

# Penerapan Data Mining Pada Ekspor Kebutuhan Ikan Dengan Menggunakan Metode Algoritma C5.0

Mhd. Reza Pratama

Program Studi Teknik Informatika, Fakultas Ilmu Komputer & Teknologi Informasi, Universitas Budi Darma, Medan, Indonesia  
Jl. Sisingamangaraja No.338, Siti Rejo I, Kec. Medan Kota, Kota Medan, Sumatera Utara, Indonesia  
Email: rezapratamamhammad24@gmail.com

**Abstrak**– Perusahaan yang berkembang pesat tentunya akan sulit akan mengelola stok barang. Penelitian ini dilakukan untuk memperkirakan adanya ekspor ikan atau tidaknya ikan. Data pemesanan akan menjadi landasan untuk menentukan angka persediaan barang, guna mengembangkan bisnis distributor yang masuk. Data mining bisa digunakan untuk mengetahui berapa banyaknya stok ikan dengan perhitungan sistem. Dan perlunya mengetahui berapa banyak mengetahui berapa banyak stok yang harus dimiliki perusahaan untuk memenuhi permintaan dari luar negeri. Prediksi dilakukan untuk memperkirakan nilai untuk masa yang akan datang untuk mengambil keputusan. Penelitian ini menggunakan metode Algoritma C5.0 diharapkan mampu untuk memprediksi stok ekspor ikan.

**Kata Kunci:** Data Mining, Algoritma C5.0, Prediksi

**Abstract**– Companies that are growing rapidly will certainly find it difficult to manage stock of goods. This research was conducted to estimate the existence of fish exports or not. Order data will be the basis for determining inventory figures, in order to develop the business of incoming distributors. Data mining can be used to find out how many fish stocks are with system calculations. And the need to know how much to know how much stock the company must have to meet demand from abroad. Prediction is done to estimate the value for the future to make a decision. This study uses the C5.0 Algorithm method, which is expected to be able to predict fish export stocks.

**Keywords:** Data Mining, C5.0 Algorithm, Prediction

## 1. PENDAHULUAN

Data *mining* adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data *mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengestraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. Dengan adanya *data mining* maka akan didapatkan suatu permata berupa pengetahuan di dalam kumpulan data-data yang banyak jumlahnya.[1]

Prediksi suatu proses memperkirakan tentang sesuatu yang terjadi di masa depan berdasarkan data informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya atau selesih antara sesuatu yang terjadi dengan hasil perkiraan dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi.[2]

Perusahaan yang bergerak dalam dunia bisnis ekspedisi muatan kapal laut. Menganalisa prediksi kebutuhan ekspor ikan dibutuhkan untuk mengetahui perusahaan lain dan bisa bersaing dalam dunia bisnis ini. Ini diakibatkan minat pembeli tidak sesuai dengan persediaan barang yang ada, jika kelebihan stock mengakibatkan perusahaan sering mengalami kerugian yang tidak sesuai harga pasaran atau tidak laku, dan jika kekurangan stock ikan membuang kesempatan yang ada. Dalam bisnis kerugian sebisa mungkin harus diminimalkan jadi dengan menganalisa prediksi kebutuhan ekspor ikan dengan menggunakan data mining, diharapkan mampu memberikan informasi yang membantu kegiatan bisnis sehingga perusahaan tidak mengalami kerugian dan mendapatkan untung.

Algoritma C5.0 ini, merupakan salah satu teknik untuk penyelesaian masalah dalam mengeksport ikan pada Perusahaan, masalah yang terjadi pada perusahaan yaitu dalam mengeksport ikan terkadang mengalami kerugian yang besar, karna tidak mengetahui kapan terjadi kerugian dan yang memesan ikan sementara stok masih banyak. Berdasarkan masalah tersebut, maka di terapkan algoritma C5.0 dalam memprediksi kebutuhan ekspor ikan bisa memberikan informasi yang akurat dan bisa dijadikan bahan pertimbangan perusahaan dalam stock kebutuhan ikan yang akan diekspor, Agar tidak terjadi penumpukan stock ikan yang membuat perusahaan rugi dan bisa memberikan informasi kepada konsumen.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Tahapan dalam penelitian yang dilaksanakan oleh penulis dengan beberapa langkah yaitu:

a. Studi Pustaka (*Library Research*)

Studi Pustaka dilakukan dengan cara mempelajari teori-teori literatur dan buku-buku yang berhubungan dengan objek kajian sebagai dasar dalam penelitian, dengan tujuan memperoleh dasar teoritis gambaran dari apa yang dilakukan.

b. Pengamatan (*observation*)

Penulis melakukan pengamatan langsung ditempat atau lokasi, kemudian dianalisis berdasarkan kebutuhan-kebutuhan yang berhubungan dengan data ekspor kebutuhan ikan.

c. Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data dengan cara menanyakan langsung kepada pemilik perusahaan atau karyawan mengenai data penerapan data mining pada ekspor kebutuhan ikan sebagai bahan untuk dianalisa dan diimplementasikan menggunakan metode c5.0.

d. Analisa

Upaya mencari dan menata secara sistematis catatan berdasarkan hasil *observasi* dan studi *literature* untuk peningkatan pemahaman penelitian tentang kasus yang diteliti sebagai temuan untuk orang lain.

e. Pengujian

Menerapkan metode Algoritma C5.0 setelah dilakukan proses analisa data untuk mendapatkan hasil penerapan data mining pada ekspor kebutuhan ikan, sesuai dengan langkah-langkah metode Algoritma C5.0.

f. Penyusunan Laporan

Setelah mendapatkan hasil analisa dengan implelementasi metode yang digunakan, maka akan diambil kesimpulan dan menyusun skripsi.

## 2.2 Data Mining

Tan (2006) mendefinisikan data mining sebagai proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. *Data mining* juga dapat diartikan sebagai pengekstraan informasi baru yang diambil dari bongkahan data yang besar yang membantu dalam pengambilan keputusan. Istilah *data mining* kadang disebut juga *knowledge discovery*. Salah satu teknik yang dibuat dalam data mining adalah bagaimana menelusuri data yang ada untuk membangun sebuah model, kemudian menggunakan model tersebut agar dapat mengenali pola data yang lain yang tidak berada dalam basis data yang tersimpan. Kebutuhan untuk prediksi juga dapat memanfaatkan teknik ini. Dalam *data mining*, pengelompokan data juga bisa dilakukan. Tujuannya adalah agar kita dapat mengetahui tindak lanjut berikutnya yang dapat diambil. Semua hal tersebut bertujuan mendukung kegiatan operasional perusahaan sehingga tujuan akhir perusahaan diharapkan dapat tercapai[5].

## 2.3 Algoritma C5.0

Algoritma C5.0 merupakan merupakan penyempurnaan dari algoritma terdahulu yang dibentuk oleh Ross Quinlan pada tahun 1987, yaitu algoritma ID3 dan C4.5. Dalam memilih atribut untuk pemecah objek dalam beberapa kelas harus dipilih atribut yang menghasilkan information gain paling besar. algoritma C5.0 menghasilkan tingkat keakuratan yang lebih tinggi dalam hal prediksi Penggunaan algoritma C5.0 dapat menghasilkan model prediksi dengan hasil tingkat akurasi yang lebih tinggi. Algoritma C5.0 diharapkan proses penggalian informasi lebih cepat dan optimal dengan kapasitas data yang lebih besar, sehingga kesalahan yang ditimbulkan dalam pengambilan keputusan lebih diminimalkan[9]. Strategi pengembangan decision tree dengan menggunakan alogritma C5.0 adalah sebagai berikut:

- Pada tahap awal, tree digambarkan sebagai node tunggal yang merepresentasikan training set.
- Jika sampel seluruhnya berisi kelas yang sama, maka node tersebut menjadi leaf dan dilabeli dengan kelas tersebut.
- Jika tidak, algoritma dengan menggunakan ukuran berbasis entropi (information gain) akan memilih variabel prediktor yang akan memisahkan record ke dalam kelas-kelas individual. Variabel tersebut menjadi variabel tes atau keputusan pada node tersebut.
- Cabang dikembangkan untuk tiap nilai yang diketahui dari variabel tes, dan sampel dipartisi berdasarkan cabang tersebut.
- Algoritma menggunakan proses yang sama secara rekursif membentuk decision tree.
- Partisi rekursif berakhir hanya ketika satu dari kondisi-kondisi berikut terpenuhi:
  - Seluruh record pada node tertentu memiliki kelas yang sama.
  - Tidak ada atribut yang tersisa pada record yang dapat dipartisi lebih lanjut. Dalam kasus ini suara mayoritas digunakan. Node tersebut menjadi leaf node dan dilabeli dengan kelas yang menjadi mayoritas dalam record yang ada.
  - Tidak ada record untuk cabang variabel tes. Dalam kasus ini, leaf terbentuk dengan mayoritas kelas sebagai label record tersebut.

Ukuran *information gain* digunakan untuk memilih atribut uji pada setiap *node* di dalam *tree*. Ukuran inidigunakan untuk memilih atribut atau *node* pada pohon. Atribut dengan nilai *information gain* tertinggi akan terpilih sebagai parent bagi *node* selanjutnya. Sebelum menghitung *gain* harus dihitung terlebih dahulu nilai *entropy*-nya. *Entropy* adalah suatu parameter untuk mengukur *heterogenitas* (keberagaman) dari suatu kumpulan data sampel. Apabila sampel data semakin *heterogen* maka nilai dari *entropy*-nya semakin besar. Formula dari *entropy* adalah:

$$Entropy (S) = \sum_{i=0}^n - \pi * \log_2 \pi i$$

Rumus diatas adalah rumus yang digunakan dalam perhitungan *entropy* yang digunakan untuk menentukan seberapa informatif atribut tersebut. Berikut keterangannya:

S = Himpunan kasus

N = Jumlah partial S

Pi = Jumlah kasus pada partisi ke-i

Langkah selanjutnya setelah nilai *entropy* diperoleh maka adalah menghitung gain untuk mengukur efektifitas suatu atribut dalam mengklasifikasi data. Gain dihitung dengan menggunakan rumus:

$$Gain(S, A) = Entropy(S) - \sum_{v \in \text{values}} (A) \frac{Sv}{S} * Entropy(Sv)$$

Rumus diatas adalah rumus yang digunakan dalam perhitungan *gain* setelah melakukan perhitungan *entropy*. Berikut keterangannya:

S = Himpunan kasus

n = Jumlah partisi atribut A

|Si| = Jumlah kasus pada partisi ke-i

|S| = Jumlah kasus dalam S.

## 2.4 Prediksi

Pengertian Prediksi adalah sama dengan ramalan atau perkiraan. Menurut Kamus Besar Bahasa Indonesia, prediksi adalah hasil dari kegiatan memprediksi atau meramal atau memperkirakan. Prediksi bisa berdasarkan metode ilmiah ataupun subjektif belaka. Kesimpulannya pengertian prediksi secara istilah akan sangat tergantung pada konteks atau permasalahannya. Berbeda dengan pengertian prediksi secara bahasa yang berarti ramalan atau perkiraan yang sudah menjadi pengertian yang baku[7].

## 3. HASIL DAN PEMBAHASAN

### 3.1 Analisa Data

Dalam menyelesaikan masalah tersebut diperlukan data dapat dilakukan analisa prediksi terhadap data persediaan ikan ekspor sebelumnya. Dengan mengetahui jenis ikan yang di ekspor apa saja yang sering di beli oleh konsumen. Dengan tujuan untuk menentukan pola pembelian konsumen sehingga dapat menentukan pola persediaan ikan ekspor pada perusahaan. Berikut ini adalah data persediaan ikan ekspor pada perusahaan yang akan dijadikan sampel yaitu Size Ikan, penjualan *ikan*, pembelian ikan untuk analisa dan pengujian tahun 2019.

Table 1. Data Ekspor

No	Kawasan Eskpor	Jenis Ikan	Size Ikan	Data Eskpor	Bahan Baku
1	Eropa	Nila Danau Toba	500 gr	Sedang	Ada
2	Asia	Nila Danau Toba	300 gr	Sedang	Tidak Ada
3	Asia	Nila Danau Toba	500 gr	Tinggi	Tidak Ada
4	Asia	Patin	700 gr	Tinggi	Ada
5	Eropa	Patin	300 gr	Sedang	Tidak Ada
6	Eropa	Patin	500 gr	Sedang	Ada
7	Eropa	Ikan Mas	300 gr	Sedang	Tidak Ada
8	Eropa	Nila Danau Toba	500 gr	Sedang	Tidak Ada
9	Asia	Patin	300 gr	Tinggi	Tidak Ada
10	Asia	Nila Danau Toba	500 gr	Sedang	Ada
11	Eropa	Ikan Mas	700 gr	Sedang	Tidak Ada
12	Eropa	Nila Danau Toba	300 gr	Tinggi	Tidak Ada
13	Asia	Patin	700 gr	Sedang	Ada

Berdasarkan tabel diatas, maka dapat dijelaskan bahwa data yang digunakan merupakan data hasil olahan perbulannya, yaitu persediaan ikan pada perusahaan karena masalah keterbatasan data, waktu, kemampuan serta data yang bersifat pribadi. Sehingga pada proses ini pembersihan data tidak dilakukan karena data dianggap sudah siap untuk dipergunakan serta data siap diolah menggunakan Algoritma C5.0.

#### 3.1.1 Algoritma

Algoritma C5.0 merupakan kelompok algoritma decision tree. Algoritma ini mempunyai input berupa training samples dan samples. Training samples berupa data contoh yang akan digunakan untuk membangun sebuah tree yang telah diuji kebenarannya. Patinkan samples merupakan field-field data yang nantinya akan kita gunakan sebagai parameter dalam melakukan klasifikasi data Algoritma C5.0 adalah pengembangan dari algoritma C4.5. Oleh karena pengembangan tersebut algoritma C5.0 mempunyai prinsip dasar kerja yang sama dengan algoritma C4.5.

Untuk penyelesaian kasus pada algoritma C5.0 ada beberapa elemen yang harus dicari nilainya yaitu:

- a. Siapkan Dataset
- b. Hitung Nilai Entropy (S)

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2(p_i)$$

Keterangan:

S = Jumlah sampel data (*Sampling*)

n = Jumlah partisi S

p<sub>i</sub> = Proporsi dari S<sub>i</sub> terhadap S.

c. Hitung Nilai Gain (S, A)

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

Keterangan:

S = Jumlah kasus (*Sampling*)

A = *attribute*

n = Jumlah partisi S

|S<sub>i</sub>| = Jumlah kasus pada partisi ke-i

|S| = Jumlah kasus dalam S

d. Hitung nilai SplitInfo

$$SplitInfo(D) = - \sum_{j=1}^n \frac{|D_j|}{|D|} * \log_2\left(\frac{|D_j|}{|D|}\right)$$

D = Jumlah kasus (*Sampling*)

n = Jumlah partisi D

|D<sub>j</sub>| = Jumlah kasus pada partisi ke-i

|D| = Jumlah kasus dalam D

e. Hitung nilai *Gain Ration*

$$GainRation = \frac{Gain(S, A)}{SplitInfo(D)}$$

f. Membentuk Node dan Cabang dari nilai Gain terIkan Mas

g. Ulangi Langkah 2 sampai langkah 5 hingga semua node terisi

Untuk dapat menentukan nilai-nilai Gain dan Entropy dari masing – masing atribut di atas pada table 4.1, maka terlebih dikonversikan data ke dalam bentuk tabel klasifikasi yang lebih detail sebagai berikut:

**Tabel 2.** Klasifikasi

Node	Arribut	Kategori	Jumlah Kasus(S)	Ada	Tidak Ada	Entropy	Gain	
1	Total		13	9	5	0,897		
		Kawasan	Eropa	7	5	2	0,863	0,894
	Asia		6	3	3	1		
	Jenis Ikan	Ikan Mas		2	2	0	0	1,694
			Patin	5	3	2	0,971	
		Nila Danau Toba		6	4	2	0,918	
	Size Ikan	500 gr		5	2	3	0,971	0,735
		300 gr		5	5	0	0	
		700 gr		3	1	2	0,918	
	Data Ekspor	Sedang		9	5	4	0,991	0,460
		Tinggi		4	3	1	0,811	

Setelah tabel 2 dibuat, maka langkah selanjutnya adalah melakukan perhitungan dengan mencari nilai Entropy menggunakan persamaan. Setelah menentukan nilai Entropy langkah selanjutnya yaitu mencari nilai Gain:

a. Nilai *Entropy(S)*

$$Total = \left(-\frac{9}{13} * \log_2 \frac{9}{13}\right) + \left(-\frac{5}{13} * \log_2 \frac{5}{13}\right) = 0.897$$

1. Kawasan

$$Eropa = \left(-\frac{5}{7} * \log_2 \frac{5}{7}\right) + \left(-\frac{2}{7} * \log_2 \frac{2}{7}\right) = 0,863$$

$$Asia = \left(-\frac{3}{6} * \log_2 \frac{3}{6}\right) + \left(-\frac{3}{6} * \log_2 \frac{3}{6}\right) = 1$$

2. Jenis Ikan

$$Ikan Mas = \left(-\frac{2}{2} * \log_2 \frac{2}{2}\right) + \left(-\frac{0}{2} * \log_2 \frac{0}{2}\right) = 0$$

$$Patin = \left(-\frac{3}{5} * \log_2 \frac{3}{5}\right) + \left(-\frac{2}{5} * \log_2 \frac{2}{5}\right) = 0,971$$

$$Nila Danau Toba = \left(-\frac{4}{6} * \log_2 \frac{4}{6}\right) + \left(-\frac{2}{6} * \log_2 \frac{2}{6}\right) = 0,918$$

3. Size Ikan

$$300 gr = \left(-\frac{2}{5} * \log_2 \frac{2}{5}\right) + \left(-\frac{3}{5} * \log_2 \frac{3}{5}\right) = 0,971$$

$$500 gr = \left(-\frac{5}{5} * \log_2 \frac{5}{5}\right) + \left(-\frac{0}{5} * \log_2 \frac{0}{5}\right) = 0$$

$$700 gr = \left(-\frac{1}{3} * \log_2 \frac{1}{3}\right) + \left(-\frac{2}{3} * \log_2 \frac{2}{3}\right) = 0,918$$

4. Data Ekspor

$$Sedang = \left(-\frac{5}{9} * \log_2 \frac{5}{9}\right) + \left(-\frac{4}{9} * \log_2 \frac{4}{9}\right) = 0,991$$

$$Tinggi = \left(-\frac{3}{4} * \log_2 \frac{3}{4}\right) + \left(-\frac{1}{4} * \log_2 \frac{1}{4}\right) = 0,811$$

b. Gain (S, A)

1. Kawasan

$$= 0,897 - \left(\left(\frac{7}{13} * 0,863\right) + \left(\frac{6}{13} * 1\right)\right) = 0,894$$

2. Jenis Ikan

$$= 0,897 - \left(\left(\frac{2}{13} * 0\right) + \left(\frac{5}{13} * 0,971\right) + \left(\frac{6}{13} * 0,918\right)\right) = 1,694$$

3. Size Ikan

$$= 0,897 - \left(\left(\frac{5}{13} * 0,971\right) + \left(\frac{5}{13} * 0\right) + \left(\frac{3}{13} * 0,918\right)\right) = 0,735$$

4. Data Ekspor

$$= 0,897 - \left(\left(\frac{9}{13} * 0,991\right) + \left(\frac{4}{13} * 0,811\right)\right) = 0,460$$

a. SplitInfo(D)

1. Kawasan

$$= \left(-\frac{7}{13} * \log_2 \frac{7}{13}\right) + \left(-\frac{6}{13} * \log_2 \frac{6}{13}\right) = 0,690$$

2. Jenis Ikan

$$= \left(-\frac{2}{13} * \log_2 \frac{2}{13}\right) + \left(-\frac{5}{13} * \log_2 \frac{5}{13}\right) + \left(-\frac{6}{13} * \log_2 \frac{6}{13}\right) = 1,012$$

3. Size Ikan

$$= \left(-\frac{5}{13} * \log_2 \frac{5}{13}\right) + \left(-\frac{5}{13} * \log_2 \frac{5}{13}\right) + \left(-\frac{3}{13} * \log_2 \frac{3}{13}\right) = 1,073$$

4. Data Ekspor

$$= \left(-\frac{9}{13} * \log_2 \frac{9}{13}\right) + \left(-\frac{4}{13} * \log_2 \frac{4}{13}\right) = 9,617$$

b. Gain Ration

1. Kawasan

$$= \frac{0,894}{0,690} = 1,294$$

2. Jenis Ikan

$$= \frac{1,694}{1,012} = 1,673$$

3. Size Ikan

$$= \frac{0,735}{1,0733} = 0,685$$

4. Data Ekspor

$$= \frac{0,460}{0,617} = 0,685$$

Setelah mendapatkan nilai *Gain Ration* pada proses diatas, maka proses selanjutnya pembentukan akar pohon keputusan berdasarkan *Gain Ration* Ikan Mas. Perhitungan *Gain* diatas diperoleh *Gain Ration* Ikan Mas yaitu atribut jenis ikan , sehingga yang menjadi node akar adalah Jenis Ikan.

Pada node akar awal atribut Jenis Ikan memiliki kategori atribut Ikan Mas, Patin dan Nila Danau Toba. Cabang atribut Ikan Mas memiliki keputusan karena nilai *Entropy*=0, Patinkan cabang atribut Patin dan Nila Danau Toba belum memiliki keputusan karena nilai *Entropy* masih belum=0. Oleh karena itu, terbentuk node baru yaitu node 1.1 dan Node

1.2 yang akan dilakukan pencarian nilai *Entropy* dan *Gain* untuk cabang Patin node 1.1 dan cabang Nila Danau Toba node 1.2 sebagai berikut:

a. Proses pembentukan node 1.1

**Tabel 3.** Dataset Ekspor Dengan Atribut Jenis Ikan Patin

No	Kawasan	Jenis Ikan	Size Ikan	Data Ekspor	Bahan Baku
1	Asia	Patin	700 gr	Tinggi	Ada
2	Eropa	Patin	300 gr	Sedang	Tidak Ada
3	Eropa	Patin	500 gr	Sedang	Ada
4	Asia	Patin	300 gr	Tinggi	Tidak Ada
5	Asia	Patin	700 gr	Sedang	Ada

Data pada tabel diatas yaitu kelompok data yang belum memiliki keputusan, oleh karena itu dihitung nilai *Entropy* dan *Gain* untuk pencarian cabang dari node 1.1. adalah sebagai berikut:

**Table 4.** Node 1.1

Node	Atribut	Kategori	Jumlah Kasus(S)	Tidak Ada	Ada	Entropy	Gain
.1	Jenis Ikan Kawasan	Patin	5	3	2	0,971	
		Eropa	2	1	1	1	0.020
		Asia	3	1	2	0,918	
	Size Ikan	500 gr	1	0	1	0	0,971
		300 gr	2	2	0	0	
		700 gr	2	0	2	0	
	Data Ekspor	Sedang	3	1	2	0,918	0.020
		Tinggi	2	1	1	1	

1. Nilai *Entropy(S)*

Jenis Ikan+ Patin

$$Total = \left(-\frac{3}{5} * \log_2 \frac{3}{5}\right) + \left(-\frac{2}{5} * \log_2 \frac{2}{5}\right) = 0,971$$

a. Kawasan

$$Eropa = \left(-\frac{1}{2} * \log_2 \frac{1}{2}\right) + \left(-\frac{1}{2} * \log_2 \frac{1}{2}\right) = 1$$

$$Asia = \left(-\frac{1}{3} * \log_2 \frac{1}{3}\right) + \left(-\frac{2}{3} * \log_2 \frac{2}{3}\right) = 0,918$$

b. Size Ikan

$$500 \text{ gr} = \left(-\frac{0}{1} * \log_2 \frac{0}{1}\right) + \left(-\frac{1}{1} * \log_2 \frac{1}{1}\right) = 0$$

$$300 \text{ gr} = \left(-\frac{2}{2} * \log_2 \frac{2}{2}\right) + \left(-\frac{0}{2} * \log_2 \frac{0}{2}\right) = 0$$

$$700 \text{ gr} = \left(-\frac{0}{2} * \log_2 \frac{0}{2}\right) + \left(-\frac{2}{2} * \log_2 \frac{2}{2}\right) = 0$$

c. Data Ekspor

$$Sedang = \left(-\frac{1}{3} * \log_2 \frac{1}{3}\right) + \left(-\frac{2}{3} * \log_2 \frac{2}{3}\right) = 0,918$$

$$Tinggi = \left(-\frac{1}{2} * \log_2 \frac{1}{2}\right) + \left(-\frac{1}{2} * \log_2 \frac{1}{2}\right) = 1$$

2. *Gain (S, A)*

a. Kawasan

$$= 0,971 - \left(\left(\frac{2}{5} * 1\right) + \left(\frac{3}{5} * 0,918\right)\right) = 0,002$$

b. Size Ikan

$$= 0,971 - \left(\left(\frac{1}{5} * 0\right) + \left(\frac{2}{5} * 0\right) + \left(\frac{2}{5} * 0\right)\right) = 0,971$$

c. Data Ekspor

$$= 0,971 - \left(\left(\frac{3}{5} * 0,918\right) + \left(\frac{2}{5} * 1\right)\right) = 0,002$$

3. *SplitInfo(D)*

a. Kawasan

$$= \left(-\frac{2}{5} * \log_2 \frac{2}{5}\right) + \left(-\frac{3}{5} * \log_2 \frac{3}{5}\right) = 0,673$$

- b. Size Ikan  
$$= \left(-\frac{1}{5} * \log_2 \frac{1}{5}\right) + \left(-\frac{2}{5} * \log_2 \frac{2}{5}\right) + \left(-\frac{2}{5} * \log_2 \frac{2}{5}\right) = 1,055$$
- c. Data Ekspor  
$$= \left(-\frac{3}{5} * \log_2 \frac{3}{5}\right) + \left(-\frac{2}{5} * \log_2 \frac{2}{5}\right) = 9,673$$
4. *Gain Ration*
- a. Kawasan  
$$= \frac{0,020}{0,673} = 0$$
- b. Size Ikan  
$$= \frac{0,971}{1,055} = 0,920$$
- c. Data Ekspor  
$$= \frac{0,020}{0,673} = 0$$

Setelah mendapatkan nilai *Gain Ration* pada proses diatas, maka proses selanjutnya pembentukan akar pohon keputusan berdasarkan *Gain Ration* Ikan Mas. Perhitungan *Gain* diatas diperoleh *Gain Ration* Ikan Mas yaitu atribut jenis ikan , sehingga yang menjadi node akar adalah Jenis Ikan.

Cabang pohon keputusan Kawasan 2(dua) nilai *attribute* yaitu Eropa dan Asia dan nilai *attribute* tersebut akan menjadi cabang keputusan cabang node Kawasan. Cabang node 1.2 Kawasan Cabang yang masing-masing sudah memiliki keputusan karena nilai *Entropy* cabang tersebut=0, dimana cabang Eropa keputusannya Tidak Ada dan Asia keputusannya Ada. Karena cabang node 1.1 Kawasan sudah memiliki keputusan maka proses berhenti, tidak ada lagi pencarian nilai *Entropy* dan *Gain*. Sehingga diperoleh pohon keputusan yang telah dibentuk, dibaca dalam bentuk *rule* / aturan. Cara membaca pohon keputusan dimulai dari node akar ke node cabang, menggunakan kata "DAN" untuk penghubung antara node dan dibentuk dalam *Jika...maka...atau Jika...dan...maka....* Bentuk *rule* pohon keputusan diatas adalah sebagai berikut:

- Jika Jenis Ikan Mas maka Ada Persediaan Ekspor Ikan
- Jika Jenis Ikan Patin dan Size Ikan 500 gr maka Tidak Ada Persediaan Ekspor Ikan
- Jika Jenis Ikan Patin dan Size Ikan 300 gr maka Ada Persediaan Ekspor Ikan
- Jika Jenis Ikan Patin dan Size Ikan 700 gr maka Ada Persediaan Ekspor Ikan
- Jika Jenis Ikan Nila Danau Toba dan Kawasan Eropa maka Ada Persediaan Ekspor Ikan
- Jika Jenis Ikan Nila Danau Toba dan Kawasan Asia maka Tidak Ada Persediaan Ekspor Ikan

#### 4. KESIMPULAN

Adapun kesimpulan pada hasil penelitian ini yaitu Proses ekspor kebutuhan ikan yang digunakan pada PT. Mutiara Anugerah Samudra dengan cara melakuak persediaan barang baku ikan yang akan di ekspor pada Kawasan Asia dan Eropa. Penerapan algoritma C5.0 digunakan untuk proses ekspor ikan pada PT. Mutiara Anugerah Samudera yang berupa rule (aturan) yang dapat menjadi informasi penting tentang pola memprediksi kebutuhan ekspor ikan. Proses pengujian prediksi kebutuhan ekspor ikan dengan menerap kan algortma C5.0 menggunakan aplikasi rapiminer yang menghasilkan pola informasi dalam melakuakn kegiatan ekspor ikan pada PT. Mutiara Anugerah Samudera.

#### REFERENCES

- [1] Kusri and Emha T.L, *Pengertian Prediksi*, Yogyakarta : Andi.2009
- [2] Kusri and Emha T.L, *Algoritma dan Data Mining*, Yogyakarta : Andi.2009
- [3] Hermawati, Fajar.Astuti, *Data Mining*, Yogyakarta: Andi. 2013
- [4] Budi Santosa, *Data Mining: Teknik Pemanfaatan Data Untuk KeperluanBisnis*, 1st ed. Yogyakarta: Graha Ilmu, 2007.
- [5] Yusuf W,Yogi.2007.Perbandingan Performansi Algoritma Decision Tree C5.0, Cart, dan Chaid:Kasus Prediksi Status Resiko Kredit Di Bank "X",Jurnal Seminar nasional Aplikasi Teknologi Informasi,1907-5022
- [6] Finn Lee and Juan Santana, *Data Mining: Meramalkan Bisnis Perusahaan*, 1st ed. Jakarta: Elex Media Komputindo, 2010.
- [7] Susanto and Dedy Suryadi, *Pengantar Data Mining: Menggali Pengetahuan dari Bungkahan Data*. Yogyakarta: Andi, 2010.
- [8] Widodo,P.P, Handayanto,R.T, Herlawati, *Penerapan Data Mining dengan MATLAB*, Bandung: Rekayasa Sains.2013
- [9] Beta Noranita and Nurdin Bahtiar, "Implementasi Data Mining Untuk Menentukan Pola Hubungan Tingkat Kelulusan Mahasiswa dengan Induk Mahasiswa," pp. 156-163, Oktober 2010.
- [10] Dana Sulistyio Kusumo, "Data Mining dengan Algoritma Apriori pada RDBMS Oracle," *Penelitian dan Pengembangan Telekomunikasi*, vol. 8, 2003
- [11] Data Mining : Klasifikasi Menggunakan Algoritma C4.5," Jurnal Edik Informatika, vol. 2, no. 2, pp. 213-219, 2017.