

KLASIFIKASI DATA *BANDWIDTH* MENGGUNAKAN METODE *NAIVE BAYES*

Zawiyah Saharuna¹⁾, Irmawati²⁾, Meylanie Olivya³⁾

^{1),2,3)} Dosen Jurusan Teknik Elektro Politeknik Negeri Ujung Pandang, Makassar

ABSTRACT

Bandwidth management is important because it serves to regulate the allocation of bandwidth to users and applications running on the network. Each user has different bandwidth requirements, therefore it is necessary to determine priorities. In addition, it is also important to set priorities to different types of data streams based on how important that data stream. This can allow the use of bandwidth to be more efficient. Bandwidth management will be maximum when the usage characteristics are known. To determine the characteristics of bandwidth usage can be done by applying classification techniques. In this study, we proposed a classification 192 bandwidth data using Naïve Bayes method. The results showed that naive bayes method was successfully used for classification and gave 73,68% accuracy when split ratio was 0,8%.

Keywords: *Bandwidth, Classification, Naive Bayes.*

1. PENDAHULUAN

Pengelolaan *bandwidth* penting dilakukan karena berfungsi untuk mengatur alokasi *bandwidth* kepada pengguna dan aplikasi yang berjalan pada jaringan berdasarkan prioritas tertentu. Tingkat kebutuhan pengguna jaringan terhadap *bandwidth* berbeda-beda, sehingga perlu ditentukan pengguna yang memiliki prioritas lebih tinggi maupun yang rendah. Selain itu, penting juga menentukan prioritas terhadap berbagai jenis aliran data berdasarkan seberapa penting atau krusialnya aliran data tersebut. Hal ini memungkinkan penggunaan *bandwidth* yang tersedia menjadi lebih efisien, dan apabila sewaktu-waktu jaringan menjadi lambat, aliran data yang memiliki prioritas yang lebih rendah dapat dihentikan, sehingga aplikasi yang penting dapat tetap berjalan. Pengelolaan dan pengaturan *bandwidth* akan maksimal dilakukan bila karakteristik penggunaannya diketahui (Singh, dkk. 2013).

Untuk mengetahui karakteristik penggunaan *bandwidth* dapat dilakukan dengan menerapkan teknik klasifikasi. Klasifikasi adalah sebuah metode dari *datamining* yang digunakan untuk memprediksi kategori dari suatu *data instance* berdasarkan sekumpulan atribut-atribut dari data tersebut. Sebagai contoh dari penerapan metode klasifikasi adalah menentukan *e-mail* yang masuk termasuk kategori *spam* atau bukan *spam* (Pratiwi dan Ulama, 2016) atau memprediksi predikat kelulusan mahasiswa (Nugroho, 2014). Beberapa metode yang familiar digunakan dalam klasifikasi adalah C4.5, CMAR, *Naïve Bayes*, K-Nearest Neighbours, CART.

Penelitian sejenis telah dilakukan oleh Vijayakumar M., dkk. (2010), mereka menganalisis aliran trafik jaringan menggunakan teknik *Hierarchical Clustering Stream Network Traffic*. Data yang digunakan adalah data set trafik jaringan dari *Internet Service Provider* (ISP) dan pada tahapan *pre processing* menggunakan Algoritma K-means untuk membentuk *cluster* trafik jaringan. Berdasarkan hasil yang diperoleh, mereka menyimpulkan bahwa perpaduan *Hierarchical Clustering Stream Network Traffic* dengan *K-means* dapat meningkatkan akurasi hasil *cluster*.

Pada tahun 2011, Arya dan Mishra juga melakukan penelitian yang sejenis. Mereka mengevaluasi performa lima Algoritma seperti J48, *Random Tree*, *Random Forest*, *Bagging* dan *Boosting Classifiers* untuk klasifikasi *dataset benchmark*. Tingkat akurasi, presisi, dan *recall* digunakan untuk membandingkan performa dari kelima classifiers. Hasil penelitian menunjukkan bahwa multi level *classifier* lebih akurat dalam proses klasifikasi jaringan internet.

Özyirmidokuz, dkk. (2012) melakukan penelitian untuk menemukan pengetahuan yang berguna dari data trafik jaringan. Kami Mereka menganalisis data trafik jaringan 150Mbps yang menghubungkan antara AS dan Jepang menggunakan metode *CART decision tree*. Penelitian ini menganalisis deteksi anomali dan menerapkan Kohonen Networks untuk mengurangi data matriks. Hasil penelitian menunjukkan bahwa metode *CART decision tree* berhasil diterapkan dalam analisis lalu lintas jaringan dan memiliki kinerja yang cukup baik.

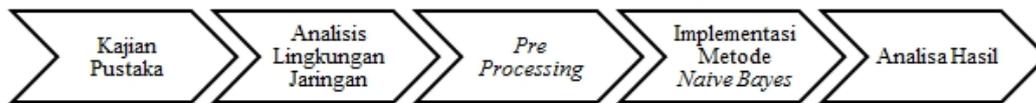
¹ Korespondensi: Zawiyah Saharuna, zawiyah@poliupg.ac.id

Selain itu, Subrata, dkk. (2017) juga melakukan penelitian sejenis. Mereka melakukan klasifikasi terhadap data capture trafik jaringan yang diolah menggunakan Algoritma K-NearestNeighbor (K-NN). Aplikasi yang digunakan untuk capture trafik jaringan yaitu aplikasi wireshark. Hasil observasi terhadap dataset trafik jaringan dan melalui proses perhitungan menggunakan Algoritma K-NN didapatkan sebuah hasil bahwa klasifikasi K-NN memiliki tingkat keakuratan yang sangat tinggi. Hal ini dibuktikan dengan hasil perhitungan yang mencapai nilai 99,14 % yaitu dengan perhitungan $k = 3$.

Sedangkan pada penelitian yang penulis lakukan yaitu klasifikasi data *bandwidth* menggunakan metode *Naïve Bayes* untuk mengetahui karakteristik penggunaan *bandwidth* di Politeknik Negeri Ujung Pandang. Metode *Naïve Bayes* dipilih karena sederhana dan memiliki tingkat akurasi yang tinggi (Rodiansyah S.F. dan Winarko E, 2012).

2. METODE PENELITIAN

Tahapan-tahapan dalam penelitian ini secara garis besar dijelaskan pada Gambar 1.



Gambar 1 Tahapan-tahapan Penelitian

Pada tahapan pertama dilakukan kajian pustaka terhadap beberapa referensi yang relevan dengan topik penelitian. Adapun referensi yang dirujuk dalam penelitian ini adalah yang berkaitan dengan *bandwidth*, trafik, teknik klasifikasi dan metodenya, dan metode *Naïve Bayes*.

Pada tahapan analisis lingkungan jaringan dilakukan analisis terkait infrastruktur dan topologi jaringan internet pada Politeknik Negeri Ujung Pandang. Pada tahapan ini dilakukan juga pengumpulan data penggunaan *bandwidth*. Data diperoleh dari noc Politeknik Negeri Ujung Pandang (PNUP) yang selanjutnya disebut “data *bandwidth*”. Data *bandwidth* yang digunakan adalah data *bandwidth* pada *port router load balancing* yang terdiri atas *port* G0/0, G0/1, dan G0/2. *Port* G0/0 adalah *port* yang terhubung dengan Astinet Telkom, *Port* G0/1 adalah *port* yang terhubung dengan Lintas Arta, dan *port* G0/2 adalah *port* yang terhubung ke jaringan lokal PNUP. Data *bandwidth* dari ketiga *port* tersebut terdiri atas atribut *Date*, *Time*, *Inbound*, dan *Outbound*.

Sebelum data yang telah dikumpulkan diolah menggunakan metode *Naïve Bayes*, terlebih dahulu dilakukan *pre-processing*. *Pre-processing* yang dilakukan pada data *bandwidth* terdiri atas beberapa tahapan:

1. Pembersihan data

Pembersihan data harus dilakukan karena data yang diperoleh dari hasil monitoring terkadang memiliki isian data yang tidak sempurna. Selain itu, sangat memungkinkan pada data terdapat atribut yang tidak relevan dengan hipotesa penelitian sehingga dianggap bisa mengurangi mutu dan akurasi. Kondisi awal data *bandwidth* terdiri atas 2013 data, kemudian dilakukan pembersihan untuk data *inbound* dan *outbound* yang “NaN. Data “NaN” terjadi karena adanya pemadaman listrik yang mengakibatkan router pada pusat pangkalan data tidak aktif sehingga nilai *inbound* dan *outbound* pada interface router tidak ada. Setelah proses pembersihan data, jumlah data *bandwidth* menjadi 1991 data.

2. Integrasi dan Transformasi

Proses integrasi dilakukan untuk menggabungkan data *bandwidth* dari setiap *port* yang dimiliki oleh *router load balancing* berdasarkan hasil pengambilan data hari pertama hingga hari kedelapan. Sedangkan untuk transformasi data dari numerik menjadi ordinal untuk atribut *inbound* dan *outbound* serta transformasi data dari *time* ke nominal untuk atribut *time* dilakukan untuk meningkatkan nilai akurasi metode *naive bayes*. Hasil transformasi data dapat dilihat pada Tabel 1 dan Tabel 2.

Tabel 1 Transformasi Atribut *Time*

<i>Time</i>	Nominal
19.00-05.59	Malam
06.00-10.59	Pagi
11.00-14.59	Siang
15.00-18.59	Sore

Tabel 2 Transformasi Atribut Nilai *Bandwidth*

Nilai <i>Bandwidth</i> (%)	Ordinal
0% - 24%	Sangat Rendah
25% - 39%	Rendah
40% - 69%	Sedang
70% - 84%	Tinggi
85% - 100%	Sangat Tinggi

3. Reduksi Data

Proses reduksi yang dilakukan adalah menghapus atribut *Date* dan mengurangi redundansi nilai *bandwidth* dengan mengambil nilai rata-rata *bandwidth* selama satu jam untuk mewakili nilai *bandwidth* perjam. Nilai redundan terjadi karena rekaman data dilakukan setiap 5 menit atau dengan kata lain terdapat 12 rekaman data setiap jam. Tahapan pre-processing menghasilkan dataset dengan 7 atribut nominal dan 192 records.

3. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan metode *Naive bayes*. *Naive bayes* bertujuan untuk melakukan klasifikasi data pada klas tertentu. Pada metode klasifikasi, semua atribut akan memberikan kontribusinya dalam pengambilan keputusan, dengan bobot atribut yang sama penting dan setiap atribut saling bebas). Dengan kata lain, dalam *Naive bayes* model yang digunakan adalah “model fitur independen” (Kusumadewi, 2009).

Prediksi *Naive bayes* didasarkan pada teorema Bayes dengan formula untuk klasifikasi sebagai berikut (Prasetyo, 2012):

$$P(Y|X) = \frac{P(Y) \prod_{i=1}^q P(X_i|Y)}{P(X)} \tag{1}$$

Keterangan:

- P(Y|X) = probabilitas data dengan vector X pada kelas Y.
- P(Y) = probabilitas awal kelas Y
- $\prod_{i=1}^q P(X_i|Y)$ = probabilitas independen kelas Y dari semua fitur dalam vector X

Data hasil pre-processing akan diklasifikasikan menggunakan metode *naive bayes*. Langkah awal yang dilakukan adalah membaca data *training*. Sampel data *training* ditunjukkan pada Tabel 3.

Tabel 3 Data Training

No	G0/0-In	G0/0-Out	G0/1-In	G0/1-Out	G0/2-In	G0/2-Out	Time
1	Sangat Rendah	Sedang	Sangat Rendah	Sangat Rendah	Sedang	Sangat Rendah	Malam
2	Sangat Rendah	Pagi					
3	Sedang	Sangat Rendah	Rendah	Sangat Rendah	Sangat Rendah	Rendah	Siang
...
192	Sangat Rendah	Malam					

Selanjutnya adalah menentukan nilai probabilitas setiap atribut yang ada pada data *training* berdasarkan teorema Bayes. Adapun nilai probabilitas setiap atribut adalah sebagai berikut:

1. Probabilitas Atribut G0/0-In

Berdasarkan data *bandwidth* pada Tabel 3 diketahui jumlah data *training* adalah sebanyak 192 data. Dari 192 data terdapat 0 data untuk kategori G0/0-In “Sangat Tinggi”, “Tinggi”, “Sedang”, dan “Rendah” dengan kategori *time* “Malam”, dan 88 data untuk kategori G0/0-In “Sangat Rendah” dengan kategori *time* “Malam”.

Tabel 4 Probabilitas Atribut G0/0-In

G0/0-In	Time				Probabilitas			
	Malam	Pagi	Siang	Sore	Malam	Pagi	Siang	Sore
Sangat Tinggi	0	0	0	0	0,00	0,00	0,00	0,00
Tinggi	0	0	3	0	0,00	0,00	0,09	0,00
Sedang	0	10	26	9	0,00	0,25	0,81	0,28
Rendah	0	7	3	7	0,00	0,18	0,09	0,22
Sangat Rendah	88	23	0	16	1,00	0,58	0,00	0,50
Jumlah	88	40	32	32	1,00	1,00	1,00	1,00

Untuk kategori *time* “Pagi”, terdapat 0 data untuk kategori G0/0-In “Sangat Tinggi” dan “Tinggi”, 10 data untuk kategori “Sedang”, 7 data untuk kategori G0/0-In “Rendah”, dan 23 data untuk kategori G0/0-In “Sangat Rendah”. Untuk kategori *time* “Siang”, terdapat 0 data untuk kategori G0/0-In “Sangat Tinggi”, dan 3 data untuk kategori “Tinggi”, 26 data untuk kategori “Sedang”, 3 data untuk kategori “Rendah”, dan 0 data untuk kategori “Sangat Rendah”. Untuk kategori *time* “Sore”, terdapat 0 data untuk kategori G0/0-In “Sangat

Tinggi” dan “Tinggi”, 9 data untuk kategori “Sedang”, 7 data untuk kategori G0/0-In “Rendah”, dan 16 data untuk kategori G0/0-In “Sangat Rendah”. Nilai probabilitas atribut G0/0-In dapat dilihat pada Tabel 4.

2. Probabilitas Atribut G0/0-Out

Untuk kategori *time* “Malam”, terdapat 0 data untuk kategori G0/0-Out “Sangat Tinggi” dan “Tinggi”, 8 data untuk kategori “Sedang”, 1 data untuk kategori G0/0- Out “Rendah”, dan 79 data untuk kategori G0/0-Out “Sangat Rendah”. Untuk kategori *time* “Pagi”, terdapat 0 data untuk kategori G0/0-Out “Sangat Tinggi”, “Tinggi”, “Sedang”, dan “Rendah”, serta 40 data untuk kategori G0/0-Out “Sangat Rendah”. Untuk kategori *time* “Siang”, terdapat 0 data untuk kategori G0/0-Out “Sangat Tinggi”, “Tinggi”, “Sedang”, dan “Rendah”, serta 32 data untuk kategori G0/0-Out “Sangat Rendah”. Untuk kategori *time* “Sore”, terdapat 0 data untuk kategori G0/0-Out “Sangat Tinggi”, “Tinggi”, “Sedang”, dan “Rendah”, serta 32 data untuk kategori G0/0-Out “Sangat Rendah”. Nilai probabilitas atribut G0/0-Out dapat dilihat pada Tabel 5.

Tabel 5 Probabilitas Atribut G0/0-Out

G0/0-Out	Time				Probabilitas			
	Malam	Pagi	Siang	Sore	Malam	Pagi	Siang	Sore
Sangat Tinggi	0	0	0	0	0,00	0,00	0,00	0,00
Tinggi	0	0	0	0	0,00	0,00	0,00	0,00
Sedang	8	0	0	0	0,09	0,00	0,00	0,00
Rendah	1	0	0	0	0,01	0,00	0,00	0,00
Sangat Rendah	79	40	32	32	0,90	1,00	1,00	1,00
Jumlah	88	40	32	32	1,00	1,00	1,00	1,00

3. Probabilitas Atribut G0/1-In

Untuk kategori *time* “Malam”, terdapat 0 data untuk kategori G0/1-In “Sangat Tinggi”, “Tinggi”, “Sedang”, dan “Rendah”, serta 88 data untuk kategori G0/1-In “Sangat Rendah”. Untuk kategori *time* “Pagi”, terdapat 0 data untuk kategori G0/1-In “Sangat Tinggi”, 1 data untuk kategori “Tinggi”, 3 data untuk kategori “Sedang”, 11 data untuk kategori G0/1-In “Rendah”, dan 25 data untuk kategori G0/1-In “Sangat Rendah”. Untuk kategori *time* “Siang”, terdapat 0 data untuk kategori G0/1-In “Sangat Tinggi”, 1 data untuk kategori “Tinggi”, 17 data untuk kategori “Sedang”, terdapat 10 data untuk kategori “Rendah”, serta 4 data untuk kategori G0/1-In “Sangat Rendah”. Untuk kategori *time* “Sore”, terdapat 0 data untuk kategori G0/1-In “Sangat Tinggi” dan “Tinggi”, 5 data untuk kategori “Sedang”, 8 data untuk kategori “Rendah”, serta 19 data untuk kategori G0/1-In “Sangat Rendah”. Nilai probabilitas atribut G0/1-In dapat dilihat pada Tabel 6.

Tabel 6 Probabilitas Atribut G0/1-In

G0/1-In	Time				Probabilitas			
	Malam	Pagi	Siang	Sore	Malam	Pagi	Siang	Sore
Sangat Tinggi	0	0	0	0	0,00	0,00	0,00	0,00
Tinggi	0	1	1	0	0,00	0,03	0,03	0,00
Sedang	0	3	17	5	0,00	0,08	0,53	0,16
Rendah	0	11	10	8	0,00	0,28	0,31	0,25
Sangat Rendah	88	25	4	19	1,00	0,63	0,13	0,59
Jumlah	88	40	32	32	1,00	1,00	1,00	1,00

4. Probabilitas Atribut G0/1-Out

Untuk kategori *time* “Malam”, terdapat 0 data untuk kategori G0/1-Out “Sangat Tinggi”, “Tinggi”, “Sedang”, dan “Rendah”, serta 88 data untuk kategori G0/1-Out “Sangat Rendah”. Untuk kategori *time* “Pagi”, terdapat 0 data untuk kategori G0/1-Out “Sangat Tinggi”, “Tinggi”, “Sedang”, dan “Rendah”, serta 40 data untuk kategori G0/1-Out “Sangat Rendah”.

Tabel 7. Probabilitas Atribut G0/1-Out

G0/1-Out	Time				Probabilitas			
	Malam	Pagi	Siang	Sore	Malam	Pagi	Siang	Sore
Sangat Tinggi	0	0	0	0	0,00	0,00	0,00	0,00
Tinggi	0	0	0	0	0,00	0,00	0,00	0,00
Sedang	0	0	0	0	0,00	0,00	0,00	0,00
Rendah	0	0	0	0	0,00	0,00	0,00	0,00
Sangat Rendah	88	40	32	32	1,00	1,00	1,00	1,00
Jumlah	88	40	32	32	1,00	1,00	1,00	1,00

Untuk kategori *time* “Siang”, terdapat 0 data untuk kategori G0/1-Out “Sangat Tinggi”, “Tinggi”, “Sedang”, dan “Rendah”, serta 32 data untuk kategori G0/1-Out “Sangat Rendah”. Untuk kategori *time* “Sore”, terdapat 0 data untuk kategori G0/1-Out “Sangat Tinggi”, “Tinggi”, “Sedang”, dan “Rendah”, serta 32 data untuk kategori G0/1-Out “Sangat Rendah”. Nilai probabilitas atribut G0/1-Out dapat dilihat pada Tabel 7.

5. Probabilitas Atribut G0/2-In

Pada atribut G0/2-In, untuk kategori *time* “Malam”, terdapat 0 data untuk kategori G0/2-In “Sangat Tinggi” dan “Tinggi”, 4 data untuk kategori G0/2-In “Sedang”, dan 4 data untuk kategori G0/2-In “Rendah”, serta 80 data untuk kategori G0/2-In “Sangat Rendah”. Untuk kategori *time* “Pagi”, terdapat 0 data untuk kategori G0/2-In “Sangat Tinggi”, “Tinggi”, “Sedang”, dan “Rendah”, serta 40 data untuk kategori G0/2-In “Sangat Rendah”. Untuk kategori *time* “Siang”, terdapat 0 data untuk kategori G0/2-In “Sangat Tinggi”, “Tinggi”, “Sedang”, dan “Rendah”, serta 32 data untuk kategori G0/2-In “Sangat Rendah”. Untuk kategori *time* “Sore”, terdapat 0 data untuk kategori G0/2-In “Sangat Tinggi”, “Tinggi”, “Sedang”, dan “Rendah”, serta 32 data untuk kategori G0/2-In “Sangat Rendah”. Nilai probabilitas atribut G0/2-In dapat dilihat pada Tabel 8.

Tabel 8 Probabilitas Atribut G0/2-In

G0/2-In	Time				Probabilitas			
	Malam	Malam	Siang	Sore	Malam	Pagi	Siang	Sore
Sangat Tinggi	0	0	0	0	0,00	0,00	0,00	0,00
Tinggi	0	0	0	0	0,00	0,00	0,00	0,00
Sedang	4	0	0	0	0,05	0,00	0,00	0,00
Rendah	4	0	0	0	0,05	0,00	0,00	0,00
Sangat Rendah	80	40	32	32	0,91	1,00	1,00	1,00
Jumlah	88	40	32	32	1,00	1,00	1,00	1,00

6. Probabilitas Atribut G0/2-Out

Untuk kategori *time* “Malam”, terdapat 0 data untuk kategori G0/2-Out “Sangat Tinggi”, “Tinggi”, “Sedang”, dan “Rendah”, serta 88 data untuk kategori G0/2-Out “Sangat Rendah”. Untuk kategori *time* “Pagi”, terdapat 0 data untuk kategori G0/2-Out “Sangat Tinggi” dan “Tinggi”, 9 data untuk kategori “Sedang”, 8 data untuk kategori G0/2-Out “Rendah”, dan 23 data untuk kategori G0/2-Out “Sangat Rendah”. Untuk kategori *time* “Siang”, terdapat 0 data untuk kategori G0/2-Out “Sangat Tinggi”, 3 data untuk kategori “Tinggi”, 23 data untuk kategori “Sedang”, terdapat 6 data untuk kategori “Rendah”, serta 0 data untuk kategori G0/2-Out “Sangat Rendah”. Untuk kategori *time* “Sore”, terdapat 0 data untuk kategori G0/2-Out “Sangat Tinggi” dan “Tinggi”, 8 data untuk kategori “Sedang”, 9 data untuk kategori “Rendah”, serta 15 data untuk kategori G0/0-In “Sangat Rendah”. Nilai probabilitas atribut G0/2-Out dapat dilihat pada Tabel 9.

Tabel 9 Probabilitas Atribut G0/2-Out

G0/2-Out	Time				Probabilitas			
	Malam	Pagi	Siang	Sore	Malam	Pagi	Siang	Sore
Sangat Tinggi	0	0	0	0	0,00	0,00	0,00	0,00
Tinggi	0	0	3	0	0,00	0,00	0,09	0,00
Sedang	0	9	23	8	0,00	0,23	0,72	0,25
Rendah	0	8	6	9	0,00	0,20	0,19	0,28
Sangat Rendah	88	23	0	15	1,00	0,58	0,00	0,47
Jumlah	88	40	32	32	1,00	1,00	1,00	1,00

7. Probabilitas Atribut *Time*

Berdasarkan Tabel 3 terdapat 192 data atribut *Time* dengan rincian 88 data kategori “Malam”, 40 data dengan kategori “Pagi”, 32 data dengan kategori “Siang”, dan 32 data dengan kategori “Sore”. Nilai probabilitas atribut *time* dapat dilihat pada Tabel 10.

Tabel 10 Probabilitas Atribut *Time*

Time				Probabilitas			
Malam	Pagi	Siang	Sore	Malam	Pagi	Siang	Sore
88	40	32	32	0,46	0,21	0,17	0,17

Untuk menilai kinerja metode *naive bayes* dilakukan pengujian terhadap 192 data dengan variasi nilai rasio 0,5; 0,6; 0,7; 0,8; dan 0,9. Hasil pengujian untuk kelima variasi data tersebut menunjukkan *split ratio*

terbaik adalah yang memiliki nilai akurasi tertinggi yaitu 73,68%. Kondisi ini tercapai saat nilai perbandingan data *training* dan data *testing* sama dengan 80:20.

Berdasarkan hasil pengujian menggunakan split ratio 0,8% yaitu 154 data *training* yang digunakan untuk melatih model *naive bayes* dan 38 data *testing* yang digunakan untuk menguji model diperoleh hasil seperti pada Tabel 11. Dari 38 data *testing*, terdapat 28 data yang diprediksi benar dan 10 data yang diprediksi salah.

Tabel 12 Hasil Pengujian *Split Ratio* 0,8%

	true Malam	true Pagi	true Siang	true Sore	class precision
pred. Malam	19	4	0	0	82.61%
pred. Pagi	0	1	0	2	33.33%
pred. Siang	0	1	8	3	66.67%
pred. Sore	0	0	0	0	0.00%
class recall	100.00%	16.67%	100.00%	0.00%	

Nilai *class recall* diperoleh dari jumlah data *true* positif dibagi total data *true* positif dan *true* negatif sehingga nilai recall untuk class malam adalah 100%, class pagi 16,67%, class siang 100%, dan class sore 0%. Nilai *class precision* diperoleh dari jumlah data *true* positif dibagi total *true* positif dan *false* positif sehingga nilai precision untuk class malam 82,61%, class pagi 33,33%, class siang 66,67%, dan class sore 0%. Sedangkan untuk nilai kinerja/akurasi diperoleh sebesar **73,68%**.

4. KESIMPULAN

Kesimpulan dari penelitian ini adalah klasifikasi data *bandwidth* menggunakan metode *naive bayes* berhasil dilakukan dan diperoleh nilai akurasi sebesar 73,68%. Nilai akurasi tersebut dicapai saat *split ratio* 0,8%.

5. DAFTAR PUSTAKA

- Arya I. B., Mishra R., *Internet Traffic Classification: An Enhancement in Performance using Classifiers* (2011) Vol. 2 pp. 663-667.
- Kusumadewi Sri, *Klasifikasi Status Gizi Menggunakan Naive Bayes Classification*, CommIT, Vol.3, hlm. 6-11, mei 2009.
- Nugroho, Y.S., 2014, *Penerapan Algoritma C4.5 Untuk Klasifikasi Predikat Kelulusan Mahasiswa Fakultas Komunikasi Dan Informatika Universitas Muhammadiyah Surakarta*, Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) 2014.
- Özyirmidokuz E.K., Gezer A., Çiflikli C., 2012, *Characterization of Network Traffic Data: A Data Preprocessing and Data Mining Application*, International Conference on Data Analytics.
- Prasetyo E. 2012. *Data Mining Konsep dan Aplikasi Menggunakan Matlab*. Yogyakarta. Andi.
- Pratiwi S.N.D., Ulama B.S.S., 2016, *Klasifikasi Email Spam dengan Menggunakan Metode Support Vector Machine dan k-Nearest Neighbor*, JURNAL SAINS DAN SENI ITS Vol. 5 No. 2.
- Rodiansyah S.F. dan Winarko E., *Klasifikasi Posting Twitter Kemacetan Lalu Lintas Kota Bandung Menggunakan Naive Bayesian Classification*, Indonesian Journal of Computing and Cybernetics Systems (IJCCS) Vol.6, No.1, Januari 2012.
- Singh K., Agrawal S., Sohi B.S., 2013, *A Near Real-time IP Traffic Classification Using Machine Learning*, IJ. Intelligent Systems and Applications, 2013 No. 03, p 83-93.
- Subrata, I.K K.A., Widyantara I.M.O., Linawati, 2017, *Klasifikasi Penggunaan Protokol Komunikasi Pada Trafik Jaringan Menggunakan Algoritma K-Nearest Neighbor*, Teknologi Elektro, Vol. 16, No1, Januari-April 2017.
- Vijayakumar M., Parvathi R.M.S., *Concept Mining of High Volume Data Streams in Network Traffic Using Hierarchical Clustering*, Academic Journal, (2010) Vol. 39 Issue 2, p234.

6. UCAPAN TERIMA KASIH

Ucapan terima kasih kepada Direktorat Riset dan Pengabdian Masyarakat (DRPM) Direktorat Jendral Penguatan Riset dan Pengembangan Kementerian Riset, Teknologi, dan Pendidikan Tinggi, Direktur Politeknik Negeri Ujung Pandang, dan Kepala Unit Penelitian dan Pengabdian Masyarakat Politeknik Negeri Ujung Pandang.