



PENERAPAN DATA MINING DALAM PEMILIHAN LAPTOP MENGGUNAKAN METODE ALGORITMA C4.5

Anggi Wijaya Sari Raja Gukguk¹
Sunarsan Sitohang²

¹Alumni Program Studi Teknik Dan Komputer, Universitas Putera Batam

²Dosen Program Studi Teknik Dan Komputer, Universitas Putera Batam
email:pb150210214@upbatam.ac.id

ABSTRACT

The development of globalization today greatly affects various fields, one of which is technology. A laptop or handheld computer is a computer that can be carried anywhere because it has a small size and light weight. The use of a laptop is the same as a computer. Good Cell Batam store is one of the stores that provide various types of laptops. Because many types of laptops make it difficult for consumers to choose the right laptop so that the store has difficulty providing information to consumers. Then we need data mining techniques and C4.5 algorithm methods in classifying the right laptop selection based on price variables, RAM capacity, processor, size and battery life. The WEKA application is used to test the accuracy of the data processed using the C4.5 algorithm. These results are obtained by finding the entropy and gain values of all variables so as to produce a decision tree where there are three variables that greatly affect the choice of laptop, namely price, processor and battery life from a total of 80 data where there are 53 choices of the right laptop and 27 choices of the right laptop. not exactly.

Keywords: *Algoritma C45., Data Mining., WEKA.*

PENDAHULUAN

Perkembangan globalisasi saat ini sangat mempengaruhi berbagai bidang, salah satunya yaitu bidang teknologi. Laptop atau komputer tentu merupakan sebuah komputer yang dapat dibawa kemana saja serta memiliki ukuran fleksibel dan enteng. Laptop mempunyai kegunaan mirip dengan komputer desktop (*Desktop Computers*) seperti umumnya yaitu dapat memudahkan mengerjakan tugas dan pekerjaan dengan kilat dan baik.

(Ginting, 2015). Toko *Good Cell* Batam yang beralamat di Lantai Dasar Blok K1 No. A2A- Mitra Mall, Batu Aji Kota Batam ini adalah salah satu toko yang menyediakan berbagai merk laptop.

Toko *Good Cell* Batam termasuk toko yang paling ramai dikunjungi oleh banyak orang yang akan membeli maupun sekedar tanya-tanya laptop. Sehingga toko *Good Cell* Batam sering kali mengalami kesulitan dalam membantu konsumen menentukan jenis dan tipe laptop yang akan dibeli. Hal ini terjadi

karena banyaknya jenis dan harga bersaing yang ditawarkan oleh pihak produsen laptop. Untuk memilih laptop yang tepat sesuai dengan kebutuhan dan dana yang disiapkan oleh konsumen, maka memerlukan suatu analisa yang akurat yang dapat dijadikan bahan pertimbangan dari banyaknya kriteria dan faktor.

Kriteria - kriteria yang dapat dijadikan acuan dalam memilih suatu laptop antara lain yaitu harga, kapasitas RAM, prosesor, ukuran dan daya tahan batrai dari laptop tersebut.

KAJIAN PUSTAKA

2.1 Knowledge Discovery In Database (KDD)

Knowledge Discovery in Database (KDD) adalah merupakan metode dalam mendapatkan keahlian basis data yang ada. Dalam *database* diperoleh grafik-grafik yang saling terhubung satu sama lainnya. Istilah *Knowledge Discovery in Database* (KDD) dan *data mining* biasanya dimanfaatkan secara bergantian dalam memperjelas tahapan penggalian informasi tertanam pada *database* besar. (Mardi, 2018)

2.2 Data Mining

Data mining memanfaatkan teknik tertentu dalam memperoleh skema maupun maklumat memikat pada sebuah data. Tahapan dari *Data mining* yang digunakan dalam mengekstrak serta mengidentifikasi suatu informasi yang dibutuhkan oleh *database* besar dengan cara teknik statistik, matematika, kecerdasan buatan, dan *machine*. (Mardi, 2018)

2.3 Metode data mining

Algoritma C4.5 merupakan metode yang menghasilkan pohon keputusan berkenaan dengan data yang telah didapatkan terlebih dahulu. Pohon keputusan yang atribut serta nilai *gain* dan *entropy* tertinggi sebagai konsep pertama algoritma ini yaitu mempunyai nilai *gain* tertinggi berdasarkan nilai *entropy* bagian dari proses pengklasifikasian. Rumus *gain* tertinggi

dapat dilihat sebagai berikut ini (Elisa & Ervina, 2018). *Waikato Environment for Knowledge Analysis* (WEKA) *Waikato Environment for Knowledge Analysis* adalah sebuah paket *tools machine learning* praktis yang disingkat WEKA. Keuntungan menggunakan WEKA yaitu khususnya klasifikasi yang dilandasi oleh pendekatan-pendekatan sangat mampu menyelesaikan masalah-masalah data *mining* dalam dunia nyata. (Pujiono, Amborowati, & Suyanto, 2013)

2.4 Penelitian Terdahulu

Penelitian terdahulu merupakan acuan dalam sebuah penelitian. Peneliti juga menemukan beberapa hasil penelitian terdahulu yang berhubungan dengan *algoritma C4.5* yaitu menurut (Novianti & Bahri, Tedy Rismawan, 2016) dalam jurnal berjudul "Implementasi Data Mining Dengan Algoritma C4.5 Untuk Penjurusan Siswa (Studi Kasus: Sma Negeri 1 Pontianak)" menyatakan bahwa klasifikasi penjurusan siswa berdasarkan penggolongan siswa berdasarkan kemampuan (nilai), bakat serta minat yang hampir sama sehingga pembelajaran yang akan diajarkan akan lebih fokus dan teratur. Salah satu institusi pendidikan yang berada di kota Pontianak yaitu SMA Negeri 1 dimana sekolah tersebut memiliki dua penjurusan yaitu IPA dan IPS. Kurikulum 2013 saat ini yang mengatur proses penyeleksian berdasarkan nilai rata-rata rapor SMP, nilai UAS SMP serta nilai tes MTK, IPA dan IPS. Metode *algoritma C4.5* dimanfaatkan sebagai proses dalam pengklasifikasian data sehingga mampu menghasilkan pohon keputusan berdasarkan proses penyeleksian data sebelumnya sehingga menghasilkan tingkat akurasi sebesar 89.74%.

METODE PENELITIAN

3.1 Desain Penelitian

Proses awal yang dilakukan yaitu pengumpulan data bertujuan untuk memperoleh informasi yang diperlukan melalui proses wawancara bersama pihak toko Good Cell Batam, melakukan observasi langsung serta study pustaka.

Data yang diperoleh selanjutnya akan diseleksi terlebih dahulu selanjutnya diolah kedalam metode algoritma c4.5 serta diuji menggunakan aplikasi WEKA sehingga memperoleh hasil mengenai pemilihan laptop yang tepat berdasarkan variabel yang telah ditentukan yaitu harga, kapasitas RAM, prosessor, ukuran dan daya tahan baterai.

3.2 Metode Perancangan Sistem

Tahapan pertama yang akan dilakukan dalam proses penyelesaian penelitian ini adalah untuk mendapatkan data penjualan laptop yang dilihat dari harga, kapasitas RAM, prosesor, ukuran dan daya tahan batrai dari laptop pada Toko *Good Cell* Batam, kemudian dilakukan data yang diperoleh diseleksi terlebih dahulu guna mendapatkan data yang sesuai dan diproses, maka selanjutnya data penjualan tersebut diproses kedalam data *mining*

menggunakan metode *algoritma* C4.5, langkah terakhir yaitu melakukan pengujian melalui aplikasi WEKA, untuk mengetahui keakuratan antara proses perhitungan manual tersebut sesuai dengan hasil pengujian melalui aplikasi WEKA dalam menentukan pilihan laptop yang tepat berdasarkan data yang diolah.

HASIL DAN PEMBAHASAN

4.1 Arsitektur sistem dan pembahasan

Data yang akan diolah kedalam penelitian ini berupa data penjualan laptop pada toko *Good cell* Batam dengan dipengaruhi oleh variabel harga, kapasitas RAM, prosessor, ukuran dan daya tahan baterai serta memiliki 2 keputusan yaitu tepat dan tidak tepat. Berikut ini adalah data dari toko *Good Cell* Batam

Tabel 1.Format Data Penjualan

No	Harga	Kapasitas Ram	Prosesor	Ukuran	Daya Tahan Batrai	Keputusan
1	Murah	Besar	Tinggi	Besar	Hemat	Tepat
2	Sedang	Besar	Tinggi	Besar	Hemat	Tepat
3	Sedang	Kecil	Tinggi	Kecil	Tidak Hemat	Tepat
4	Sedang	Besar	Tinggi	Besar	Hemat	Tepat
5	Sedang	Besar	Tinggi	Besar	Hemat	Tepat

Sumber: Data Peneliti, 2021

Tabel 1 menjelaskan bahwa data yang diperoleh tersebut berupa data penjualan yang bersumber dari toko *Good Cell* Batam yang sudah melalui proses pengumpulan data seperti wawancara dan observasi serta sudah melalui proses analisa data, seleksi data serta pembersihan data sehingga menghasilkan variabel harga, kapasitas

RAM, prosessor, ukuran dan daya tahan baterai. Namun, data yang akan diolah sebagai sampel dalam penelitian ini harus melakukan proses transformasi terlebih dahulu dimana setiap variabel diubah kebentuk khusus sebelum diaplikasikan. Maka hasil pra proses data Klasifikasi atribut dari masing-masing variabel dapat dilihat pada tabel berikut:

Tabel 2.Format Data Pra-Proses Penjualan

No	Harga	Kapasitas Ram	Prosesor	Ukuran	Daya Tahan Batrai	Keputusan
1	MUR	BSR	T	B	H	Tepat

2	S	BSR	T	B	H	Tepat
3	S	KEC	T	K	TH	Tepat
4	S	BSR	T	B	H	Tepat
5	S	BSR	T	B	H	Tepat

Sumber: Data Peneliti, 2021

Tabel 2 menjelaskan bahwa semua data pejualan yang bersumber dari toko Good Cell Batam sudah diubah kedalam bentuk khusus melalui proses transformasi data dimana variabel harga memiliki 3 klasifikasi atribut yaitu mahal (MAH), murah (MUR) dan sedang (S), selanjutnya variabel kapasitas RAM memiliki 2 klasifikasi atribut yaitu besar (BSR) dan kecil (KEC), selanjutnya variabel prosessor memiliki 2 klasifikasi atribut yaitu tinggi (T) dan rendah (R), selanjutnya variabel ukuran memiliki 2

klasifikasi atribut yaitu besar (B) dan kecil (K), selanjutnyavariabel terakhir yaitu daya tahan baterai memiliki 2 klasifikasi atribut yaitu hemat (H) dan tidak hemat (TH), dengan kputusan tepat dan tidak tepat. Setelah proses transformasi di lakukan maka langkah selanjutnya yaitu melakukan perhitungan *node* 1 dengan mencari nilai *entropy* dan *gain* dari masing-masing variabel seperti yang ditunjukkan pada tabel berikut ini:

Tabel 3.Hasil Perhitungan *Node* 1

		Jumlah (s)	Ya (si)	Tidak (si)	Entropy	Gain
Total		80	53	27	0.922406262	
Harga	Mahal	21	11	10	0.998363673	0.179920462
	Sedang	46	39	7	0.615253891	
	Murah	13	3	10	0.779349837	
Kapasitas RAM	Besar	52	43	9	0.664695305	0.161254228
	Kecil	28	10	18	0.940285959	
Prosesor	Tinggi	55	53	2	0.225363639	0.76746876
	Rendah	25	0	25	0	
Ukuran	Besar	52	43	9	0.664695305	0.161254228
	Kecil	28	10	18	0.940285959	
Daya Tahan batrai	Hemat	54	45	9	0.650022422	0.194231344
	Tidak Hemat	26	8	18	0.89049164	

Sumber: Data Peneliti, 2021

Tabel 3 menjelaskan hasil dari perhitungan *node* 1 dimana total data terdapat 80 data pejualan dengan jumlah pemilihan tepat sebanyak 53 data

dan jumlah pemilihan tidak tepat terdapat 27 data dan nilai entropy total 0.922406262. maka perhitungan masing-masing variabel sebagai berikut:

Entropy Harga

$$Entropy(MAH) = \left(\frac{-11}{21} * \log_2\left(\frac{11}{21}\right)\right) +$$

$$\left(\frac{-10}{21} * \log_2\left(\frac{10}{21}\right)\right) = 0.998363673$$

$$Entropy(S) = \left(\frac{-39}{46} * \log_2\left(\frac{39}{46}\right)\right) +$$

$$\left(\frac{-7}{46} * \log_2\left(\frac{7}{46}\right)\right) = 0.615253891$$

$$Entropy(MUR) = \left(\frac{-3}{13} * \log_2\left(\frac{3}{13}\right)\right) +$$

$$\left(\frac{-10}{13} * \log_2\left(\frac{10}{13}\right)\right) = 0.779349837$$

Entropy Kapasitas RAM

$$Entropy(BSR) = \left(\frac{-43}{52} * \log_2\left(\frac{43}{52}\right)\right) +$$

$$\left(\frac{-9}{52} * \log_2\left(\frac{9}{52}\right)\right) = 0.664695305$$

$$Entropy(KEC) = \left(\frac{-10}{28} * \log_2\left(\frac{10}{28}\right)\right) +$$

$$\left(\frac{-18}{28} * \log_2\left(\frac{18}{28}\right)\right) = 0.940285959$$

Entropy Processor

$$Entropy(T) = \left(\frac{-53}{55} * \log_2\left(\frac{53}{55}\right)\right) +$$

$$\left(\frac{-2}{55} * \log_2\left(\frac{2}{55}\right)\right) = 0.225363639$$

$$Entropy(R) = \left(\frac{-0}{25} * \log_2\left(\frac{0}{25}\right)\right) +$$

$$\left(\frac{-25}{25} * \log_2\left(\frac{25}{25}\right)\right) = 0$$

Entropy Ukuran

$$Entropy(B) = \left(\frac{-43}{52} * \log_2\left(\frac{43}{52}\right)\right) +$$

$$\left(\frac{-9}{52} * \log_2\left(\frac{9}{52}\right)\right) = 0.664695305$$

$$Entropy(K) = \left(\frac{-10}{28} * \log_2\left(\frac{10}{28}\right)\right) +$$

$$\left(\frac{-18}{28} * \log_2\left(\frac{18}{28}\right)\right) = 0.940285959$$

Entropy Daya Tahan Baterai

$$Entropy(H) = \left(\frac{-45}{54} * \log_2\left(\frac{45}{54}\right)\right) +$$

$$\left(\frac{-9}{54} * \log_2\left(\frac{9}{54}\right)\right) = 0.650022422$$

$$Entropy(TH) = \left(\frac{-8}{26} * \log_2\left(\frac{8}{26}\right)\right) +$$

$$\left(\frac{-18}{26} * \log_2\left(\frac{18}{26}\right)\right) = 0.89049164$$

Gain Total Atribut Harga

$$Gain (total) = 0.922406262 - \left(\left(\frac{21}{80} * 0.998363673\right) + \left(\frac{46}{80} * 0.615253891\right) + \left(\frac{13}{80} * 0.779349837\right)\right) = 0,179920462$$

Gain Total Kapasitas RAM

$$Gain (total) = 0,922406262 - \left(\left(\frac{52}{80} * 0,664695305\right) + \left(\frac{28}{80} * 0,940285959\right)\right) = 0,161254228$$

Gain Total Processor

$$Gain (total) = 0.922406262 - \left(\left(\frac{55}{80} * 0.225363639\right) + \left(\frac{25}{80} * 0\right)\right) = 0,76746876$$

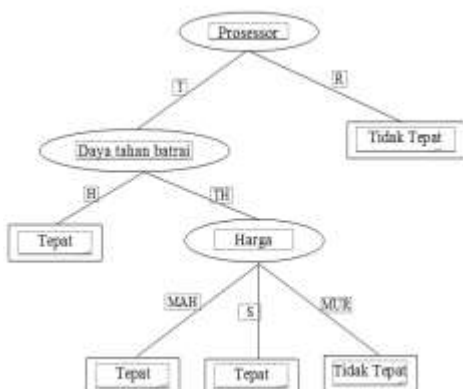
Gain Total ukuran

$$Gain (total) = 0,922406262 - \left(\left(\frac{52}{80} * 0,664695305\right) + \left(\frac{28}{80} * 0,940285959\right)\right) = 0,161254228$$

Gain Total Daya Tahan Batrai

$$Gain (total) = 0,922406262 - \left(\left(\frac{54}{80} * 0,664695305\right) + \left(\frac{26}{80} * 0.89049164\right)\right) = 0,194231344$$

Berdasarkan pencarian nilai *node* terakhir dari tabel 5, semua variabel dan atribut menghasilkan nilai *entropy* dan *gain* yaitu 0, Sehingga harga akan menjadi turunan terakhir dari pohon keputusan. Sehingga pohon keputusan yang dihasilkan adalah sebagai berikut:



Gambar 1. Hasil Pohon keputusan
Sumber: Data Peneliti, 2021

Berdasarkan gambar 1 hasil akhir dari pohon keputusan, maka dapat disimpulkan sebagai berikut

1. Jika prosessor rendah maka pemilihan tidak tepat
2. Jika prosessor tinggi dan daya tahan baterai hemat maka pemilihan tepat
3. Jika prosessor tinggi, daya tahan baterai tidak hemat dan harga murah maka pemilihan tidak tepat
4. Jika prosessor tinggi, daya tahan baterai tidak hemat dan harga sedang maka pemilihan tepat
5. Jika prosessor tinggi, daya tahan baterai tidak hemat dan harga mahal maka pemilihan tepat.

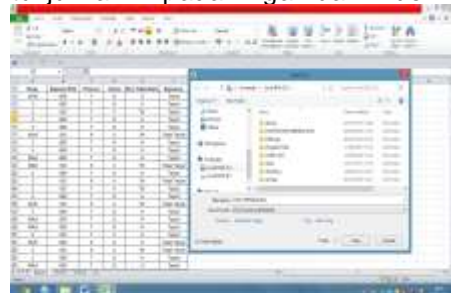
Berdasarkan kesimpulan diatas maka dapat dijelaskan bahwa yang menjadi faktor-faktor utama dalam pemilihan laptop padaa toko *Good Cell* Batam dapat dilihat dari segi prosessor, daya tahan baterai serta harga. Langkah selanjutnya yaitu menguji kebenaran serta keakuratan data tersebut melalui aplikasi yaitu menggunakan aplikasi WEKA.

4.2 Pengujian Menggunakan Aplikasi WEKA

Setelah pencarian nilai *gain* dan *entropy* dilakukan sehingga menghasilkan pohon keputusan, selanjutnya adalah melalukan pengetesan untuk mengetahui fakta serta keakuratan data-data apakah hasil analisa tersebut sudah sesuai atau tidak

sesuai dengan kebutuhan dalam penelitian ini. Maka langkah selanjutnya peneliti akan melakukan pengujian menggunakan aplikasi yaitu WEKA dan langkah-langkah sebagai berikut:

Variabel-variabel dari semua data dibuat kedalam aplikasi *microsoft excel* dengan format penyimpanan *save as type CSV (Comma Delimited)* kemudian klik *save* . Seperti yang tunjukkan pada gambar berikut:



Gambar 2. Format Data *Microsoft Excel*
 (Sumber: Data Peneliti, 2021)

2. Kemudian buka aplikasi WEKA pada laptop atau komputer, lalu klik *Explorer*. Dapat dilihat pada dibawah ini:



Gambar 3.Halaman Awal aplikasi WEKA
 (Sumber: Data Peneliti, 2021)

3. Selanjutnya peneliti akan mengambil data yang sudah di simpan sebelumnya, dengan cara klik *open file*, kemudian ganti *Files of Type* kedalam format *CSV data files (*.csv)* pada data yang tersimpan. lalu klik *Open*



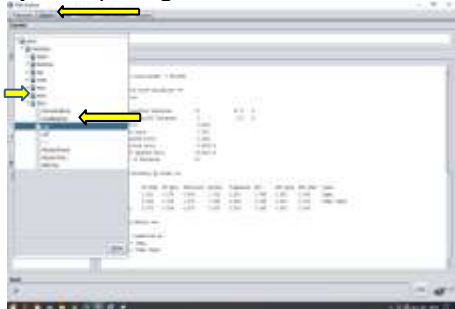
Gambar 4 .Membuka Data Pada Aplikasi WEKA
(Sumber: Data Peneliti, 2021)

4. Maka semua variabel yang digunakan dalam penelitian akan muncul, lalu klik *All* untuk memasukkan variabel ke proses klasifikasi. seperti gambar sberikut:



Gambar 5.Pemilihan Variabel
(Sumber: Data Peneliti, 2021)

5. Kemudian klik menu *Classify* kemudian klik *choose*, selanjutnya klik *trees* dan pilih *J48*. Maka tampilan akan menjadi seperti gambar dibawah ini:



Gambar 6.Pemilihan Model *Classify*
(Sumber: Data Peneliti, 2021)

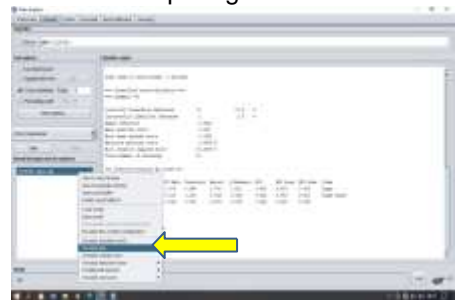
6. Klik *start* kemudian keluar *output* dibagian kanan yaitu hasil dari pengetesan *WEKA*. berita yang ditampilkan oleh *classify output* yaitu merupakan total perkara dan hasilnya

serta banyaknya bagian dari potoh keputusan tersebut. Dapat dilihat pada gambar dibawah ini:



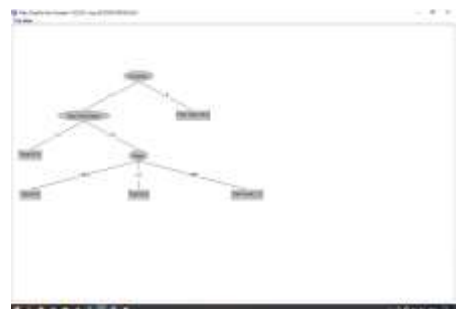
Gambar 7.*Classify Output* pada WEKA
(Sumber: Data Peneliti, 2021)

7. Langkah terakhir dalam pengujian ini adalah menyaksikan pohon keputusan dengan mengklik kanan pada kolom *result list*, lalu pilih dan klik *visualize tree* seperti gambar berikut.



Gambar 8.*Visualize Tree* Pada WEKA
(Sumber: Data Peneliti, 2021)

8. Maka hasil akhir dari pohon keputusan yang telah diproses dengan aplikasi WEKA yaitu sebagai berikut:



Gambar 9.Hasil *Visualize Tree* WEKA
(Sumber: Data Peneliti, 2021)



KESIMPULAN

5.1 Kesimpulan

Berkenaan dengan hasil yang telah didapatkan maka dapat disimpulkan bahwa:

1. Dengan adanya teknik data mining khususnya menggunakan metode algoritma C45 yang melibatkan 80 data penjualan laptop yang telah dihitung maka memperoleh hasil yaitu 53 data dengan pemilihan tepat dan terdapat 27 data dengan pemilihan tidak tepat dari pemilihan laptop pada toko *Good Cell* Batam dengan dipengaruhi oleh 5 variabel yaitu variable harga, kapasitas RAM, processor, ukuran dan daya tahan batrai serta memiliki 2 keputusan yaitu tepat dan tidak tepat.

1. Melalui perhitungan manual dengan menggunakan metode algoritma C45 dan diuji melalui aplikasi WEKA dengan 80 data penjualan berdasarkan 5 variabel, sehingga memperoleh kebenaran serta keakuran sesuai dengan perhitungan manual dan aplikasi WEKA dalam pemilihan laptop pada toko *Good Cell* Batam.

2. Dengan adanya teknik data mining dan metode algoritma C45 akan dapat mempercepat pengolahan data penjualan sehingga memiliki fungsi penting dalam membantu memberikan informasi dimasa yang akan datang dalam pemilihan laptop yang tepat.

DAFTAR PUSTAKA

- Azwanti, N. (2018). ANALISA ALGORITMA C4.5 UNTUK MEMPREDIKSI PENJUALAN MOTOR PADA PT. CAPELLA DINAMIK NUSANTARA CABANG MUKA KUNING. *Informatika Mulawarman: Jurnal Ilmiah Ilmu Komputer*, Vol. 13, N.
- Elisa, E., & Ervina. (2018). MOTOR BEKAS YANG AKAN DIJUAL DI

AMAZONE MOTOR Private vehicles are very important for Indonesian people nowadays . Most Indonesians prefer motorbikes as vehicles to support their daily activities . Many people also buy used motorbikes for various reasons.

- Mardi, Y. (2018). Data Mining : Klasifikasi Menggunakan Algoritma C4.5. *Jurnal Edik Informatika Penelitian Bidang Komputer Sains Dan Pendidikan Informatika*, V2.i2(213-.
- Novianti, B., & Bahri, Tedy Rismawan, S. (2016). IMPLEMENTASI DATA MINING DENGAN ALGORITMA C4.5 UNTUK PENJURUSAN SISWA (STUDI KASUS: SMA NEGERI 1 PONTIANAK). *Jurnal Coding, Sistem Komputer Untan, Volume 04*,.
- Pujiono, S., Amborowati, A., & Suyanto, M. (2013). ANALISIS KEPUASAN PUBLIK MENGGUNAKAN WEKA DALAM MEWUJUDKAN GOOD GOVERNANCE DI KOTA YOGYAKARTA. *JURNAL DASIS ISSN*, Vol. 14 No.

	<p>Biodata Penulis Pertama Nama: Anggi Wijaya SariRaja Gukguk Merupakan Mahasiswa dari Prodi Teknik Dan Komputer Universitas Putera Batam</p>
	<p>Biodata Penulis Kedua Nama: Sunarsan Sitohang Merupakan dosen dari Prodi Teknik Dan Komputer Universitas Putera Batam</p>