

KLASIFIKASI TRAFIK INTERNET KAMPUS BERBASIS PROTOKOL JARINGAN MENGGUNAKAN ALGORITMA NAÏVE BAYES

I Made Bayu Dibawan¹, I Made Oka Widyantara², Linawati³

Jurusan Teknik Elektro dan Komputer, Fakultas Teknik, Universitas Udayana
Email: imadebayudibawan@gmail.com¹, okawidyantara@unud.ac.id², linawati@unud.ac.id³

ABSTRAK

Komunikasi berbasis teknologi jaringan komputer merupakan teknologi yang sudah banyak digunakan. Komunikasi pada jaringan komputer, menggunakan aturan komunikasi yang sesuai dikenal dengan nama *network protocol*. Penggunaan *network protocol* dalam komunikasi di jaringan komputer, terkadang menuntut adanya prioritas komunikasi. Dasar dari pemberian prioritas adalah dengan menganalisa trafik jaringan internet dan pemberian prioritas yang dilakukan pada trafik jaringan. Prioritas tersebut bertujuan untuk mengoptimalkan penggunaan jaringan internet. Permasalahan yang didapat pada Kampus UNUD Jimbaran, antara lain data yang dikirim lambat, rusak dan bahkan tidak sampai ke tujuan dikarenakan trafik jaringan tersebut belum optimal. Pada penelitian ini melakukan klasifikasi terhadap data *capture* trafik jaringan yang diproses menggunakan Algoritma *Naïve Bayes*. Aplikasi yang digunakan untuk *capture* trafik jaringan yaitu aplikasi *Wireshark*. Hasil observasi terhadap dataset jaringan melalui proses perhitungan menggunakan Algoritma *Naïve Bayes* memiliki tingkat keakuratan yang sangat tinggi, yaitu sebesar 92,34 %. Hasil proses klasifikasi dijadikan sebagai acuan dalam penentuan pemberian prioritas terhadap *network protocol* yang sering digunakan dalam komunikasi jaringan.

Kata Kunci: *Trafik Jaringan, Algoritma Naïve Bayes*

ABSTRACT

Communication based on network computer technology is a technology that already widely used. Communications on computer networks, using appropriate communication rules known as the network protocol. The usage of network protocol in communication of computer networks, sometimes demand for communication priorities. The basic of priorities is conducted by analyzing the internet network traffic and the given priority to network traffic. The priority aims to optimize the use of the Internet network. Problems encountered in Udayana University Jimbaran such as the failure and tardiness of data and also the data has not reach its destination due to the not optimal use of network traffic. The study undertakes a classification of the data capture network traffic that is processed using Naïve Bayes Algorithm. The application used to capture the network traffic is Wireshark application. The observation' results of the network dataset through calculation process using Naïve Bayes algorithm has a very high level of accuracy, that reaches 92.34%. The result of the classification process will be used as a reference in determining the priority of the network protocol that is often used in network communications.

Keywords: *Network protocol, Naïve Bayes, network capture*

1. PENDAHULUAN

Pemanfaatan teknologi jaringan sebagai media komunikasi data terus meningkat dan berkembang terutama dalam bidang jaringan internet. Jaringan komputer saat ini sangat dibutuhkan untuk menghubungkan berbagai instansi, seperti pemerintahan, perkantoran dan kampus, Seiring dengan tingginya tingkat kebutuhan serta pemanfaatan teknologi jaringan menyebab-

kan para pengguna menginginkan sebuah jaringan yang maksimal baik dari segi efisiensi maupun tingkat keamanan. seperti permasalahan yang ada pada kampus Unud jimbaran antara lain data yang dikirimkan lambat, rusak dan bahkan tidak sampai ke tujuan dikarenakan trafik jaringan tersebut belum optimal. Oleh sebab itu, analisis terhadap suatu trafik jaringan adalah salah satu cara mengetahui penggunaan protokol komunikasi jaringan. Se-

hingga dapat menjadi dasar untuk penentuan prioritas suatu trafik jaringan.

Terkait dengan klasifikasi trafik internet, beberapa penelitian telah dilakukan dengan menggunakan metode data mining, filterisasi untuk keamanan aplikasi menggunakan *Algoritma Naïve Bayes* yang diajukan oleh [1], adalah teknik klasifikasi menggunakan algoritma *Naïve Bayes* untuk memfilter layanan email pada kelas email spam dan kelas email *valid/legitimate*. *Naïve Bayes Filter* mengklasifikasi email dengan menghitung probabilitas email berdasarkan nilai probabilitas token pada database filter yang dibangun.

Selanjutnya, penerapan *Algoritma Naïve Bayes* didalam Manajemen prioritas layanan (QoS) oleh [2], dalam penelitian ini teknik klasifikasi *Naïve Bayes* menggunakan data set yang terdiri dari beberapa kumpulan kategori klasifikasi trafik jaringan seperti (*length, port number, dan time*). Data set kemudian dibandingkan dengan data yang sudah terpola dengan melalui tahap inialisasi asumsi.

Penerapan *Algoritma Naïve Bayes* didalam Manajemen prioritas layanan QoS diajukan oleh [3], dimana dalam penelitian ini pemberian prioritas QoS yaitu dengan menganalisa terhadap data trafik jaringan internet menggunakan *Algoritma Naïve Bayes*,

Selanjutnya Analisis konten pada aplikasi diajukan oleh [4], adalah penerapan algoritma *Naïve Bayes* dengan mengklasifikasi beragam topik pembicaraan yang populer pada aplikasi twitter dengan cara membandingkan setiap fitur yang dimiliki oleh tiap kategori yang telah ditentukan. Selanjutnya, penerapan *Algoritma Naïve Bayes* untuk mengklasifikasi data trafik internet diajukan oleh [5], dimana dalam penelitian ini pengambilan data trafik internet yang dilakukan adalah trafik internet kampus menggunakan tiga variabel yaitu *destination, protocol, dan length*. Tujuan dari penelitian ini adalah mengklasifikasi data trafik internet kampus sehingga dari klasifikasi tersebut dapat diketahui *destination network, protocol dan lebar bandwidth* yang banyak diakses pada waktu tertentu.

Pemaparan diatas menunjukkan bahwa penerapan algoritma *Naïve Bayes* adalah solusi yang bisa diterapkan pada mekanisme klasifikasi trafik internet untuk penentuan prioritas layanan QoS. Selanjutnya, paper ini akan mengevaluasi lebih lanjut ki-

nerja penerapan algoritma *Naïve Bayes* untuk klasifikasi trafik internet kampus dengan studi kasus pada trafik internet universitas Udayana. Sasarannya adalah meningkatkan kinerja pada jaringan unud dan sebagai acuan *administrator* didalam manajemen jaringan. Mekanisme klasifikasi yang digunakan adalah menentukan probabilitas trafik internet unud yang *dicapture* pada *router* jaringan unud dengan menggunakan aplikasi *wireshark*.

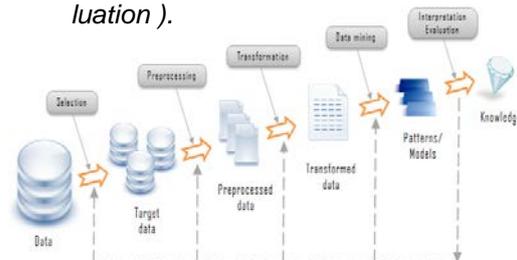
2. TINJAUAN PUSTAKA

2.1. Data Mining

Data Mining secara singkat dapat diartikan sebagai mengekstraksi atau menggali pengetahuan dari data yang berjumlah besar. Sedangkan dari beberapa sumber, definisi secara umum mengenai *Data Mining* adalah proses menemukan sesuatu yang bermakna dari suatu korelasi baru, pola dan tren yang ada dengan cara memilah-memilah data berukuran besar. Sedangkan definisi secara umum dari beberapa sumber, *Data Mining* adalah proses menemukan sesuatu yang bermakna dari suatu korelasi baru, pola dan tren yang ada dengan cara memilah-memilah data berukuran besar. *Data Mining* adalah analisis pengamatan data set untuk menemukan hubungan yang tidak berduga. *Data Mining* merupakan bidang ilmu interdisipliner yang menyatakan teknik pembelajaran dari mesin. *Data Mining* diartikan sebagai suatu proses ekstraksi informasi berguna dan potensial dari sekumpulan data yang terdapat secara implisit dalam suatu basis data.

Tahapan data mining dapat dilihat pada Gambar 1. Pada Gambar 1 mengilustrasikan bagaimana tiap proses bersifat interaktif dimana pemakaian terlibat langsung atau dengan perantara knowledge base.

- 1) Pembersihan data (*Selection*).
- 2) Pra pemrosesan (*Preprocessing*).
- 3) Transformasi data (*Transformation*).
- 4) Teknik Data Mining (*Data Mining*).
- 5) Evaluasi pola (*Interpretation / Evaluation*).



Gambar 1. Tahapan data minin

2.2. Klasifikasi

Klasifikasi merupakan proses menemukan sebuah model atau fungsi yang mendeskripsikan dan membedakan data kedalam kelas-kelas. Klasifikasi melibatkan proses pemeriksaan karakteristik dari objek dan memasukkan objek kedalam salah satu kelas.

Secara umum klasifikasi merupakan sebuah model dibuat untuk menggambarkan himpunan kelas atau konsep data yang telah ditentukan sebelumnya. Model tersebut dibangun dengan menganalisis *record-record* diasumsikan kedalam satu kelas yang telah ditentukan sebelumnya, yang dinamakan atribut kelas.

2.3. Algoritma Naïve Bayes

Algoritma *Naïve Bayes* merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. *Naïve Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris *Thomas Bayes*, yaitu memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai *Teorema Bayes*. Persamaan dari *Teorema Bayes* adalah sebagai berikut [6].

$$P(H|X) = \frac{P(X|H) * P(H)}{P(X)} \quad (1)$$

Keterangan :

- X : data dengan kelas yang belum diketahui
- H : Hipotesa data X suatu kelas
- $P(H|X)$: Probabilitas hipotesa H berdasarkan kondisi X (*Posterior*)
- $P(H)$: Probabilitas hipotesa H (*prior*)
- $P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesa H
- $P(X)$: Probabilitas X

2.4. Pengukuran Kinerja Klasifikasi

Pengujian kinerja sistem klasifikasi pada algoritma *Naive Bayes* ini dapat dilakukan dengan menggunakan *confusion matrix*. *Confusion matrix* ini alat yang berguna untuk menganalisis seberapa baiknya klasifikasi yang kita pakai dapat mengenali pola dari kelas yang berbeda seperti ditunjukkan pada Table 1.

Tabel 1. *Confusion Matrix*s

		Kelas Hasil Klasifikasi	
		f_0	f_1
Kelas Asli	f_0	f_{00}	f_{01}
	f_1	f_{10}	f_{11}

Keterangan:

f_{00} = jumlah prediksi yang tepat bahwa

instance bersifat negatif

f_{01} = jumlah prediksi yang salah bahwa

instance bersifat positif

f_{10} = jumlah prediksi yang salah bahwa

instance bersifat negatif

f_{11} = jumlah prediksi yang tepat bahwa

Instance bersifat positif.

Jumlah data yang diklasifikasi secara benar maupun diklasifikasi secara salah, dapat dihitung hasil akurasi, dan laju error dari prediksi sesuai Persamaan (2) dan (3) [7].

$$Akurasi = \frac{\text{Jumlah data yang diprediksi secara benar}}{\text{Jumlah prediksi yang dilakukan}}$$

$$= \frac{f_{11} + f_{00}}{f_{11} + f_{10} + f_{01} + f_{00}} \quad (2)$$

Untuk menghitung laju error (kesalahan prediksi) digunakan formula

$$Laju Error = \frac{\text{Jumlah prediksi yang salah}}{\text{Jumlah prediksi keseluruhan}}$$

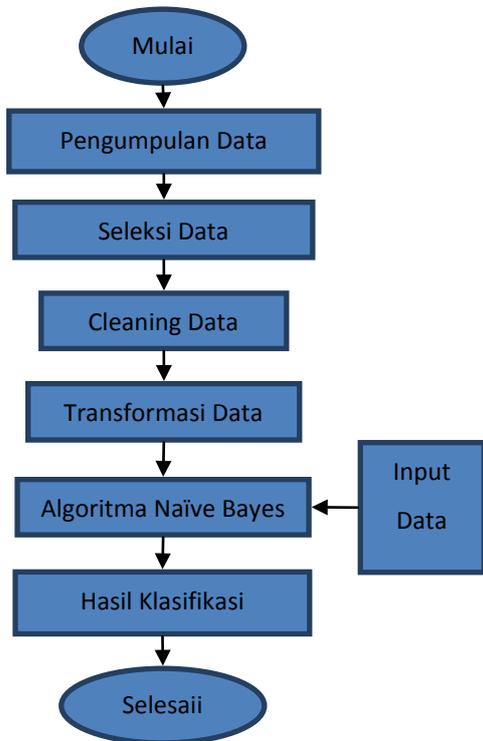
$$= \frac{f_{10} + f_{01}}{f_{11} + f_{10} + f_{01} + f_{00}} \quad (3)$$

3. METODELOGI PENELITIAN

Pada tahap ini merupakan analisis bagaimana penerapan algoritma *Naïve Bayes*. *Naïve Bayes* merupakan pengklasifikasian dengan metode propabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris *Thomas Bayes*, yaitu memprediksi dimasa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai *Teorema Bayes*.

Dalam proses klasifikasi, *Naive Bayes* akan menghitung nilai probabilitas kemunculan kelas label, Gambar 2 meru-

pakan gambaran umum istem klasifikasi menggunakan algoritma *Naïve Bayes*.



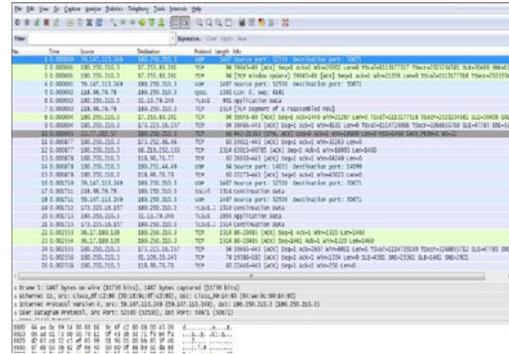
Gambar 2 Metodologi Penelitian

Berdasarkan tahapan pada Gambar 2, maka klasifikasi menggunakan algoritma *Naïve Bayes* dapat dijabarkan secara lebih terperinci sebagai berikut.

A. Pengambilan Data Trafik Internet

Pengambilan data trafik jaringan menghasilkan kurang lebih hingga puluhan juta *record traffic*. Jumlah yang dihasilkan tiap recordnya tidak sama. Dikarenakan ketidaksamaan model komunikasi dalam jaringan komputer yang dilakukan oleh user.

Model *capturing* data *traffic* dapat dilihat pada Gambar 3.



Gambar 3 Capturing Wireshark

B. Data Filtering

Data filtering merupakan pemilihan atribut data yang akan digunakan sebagai proses perhitungan klasifikasi. Hasil *capturing* data *traffic* berupa data mentah diproses dan difilter menggunakan Tool *Pentaho*. Atribut data yang akan dihitung nanti menggunakan algoritma *Naïve Bayes* hanya menggunakan atribut Protokol, Length dan jumlah *Counting*.

Dari informasi 3 atribut mampu dijadikan sebagai acuan pemberian prioritas dengan menganalisa jumlah *counting* (kemunculan data yang sama).

C. Klasifikasi Naïve Bayes

Pada tahapan pembentukan klasifikasi menggunakan *Naïve Bayes* adalah dengan menghitung probabilitas dan *statistic*, yaitu memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai *Teorema Bayes*.

Data trafik internet yang dipetakan dalam algoritma *Naïve Bayes* sebagai kelas klasifikasi dapat dilihat pada Tabel 2. berikut.

Tabel 2. Atribut Klas Klasifikasi

Protokol	Length Range	Counting Range	Prioritas
ARP, DHCP, DHCPv6, DNS, HTTP, ICMP, ICMPv6, IGMP, MDNS, MNDP, NBNS, NTP, SSDP, SSHv2, TCP	0–32, 33–64, 65–128, 129–256, 257–512, 513–1024, 1025–2048	0–500, 501–1000, 1501–2000, 2001–2500, 2501–3000, x > 3000	Rendah, Menengah dan Tinggi

Penelitian ini klasifikasi yang dilakukan yaitu dengan atribut *protocol*, *length*, *counting*, *length range* dan *counting range* dan prioritas. Untuk melakukan inisialisasi, penelitian ini mengusulkan model untuk menentukan kelas label yaitu sebagai berikut .

Tabel 3. Model Klas Label

Range	Prioritas		
	Rendah	Menengah	Tinggi
Length	33 – 64	129 – 256	513 - 1024
	65 - 128	257 - 512	1025 - 2048
Counting	0 - 500	1501 – 2000	2501 - 3000
	501 - 1000	2001 – 2500	x > 3000
	1001 - 1500		

Pada Tabel 3, yaitu merupakan model kelas label yang akan digunakan untuk melakukan inisialisasi pembentukan data latih, rentang batas minimum dan maksimum setiap *prioritas* didasarkan pada hasil data trafik jaringan yang didapat dari hasil *capture* trafik jaringan di kampus unud. Dari aturan model tersebut penelitian ini mengajukan *inisialisasi* label kelas dalam pembentukan data latih yang dilakukan secara manual, seperti ditunjukkan pada Table 4.

Pada Tabel 4. merupakan hasil dari inisialisasi yang dilakukan secara manual. Dasar dari pemberian inisial terhadap masing-masing protokol melalui hasil wawancara dari administrator jaringan. Sebagai acuan dalam pemberian batas *Length Range* dan *Counting Range* adalah melalui data *capturing* setelah melalui proses data filtering.

Tabel 4. Inisialisasi Kelas Label

Protokol	Length	Counting	Length Range	Counting Range	Prioritas
ARP	60	2884	33-64	2501 - 3000	Tinggi
ARP	118	1544	65-128	1501 - 2000	Tinggi
DHCP	346	2	257- 512	0 - 500	Rendah
DHCP	362	5	257-512	0 - 500	Rendah
DHCP	414	6	257 - 512	0 - 500	Rendah
DHCP	426	6	257 - 512	0 - 500	Rendah
DHCPv6	112	1	65 - 128	0 - 500	Menengah
DHCPv6	142	1	129 - 256	0 - 500	Menengah
DHCPv6	146	30	129 - 256	0 - 500	Menengah
DHCPv6	148	13	129 - 256	0 - 500	Menengah
DHCPv6	150	39	129 - 256	0 - 500	Menengah
DHCPv6	152	10	129 - 256	0 - 500	Menengah
DHCPv6	154	6	129 - 256	0 - 500	Menengah
DHCPv6	156	2	129 - 256	0 - 500	Menengah
DHCPv6	190	2	129 - 256	0 - 500	Menengah
DHCPv6	224	4	129 - 256	0 - 500	Menengah
DHCPv6	227	32	129 - 256	0 - 500	Menengah
DNS	60	4	33 - 64	0 - 500	Rendah
DNS	64	5	33 - 64	0 - 500	Rendah
DNS	66	71	65 - 128	0 - 500	Rendah
DNS	68	275	65 - 128	0 - 500	Rendah
DNS	70	727	65 - 128	501 - 1000	Rendah

4. HASIL DAN PEMBAHASAN

4.1. EVALUASI MEKANISME PELATIHAN

Mekanisme yang digunakan adalah setengah data digunakan untuk data latih dan setengah digunakan untuk data uji. Detail perhitungan manual dapat dilihat pada Tabel 5, yaitu dengan *length* 414, *counting* 6, dan proritias belum diketahui (?).

Tabel 5. Sample Data

Protokol	Length	Counting	Prioritas
DHCP	414	6	?

a) Mekanisme pelatihan

- 1) Menghitung jumlah setiap kelas berdasarkan klasifikasi yang terbentuk (*prior probability*):

$$P(\text{Rendah}) : 2200 / 3447 = 0.639$$

$$P(\text{Menengah}) : 1113 / 3447 = 0.323$$

$$P(\text{Tinggi}) : 132 / 3447 = 0.039$$

- 2) Perhitungan P terhadap nama Protocol :

$$\text{Nilai } P(\text{DHCP} | \text{Rendah}) \\ 4 / 2200 = 0.00181818$$

$$\text{Nilai } P(\text{DHCP} | \text{Menengah}) \\ 0 / 1113 = 0$$

$$\text{Nilai } P(\text{DHCP} | \text{Tinggi}) \\ 0 / 132 = 0$$

- 3) Perhitungan P terhadap Length :

$$\text{Nilai } P(257 - 512 | \text{Rendah}) \\ 4 / 2200 = 0.00181818$$

$$\text{Nilai } P(257 - 512 | \text{Menengah}) \\ 0 / 1113 = 0$$

$$\text{Nilai } P(157 - 512 | \text{Tinggi}) \\ 4 / 132 = 0$$

- 4) Perhitungan P terhadap Counting :

$$\text{Nilai } P(0 - 500 | \text{Rendah}) \\ 4 / 2200 = 0.00181818$$

$$\text{Nilai } P(0 - 500 | \text{Menengah}) \\ 0 / 1113 = 0$$

$$\text{Nilai } P(0 - 500 | \text{Tinggi}) \\ 0 / 132 = 0$$

- 5) Perhitungan *Posterior* Rendah
 $0.63823615 \times 0.00181818 \times 0.00181818 \times 0.00181818 = 8.4395E-06$

- 6) Perhitungan *Posterior* Menengah
 $0.32288947 \times 0 \times 0 \times 0 = 0$

- 7) Perhitungan *Posterior* Tinggi
 $0.03829417 \times 0 \times 0 \times 0 = 0$

Berdasarkan perhitungan nilai *Posterior* masing-masing kelas prioritas terlihat bahwa *Posterior* "Rendah" yang memiliki nilai tertinggi, maka hasil perhitungan dengan algoritma klasifikasi *Naïve Bayes* menghasilkan klasifikasi prioritas "rendah". Yang ditunjukkan pada Tabel 4.

b) Analisis kinerja Sistem

Perhitungan manual terhadap data latih dengan data uji yang diambil secara acak. Maka didapatkan tingkat akurasi dengan membaca hasil dari *confusion matrik*, seperti yang ditunjukkan pada Tabel 6.

Tabel 6. Hasil Confusion Matriks

Data Asli	Data Hasil		
	Rendah	Menengah	Tinggi
Rendah	2073	100	27
Menengah	61	1035	7
Tinggi	32	27	73

Persamaan yang digunakan adalah persamaan (2),

$$\text{Akurasi} = \frac{2073 + 1035 + 73}{2073 + 100 + 27 + 61 + 1035 + 17 + 32 + 27 + 73}$$

$$\text{Akurasi} = \frac{3181}{3445}$$

$$\text{Akurasi} = 0.92336720$$

$$\text{Akurasi} = 92.34\%$$

Error rate yang didapat dengan menggunakan persamaan (3)

$$\text{Error} = \frac{100 + 27 + 61 + 17 + 32 + 73}{2073 + 100 + 27 + 61 + 1035 + 17 + 32 + 27 + 73}$$

$$\text{Error} = \frac{310}{3445}$$

$$\text{Error} = 0.08998549$$

Hasil perhitungan manual maka hasil akurasi yang didapat adalah 92,34 %. Dari nilai 92,34% dapat ditarik simpulan bahwa model yang diujikan memiliki tingkat keakuratan yang tinggi. Dengan *errorrate* senilai 0,08998549.

4.2. HASIL KLASIFIKASI

Berdasarkan hasil pengujian yang dilakukan terhadap data uji baru maka dihasilkan data hasil klasifikasi trafik jaringan

menggunakan algoritma *Naïve Bayes* yang ditunjukkan pada Tabel 7.

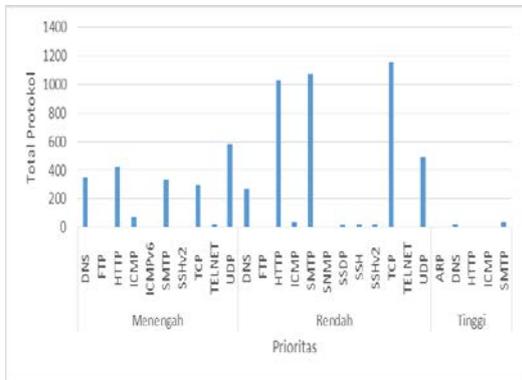
Tabel 7 menunjukkan hasil pengujian yang dilakukan terhadap data uji baru maka dihasilkan data hasil klasifikasi menggunakan algoritma *Naïve Bayes* menggunakan *MATLAB*. Bahwa protocol *ARP* dengan *length 60* dan *counting 246* menghasilkan klasifikasi dengan prioritas tinggi. Hal ini

disebabkan karena pengguna internet yang menggunakan protokol *ARP* lebih sering dengan *ukuran length 60 byte* oleh karena itu, protokol *ARP* akan diberikan prioritas tinggi. Dengan diberikan prioritas tinggi kepada pengguna internet maka pengguna internet merasa nyaman menggunakan protokol komunikasi *ARP* tersebut.

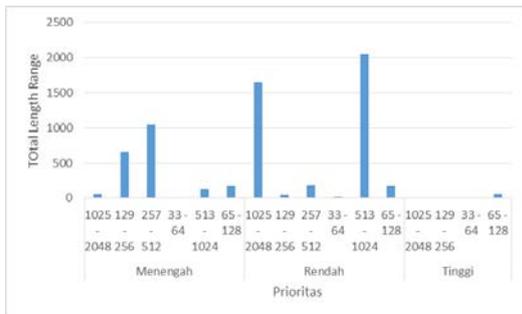
Tabel 7. Hasil Klasifikasi *Naïve Bayes*

Protokol	Length	Counting	Length Range	Counting Range	Prioritas
ARP	60	264	33 - 64	0 - 500	Tinggi
DNS	60	6	33 - 64	0 - 500	Rendah
DNS	61	6	33 - 64	0 - 500	Rendah
DNS	64	1	33 - 64	0 - 500	Rendah
DNS	65	13	65 - 128	0 - 500	Rendah
DNS	66	5	65 - 128	0 - 500	Rendah
DNS	67	34	65 - 128	0 - 500	Rendah
DNS	68	14	65 - 128	0 - 500	Rendah
DNS	69	125	65 - 128	0 - 500	Tinggi
DNS	70	409	65 - 128	0 - 500	Tinggi
DNS	71	691	65 - 128	501 - 1000	Menengah
DNS	72	850	65 - 128	501 - 1000	Tinggi
DNS	73	662	65 - 128	501 - 1000	Menengah
DNS	74	1830	65 - 128	1501 - 2000	Tinggi
DNS	75	1229	65 - 128	1001 - 1500	Tinggi
DNS	76	1216	65 - 128	1001 - 1500	Tinggi
DNS	77	1329	65 - 128	1001 - 1500	Tinggi
DNS	78	1365	65 - 128	1001 - 1500	Tinggi
DNS	79	1370	65 - 128	1001 - 1500	Tinggi
DNS	80	1222	65 - 128	1001 - 1500	Tinggi
DNS	81	1341	65 - 128	1001 - 1500	Tinggi
DNS	82	1183	65 - 128	1001 - 1500	Menengah
DNS	83	1679	65 - 128	1501 - 2000	Tinggi
DNS	84	1798	65 - 128	1501 - 2000	Tinggi

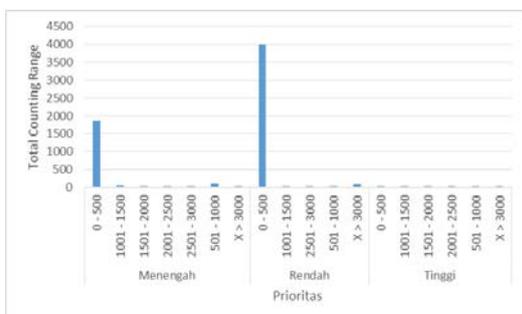
Pada Gambar 4, dapat dilihat bahwa hasil klasifikasi *networktraffic* dari sisi protokol terhadap prioritas menjelaskan bahwa protokol *TCP* memiliki tingkat kemunculan yang tinggi maka prioritas yang diberikan oleh sistem adalah prioritas rendah. Gambar 5 menunjukkan bahwa *length* yang yang diberikan prioritas rendah adalah terhadap komunikasi yang memiliki ukuran data dari 1025 – 2048 *byte* dan 513 – 1024 *byte*. Sedangkan pada Gambar 6 dapat dilihat bahwa *counting* range 0 – 500 memiliki tingkat kemunculan (banyaknya data *record* aktifitas yang sama) yang tinggi maka, akan diberikan prioritas rendah pada *countrange* 0 – 500.



Gambar 4. Protokol berdasarkan prioritas



Gambar 5. Length range berdasarkan prioritas



Gambar 6. Counting range berdasarkan prioritas

5. SIMPULAN

Paper ini telah memaparkan penerapan algoritma *Naïve Bayes* untuk klasifikasi internet pada jaringan kampus Universitas Udayana. Penerapan dalam klasifikasi berdasarkan penggunaan protokol, dengan ukuran data transmisi dalam bentuk *Length byte* dan *counting* (aktifitas kemunculan data yang sama dalam *record*).

Hasil analisis menunjuk bahwa pada trafik jaringan Unud didapatkan sebuah hasil nilai akurasi terhadap data latih sebesar 92,34%. Dimana 92,34% merupakan hasil akurasi dengan prediksi cukup akurat.

6. DAFTAR PUSTAKA

- [1] M. Rachli, "Email Filtering menggunakan Naïve Bayesian," *Tugas Akhir Jur. Tek. Elektro Inst. Teknol. Bdg. Bdg.*, 2007.
- [2] A. W. Moore and D. Zuev, "Internet traffic classification using bayesian analysis techniques," in *ACM SIGMETRICS Performance Evaluation Review*, 2005, vol. 33, pp. 50–60.
- [3] M. Sudarma and D. P. Hostiadi, "Klasifikasi Penggunaan Protokol Komunikasi Pada Network Traffic Menggunakan Naïve Bayes Sebagai Penentuan QoS," *Pros. CSGTEIS 2013*, vol. 0, no. 0, 2013.
- [4] P. A. Agustina, T. Matulatan, M. info Tech, and M. B. S. Si, "KLASIFIKASI TRENDING TOPIC TWITTER DENGAN PENERAPAN METODE NAÏVE BAYES."
- [5] R. Prathivi, "Klasifikasi Data Trafik Internet Menggunakan Metode Bayes Network (Studi Kasus Jaringan Internet Universitas Semarang)," *J. Tr NsForMat Ka*, vol. 12, no. 2, pp. 42–45, 2015.
- [6] C. C. Aggarwal and C. Zhai, *Mining text data*. Springer Science & Business Media, 2012.
- [7] F. GIAN, "Perbandingan Kinerja Metode Klasifikasi Data Mining Menggunakan Naive Bayes dan Algoritma C4. 5 untuk Prediksi Ketepatan Waktu Kelulusan Mahasiswa," *Skripsi Fak. Ilmu Komput.*, 2014.