

# IMPLEMENTASI METODE *ROUGH SET* UNTUK MENGANALISA LABA/RUGI PADA SUATU PERUSAHAAN DISTRIBUTOR (STUDI KASUS : USAHA KITA PS PAYAKUMBUH)

M.Ardiansyah Sembiring<sup>\*1)</sup>, Zulfi Azhar<sup>2)</sup>

Program Studi Sistem Informasi, STMIK Royal Kisaran

Jl. Prof. M. Yamin 173 Kisaran, Sumatera Utara 21222

Telp: (0623) 41079

<sup>\*1)</sup>E-mail : adinmantap88@gmail.com<sup>1)</sup>

---

## Abstrak

Penelitian ini dilakukan untuk menemukan *New Knowledge* (pengetahuan baru) berupa *rule*(aturan) melalui implementasi *Data Mining* metode *Rough Set* dalam menganalisa hasil perhitungan laba/rugi pada suatu periode akuntansi di dalam sebuah perusahaan. Penelitian ini akan dapat membantu pihak perusahaan di dalam menganalisa hasil kinerja pengelola perusahaan secara objektif dan utuh. Hasil dari penelitian ini berupa *rule*(aturan) yang akan menentukan kondisi pencapaian target *pf*ofit yang dapat dikategorikan tercapai atau tidak tercapai. Proses penemuan *rule*(aturan) pada metode *Rough Set* dimulai dari pembentukan *Decision System* yang merupakan data awal menjadi klasifikasi periode akuntansi yang memiliki nilai atribut kondisi dan keputusan yang sama disebut *Equivalence Class*. Setelah didapat *Equivalence Class* dilakukan proses *Discernibility Matrix* dan *Discernibility Matrix Modulo D*. Selanjutnya melakukan proses *Reduction* yang akan digunakan sebagai acuan pembuatan *General Rules*. Hasil dari *General Rules* kemudian yang akan menjadi pengetahuan baru dalam penelitian ini. Penulis menyadari bahwa masih terdapat kekurangan pada penelitian ini. Namun, metode *Rough Set* merupakan metode yang tepat untuk digunakan dalam menganalisa hasil kinerja suatu perusahaan.

**Kata kunci :** *Data Mining, Decision System, General Rules, Rough Set dan Reduction.*

## 1. Pendahuluan

Perusahaan distributor merupakan perusahaan dagang yang berorientasi pada keuntungan. Tingginya tingkat persaingan di dalam dunia usaha tersebut, menjadikan setiap perusahaan harus meningkatkan kinerja dan terus mengevaluasi diri secara tepat agar dapat bertahan bahkan memenangkan persaingan. Proses evaluasi yang dilakukan seharusnya dilakukan secara objektif. Salah satu cara objektif yang dapat menggambarkan kondisi perusahaan dan kemudian dilakukan evaluasi tersebut yaitu melalui perhitungan laba/rugi. Hasil perhitungan laba/rugi akan dianalisa berdasarkan variabel yang menyusunnya dalam satu periode. Satu periode biasanya dihitung pada akhir setiap bulan pada tahun masehi. Terdapat beberapa variabel dalam perhitungan di atas yaitu penjualan, pembelian, stok gudang, dan beban usaha. Banyaknya variabel yang mempengaruhi hasil perhitungan laba/rugi setiap periodenya, maka dibutuhkan analisa yang mendalam agar mendapatkan informasi yang benar dan utuh yang akan menjadi acuan dalam pengambilan keputusan.

Penelitian ini dilakukan penulis pada sebuah perusahaan distributor swasta yaitu Usaha Kita PS Kota Payakumbuh yang bergerak pada penjualan

pakan ternak yang berasal dari PT. Sabas Medan kepada konsumen. Pada perusahaan ini, penulis menemukan masalah dalam menganalisa laporan laba/rugi yaitu variabel yang dijadikan bahan evaluasi tidak utuh karena hanya fokus kepada omset penjualan. Perusahaan tersebut menganggap keuntungan perusahaan akan besar hanya jika omset penjualan besar. Hal ini menyebabkan evaluasi terhadap variabel lain terabaikan. Akibatnya, penyebab - penyebab tercapai atau tidak tercapainya target profit tidak diketahui secara utuh. Hal tersebut di atas dapat menyebabkan keuntungan menjadi tidak maksimal bahkan dapat menyebabkan kerugian bagi perusahaan.

Metode yang penulis gunakan untuk menyelesaikan masalah di atas adalah metode *Rough Set*. Melalui metode *Rough Set* dengan *tools* yang digunakan di dalam pengolahan data adalah *Rosetta versi 1.4.41*, penulis akan membandingkan setiap variabel yang menyusun laporan laba/rugi yang berasal dari data keuangan perusahaan.

Pada penerapannya Metode *Rough Set* pernah dilakukan oleh Rodríguez (2013) dengan tujuan utama dari penelitian ini adalah untuk mempelajari

keadaan seni teknik saat ini dan algoritma untuk meningkatkan efisiensi terisolasi sistem surya fotovoltaik dengan menerapkan rough set untuk identifikasi variabel yang signifikan dalam produksi energi photovoltaic dengan sistem terisolasi.

Berdasarkan hal-hal yang telah dijelaskan di atas maka tujuan yang ingin dicapai pada penelitian ini adalah menemukan *rule* atau aturan dari variabel-variabel penyusun laba/rugi sesuai dengan proses *data mining* metode *rough set* melalui pengujian menggunakan *Tools Rosetta versi 1.4.41* untuk mendapatkan pengetahuan baru.

Batasan masalah pada penelitian ini adalah :

1. Mengimplementasikan metode *rough set* untuk menganalisa perhitungan laba/rugi perusahaan distributor di Usaha Kita PS Payakumbuh.
2. Menemukan *rule* atau aturan dari variabel-variabel penyusun laba/rugi sesuai dengan proses *data mining* metode *rough set*.
3. Menguji *decision system* dari proses ekstraksi menggunakan *Tools Rosetta versi 1.4.41*.

## 2. TINJAUAN PUSTAKA

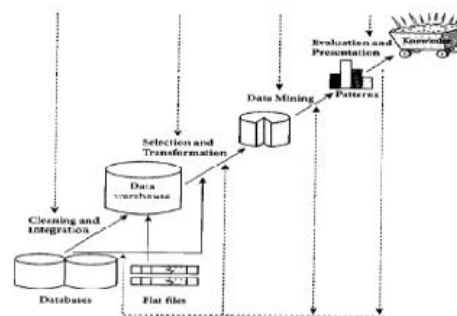
Menurut Pernyataan Standar Akuntansi Keuangan No. 25 (IAI, 2007), laporan laba rugi merupakan laporan utama untuk melaporkan kinerja dari suatu perusahaan selama suatu periode tertentu. Informasi tentang kinerja perusahaan, terutama tentang profitabilitas, dibutuhkan untuk mengambil keputusan tentang sumber ekonomi yang akan dikelola oleh suatu perusahaan di masa depan (Meythi, 2012).

*Data mining* sebenarnya merupakan salah satu bagian proses *Knowledge Discovery in Database* (KDD) yang bertugas untuk mengekstrak pola atau model dari data dengan menggunakan suatu algoritma yang spesifik. Adapun proses KDD sebagai berikut :

1. *Data Selection* : pemilihan data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalan informasi dalam KDD dimulai.
2. *Preprocessing* : sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* dengan tujuan untuk membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

3. *Transformation* : yaitu proses *coding* pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam database.
4. *Data mining* : proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu.
5. *Interpretation / Evaluation* : pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya atau tidak.

*Data mining* juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan (Prasetyo, 2012). Tahapan-tahapan pada data mining dapat digambarkan seperti gambar 1 berikut :

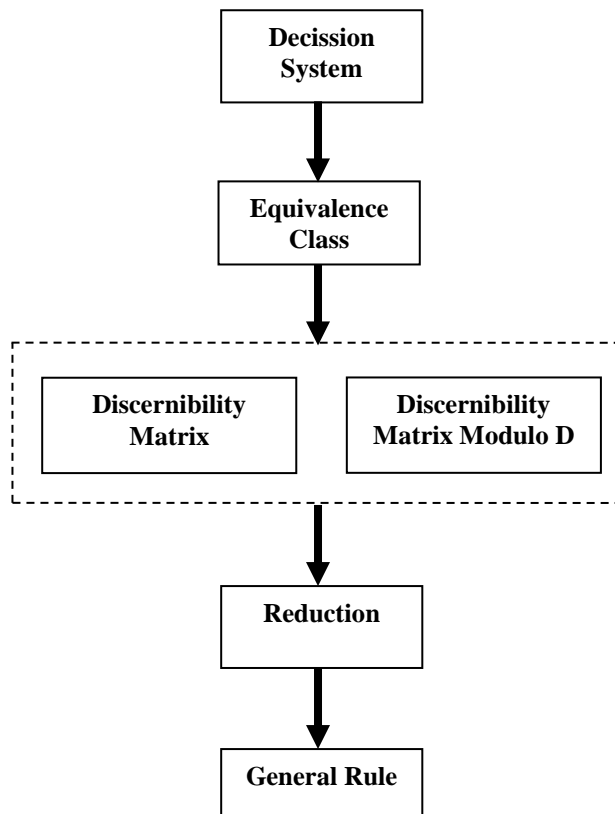


Gambar 1. Tahapan Data Mining

Metode *Rough Set* dikembangkan oleh Zdzislaw Pawlak sebagai alat matematis untuk menangani ketidakjelasan dan ketidakpastian. Telah berhasil diterapkan dalam berbagai tugas, seperti fitur seleksi/ekstraksi, sintesis aturan dan klasifikasi, penemuan pengetahuan, dan lain-lain. Di dalam Metode *Rough Set* terdapat beberapa langkah - langkah penyelesaian masalah, yaitu sebagai berikut:

1. *Decision System* tersebut dilakukan teknik klasifikasi kriteria yang disebut “*Equivalen Class*”
2. Kemudian dilakukan proses *Discernibility Matrix* atau *Discernibility Matrix Modulo D*
3. Proses “*Reduction*”
4. Untuk memperoleh hasil akhir dilakukan proses “*General Rules*”

Langkah-langkah dalam menjalankan metode *rough set* di atas dapat digambarkan pada gambar berikut ini.



Gambar 2. Proses Metode Rough Set

a. Informasi System

Dalam *rough set*, sebuah *set* data direpresentasikan sebagai sebuah tabel, dimana baris dalam tabel merepresentasikan objek dan kolom-kolom merepresentasikan *attribute* dari objek-objek tersebut. Tabel tersebut disebut dengan *information system* yang dapat digambarkan sebagai berikut :

$$IS = \{U, A\}$$

Dimana  $U$  adalah set terhingga yang tidak kosong dari objek yang disebut dengan *universe* dan  $A$  set terhingga tidak kosong dari atribut dimana:

$$a: U \longrightarrow V_a$$

untuk tiap  $a \in A$ . Set  $V_a$  disebut *value* set dari  $a$ .  $U = \{e_1, e_2, \dots, e_m\}$  merupakan sekumpulan *example* dan  $A = \{a_1, a_2, \dots, a_n\}$  yang merupakan *attribute* kondisi secara berurutan.

b. Decision System

Dalam penggunaan *information system*, terdapat *outcome* dari klasifikasi yang telah diketahui yang disebut dengan atribut keputusan. *Information system* tersebut disebut dengan *Decision system*. *Decision system* dapat digambarkan sebagai: (3)

$$IS = (U, \{A, C\})$$

Dimana :

$U = \{x_1, x_2, \dots, x_m\}$  yang merupakan sekumpulan *example*.

$A = \{a_1, a_2, \dots, a_n\}$  yang merupakan sekumpulan *attribute* kondisi secara berurutan atau *attribute*.

$C = \text{Decision attributes}$  (keputusan).

c. Equivalence Class

*Equivalence class* adalah mengelompokan objek-objek yang sama untuk *attribute*  $A \in (U, A)$ .

d. Discernibility Matrix

Definisi *Discernibility Matrix*: Diberikan sebuah  $IS A = (U, A)$  dan  $B \subseteq A$ , *discernibility matrix* dari  $A$  adalah  $MB$ , dimana tiap-tiap entry  $MB(I, j)$  terdiri dari sekumpulan *attribute* yang berbeda antara objek  $X_i$  dan  $X_j$ .

e. Discernibility Matrix Modulo D

Didefinisikan seperti berikut dimana  $MB(I, j)$  adalah sekumpulan *attribute* yang berbeda antara objek  $X_i$  dan  $X_j$  dan juga berbeda *attribute* keputusan. Diberikan sebuah  $DS A = (U, A \setminus \{d\})$  dan subset dari *attribute*  $B \subseteq A$ , *Discernibility Matrix Modulo D* dari  $A$ ,  $MB_d$ .

f. Reduct

*Reduct* adalah penyeleksian *attribute* minimal (*interesting attribute*) dari sekumpulan *attribute* kondisi dengan menggunakan *Prime Implicant* fungsi *Boolean*. Kumpulan dari semua *Prime Implicant* mendeterminasikan *sets of reduct*.

g. Gerated Rules

Proses selanjutnya yaitu mendapatkan pengetahuan dalam *database* melalui ekstraksi aturan dari sistem keputusan. Hasil keputusan tersebut didasarkan pada proses *reduct*. (2)

### 3. ANALISIS SISTEM

Decision system yang akan diproses pada sistem adalah sebagai berikut :

Tabel 1. Sampel Decision System

PENJUALAN	PEMBELIAN	STOK GUDANG	BEBAN USAHA	PROFIT
DIBAWAH	KURANG	SEDIKIT	TIDAK EFISIEN	TIDAK TERCAPAI
DIBAWAH	BANYAK	PENUH	TIDAK EFISIEN	TIDAK TERCAPAI
DIBAWAH	KURANG	SEDIKIT	TIDAK EFISIEN	TIDAK TERCAPAI
TARGET	BANYAK	PENUH	TIDAK EFISIEN	TIDAK TERCAPAI
DIBAWAH	BANYAK	PENUH	TIDAK EFISIEN	TIDAK TERCAPAI
DIATAS	BANYAK	PENUH	EFISIEN	TERCAPAI
TARGET	BANYAK	PENUH	EFISIEN	TERCAPAI
.....	.....	.....	.....	.....
TARGET	BANYAK	PENUH	EFISIEN	TERCAPAI

Tabel 1. Memperlihatkan sebuah Decision Systems yang pada penelitian ini. Ia hanyaterdiri dari  $n$  objek,  $E1, E2, E3, E4, \dots, E48$  dan attribute kondisi yaitu penjualan, pembelian, stok gudang, serta beban usaha. Sedangkan profit adalah decision attribute.

Decision System dari Rosetta yang telah diimpor masuk ke dalam system melalui prosedur input data ditampilkan pada gambar berikut.

Gambar 3. Decision System Rosetta

Tahapan selanjutnya di dalam tools Rosetta akan melakukan proses Equivalence Class dan Discernibility Matrix Modulo  $D$ , namun tidak secara langsung ditampilkan hasil berupa tabel Equivalence Class dan Discernibility Matrix Modulo  $D$  Melainkan langsung menampilkan proses selanjutnya didalam metode Rough Set yaitu hasil reduct. Hasil reduct pada penelitian ini ditunjukkan pada gambar 4.

	Reduct	Support	Length
1	{BEBAN USAHA}	60	1
2	{PENJUALAN}	60	1
3	{PENJUALAN, STOK GUDANG}	59	2
4	{PENJUALAN, BEBAN USAHA}	58	2
5	{PEMBELIAN, BEBAN USAHA}	60	2
6	{STOK GUDANG, BEBAN USAHA}	60	2
7	{PENJUALAN, PEMBELIAN}	59	2
8	{PEMBELIAN, STOK GUDANG}	57	2

Gambar 4. Hasil Reduct Rosetta

Gambar 4 menunjukkan hasil reduct dari proses yang dilakukan dengan hasil berupa attribute minimal sebanyak 8 (delapan) pasangan. Nantinya attribute-attribute ini akan menjadi acuan dalam melakukan General Rule. Hasil general rule pada penelitian ini ditampilkan melalui proses Rosetta pada gambar 5.

	Rule
1	BEBAN USAHA(EFISIEN) => PROFIT(TERCAPAI)
2	PENJUALAN(DIATAS) => PROFIT(TERCAPAI)
3	PENJUALAN(DIBAWAH) => PROFIT(TIDAK TERCAPAI) OR PROFIT(TERCAPAI)
4	PENJUALAN(DIBAWAH) AND STOK GUDANG(SEDIKIT) => PROFIT(TIDAK TERCAPAI)
5	PENJUALAN(TARGET) AND STOK GUDANG(SEDIKIT) => PROFIT(TERCAPAI)
6	PENJUALAN(DIBAWAH) AND STOK GUDANG(PENUH) => PROFIT(TIDAK TERCAPAI) OR PROFIT(TERCAPAI)
7	PENJUALAN(DIBAWAH) AND BEBAN USAHA(TIDAK EFISIEN) => PROFIT(TIDAK TERCAPAI)
8	PENJUALAN(TARGET) AND BEBAN USAHA(TIDAK EFISIEN) => PROFIT(TIDAK TERCAPAI) OR PROFIT(TERCAPAI)
9	PEMBELIAN(BANYAK) AND BEBAN USAHA(TIDAK EFISIEN) => PROFIT(TIDAK TERCAPAI)
10	STOK GUDANG(PENUH) AND BEBAN USAHA(TIDAK EFISIEN) => PROFIT(TIDAK TERCAPAI)
11	PENJUALAN(TARGET) AND PEMBELIAN(KURANG) => PROFIT(TERCAPAI)
12	PENJUALAN(DIBAWAH) AND PEMBELIAN(BANYAK) => PROFIT(TIDAK TERCAPAI)
13	PEMBELIAN(KURANG) AND STOK GUDANG(PENUH) => PROFIT(TERCAPAI)

Gambar 5. Hasil General Rule Rosetta

Setelah kita menyelesaikan keseluruhan tahapan dari proses penginputan Decision System, Equivalence Class, Discernibility Matrix Modulo  $D$ , Reduction sampai terbentuknya Knowledge baru, maka kita dapat melihat output yang dihasilkan dari Generated Rules seperti pada gambar.5 Adapun rule yang dihasilkan adalah sebagai berikut :

1. *If* Beban Usaha = Efisien *Then* Profit = Tercapai
2. *If* Penjualan = Diatas *Then* Profit = Tercapai
3. *If* Penjualan = Dibawah *Then* Profit = Tidak Tercapai *or* Profit = Tercapai
4. *If* Penjualan = Dibawah *and* Stok Gudang = Sedikit *Then* Profit = Tidak Tercapai
5. *If* Penjualan = Target *and* Stok Gudang = Sedikit *Then* Profit = Tercapai
6. *If* Penjualan = Dibawah *and* stok Gudang Penuh *Then* Profit = Tercapai *or* Profit = Tidak Tercapai
7. *If* Penjualan = Dibawah *and* Beban Usaha = Tidak Efisien *Then* Profit = Tidak Tercapai
8. *If* Penjualan = Target *and* Beban Usaha = Tidak Efisien *Then* Profit = Tidak Tercapai *or* Profit = Tercapai
9. *If* Pembelian = Banyak *and* Beban Usaha = Tidak Efisien *Then* Profit = Tidak Tercapai
10. *If* Stok Gudang = Penuh *and* Beban Usaha = Tidak Efisien *Then* Profit = Tidak Tercapai
11. *If* Penjualan = Target *and* Pembelian = Kurang *Then* Profit = Tercapai
12. *If* Penjualan = Dibawah *and* Pembelian = Banyak *Then* Profit = Tidak tercapai
13. *If* Pembelian = Kurang *and* Stok Gudang = Penuh *Then* Profit = Tercapai

Berdasarkan hasil *generated rule* menunjukkan 13 (tiga belas) *rule* atau pengetahuan baru yang tersusun dari *attribute-attribute* penyusunnya. Berdasarkan analisa yang peneliti lakukan dari *rule* yang didapatkan, jumlah kemunculan *attribute* Penjualan sebanyak 9(sembilan) kali, *attribute* Beban Usaha sebanyak 5(lima) kali, *attribute* Pembelian sebanyak 4(empat) kali, dan *attribute* Stok Gudang sebanyak 5(lima) kali. Sehingga dapat diketahui bahwa *attribute* yang paling berpengaruh dalam ketercapaian profit dari hasil perhitungan laba/rugi yaitu *attribute* penjualan karena memiliki jumlah kemunculan terbanyak. *Attribute* yang berpengaruh berikutnya setelah Penjualan yaitu Beban Usaha dan Stok Gudang yang memiliki jumlah kemunculan yang sama dengan kemunculan terbanyak kedua. Kemudian *attribute* berpengaruh yang terakhir yaitu Pembelian yaitu karena memiliki jumlah kemunculan terkecil.

#### 4. EVALUASI SISTEM

Pada awal penelitian penulis telah menjelaskan bahwasannya yang menjadi permasalahan selama ini yaitu dalam menganalisa laba/rugi secara objektif. Selama ini perusahaan yang menjadi studi kasus penelitian hanya menggunakan 1(satu) variabel yaitu penjualan untuk menganalisa laba/rugi dalam menentukan ketercapaian target profit. Berikut ini adalah

deskripsi tentang perbedaan sistem yang lama dengan sistem yang baru.

##### a. Sistem Yang Lama

Di dalam sistem yang lama, penentuan variabel yang menentukan hasil profit tercapai atau tidak tercapai hanya variabel penjualan saja. Hal ini tentu saja mengabaikan variabel-variabel lain yang juga memberikan pengaruh kepada profit. Kemudian dampak yang akan terjadi akan mengakibatkan evaluasi kinerja perusahaan Usaha Kita PS menjadi tidak utuh. Akibat yang akan timbul dari ketidak utuhan tersebut akan menjadi celah keuntungan perusahaan menjadi tidak maksimal dan bahkan dapat menyebabkan kerugian.

##### b. Sistem Yang Baru

Beberapa manfaat yang dirasakan sejak adanya sistem pengolahan *data mining* menggunakan *Rough Set* sebagai metode pemecahan masalah dan *Rosetta 1.4.41* sebagai *tools* untuk menguji data dan menghasilkan parameter- parameter yang menentukan ketercapain target profit serta *knowledge* baru yang dihasilkannya yaitu sebagai berikut:

##### 1. Pencarian Informasi

Pada proses pencarian informasi menggunakan *tools Rosetta 1.4.41* lebih mudah. Hal ini dikarenakan kita dibantu pada seluruh proses di dalam metode *rough set* dari *decission system* data laba/rugi sesuai dengan *attribute* yang dimiliki pada setiap periode akuntansi. Kemudian kita juga akan menemukan hasil yang lebih akurat jika dibandingkan secara manual karena terhindar dari *human error*.

##### 2. Knowledge Yang Dihasilkan

Dengan memanfaatkan *tools Rosetta 1.4.41*, *Knowledge* yang dihasilkan dapat terlihat dengan jelas, baik parameter penentu yang disebut dengan *Reduct* dan dapat *General Rules* dari hasil penelitian yang lebih terperinci.

#### 5. KESIMPULAN

Berdasarkan hasil penulisan pada Bab I sampai dengan Bab V yang telah dilakukan, maka dapat diambil kesimpulan yang dapat berguna bagi para pembaca sehingga penelitian ini dapat lebih bermanfaat. Adapun kesimpulan dari penulisan penelitian ini adalah sebagai berikut :

1. Implementasi metode *rough set* mampu menjawab permasalahan dalam menganalisa hasil perhitungan laba/rugi dari penemuan

- rule* (aturan) tentang pencapaian target profit pada suatu periode akuntansi.
2. Berdasarkan *rule* yang ditemukan dapat diketahui bahwa *attribute* kondisi yang paling berpengaruh dalam ketercapaian profit berdasarkan hasil perhitungan laba/rugi yaitu Penjualan. Hal ini dikarenakan *attribute* Penjualan merupakan *attribute* yang paling sering muncul dari seluruh *attribute* yang terdapat pada *rule* yang dihasilkan. *Attribute* kondisi yang berpengaruh berikutnya setelah Penjualan yaitu Beban Usaha. Setelah Beban Usaha *Attribute* kondisi yang berpengaruh berikutnya yaitu Pembelian dan Stok Gudang.
  3. Pengujian *decision system* dengan menggunakan *Tools Rosetta versi 1.4.41* dapat mempermudah perusahaan dalam mendapatkan pengetahuan baru sebagai acuan pengambilan keputusan.
- Memprediksi Kriteria Nasabah Kredit". *Jurnal Komputa*. Volume 1. 53 – 57.
- Rodriguez, A.P.R dan Riverola, F.F. (2013). "Applying Rough Set for the Identification of Significant Variables in Photovoltaic Energy Production with Isolated System". *Jurnal Teknologi*. ISSN:0127-9696. Hlm. 9 – 16.
- Ryan S.J.D. Bakers, (2008). "Data Mining for Education". USA: Carniege Mellon University Pittsburgh.
- Subekti Mujiasih, (2011). "Pemanfaatan Data Mining untuk Prakiraan Cuaca". Volume 12. 189 – 195.
- Thakare, V.M dan Deshpande, S.P. (2010). "Data Mining System And Applications: A Review". *IJDPS*. Volume 1. 32 - 44.
- Yellarisi Ramadevi dan Rao, C.R. (2007). "Decission Tree Induction Using Rough Set Theory-Comparative Study". *JATIT*. Hlm. 110 – 114.

## DAFTAR PUSTAKA

- Dicky Nofriyansah (2013). "Implementasi Metode *Rough Set* Untuk Menganalisa Kelulusan Mahasiswa Pada Suatu Mata Kuliah (Studi Kasus: STMIK Budi Dharma Medan)". UPI YPTK Padang: Tesis.
- Eko Prasetyo, 2012. "Data Mining Konsep dan Aplikasi Menggunakan Matlab". Ed.I. Yogyakarta : Andi. Hlm.3–7.
- Fajar Astuti Hermawati. (2009). "Data Mining". Ed.I. Yogyakarta : Andi. Hlm.2-3.
- Gogoi Prasanta, (2011). "Efficient Rule Set Generation Using Rough Set Theory for Classification of High Dimensional Data". *IJSSAN*. ISSN: 2248–9738 Volume 1. Hlm.13 – 60.
- Goldie Gunadi dan Dana Indra Sensuse, (2012). "Penerapan Metode *Data Mining Market Basket Analysis* Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori dan *Frequent Pattern Growth (FP-Growth)* Studi Kasus Percetakan PT. Gramedia". *Jurnal Telematika*. ISSN:2085-725X. Volume 4. Hlm. 118 – 132.
- Meythi dan Selvy Hartono, (2012). "Pengaruh Informasi Laba dan Arus Kas Terhadap Harga Saham". *Jurnal Akuntansi*. ISSN:2086-4159. Hlm. 57– 59.
- Obadi Gamila, (2010). "A Tolerance Rough Set Based Overlapping Clustering for the DBLP Data". *Computer Society IEE*. Hlm.57 – 59.
- Riani Lubis dan Angga Ginanjar Maburur, (2012). "Penerapan Data Mining untuk