitian ini peneliti melakukan seleksi fitur menggunakan *InfoGainAttributEval* dengan metode *ranked* atau pemilihan fitur dengan *ranking* tertinggi, adapun hasil seleksi fitur dengan metode yang diusulkan disajikan pada tabel 4.5.

Tabel 4.5 Hasil Seleksi Fitur *Information Gain*

| Rank | Bobot | ID | Nama Fitur |
|-------------|--------------|-----------|-------------------------|
| 1 | 0.084638 | 67 | Init_Win_bytes_backward |
| 2 | 0.075347 | 25 | Fwd IAT Min |
| 3 | 0.057179 | 36 | Fwd Packets/s |
| 4 | 0.056886 | 55 | Fwd Header Length |
| 5 | 0.056646 | 16 | Flow Packets/s |
| 6 | 0.055857 | 22 | Fwd IAT Mean |
| 7 | 0.05437 | 18 | Flow IAT Std |
| 8 | 0.053655 | 17 | Flow IAT Mean |
| 9 | 0.051543 | 19 | Flow IAT Max |
| 10 | 0.051265 | 23 | Fwd IAT Std |
| 11 | 0.050011 | 37 | Bwd Packets/s |
| 12 | 0.049004 | 24 | Fwd IAT Max |
| 13 | 0.048183 | 2 | Flow Duration |
| 14 | 0.047764 | 21 | Fwd IAT Total |
| 15 | 0.046978 | 20 | Flow IAT Min |
| 16 | 0.044595 | 39 | Max Packet Length |

| | | | |
|----|----------|----|-----------------------------|
| 17 | 0.044094 | 7 | Fwd Packet Length Max |
| 18 | 0.043605 | 9 | Fwd Packet Length Mean |
| 19 | 0.043605 | 53 | Avg Fwd Segment Size |
| 20 | 0.04353 | 63 | Subflow Fwd Bytes |
| 21 | 0.04353 | 5 | Total Length of Fwd Packets |
| 22 | 0.043512 | 66 | Init_Win_bytes_forward |
| 23 | 0.042971 | 52 | 0. Average Packet Size |
| 24 | 0.042525 | 40 | Packet Length Mean |
| 25 | 0.041813 | 3 | Total Fwd Packets |
| 26 | 0.041813 | 62 | Subflow Fwd Packets |
| 27 | 0.039649 | 1 | Destination Port |
| 28 | 0.039224 | 15 | Flow Bytes/s |
| 29 | 0.031639 | 11 | Bwd Packet Length Max |
| 30 | 0.029485 | 65 | Subflow Bwd Bytes |
| 31 | 0.029485 | 6 | Total Length of Bwd Packets |
| 32 | 0.029217 | 13 | Bwd Packet Length Mean |
| 33 | 0.029217 | 54 | Avg Bwd Segment Size |
| 34 | 0.027891 | 42 | Packet Length Variance |
| 35 | 0.027891 | 41 | Packet Length Std |
| 36 | 0.027646 | 35 | Bwd Header Length |
| 37 | 0.021718 | 4 | Total Backward Packets |
| 38 | 0.021718 | 64 | Subflow Bwd Packets |
| 39 | 0.020148 | 46 | PSH Flag Count |
| 40 | 0.020103 | 68 | act_data_pkt_fwd |
| 41 | 0.019253 | 8 | Fwd Packet Length Min |
| 42 | 0.019069 | 38 | Min Packet Length |
| 43 | 0.018101 | 29 | Bwd IAT Max |
| 44 | 0.017448 | 26 | Bwd IAT Total |
| 45 | 0.017067 | 27 | Bwd IAT Mean |
| 46 | 0.016207 | 12 | Bwd Packet Length Min |
| 47 | 0.015851 | 69 | min_seg_size_forward |
| 48 | 0.012948 | 51 | Down/Up Ratio |
| 49 | 0.012359 | 30 | Bwd IAT Min |
| 50 | 0.011553 | 10 | Fwd Packet Length Std |
| 51 | 0.00964 | 14 | Bwd Packet Length Std |
| 52 | 0.00838 | 28 | Bwd IAT Std |
| 53 | 0.003338 | 77 | Idle Min |
| 54 | 0.003338 | 76 | Idle Max |
| 55 | 0.003338 | 74 | Idle Mean |
| 56 | 0.003317 | 70 | Active Mean |
| 57 | 0.003317 | 72 | Active Max |
| 58 | 0.003317 | 73 | Active Min |

| | | | |
|----|----------|----|----------------------|
| 59 | 0.002387 | 47 | ACK Flag Count |
| 60 | 0.00204 | 75 | Idle Std |
| 61 | 0.002031 | 71 | Active Std |
| 62 | 0.000834 | 44 | SYN Flag Count |
| 63 | 0.000834 | 31 | Fwd PSH Flags |
| 64 | 0.000562 | 48 | URG Flag Count |
| 65 | 0.000203 | 43 | FIN Flag Count |
| 66 | 0 | 50 | ECE Flag Count |
| 67 | 0 | 61 | Bwd Avg Bulk Rate |
| 68 | 0 | 56 | Fwd Avg Bytes/Bulk |
| 69 | 0 | 45 | RST Flag Count |
| 70 | 0 | 33 | Fwd URG Flags |
| 71 | 0 | 34 | Bwd URG Flags |
| 72 | 0 | 32 | Bwd PSH Flags |
| 73 | 0 | 60 | Bwd Avg Packets/Bulk |
| 74 | 0 | 59 | Bwd Avg Bytes/Bulk |
| 75 | 0 | 58 | Fwd Avg Bulk Rate |
| 76 | 0 | 49 | CWE Flag Count |
| 77 | 0 | 57 | Fwd Avg Packets/Bulk |

Dari hasil seleksi fitur yang telah dilakukan menggunakan *information gain*, berdasarkan hasil eksperimen diambil fitur dengan nilai bobot >0.05 dan didapat 11 fitur terpilih yang dapat dilihat pada tabel 4.6.

Tabel 4.6 Fitur Terpilih

| Rank | Bobot | ID | Nama Fitur |
|------|----------|----|-------------------------|
| 1 | 0.084638 | 67 | Init_Win_bytes_backward |
| 2 | 0.075347 | 25 | Fwd IAT Min |
| 3 | 0.057179 | 36 | Fwd Packets/s |
| 4 | 0.056886 | 55 | Fwd Header Length |
| 5 | 0.056646 | 16 | Flow Packets/s |
| 6 | 0.055857 | 22 | Fwd IAT Mean |
| 7 | 0.05437 | 18 | Flow IAT Std |
| 8 | 0.053655 | 17 | Flow IAT Mean |
| 9 | 0.051543 | 19 | Flow IAT Max |
| 10 | 0.051265 | 23 | Fwd IAT Std |
| 11 | 0.050011 | 37 | Bwd Packets/s |

4.2 KLASIFIKASI DENGAN TOOL WEKA

Pada tahapan ini *dataset* yang telah siap akan dilakukan proses pengklasifikasian atau perhitungan menggunakan *tools* bantu WEKA. Dalam penelitian ini peneliti melakukan 3 pembagian jumlah *dataset*. Kemudian setiap *dataset* diuji dengan 3 mode pengujian (validasi) yaitu menggunakan *Use Training Set*, *5-Cross Validation* dan *10-Cross Validation*.

4.2.1 KLASIFIKASI TANPA SELEKSI ATRIBUT

4.2.1.1 Klasifikasi Pada Split Data 60 : 40 Tanpa Seleksi Fitur

Perbandingan split data yang digunakan yaitu 60:40, 60% dari 170.366 *record* data atau sebanyak 102.220 *record* data digunakan sebagai data *training* dan 40% atau 68.146 *record* data sebagai *data testing* tanpa seleksi fitur atau menggunakan 77 fitur awal.

1. Naive Bayes

a. Use Training Set

Tabel 4.7 Hasil Klasifikasi NB UTS Split Data 60:40 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,978 | 0,011 | 1,000 | 0,978 | 0,989 | 0,597 | 0,997 | 1,000 |
| Web Attack Brute Force | 0,588 | 0,019 | 0,220 | 0,588 | 0,320 | 0,351 | 0,980 | 0,179 |
| Web Attack XSS | 0,442 | 0,005 | 0,259 | 0,442 | 0,327 | 0,335 | 0,988 | 0,226 |
| Web Attack Sql Injection | 1,000 | 0,004 | 0,031 | 1,000 | 0,061 | 0,176 | 0,998 | 0,038 |
| Weighted Avg. | 0,973 | 0,012 | 0,990 | 0,973 | 0,980 | 0,594 | 0,997 | 0,990 |

Tabel 4.7 merupakan hasil klasifikasi dengan menggunakan algoritma *Naive Bayes* pada mode pengujian *Use Training Set*, menunjukkan hasil 99.423 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan

akurasi sebesar 97,2637% dan 2.797 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 2,7363% , dengan waktu proses klasifikasi selama 29,06 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 158 *point a*.

Tabel 4.8 Confusion Matrix NB UTS Split Data 60:40 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|-----|-----|------------------------------|--------|
| a | b | c | d | | |
| 98705 | 1687 | 212 | 308 | a = Benign | Aktual |
| 2 | 532 | 283 | 87 | b = Web Attack Brute Force | |
| 13 | 197 | 173 | 8 | c = Web Attack XSS | |
| 0 | 0 | 0 | 13 | d = Web Attack Sql Injection | |

Tabel 4.8 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 98.705 data diprediksi benar kategori *benign*, 1.687 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 212 data salah dikategorikan *web attack xss* seharusnya *benign* dan 308 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 2 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 532 data benar diprediksi sebagai *web attack brute force*, 283 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 87 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 13 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 197 data salah

dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 173 data benar diprediksi sebagai *web attack xss* dan 8 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 13 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.9 Hasil Klasifikasi NB 5-CV Split Data 60:40 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,973 | 0,010 | 1,000 | 0,973 | 0,986 | 0,556 | 0,995 | 1,000 |
| Web Attack Brute Force | 0,386 | 0,023 | 0,132 | 0,386 | 0,197 | 0,215 | 0,975 | 0,147 |
| Web Attack XSS | 0,547 | 0,008 | 0,218 | 0,547 | 0,311 | 0,341 | 0,984 | 0,196 |
| Web Attack Sql Injection | 0,923 | 0,004 | 0,029 | 0,923 | 0,057 | 0,164 | 0,998 | 0,034 |
| Weighted Avg. | 0,966 | 0,010 | 0,989 | 0,966 | 0,977 | 0,552 | 0,995 | 0,989 |

Tabel 4.9 merupakan hasil klasifikasi dengan menggunakan algoritma *Naive Bayes* pada mode pengujian *5 Cross Validation*, menunjukkan hasil 98.750 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 96,6054% dan 3.470 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 3,3946% , dengan waktu proses klasifikasi selama 3,1 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 158 *point* b.

Tabel 4.10 Confusion Matrix NB 5-CV Split Data 60:40 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|-----|-----|------------------------------|---------------|
| a | b | c | d | | |
| 98175 | 2134 | 294 | 309 | a = Benign | Aktual |
| 1 | 349 | 475 | 79 | b = Web Attack Brute Force | |
| 12 | 156 | 214 | 9 | c = Web Attack XSS | |
| 0 | 1 | 0 | 12 | d = Web Attack Sql Injection | |

Tabel 4.10 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 98.175 data diprediksi benar kategori *benign*, 2.134 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 294 data salah dikategorikan *web attack xss* seharusnya *benign* dan 309 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 1 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 349 data benar diprediksi sebagai *web attack brute force*, 475 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 79 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 12 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 156 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 214 data benar diprediksi sebagai *web attack xss* dan 9 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 1 data salah dikategorikan

sebagai *web attack brute force* seharusnya *web attack sql injection* dan 12 data diprediksi benar kategori *web attack sql injection*.

c. *10 Cross Validation*

Tabel 4.11 Hasil Klasifikasi NB 10-CV Split Data 60:40 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,978 | 0,011 | 1,000 | 0,978 | 0,989 | 0,595 | 0,997 | 1,000 |
| Web Attack Brute Force | 0,378 | 0,018 | 0,160 | 0,378 | 0,225 | 0,236 | 0,979 | 0,174 |
| Web Attack XSS | 0,529 | 0,008 | 0,213 | 0,529 | 0,304 | 0,332 | 0,985 | 0,188 |
| Web Attack Sql Injection | 1,000 | 0,004 | 0,031 | 1,000 | 0,060 | 0,175 | 0,998 | 0,035 |
| Weighted Avg. | 0,971 | 0,012 | 0,989 | 0,971 | 0,979 | 0,591 | 0,996 | 0,989 |

Tabel 4.11 merupakan hasil klasifikasi dengan menggunakan algoritma *Naive Bayes* pada mode pengujian *10 Cross Validation*, menunjukkan hasil 99.244 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 97,0886% dan 2.976 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 2,9114% , dengan waktu proses klasifikasi selama 2,69 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 159 *point c*.

Tabel 4.12 Confusion Matrix NB 10-CV Split Data 60:40 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|-----|-----|------------------------------|---------------|
| a | b | c | d | | |
| 98682 | 1626 | 291 | 313 | a = Benign | Aktual |
| 2 | 342 | 473 | 87 | b = Web Attack Brute Force | |
| 13 | 163 | 207 | 8 | c = Web Attack XSS | |
| 0 | 0 | 0 | 13 | d = Web Attack Sql Injection | |

Tabel 4.12 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 98.682 data diprediksi benar kategori *benign*, 1.626 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 291 data salah dikategorikan *web attack xss* seharusnya *benign* dan 313 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 2 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 342 data benar diprediksi sebagai *web attack brute force*, 473 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 87 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 13 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 163 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 207 data benar diprediksi sebagai *web attack xss* dan 8 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 13 data diprediksi benar kategori *web attack sql injection*.

2. C4.5

a. Use Training Set

Tabel 4.13 Hasil Klasifikasi C45 UTS Split Data 60:40 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,006 | 1,000 | 1,000 | 1,000 | 0,995 | 0,999 | 1,000 |
| Web Attack Brute Force | 0,990 | 0,003 | 0,729 | 0,990 | 0,840 | 0,848 | 0,999 | 0,833 |
| Web Attack XSS | 0,146 | 0,000 | 0,864 | 0,146 | 0,249 | 0,354 | 0,998 | 0,497 |
| Web Attack Sql Injection | 0,692 | 0,000 | 0,900 | 0,692 | 0,703 | 0,709 | 0,935 | 0,603 |
| Weighted Avg. | 0,997 | 0,006 | 0,997 | 0,997 | 0,996 | 0,992 | 0,999 | 0,997 |

Tabel 4.13 merupakan hasil klasifikasi dengan menggunakan algoritma C4.5 pada mode pengujian *Use Training Set*, menunjukkan hasil 101.869 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,6566% dan 351 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,3434%, dengan waktu proses klasifikasi selama 48,55 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 159 *point a*.

Tabel 4.14 Confusion Matrix C45 UTS Split Data 60:40 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|----|---|------------------------------|--------|
| a | b | c | d | | |
| 100908 | 2 | 2 | 0 | a = Benign | Aktual |
| 2 | 895 | 7 | 0 | b = Web Attack Brute Force | |
| 2 | 331 | 57 | 1 | c = Web Attack XSS | |
| 4 | 0 | 0 | 9 | d = Web Attack Sql Injection | |

Tabel 4.14 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “*Benign*” terdapat sebanyak 100.908 data diprediksi benar kategori *benign*, 2 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, dan 2 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “*Web Attack Brute Force*” terdapat sebanyak 2 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 895 data benar diprediksi sebagai *web attack brute force*, dan 7 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “*Web Attack XSS*” terdapat sebanyak 2 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 331 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 57 data benar diprediksi sebagai *web attack xss* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada *class* “*Web Attack Sql Injection*” terdapat sebanyak 4 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection* dan 9 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.15 Hasil Klasifikasi C45 5-CV Split Data 60:40 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,013 | 1,000 | 1,000 | 1,000 | 0,986 | 0,997 | 1,000 |
| Web Attack Brute Force | 0,951 | 0,004 | 0,707 | 0,951 | 0,811 | 0,818 | 0,995 | 0,785 |
| Web Attack XSS | 0,107 | 0,000 | 0,483 | 0,107 | 0,176 | 0,226 | 0,980 | 0,403 |
| Web Attack Sql Injection | 0,462 | 0,000 | 0,857 | 0,462 | 0,600 | 0,629 | 0,688 | 0,534 |
| Weighted Avg. | 0,996 | 0,013 | 0,995 | 0,996 | 0,995 | 0,982 | 0,997 | 0,996 |

Tabel 4.15 merupakan hasil klasifikasi dengan menggunakan algoritma *C4.5* pada mode pengujian *5 Cross Validation*, menunjukkan hasil 101.801 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5901% dan 419 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,4099% , dengan waktu proses klasifikasi selama 49,14 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 160 *point* b.

Tabel 4.16 Confusion Matrix C45 5-CV Split Data 60:40 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|----|---|------------------------------|--------|
| a | b | c | d | | |
| 100893 | 15 | 4 | 0 | a = Benign | Aktual |
| 4 | 860 | 40 | 0 | b = Web Attack Brute Force | |
| 10 | 338 | 42 | 1 | c = Web Attack XSS | |
| 3 | 3 | 1 | 6 | d = Web Attack Sql Injection | |

Tabel 4.16 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “*Benign*” terdapat sebanyak 100.893 data diprediksi benar kategori *benign*, 15 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, dan 4 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “*Web Attack Brute Force*” terdapat sebanyak 4 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 860 data benar diprediksi sebagai *web attack brute force*, dan 40 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “*Web Attack XSS*” terdapat sebanyak 10 data salah

dikategorikan sebagai *benign* seharusnya *web attack xss*, 338 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 42 data benar diprediksi sebagai *web attack xss* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 3 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 3 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection*, 1 data salah dikategorikan sebagai *web attack xss* seharusnya *web attack sql injection* dan 6 data diprediksi benar kategori *web attack sql injection*.

c. *10 Cross Validation*

Tabel 4.17 Hasil Klasifikasi C45 10-CV Split Data 60:40 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,013 | 1,000 | 1,000 | 1,000 | 0,988 | 0,996 | 1,000 |
| Web Attack Brute Force | 0,995 | 0,004 | 0,702 | 0,955 | 0,809 | 0,817 | 0,996 | 0,815 |
| Web Attack XSS | 0,069 | 0,000 | 0,422 | 0,069 | 0,119 | 0,170 | 0,989 | 0,419 |
| Web Attack Sql Injection | 0,615 | 0,000 | 0,727 | 0,615 | 0,667 | 0,669 | 0,675 | 0,575 |
| Weighted Avg. | 0,996 | 0,013 | 0,995 | 0,996 | 0,995 | 0,984 | 0,996 | 0,996 |

Tabel 4.17 merupakan hasil klasifikasi dengan menggunakan algoritma *C4.5* pada mode pengujian *10 Cross Validation*, menunjukkan hasil 101.797 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5862% dan 423 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,4138% , dengan waktu proses klasifikasi selama 46,2 detik. Untuk *screen shot* hasil

pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 160 *point c*.

Tabel 4.18 Confusion Matrix C45 10-CV Split Data 60:40 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|----|---|------------------------------|--------|
| a | b | c | d | | |
| 100899 | 10 | 2 | 1 | a = Benign | Aktual |
| 5 | 863 | 35 | 1 | b = Web Attack Brute Force | |
| 8 | 355 | 27 | 1 | c = Web Attack XSS | |
| 4 | 1 | 0 | 8 | d = Web Attack Sql Injection | |

Tabel 4.18 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 100.899 data diprediksi benar kategori *benign*, 10 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 2 data salah dikategorikan *web attack xss* seharusnya *benign* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 5 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 863 data benar diprediksi sebagai *web attack brute force*, 35 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 8 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 355 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 27 data benar diprediksi sebagai *web attack xss* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada

class “*Web Attack Sql Injection*” terdapat sebanyak 4 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 1 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 8 data diprediksi benar kategori *web attack sql injection*.

3. Random Forest

a. Use Training Set

Tabel 4.19 Hasil Klasifikasi RF UTS Split Data 60:40 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| Web Attack Brute Force | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| Web Attack XSS | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| Web Attack Sql Injection | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| Weighted Avg. | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |

Tabel 4.19 merupakan hasil klasifikasi dengan menggunakan algoritma *Random Forest* pada mode pengujian *Use Training Set*, menunjukkan hasil 102.220 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 100% dan 0 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0% , dengan waktu proses klasifikasi selama 153,94 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 161 *point a*.

Tabel 4.20 Confusion Matrix RF UTS Split Data 60:40 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|-----|----|------------------------------|--------|
| a | b | c | d | | |
| 100912 | 0 | 0 | 0 | a = Benign | Aktual |
| 0 | 904 | 0 | 0 | b = Web Attack Brute Force | |
| 0 | 0 | 391 | 0 | c = Web Attack XSS | |
| 0 | 0 | 0 | 13 | d = Web Attack Sql Injection | |

Tabel 4.20 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 100.912 data diprediksi benar kategori *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 904 data diprediksi benar *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 391 data diprediksi benar *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 13 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.21 Hasil Klasifikasi RF 5-CV Split Data 60:40 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,028 | 1,000 | 1,000 | 1,000 | 0,986 | 1,000 | 1,000 |
| Web Attack Brute Force | 0,794 | 0,002 | 0,739 | 0,794 | 0,766 | 0,764 | 0,999 | 0,827 |
| Web Attack XSS | 0,304 | 0,002 | 0,403 | 0,304 | 0,347 | 0,348 | 0,988 | 0,406 |
| Web Attack Sql Injection | 0,385 | 0,000 | 1,000 | 0,385 | 0,556 | 0,620 | 0,961 | 0,739 |
| Weighted Avg. | 0,995 | 0,028 | 0,995 | 0,995 | 0,995 | 0,981 | 1,000 | 0,996 |

Tabel 4.21 merupakan hasil klasifikasi dengan menggunakan algoritma *Random Forest* pada mode pengujian *5 Cross Validation*, menunjukkan hasil 101.754 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5441% dan 466 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,4559% , dengan waktu proses klasifikasi selama 147,78 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 161 *point* b.

Tabel 4.22 Confusion Matrix RF 5-CV Split Data 60:40 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|-----|---|------------------------------|--------|
| a | b | c | d | | |
| 100912 | 0 | 0 | 0 | a = Benign | Aktual |
| 10 | 718 | 176 | 0 | b = Web Attack Brute Force | |
| 22 | 250 | 119 | 0 | c = Web Attack XSS | |
| 5 | 3 | 0 | 5 | d = Web Attack Sql Injection | |

Tabel 4.22 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 100.912 data diprediksi benar kategori *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 10 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 718 data benar diprediksi sebagai *web attack brute force*, dan 176 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 22 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 250 data salah dikategorikan sebagai *web attack brute force* seharusnya *web*

attack xss, dan 119 data benar diprediksi sebagai *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 5 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 3 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 5 data diprediksi benar kategori *web attack sql injection*.

c. *10 Cross Validation*

Tabel 4.23 Hasil Klasifikasi RF 10-CV Split Data 60:40 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,026 | 1,000 | 1,000 | 1,000 | 0,986 | 1,000 | 1,000 |
| Web Attack Brute Force | 0,802 | 0,003 | 0,738 | 0,802 | 0,769 | 0,767 | 0,998 | 0,825 |
| Web Attack XSS | 0,297 | 0,002 | 0,404 | 0,297 | 0,342 | 0,344 | 0,985 | 0,405 |
| Web Attack Sql Injection | 0,462 | 0,000 | 1,000 | 0,462 | 0,632 | 0,679 | 0,961 | 0,725 |
| Weighted Avg. | 0,995 | 0,026 | 0,995 | 0,995 | 0,995 | 0,982 | 1,000 | 0,996 |

Tabel 4.23 merupakan hasil klasifikasi dengan menggunakan algoritma *Random Forest* pada mode pengujian *10 Cross Validation*, menunjukkan hasil 101.758 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,548% dan 462 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,452% , dengan waktu proses klasifikasi selama 148 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 162 *point c*.

Tabel 4.24 Confusion Matrix RF 10-CV Split Data 60:40 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|-----|---|------------------------------|--------|
| a | b | c | d | | |
| 100911 | 0 | 1 | 0 | a = Benign | Aktual |
| 9 | 725 | 170 | 0 | b = Web Attack Brute Force | |
| 21 | 254 | 116 | 0 | c = Web Attack XSS | |
| 4 | 3 | 0 | 6 | d = Web Attack Sql Injection | |

Tabel 4.24 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 100.911 data diprediksi benar kategori *benign* dan 1 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 9 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 725 data benar diprediksi sebagai *web attack brute force* dan 170 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 21 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 254 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss* dan 116 data benar diprediksi sebagai *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 4 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 3 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 6 data diprediksi benar kategori *web attack sql injection*.

4.2.1.2 Klasifikasi Pada Split Data 70 : 30 Tanpa Seleksi Fitur

Perbandingan split data yang digunakan yaitu 70:30, 70% dari 170.366 *record* data atau sebanyak 119.256 *record* data digunakan sebagai data *training* dan 30% atau 51.110 *record* data sebagai *data testing* tanpa seleksi fitur atau menggunakan 77 fitur awal.

1. Naive Bayes

a. Use Training Set

Tabel 4.25 Hasil Klasifikasi UTS Split Data 70:30 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,979 | 0,010 | 1,000 | 0,979 | 0,989 | 0,608 | 0,996 | 1,000 |
| Web Attack Brute Force | 0,065 | 0,014 | 0,041 | 0,065 | 0,050 | 0,041 | 0,981 | 0,182 |
| Web Attack XSS | 0,917 | 0,012 | 0,229 | 0,917 | 0,367 | 0,455 | 0,989 | 0,239 |
| Web Attack Sql Injection | 1,000 | 0,004 | 0,034 | 1,000 | 0,065 | 0,183 | 0,999 | 0,043 |
| Weighted Avg. | 0,971 | 0,011 | 0,988 | 0,971 | 0,979 | 0,602 | 0,996 | 0,990 |

Tabel 4.25 merupakan hasil klasifikasi dengan menggunakan algoritma *Naive Bayes* pada mode pengujian *Use Training Set*, menunjukkan hasil 115.786 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 97,0903% dan 3.470 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 2,9097% , dengan waktu proses klasifikasi selama 26,52 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 162 *point* a.

Tabel 4.26 Confusion Matrix NB UTS Split Data 70:30 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|-----|-----|------------------------------|--------|
| a | b | c | d | | |
| 115284 | 1604 | 526 | 316 | a = Benign | Aktual |
| 2 | 69 | 879 | 105 | b = Web Attack Brute Force | |
| 14 | 16 | 418 | 8 | c = Web Attack XSS | |
| 0 | 0 | 0 | 15 | d = Web Attack Sql Injection | |

Tabel 4.26 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 115.284 data diprediksi benar kategori *benign*, 1.604 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 526 data salah dikategorikan *web attack xss* seharusnya *benign* dan 316 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 2 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 69 data benar diprediksi sebagai *web attack brute force*, 879 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 105 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 14 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 16 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 418 data benar diprediksi sebagai *web attack xss* dan 8 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 15 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.27 Hasil Klasifikasi NB 5-CV Split Data 70:30 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,978 | 0,010 | 1,000 | 0,978 | 0,989 | 0,601 | 0,994 | 1,000 |
| Web Attack Brute Force | 0,069 | 0,014 | 0,041 | 0,069 | 0,052 | 0,042 | 0,980 | 0,178 |
| Web Attack XSS | 0,917 | 0,012 | 0,233 | 0,917 | 0,372 | 0,459 | 0,986 | 0,213 |
| Web Attack Sql Injection | 1,000 | 0,004 | 0,031 | 1,000 | 0,061 | 0,176 | 0,998 | 0,034 |
| Weighted Avg. | 0,970 | 0,011 | 0,988 | 0,970 | 0,978 | 0,596 | 0,994 | 0,989 |

Tabel 4.27 merupakan hasil klasifikasi dengan menggunakan algoritma *Naive Bayes* pada mode pengujian *5 Cross Validation*, menunjukkan hasil 115.703 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 97,0207% dan 3.553 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 2,9793% , dengan waktu proses klasifikasi selama 3,12 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 163 *point* b.

Tabel 4.28 Confusion Matrix NB 5-CV Split Data 70:30 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|-----|-----|------------------------------|--------|
| a | b | c | d | | |
| 115197 | 1679 | 501 | 353 | a = Benign | Aktual |
| 2 | 73 | 875 | 105 | b = Web Attack Brute Force | |
| 14 | 17 | 418 | 7 | c = Web Attack XSS | |
| 0 | 0 | 0 | 15 | d = Web Attack Sql Injection | |

Tabel 4.28 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class*

“Benign” terdapat sebanyak 115.197 data diprediksi benar kategori *benign*, 1.679 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 501 data salah dikategorikan *web attack xss* seharusnya *benign* dan 353 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada class “*Web Attack Brute Force*” terdapat sebanyak 2 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 73 data benar diprediksi sebagai *web attack brute force*, 875 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 105 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada class “*Web Attack XSS*” terdapat sebanyak 14 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 17 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 418 data benar diprediksi sebagai *web attack xss* dan 7 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 15 data diprediksi benar kategori *web attack sql injection*.

c. 10 Cross Validation

Tabel 4.29 Hasil Klasifikasi NB 10-CV Split Data 70:30 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,979 | 0,010 | 1,000 | 0,979 | 0,989 | 0,605 | 0,994 | 1,000 |
| Web Attack Brute Force | 0,118 | 0,014 | 0,069 | 0,118 | 0,087 | 0,080 | 0,980 | 0,179 |
| Web Attack XSS | 0,888 | 0,011 | 0,237 | 0,888 | 0,374 | 0,455 | 0,985 | 0,201 |
| Web Attack Sql Injection | 1,000 | 0,004 | 0,033 | 1,000 | 0,063 | 0,180 | 0,998 | 0,035 |
| Weighted Avg. | 0,971 | 0,011 | 0,989 | 0,971 | 0,979 | 0,600 | 0,994 | 0,989 |

Tabel 4.29 merupakan hasil klasifikasi dengan menggunakan algoritma *Naive Bayes* pada mode pengujian *10 Cross Validation*, menunjukkan hasil 115.796 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 97,0987% dan 3.460 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 2,9013% , dengan waktu proses klasifikasi selama 3,62 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 163 *point c*.

Tabel 4.30 Confusion Matrix NB 10-CV Split Data 70:30 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|-----|-----|------------------------------|--------|
| a | b | c | d | | |
| 115251 | 1664 | 483 | 332 | a = Benign | Aktual |
| 2 | 125 | 823 | 105 | b = Web Attack Brute Force | |
| 14 | 29 | 405 | 8 | c = Web Attack XSS | |
| 0 | 0 | 0 | 15 | d = Web Attack Sql Injection | |

Tabel 4.30 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “*Benign*” terdapat sebanyak 115.251 data diprediksi benar kategori *benign*, 1.664 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 483 data salah dikategorikan *web attack xss* seharusnya *benign* dan 332 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “*Web Attack Brute Force*” terdapat sebanyak 2 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 125 data benar diprediksi sebagai *web attack brute force*, 823 data salah dikategorikan *web attack xss* seharusnya *web*

attack brute force dan 105 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada class “*Web Attack XSS*” terdapat sebanyak 14 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 29 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 405 data benar diprediksi sebagai *web attack xss* dan 8 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 15 data diprediksi benar kategori *web attack sql injection*.

2. C4.5

a. Use Training Set

Tabel 4.31 Hasil Klasifikasi C45 UTS Split Data 70:30 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,004 | 1,000 | 1,000 | 1,000 | 0,997 | 0,999 | 1,000 |
| Web Attack Brute Force | 0,984 | 0,003 | 0,735 | 0,984 | 0,842 | 0,849 | 0,999 | 0,839 |
| Web Attack XSS | 0,178 | 0,000 | 0,827 | 0,178 | 0,292 | 0,382 | 0,998 | 0,514 |
| Web Attack Sql Injection | 0,800 | 0,000 | 0,857 | 0,800 | 0,828 | 0,828 | 0,944 | 0,753 |
| Weighted Avg. | 0,997 | 0,004 | 0,997 | 0,997 | 0,996 | 0,993 | 0,999 | 0,997 |

Tabel 4.31 merupakan hasil klasifikasi dengan menggunakan algoritma *C4.5* pada mode pengujian *Use Training Set*, menunjukkan hasil 118.857 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,6654% dan 399 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,3346% , dengan waktu proses klasifikasi selama 36,4 detik. Untuk *screen shot* hasil

pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 164 *point* a.

Tabel 4.32 Confusion Matrix C45 UTS Split Data 70:30 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|----|----|------------------------------|--------|
| a | b | c | d | | |
| 117726 | 2 | 2 | 0 | a = Benign | Aktual |
| 2 | 1038 | 15 | 0 | b = Web Attack Brute Force | |
| 1 | 372 | 81 | 2 | c = Web Attack XSS | |
| 3 | 0 | 0 | 12 | d = Web Attack Sql Injection | |

Tabel 4.32 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 117.726 data diprediksi benar kategori *benign*, 2 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, dan 2 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 2 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 1.038 data benar diprediksi sebagai *web attack brute force*, dan 15 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 1 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 372 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 81 data benar diprediksi sebagai *web attack xss* dan 2 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 3 data salah dikategorikan sebagai *benign*

seharusnya *web attack sql injection* dan 12 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.33 Hasil Klasifikasi C45 5-CV Split Data 70:30 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,010 | 1,000 | 1,000 | 1,000 | 0,986 | 0,993 | 1,000 |
| Web Attack Brute Force | 0,972 | 0,004 | 0,706 | 0,972 | 0,818 | 0,826 | 0,989 | 0,782 |
| Web Attack XSS | 0,081 | 0,000 | 0,507 | 0,081 | 0,140 | 0,202 | 0,984 | 0,379 |
| Web Attack Sql Injection | 0,400 | 0,000 | 0,500 | 0,400 | 0,444 | 0,447 | 0,790 | 0,429 |
| Weighted Avg. | 0,996 | 0,010 | 0,995 | 0,996 | 0,995 | 0,982 | 0,993 | 0,995 |

Tabel 4.33 merupakan hasil klasifikasi dengan menggunakan algoritma *C4.5* pada mode pengujian *5 Cross Validation*, menunjukkan hasil 118.772 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5942% dan 484 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,4058% , dengan waktu proses klasifikasi selama 33,08 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 164 *point* b.

Tabel 4.34 Confusion Matrix C45 5-CV Split Data 70:30 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|----|---|------------------------------|--------|
| a | b | c | d | | |
| 117704 | 15 | 6 | 5 | a = Benign | Aktual |
| 5 | 1025 | 25 | 0 | b = Web Attack Brute Force | |
| 6 | 412 | 37 | 1 | c = Web Attack XSS | |
| 4 | 0 | 5 | 6 | d = Web Attack Sql Injection | |

Tabel 4.34 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 117.704 data diprediksi benar kategori *benign*, 15 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 6 data salah dikategorikan *web attack xss* seharusnya *benign* dan 5 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 5 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 1.025 data benar diprediksi sebagai *web attack brute force*, dan 25 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 6 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 412 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 37 data benar diprediksi sebagai *web attack xss* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 4 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 5 data salah dikategorikan sebagai *web attack*

xss seharusnya *web attack sql injection* dan 6 data diprediksi benar kategori *web attack sql injection*.

c. *10 Cross Validation*

Tabel 4.35 Hasil Klasifikasi C45 10-CV Split Data 70:30 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,012 | 1,000 | 1,000 | 1,000 | 0,987 | 0,995 | 1,000 |
| Web Attack Brute Force | 0,965 | 0,004 | 0,710 | 0,965 | 0,818 | 0,826 | 0,990 | 0,816 |
| Web Attack XSS | 0,094 | 0,000 | 0,544 | 0,094 | 0,161 | 0,225 | 0,984 | 0,415 |
| Web Attack Sql Injection | 0,733 | 0,000 | 0,688 | 0,733 | 0,710 | 0,710 | 0,779 | 0,575 |
| Weighted Avg. | 0,996 | 0,012 | 0,996 | 0,996 | 0,995 | 0,983 | 0,995 | 0,996 |

Tabel 4.35 merupakan hasil klasifikasi dengan menggunakan algoritma *C4.5* pada mode pengujian *10 Cross Validation*, menunjukkan hasil 118.781 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,6017% dan 475 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,3983% , dengan waktu proses klasifikasi selama 35,27 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 165 *point c*.

Tabel 4.36 Confusion Matrix C45 10-CV Split Data 70:30 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|----|----|------------------------------|--------|
| a | b | c | d | | |
| 117709 | 13 | 5 | 3 | a = Benign | Aktual |
| 6 | 1018 | 31 | 0 | b = Web Attack Brute Force | |
| 8 | 403 | 43 | 2 | c = Web Attack XSS | |
| 4 | 0 | 0 | 11 | d = Web Attack Sql Injection | |

Tabel 4.36 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 117.709 data diprediksi benar kategori *benign*, 13 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 5 data salah dikategorikan *web attack xss* seharusnya *benign* dan 3 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 6 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 1.018 data benar diprediksi sebagai *web attack brute force* dan 31 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 8 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 403 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 43 data benar diprediksi sebagai *web attack xss* dan 2 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 4 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, dan 11 data diprediksi benar kategori *web attack sql injection*.

3. Random Forest

a. Use Training Set

Tabel 4.37 Hasil Klasifikasi RF UTS Split Data 70:30 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| Web Attack Brute Force | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| Web Attack XSS | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| Web Attack Sql Injection | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| Weighted Avg. | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |

Tabel 4.37 merupakan hasil klasifikasi dengan menggunakan algoritma *Random Forest* pada mode pengujian *Use Training Set*, menunjukkan hasil 119.256 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 100% dan 0 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0% , dengan waktu proses klasifikasi selama 120,15 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 165 *point a*.

Tabel 4.38 Confusion Matrix RF UTS Split Data 70:30 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|-----|----|------------------------------|---------------|
| a | b | c | d | | |
| 117730 | 0 | 0 | 0 | a = Benign | Aktual |
| 0 | 1055 | 0 | 0 | b = Web Attack Brute Force | |
| 0 | 0 | 456 | 0 | c = Web Attack XSS | |
| 0 | 0 | 0 | 15 | d = Web Attack Sql Injection | |

Tabel 4.38 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 117.730 data diprediksi benar kategori *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 1.055 data benar diprediksi sebagai *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 456 data benar diprediksi sebagai *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 15 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.39 Hasil Klasifikasi RF 5-CV Split Data 70:30 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,026 | 1,000 | 1,000 | 1,000 | 0,986 | 1,000 | 1,000 |
| Web Attack Brute Force | 0,813 | 0,003 | 0,736 | 0,813 | 0,773 | 0,772 | 0,999 | 0,805 |
| Web Attack XSS | 0,289 | 0,002 | 0,415 | 0,289 | 0,341 | 0,345 | 0,977 | 0,388 |
| Web Attack Sql Injection | 0,400 | 0,000 | 1,000 | 0,400 | 0,571 | 0,632 | 0,967 | 0,783 |
| Weighted Avg. | 0,996 | 0,025 | 0,995 | 0,996 | 0,995 | 0,982 | 1,000 | 0,996 |

Tabel 4.39 merupakan hasil klasifikasi dengan menggunakan algoritma *Random Forest* pada mode pengujian *5 Cross Validation*, menunjukkan hasil 118.724 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5539% dan 532 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,4461% , dengan waktu proses klasifikasi selama 108,63 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 166 *point* b.

Tabel 4.40 Confusion Matrix RF 5-CV Split Data 70:30 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|-----|---|------------------------------|--------|
| a | b | c | d | | |
| 117728 | 1 | 1 | 0 | a = Benign | Aktual |
| 12 | 858 | 185 | 0 | b = Web Attack Brute Force | |
| 22 | 302 | 132 | 0 | c = Web Attack XSS | |
| 5 | 4 | 0 | 6 | d = Web Attack Sql Injection | |

Tabel 4.40 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 117.728 data diprediksi benar kategori *benign*, 1 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign* dan 1 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 12 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 858 data benar diprediksi sebagai *web attack brute force*, dan 185 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 22 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 302 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 132 data benar diprediksi sebagai *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 5 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 4 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 6 data diprediksi benar kategori *web attack sql injection*.

c. 10 Cross Validation

Tabel 4.41 Hasil Klasifikasi RF 10-CV Split Data 70:30 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,025 | 1,000 | 1,000 | 1,000 | 0,987 | 1,000 | 1,000 |
| Web Attack Brute Force | 0,808 | 0,003 | 0,734 | 0,808 | 0,769 | 0,768 | 0,999 | 0,827 |
| Web Attack XSS | 0,287 | 0,002 | 0,406 | 0,287 | 0,336 | 0,339 | 0,988 | 0,401 |
| Web Attack Sql Injection | 0,400 | 0,000 | 1,000 | 0,400 | 0,571 | 0,632 | 0,967 | 0,800 |
| Weighted Avg. | 0,995 | 0,025 | 0,995 | 0,995 | 0,995 | 0,983 | 1,000 | 0,996 |

Tabel 4.41 merupakan hasil klasifikasi dengan menggunakan algoritma *Random Forest* pada mode pengujian *10 Cross Validation*, menunjukkan hasil 118.718 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5489% dan 538 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,4511% , dengan waktu proses klasifikasi selama 99,46 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 166 *point c*.

Tabel 4.42 Confusion Matrix RF 10-CV Split Data 70:30 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|-----|---|------------------------------|---------------|
| a | b | c | d | | |
| 117729 | 0 | 1 | 0 | a = Benign | Aktual |
| 12 | 852 | 191 | 0 | b = Web Attack Brute Force | |
| 22 | 303 | 131 | 0 | c = Web Attack XSS | |
| 4 | 5 | 0 | 6 | d = Web Attack Sql Injection | |

Tabel 4.42 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “*Benign*” terdapat sebanyak 117.729 data diprediksi benar kategori *benign*, dan 1 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “*Web Attack Brute Force*” terdapat sebanyak 12 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 852 data benar diprediksi sebagai *web attack brute force*, dan 191 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “*Web Attack XSS*” terdapat sebanyak 22 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 303 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 131 data benar diprediksi sebagai *web attack xss*. Pada *class* “*Web Attack Sql Injection*” terdapat sebanyak 4 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 5 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 6 data diprediksi benar kategori *web attack sql injection*.

4.2.1.3 Klasifikasi Pada Split Data 80 : 20 Tanpa Seleksi Fitur

Perbandingan split data yang digunakan yaitu 80:20, 80% dari 170.366 *record* data atau sebanyak 136.294 *record* data digunakan sebagai data *training* dan 20% atau 34.072 *record* data sebagai *data testing*, tanpa seleksi fitur atau menggunakan 77 fitur awal.

1. Naive Bayes

a. Use Training Set

Tabel 4.43 Hasil Klasifikasi NB UTS Split Data 80:20 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,981 | 0,011 | 1,000 | 0,981 | 0,990 | 0,621 | 0,994 | 1,000 |
| Web Attack Brute Force | 0,099 | 0,013 | 0,063 | 0,099 | 0,077 | 0,069 | 0,982 | 0,188 |
| Web Attack XSS | 0,916 | 0,011 | 0,238 | 0,916 | 0,378 | 0,464 | 0,988 | 0,266 |
| Web Attack Sql Injection | 1,000 | 0,003 | 0,039 | 1,000 | 0,076 | 0,198 | 0,999 | 0,082 |
| Weighted Avg. | 0,973 | 0,011 | 0,989 | 0,973 | 0,980 | 0,616 | 0,994 | 0,990 |

Tabel 4.43 merupakan hasil klasifikasi dengan menggunakan algoritma *Naive Bayes* pada mode pengujian *Use Training Set*, menunjukkan hasil 132.558 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 97,2589% dan 3.736 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 2,7411% ,dengan waktu proses klasifikasi selama 45,98 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 167 *point a*.

Tabel 4.44 Confusion Matrix NB UTS Split Data 80:20 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|-----|-----|------------------------------|--------|
| a | b | c | d | | |
| 131944 | 1751 | 568 | 286 | a = Benign | Aktual |
| 2 | 119 | 963 | 122 | b = Web Attack Brute Force | |
| 18 | 18 | 478 | 8 | c = Web Attack XSS | |
| 0 | 0 | 0 | 17 | d = Web Attack Sql Injection | |

Tabel 4.44 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 131.944 data diprediksi benar kategori *benign*, 1.751

data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 568 data salah dikategorikan *web attack xss* seharusnya *benign* dan 286 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada class “*Web Attack Brute Force*” terdapat sebanyak 2 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 119 data benar diprediksi sebagai *web attack brute force*, 963 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 122 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada class “*Web Attack XSS*” terdapat sebanyak 18 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 18 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 478 data benar diprediksi sebagai *web attack xss* dan 8 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 17 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.45 Hasil Klasifikasi NB 5-CV Split Data 80:20 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,979 | 0,010 | 1,000 | 0,979 | 0,989 | 0,608 | 0,996 | 1,000 |
| Web Attack Brute Force | 0,082 | 0,014 | 0,050 | 0,082 | 0,062 | 0,054 | 0,981 | 0,179 |
| Web Attack XSS | 0,920 | 0,012 | 0,234 | 0,920 | 0,374 | 0,461 | 0,987 | 0,217 |
| Web Attack Sql Injection | 1,000 | 0,004 | 0,034 | 1,000 | 0,065 | 0,183 | 0,998 | 0,034 |
| Weighted Avg. | 0,971 | 0,010 | 0,988 | 0,971 | 0,979 | 0,603 | 0,996 | 0,990 |

Tabel 4.45 merupakan hasil klasifikasi dengan menggunakan algoritma *Naive Bayes* pada mode pengujian *5 Cross Validation*, menunjukkan hasil 132.349 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 97,1055% dan 3.945 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 2,8945% , dengan waktu proses klasifikasi selama 12,31 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 167 *point* b.

Tabel 4.46 Confusion Matrix NB 5-CV Split Data 80:20 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|-----|-----|------------------------------|---------------|
| a | b | c | d | | |
| 131753 | 1855 | 579 | 362 | a = Benign | Aktual |
| 2 | 99 | 988 | 117 | b = Web Attack Brute Force | |
| 15 | 17 | 480 | 10 | c = Web Attack XSS | |
| 0 | 0 | 0 | 17 | d = Web Attack Sql Injection | |

Tabel 4.46 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “*Benign*” terdapat sebanyak 131.753 data diprediksi benar kategori *benign*, 1.855 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 579 data salah dikategorikan *web attack xss* seharusnya *benign* dan 362 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “*Web Attack Brute Force*” terdapat sebanyak 2 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 99 data benar diprediksi sebagai *web attack brute force*, 988 data salah dikategorikan *web attack xss* seharusnya *web*

attack brute force dan 117 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada class “*Web Attack XSS*” terdapat sebanyak 15 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 17 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 480 data benar diprediksi sebagai *web attack xss* dan 10 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 17 data diprediksi benar kategori *web attack sql injection*.

c. *10 Cross Validation*

Tabel 4.47 Hasil Klasifikasi NB 10-CV Split Data 80:20 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,980 | 0,009 | 1,000 | 0,980 | 0,990 | 0,614 | 0,996 | 1,000 |
| Web Attack Brute Force | 0,076 | 0,014 | 0,048 | 0,076 | 0,059 | 0,050 | 0,981 | 0,181 |
| Web Attack XSS | 0,920 | 0,012 | 0,234 | 0,920 | 0,373 | 0,460 | 0,986 | 0,216 |
| Web Attack Sql Injection | 1,000 | 0,003 | 0,035 | 1,000 | 0,068 | 0,187 | 0,999 | 0,043 |
| Weighted Avg. | 0,972 | 0,009 | 0,988 | 0,972 | 0,979 | 0,608 | 0,995 | 0,990 |

Tabel 4.47 merupakan hasil klasifikasi dengan menggunakan algoritma *Naive Bayes* pada mode pengujian *10 Cross Validation*, menunjukkan hasil 132.412 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 97,1517% dan 3.882 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 2,8483% , dengan waktu proses klasifikasi selama 12,07 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 168 *point c*.

Tabel 4.48 Confusion Matrix NB 10-CV Split Data 80:20 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|-----|-----|------------------------------|--------|
| a | b | c | d | | |
| 131823 | 1811 | 583 | 332 | a = Benign | Aktual |
| 2 | 92 | 990 | 122 | b = Web Attack Brute Force | |
| 14 | 17 | 480 | 11 | c = Web Attack XSS | |
| 0 | 0 | 0 | 17 | d = Web Attack Sql Injection | |

Tabel 4.48 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 131.823 data diprediksi benar kategori *benign*, 1.811 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 583 data salah dikategorikan *web attack xss* seharusnya *benign* dan 332 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 2 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 92 data benar diprediksi sebagai *web attack brute force*, 990 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 122 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 14 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 17 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 480 data benar diprediksi sebagai *web attack xss* dan 11 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 17 data diprediksi benar kategori *web attack sql injection*.

2. C4.5

a. Use Training Set

Tabel 4.49 Hasil Klasifikasi C45 UTS Split Data 80:20 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,003 | 1,000 | 1,000 | 1,000 | 0,998 | 1,000 | 1,000 |
| Web Attack Brute Force | 0,995 | 0,003 | 0,729 | 0,995 | 0,842 | 0,851 | 0,999 | 0,839 |
| Web Attack XSS | 0,148 | 0,000 | 0,928 | 0,148 | 0,255 | 0,369 | 0,998 | 0,506 |
| Web Attack Sql Injection | 0,824 | 0,000 | 0,933 | 0,824 | 0,875 | 0,877 | 0,950 | 0,797 |
| Weighted Avg. | 0,997 | 0,003 | 0,997 | 0,997 | 0,996 | 0,994 | 0,999 | 0,997 |

Tabel 4.49 merupakan hasil klasifikasi dengan menggunakan algoritma C4.5 pada mode pengujian *Use Training Set*, menunjukkan hasil 135.837 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,6647% dan 457 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,3353% , dengan waktu proses klasifikasi selama 75,48 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 168 *point* a.

Tabel 4.50 Confusion Matrix C45 UTS Split Data 80:20 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|----|----|------------------------------|---------------|
| a | b | c | d | | |
| 134546 | 2 | 1 | 0 | a = Benign | Aktual |
| 1 | 1200 | 5 | 0 | b = Web Attack Brute Force | |
| 1 | 443 | 77 | 1 | c = Web Attack XSS | |
| 3 | 0 | 0 | 14 | d = Web Attack Sql Injection | |

Tabel 4.50 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “*Benign*” terdapat sebanyak 134.546 data diprediksi benar kategori *benign*, 2 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign* dan 1 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “*Web Attack Brute Force*” terdapat sebanyak 1 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 1200 data benar diprediksi sebagai *web attack brute force*, dan 5 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “*Web Attack XSS*” terdapat sebanyak 1 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 443 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 77 data benar diprediksi sebagai *web attack xss* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada *class* “*Web Attack Sql Injection*” terdapat sebanyak 3 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection* dan 14 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.51 Hasil Klasifikasi C45 5-CV Split Data 80:20 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,014 | 1,000 | 1,000 | 1,000 | 0,990 | 0,996 | 1,000 |
| Web Attack Brute Force | 0,978 | 0,004 | 0,714 | 0,978 | 0,825 | 0,834 | 0,992 | 0,817 |
| Web Attack XSS | 0,084 | 0,000 | 0,688 | 0,084 | 0,150 | 0,240 | 0,982 | 0,424 |
| Web Attack Sql Injection | 0,706 | 0,000 | 0,857 | 0,706 | 0,774 | 0,778 | 0,815 | 0,635 |
| Weighted Avg. | 0,996 | 0,014 | 0,996 | 0,996 | 0,995 | 0,986 | 0,996 | 0,996 |

Tabel 4.51 merupakan hasil klasifikasi dengan menggunakan algoritma C4.5 pada mode pengujian 5 *Cross Validation*, menunjukkan hasil 135.775 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,6192% dan 519 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,3808% , dengan waktu proses klasifikasi selama 70,71 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 169 *point* b.

Tabel 4.52 Confusion Matrix C45 5-CV Split Data 80:20 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|----|----|------------------------------|---------------|
| a | b | c | d | | |
| 134539 | 7 | 3 | 0 | a = Benign | Aktual |
| 9 | 1180 | 17 | 0 | b = Web Attack Brute Force | |
| 10 | 466 | 44 | 2 | c = Web Attack XSS | |
| 5 | 0 | 0 | 12 | d = Web Attack Sql Injection | |

Tabel 4.52 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 134.539 data diprediksi benar kategori *benign*, 7 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign* dan 3 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 9 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 1.180 data benar diprediksi sebagai *web attack brute force* dan 17 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 10 data salah

dikategorikan sebagai *benign* seharusnya *web attack xss*, 466 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 44 data benar diprediksi sebagai *web attack xss* dan 2 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 5 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection* dan 12 data diprediksi benar kategori *web attack sql injection*.

c. *10 Cross Validation*

Tabel 4.53 Hasil Klasifikasi C45 10-CV Split Data 80:20 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,013 | 1,000 | 1,000 | 1,000 | 0,991 | 0,996 | 1,000 |
| Web Attack Brute Force | 0,980 | 0,004 | 0,714 | 0,980 | 0,826 | 0,835 | 0,993 | 0,833 |
| Web Attack XSS | 0,082 | 0,000 | 0,717 | 0,082 | 0,148 | 0,242 | 0,988 | 0,441 |
| Web Attack Sql Injection | 0,765 | 0,000 | 0,867 | 0,765 | 0,813 | 0,814 | 0,805 | 0,645 |
| Weighted Avg. | 0,996 | 0,013 | 0,996 | 0,996 | 0,995 | 0,986 | 0,996 | 0,996 |

Tabel 4.53 merupakan hasil klasifikasi dengan menggunakan algoritma C4.5 pada mode pengujian *10 Cross Validation*, menunjukkan hasil 135.778 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,6214% dan 516 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,3786% , dengan waktu proses klasifikasi selama 70,73 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 169 *point c*.

Tabel 4.54 Confusion Matrix C45 10-CV Split Data 80:20 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|----|----|------------------------------|---------------|
| a | b | c | d | | |
| 134540 | 6 | 3 | 0 | a = Benign | Aktual |
| 10 | 1182 | 14 | 0 | b = Web Attack Brute Force | |
| 9 | 468 | 43 | 1 | c = Web Attack XSS | |
| 4 | 0 | 0 | 13 | d = Web Attack Sql Injection | |

Tabel 4.54 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 134.540 data diprediksi benar kategori *benign*, 6 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign* dan 3 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 10 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 1.182 data benar diprediksi sebagai *web attack brute force*, dan 14 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 9 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 468 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 43 data benar diprediksi sebagai *web attack xss* dan 2 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 4 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, dan 13 data diprediksi benar kategori *web attack sql injection*.

3. Random Forest

a. Use Training Set

Tabel 4.55 Hasil Klasifikasi RF UTS Split Data 80:20 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| Web Attack Brute Force | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| Web Attack XSS | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| Web Attack Sql Injection | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| Weighted Avg. | 1,000 | 0,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |

Tabel 4.55 merupakan hasil klasifikasi dengan menggunakan algoritma *Random Forest* pada mode pengujian *Use Training Set*, menunjukkan hasil 136.294 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 100% dan 0 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0% , dengan waktu proses klasifikasi selama 215,44 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 170 *point a*.

Tabel 4.56 Confusion Matrix RF UTS Split Data 80:20 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|------|-----|----|------------------------------|--------|
| a | b | c | d | | |
| 134549 | 0 | 0 | 0 | a = Benign | Aktual |
| 0 | 1206 | 0 | 0 | b = Web Attack Brute Force | |
| 0 | 0 | 522 | 0 | c = Web Attack XSS | |
| 0 | 0 | 0 | 17 | d = Web Attack Sql Injection | |

Tabel 4.56 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 134.549 data diprediksi benar kategori *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 1.206 data benar diprediksi sebagai *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 522 data benar diprediksi sebagai *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 17 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.57 Hasil Klasifikasi RF 5-CV Split Data 80:20 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,024 | 1,000 | 1,000 | 1,000 | 0,987 | 1,000 | 1,000 |
| Web Attack Brute Force | 0,802 | 0,003 | 0,735 | 0,802 | 0,767 | 0,765 | 0,998 | 0,815 |
| Web Attack XSS | 0,295 | 0,002 | 0,403 | 0,295 | 0,341 | 0,343 | 0,983 | 0,395 |
| Web Attack Sql Injection | 0,412 | 0,000 | 1,000 | 0,412 | 0,583 | 0,642 | 0,970 | 0,802 |
| Weighted Avg. | 0,995 | 0,024 | 0,995 | 0,995 | 0,995 | 0,983 | 1,000 | 0,996 |

Tabel 4.57 merupakan hasil klasifikasi dengan menggunakan algoritma *Random Forest* pada mode pengujian *5 Cross Validation*, menunjukkan hasil 135.675 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5458% dan 619 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,4542% , dengan waktu proses klasifikasi selama 202,3 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 170 *point* b.

Tabel 4.58 Confusion Matrix RF 5-CV Split Data 80:20 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|-----|---|------------------------------|---------------|
| a | b | c | d | | |
| 134547 | 0 | 2 | 0 | a = Benign | Aktual |
| 13 | 967 | 226 | 0 | b = Web Attack Brute Force | |
| 25 | 343 | 154 | 0 | c = Web Attack XSS | |
| 4 | 6 | 0 | 7 | d = Web Attack Sql Injection | |

Tabel 4.58 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 134.547 data diprediksi benar kategori *benign*, dan 2 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 13 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 967 data benar diprediksi sebagai *web attack brute force*, dan 226 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 25 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 343 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 154 data benar diprediksi sebagai *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 4 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 6 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 7 data diprediksi benar kategori *web attack sql injection*.

c. 10 Cross Validation

Tabel 4.59 Hasil Klasifikasi RF 10-CV Split Data 80:20 Tanpa Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,021 | 1,000 | 1,000 | 1,000 | 0,988 | 1,000 | 1,000 |
| Web Attack Brute Force | 0,794 | 0,003 | 0,731 | 0,794 | 0,761 | 0,760 | 0,999 | 0,816 |
| Web Attack XSS | 0,297 | 0,002 | 0,393 | 0,297 | 0,338 | 0,340 | 0,983 | 0,389 |
| Web Attack Sql Injection | 0,353 | 0,000 | 1,000 | 0,353 | 0,522 | 0,594 | 0,970 | 0,787 |
| Weighted Avg. | 0,995 | 0,021 | 0,995 | 0,995 | 0,995 | 0,984 | 1,000 | 0,996 |

Tabel 4.59 merupakan hasil klasifikasi dengan menggunakan algoritma *Random Forest* pada mode pengujian *10 Cross Validation*, menunjukkan hasil 135.665 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5385% dan 629 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,4615% , dengan waktu proses klasifikasi selama 199,34 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 171 *point c*.

Tabel 4.60 Confusion Matrix RF 10-CV Split Data 80:20 Tanpa Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|-----|---|------------------------------|--------|
| a | b | c | d | | |
| 134546 | 0 | 3 | 0 | a = Benign | Aktual |
| 12 | 958 | 236 | 0 | b = Web Attack Brute Force | |
| 20 | 347 | 155 | 0 | c = Web Attack XSS | |
| 5 | 6 | 0 | 6 | d = Web Attack Sql Injection | |

Tabel 4.60 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “*Benign*” terdapat sebanyak 134.546 data diprediksi benar kategori *benign*, dan 3 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “*Web Attack Brute Force*” terdapat sebanyak 12 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 958 data benar diprediksi sebagai *web attack brute force*, dan 236 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “*Web Attack XSS*” terdapat sebanyak 20 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 347 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 155 data benar diprediksi sebagai *web attack xss*. Pada *class* “*Web Attack Sql Injection*” terdapat sebanyak 5 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 6 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 6 data diprediksi benar kategori *web attack sql injection*.

4.2.2 KLASIFIKASI DENGAN SELEKSI FITUR

4.2.2.1 Klasifikasi Pada Split Data 60 : 40 Dengan Seleksi Fitur

Perbandingan split data yang digunakan yaitu 60:40, 60% dari 170.366 *record* data atau sebanyak 102.220 *record* data digunakan sebagai data *training* dan 40% atau 68.146 *record* data sebagai *data testing* dengan seleksi fitur atau menggunakan 11 fitur hasil seleksi fitur *information gain*.

1. Naive Bayes

a. Use Training Set

Tabel 4.61 Hasil Klasifikasi NB UTS Split Data 60:40 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,228 | 0,031 | 0,998 | 0,228 | 0,371 | 0,053 | 0,947 | 0,999 |
| Web Attack Brute Force | 0,085 | 0,699 | 0,001 | 0,085 | 0,002 | 0,125 | 0,291 | 0,006 |
| Web Attack XSS | 0,916 | 0,009 | 0,291 | 0,916 | 0,442 | 0,514 | 0,970 | 0,290 |
| Web Attack Sql Injection | 1,000 | 0,068 | 0,002 | 1,000 | 0,004 | 0,042 | 0,960 | 0,002 |
| Weighted Avg. | 0,230 | 0,036 | 0,987 | 0,230 | 0,368 | 0,053 | 0,941 | 0,987 |

Pada tabel 4.61 merupakan hasil klasifikasi algoritma *Naive Bayes* pada *tools* Weka menggunakan mode pengujian *Use Training Set*, menunjukkan hasil 23.472 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 22,9622% dan 78.748 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 77,0378%, dengan waktu proses klasifikasi selama 5,64 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 171 *point* a.

Tabel 4.62 Confusion Matrix NB UTS Split Data 60:40 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-------|-----|------|------------------------------|---------------|
| a | b | c | d | | |
| 23024 | 70823 | 146 | 6919 | a = Benign | Aktual |
| 37 | 77 | 726 | 64 | b = Web Attack Brute Force | |
| 3 | 24 | 358 | 6 | c = Web Attack XSS | |
| 0 | 0 | 0 | 13 | d = Web Attack Sql Injection | |

Tabel 4.62 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class*

“Benign” terdapat sebanyak 23.024 data diprediksi benar kategori *benign*, 70.823 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 146 data salah dikategorikan *web attack xss* seharusnya *benign* dan 6.919 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada class “*Web Attack Brute Force*” terdapat sebanyak 37 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 77 data benar diprediksi sebagai *web attack brute force*, 726 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 64 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada class “*Web Attack XSS*” terdapat sebanyak 3 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 24 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 358 data benar diprediksi sebagai *web attack xss* dan 6 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 13 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.63 Hasil Klasifikasi NB 5-CV Split Data 60:40 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,224 | 0,033 | 0,998 | 0,224 | 0,365 | 0,052 | 0,943 | 0,999 |
| Web Attack Brute Force | 0,08 | 0,701 | 0,001 | 0,080 | 0,002 | -0,126 | 0,284 | 0,006 |
| Web Attack XSS | 0,91 | 0,011 | 0,236 | 0,910 | 0,375 | 0,46 | 0,965 | 0,253 |
| Web Attack Sql Injection | 0,923 | 0,069 | 0,002 | 0,923 | 0,003 | 0,038 | 0,951 | 0,001 |
| Weighted Avg. | 0,225 | 0,039 | 0,986 | 0,225 | 0,362 | 0,052 | 0,938 | 0,987 |

Pada tabel 4.63 merupakan hasil klasifikasi algoritma *Naive Bayes* pada *tools* Weka menggunakan mode pengujian *5 Cross Validation*, menunjukkan hasil 22.999 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 22,4995% dan 79.221 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 77,5005%, dengan waktu proses klasifikasi selama 0,55 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 172 *point* b.

Tabel 4.64 Confusion Matrix NB 5-CV Split Data 60:40 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-------|-----|------|------------------------------|--------|
| a | b | c | d | | |
| 22559 | 70968 | 428 | 6957 | a = Benign | Aktual |
| 40 | 72 | 726 | 66 | b = Web Attack Brute Force | |
| 3 | 26 | 356 | 6 | c = Web Attack XSS | |
| 0 | 1 | 0 | 12 | d = Web Attack Sql Injection | |

Tabel 4.64 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 22.559 data diprediksi benar kategori *benign*, 70.968 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 428 data salah dikategorikan sebagai *web attack xss* seharusnya *benign* dan 6.957 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 40 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 72 data benar diprediksi sebagai *web attack brute force*, 726 data salah dikategorikan sebagai *web attack xss* seharusnya *web*

attack brute force dan 66 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada class “*Web Attack XSS*” terdapat sebanyak 3 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 26 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 356 data benar diprediksi sebagai *web attack xss* dan 6 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 1 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection*, dan 12 data diprediksi benar kategori *web attack sql injection*.

c. 10 Cross Validation

Tabel 4.65 Hasil Klasifikasi NB 10-CV Split Data 60:40 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,229 | 0,032 | 0,998 | 0,229 | 0,373 | 0,053 | 0,946 | 0,999 |
| Web Attack Brute Force | 0,083 | 0,700 | 0,001 | 0,083 | 0,002 | 0,125 | 0,287 | 0,006 |
| Web Attack XSS | 0,916 | 0,009 | 0,288 | 0,916 | 0,439 | 0,511 | 0,969 | 0,297 |
| Web Attack Sql Injection | 1,000 | 0,066 | 0,002 | 1,000 | 0,004 | 0,042 | 0,961 | 0,002 |
| Weighted Avg. | 0,231 | 0,038 | 0,987 | 0,231 | 0,370 | 0,053 | 0,940 | 0,987 |

Pada tabel 4.65 merupakan hasil klasifikasi algoritma *Naive Bayes* pada *tools Weka* menggunakan mode pengujian *10 Cross Validation*, menunjukkan hasil 23.597 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 23,0845% dan 78.623 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 76,9155%, dengan waktu proses klasifikasi selama 0,58 detik. Untuk *screen shot*

hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 172 *point c*.

Tabel 4.66 Confusion Matrix NB 10-CV Split Data 60:40 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-------|-----|------|------------------------------|--------|
| a | b | c | d | | |
| 23151 | 70902 | 157 | 6702 | a = Benign | Aktual |
| 39 | 75 | 726 | 64 | b = Web Attack Brute Force | |
| 3 | 25 | 358 | 5 | c = Web Attack XSS | |
| 0 | 0 | 0 | 13 | d = Web Attack Sql Injection | |

Tabel 4.66 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 23.151 data diprediksi benar kategori *benign*, 70.902 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 157 data salah dikategorikan *web attack xss* seharusnya *benign* dan 6.702 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 39 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 75 data benar diprediksi sebagai *web attack brute force*, 726 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 64 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 3 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 25 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 358 data benar diprediksi sebagai *web attack xss* dan 5 data salah

dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 13 data diprediksi benar kategori *web attack sql injection*.

2. C45

a. Use Training Set

Tabel 4.67 Hasil Klasifikasi C45 UTS Split Data 60:40 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,021 | 1,000 | 1,000 | 1,000 | 0,989 | 0,995 | 1,000 |
| Web Attack Brute Force | 0,986 | 0,003 | 0,726 | 0,986 | 0,836 | 0,844 | 0,998 | 0,825 |
| Web Attack XSS | 0,100 | 0,000 | 0,867 | 0,100 | 0,179 | 0,293 | 0,983 | 0,447 |
| Web Attack Sql Injection | 0,692 | 0,000 | 1,000 | 0,692 | 0,818 | 0,832 | 1,000 | 0,760 |
| Weighted Avg. | 0,996 | 0,020 | 0,997 | 0,996 | 0,996 | 0,985 | 0,995 | 0,996 |

Pada tabel 4.67 merupakan hasil klasifikasi algoritma C45 pada *tools* Weka menggunakan mode pengujian *Use Training Set*, menunjukkan hasil 101.850 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,638% dan 370 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,362%, dengan waktu proses klasifikasi selama 5,1 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 173 *point* a.

Tabel 4.68 Confusion Matrix C45 UTS Split Data 60:40 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|----|---|------------------------------|---------------|
| a | b | c | d | | |
| 100911 | 1 | 0 | 0 | a = Benign | Aktual |
| 7 | 891 | 6 | 0 | b = Web Attack Brute Force | |
| 20 | 332 | 39 | 0 | c = Web Attack XSS | |
| 0 | 4 | 0 | 9 | d = Web Attack Sql Injection | |

Tabel 4.68 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 100.911 data diprediksi benar kategori *benign*, dan 1 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 7 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 891 data benar diprediksi sebagai *web attack brute force*, dan 6 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 20 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 332 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 39 data benar diprediksi sebagai *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 4 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 9 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.69 Hasil Klasifikasi C45 5-CV Split Data 60:40 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,024 | 1,000 | 1,000 | 1,000 | 0,984 | 0,993 | 1,000 |
| Web Attack Brute Force | 0,956 | 0,004 | 0,708 | 0,956 | 0,814 | 0,821 | 0,996 | 0,793 |
| Web Attack XSS | 0,066 | 0,000 | 0,464 | 0,066 | 0,116 | 0,175 | 0,974 | 0,367 |
| Web Attack Sql Injection | 0,615 | 0,000 | 0,800 | 0,615 | 0,696 | 0,702 | 0,961 | 0,558 |
| Weighted Avg. | 0,996 | 0,023 | 0,995 | 0,996 | 0,995 | 0,980 | 0,993 | 0,996 |

Pada tabel 4.69 merupakan hasil klasifikasi algoritma C45 pada *tools* Weka menggunakan mode pengujian 5 *Cross Validation*, menunjukkan hasil 101.801 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5901% dan 419 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,4099%, dengan waktu proses klasifikasi selama 3,84 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 173 *point* b.

Tabel 4.70 Confusion Matrix C45 5-CV Split Data 60:40 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|----|---|------------------------------|---------------|
| a | b | c | d | | |
| 100903 | 8 | 0 | 1 | a = Benign | Aktual |
| 9 | 864 | 30 | 1 | b = Web Attack Brute Force | |
| 21 | 344 | 26 | 0 | c = Web Attack XSS | |
| 1 | 4 | 0 | 8 | d = Web Attack Sql Injection | |

Tabel 4.70 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class*

“Benign” terdapat sebanyak 100.903 data diprediksi benar kategori *benign*, 8 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada class “Web Attack Brute Force” terdapat sebanyak 9 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 864 data benar diprediksi sebagai *web attack brute force*, 30 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada class “Web Attack XSS” terdapat sebanyak 21 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 344 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 26 data benar diprediksi sebagai *web attack xss*. Pada class “Web Attack Sql Injection” terdapat sebanyak 1 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 4 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection*, dan 8 data diprediksi benar kategori *web attack sql injection*.

c. 10 Cross Validation

Tabel 4.71 Hasil Klasifikasi C45 10-CV Split Data 60:40 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,021 | 1,000 | 1,000 | 1,000 | 0,986 | 0,994 | 1,000 |
| Web Attack Brute Force | 0,959 | 0,004 | 0,703 | 0,959 | 0,811 | 0,819 | 0,996 | 0,800 |
| Web Attack XSS | 0,038 | 0,000 | 0,333 | 0,038 | 0,069 | 0,112 | 0,973 | 0,357 |
| Web Attack Sql Injection | 0,692 | 0,000 | 0,818 | 0,692 | 0,750 | 0,750 | 1,000 | 0,631 |
| Weighted Avg. | 0,996 | 0,021 | 0,995 | 0,996 | 0,995 | 0,981 | 0,994 | 0,996 |

Pada tabel 4.71 merupakan hasil klasifikasi algoritma C45 pada *tools* Weka menggunakan mode pengujian 10 *Cross Validation*, menunjukkan hasil 101.794 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5833% dan 426 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,4167%, dengan waktu proses klasifikasi selama 4,02 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 174 *point c*.

Tabel 4.72 Confusion Matrix C45 10-CV Split Data 60:40 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|----|---|------------------------------|--------|
| a | b | c | d | | |
| 100903 | 7 | 1 | 1 | a = Benign | Aktual |
| 7 | 867 | 29 | 1 | b = Web Attack Brute Force | |
| 21 | 355 | 15 | 0 | c = Web Attack XSS | |
| 0 | 4 | 0 | 9 | d = Web Attack Sql Injection | |

Tabel 4.72 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 100.903 data diprediksi benar kategori *benign*, 7 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 1 data salah dikategorikan *web attack xss* seharusnya *benign* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 7 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 867 data benar diprediksi sebagai *web attack brute force*, 29 data salah dikategorikan *web attack xss* seharusnya *web*

attack brute force dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada class “*Web Attack XSS*” terdapat sebanyak 21 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 355 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 15 data benar diprediksi sebagai *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 4 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 9 data diprediksi benar kategori *web attack sql injection*.

3. Random Forest

a. Use Training Set

Tabel 4.73 Hasil Klasifikasi RF UTS Split Data 60:40 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,006 | 1,000 | 1,000 | 1,000 | 0,997 | 1,000 | 1,000 |
| Web Attack Brute Force | 0,999 | 0,000 | 0,998 | 0,998 | 0,998 | 0,998 | 1,000 | 1,000 |
| Web Attack XSS | 0,980 | 0,000 | 1,000 | 0,990 | 0,990 | 0,990 | 1,000 | 0,999 |
| Web Attack Sql Injection | 0,923 | 0,000 | 1,000 | 0,923 | 0,960 | 0,961 | 1,000 | 0,995 |
| Weighted Avg. | 1,000 | 0,006 | 1,000 | 1,000 | 1,000 | 0,997 | 1,000 | 1,000 |

Pada tabel 4.73 merupakan hasil klasifikasi algoritma *Random Forest* pada *tools Weka* menggunakan mode pengujian *Use Training Set*, menunjukkan hasil 102.210 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,9902% dan 10 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,0098%, dengan waktu proses klasifikasi selama 72,16 detik. Untuk *screen shot* hasil

pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 174 *point* a.

Tabel 4.74 Confusion Matrix RF UTS Split Data 60:40 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|-----|----|------------------------------|--------|
| a | b | c | d | | |
| 100912 | 0 | 0 | 0 | a = Benign | Aktual |
| 1 | 903 | 0 | 0 | b = Web Attack Brute Force | |
| 7 | 1 | 383 | 0 | c = Web Attack XSS | |
| 0 | 1 | 0 | 12 | d = Web Attack Sql Injection | |

Tabel 4.74 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 100.912 data diprediksi benar kategori *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 1 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force* dan 903 data benar diprediksi sebagai *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 7 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 1 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 383 data benar diprediksi sebagai *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 1 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 12 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.75 Hasil Klasifikasi RF 5-CV Split Data 60:40 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,029 | 1,000 | 1,000 | 1,000 | 0,982 | 0,994 | 1,000 |
| Web Attack Brute Force | 0,793 | 0,002 | 0,746 | 0,793 | 0,769 | 0,767 | 0,996 | 0,807 |
| Web Attack XSS | 0,330 | 0,002 | 0,423 | 0,330 | 0,371 | 0,371 | 0,966 | 0,374 |
| Web Attack Sql Injection | 0,538 | 0,000 | 0,583 | 0,538 | 0,560 | 0,560 | 0,846 | 0,629 |
| Weighted Avg. | 0,995 | 0,029 | 0,995 | 0,995 | 0,995 | 0,978 | 0,994 | 0,996 |

Pada tabel 4.75 merupakan hasil klasifikasi algoritma *Random Forest* pada tools Weka menggunakan mode pengujian *5 Cross Validation*, menunjukkan hasil 101.757 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5471% dan 463 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,4529% , dengan waktu proses klasifikasi selama 63,14 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 175 *point* b.

Tabel 4.76 Confusion Matrix RF 5-CV Split Data 60:40 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|-----|---|------------------------------|--------|
| a | b | c | d | | |
| 100904 | 3 | 5 | 0 | a = Benign | Aktual |
| 11 | 717 | 171 | 5 | b = Web Attack Brute Force | |
| 24 | 238 | 129 | 0 | c = Web Attack XSS | |
| 3 | 3 | 0 | 7 | d = Web Attack Sql Injection | |

Tabel 4.76 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class*

“Benign” terdapat sebanyak 100.904 data diprediksi benar kategori *benign*, 3 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, dan 5 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada class “*Web Attack Brute Force*” terdapat sebanyak 11 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 717 data benar diprediksi sebagai *web attack brute force*, 171 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 5 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada class “*Web Attack XSS*” terdapat sebanyak 24 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 238 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 129 data benar diprediksi sebagai *web attack*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 3 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 3 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 7 data diprediksi benar kategori *web attack sql injection*.

c. 10 Cross Validation

Tabel 4.77 Hasil Klasifikasi RF 10-CV Split Data 60:40 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,028 | 1,000 | 1,000 | 1,000 | 0,982 | 0,995 | 1,000 |
| Web Attack Brute Force | 0,792 | 0,003 | 0,738 | 0,792 | 0,764 | 0,762 | 0,996 | 0,806 |
| Web Attack XSS | 0,304 | 0,002 | 0,401 | 0,304 | 0,347 | 0,347 | 0,970 | 0,375 |
| Web Attack Sql Injection | 0,615 | 0,000 | 0,615 | 0,615 | 0,615 | 0,615 | 0,884 | 0,637 |
| Weighted Avg. | 0,995 | 0,028 | 0,995 | 0,995 | 0,995 | 0,978 | 0,994 | 0,996 |

Pada tabel 4.77 merupakan hasil klasifikasi algoritma *Random Forest* pada *tools* Weka menggunakan mode pengujian 10 *Cross Validation*, menunjukkan hasil 101.746 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5363% dan 474 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,4637% , dengan waktu proses klasifikasi selama 65,56 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 175 *point c*.

Tabel 4.78 Confusion Matrix RF 10-CV Split Data 60:40 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|-----|---|------------------------------|--------|
| a | b | c | d | | |
| 100903 | 3 | 6 | 0 | a = Benign | Aktual |
| 11 | 716 | 172 | 5 | b = Web Attack Brute Force | |
| 24 | 248 | 119 | 0 | c = Web Attack XSS | |
| 2 | 3 | 0 | 8 | d = Web Attack Sql Injection | |

Tabel 4.78 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “*Benign*” terdapat sebanyak 100.903 data diprediksi benar kategori *benign*, 3 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, dan 6 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “*Web Attack Brute Force*” terdapat sebanyak 11 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 716 data benar diprediksi sebagai *web attack brute force*, 172 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 5 data salah dikategorikan sebagai *web attack sql injection*

seharusnya *web attack brute force*. Pada class “*Web Attack XSS*” terdapat sebanyak 24 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 248 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 119 data benar diprediksi sebagai *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 2 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 3 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 8 data diprediksi benar kategori *web attack sql injection*.

4.2.2.2 Klasifikasi Pada Split Data 70 : 30 Dengan Seleksi Fitur

Perbandingan split data yang digunakan yaitu 70:30, 70% dari 170.366 *record* data atau sebanyak 119.256 *record* data digunakan sebagai data *training* dan 30% atau 51.110 *record* data sebagai *data testing* dengan seleksi fitur atau menggunakan 11 fitur hasil seleksi fitur *information gain*.

1. Naive Bayes

a. Use Training Set

Tabel 4.79 Hasil Klasifikasi NB UTS Split Data 70:30 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,229 | 0,029 | 0,998 | 0,229 | 0,372 | 0,054 | 0,950 | 0,999 |
| Web Attack Brute Force | 0,095 | 0,698 | 0,001 | 0,095 | 0,002 | -0,122 | 0,293 | 0,006 |
| Web Attack XSS | 0,921 | 0,008 | 0,294 | 0,921 | 0,446 | 0,518 | 0,970 | 0,297 |
| Web Attack Sql Injection | 1,000 | 0,069 | 0,002 | 1,000 | 0,004 | 0,041 | 0,963 | 0,005 |
| Weighted Avg. | 0,230 | 0,035 | 0,987 | 0,230 | 0,369 | 0,054 | 0,944 | 0,987 |

Pada tabel 4.79 merupakan hasil klasifikasi algoritma *Naive Bayes* pada *tools* Weka menggunakan mode pengujian *Use Training Set*, menunjukkan hasil 27.472 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 23,0362% dan 91.784 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 76,9638% , dengan waktu proses klasifikasi selama 5,63 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 176 *point* a.

Tabel 4.80 Confusion Matrix NB UTS Split Data 70:30 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-------|-----|------|------------------------------|--------|
| a | b | c | d | | |
| 26937 | 82503 | 160 | 8130 | a = Benign | Aktual |
| 41 | 100 | 847 | 67 | b = Web Attack Brute Force | |
| 4 | 26 | 420 | 6 | c = Web Attack XSS | |
| 0 | 0 | 0 | 15 | d = Web Attack Sql Injection | |

Tabel 4.80 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “*Benign*” terdapat sebanyak 26.937 data diprediksi benar kategori *benign*, 82.503 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 160 data salah dikategorikan *web attack xss* seharusnya *benign* dan 8.390 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “*Web Attack Brute Force*” terdapat sebanyak 41 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 100 data benar diprediksi sebagai *web attack brute force*, 847 data salah dikategorikan *web attack xss* seharusnya *web*

attack brute force dan 67 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada class “*Web Attack XSS*” terdapat sebanyak 4 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 26 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 420 data benar diprediksi sebagai *web attack xss* dan 6 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 15 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.81 Hasil Klasifikasi NB 5-CV Split Data 70:30 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,231 | 0,030 | 0,998 | 0,231 | 0,375 | 0,054 | 0,947 | 0,999 |
| Web Attack Brute Force | 0,091 | 0,694 | 0,001 | 0,091 | 0,002 | -0,122 | 0,291 | 0,006 |
| Web Attack XSS | 0,921 | 0,009 | 0,291 | 0,921 | 0,442 | 0,515 | 0,969 | 0,290 |
| Web Attack Sql Injection | 0,933 | 0,070 | 0,002 | 0,933 | 0,003 | 0,038 | 0,956 | 0,002 |
| Weighted Avg. | 0,233 | 0,036 | 0,987 | 0,233 | 0,372 | 0,054 | 0,941 | 0,987 |

Pada tabel 4.81 merupakan hasil klasifikasi algoritma *Naive Bayes* pada *tools Weka* menggunakan mode pengujian *5 Cross Validation*, menunjukkan hasil 27.731 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 23,2533% dan 91.525 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 76,7467%, dengan waktu proses klasifikasi selama 0,64 detik. Untuk *screen shot*

hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 176 *point* b.

Tabel 4.82 Confusion Matrix NB 5-CV Data 70:30 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-------|-----|------|------------------------------|--------|
| a | b | c | d | | |
| 27201 | 82036 | 177 | 8316 | a = Benign | Aktual |
| 43 | 96 | 846 | 70 | b = Web Attack Brute Force | |
| 3 | 27 | 420 | 6 | c = Web Attack XSS | |
| 0 | 1 | 0 | 14 | d = Web Attack Sql Injection | |

Tabel 4.82 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 27.201 data diprediksi benar kategori *benign*, 82.036 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 177 data salah dikategorikan sebagai *web attack xss* seharusnya *benign* dan 8.316 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 43 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 96 data benar diprediksi sebagai *web attack brute force*, 846 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 70 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 3 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 27 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 420 data benar diprediksi sebagai *web attack xss* dan 6 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada

class “*Web Attack Sql Injection*” terdapat sebanyak 1 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection*, dan 14 data diprediksi benar kategori *web attack sql injection*.

c. 10 Cross Validation

Tabel 4.83 Hasil Klasifikasi NB 10-CV Split Data 70:30 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,229 | 0,029 | 0,998 | 0,229 | 0,372 | 0,054 | 0,948 | 0,999 |
| Web Attack Brute Force | 0,092 | 0,699 | 0,001 | 0,092 | 0,002 | -0,123 | 0,289 | 0,006 |
| Web Attack XSS | 0,921 | 0,009 | 0,293 | 0,921 | 0,444 | 0,517 | 0,969 | 0,288 |
| Web Attack Sql Injection | 0,867 | 0,068 | 0,002 | 0,867 | 0,003 | 0,036 | 0,961 | 0,002 |
| Weighted Avg. | 0,230 | 0,035 | 0,987 | 0,230 | 0,369 | 0,054 | 0,942 | 0,987 |

Pada tabel 4.83 merupakan hasil klasifikasi algoritma *Naive Bayes* pada *tools* Weka menggunakan mode pengujian 10 *Cross Validation*, menunjukkan hasil 27.457 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 23,0236% dan 91.799 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 76,9764% , dengan waktu proses klasifikasi selama 0,58 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 177 *point c*.

Tabel 4.84 Confusion Matrix NB 10-CV Split Data 70:30 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-------|-----|------|------------------------------|--------|
| a | b | c | d | | |
| 26927 | 82656 | 167 | 7980 | a = Benign | Aktual |
| 42 | 97 | 847 | 69 | b = Web Attack Brute Force | |
| 3 | 27 | 420 | 6 | c = Web Attack XSS | |
| 0 | 2 | 0 | 13 | d = Web Attack Sql Injection | |

Tabel 4.84 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 26.927 data diprediksi benar kategori *benign*, 82.645 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 167 data salah dikategorikan *web attack xss* seharusnya *benign* dan 7.980 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 42 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 97 data benar diprediksi sebagai *web attack brute force*, 847 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 69 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 3 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 27 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 420 data benar diprediksi sebagai *web attack xss* dan 6 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 2 data salah dikategorikan

sebagai *web attack brute force* seharusnya *web attack sql injection*, dan 13 data diprediksi benar kategori *web attack sql injection*.

2. C45

a. Use Training Set

Tabel 4.85 Hasil Klasifikasi C45 UTS Split Data 70:30 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,020 | 1,000 | 1,000 | 1,000 | 0,990 | 0,995 | 1,000 |
| Web Attack Brute Force | 0,975 | 0,003 | 0,728 | 0,986 | 0,834 | 0,841 | 0,998 | 0,824 |
| Web Attack XSS | 0,125 | 0,000 | 0,760 | 0,100 | 0,215 | 0,307 | 0,982 | 0,448 |
| Web Attack Sql Injection | 0,600 | 0,000 | 1,000 | 0,692 | 0,750 | 0,775 | 1,000 | 0,685 |
| Weighted Avg. | 0,996 | 0,019 | 0,996 | 0,996 | 0,995 | 0,986 | 0,995 | 0,996 |

Pada tabel 4.85 merupakan hasil klasifikasi algoritma C45 pada *tools* Weka menggunakan mode pengujian *Use Training Set*, menunjukkan hasil 118.824 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,6378 % dan 432 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,3622%, dengan waktu proses klasifikasi selama 5,86 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 177 *point a*.

Tabel 4.86 Confusion Matrix C45 UTS Split Data 70:30 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|------|----|---|------------------------------|--------|
| a | b | c | d | | |
| 117729 | 1 | 0 | 0 | a = Benign | Aktual |
| 8 | 1029 | 18 | 0 | b = Web Attack Brute Force | |
| 22 | 377 | 57 | 0 | c = Web Attack XSS | |
| 0 | 6 | 0 | 9 | d = Web Attack Sql Injection | |

Tabel 4.86 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 117.729 data diprediksi benar kategori *benign*, dan 1 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 8 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 1.029 data benar diprediksi sebagai *web attack brute force*, dan 18 data salah dikategorikan *web attack xss* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 22 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 377 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 57 data benar diprediksi sebagai *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 6 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection*, dan 9 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.87 Hasil Klasifikasi C45 5-CV Split Data 70:30 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,022 | 1,000 | 1,000 | 1,000 | 0,986 | 0,992 | 1,000 |
| Web Attack Brute Force | 0,972 | 0,004 | 0,708 | 0,972 | 0,819 | 0,828 | 0,996 | 0,793 |
| Web Attack XSS | 0,053 | 0,000 | 0,533 | 0,053 | 0,096 | 0,167 | 0,969 | 0,367 |
| Web Attack Sql Injection | 0,600 | 0,000 | 0,818 | 0,600 | 0,692 | 0,701 | 1,000 | 0,592 |
| Weighted Avg. | 0,996 | 0,021 | 0,995 | 0,996 | 0,995 | 0,981 | 0,992 | 0,995 |

Pada tabel 4.87 merupakan hasil klasifikasi algoritma C45 pada *tools* Weka menggunakan mode pengujian 5 *Cross Validation*, menunjukkan hasil 118.778 data dengan prediksi benar terklasifikasi dengan akurasi sebesar 99,5992% dan 478 data dengan prediksi salah terklasifikasi dengan persentasi sebesar 0,4008 %, dengan waktu proses klasifikasi selama 4,76 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 178 *point* b.

Tabel 4.88 Confusion Matrix C45 5-CV Split Data 70:30 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|------|----|---|------------------------------|---------------|
| a | b | c | d | | |
| 117720 | 7 | 2 | 1 | a = Benign | Aktual |
| 10 | 1025 | 19 | 1 | b = Web Attack Brute Force | |
| 23 | 409 | 24 | 0 | c = Web Attack XSS | |
| 0 | 6 | 0 | 9 | d = Web Attack Sql Injection | |

Tabel 4.88 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class*

“Benign” terdapat sebanyak 117.720 data diprediksi benar kategori *benign*, 7 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 2 data salah dikategorikan *web attack xss* seharusnya *benign* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada class “Web Attack Brute Force” terdapat sebanyak 10 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 1.025 data benar diprediksi sebagai *web attack brute force*, 19 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada class “Web Attack XSS” terdapat sebanyak 23 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 409 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 24 data benar diprediksi sebagai *web attack xss*. Pada class “Web Attack Sql Injection” terdapat sebanyak 6 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection*, dan 9 data diprediksi benar kategori *web attack sql injection*.

c. 10 Cross Validation

Tabel 4.89 Hasil Klasifikasi C45 10-CV Split Data 70:30 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,021 | 1,000 | 1,000 | 1,000 | 0,986 | 0,991 | 1,000 |
| Web Attack Brute Force | 0,972 | 0,004 | 0,710 | 0,972 | 0,821 | 0,829 | 0,997 | 0,815 |
| Web Attack XSS | 0,061 | 0,000 | 0,571 | 0,061 | 0,111 | 0,186 | 0,961 | 0,398 |
| Web Attack Sql Injection | 0,600 | 0,000 | 0,818 | 0,600 | 0,692 | 0,701 | 1,000 | 0,575 |
| Weighted Avg. | 0,996 | 0,021 | 0,996 | 0,996 | 0,995 | 0,982 | 0,991 | 0,996 |

Pada gambar 4.89 merupakan hasil klasifikasi algoritma C45 pada *tools* Weka menggunakan mode pengujian 10 *Cross Validation*, menunjukkan hasil 118.783 data dengan prediksi benar terklasifikasi dengan akurasi sebesar 99,6034% dan 473 data dengan prediksi salah terklasifikasi dengan persentasi sebesar 0,3966% , dengan waktu proses klasifikasi selama 4,79 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 178 *point c*.

Tabel 4.90 Confusion Matrix C45 10-CV Split Data 70:30 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|------|----|---|------------------------------|--------|
| a | b | c | D | | |
| 117721 | 7 | 1 | 1 | a = Benign | Aktual |
| 9 | 1025 | 20 | 1 | b = Web Attack Brute Force | |
| 23 | 405 | 28 | 0 | c = Web Attack XSS | |
| 0 | 6 | 0 | 9 | d = Web Attack Sql Injection | |

Tabel 4.90 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 117.721 data diprediksi benar kategori *benign*, 7 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 1 data salah dikategorikan *web attack xss* seharusnya *benign* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 9 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 1.025 data benar diprediksi sebagai *web attack brute force*, 20 data salah dikategorikan *web attack xss* seharusnya *web*

attack brute force dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada class “*Web Attack XSS*” terdapat sebanyak 23 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 405 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 28 data benar diprediksi sebagai *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 6 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection*, dan 9 data diprediksi benar kategori *web attack sql injection*.

3. Random Forest

a. Use Training Set

Tabel 4.91 Hasil Klasifikasi RF UTS Split Data 70:30 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,006 | 1,000 | 1,000 | 1,000 | 0,997 | 1,000 | 1,000 |
| Web Attack Brute Force | 0,998 | 0,000 | 0,997 | 0,998 | 0,998 | 0,998 | 1,000 | 1,000 |
| Web Attack XSS | 0,980 | 0,000 | 1,000 | 0,990 | 0,990 | 0,990 | 1,000 | 0,999 |
| Web Attack Sql Injection | 0,867 | 0,000 | 0,929 | 0,867 | 0,897 | 0,897 | 1,000 | 0,973 |
| Weighted Avg. | 1,000 | 0,006 | 1,000 | 1,000 | 1,000 | 0,997 | 1,000 | 1,000 |

Pada tabel 4.91 merupakan hasil klasifikasi algoritma *Random Forest* pada *tools* Weka menggunakan mode pengujian *Use Training Set*, menunjukkan hasil 119.243 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,9891% dan 13 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,0109%, dengan waktu proses klasifikasi selama 92,22 detik. Untuk *screen shot* hasil

pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 179 *point* a.

Tabel 4.92 Confusion Matrix RF UTS Split Data 70:30 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|------|-----|----|------------------------------|--------|
| a | b | c | d | | |
| 117730 | 0 | 0 | 0 | a = Benign | Aktual |
| 1 | 1053 | 0 | 1 | b = Web Attack Brute Force | |
| 8 | 1 | 447 | 0 | c = Web Attack XSS | |
| 0 | 2 | 0 | 13 | d = Web Attack Sql Injection | |

Tabel 4.92 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 117.730 data diprediksi benar kategori *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 1 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 1.053 data benar diprediksi sebagai *web attack brute force*, dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 8 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 1 data salah dikategorikan sebagai sebagai *web attack brute force* seharusnya *web attack xss*, dan 447 data benar diprediksi sebagai *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 2 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection*, dan 13 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.93 Hasil Klasifikasi RF 5-CV Split Data 70:30 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,026 | 1,000 | 1,000 | 1,000 | 0,984 | 0,995 | 1,000 |
| Web Attack Brute Force | 0,783 | 0,003 | 0,728 | 0,783 | 0,755 | 0,753 | 0,998 | 0,795 |
| Web Attack XSS | 0,294 | 0,002 | 0,382 | 0,294 | 0,332 | 0,333 | 0,968 | 0,366 |
| Web Attack Sql Injection | 0,467 | 0,000 | 0,583 | 0,467 | 0,519 | 0,522 | 0,866 | 0,565 |
| Weighted Avg. | 0,995 | 0,025 | 0,995 | 0,995 | 0,995 | 0,979 | 0,995 | 0,996 |

Pada tabel 4.93 merupakan hasil klasifikasi algoritma *Random Forest* pada *tools* Weka menggunakan mode pengujian *5 Cross Validation*, menunjukkan hasil 118.687 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5229% dan 569 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,4771%, dengan waktu proses klasifikasi selama 82,04 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 179 *point* b.

Tabel 4.94 Confusion Matrix RF 5-CV Data 70:30 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|-----|---|------------------------------|--------|
| a | b | c | d | | |
| 117720 | 4 | 6 | 0 | a = Benign | Aktual |
| 13 | 826 | 211 | 5 | b = Web Attack Brute Force | |
| 23 | 299 | 134 | 0 | c = Web Attack XSS | |
| 3 | 5 | 0 | 7 | d = Web Attack Sql Injection | |

Tabel 4.94 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class*

“Benign” terdapat sebanyak 117.720 data diprediksi benar kategori *benign*, 4 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, dan 6 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada class “Web Attack Brute Force” terdapat sebanyak 13 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 826 data benar diprediksi sebagai *web attack brute force*, 211 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 5 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada class “Web Attack XSS” terdapat sebanyak 23 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 299 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 134 data benar diprediksi sebagai *web attack xss*. Pada class “Web Attack Sql Injection” terdapat sebanyak 3 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 5 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection*, dan 7 data diprediksi benar kategori *web attack sql injection*.

c. 10 Cross Validation

Tabel 4.95 Hasil Klasifikasi RF 10-CV Split Data 70:30 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,026 | 1,000 | 1,000 | 1,000 | 0,984 | 0,995 | 1,000 |
| Web Attack Brute Force | 0,780 | 0,003 | 0,728 | 0,78 | 0,753 | 0,751 | 0,997 | 0,805 |
| Web Attack XSS | 0,294 | 0,002 | 0,380 | 0,294 | 0,331 | 0,332 | 0,970 | 0,369 |
| Web Attack Sql Injection | 0,533 | 0,000 | 0,667 | 0,533 | 0,593 | 0,596 | 0,933 | 0,518 |
| Weighted Avg. | 0,995 | 0,025 | 0,995 | 0,995 | 0,995 | 0,979 | 0,995 | 0,996 |

Pada tabel 4.95 merupakan hasil klasifikasi algoritma *Random Forest* pada *tools* Weka menggunakan mode pengujian *10 Cross Validation*, menunjukkan hasil 118.686 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,522% dan 570 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi akurasi sebesar 0,478%, dengan waktu proses klasifikasi selama 83,33 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 180 *point c*.

Tabel 4.96 Confusion Matrix RF 10-CV Split Data 70:30 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|-----|---|------------------------------|--------|
| a | b | c | d | | |
| 117721 | 4 | 5 | 0 | a = Benign | Aktual |
| 14 | 823 | 214 | 4 | b = Web Attack Brute Force | |
| 23 | 299 | 134 | 0 | c = Web Attack XSS | |
| 2 | 5 | 0 | 8 | d = Web Attack Sql Injection | |

Tabel 4.96 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “*Benign*” terdapat sebanyak 117.721 data diprediksi benar kategori *benign*, 4 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, dan 5 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “*Web Attack Brute Force*” terdapat sebanyak 14 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 823 data benar diprediksi sebagai *web attack brute force*, 214 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 4 data salah dikategorikan sebagai *web attack sql injection*

seharusnya *web attack brute force*. Pada class “*Web Attack XSS*” terdapat sebanyak 23 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 299 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 134 data benar diprediksi sebagai *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 2 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 5 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection*, dan 8 data diprediksi benar kategori *web attack sql injection*.

4.2.2.3 Klasifikasi Pada Split Data 80 : 20 Dengan Seleksi Fitur

Perbandingan split data yang digunakan yaitu 80:20, 80% dari 170.366 *record* data atau sebanyak 136.294 *record* data digunakan sebagai data *training* dan 20% atau 34.072 *record* data sebagai data *testing* dengan seleksi fitur atau menggunakan 11 fitur hasil seleksi fitur *information gain*.

1. Naive Bayes

a. Use Training Set

Tabel 4.97 Hasil Klasifikasi NB UTS Split Data 80:20 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,233 | 0,026 | 0,999 | 0,233 | 0,378 | 0,055 | 0,936 | 0,998 |
| Web Attack Brute Force | 0,104 | 0,699 | 0,001 | 0,104 | 0,003 | 0,121 | 0,278 | 0,006 |
| Web Attack XSS | 0,921 | 0,008 | 0,297 | 0,921 | 0,449 | 0,520 | 0,971 | 0,304 |
| Web Attack Sql Injection | 1,000 | 0,064 | 0,002 | 1,000 | 0,004 | 0,043 | 0,968 | 0,006 |
| Weighted Avg. | 0,235 | 0,032 | 0,987 | 0,235 | 0,375 | 0,056 | 0,931 | 0,987 |

Pada tabel 4.97 merupakan hasil klasifikasi algoritma *Naive Bayes* pada *tools* Weka menggunakan mode pengujian *Use Training Set*, menunjukkan hasil

32.032 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 23,5021% dan 104.262 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 76,4979%, dengan waktu proses klasifikasi selama 5,06 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 180 *point* a.

Tabel 4.98 Confusion Matrix NB UTS Split Data 80:20 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-------|-----|------|------------------------------|--------|
| a | b | c | d | | |
| 31409 | 94350 | 173 | 8617 | a = Benign | Aktual |
| 42 | 125 | 966 | 73 | b = Web Attack Brute Force | |
| 3 | 33 | 481 | 5 | c = Web Attack XSS | |
| 0 | 0 | 0 | 17 | d = Web Attack Sql Injection | |

Tabel 4.98 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 31.409 data diprediksi benar kategori *benign*, 94.350 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 173 data salah dikategorikan *web attack xss* seharusnya *benign* dan 8.617 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 42 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 125 data benar diprediksi sebagai *web attack brute force*, 966 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 73 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat

sebanyak 3 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 33 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 481 data benar diprediksi sebagai *web attack xss* dan 5 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada class “*Web Attack Sql Injection*” terdapat sebanyak 17 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.99 Hasil Klasifikasi NB 5-CV Split Data 80:20 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,231 | 0,026 | 0,999 | 0,231 | 0,375 | 0,055 | 0,947 | 0,999 |
| Web Attack Brute Force | 0,103 | 0,696 | 0,001 | 0,103 | 0,003 | -0,120 | 0,293 | 0,006 |
| Web Attack XSS | 0,921 | 0,008 | 0,295 | 0,921 | 0,447 | 0,519 | 0,969 | 0,302 |
| Web Attack Sql Injection | 0,824 | 0,068 | 0,001 | 0,824 | 0,003 | 0,033 | 0,962 | 0,002 |
| Weighted Avg. | 0,233 | 0,032 | 0,987 | 0,233 | 0,372 | 0,055 | 0,942 | 0,987 |

Pada tabel 4.99 merupakan hasil klasifikasi algoritma *Naive Bayes* pada *tools* Weka menggunakan mode pengujian *5 Cross Validation*, menunjukkan hasil 31.707 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 23,2637% dan 104.587 data dengan prediksi salah terklasifikasi (*Inccorrectly Classified Instances*) dengan persentasi sebesar 76,7363% , dengan waktu proses klasifikasi selama 0,93 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 181 *point* b.

Tabel 4.100 Confusion Matrix NB 5-CV Split Data 80:20 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-------|-----|------|------------------------------|--------|
| a | b | c | d | | |
| 31088 | 94028 | 184 | 9249 | a = Benign | Aktual |
| 43 | 124 | 966 | 73 | b = Web Attack Brute Force | |
| 3 | 32 | 481 | 6 | c = Web Attack XSS | |
| 0 | 3 | 0 | 14 | d = Web Attack Sql Injection | |

Tabel 4.100 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 31.088 data diprediksi benar kategori *benign*, 94.028 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 184 data salah dikategorikan *web attack xss* seharusnya *benign* dan 9.249 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 43 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 124 data benar diprediksi sebagai *web attack brute force*, 966 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 73 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 3 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 32 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 481 data benar diprediksi sebagai *web attack xss* dan 6 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 3 data salah dikategorikan

sebagai *web attack brute force* seharusnya *web attack sql injection*, dan 14 data diprediksi benar kategori *web attack sql injection*.

c. 10 *Cross Validation*

Tabel 4.101 Hasil Klasifikasi NB 10-CV Split Data 80:20 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 0,230 | 0,027 | 0,998 | 0,230 | 0,374 | 0,055 | 0,948 | 0,999 |
| Web Attack Brute Force | 0,101 | 0,696 | 0,001 | 0,101 | 0,003 | -0,120 | 0,295 | 0,006 |
| Web Attack XSS | 0,921 | 0,008 | 0,296 | 0,921 | 0,448 | 0,519 | 0,967 | 0,284 |
| Web Attack Sql Injection | 0,882 | 0,070 | 0,002 | 0,882 | 0,003 | 0,036 | 0,962 | 0,002 |
| Weighted Avg. | 0,232 | 0,033 | 0,987 | 0,232 | 0,371 | 0,055 | 0,942 | 0,987 |

Pada tabel 4.101 merupakan hasil klasifikasi algoritma *Naive Bayes* pada *tools* Weka menggunakan mode pengujian 10 *Cross Validation*, menunjukkan hasil 31.623 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 23,202% dan 104.671 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 76,798%, dengan waktu proses klasifikasi selama 1,31 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 181 *point c*.

Tabel 4.102 Confusion Matrix NB 10-CV Split Data 80:20 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-------|-----|------|------------------------------|--------|
| a | b | c | d | | |
| 31005 | 93938 | 180 | 9426 | a = Benign | Aktual |
| 42 | 122 | 966 | 76 | b = Web Attack Brute Force | |
| 5 | 29 | 481 | 7 | c = Web Attack XSS | |
| 0 | 2 | 0 | 15 | d = Web Attack Sql Injection | |

Tabel 4.102 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 31.005 data diprediksi benar kategori *benign*, 93.938 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 180 data salah dikategorikan *web attack xss* seharusnya *benign* dan 9.426 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 42 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 122 data benar diprediksi sebagai *web attack brute force*, 966 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 76 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 5 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 29 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, 481 data benar diprediksi sebagai *web attack xss* dan 7 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 2 data salah dikategorikan

sebagai *web attack brute force* seharusnya *web attack sql injection*, dan 15 data diprediksi benar kategori *web attack sql injection*.

2. C45

a. Use Training Set

Tabel 4.103 Hasil Klasifikasi C45 UTS Split Data 80:20 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,019 | 1,000 | 1,000 | 1,000 | 0,990 | 0,995 | 1,000 |
| Web Attack Brute Force | 0,989 | 0,003 | 0,720 | 0,989 | 0,833 | 0,842 | 0,998 | 0,816 |
| Web Attack XSS | 0,079 | 0,000 | 0,932 | 0,079 | 0,145 | 0,270 | 0,982 | 0,430 |
| Web Attack Sql Injection | 0,529 | 0,000 | 0,900 | 0,529 | 0,667 | 0,690 | 1,000 | 0,575 |
| Weighted Avg. | 0,996 | 0,019 | 0,997 | 0,996 | 0,995 | 0,986 | 0,995 | 0,996 |

Pada tabel 4.103 merupakan hasil klasifikasi algoritma C45 pada *tools* Weka menggunakan mode pengujian *Use Training Set*, menunjukkan hasil 135.791 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,6309% dan 503 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,3691%, dengan waktu proses klasifikasi selama 8,54 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 182 *point* a.

Tabel 4.104 Confusion Matrix C45 UTS Split Data 80:20 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|------|----|---|------------------------------|--------|
| a | b | c | d | | |
| 134548 | 1 | 0 | 0 | a = Benign | Aktual |
| 9 | 1193 | 3 | 1 | b = Web Attack Brute Force | |
| 25 | 456 | 41 | 0 | c = Web Attack XSS | |
| 0 | 8 | 0 | 9 | d = Web Attack Sql Injection | |

Tabel 4.104 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 134.548 data diprediksi benar kategori *benign*, dan 1 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 9 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 1193 data benar diprediksi sebagai *web attack brute force*, 3 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 25 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 456 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 41 data benar diprediksi sebagai *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 8 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 9 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.105 Hasil Klasifikasi C45 5-CV Split Data 80:20 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,021 | 1,000 | 1,000 | 1,000 | 0,986 | 0,993 | 1,000 |
| Web Attack Brute Force | 0,988 | 0,004 | 0,708 | 0,988 | 0,825 | 0,834 | 0,997 | 0,808 |
| Web Attack XSS | 0,034 | 0,000 | 0,720 | 0,034 | 0,066 | 0,157 | 0,967 | 0,385 |
| Web Attack Sql Injection | 0,529 | 0,000 | 0,818 | 0,529 | 0,643 | 0,658 | 1,000 | 0,548 |
| Weighted Avg. | 0,996 | 0,021 | 0,996 | 0,996 | 0,995 | 0,981 | 0,993 | 0,996 |

Pada tabel 4.105 merupakan hasil klasifikasi algoritma C45 pada *tools* Weka menggunakan mode pengujian 5 *Cross Validation*, menunjukkan hasil 135.756 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,6053% dan 538 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,3947%, dengan waktu proses klasifikasi selama 6,4 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 182 *point* b.

Tabel 4.106 Confusion Matrix C45 5-CV Split Data 80:20 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|------|----|---|------------------------------|---------------|
| a | b | c | d | | |
| 134538 | 8 | 2 | 1 | a = Benign | Aktual |
| 9 | 1191 | 5 | 1 | b = Web Attack Brute Force | |
| 28 | 476 | 18 | 0 | c = Web Attack XSS | |
| 0 | 8 | 0 | 9 | d = Web Attack Sql Injection | |

Tabel 4.106 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “*Benign*” terdapat sebanyak 134.538 data diprediksi benar kategori *benign*, 8 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, 2 data salah dikategorikan *web attack xss* seharusnya *benign* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada *class* “*Web Attack Brute Force*” terdapat sebanyak 9 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 1.191 data benar diprediksi sebagai *web attack brute force*, 5 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “*Web Attack XSS*” terdapat sebanyak 28 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 476 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 18 data benar diprediksi sebagai *web attack xss*. Pada *class* “*Web Attack Sql Injection*” terdapat sebanyak 8 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection*, dan 9 data diprediksi benar kategori *web attack sql injection*.

c. 10 Cross Validation

Tabel 4.107 Hasil Klasifikasi C45 10-CV Split Data 80:20 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,021 | 1,000 | 1,000 | 1,000 | 0,987 | 0,992 | 1,000 |
| Web Attack Brute Force | 0,981 | 0,004 | 0,710 | 0,981 | 0,824 | 0,833 | 0,997 | 0,819 |
| Web Attack XSS | 0,046 | 0,000 | 0,649 | 0,046 | 0,086 | 0,172 | 0,970 | 0,395 |
| Web Attack Sql Injection | 0,529 | 0,000 | 0,818 | 0,529 | 0,643 | 0,658 | 1,000 | 0,510 |
| Weighted Avg. | 0,996 | 0,021 | 0,996 | 0,996 | 0,995 | 0,983 | 0,992 | 0,996 |

Pada tabel 4.107 merupakan hasil klasifikasi algoritma C45 pada *tools* Weka menggunakan mode pengujian 10 *Cross Validation*, menunjukkan hasil 135.759 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,6075% dan 535 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,3925%, dengan waktu proses klasifikasi selama 6,42 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 183 *point c*.

Tabel 4.108 Confusion Matrix C45 10-CV Split Data 80:20 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|------|----|---|------------------------------|--------|
| a | b | c | d | | |
| 134543 | 5 | 0 | 1 | a = Benign | Aktual |
| 9 | 1183 | 13 | 1 | b = Web Attack Brute Force | |
| 28 | 470 | 24 | 0 | c = Web Attack XSS | |
| 0 | 8 | 0 | 9 | d = Web Attack Sql Injection | |

Tabel 4.108 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class*

“Benign” terdapat sebanyak 134.543 data diprediksi benar kategori *benign*, 5 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *benign*. Pada class “Web Attack Brute Force” terdapat sebanyak 9 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 1.183 data benar diprediksi sebagai *web attack brute force*, 13 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 1 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada class “Web Attack XSS” terdapat sebanyak 28 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 470 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 24 data benar diprediksi sebagai *web attack xss*. Pada class “Web Attack Sql Injection” terdapat sebanyak 8 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 9 data diprediksi benar kategori *web attack sql injection*.

3. Random Forest

a. Use Training Set

Tabel 4.109 Hasil Klasifikasi RF UTS Split Data 80:20 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,006 | 1,000 | 1,000 | 1,000 | 0,997 | 1,000 | 1,000 |
| Web Attack Brute Force | 0,999 | 0,000 | 0,997 | 0,999 | 0,998 | 0,998 | 1,000 | 1,000 |
| Web Attack XSS | 0,979 | 0,000 | 1,000 | 0,979 | 0,989 | 0,989 | 1,000 | 0,998 |
| Web Attack Sql Injection | 0,824 | 0,000 | 1,000 | 0,824 | 0,903 | 0,907 | 1,000 | 0,979 |
| Weighted Avg. | 1,000 | 0,006 | 1,000 | 1,000 | 1,000 | 0,997 | 1,000 | 1,000 |

Pada tabel 4.109 merupakan hasil klasifikasi algoritma *Random Forest* pada *tools* Weka menggunakan mode pengujian *Use Training Set*, menunjukkan hasil 136.279 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,989% dan 15 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi sebesar 0,011%, dengan waktu proses klasifikasi selama 102,47 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 183 *point* a.

Tabel 4.110 Confusion Matrix RF UTS Split Data 80:20 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|------|-----|----|------------------------------|---------------|
| a | b | c | d | | |
| 134549 | 0 | 0 | 0 | a = Benign | Aktual |
| 1 | 1205 | 0 | 0 | b = Web Attack Brute Force | |
| 10 | 1 | 511 | 0 | c = Web Attack XSS | |
| 0 | 3 | 0 | 14 | d = Web Attack Sql Injection | |

Tabel 4.110 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “*Benign*” terdapat sebanyak 134.549 data diprediksi benar kategori *benign*. Pada *class* “*Web Attack Brute Force*” terdapat sebanyak 1 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, dan 1.205 data benar diprediksi sebagai *web attack brute force*. Pada *class* “*Web Attack XSS*” terdapat sebanyak 10 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 1 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 511 data benar diprediksi sebagai *web attack xss*. Pada *class* “*Web Attack Sql*”

Injection” terdapat sebanyak 3 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection*, dan 14 data diprediksi benar kategori *web attack sql injection*.

b. 5 Cross Validation

Tabel 4.111 Hasil Klasifikasi RF 5-CV Split Data 80:20 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,025 | 1,000 | 1,000 | 1,000 | 0,984 | 0,994 | 1,000 |
| Web Attack Brute Force | 0,792 | 0,003 | 0,727 | 0,792 | 0,758 | 0,756 | 0,756 | 0,799 |
| Web Attack XSS | 0,284 | 0,002 | 0,380 | 0,284 | 0,326 | 0,326 | 0,326 | 0,373 |
| Web Attack Sql Injection | 0,412 | 0,000 | 0,636 | 0,412 | 0,512 | 0,512 | 0,512 | 0,501 |
| Weighted Avg. | 0,995 | 0,024 | 0,995 | 0,995 | 0,979 | 0,979 | 0,974 | 0,996 |

Pada tabel 4.111 merupakan hasil klasifikasi algoritma *Random Forest* pada *tools* Weka menggunakan mode pengujian *5 Cross Validation*, menunjukkan hasil 135.647 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5253% dan 647 data dengan prediksi salah (*Incorrectly Classified Instances*) terklasifikasi dengan persentasi sebesar 0,4747%, dengan waktu proses klasifikasi selama 94,47 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 184 *point* b.

Tabel 4.112 Confusion Matrix RF 5-CV Split Data 80:20 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|-----|---|------------------------------|--------|
| a | b | c | d | | |
| 134537 | 5 | 7 | 0 | a = Benign | Aktual |
| 13 | 955 | 234 | 4 | b = Web Attack Brute Force | |
| 27 | 347 | 148 | 0 | c = Web Attack XSS | |
| 3 | 7 | 0 | 7 | d = Web Attack Sql Injection | |

Tabel 4.112 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “Benign” terdapat sebanyak 134.537 data diprediksi benar kategori *benign*, 5 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, dan 7 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “Web Attack Brute Force” terdapat sebanyak 13 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 955 data benar diprediksi sebagai *web attack brute force*, 234 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 4 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “Web Attack XSS” terdapat sebanyak 27 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 347 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 148 data benar diprediksi sebagai *web attack xss*. Pada *class* “Web Attack Sql Injection” terdapat sebanyak 3 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 7 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 7 data diprediksi benar kategori *web attack sql injection*.

c. 10 Cross Validation

Tabel 4.113 Hasil Klasifikasi RF 10-CV Split Data 80:20 Dengan Seleksi Fitur

| Class | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area |
|--------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Benign | 1,000 | 0,026 | 1,000 | 1,000 | 1,000 | 1,000 | 0,995 | 1,000 |
| Web Attack Brute Force | 0,779 | 0,003 | 0,729 | 0,779 | 0,753 | 0,751 | 0,997 | 0,797 |
| Web Attack XSS | 0,295 | 0,002 | 0,376 | 0,295 | 0,330 | 0,331 | 0,966 | 0,366 |
| Web Attack Sql Injection | 0,412 | 0,000 | 0,583 | 0,412 | 0,483 | 0,490 | 0,912 | 0,529 |
| Weighted Avg. | 0,995 | 0,025 | 0,995 | 0,995 | 0,995 | 0,979 | 0,995 | 0,996 |

Pada tabel 4.113 merupakan hasil klasifikasi algoritma *Random Forest* pada *tools* Weka menggunakan mode pengujian 10 Cross Validation, menunjukkan hasil 135.638 data dengan prediksi benar terklasifikasi (*Correctly Classified Instances*) dengan akurasi sebesar 99,5187% dan 656 data dengan prediksi salah terklasifikasi (*Incorrectly Classified Instances*) dengan persentasi akurasi sebesar 0,4813%, dengan waktu proses klasifikasi selama 98,43 detik. Untuk *screen shot* hasil pengklasifikasian menggunakan *tools* WEKA dapat dilihat pada bagian lampiran B halaman 184 *point c*.

Tabel 4.114 Confusion Matrix RF 10-CV Split Data 80:20 Dengan Seleksi Fitur

| Prediksi | | | | | |
|----------|-----|-----|---|------------------------------|--------|
| a | b | c | d | | |
| 134537 | 3 | 9 | 0 | a = Benign | Aktual |
| 14 | 940 | 247 | 5 | b = Web Attack Brute Force | |
| 28 | 340 | 154 | 0 | c = Web Attack XSS | |
| 3 | 7 | 0 | 7 | d = Web Attack Sql Injection | |

Tabel 4.114 merupakan *confusion matrix* yang didapat dari perhitungan yang dilakukan pada Weka, dari data yang ada diketahui bahwa pada *class* “*Benign*” terdapat sebanyak 134.537 data diprediksi benar kategori *benign*, 3 data salah dikategorikan sebagai *web attack brute force* seharusnya *benign*, dan 9 data salah dikategorikan *web attack xss* seharusnya *benign*. Pada *class* “*Web Attack Brute Force*” terdapat sebanyak 14 data salah dikategorikan sebagai *benign* seharusnya *web attack brute force*, 940 data benar diprediksi sebagai *web attack brute force*, 247 data salah dikategorikan *web attack xss* seharusnya *web attack brute force* dan 5 data salah dikategorikan sebagai *web attack sql injection* seharusnya *web attack brute force*. Pada *class* “*Web Attack XSS*” terdapat sebanyak 28 data salah dikategorikan sebagai *benign* seharusnya *web attack xss*, 340 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack xss*, dan 154 data benar diprediksi sebagai *web attack xss*. Pada *class* “*Web Attack Sql Injection*” terdapat sebanyak 3 data salah dikategorikan sebagai *benign* seharusnya *web attack sql injection*, 7 data salah dikategorikan sebagai *web attack brute force* seharusnya *web attack sql injection* dan 7 data diprediksi benar kategori *web attack sql injection*.

4.3 HASIL PERBANDINGAN 3 MODE VALIDASI PADA SETIAP SPLIT DATA

4.3.1 HASIL PERBANDINGAN TANPA SELEKSI FITUR

Setelah dilakukan analisis klasifikasi pada *tools* WEKA menggunakan *Use Training Set*, *5-Fold Cross Validation* dan *10-Fold Cross Validation*, maka didapatkan akurasi untuk setiap pembagian data sebagai berikut :

Tabel 4.115 Evaluasi Akurasi Split Data 60 : 40 Tanpa Seleksi Fitur

| Algoritma | Mode Validasi | Correctly Classified Instances | | Incorrectly Classified Instances | |
|---------------|---------------------|--------------------------------|----------------|----------------------------------|--------|
| | | Jumlah | (%) | Jumlah | (%) |
| Naive Bayes | Use Training Set | 99423 | 97,2637 | 2797 | 2,7363 |
| | 5 Cross Validation | 98175 | 96,6054 | 3.470 | 3,3946 |
| | 10 Cross Validation | 99244 | 97,0886 | 2.976 | 2.976 |
| C45 | Use Training Set | 101869 | 99,6566 | 351 | 0,3434 |
| | 5 Cross Validation | 101801 | 99,5901 | 419 | 0,4099 |
| | 10 Cross Validation | 101797 | 99,5862 | 423 | 0,4138 |
| Random Forest | Use Training Set | 102220 | 100 | 0 | 0 |
| | 5 Cross Validation | 101754 | 99,5441 | 466 | 0,4559 |
| | 10 Cross Validation | 101758 | 99,548 | 462 | 0,452 |

Pada tabel 4.115 merupakan hasil perbandingan akurasi pada split data 60:40 tanpa seleksi fitur dengan menggunakan 3 mode pengujian (validasi), didapatkan tingkat akurasi tertinggi berada pada algoritma *Random Forest* dengan menggunakan *Use Training Set* dengan persentasi akurasi 100 % untuk *Correctly*

Classified Instances dengan jumlah 102.220 data benar terklasifikasi dan 0% untuk *Incorrectly Classified Instances* dengan jumlah 0 data salah terklasifikasi.

Tabel 4.116 Evaluasi Akurasi Weka Split Data 70 : 30 Tanpa Seleksi Fitur

| Algoritma | Mode Validasi | Correctly Classified Instances | | Incorrectly Classified Instances | |
|---------------|---------------------|--------------------------------|----------------|----------------------------------|---------|
| | | Jumlah | (%) | Jumlah | (%) |
| Naive Bayes | Use Training Set | 157856 | 97,0903 | 3470 | 2,90097 |
| | 5 Cross Validation | 115703 | 97,0207 | 3553 | 2,9793 |
| | 10 Cross Validation | 115796 | 97,0987 | 3460 | 2,9013 |
| C45 | Use Training Set | 118857 | 99,6654 | 399 | 0,3346 |
| | 5 Cross Validation | 118772 | 99,5942 | 484 | 0,4058 |
| | 10 Cross Validation | 118781 | 99,6017 | 475 | 0,3983 |
| Random Forest | Use Training Set | 119256 | 100 | 0 | 0 |
| | 5 Cross Validation | 118724 | 99,5539 | 532 | 0,4461 |
| | 10 Cross Validation | 118718 | 99,5489 | 538 | 0,4511 |

Pada tabel 4.116 merupakan hasil perbandingan akurasi pada split data 70:30 tanpa seleksi fitur dengan menggunakan 3 mode pengujian (validasi), dan didapatkan tingkat akurasi tertinggi berada pada algoritma *Random Forest* dengan menggunakan *Use Training Set* dengan persentasi akurasi 100% untuk *Correctly Classified Instances* dengan jumlah 119.256 data benar terklasifikasi dan 0% untuk *Incorrectly Classified Instances* dengan jumlah 0 data salah terklasifikasi.

Tabel 4.117 Evaluasi Akurasi Weka Split Data 80 : 20 Tanpa Seleksi Fitur

| Algoritma | Mode Validasi | Correctly Classified Instances | | Incorrectly Classified Instances | |
|---------------|---------------------|--------------------------------|----------------|----------------------------------|--------|
| | | Jumlah | (%) | Jumlah | (%) |
| Naive Bayes | Use Training Set | 132558 | 97,2589 | 3736 | 2,7411 |
| | 5 Cross Validation | 132349 | 97,1055 | 3945 | 2,8945 |
| | 10 Cross Validation | 132412 | 97,1517 | 3882 | 2,8483 |
| C45 | Use Training Set | 135837 | 99,6647 | 457 | 0,3353 |
| | 5 Cross Validation | 135775 | 99,6192 | 519 | 0,3808 |
| | 10 Cross Validation | 135778 | 99,6214 | 516 | 0,3786 |
| Random Forest | Use Training Set | 136294 | 100 | 0 | 0 |
| | 5 Cross Validation | 135675 | 99,5458 | 619 | 0,4542 |
| | 10 Cross Validation | 135665 | 99,5385 | 629 | 0,4615 |

Pada tabel 4.117 merupakan hasil perbandingan akurasi pada split data 80:20 tanpa seleksi fitur dengan menggunakan 3 mode pengujian (validasi), dan didapatkan tingkat akurasi tertinggi berada pada algoritma *Random Forest* dengan menggunakan *Use Training Set* dengan persentasi akurasi 100% untuk *Correctly Classified Instances* dengan jumlah 136.294 data benar terklasifikasi dan 0% untuk *Incorrectly Classified Instances* dengan 0 data salah terklasifikasi.

4.3.2 HASIL PERBANDINGAN DENGAN SELEKSI FITUR

Setelah dilakukan analisis klasifikasi pada *tools WEKA* menggunakan *Use Training Set*, *5-Fold Cross Validation* dan *10-Fold Cross Validation*, maka didapatkan akurasi untuk setiap pembagian data sebagai berikut :

Tabel 4.118 Evaluasi Akurasi Weka Split Data 60 : 40 Dengan Seleksi Fitur

| Algoritma | Mode Validasi | Correctly Classified Instances | | Incorrectly Classified Instances | |
|---------------|---------------------|--------------------------------|----------------|----------------------------------|---------|
| | | Jumlah | (%) | Jumlah | (%) |
| Naive Bayes | Use Training Set | 23472 | 22,9622 | 78748 | 77,0378 |
| | 5 Cross Validation | 22999 | 22,4995 | 79221 | 77,5005 |
| | 10 Cross Validation | 23597 | 23,0845 | 78623 | 76,9155 |
| C45 | Use Training Set | 101850 | 99,638 | 370 | 0,362 |
| | 5 Cross Validation | 101801 | 99,5901 | 419 | 0,4099 |
| | 10 Cross Validation | 101794 | 99,5833 | 426 | 0,4167 |
| Random Forest | Use Training Set | 102210 | 99,9902 | 10 | 0,0098 |
| | 5 Cross Validation | 101757 | 99,5471 | 463 | 0,4529 |
| | 10 Cross Validation | 101746 | 99,5363 | 474 | 0,4637 |

Pada tabel 4.118 merupakan hasil perbandingan akurasi pada split data 60:40 dengan seleksi fitur pada 3 mode pengujian (validasi), dan didapatkan tingkat akurasi tertinggi berada pada algoritma *Random Forest* dengan menggunakan *Use Training Set* dengan persentasi akurasi 99,9902 % untuk *Correctly Classified Instances* dengan jumlah 102.210 data benar terklasifikasi dan 0,0098 % untuk *Incorrectly Classified Instances* dengan jumlah 10 data salah terklasifikasi.

Tabel 4.119 Evaluasi Akurasi Weka Split Data 70 : 30 Dengan Seleksi Fitur

| Algoritma | Mode Validasi | Correctly Classified Instances | | Incorrectly Classified Instances | |
|---------------|---------------------|--------------------------------|----------------|----------------------------------|---------|
| | | Jumlah | (%) | Jumlah | (%) |
| Naive Bayes | Use Training Set | 27472 | 23,0362 | 91784 | 76,9638 |
| | 5 Cross Validation | 27731 | 23,2533 | 91525 | 76,7467 |
| | 10 Cross Validation | 27457 | 23,0236 | 91799 | 76,9764 |
| C45 | Use Training Set | 118824 | 99,6378 | 432 | 0,3622 |
| | 5 Cross Validation | 118778 | 99,5992 | 478 | 0,4008 |
| | 10 Cross Validation | 118783 | 99,6034 | 473 | 0,3966 |
| Random Forest | Use Training Set | 119243 | 99,9891 | 13 | 0,0109 |
| | 5 Cross Validation | 118687 | 99,5229 | 569 | 0,4771 |
| | 10 Cross Validation | 118686 | 99,522 | 570 | 0,478 |

Pada tabel 4.119 merupakan hasil perbandingan akurasi pada split data 70:30 dengan menggunakan 3 mode pengujian (validasi), dan didapatkan tingkat akurasi tertinggi berada pada algoritma *Random Forest* dengan menggunakan *Use Training Set* dengan persentasi akurasi 99,9891% untuk *Correctly Classified Instances* dengan jumlah 119.243 data benar terklasifikasi dan 0,0302 % untuk *Incorrectly Classified Instances* dengan jumlah 13 data salah terklasifikasi.

Tabel 4.120 Evaluasi Akurasi Weka Split Data 80 : 20 Dengan Seleksi Fitur

| Algoritma | Mode Validasi | Correctly Classified Instances | | Incorrectly Classified Instances | |
|-------------|---------------------|--------------------------------|---------------|----------------------------------|---------|
| | | Jumlah | (%) | Jumlah | (%) |
| Naive Bayes | Use Training Set | 32032 | 23.5021 | 104262 | 76.4979 |
| | 5 Cross Validation | 31707 | 23.2637 | 104587 | 76.7363 |
| | 10 Cross Validation | 31623 | 23.202 | 104671 | 76.798 |
| J48 | Use Training Set | 135791 | 99.6309 | 503 | 0.3691 |

| | | | | | |
|---------------|---------------------|--------|---------------|-----|--------|
| | 5 Cross Validation | 135756 | 99.6053 | 538 | 0.3947 |
| | 10 Cross Validation | 135759 | 99.6075 | 535 | 0.3925 |
| Random Forest | Use Training Set | 136279 | 99.989 | 15 | 0.011 |
| | 5 Cross Validation | 135647 | 99.5253 | 647 | 0.4747 |
| | 10 Cross Validation | 135638 | 99.5187 | 658 | 0.4813 |

Pada gambar 4.120 merupakan hasil perbandingan akurasi pada split data 80:20 dengan menggunakan 3 mode pengujian (validasi), dan didapatkan tingkat akurasi tertinggi berada pada algoritma *Random Forest* dengan menggunakan *Use Training Set* dengan persentasi akurasi 99,989% untuk *Correctly Classified Instances* dengan jumlah 136.279 data benar terklasifikasi dan 0.011 % untuk *Incorrectly Classified Instances* dengan 15 data salah terklasifikasi.

4.4 KOMPARASI PERFORMA KLASIFIKASI

Untuk menguji performa algoritma klasifikasi dalam mendeteksi anomali trafik dilakukan dengan 3 mode pengujian atau validasi pada setiap split data yang digunakan, untuk menentukan nilai performa yang baik dilakukan komparasi terhadap akurasi, *True Positive Rate (TRP)*, *False Positive Rate (FPR)*, *Precision*, *Recall* dan ROC(AUC), berikut merupakan perbandingan hasil pengujian klasifikasi dengan menggunakan WEKA :

Tabel 4.121 Perbandingan Akurasi

| | Algoritma Klasifikasi | Akurasi | |
|-------------------------|-----------------------|---------------------|----------------------|
| | | Tanpa Seleksi Fitur | Dengan Seleksi Fitur |
| Split Data 60:40 | | | |
| Use Training | NB | 97,267 | 22,9622 |
| | C45 | 99,6566 | 99,638 |
| | RF | 100 | 99,9902 |
| 5 Cross Validation | NB | 96,6054 | 22,4995 |
| | C45 | 99,5901 | 99,5901 |
| | RF | 99,5441 | 99,5471 |
| 10 Cross Validation | NB | 97,0886 | 23,0845 |
| | C45 | 99,5862 | 99,5833 |
| | RF | 99,548 | 99,5363 |
| Split Data 70:30 | | | |
| Use Training | NB | 97,0903 | 23,0362 |
| | C45 | 99,6654 | 99,6378 |
| | RF | 100 | 99,9891 |
| 5 Cross Validation | NB | 97,0207 | 23,2533 |
| | C45 | 99,5942 | 99,5992 |
| | RF | 99,5539 | 99,5229 |
| 10 Cross Validation | NB | 97,0987 | 23,0236 |
| | C45 | 99,6017 | 99,6034 |
| | RF | 99,5489 | 99,522 |
| Split Data 80:20 | | | |
| Use Training | NB | 97,2589 | 23,501 |
| | C45 | 99,6647 | 99,6309 |
| | RF | 100 | 99,989 |
| 5 Cross Validation | NB | 97,1055 | 23,2637 |
| | C45 | 99,6192 | 99,6053 |
| | RF | 99,5458 | 99,5253 |
| 10 Cross Validation | NB | 97,1517 | 23,202 |
| | C45 | 99,6214 | 99,6075 |
| | RF | 99,5385 | 99,5187 |

Tabel 4.121 merupakan tabel komparasi tingkat akurasi dari tiga algoritma klasifikasi, dan didapatkan akurasi tertinggi pada algoritma *Random Forest* dengan menggunakan *Use Training Set* lebih tinggi dibandingkan C45 dan *Naive Bayes*, sedangkan pada pengujian *5 cross validation* dan *10 cross validation*

algoritma C45 lebih unggul dibandingkan *Random Forest* dan *Naive Bayes*. Algoritma *Naive Bayes* memiliki tingkat akurasi terendah dari *Random Forest* dan C45. Penggunaan fitur terpilih dengan menggunakan seleksi fitur *information gain* tidak bisa menjadi dasar bahwa dengan penggunaannya dapat meningkatkan performa dari algoritma, dilihat dari tabel 4.121 nilai akurasi algoritma tanpa seleksi fitur lebih tinggi dibandingkan nilai akurasi dengan menggunakan seleksi fitur, dengan menggunakan fitur hasil seleksi fitur *information gain* algoritma mengalami pengurangan performa akurasi terutama algoritma pada *Naive Bayes* yang pada awalnya (tanpa seleksi fitur) memiliki akurasi sekitar 97% berkurang menjadi 23%.

Tabel 4.122 True Positive Rate (TPR)

| Algoritma Klasifikasi | Mode Pengujian | True Positive Rate | |
|-------------------------|---------------------|---------------------|----------------------|
| | | Tanpa Seleksi Fitur | Dengan Seleksi Fitur |
| Split Data 60:40 | | | |
| Naive Bayes | Use Training | 0,973 | 0,230 |
| | 5 Cross Validation | 0,966 | 0,225 |
| | 10 Cross Validation | 0,971 | 0,231 |
| C45 | Use Training | 0,997 | 0,996 |
| | 5 Cross Validation | 0,996 | 0,996 |
| | 10 Cross Validation | 0,996 | 0,996 |
| Random Forest | Use Training | 1,000 | 1,000 |
| | 5 Cross Validation | 0,995 | 0,995 |
| | 10 Cross Validation | 0,995 | 0,995 |
| Split Data 70:30 | | | |
| Naive Bayes | Use Training | 0,971 | 0,230 |
| | 5 Cross Validation | 0,970 | 0,233 |
| | 10 Cross Validation | 0,971 | 0,230 |
| C45 | Use Training | 0,997 | 0,996 |
| | 5 Cross Validation | 0,996 | 0,996 |
| | 10 Cross Validation | 0,996 | 0,996 |
| Random Forest | Use Training | 1,000 | 1,000 |
| | 5 Cross Validation | 0,996 | 0,995 |
| | 10 Cross Validation | 0,995 | 0,995 |

| Split Data 80:20 | | | |
|-------------------------|---------------------|--------------|--------------|
| Naive Bayes | Use Training | 0,973 | 0,235 |
| | 5 Cross Validation | 0,971 | 0,233 |
| | 10 Cross Validation | 0,972 | 0,232 |
| C45 | Use Training | 0,997 | 0,996 |
| | 5 Cross Validation | 0,996 | 0,996 |
| | 10 Cross Validation | 0,996 | 0,996 |
| Random Forest | Use Training | 1,000 | 1,000 |
| | 5 Cross Validation | 0,995 | 0,995 |
| | 10 Cross Validation | 0,995 | 0,995 |

Pada tabel 4.122 merupakan tabel komparasi *True Positive Rate* (TPR) dari masing-masing algoritma pada 3 mode pengujian (validasi), diperoleh hasil TPR tertinggi didapatkan pada algoritma *Random Forest* pada *Use Training Set* dengan nilai $TPR = 1$ disetiap pembagian data (split data), sedangkan pada pengujian *5-fold cross validation* dan *10-fold cross validation* algoritma *Random Forest* dan C45 memiliki nilai TRP yang sama dan lebih tinggi dibanding TPR yang dihasilkan *Naive Bayes*. Untuk penggunaan seleksi fitur, tidak terlalu berpengaruh terhadap hasil TPR dari algoritma *Random Forest* dan C45, akan tetapi berpengaruh terhadap hasil yang didapat dari algoritma *Naive Bayes* dimana nilai TPR yang dihasilkan menggunakan seleksi fitur menjadi sangat rendah dibandingkan nilai TPR tanpa seleksi fitur atau menggunakan keseluruhan fitur pada *dataset*.

Tabel 4.123 False Positive Rate (FPR)

| Algoritma Klasifikasi | Mode Pengujian | False Positive Rate | |
|-------------------------|---------------------|---------------------|----------------------|
| | | Tanpa Seleksi Fitur | Dengan Seleksi Fitur |
| Split Data 60:40 | | | |
| Naive Bayes | Use Training | 0,012 | 0,036 |
| | 5 Cross Validation | 0,010 | 0,039 |
| | 10 Cross Validation | 0,012 | 0,038 |

| | | | |
|-------------------------|---------------------|--------------|--------------|
| C45 | Use Training | 0,006 | 0,020 |
| | 5 Cross Validation | 0,013 | 0,023 |
| | 10 Cross Validation | 0,013 | 0,021 |
| Random Forest | Use Training | 0,000 | 0,006 |
| | 5 Cross Validation | 0,028 | 0,029 |
| | 10 Cross Validation | 0,026 | 0,028 |
| Split Data 70:30 | | | |
| Naive Bayes | Use Training | 0,011 | 0,035 |
| | 5 Cross Validation | 0,011 | 0,036 |
| | 10 Cross Validation | 0,011 | 0,035 |
| C45 | Use Training | 0,004 | 0,019 |
| | 5 Cross Validation | 0,010 | 0,021 |
| | 10 Cross Validation | 0,012 | 0,021 |
| Random Forest | Use Training | 0,000 | 0,006 |
| | 5 Cross Validation | 0,025 | 0,025 |
| | 10 Cross Validation | 0,025 | 0,025 |
| Split Data 80:20 | | | |
| Naive Bayes | Use Training | 0,011 | 0,032 |
| | 5 Cross Validation | 0,010 | 0,032 |
| | 10 Cross Validation | 0,009 | 0,033 |
| C45 | Use Training | 0,003 | 0,019 |
| | 5 Cross Validation | 0,014 | 0,021 |
| | 10 Cross Validation | 0,013 | 0,021 |
| Random Forest | Use Training | 0,000 | 0,006 |
| | 5 Cross Validation | 0,024 | 0,024 |
| | 10 Cross Validation | 0,021 | 0,025 |

Pada tabel 4.123 merupakan komparasi nilai *False Positive Rate* (FPR) pada masing-masing algoritma, dari hasil eksperimen yang telah dilakukan diperoleh nilai FPR terendah pada algoritma *Random Forest* dengan nilai FPR = 0 dengan mode pengujian *Use Training Set* tanpa seleksi fitur, serta FPR = 0,006 dengan seleksi fitur. Pada mode pengujian *5-fold cross validation* dan *10-fold cross validation*, C45 dan *Naive Bayes* memiliki nilai FPR yang hampir sama dan lebih rendah dibandingkan *Random Forest*. Dengan menggunakan seleksi fitur nilai FPR pada setiap algoritma menjadi semakin tinggi dibandingkan tanpa

menggunakan seleksi fitur, dimana semakin tinggi FPR atau mendekati 1 maka semakin banyak *error* atau kesalahan dalam pengklasifikasian data. *Random Forest*, *C45* dan *Naive Bayes* mencapai performa FPR terbaik pada skenario klasifikasi tanpa seleksi fitur dan dapat dilihat pada tabel 4.123.

Tabel 4.124 Precision

| Algoritma Klasifikasi | Mode Pengujian | Precision | |
|-------------------------|---------------------|---------------------|----------------------|
| | | Tanpa Seleksi Fitur | Dengan Seleksi Fitur |
| Split Data 60:40 | | | |
| Naive Bayes | Use Training | 0,990 | 0,987 |
| | 5 Cross Validation | 0,989 | 0,986 |
| | 10 Cross Validation | 0,989 | 0,987 |
| C45 | Use Training | 0,997 | 0,997 |
| | 5 Cross Validation | 0,995 | 0,995 |
| | 10 Cross Validation | 0,995 | 0,995 |
| Random Forest | Use Training | 1,000 | 1,000 |
| | 5 Cross Validation | 0,995 | 0,995 |
| | 10 Cross Validation | 0,995 | 0,955 |
| Split Data 70:30 | | | |
| Naive Bayes | Use Training | 0,988 | 0,987 |
| | 5 Cross Validation | 0,988 | 0,987 |
| | 10 Cross Validation | 0,989 | 0,987 |
| C45 | Use Training | 0,997 | 0,996 |
| | 5 Cross Validation | 0,995 | 0,995 |
| | 10 Cross Validation | 0,996 | 0,996 |
| Random Forest | Use Training | 1,000 | 1,000 |
| | 5 Cross Validation | 0,995 | 0,995 |
| | 10 Cross Validation | 0,995 | 0,995 |
| Split Data 80:20 | | | |
| Naive Bayes | Use Training | 0,989 | 0,987 |
| | 5 Cross Validation | 0,988 | 0,987 |
| | 10 Cross Validation | 0,988 | 0,987 |
| C45 | Use Training | 0,997 | 0,997 |
| | 5 Cross Validation | 0,996 | 0,996 |
| | 10 Cross Validation | 0,996 | 0,996 |
| Random Forest | Use Training | 1,000 | 1,000 |
| | 5 Cross Validation | 0,995 | 0,995 |
| | 10 Cross Validation | 0,995 | 0,995 |

Pada tabel 4.124 merupakan komparasi nilai *precision* masing-masing algoritma dengan 3 mode pengujian (validasi). Data pada tabel memperlihatkan nilai *precision* tertinggi berada pada algoritma *Random Forest* dengan *Use Training Set*, sedangkan pada pengujian *5-fold cross validation* dan *10-fold cross validation* *Random Forest* dan *C45* memiliki nilai *precision* yang sama serta lebih tinggi dibandingkan dengan *Naive Bayes*. Untuk penggunaan fitur baik tanpa seleksi fitur atau bahkan dengan seleksi fitur, tidak terlalu memberikan pengaruh terhadap nilai *precision*.

Tabel 4.125 Recall

| Algoritma Klasifikasi | Mode Pengujian | Recall | |
|-------------------------|---------------------|---------------------|----------------------|
| | | Tanpa Seleksi Fitur | Dengan Seleksi Fitur |
| Split Data 60:40 | | | |
| Naive Bayes | Use Training | 0,973 | 0,230 |
| | 5 Cross Validation | 0,966 | 0,225 |
| | 10 Cross Validation | 0,971 | 0,231 |
| C45 | Use Training | 0,997 | 0,996 |
| | 5 Cross Validation | 0,996 | 0,996 |
| | 10 Cross Validation | 0,996 | 0,996 |
| Random Forest | Use Training | 1,000 | 1,000 |
| | 5 Cross Validation | 0,995 | 0,995 |
| | 10 Cross Validation | 0,995 | 0,995 |
| Split Data 70:30 | | | |
| Naive Bayes | Use Training | 0,971 | 0,230 |
| | 5 Cross Validation | 0,970 | 0,233 |
| | 10 Cross Validation | 0,971 | 0,230 |
| C45 | Use Training | 0,997 | 0,996 |
| | 5 Cross Validation | 0,996 | 0,996 |
| | 10 Cross Validation | 0,996 | 0,996 |
| Random Forest | Use Training | 1,000 | 1,000 |
| | 5 Cross Validation | 0,996 | 0,995 |
| | 10 Cross Validation | 0,995 | 0,995 |
| Split Data 80:20 | | | |
| Naive Bayes | Use Training | 0,973 | 0,235 |
| | 5 Cross Validation | 0,971 | 0,233 |

| | | | |
|---------------|---------------------|--------------|--------------|
| | 10 Cross Validation | 0,972 | 0,232 |
| C45 | Use Training | 0,997 | 0,996 |
| | 5 Cross Validation | 0,996 | 0,996 |
| | 10 Cross Validation | 0,996 | 0,996 |
| Random Forest | Use Training | 1,000 | 1,000 |
| | 5 Cross Validation | 0,995 | 0,995 |
| | 10 Cross Validation | 0,995 | 0,955 |

Pada tabel 4.125 memperlihatkan komparasi nilai *Recall* pada masing-masing algoritma yang digunakan dengan 3 mode pengujian (validasi) dan didapatkan hasil nilai *recall* tertinggi berada pada algoritma *Random Forest* dengan mode pengujian *Use Training Set*, pada pengujian *5-fold cross validation* dan *10-fold validation Random Forest* dan C45 memiliki nilai yang sama dan lebih unggul dibanding *Naive Bayes*. Performa *recall* tertinggi didapat dengan pengklasifikasian tanpa seleksi fitur, penggunaan seleksi fitur mengurangi nilai *recall* yang dihasilkan terutama pada algoritma *Naive Bayes* yang semula memiliki nilai sekitar 0,970 menggunakan 77 atribut awal (tanpa seleksi fitur) mengalami pengurangan menjadi sekitar 0,230 (dengan seleksi fitur).

Tabel 4.126 ROC (AUC)

| Algoritma Klasifikasi | Mode Pengujian | ROC Area | |
|-------------------------|---------------------|---------------------|----------------------|
| | | Tanpa Seleksi Fitur | Dengan Seleksi Fitur |
| Split Data 60:40 | | | |
| Naive Bayes | Use Training | 0,997 | 0,941 |
| | 5 Cross Validation | 0,995 | 0,938 |
| | 10 Cross Validation | 0,996 | 0,940 |
| C45 | Use Training | 0,999 | 0,995 |
| | 5 Cross Validation | 0,997 | 0,993 |
| | 10 Cross Validation | 0,996 | 0,994 |
| Random Forest | Use Training | 1,000 | 1,000 |
| | 5 Cross Validation | 1,000 | 0,994 |
| | 10 Cross Validation | 1,000 | 0,994 |

| Split Data 70:30 | | | |
|-------------------------|---------------------|--------------|-------|
| Naive Bayes | Use Training | 0,996 | 0,944 |
| | 5 Cross Validation | 0,994 | 0,941 |
| | 10 Cross Validation | 0,994 | 0,942 |
| C45 | Use Training | 0,999 | 0,995 |
| | 5 Cross Validation | 0,993 | 0,992 |
| | 10 Cross Validation | 0,995 | 0,991 |
| Random Forest | Use Training | 1,000 | 1,000 |
| | 5 Cross Validation | 1,000 | 0,995 |
| | 10 Cross Validation | 1,000 | 0,995 |
| Split Data 80:20 | | | |
| Naive Bayes | Use Training | 0,994 | 0,931 |
| | 5 Cross Validation | 0,996 | 0,942 |
| | 10 Cross Validation | 0,995 | 0,942 |
| C45 | Use Training | 0,999 | 0,995 |
| | 5 Cross Validation | 0,996 | 0,993 |
| | 10 Cross Validation | 0,996 | 0,992 |
| Random Forest | Use Training | 1,000 | 1,000 |
| | 5 Cross Validation | 1,000 | 0,974 |
| | 10 Cross Validation | 1,000 | 0,995 |

Pada tabel 4.126 didapatkan nilai *ROC (Receiver Operating Characteristic)* tertinggi didapat pada algoritma *Random Forest* dengan nilai 1,00 dikategorikan “*Excellent Classification*”. Algoritma *Random Forest*, *C45* dan *Naive Bayes* termasuk dalam kategori *Excellent Classification* berdasarkan tabel nilai *ROC (AUC)* yang telah disajikan pada tabel 3.2 apabila nilai *AUC* bernilai 0,90 – 1,00 maka tergolong *Excellent Classification*. Nilai *ROC* mengalami penurunan pada klasifikasi menggunakan seleksi fitur, nilai terbaik didapat pada klasifikasi tanpa seleksi fitur.

Tabel 4.127 Waktu Proses

| Algoritma Klasifikasi | Mode Pengujian | Waktu Proses (detik) | |
|-------------------------|---------------------|----------------------|----------------------|
| | | Tanpa Seleksi Fitur | Dengan Seleksi Fitur |
| Split Data 60:40 | | | |
| Naive Bayes | Use Training | 29,06 | 5,64 |
| | 5 Cross Validation | 3,1 | 0,55 |
| | 10 Cross Validation | 2,69 | 0,58 |
| C45 | Use Training | 48,55 | 5,1 |
| | 5 Cross Validation | 49,14 | 3,84 |
| | 10 Cross Validation | 46,2 | 4,02 |
| Random Forest | Use Training | 153,94 | 72,16 |
| | 5 Cross Validation | 147,78 | 63,14 |
| | 10 Cross Validation | 148 | 65,56 |
| Split Data 70:30 | | | |
| Naive Bayes | Use Training | 16,52 | 5,63 |
| | 5 Cross Validation | 3,12 | 0,64 |
| | 10 Cross Validation | 3,62 | 0,58 |
| C45 | Use Training | 36,4 | 5,86 |
| | 5 Cross Validation | 33,08 | 4,76 |
| | 10 Cross Validation | 35,27 | 4,79 |
| Random Forest | Use Training | 120,15 | 92,22 |
| | 5 Cross Validation | 108,63 | 82,04 |
| | 10 Cross Validation | 99,46 | 83,33 |
| Split Data 80:20 | | | |
| Naive Bayes | Use Training | 45,98 | 5,06 |
| | 5 Cross Validation | 12,31 | 0,93 |
| | 10 Cross Validation | 12,07 | 1,31 |
| C45 | Use Training | 75,48 | 8,54 |
| | 5 Cross Validation | 70,71 | 6,4 |
| | 10 Cross Validation | 70,73 | 6,42 |
| Random Forest | Use Training | 215,44 | 102,47 |
| | 5 Cross Validation | 202,3 | 94,47 |
| | 10 Cross Validation | 199,34 | 98,43 |

Pada tabel 4.127 merupakan tabel komparasi waktu proses (dalam satuan detik) klasifikasi dari masing-masing algoritma, *Random Forest* memiliki waktu proses paling lama dibandingkan C45 dan *Naive Bayes*. Meskipun *Random Forest*

unggul dalam perbandingan akurasi, TPR, FPR, *Precision*, *Recall* dan ROC tetapi memiliki komputasi yang relatif lama. *Naive Bayes* memiliki waktu proses paling cepat dibanding *Random Forest* dan C45. Penggunaan seleksi fitur dapat mengurangi waktu proses karena fitur yang diklasifikasi berkurang. Pembagian data atau jumlah data yang dipakai juga mempengaruhi lamanya waktu komputasi, dari pembagian data yang dilakukan waktu komputasi terbaik terdapat pada split data 70 : 30.