



## Computer Based Information System Journal

ISSN (Print): 2337-8794 | E- ISSN : 2621-5292  
 web jurnal : <http://ejournal.upbatam.ac.id/index.php/cbis>



# PENERAPAN DATA MINING DALAM MEMPREDIKSI PEMBELIAN SEMEN

**Rika Harman**

Universitas Putera Batam, Indonesia.

### INFORMASI ARTIKEL

Diterima Redaksi: Februari 2019  
 Diterbitkan Online: Maret 2019

### KATA KUNCI

Pembelian, Semen, Pilihan, Bahan Bangunan

### KORESPONDENSI

E-mail: [rika.harman@gmail.com](mailto:rika.harman@gmail.com)

### A B S T R A C T

*When faced with a decision to determine one choice it is sometimes a very difficult thing, this can occur due to lack of knowledge and supporting data that exists. Supporting data is usually not too difficult to find because it can be taken from our own environment or on our own, because without realizing it around us it is basically a very large and large collection of data but sometimes it is not utilized because it cannot use it. For example, this can be seen in the Batam Saddhana Legenda Building Shop, in principle, the building store is the same as the other building shops. But the problem arising from shop owners was when determining the purchase of cement as the main ingredient in building construction. If you prioritize closeness or good relations with sponsors and sponsors' contributions to the store, the shop owner will not be difficult to determine which supplier to contact, but because you have to learn the tendency of the buyer which brand they prefer, the problem will be complicated. Because the shop owner must order products that are tailored to consumer demand, if this is not done then what will happen to the existing customers will be disappointed because the product they want does not exist, even the shop owner must put stock for a long time that can harm the shop owner own. By applying the data mining classification technique to effective cement purchases at the Saddhana Shop, it is hoped that later it can produce a knowledge that can be used for decision making in effective cement purchases from suppliers.*

### I. Latar Belakang

Pada kegiatan hariannya sebuah toko yang mempunyai konsumen tentu banyak menggunakan, menyimpan bahkan membuang data yang pada dasarnya tercipta begitu saja tanpa disadari, padahal data tersebut jika dikumpulkan merupakan data yang sangat penting baik bagi toko itu sendiri maupun bagi orang lain disekitarnya, setuju dan tidak setuju data yang ada tersebut tidak semuanya harus dibuang begitu saja, karena dari beberapa data tersebut masih bisa dimanfaatkan untuk kepentingan toko tersebut, misalkan dalam

kesehariannya toko tersebut pasti akan melakukan transaksi, dari transaksi tersebut tentu bisa dipelajari apa saja yang dibeli konsumen, misalkan satu produk yang sama dengan merek yang berbeda atau satu produk yang sama dengan harga yang berbeda atau lagi produk yang sama dengan kualitas yang berbeda mana konsumen yang lebih menyukainya, hal ini jika dibuat pemetaannya akan sangat membantu pemilik toko dalam pengambilan keputusan karena pemilik toko akan mengetahui keinginan konsumen tentang produk yang tersedia ditokonya. Jika hal ini dilakukan

untuk setiap merek produk yang ada di toko jelas data yang cukup besar sedang dibuang dan tidak dimanfaatkan dengan baik, tapi untuk mengumpulkan data tersebut juga merupakan hal yang tidak mudah karena jumlah data yang sangat besar.

Dalam penelitian yang akan dilakukan ini jelas tidak semua produk pada toko akan diuji, diolah dan di ekstrak menggunakan datamining tapi cukup untuk produk yang tingkat kebutuhan setiap harinya sangat tinggi saja yaitu semen. Sebagai bahan dasar bangunan semen merupakan produk yang paling banyak dicari oleh konsumen dan paling banyak juga di sediakan oleh toko bangunan. Walaupun semen merupakan produk yang paling banyak dicari semen tidak bisa di stok dalam waktu yang cukup lama karena sangat beresiko, seperti bisa menjadi keras atau pada saat digunakan tidak baik lagi. Hal ini yang harus dipandang dengan seksama oleh pemilik toko, pemilik toko harus menyediakan produk semen tapi tidak bisa lama, ditambah lagi semen yang tersedia tidak hanya satu produk atau satu merek saja, yang umum beredar ditoko bangunan di Batam ada 4 seperti Semen Padang, Tiga Roda, Holcim dan Bosowa. Salah dalam dalam memutuskan dalam proses pembelian maka akan menjadi pembelian yang tidak efektif dan akan dapat menimbulkan kekecewaan dari konsumen karena produk yang tersedia tidak atau kurang diminati.

Dari hasil studi lapangan yang penulis lakukan pada hari kamis tanggal 28 Juni 2018 maka didapat dari pemilik toko bangunan tentang jumlah semen yang dipesan, terjual dan yang tersedia sekarang yaitu adalah sebagai berikut:

**Tabel 1.1.** Data Penjualan Semen Juni 2018

No	Semen	Pesanan	Jual	Persediaan
1	Padang	160 Sak	145 Sak	15 Sak
2	Tiga Roda	100 Sak	80 Sak	20 Sak
3	Holcim	80 Sak	70 Sak	10 Sak
4	Bosowa	60 Sak	53 Sak	7 Sak

**Sumber:** Toko Sadana Legenda Batam (2018)

Dari tabel diatas dapat dilihat bahwa untuk periode bulan Juni mulai dari tanggal satu Juni sampai dengan tanggal 28 Juni sudah menghabiskan hampir sekitar 348 Sak semen sedangkan yang tersedia sekarang adalah 52 Sak semen.

Pada dasarnya dalam melakukan proses pembelian semen ke supplier bukanlah hal sulit karena cukup menghubungi supplier mana yang paling banyak memberikan kontribusi, kemudahan serta keuntungan kepada manajemen toko, hal ini tentu bisa dilihat berdasarkan hubungan keseharian antara manajemen toko dengan supplier, tapi yang sulit adalah menentukan semen yang sesuai dengan keinginan konsumenlah yang sulit karena salah dalam membeli konsumen akan kecewa dan stok akan menumpuk serta kerugian akan datang kepada pemilik toko. Hal inilah yang mendorong proses pengolahan data dengan teknik kalsifikasi menggunakan data mining, sehingga dari empat produk semen tersebut (Semen Padang, Tiga Roda, Holcim dan Bosowa) dapat dibeli sesuai dengan kebutuhan konsumen sehingga menjadi pembelian yang efektif. Dari empat produk semen ini nantinya akan dipesan ke supplier berdasarkan tingkat kebutuhan konsumen bukanlah berdasarkan kebutuhan dan hubungan baik antara manajemen toko dengan supplier. Dari pembelian yang efektif tersebut diharapkan tidak adanya stok yang terlalu lama dan kerusakan semen di toko dapat dicegah sedini mungkin dan konsumen juga terpuaskan.

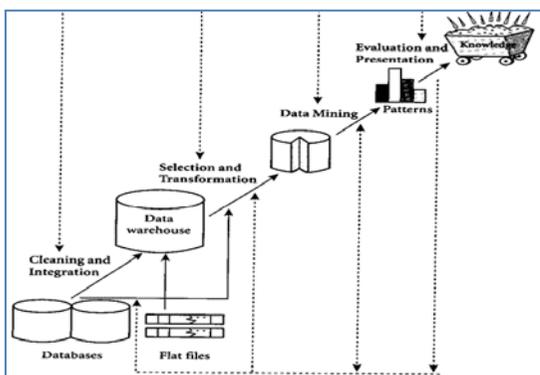
## II. Kajian Literatur

*Knowledge Discovery in Database* (KDD) merupakan proses terorganisir untuk mengidentifikasi pola dalam data yang besar dan kompleks di mana pola data tersebut ditemukan yang bersifat sah, baru, dan dapat bermanfaat serta dapat dimengerti. *Data Mining*, juga dikenal sebagai *Knowledge Discovery in Database* yang mengacu penggalan atau penambahan pengetahuan dari data dalam jumlah besar. Jadi dari kedua pendapat di atas dapat disimpulkan bahwa KDD adalah proses pencarian informasi atau pengetahuan dari kumpulan data yang besar (*database*). Menurut Feelders, Daniels, dan Holsheimer dalam Gunadi (2008) *Data Mining* adalah proses ekstraksi informasi dari kumpulan data melalui penggunaan algoritma dan teknik yang melibatkan bidang teknik statistik, matematika, mesin pembelajaran, dan sistem manajemen *database*. *Data Mining* dapat digunakan untuk mengekstrak informasi dan pengetahuan penting yang tersembunyi dari *dataset* yang besar.

Dengan adanya *Data Mining*, maka akan dimungkinkan untuk mendapatkan harta berharga berupa pengetahuan didalam kumpulan data. Hampir senada dengan Gunadi, Benni (2014) mendefinisikan *Data Mining* sebagai kegiatan menemukan pola yang menarik dari data yang berjumlah besar, data dapat disimpan dalam *database*, *data warehouse*, atau penyimpanan informasi lainnya. Pola yang ditemukan merupakan pola yang menarik, berharga dan penting dalam memberikan keuntungan kepada pemilik data. Terdapat yang mirip diungkapkan juga oleh Geetha dan Manimekalai (2013) bahwa *Data mining* adalah proses yang menggunakan berbagai alat analisis data untuk menemukan pola dan hubungan dalam data yang dapat digunakan untuk membuat prediksi yang *valid*.

### A. Tahapan *Data Mining*

Karena *Data Mining* merupakan rangkaian proses, maka *Data Mining* dapat dibagi menjadi beberapa tahap. Tahap-tahap tersebut bersifat *interaktif* di mana pemakai terlibat langsung atau dengan perantara *knowledge base*. Tahap-tahap ini diilustrasikan pada Gambar 2.1.



Gambar 1. Tahap-Tahap *Data Mining*

### B. Pengelompokan *Data Mining*

Pengelompokan *Data Mining* dibagi menjadi beberapa kelompok, menurut Kusriani dan Luthfi (2009) dalam Jumanto (2014), pengelompokan tersebut didasarkan kepada 6 (enam) faktor, diantaranya adalah sebagai berikut:

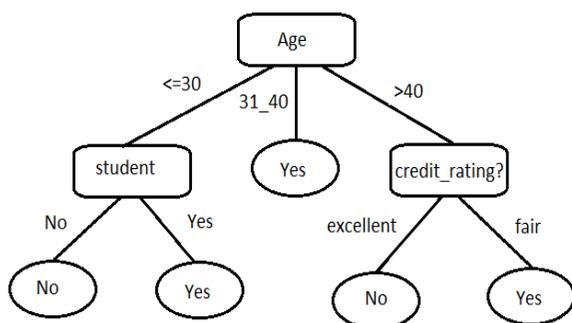
1. Deskripsi adalah cara untuk menggambarkan pola.

2. Estimasi, model yang dibangun menggunakan *record* lengkap yang menyediakan nilai *variable* target sebagai nilai prediksi.
3. Prediksi, menerka sebuah nilai yang belum diketahui termasuk masa mendatang. Klasifikasi, terdapat target *variable* yang didefinisikan dalam situasi tinggi, sedang, dan rendah.
4. Pengklasteran, pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan.
5. Asosiasi, menemukan atribut yang muncul dalam satu waktu.

### C. Pohon Keputusan (*Decision Tree*)

Metode ini merupakan salah satu metode yang ada pada teknik klasifikasi dalam *Data Mining*. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang mempresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami. Dan mereka juga dapat diekspresikan dalam bentuk bahasa basis data seperti *Structured Query Language* untuk mencari *record* pada kategori tertentu Utari (2015). Pohon keputusan menggunakan representasi struktur pohon (*tree*) di mana setiap *node* merepresentasikan atribut, cabangnya merepresentasikan nilai dari atribut dan daun merepresentasikan kelas. *Node* yang paling atas dari pohon keputusan disebut sebagai *root*. Pohon keputusan merupakan metode klasifikasi yang paling populer digunakan. Selain karena pembangunannya relatif cepat, hasil dari model yang dibangun mudah untuk dipahami. Data dalam pohon keputusan biasanya dinyatakan dalam bentuk tabel dengan atribut dan *record*. Atribut menyatakan suatu parameter yang disebut sebagai kriteria dalam pembentukan pohon. Manfaat utama dari penggunaan pohon keputusan adalah kemampuannya untuk *break down* proses pengambilan keputusan yang kompleks menjadi lebih simpel sehingga pengambil keputusan akan lebih menginterpretasikan solusi dari permasalahan.

Contoh dari pohon keputusan dapat dilihat di Gambar dibawah ini:



**Gambar 2.** Pohon Keputusan Identifikasi

Pada pohon keputusan terdapat 3 jenis *node*, yaitu:

1. *Root Node*, merupakan *node* paling atas, pada *node* ini tidak ada *input* dan bisa tidak mempunyai *output* atau mempunyai *output* lebih dari satu.
2. *Internal Node*, merupakan *node* percabangan, pada *node* ini hanya terdapat satu *input* dan mempunyai *output* minimal dua.
3. *Leaf node atau terminal node*, merupakan *node* akhir, pada *node* ini hanya terdapat satu *input* dan tidak mempunyai *output* Andriani (2013).

#### D. Algoritma C.45 dan Pembentukan Pohon Keputusan

Algoritma C.45 yaitu sebuah algoritma yang digunakan untuk membangun *decision tree* (pengambilan keputusan). Algoritma C.45 adalah

1. Simpul induknya dan kesalahan klasifikasi bernilai 0.
2. Apabila  $T_i$  memiliki cabang lagi maka pemilahan diproses kembali menggunakan kasus-kasus dalam  $T_i$ .
3. Kesalahan klasifikasi simpul dihitung dari penjumlahan dari kesalahan-kesalahan simpul anak yang dibandingkan dengan simpul induknya.

salah satu algoritma induksi pohon keputusan yaitu ID3 (*Iterative Dichotomiser 3*). ID3 dikembangkan oleh J. Ross Quinlan. Dalam prosedur algoritma ID3, input berupa sampel *training*, label *training* dan atribut. Algoritma C.45 merupakan pengembangan dari ID3. Beberapa pengembangan yang dilakukan pada C.45 adalah sebagai antara lain bisa mengatasi *missing value*, bisa mengatasi *continu data*, dan *pruning* Fadillah (2013). Berikut ini langkah-langkah konstruksi pohon keputusan menggunakan Algoritma C.45:

1. Misalkan  $T$  adalah himpunan kasus-kasus yang akan dibuat simpul di mana kasus-kasus tersebut memiliki kelas dan atribut-atribut. Frekuensi terboboti  $freq(C_j, T)$  diperoleh dari perhitungan  $T$  dan kelas yang dihasilkan adalah  $C_j$ , untuk setiap  $j \in \{1, 2, 3, \dots, n\}$
2. Jika semua kasus berada dalam kelas  $C_j$  yang sama maka simpul yang dihasilkan adalah simpul daun yang diberi label dengan kelas  $C_j$  sebagai kelas terbanyak.
3. Jika  $T$  berisi kasus yang memiliki dua atau lebih kelas maka dapat dihitung  $j$  *information gain* dari setiap atribut tersebut.
4. Atribut dengan nilai *information gain* tertinggi terpilih sebagai pemilah dalam simpul tersebut.
5. Simpul keputusan memiliki cabang sebanyak  $s$  yaitu  $T_1, \dots, T_s$  di mana  $s = 2$  untuk atribut kontinu dan  $s = h$  untuk atribut diskret dengan nilai  $h$  yang sudah diketahui.
6. Untuk setiap  $i = \{1, 2, \dots, s\}$ , jika  $T_i$  tidak memiliki cabang lagi maka simpul tersebut secara langsung menjadi simpul daun yang diberi label kelas terbanyak di bawah

#### E. Entropy dan Information Gain

Sebuah obyek yang diklasifikasikan dalam pohon harus dipes nilai *Entropy* -nya. *Entropy* adalah ukuran dari teori informasi yang dapat mengetahui karakteristik dari *impurity* dan *homogeneity* dari kumpulan data. Dari nilai *Entropy* tersebut kemudian dihitung nilai *information gain* (IG) masing-masing atribut. *Entropy* ( $S$ ) merupakan jumlah bit yang diperkirakan dibutuhkan untuk dapat mengekstrak suatu kelas (+ atau -) dari sejumlah

data acak pada ruang sampel  $S$ . *Entropy* dapat dikatakan sebagai kebutuhan bit untuk menyatakan suatu kelas. Semakin kecil nilai *Entropy* maka akan semakin *Entropy* digunakan dalam mengekstrak suatu kelas. *Entropy* digunakan untuk mengukur ketidakkaslian  $S$ .

1. Untuk perhitungan Gain adalah sebagai berikut:

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

2. Untuk menghitung nilai Entropy adalah sebagai berikut:

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

### III. Metodologi

Metodologi penelitian diperlukan sebagai kerangka dan panduan dalam melakukan proses penelitian, sehingga penelitian yang dilakukan menjadi lebih terarah, teratur, dan sistematis. Metodologi penelitian adalah suatu cabang ilmu pengetahuan yang menjelaskan mengenai cara-cara melaksanakan penelitian yang dimulai dari kegiatan mencari, mencatat, merumuskan, menganalisis, hingga menyusun laporan berdasarkan fakta-fakta atau gejala-gejala secara ilmiah. Motivasi dan tujuan penelitian secara umum pada dasarnya adalah sama, yaitu bahwa penelitian merupakan refleksi dari keinginan manusia yang selalu berusaha untuk mengetahui sesuatu. Keinginan untuk memperoleh dan mengembangkan pengetahuan merupakan kebutuhan dasar manusia yang umumnya menjadi motivasi untuk melakukan penelitian. Metodologi penelitian ini dilakukan secara sistematis yang dapat digunakan sebagai pedoman untuk peneliti dalam melaksanakan penelitian agar hasil yang dicapai tidak menyimpang dan tujuan yang diinginkan dapat terlaksana dengan baik dan sesuai dengan tujuan yang telah ditetapkan sebelumnya.

### IV. Pembahasan

Berdasarkan data yang diperoleh dari toko Shadana baik yang berdasarkan pengamatan atau pembelajaran terhadap data yang sudah ada ataupun data sebelumnya ditambah dengan hasil wawancara singkat dengan beberapa konsumen

maka dapat ditentukan format data dalam memprediksi pembelian semen pada toko Shadana yang antara lain adalah sebagai berikut:

1. Merek (terkenal dan tidak terkenal)
2. Jenis atau berat (berat dan sedang)
3. Harga (mahal, sedang dan murah).
4. Rekomendasi (ya dan tidak).
5. Animo (tinggi sedang dan rendah)

Adapun yang menjadi variabel tujuan (GOAL) dalam hal ini variabel yang akan menjadi keputusan adalah laris (Y) dan tidak laris suatu produk (T)

#### A. Pohon Keputusan

Menentukan atribut sebagai akar dan menghitung nilai informasi *gain* atribut. Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Dibutuhkan nilai *Entropy* untuk menentukan *gain* tertinggi.

Entropy total Atribut

$$\text{Entropy}(\text{total}) = \left(-\frac{10}{16} * \log_2 \left(\frac{10}{16}\right)\right) + \left(-\frac{6}{16} * \log_2 \left(\frac{6}{16}\right)\right) = 0,954434$$

Entropy (total) adalah menghitung nilai total keputusan Ya (9) dan Tidak (8), sedangkan 17 adalah jumlah keseluruhan kasus.

1. Atribut Merek

$$- \text{Entropy} \text{TKNL} = \left(-\frac{1}{6} * \log_2 \left(\frac{1}{6}\right)\right) + \left(-\frac{5}{6} * \log_2 \left(\frac{5}{6}\right)\right) = 0.650022$$

$$- \text{Entropy} \text{SDNG} = \left(-\frac{5}{7} * \log_2 \left(\frac{5}{7}\right)\right) + \left(-\frac{2}{7} * \log_2 \left(\frac{2}{7}\right)\right) = 0.863121$$

$$- \text{Entropy} \text{TTKL} = \left(-\frac{3}{4} * \log_2 \left(\frac{3}{4}\right)\right) + \left(-\frac{1}{4} * \log_2 \left(\frac{1}{4}\right)\right) = 0.8112780$$

Sedangkan untuk nilai Gain Total Merek adalah sebagai berikut, seperti dibawah ini:

$$= 0.99750 - \left(\left(\frac{6}{17} * 0.65002\right) + \left(\frac{7}{17} * 0.863121\right) + \left(\frac{4}{17} * 0.8112780\right)\right) = 0.084674$$

2. Atribut Jenis

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

$$- \text{Entropy} \text{BRT} = \left(-\frac{6}{11} * \log_2 \left(\frac{6}{11}\right)\right) + \left(-\frac{5}{11} * \log_2 \left(\frac{5}{11}\right)\right) = 0.99403$$

Sedangkan untuk nilai Gain Total Jenis adalah sebagai berikut, seperti dibawah ini:

$$= 0.99750 - \left( \left( \frac{11}{17} * 0.99403 \right) + \left( \frac{6}{17} * 1 \right) \right) = 0.0136551$$

3. Atribut Harga

$$- Entropy MHL = \left( -\frac{2}{7} * \log_2 \left( \frac{2}{7} \right) \right) + \left( -\frac{5}{7} * \log_2 \left( \frac{5}{7} \right) \right) = 0.863121$$

$$- Entropy SDNG = \left( -\frac{6}{6} * \log_2 \left( \frac{6}{6} \right) \right) + \left( -\frac{0}{6} * \log_2 \left( \frac{0}{6} \right) \right) = 0$$

$$- Entrop MRH = \left( -\frac{3}{4} * \log_2 \left( \frac{3}{4} \right) \right) + \left( -\frac{1}{4} * \log_2 \left( \frac{1}{4} \right) \right) = 0.8112780$$

Sedangkan untuk nilai Gain Total Harga adalah sebagai berikut, seperti dibawah ini:

$$= 0.99750 - \left( \left( \frac{7}{17} * 0.863121 \right) + \left( \frac{6}{17} * 0 \right) + \left( \frac{4}{17} * 0.8112780 \right) \right) = 0.451211$$

4. Atribut Rekomendasi

$$- Entropy Ya = \left( -\frac{5}{8} * \log_2 \left( \frac{5}{8} \right) \right) + \left( -\frac{3}{8} * \log_2 \left( \frac{3}{8} \right) \right) = 0.954434$$

$$- Entropy Tdk = \left( -\frac{4}{9} * \log_2 \left( \frac{4}{9} \right) \right) + \left( -\frac{5}{9} * \log_2 \left( \frac{5}{9} \right) \right) = 0.991076$$

Sedangkan untuk nilai Gain Total Rekomendasi adalah sebagai berikut, seperti dibawah ini:

$$= 0.99750 - \left( \left( \frac{8}{17} * 0.954434 \right) + \left( \frac{9}{17} * 0.991076 \right) \right) = 0.02366$$

$$- Entropy TGG = \left( -\frac{4}{9} * \log_2 \left( \frac{4}{9} \right) \right) + \left( -\frac{5}{9} * \log_2 \left( \frac{5}{9} \right) \right) = 0.991076$$

5. Atribut Animo

$$- Entropy TGG = \left( -\frac{4}{9} * \log_2 \left( \frac{4}{9} \right) \right) + \left( -\frac{5}{9} * \log_2 \left( \frac{5}{9} \right) \right) = 0.991076$$

$$- Entrop RNDH = \left( -\frac{2}{4} * \log_2 \left( \frac{2}{4} \right) \right) + \left( -\frac{2}{4} * \log_2 \left( \frac{2}{4} \right) \right) = 1$$

$$- Entropy SDNG = \left( -\frac{3}{4} * \log_2 \left( \frac{3}{4} \right) \right) + \left( -\frac{1}{4} * \log_2 \left( \frac{1}{4} \right) \right) = 0.811278$$

Sedangkan untuk nilai Gain Total Animo adalah sebagai berikut, seperti dibawah ini:

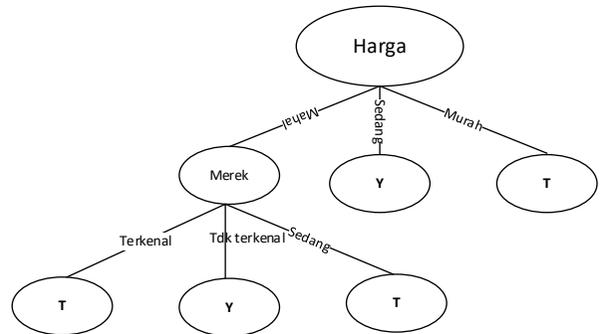
$$= 0.99750 - \left( \left( \frac{9}{17} * 0.991076 \right) + \left( \frac{4}{17} * 0.811278 \right) + \left( \frac{4}{17} * 1 \right) \right) = 0.046632$$

Lakukan hal yang sama secara berulang sampai ditemukan nilai gain tertinggi yang akan dijadikan sebagai node akar.

Keterangan:

Dari perhitungan pada tabel 4.9 dapat diketahui bahwa atribut dengan gain tertinggi adalah Merek dengan atributnya yakni 0.86312. Merek dapat menjadi node akar, dimana nilai atribut TKNL (terkenal) dikatakan Tidak Layak dengan

jumlah tidak layak adalah 3, dan Merek SDG (sedang) tetap dikatakan tidak layak dengan jumlah tidak layak adalah 2, dan yang terakhir adalah merek dengan atribut TTKL (tidak terkenal) dapat dikatakan layak dengan nilai kelayakannya adalah 2. Sehingga Pohon keputusan yang terbentuk seperti pada gambar dibawah ini::



Gambar 3. Pohon Keputusan Hasil

Keterangan:

Penarikan kesimpulan dapat dilakukan setelah kedua proses pengujian dilakukan baik itu pengujian manual menggunakan proses perhitungan matematika. Dimana lebih lanjut aturan atau rule yang terbentuk berdasarkan pohon keputusan terakhir seperti pada gambar 3 diatas adalah sebagai berikut:

1. IF (harga = murah) maka (semen = tidak akan dibeli).
2. IF (harga = sedang) maka (semen = dibeli).
3. IF (harga = mahal) dan (merek = tidak terkenal) maka semen akan di beli.
4. IF (harga = mahal) dan (merek=terkenal) serta (harga=mahal) dan (merek = sedang) maka semen tidak akan dibeli.

Berdasarkan rule atau knowledge yang dihasilkan terdapat rules yang sesuai dengan kejadian yang terjadi dalam memprediksi pembelian semen pada toko Shadana Legenda Malaka Batam. Sebagaimana contoh pembelian semen akan dilakukan jika harga semen tersebut adalah harga sedang dan mahal jika harga murah maka sipembeli tidak akan melakukannya karena menyangkut masalah kualitas, dan rata-rata pelanggan yang membeli semen adalah untuk rumahnya sendiri, artinya walaupun harganya murah sipembeli belum tentu tertarik untuk

membelinya karena faktor kualitas dan keperluan spesifik. Sedangkan untuk harga semen yang mahal maka sipembeli akan melihat kembali merek dari semen tersebut jika mereknya terkenal dan sedang sipembeli tidak akan melakukan pembelian karena sipembeli lebih mengutamakan membeli semen dengan merek tidak terkenal. Hal ini bisa terjadi karena pada saat pengambilan data stok semen yang tersedia adalah semen dengan harga mahal dan merek tidak terkenal sedangkan untuk merek sangat terkenal tidak ada di stok toko tersebut atau habis, sebaliknya untuk semen dengan kategori harga mahal merek sedang memang tidak dicari oleh pelanggan walaupun ada itupun jumlah sedikit.

## V. Kesimpulan

Secara garis besar penerapan dari Data Mining khususnya teknik klasifikasi C.45 didalam penelitian ini dalam hal memprediksi pembelian semen adalah sukses (berhasil) hal ini dapat dibuktikan dengan terbentuknya beberapa jalur keputusan yang dapat diambil yang antara lain adalah sebagai berikut:

1. Semen dengan harga murah tidak akan dibeli oleh konsumen karena semen dengan harga ini jarang diminati karena dianggap dari segi kualitas kurang baik, ditambah lagi rata-rata sipembeli membeli untuk rumahnya sendiri.
2. Semen yang akan dibeli oleh konsumen berdasarkan survei yang dilakukan adalah semen dengan harga sedang dan mahal, hal ini terjadi karena anggapan sebagian konsumen untuk masalah kualitas semen yang baik hanya ada pada semen dengan harga sedang dan mahal.
3. Untuk kategori harga yang mahal dari semen yang akan dibeli tersebut adalah khusus merek yang tidak terkenal. Hal juga dipengaruhi oleh stok yang ada pada saat survei dilakukan.
4. Semen dengan harga mahal dan merek terkenal tidak akan dibeli oleh sipembeli hal ini dipengaruhi oleh stok yang ada pada saat survei dilakukan yang dalam hal ini tidak lebih dominan, artinya stok dengan merek terkenal tidak tersedia.
5. Semen dengan harga mahal dan merek sedang tingkat terkenalnya juga tidak akan dibeli oleh konsumen, karena kebanyakan konsumen akan melakukan pembelian harga

mahal merek pasti pilih yang terkenal. Hal ini juga tergantung dari keadaan stok di Gudang pada saat proses pengambilan data dilakukan.

## Ucapan Terima Kasih

Pada penulisan jurnal ini peneliti mengucapkan terimakasih yang sebesar-besarnya kepada pemilik toko Sadana beserta staf dan karyawan dalam hal ini adalah bapak Johan Arista, yang telah membantu dalam menyiapkan dan memberikan data selama proses penelitian berlangsung, seterusnya kepada yang terhormat seluruh pelanggan toko Sadana yang telah berkontribusi sehingga terkumpulnya data yang diperlukan selama proses pengambilan data berlangsung. Terakhir ucapan kepada pihak LPPM Universitas Putera Batam yang telah mendukung baik secara finansial maupun nonfinansial. Terakhir mahasiswa yang telah membantu dalam proses penelitian ini mulai dari proses pembuatan proposal dan pengolahan data penelitian.

## Daftar Pustaka

- [1] Benni R Suburian. (2014). Aplikasi Data Mining Untuk Menampilkan Tingkat Kelulusan Mahasiswa Dengan Algoritma Apriori. *Teknik Informatika STMIK Budi Darma Medan*.
- [2] Eki Roziqa Maris. (2017). Analisis Kepuasan Pelanggan Menggunakan Algoritma C4, Semarang. *Program Studi Teknik Informatika Universitas Nurswantoro Semarang*.
- [3] Haryanto, F. F. (2017). Penerapan Algoritma C4.5 Untuk Memprediksi Penerimaan Calon Pegawai Pada PT WISE. *Program Studi Teknik Informatika Universitas Multi Media Nusantara, Tangerang Banten*.
- [4] Ivandari. (2015). Model Keputusan Untuk Klasifikasi Persetujuan Kredit Menggunakan Algoritma C4.5. *Jurnal Litbang Kota Pekalongan*.
- [5] Jumanto. (2014). Implementasi Data Mining Algoritma C.45 Untuk Memprediksi Perilaku Mahasiswa Diploma 3 Melanjutkan Strata 1 Di STMIK AMIKOM Yogyakarta. *STMIK AMIKOM Yogyakarta*.
- [6] Larissa Maharani. (2015). Klasifikasi Nasabah Bank Menggunakan Algoritma

- C4.5 Sebagai Dasar Pemberian Kredit. *Padang Fakultas Ilmu Komputer Universitas Putera Indonesia Padang.*
- [7] Liliana Swastina. (2013). Penerapan Algoritma C4.5 Untuk Penentuan Jurusan Mahasiswa, Banjarmasin. *Program Teknik Informatika STMIK Indonesia Banjarmasin.*
- [8] Nurcahyo, G. W. (2008). Penerapan Data Mining dengan Algoritma Apriori untuk Mendukung Strategi Promosi Pendidikan. *Universitas Putra Indonesia "YPTK" Padang Sumatera Barat.*
- [9] Sukma Putri Utari. (2015). Implementasi Metode C.45 Untuk Menentukan Guru Terbaik Pada SMK 1 Percut Sei Tuan Medan. *Teknik Informatika STMIK Budi Darma Medan.*