

Implementasi Algoritma C4.5 Untuk Klasifikasi Tingkat Kepuasan Pembeli Online Shop

Dimas Bayu Febriyanto, Lovi Handoko, Wahyuli, Hanif Aisyah, Rumini

Jurusan Informatika, Fakulitas Ilmu Komputer, Universitas AMIKOM Yogyakarta, Indonesia Jl. Ring Road Utara, Condong Catur, Sleman, Yogyakarta, Indonesia

Abstrak

Penelitian ini menerapkan algoritma C4.5 dalam memutuskan kepuasan pelanggan. Jenis penelitian ini yakni klasifikasi dengan konsep data mining dengan melibatkan sebanyak 100 data yang kemudian diambil sampel sebanyak 15 data konsumen Polaroid Vektor & Gift Jogja yang dikategorikan dengan: puas dan tidak puas. Metode yang digunakan dalam penelitian ini yakni metode klasifikasi Klasifikasi adalah proses menemukan kumpulan pola atau fungsi-fungsi yang mendeskripsikan dan memisahkan kelas data satu dengan lainnya, untuk dapat digunakan, untuk memprediksi data yang belum memiliki kelas data tertentu. Hasil penelitian terdapat tingkat akurasi sebesar 91%, dengan nilai presisi pada prediksi puas sebesar 66.67% dan nilai presisi pada prediksi tidak puas sebesar 33.3%

Kata kunci: Kepuasan Pelanggan, Algoritma C4.5, Klasifikasi

Abstract

This study applies the C4.5 algorithm in deciding customer satisfaction. The type of this research is classification with the concept of data mining by involving as many as 100 data which then take a sample of 15 data consumers of Vector Polaroid & Gift Jogja categorized by satisfied and dissatisfied. The method used in this research is classification method Classification is the process of finding a collection of patterns or functions that describe and separate data classes from one another, to be used, to predict data that does not yet have a particular data class. The results of the study have an accuracy rate of 91%, with the value of precision at predictive satisfaction of 66.67% and the value of precision in predictions not satisfied at 33.3%

Keywords: Customer Satisfaction, C4.5 Algorithm, Classification

1. PENDAHULUAN

Globalisasi sekarang ini telah terjadi dan menyebabkan zaman semakin maju, akan tetapi tidak diikuti dengan perekonomian negara kita yang jauh tertinggal. Banyak sekali perkembangan zaman yang bisa kita lihat di sekitar kita saat ini. Salah satunya adalah penggunaan ponsel pintar yang semakin meningkat setiap tahunnya. Banyaknya fitur-fitur media sosial yang mudah diakses dan ditawarkan dalam sistem komunikasi ponsel pintar ini tentunya semakin mempermudah kita dalam berkomunikasi dan tetap terhubung dengan orang-orang. Hal ini jelas dimanfaatkan oleh para produsen atau affiliate (orang yang menjualkan produk orang lain) untuk mencari celah atau peluang usaha dengan mengambil keuntungan yang didapat dari banyak pengguna ponsel pintar saat ini. Dengan hanya bermodalkan ponsel pintar yang kita gunakan dalam keseharian ditunjang pula dengan fasilitas internet yang semakin luas, kita dapat mencoba suatu bentuk usaha yang dapat ditawarkan kepada konsumen termasuk dalam interaksi jual belinya. Saat ini ada satu tren yang sedang mengemuka di Indonesia, bahkan di dunia, yakni belanja online atau sering disebut online shop.

Persaingan online shop saat ini amat ketat, ada banyak online shop yang menawarkan berbagai promo, hadiah, bahkan dengan harga murah, namun jika pelanggan kurang puas semua itu hasilnya akan sia-sia. Kepuasan pelanggan sangat dipengaruhi oleh kualitas layanan. Kualitas layanan mendorong pelanggan untuk komitmen kepada produk dan layanan suatu perusahaan sehingga berdampak kepada peningkatan penjualan suatu produk. Hal tersebut membuat pelaku usaha online shop menjadi harus berfikir dengan baik serta memiliki taktik dan strategi dalam mengembangkan dan menjaga bisnisnya, dan berusaha selalu meningkatkan kepuasan konsumen.

Kepuasan konsumen dapat diartikan sebagai pendapat dari konsumen yang mengkonsumsi atau membeli suatu barang sesuai apa yang diinginkan atau diharapkan. Pada umumnya kepuasan pada tiap konsumen selalu berbeda-beda. Tak jarang ada konsumen yang mendapat begitu banyak promosi atau kemudahan yang ditawarkan namun tidak puas, dan tak jarang pula dengan kondisi dan keadaan yang sederhana konsumen juga cukup puas.

Online shop yang dapat memberikan pelayanan kepuasan konsumen tentunya akan memiliki nilai tambah tersendiri, hal itu dikarenakan pada konsumen yang merasa sesuai dengan keinginannya dalam membeli produk dari online shop merasa sesuai dengan apa yang dikeluarkan konsumen. Konsumen yang puas tentunya akan terus kembali pada online shop karena sesuai dengan peresepsi konsumen dan merasa memberikan hasil yang baik bagi konsumen, namun bagi konsumen yang tidak puas akan pergi begitu saja tanpa kembali lagi sehingga hal tersebut perlahan akan membuat pelaku usaha online shop menjadi sepi pelanggan.

Berdasarkan kondisi di atas, peneliti tertarik untuk mengkaji mengenai kepuasan konsumen online shop dengan berbagai attribute dengan judul penelitian adalah : Implementasi Algoritma C4.5 Untuk Klasifikasi Tingkat Kepuasan Pembeli Online Shop". Objek penelitian adalah konsumen dari online shop Polaroid Vektor & Gift Jogja yang berada di kota Yogyakarta dengan pertimbangan bahwa konsumen merupakan salah satu faktor utama yang sangat dipertimbangkan oleh penjual jika ingin kelangsungan bisnisnya terus berjalan. Dalam hal ini



kepuasan konsumen menjadi hal yang sangat penting diatas segalanya. Untuk itu perlunya dukungan dari beberapa unsur lain agar memberikan pelayanan yang maksimal kepada konsumen. Sehingga dengan jumlah konsumen yang cukup banyak menjadi tolak ukur sebuah online shop dapat memenangkan persaingan diantara para pelaku online shop lainnya.

Penerapan algoritma C4.5 merupakan salah satu solusi pemecahan kasus yang sering digunakan dalam pemecahan masalah pada teknik klasifikasi yang memiliki karakteristik yaitu dengan proses penentuan nilai entropy dan nilai gain dari kemungkinan setiap kriteria yang menjadi acuan keputusan yang dilanjutkan dengan proses perangkingan dari hasil keputusan. Keluaran dari algoritma C4.5 yaitu berupa sebuah pohon keputusan (decision tree). Pohon keputusan adalah sebuah struktur yang dapat digunakan untuk mengubah data menjadi pohon keputusan yang akan menghasilkan aturan-aturan keputusan.

2. TEORITIS

2.1 Algoritma C4.5

Menurut (Sukma Putri Utari, 2015 Volume: IX, Nomor: 3) Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Sedangkan pohon keputusan dapat diartikan sebagai suatu cara untuk memprediksi atau menngelompokkan yang sangat kuat. Rumus algoritma C4.5 terbagi menjadi 2 rumus yaitu untuk mencari nilai gain dan untuk mencari nilai entropy

Rumus untuk mencari nilai gain adalah sebagai berikut:

$$Gaint(S, A) = Entropy(S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} * Entropy(Si)$$
 ...(1)

Rumus untuk mencari nilai entropy adalah sebagai berikut:

$$Entropy(S) = \sum_{i=1}^{n} -pi * log_2 pi \qquad \dots (2)$$

Keterangan:

S : Himpunan Kasus

A : Atribut

n : Jumlah partisi atribut A |Si| : Jumlah kasus pada partisi ke I

|S| : Jumlah kasus dalam S

2.2 Decision Tree

Secara singkat bahwa *decision tree* merupakan salah satu metode klasifikasi pada *text mining*. Klasifikasi adalah proses menemukan kumpulan pola atau fungsi-fungsi yang mendeskripsikan dan memisahkan kelas data satu dengan lainnya, untuk dapat digunakan untuk memprediksi data yang belum memiliki kelas data tertentu (Jianwei Han, 2001). Pohon keputusan dikembangkan untuk membantu pengambil keputusan membuat serangkaian keputusan yang melibatkan peristiwa ketidakpastian. Pohon keputusan adalah suatu peralatan yang mengambarkan secara grafik berbagai kegiatan yang dapat diambil dan dihubungkan dengan kegiatan ini dengan berbagai peristiwa di waktu mendatang yang dapat terjadi. Seperti dalam teknik riset operasi, pohon keputusan tidak akan membuat keputusan bagi pengambil keputusan, kebijakan masih akan diperlukan. Bagaimanapun dalam berbagai situasi yang tepat, penggunaan pohon keputusan akan mengurangi kekacauan potensial dalam suatu masalah kompleks dan memungkinkan pengambil keputusan menganalisis masalah secara rasional (Sutabri, 2005).

3. ANALISA DAN PEMBAHASAN

Untuk mencapai keakuratan dan ketelitian serta informasi dalam penelitian ini, maka pengumpulan data dilakukan dengan cara melihat dari testimoni 100 pelanggan online shop di kota Yogyakarta. Attribute yang digunakan dalam testimoni ini yaitu: (1) Puas, dan (2) Tidak Puas. Metode pengumpulan data yang penulis lakukan di dalam penelitian ini adalah:

1. Wawancara

Pada tahap ini kegiatan yang dilakukan adalah pengumpulan data dengan mengadakan tanya jawab secara langsung dengan owner/pemilik online shop.

2. Testimoni

Pada tahap ini kegiatan yang dilakukan adalah peneliti mengumpulkan data dengan membaca testimoni pelanggan yang telah mereka berikan setelah pelanggan tersebut menerima barang dari online shop.



Pengambilan data ini disesuaikan dengan tujuan penelitian, yaitu membuat model data mining dan mencari tingkat akurasi data.

3.1 Analisa Data

Data kepuasan konsumen online shop memiliki format sebagai berikut:

- 1. Puas
- 2. Tidak Puas

Dari data tersebut maka dapat diketahui bahwa konsumen yang merasa puas akan memberikan testimoni berupa pujian dan kesan-kesan yang positif, sementara konsumen yang merasa tidak puas akan memberikan testimoni namun dengan kesan-kesan yang negatif, sementara untuk konsumen yang tidak memberikan testimoni dianggap tidak menjawab.

Berikut adalah simulasi perhitungan klasifikasi kepuasan pembeli online shop dengan algoritma C4.5 yang menggunakan 15 sampel data testimoni sebagai data training.

DD 1 1 1	D , r	п	T.7
Tabel L	Data	Lestimoni	Konsumen

			Kecepatan	outu Testimoi				_
	Promo	Konsep	Akses	Layout	Kualitas	Pelayanan	Harga	Klasifikasi
	Ya	Menarik	Cepat	Menarik	Baik	Baik	Ya	Puas
	Ya	Tidak	Tidak	Menarik	Tidak	Kurang	Ya	Tidak
	Ya	Menarik	Cepat	Menariuk	Baik	Baik	Ya	Puas
	Ya	Tidak	Cepat	Lumayan	Lumayan	Baik	Kurang	Puas
	Ya	Menarik	Cepat	Lumayan	Lumayan	Kurang	Kurang	Tidak
	Ya	Tidak	Cepat	Tidak	Kurang	Lumayan	Tidak	Tidak
	Tidak	Tidak	Cepat	Tidak	Kurang	Baik	Tidak	Puas
	Ya	Menarik	Cepat	Lumayan	Lumayan	Lumayan	Ya	Puas
	Tidak	Tidak	Cepat	Lumayan	Lumayan	Kurang	Tidak	Tidak
	Ya	Tidak	Cepat	Menarik	Lumayan	Kurang	Lumayan	Puas
	Ya	Menarik	Tidak	Tidak	Tidak	Lumayan	Tidak	Tidak
	Ya	Menarik	Tidak	Tidak	Tidak	Baik	Tidak	Puas
	Ya	Menarik	Cepat	Menarik	Lumayan	Lumayan	Lumayan	Puas
	Ya	Menarik	Cepat	Lumayan	Baik	Baik	Lumayan	Puas
	Ya	Menarik	Cepat	Tidak	Lumayan	Baik	Lumayan	Puas
Presentasi Kepuasan	86.67%	60%	80%	66.67%	80%	73.33%	66.67%	66.67%

Keterangan:

Untuk testimoni terbagi menjadi 2 (dua) yaitu:

- Puas
 - Testimoni puas apabila konsumen memberikan testimoni berupa pujian dan kesan pesan yang positif seperti (puas, ramah, cukup).
- 2. Tidak Puas

Testimoni tidak puas apabila konsumen memberikan testimoni tanpa pujian dan kesan pesan yang negatif seperti (tidak, kurang)

Tabel 2. Penyederhanaan Data Testimoni Pembeli

Produk	Pelayanan	Fasilitas	Klasifikasi
Baik	Baik	Ya	Puas
Baik	Kurang	Tidak	Tidak
Baik	Baik	Ya	Puas
Lumayan	Baik	Ya	Puas
Lumayan	Kurang	Ya	Tidak
Kurang	Lumayan	Tidak	<mark>Tidak</mark>



Kurang	Baik	Ya	Puas
Baik	Lumayan	Ya	Puas
Lumayan	Kurang	Ya	<mark>Tidak</mark>
Baik	Kurang	Ya	Puas
Kurang	Lumayan	Ya	Tidak
Kurang	Baik	Ya	Puas
Baik	Lumayan	Ya	Puas
Baik	Baik	Ya	Puas
Lumayan	Baik	Ya	Puas

Setelah dianalisis, dataset tersebut memiliki 15 kasus yang terdiri dari 10 "PUAS" dan 5 "TIDAK PUAS" yang terdapat pada kolom klasifikasi, langkah selanjutnya yang harus dilakukan adalah menghitung entropinya.

Total Kasus	Sum (Puas)	Sum (Tidak)	Entropi	Gain
15	10	5		

$$\begin{aligned} \text{Entropi(S)} &= \frac{\text{Entropi total (S)} = 10}{15} \text{ x } (2 \log 10 - 2 \log 15) + -\frac{5}{15} \text{ x } (2 \log 5 - 2 \log 15) \\ &= 0,666666666 \text{ x } (3,32192809 - 3,90689059) + 0,33333333 \text{ x } (2,32192809 - 3,90689059) \\ &= 0,666666666 \text{ x } (-0,58496286) + -0,333333333 \text{ x } (-1,5849625) \\ &= 0,38997523 + 0,52832082 \\ &= 0,91830325 \end{aligned}$$

Tabel 3. Hasil Perhitungan Entropi

Total		Tidak Puas	•
Kasus	Puas(S1)	(S2)	Entropi total
15	10	5	0,91830325

Setelah entropi total sudah dihitung, maka langkah selanjutnya adalah melakukan analisis dan perhitungan setiap atribut dan nilai-nilainya.

Tabel 4. Hasil perhitungan total Sum(Puas) Sum(Tidak) Atribut Nilai Gain Sum **Entropi Produk** 0,55990241 Baik 2 Lumayan 2 2 2 Kurang 0,08169673 7 0 Pelayanan Baik Lumayan 2 2 3 Kurang 4 1 0,81127812 0,43529577 **Fasilitas** Ya 13 10 3 0,78956963 Tidak 0 0,23400958

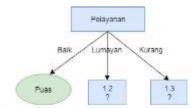
Langkah selanjutnya adalah menghitung gain pada setiap atribut dengan menggunakan rumus berikut ini :

Gain(produk) =
$$\left(\frac{Gain6}{15} \times 0\right) + \left(\frac{4}{15} \times 1\right) + \frac{5}{15} \times 0,97095058$$

= 0 + 0.2666666666 + 0.32365019

= 0,59031685

Setelah nilai gain pada setiap atribut diketahui maka langkah selanjutnya adalah membuat pohon keputusannya. Umtuk membuat pohon keputusan tersebut langkah awalnya kita harus mencari nilai gain yang terbesar. Karena nilai gain (Pelayanan)yang paling besar maka pelayanan menjadi node akar (*root nude*)



Gambar 1. Pohon keputusan node 1 (root node)



Berdasarkan pembentukan pohon keputusan node 1 (root node), node 1.2 dan node 1.3 akan dianalisis lebih lanjut, untuk lebih mudahnya maka node 1.2 akan dianalisis terlebih dahulu. Untuk mempermudahnya tabel 4 difilter dengan data yang memiliki Pelayanan = Lumayan sehingga jadilah tabel berikut.

Tabel 5. Penyederhanaan atribut pelayanan

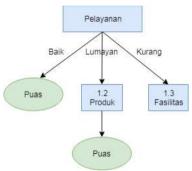
Pelayanan	Fasilitas	Produk	Klasifikasi
Lumayan	Tidak	Kurang	Tidak Puas
Lumayan	Ya	Baik	Puas
Lumayan	Ya	Kurang	Tidak Puas
Lumayan	Ya	Baik	Puas

Selanjutnya, data pada tabel 5 dianalisis dan dihitung kembali entropi atribut pelayanan lumayan serta entropi dan gain setiap atribut lainnya dengan mengecualikan atribut pelayanan yang sudah berada pada jalur diatasnya sehingga membentuk tabel baru yaitu tabel 6.

Tabel 6. Hasil perhitungan node 1.2

Pelayanan	Sum	Sum (Tdk					
Lumayan	(Puas)	Puas)	Entropi				
4	2	2	1				_
Node	Atribut	Nilai	Sum	Puas	Tidak Puas	Entropi	Gain
	Fasilitas	Ya	3	2	1	0,91829581	_
		Tidak	1	0	1	0	0,31127815
1.2	Produk	Baik	2	2	0	0	
		Lumayan	0	0	0	0	
		Kurang	2	2	0	0	1

Dari tabel hasil perhitungan tersebut dapat diketahui bahwa gain Produk memiliki nilai yang paling tinggi. Oleh karena itu node selanjutnya adalah Produk dengan klasifikasi puas.



Gambar 2. Pohon keputusan node 2

Karena node 1.2 sudah menemukan ujung dari setiap cabangnya, maka node 1.2 sudah selesai. Maka selanjutnya adalah menganalisis node 1.3 menggunakan data pada tabel 5

Tabel 7. Penyederhanaan atribut pelayanan

Pelayanan	Fasilitas	Produk	Klasifikasi
Lumayan	Tidak	Kurang	Tidak Puas
Lumayan	Ya	Baik	Puas
Lumayan	Ya	Kurang	Tidak Puas
Lumayan	Ya	Baik	Puas

Selanjutnya, data pada tabel 7 dianalisis dan dihitung kembali entropi atribut fasilitas serta entropi dan gain setiap atribut lainnya dengan mengecualikan atribut pelayanan yang sudah berada pada jalur diatasnya sehingga membentuk tabel baru yaitu tabel 8.

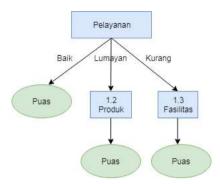
Tabel 8 Hasil perhitungan node 1.3

Pelayanan	Sum	Sum (Tdk	
Kurang	(Puas)	Puas)	Entropi
4	1	3	



Node	Atribut	Nilai	Sum	Puas	Tidak Puas	Entropi	Gain
	Fasilitas	Ya	3	2	1	0,91829581	_
		Tidak	1	0	1	0	0,31127815
1.2	Produk	Baik	2	2	0	0	
		Lumayan	0	0	0	0	
		Kurang	2	2	0	0	1

Dari tabel hasil perhitungan tersebut dapat diketahui bahwa gain Fasilitas memiliki nilai 0,31127815. Karena atribut Produk telah selesai maka selanjutnya menggunakan nilai gain dari atribut Fasilitas untuk membentuk klasifikasi puas pada node 1.3.



Gambar 3. Pohon keputusan node 3

Adapun aturan atau rule yang terbentuk berdasarkan pohon keputusan terakhir seperti pada gambar 3 di atas adalah sebagai berikut :

- a. Jika Pelayanan online shop tersebut = baik, maka secara otomatis pelanggan akan merasa puas.
- b. Jika Pelayanan online shop tersebut = lumayan, maka pada atribut Produk harus bernilai lumayan ataupun baik supaya pelanggan merasa puas.
- c. Jika Pelayanan online shop tersebut = kurang, maka pada atribut Fasilitas harus bernilai Ya/ Baik supaya pelanggan merasa puas.

Untuk menghitung presentase "Puas" dan "Tidak Puas" menggunakan rumus :

Puas = 10/15*100% = 66.7%

Tidak Puas = 5/15*100% = 33.3%

4. KESIMPULAN

Dari uraian pada bab-bab yang sudah dibahas sebelumnya dapat ditarik kesimpulan sebagai berikut:

- 1. Dengan menggunakan algoritma C4.5 untuk melakukan perhitungan, ternyata mampu mengklasifikasi tingkat kepuasan pelanggan pada online shop tersebut dengan tingkat keakuratan sebesar91%, dengan nilai presisi pada prediksi puas sebesar 66.67% dan nilai presisi pada prediksi tidak puas sebesar 33.3%.
- 2. Pohon Keputusan yang dihasilkan oleh teknik klasifikasi Algoritma C4,5 dari variabel yang memiliki gain tertinggi yaitu Pelayanan, sehingga pemilik online shop dapat menjadikan acuan untuk memperbaiki kualitas online shop tersebut.

Berdasarkan kesimpulan di atas, peneliti mengusulkan beberapa rekomendasi bagi pengembang pendidikan, peneliti lebih lanjut, dan pihak terkait sebagai berikut: Melakukan evaluasi internal terhadapat data konsumen dan membandingkat dengan beberapa algoritma lainnya sehingga terus mendapatkan model terbaru dan rule terbaru yang lebih baik lagi.

REFERENCES

- [1] E. Buulolo, N. Silalahi, Fadlina, R. Rahim, C4.5 Algorithm To Predict the Impact of the Earthquake, Int. J. Eng. Res. Technol. 6 (2017) 10–15
- [2] Harry Dhika, Fitriana Destiawati, Aswin Fitriansyah, IMPLEMENTASI ALGORITMA C4.5 TERHADAP KEPUASAN PELANGGAN, Vol 6, No.1, Th, 2016
- [3] Darudiato, S., Santoso, S. W., & Wiguna, S. (2002). BUSINESS INTELLIGENCE: KONSEP DAN METODE, (9), 63-67.
- [4] H. Widayu, S.D. Nasution, N. Silalahi, M. Mesran, DATA MINING UNTUK MEMPREDIKSI JENIS TRANSAKSI NASABAH PADA KOPERASI SIMPAN PINJAM DENGAN ALGORITMA C4.5, MEDIA Inform. BUDIDARMA. 1 (2017). http://ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/323/273 (accessed June 27, 2017).



- [5] Gorunescu, F. (2011). Data mining: Concepts, models and techniques. *Intelligent Systems Reference Library*, 12. http://doi.org/10.1007/978-3-642-19721-5
- [6] M. Maharani, N.A. Hasibuan, N. Silalahi, S.D. Nasution, M. Mesran, S. Suginam, D.U. Sutiksno, H. Nurdiyanto, E. Buulolo, Y. Yuhandri, IMPLEMENTASI DATA MINING UNTUK PENGATURAN LAYOUT MINIMARKET DENGAN MENERAPKAN ASSOCIATION RULE, J. Ris. Komput. 4 (2017) 6–11. https://www.researchgate.net/publication/312495968.
- [7] Han, J., Kamber, M., & Pei, J. (2012). Data Mining. Data Mining. http://doi.org/10.1016/B978-0-12-381479-1.00001-0
- [8] R. Rahim, I. Zufria, N. Kurniasih, M.Y. Simargolang, A. Hasibuan, D.U. Sutiksno, R.F. Nanuru, J.N. Anamofa, A.S. Ahmar, A. Daengs GS, C4.5 Classification Data Mining for Inventory Control, Int. J. Eng. Technol. 7 (2018) 68–72. doi:10.14419/ijet.v7i2.3.12618.
- [9] S. Defit, G.W. Nurcahyo, Penerapan Algoritma C4 . 5 Pada NUPTK Untuk Menentukan Pola Sertifikasi Guru Dengan Menggunakan Metode Klasifikasi Decision Tree (Studi Kasus : Pada Dinas Pendidikan Pemuda dan Olahraga Kabupaten Rokan Hulu) Of Computer Science, 1 (2015) 69–83.
- [10] Y.A. Fiandra, S. Defit, Yuhandri, Penerapan Algoritma C4.5 untuk Klasifikasi Data Rekam Medis berdasarkan International Classification Diseases (ICD-10), J. RESTI (Rekayasa Sist. Dan Teknol. Informasi). 1 (2017) 82–89. doi:https://doi.org/10.29207/resti.v1i2.48.