ANALISIS *INFORMATION GAIN ATTRIBUTE EVALUATION* UNTUK KLASIFIKASI SERANGAN INTRUSI

Aulia Essra[(1)] , Rahmadani[(2)] , Safriadi[(3)]

Magister Teknik Informatika, Universitas Sumatera Utara
Jl. Universitas No.24A Kampus USU, Medan, 20155, Indonesia
E-mail : aulia.e@live.com [(1)]

ABSTRACT

In this study applied Attribute Information Gain Evaluation techniques for preprocessing the Naïve Bayes classification algorithm. Characteristic of Naïve Bayes algorithm within the classification is based on probability theory that sees all the features of the data as evidence in probabilitias. By doing feature selection, the results of this study indicate that Information Gain Attribute Evaluation not significantly improve performance and computation time Naïve Bayes classification algorithm.

Keywords : *Information Gain Attribute Evaluation , Naïve Bayes , Classification , Performance*

ABSTRAK

Pada penelitian ini diterapkan teknik Information Gain Attribute Evaluation untuk preprocessing dalam klasifikasi menggunakan algoritma Naïve Bayes. Karakteristik Naïve Bayes dalam melakukan klasifikasi didasarkan pada teori probabilitas yang memandang semua fitur dari data sebagai bukti dalam probabilitias. Dengan melakukan seleksi fitur, hasil penelitian ini menunjukkan bahwa Information Gain Attribute Evaluation tidak begitu signifikan meningkatkan kinerja dan waktu komputasi klasifikasi dengan algoritma Naïve Bayes.

Kata kunci: *Information Gain Attribute Evaluation, Naïve Bayes, Klasifikasi, Performansi*

PENDAHULUAN

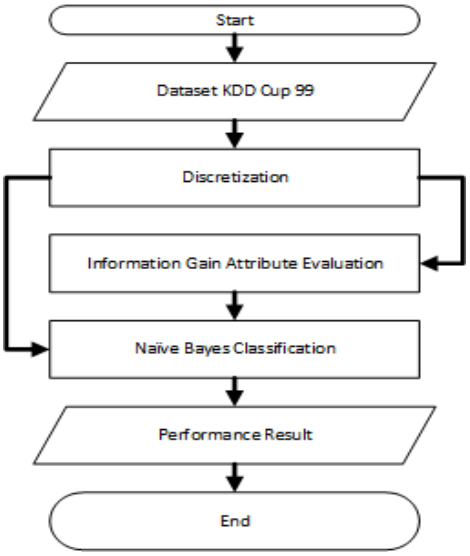
Klasifikasi digunakan untuk menilai data dengan memasukkan data tersebut ke dalam kelas tertentu dari sejumlah kelas yang tersedia. Ada dua pekerjaan utama dalam klasifikasi yaitu pembangunan model dengan melakukan pelatihan (*training*) sebagai *prototype* untuk disimpan sebagai memori; penggunaan atau penerapan model tersebut untuk melakukan klasifikasi data lain agar diketahui di kelas mana data tersebut dalam model yang sudah disimpan pada proses pembangunan model [1].

Karakteristik Naïve Bayes dalam melakukan klasifikasi didasarkan pada teori probabilitas yang memandang semua fitur/atribut dari data sebagai bukti dalam probabilitias. Model yang digunakan dalam Naïve Bayes adalah model fitur/atribut independen. Dengan kata lain, Bayes merupakan teknik prediksi berbasis probabilistik dengan asumsi independensi yang kuat pada fitur/atribut. Maksudnya adalah bahwa fitur/atribut pada sebuah data tidak berkaitan dengan ada atau tidak adanya fitur/atribut lain dalam data yang sama [1].

Dengan lebih menitik beratkan pada atribut data, seleksi atribut dapat menjadi sebuah masalah fundamental terhadap kinerja suatu algoritma. Untuk beberapa algoritma, setiap atribut data memiliki peranan penting, tapi untuk beberapa algoritma dalam mencapai targetnya, hanya beberapa atribut yang relevan untuk tindakan lebih lanjutnya. Beberapa atribut pada data mungkin memiliki nilai yang tidak relevan untuk tugas mining dan jika mengikutsertakannya dapat merugikan dan mengacaukan tugas algoritma data mining [2]. Untuk itu perlu dilakukan seleksi atribut yang merupakan proses untuk mengidentifikasi dan menghilangkan atribut dengan nilai yang tidak relevan atau berlebihan [3].
Seleksi atribut dapat dilakukan dengan metode penyaringan dimana atribut diurutkan sesuai dengan peringkat (ranking) berdasarkan evaluasi dengan kriteria tertentu seperti akurasi dan konsistensi data [4]. Salah satu algoritma yang digunakan dalam seleksi atribut dengan metode penyaringan adalah informasi Gain. Informasi Gain menggunakan entropy untuk menentukan atribut terbaik. Entropy merupakan ukuran ketidakpastian. Semakin besar nilai informasi gain dari suatu atribut, maka semakin signifikan atribut tersebut untuk tugas prediksi [5].
Penelitian ini menyajikan penerapan algoritma Naïve Bayes untuk mengetahui perubahan nilai akurasi dan hasil prediksi serangan yang ada dalam dataset KDD Cup 99 dengan dan tanpa seleksi atribut *Information Gain Attribute Evaluation* yang dipilih berdasarkan nilai informasi gainnya.

METODE PENELITIAN

Berikut ini *flow* diagram untuk metode penelitian yang dilakukan:



Gambar 1. Metode Penelitian

Dataset

Pada umumnya, dataset yang digunakan untuk penelitian mengenai klasifikasi serangan intrusi adalah dataset KDD Cup 99 [6]. Dataset ini merupakan data rekam koneksi yang terdiri dari 1 jenis data normal dan 22 jenis data serangan yang dikelompokkan kedalam empat tipe intrusi. Dataset yang diperoleh dari UCI Repository ini memiliki 41 atribut/fitur yang dibagi ke dalam tiga kelompok yaitu atribut basic, atribut konten dan atribut trafik.
Atribut basic (atribut nomor 1 sampai 9) merupakan hasil ekstraksi dari sistem log tcpdump dalam jaringan komputer. Atribut konten (atribut nomor 10 sampai 22) merupakan atribut-atribut yang diambil dari kegiatan yang berlangsung dalam sistem jaringan komputer. Sedangkan atribut trafik terbagi menjadi dua bagian, pertama terdiri dari atribut nomor 23 sampai 31 merupakan atribut trafik jaringan yang dihitung menggunakan waktu dua detik time window, dan kedua terdiri dari atribut nomor 32 sampai 41 dihitung menggunakan waktu dua detik time window dari tujuan ke host.

Discretization

Umumnya, Bayes mudah dihitung untuk atribut bertipe kategoris, namun pada dataset KDD Cup 99 terdapat beberapa atribut bertipe kontinu (numerik) sehingga ada perlakuan khusus sebelum dimasukkan dalam Naïve Bayes. Untuk mengubah fitur bertipe numerik menjadi kategoris dapat dilakukan dengan cara diskritisasi pada setiap fitur kontinu tersebut dengan nilai interval diskret. Dalam hal ini, pendekatan yang digunakan adalah metode *Supervised Entropy Based Discretization* yang diusulkan oleh Fayyad dan Irani [7].

Information Gain Attribute Evaluation

Penggunaan teknik seleksi atribut/fitur bertujuan untuk membuang fitur yang tidak relevan atau berlebihan dari segi vektor fitur yang diberikan. Dalam literatur *machine learning* telah banyak diusulkan teknik-teknik seleksi atribut. Beberapa seleksi atribut di antaranya adalah: IG (Information Gain), Gain Ratio, Symmetrical Uncertainty, Relief-F, One-R and Chi-Squared. Dalam tulisan ini dilakukan evaluasi terhadap Information Gain Attribute Evaluation.
Entropi pada umumnya digunakan untuk mengukur ketidakpastian dari sekumpulan atribut dari suatu set data. Ukuran entropi dianggap sebagai ukuran ketidakpastian dimana semakin tinggi entropy suatu atribut maka semakin tinggi ketidakpastian. Berikut ini rumus entropi [8].

$$E(Y) = -\sum_{y \in Y} p(y) log_2(p(y)) \quad (1)$$

dimana $p(y)$ sama dengan fungsi probabilitas marginal untuk variabel acak Y. Jika nilai-nilai Y yang diamati dalam dataset S dipartisi sesuai dengan nilai-nilai dari fitur kedua X, dan entropi Y terhadap partisi yang disebabkan oleh X kurang dari entropi Y sebelum partisi, maka ada hubungan antara fitur Y dan X. kemudian entropi Y setelah mengamati X adalah:

$$E(Y|X) = -\sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) \quad (2)$$

dimana $p(y|x)$ sama dengan probabilitas bersyarat dari y terhadap x. Mengingat entropi sebagai kriteria pengotor dalam pelatihan set S, kita dapat mendefinisikan ukuran mencerminkan informasi tambahan tentang Y disediakan oleh X yang mewakili jumlah dimana entropi Y menurun. Langkah ini dikenal sebagai IG. Hal ini diberikan oleh

$$IG = H(Y) - H(Y|X) = H(X) - H(X|Y) \quad (3)$$

Seperti yang terlihat pada persamaan 3 bahwa informasi yang diperoleh tentang

Y setelah mengamati X sama dengan informasi yang diperoleh tentang X setelah mengamati Y.

Untuk hasil perhitungan nilai Information Gain terhadap dataset KDD Cup 99 menunjukkan bahwa ada empat atribut yang memiliki nilai gain 0 yang berarti bahwa atribut tersebut tidak memberikan kontribusi pada karakteristik trafik jaringan dan 37 atribut memiliki nilai gain lebih besar dari 0 yang berarti relevan dengan karakteristik trafik jaringan.

Adapun empat atribut yang direduksi (nilai gain 0) tersebut adalah *lnum_outbound_cmds, lsu_attempted, urgent, is_host_login*. Hal ini disebabkan karena adanya kesamaan nilai pada setiap data yang dimiliki oleh masing-masing atribut sehingga tidak memberikan kontribusi terhadap karakteristik jaringan. Sebagai contoh atribut is_host_login bernilai 1 jika login dari hot list dan bernilai 0 jika sebaliknya. Di dalam dataset KDD Cup 99, semua data untuk atribut ini bernilai 0 sehingga perhitungan gain yang dihasilkan sama dengan 0. Dengan demikian, maka atribut ini dianggap tidak memberi kontribusi terhadap karakteristik jaringan. Menurut peneliti Suthaharan & Panchagnula [9], tidak semua atribut yang ada di dalam dataset KDD Cup 99 memberikan kontribusi pada karakteristik trafik jaringan

Naïve Bayes

Kaitan antara Naïve Bayes dengan klasifikasi, korelasi hipotesis, dan bukti dengan klasifikasi adalah bahwa hipotesis dalam teorema Bayes merupakan label kelas yang menjadi target pemetaan dalam klasifikasi, sedangkan bukti merupakan fitur-fitur yang menjadi masukan dalam model klasifikasi [1]. Jika X adalah vektor masukan yang berisi fitur dan Y adalah label kelas, Naïve Bayes dituliskan dengan P(Y|X). Notasi tersebut berarti probabilitas label kelas Y didapatkan setelah fitur-fitur X diamati. Notasi ini disebut juga probabilitas akhir (*posterior probability*) untuk Y, sedangkan P(Y) disebut probabilitas awal (*prior probability*) Y.

Selama proses pelatihan harus dilakukan pembelajaran probabilitas akhir P(Y|X) pada model untuk setiap kombinasi X dan Y berdasarkan informasi yang didapat dari data latih. Dengan membangun model tersebut, suatu data uji X' dapat diklasifikasikan dengan mencari nilai Y' dengan memaksimalkan nilai P(Y'|X') yang didapat.

Formulasi Naïve Bayes untuk klasifikasi adalah:

$$P(Y|X) = \frac{P(Y)\prod_{i=1}^{q} P(Y_i|X)}{P(X)} \quad (4)$$

P(Y|X) adalah probabilitas data dengan vektor X pada kelas Y. P(Y) adalah probabilitas awal kelas Y. $\prod_{i=1}^{q} P(Y_i|X)$ adalah probabilitas independen kelas Y dari semua fitur dalam vektor X. Nilai P(X) selalu tetap sehingga dalam perhitungan prediksi hanya menghitung bagian $P(Y)\prod_{i=1}^{q} P(Y_i|X)$ dengan memilih yang terbesar sebagai kelas yang dipilih sebagai hasil prediksi. Sementara probabilitas independen $\prod_{i=1}^{q} P(Y_i|X)$ tersebut merupakan pengaruh semua fitur dari data terhadap setiap kelas Y, yang difomulasikan dengan:

$$P(X|Y = y) = \prod_{i=1}^{q} P(X_i|Y = y) \quad (5)$$

dimana, setiap set fitur $X = \{X_1, X_2, X_3, \dots, X_q\}$ terdiri atas q atribut.

Pengukuran Performansi

Pada umumnya pengukuran performansi klasifikasi menggunakan matriks konfusi (*confusion matrix*). *Confusion matrix* merupakan tabel pencatat hasil kerja klasifikasi. Tabel

confusion matrix berisikan empat kemungkinan keluaran sebagai bahan acuan dalam membandingan antara kejadian yang sebenarnya (aktual) dengan kejadian yang terprediksi seperti yang ditunjukan pada tabel 1.

Table 1. Confusion Matrix

		Prediksi	
		Normal	Intrusi
Aktual	Normal	TP	TN
	Intrusi	FP	FN

dimana:
True Positive (*TP*) sama dengan jumlah data normal yang diprediksi normal.
True Negative (*TN*) sama dengan jumlah data normal yang diprediksi serangan.
False Positive (*FP*) sama dengan jumlah data serangan yang diprediksi normal.
False Negative (*FN*) sama dengan jumlah data serangan yang diprediksi serangan.
Jumlah data yang diklasifikasikan benar, dapat digunakan untuk menghitung performansi akurasi dan *Detection Rate*, sedangkan jumlah data yang diklasifikasikan salah, dapat digunakan untuk menghitung performansi *False Positive Rate* dan *True Negative Rate*. Berikut persamaan untuk menghitung performansi klasifikasi:

$$Akurasi = \frac{TP+FN}{TP+FN+TN+FP}\ 100\% \quad (4)$$

$$DR = \frac{FN}{FN+FP}\ 100\% \quad (5)$$

$$FPR = \frac{FP}{FP+FN}\ 100\% \quad (6)$$

$$TNR = \frac{TN}{TN+TP}\ 100\% \quad (7)$$

HASIL DAN PEMBAHASAN

Hasil Pengujian

Pengujian dilakukan dengan menggunakan seluruh data yang ada di dalam *dataset KDD Cup 99* 10% dengan komposisi data normal sebanyak 97.277 data dan data intrusi sebanyak 396.743 data. Sama dengan pengujian sebelumnya, hasil proses klasifikasi pola didistribusikan ke dalam *confusion matrix* seperti ditunjukkan pada Tabel 2 dan 3. Kemudian, nilai tersebut digunakan untuk mengukur performansi.

Table 2. Confusion Matrix Pengujian Tanpa Seleksi Atribut

		Prediksi	
		Normal	Intrusi
Aktual	Normal	97.176	101
	Intrusi	4.533	392.210

Berikut ini hasil perhitungan performansi untuk pengujian tanpa menggunakan seleksi atribut Information Gain Attribut Evaluation:
$Akurasi = 99{,}06198\ \%$
$DR = 98{,}85745\ \%$
$FPR = 1{,}14255\ \%$
$TNR = 0{,}10383\ \%$

Table 3. Confusion Matrix Pengujian Dengan Seleksi Atribut

		Prediksi	
		Normal	Intrusi
Aktual	Normal	97.176	101
	Intrusi	4.521	392.222

Berikut ini hasil perhitungan performansi untuk pengujian menggunakan seleksi atribut Information Gain Attribut Evaluation:
$Akurasi = 99{,}06441\ \%$
$DR = 98{,}86047\ \%$
$FPR = 1{,}13953\ \%$
$TNR = 0{,}10383\ \%$

Pembahasan

Dari dua eksperimen yang dilakukan dapat dilihat bahwa penggunaan teknik *Information Gain Evaluation Attribute* untuk menyeleksi dan mengurangi atribut yang tidak relevan atau nilai gainnya sama dengan 0, tidak memberi pengaruh yang signifikan terhadap performansi klasifikasi Naive Bayes. Perubahan hasil klasifikasi hanya terlihat pada peningkatan deteksi data intrusi dari 392.210 menjadi 392.222, sedangkan untuk data normal masih tetap

KESIMPULAN

Information Gain Attribute Evaluation tidak mampu secara signifikan meningkatkan performansi akurasi, deteksi dan false positive algoritma Naïve Bayes dalam melakukan klasifikasi data serangan yang ada pada dataset KDD Cup 99. Begitu juga dengan waktu pembangunan model yang tidak berpengaruh signifikan.

DAFTAR PUSTAKA

[1] E. Prasetyo, Data Mining: Konsep dan aplikasi menggunakan matlab, Yogyakarta: Andi Yogyakarta, 2012.

[2] B. Azhagusundari and A. Thanami, "Feature selection based on information gain," *International Journal and Innovative Technology and Exploring Engineering (IJITEE),* 2013.

[3] R. Abraham, J. Simha and S. Iyengar, "Effective discretization and hybrid feature selection using naive bayessian classifier for medical data mining," *International Journal of Computational Intelligence Research,* 2013.

[4] M. Danubianu, S. Pentiuc and D. Danubianu, "Data Dimensionality Reduction for Data Mining: A combined filter-wrapper framework," *International Journal of Computers, Communications & Control,* 2012.

[5] L. Ladha and T. Deepa, "Feature selection methods and algorithms," *International Journal on Computer Science and Engineering (IJCSE),* 2013.

[6] "UCI Machine Learning Repository," 1999. [Online]. Available: http://kdd.ics.uci.edu/.... [Accessed 25 Juni 2016].

[7] U. M. Fayyad and K. B. Irani, "Multi-interval discretization of continuous-valued attributes for classification learning," in *Proceedings of the 13th International Joint Conference on Artificial Intelligence*, 1993.

[8] M. Slocum, "Decision making using ID3 algorithm," *InSight: Rivier Academic Journal,* vol. Volume 8, p. Number 2, FALL 2012.

[9] S. Suthaharan and T. Panchagnula, "Relevance feature selection with data cleaning for intrusion detection system," in *Proceedings of IEEE (pp. 1-6)*, in Southeastcon: IEEE, 2012.