

Penentuan Kelayakan Penggunaan Barang dengan Model Aturan Algoritma C4.5

Determining Product Suitability using Rule-Based Model with C4.5 Algorithm

Chintya Carolina Situmorang ¹, Dedy Hartama ², Irfan Sudahri Damanik ³, Jaya Tata Hardinata ⁴

^{1,2,3} STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

⁴ Universitas HKBP Nommensen Pematangsiantar, Pematangsiantar, Indonesia

Article Info

ABSTRAK

Genesis Artikel:

Diterima, 23 Februari 2023

Direvisi, 20 Maret 2023

Disetujui, 31 Maret 2023

Kata Kunci:

Data Mining
Pohon Keputusan
Model Aturan
C4.5
Kelayakan Barang

Sebuah gudang perhotelan harus memiliki pengadaan barang yang rapi, baik, aman, nyaman dan layak digunakan. Kendala yang sering terjadi pada gudang adalah barang rusak dan tidak layak untuk dipakai. Jumlah produksi permintaan barang yang terkadang tidak stabil membuat barang pada gudang tidak diperhatikan lagi kelayakan barangnya. Untuk mengetahui barang yang layak pakai maka dibutuhkan rekomendasi yang tepat. Algoritma C4.5 dengan teknik data mining merupakan perekomendasi yang tepat dalam menganalisis jumlah data yang banyak untuk diklasifikasikan. Data yang digunakan pada penelitian ini adalah data barang pada gudang Hotel Sapadia Pematangsiantar. Dengan implementasi algoritma C4.5 yang menghasilkan pohon keputusan dapat membantu pihak gudang dalam menentukan barang yang masih layak di gunakan pada aktifitas di perhotelan. Penelitian ini menghasilkan variabel terbaik dari model aturan yang digunakan untuk menentukan kelayakan barang adalah fisik barang. Akurasi model aturan yang dihasilkan dari pemodelan Algoritma C4.5 sebesar 99,02% terhadap kelayakan barang.

ABSTRACT

Keywords:

Data Mining
Decision Tree
Rules Model
C4.5
Goods Eligibility

A hotel warehouse must have orderly, good, safe, comfortable, and usable procurement of goods. The common issue that occurs in a warehouse is damaged and unusable goods. The fluctuating production demand for goods sometimes leads to neglecting the quality of the goods in the warehouse. To determine usable goods, appropriate recommendations are needed. The C4.5 algorithm with data mining techniques is an appropriate recommendation for analyzing a large amount of data for classification. The data used in this study is the inventory data of Hotel Sapadia Pematangsiantar's warehouse. Implementing the C4.5 algorithm that produces a Decision Tree can assist the warehouse in determining which goods are still usable for hotel activities. This study resulted in the best variable from the rule model used to determine the feasibility of goods being the physical condition of the goods. The accuracy of the rule model generated from the C4.5 Algorithm modeling is 99.02% against the feasibility of goods.

This is an open access article under the [CC BY-SA](#) license.



Penulis Korespondensi:

Chintya Carolina Situmorang,
Program Studi Sistem Informasi,
STIKOM Tunas Bangsa, Pematangsiantar, Indonesia
Email: chintyasitumorang@gmail.com

1. PENDAHULUAN

Manajemen barang yang efektif merupakan faktor penting dalam menjalankan kegiatan bisnis, termasuk di dalamnya kegiatan di bidang perhotelan [1]. Pengelolaan barang yang tepat dapat membantu meminimalkan biaya persediaan, memaksimalkan penggunaan barang, dan memastikan ketersediaan barang yang dibutuhkan saat dibutuhkan [2]. Gudang

perhotelan merupakan salah satu bagian penting dalam manajemen bisnis perhotelan yang bertanggung jawab untuk mengelola persediaan barang dan memastikan ketersediaan barang yang tepat pada waktu yang tepat.

Namun, salah satu permasalahan yang sering terjadi pada gudang perhotelan adalah terdapatnya barang yang rusak atau tidak layak digunakan lagi [3]. Masalah ini dapat terjadi karena beberapa faktor, seperti kurangnya perhatian terhadap kondisi barang di gudang, jumlah produksi permintaan barang yang tidak stabil, atau ketidaktahuan tentang cara menentukan kelayakan penggunaan barang.

Ketika barang yang tidak layak digunakan ditemukan di gudang perhotelan, hal ini dapat berdampak negatif pada kelancaran operasional perhotelan. Penggunaan barang yang rusak atau tidak layak dapat mengurangi kualitas layanan yang diberikan dan mempengaruhi citra perusahaan. Oleh karena itu, diperlukan suatu mekanisme atau model yang dapat membantu pihak gudang dalam menentukan kelayakan barang yang akan digunakan dalam kegiatan perhotelan. Banyak teknologi berbasis komputer yang dapat digunakan untuk menyelesaikan masalah, diantaranya teknologi *Machine Learning* teknologi *Machine Learning* [4]–[7], Jaringan Saraf Tiruan [8]–[10], dan Data Mining [11]–[16] dapat juga digunakan untuk membangun model prediksi dan klasifikasi yang dapat membantu dalam proses pengambilan keputusan. Salah satu algoritma terakhir yang disebutkan (Data Mining), dianggap cocok untuk menyelesaikan masalah ini yakni algoritma C4.5.

Penelitian ini bertujuan untuk mengembangkan model aturan dengan menggunakan algoritma C4.5 dalam menentukan kelayakan penggunaan barang pada gudang perhotelan. Algoritma C4.5 merupakan salah satu algoritma pembelajaran mesin yang digunakan dalam analisis data dan memungkinkan pembuatan model klasifikasi berbasis pohon keputusan. Model aturan yang dihasilkan dari algoritma ini dapat membantu dalam menentukan kelayakan penggunaan barang berdasarkan berbagai variabel atau faktor yang mempengaruhi kondisi barang, seperti kondisi fisik, umur barang, dan frekuensi penggunaan.

Data yang digunakan dalam penelitian ini adalah data inventaris barang pada gudang Hotel Sapadia Pematangsiantar. Data ini dianalisis menggunakan algoritma C4.5 dan teknik data mining untuk menghasilkan model aturan yang dapat digunakan dalam menentukan kelayakan penggunaan barang pada gudang perhotelan. Hasil penelitian diharapkan dapat memberikan kontribusi dan solusi dalam meningkatkan efektivitas manajemen barang pada gudang perhotelan dan membantu dalam pengambilan keputusan terkait dengan kelayakan penggunaan barang.

2. METODE PENELITIAN

2.1. Pengumpulan Data

Data untuk penelitian ini diambil dari data primer dan data sekunder. Data primer di dapat dari wawancara dan observasi lapangan. Data sekunder di dapat dari data objek penelitian (database), studi literatur dan tulisan ilmiah tentang data mining. Metode pengumpulan data yang dilakukan dalam penelitian ini sebagai berikut:

1. Wawancara

Wawancara dilakukan dengan pegawai gudang Hotel Sapadia Pematangsiantar, hal yang ditanyakan penulis adalah tentang sistem dan metode yang digunakan dalam menentukan kelayakan barang.

2. Observasi

Penulis melakukan observasi langsung dengan melihat sistem yang digunakan dan mengumpulkan data-data yang diperlukan dalam penelitian seperti data barang yang diperoleh pada gudang hotel Sapadia.

3. Studi Literatur

Pada metode literatur, diharapkan penelitian yang penulis lakukan dapat mempermudah dalam menyelesaikan penelitian. Penulis mengumpulkan, membaca, mempelajari, dan mencatat literatur dari jurnal maupun buku yang berkaitan dengan algoritma C4.5.

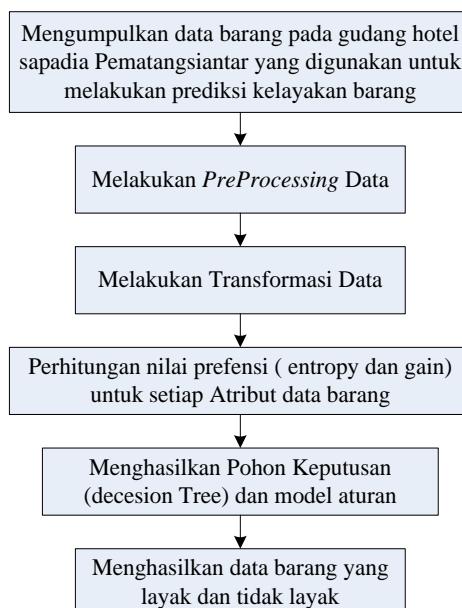
Tabel 1. Data Penelitian

Kode Barang	Nama Barang	Bagian/Departemen	Kategori	Masa Penggunaan Barang/bulan	Fisik Barang	Fisik Kemasan	Status Kelayakan
1101001	Sari Wangi isi 25	F&B	Minuman	Tiga Bulan	Baik	Baik	Layak
1101002	Coff Gold SPM Roasted @500	F&B	Minuman	Tiga Bulan	Baik	Baik	Layak
1101003	Kopi Merah Putih Classic	F&B	Minuman	Tiga Bulan	Baik	Baik	Layak
1101004	tepung Beras	F&B	Makanan	Tiga Bulan	Buruk	Baik	Tidak layak
1101005	tepung Kanji	F&B	Makanan	Tiga Bulan	Buruk	Baik	Tidak layak
1101034	Fish Sauce	F&B	Makanan	Tiga Bulan	Baik	Baik	Layak
1101035	Minyak Wijen	F&B	Makanan	Tiga Bulan	Baik	Baik	Layak
1101036	Minyak Goreng Fortune(new)	F&B	Makanan	Tiga Bulan	Buruk	Baik	Tidak layak
1101037	Kecap Bangau (Jrg)	F&B	Makanan	Tiga Bulan	Baik	Baik	Layak
1101038	Syrup Kurnia	F&B	Minuman	Tiga Bulan	Buruk	Baik	Tidak layak
1101039	Gula Tropicana Isi 100	F&B	Makanan	Tiga Bulan	Buruk	Baik	Tidak layak
1101040	Syrup Marjan Melon 600ML	F&B	Minuman	Tiga Bulan	Baik	Baik	Layak

Kode Barang	Nama Barang	Bagian/Departemen	Kategori	Masa Penggunaan Barang/bulan	Fisik Barang	Fisik Kemasan	Status Kelayakan
1101041	Gula 50 Kg (1 Goni)	F&B	Makanan	Tiga Bulan	Buruk	Rusak	Tidak layak
1101042	Aqua 600 Ml	F&B	Minuman	Tiga Bulan	Baik	Baik	Layak
1101043	Aqua Gelas/ cup	F&B	Minuman	Tiga Bulan	Baik	Baik	Layak
1101044	Mineral OH5 Cup	F&B	Minuman	Tiga Bulan	Baik	Baik	Layak
1101045	Air Mineral Terapi Oxygen	F&B	Minuman	Tiga Bulan	Buruk	Baik	Tidak layak
1101046	Air Mineral OH5 Botol 600ml	F&B	Minuman	Tiga Bulan	Buruk	Rusak	Tidak layak
1101047	Beers Bir Bintang Botol	F&B	Minuman	Tiga Bulan	Buruk	Baik	Tidak layak
1101048	Beers Bir Hitam Botol Kecil	F&B	Minuman	Tiga Bulan	Buruk	Rusak	Tidak layak
1101049	Bayclin 1000 ml	HK	Bahan kimia	Enam Bulan	Baik	Baik	Layak
1101050	Polish Furniture 20 Ltr	HK	Bahan kimia	Enam Bulan	Baik	Baik	Layak
1101051	Epro Rinse Add	HK	Bahan kimia	Enam Bulan	Baik	Baik	Layak
1101052	Floordress T510 2x5ltr	HK	Bahan kimia	Enam Bulan	Baik	Baik	Layak
1101053	Glass & Stainless Cleaner	HK	Bahan kimia	Enam Bulan	Baik	Rusak	Tidak layak
1101054	Bendurol G105 2x5Ltr	HK	Bahan kimia	Enam Bulan	Baik	Baik	Layak
1101055	Clean Carpet Shampoo	HK	Bahan kimia	Enam Bulan	Baik	Baik	Layak
1101056	Clean 202	HK	Bahan kimia	Enam Bulan	Baik	rusak	Tidak layak
1101087	Plastik Gula	HK	Bahan kimia	Enam Bulan	Baik	Baik	Layak
1101088	Plastik Sampah Hitam 100x12	HK	Bahan kimia	Enam Bulan	Buruk	Baik	Tidak layak
...
1101199	Kertas Glossy paper @100Lb	FO	ATK	Tak Terhingga	Baik	Baik	Layak

2.2. Tahapan Penelitian

Tahapan penelitian untuk menentukan kelayakan barang menggunakan metode klasifikasi dengan algoritma C4.5 adalah sebagai berikut:



Gambar 1. Alur Penelitian yang Diusulkan

Keterangan :

1. Melakukan pengumpulan data pada gudang hotel Sapadia yang akan digunakan dalam proses algoritma klasifikasi C4.5.
2. Melakukan Pre-Processing data dengan memilih data-data pada gudang hotel Sapadia yang akan digunakan untuk dilakukan proses klasifikasi algoritma C4.5.
3. Transformasi data dilakukan untuk mentransformasi atau mengubah data ke dalam bentuk yang dapat di proses dengan perhitungan algoritma C4.5.
4. Melakukan perhitungan *Entropy* dan *Gain* pada setiap atribut barang yang digunakan untuk dijadikan simpul akar pada pembuatan pohon keputusan.

5. Menghasilkan pohon keputusan (*Decision Tree*) yang didapat dari proses perhitungan *Entropy* dan information gain dan menghasilkan penjelasan dalam merepresentasikan sebuah pohon keputusan yang terbentuk.
6. Menghasilkan keputusan tentang data barang yang layak dan tidak layak digunakan.

2.3. Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Didalam algoritma C4.5 ini, pohon-pohon keputusan yang dibentuk berdasarkan kriteria-kriteria pembentuk keputusan. Berikut ini adalah langkah-langkah penyelesaian yang dilakukan dalam pengelompokan data barang berdasarkan atribut yang diperoleh dari data barang menggunakan algoritma C4.5.



Gambar 2. Alur Penelitian dengan Algoritma C4.5

Keterangan :

1. Menentukan jumlah data yang akan dimasukkan yang terdapat pada gudang dengan jumlah data sebanyak 199 data. Kriteria atau variabel yang digunakan dalam prediksi kelayakan barang adalah nama barang, bagian / departemen, satuan, kategori, fisik, masa tenggang dan status barang.
2. Memilih salah satu atribut sebagai akar untuk menghitung gain dan *Entropy* dari masing-masing atribut.
3. Memilih gain terbesar dalam membuat node akar dari pemilihan atribut.
4. Menghitung *Entropy* dan gain dari setiap atribut dengan menghilangkan atribut yang telah dipilih sebelumnya.
5. Melakukan pengecekan pada semua atribut yang sudah masuk pada kelas yang sama dengan memiliki nilai gain tertinggi.
6. Hasil aturan keputusan dibuat mengikuti pohon yang telah dibentuk sebelumnya.

3. HASIL DAN ANALISIS

3.1. Perhitungan Gain dan *Entropy*

Perhitungan gain dan *Entropy* dilakukan untuk membentuk suatu pohon keputusan. Dalam membuat pohon keputusan terlebih dahulu menghitung jumlah kasus untuk keputusan Layak dan jumlah kasus keputusan Tidak Layak. Berikut ini langkah-langkah dalam membentuk pohon keputusan dengan menggunakan algoritma klasifikasi C4.5.

1. Menghitung *Entropy*. Dari data *training* diketahui jumlah kasus ada 199 *record*, diketahui status kelayakan barang yang layak ada 103 *record* dan status kelayakan barang yang tidak layak ada 96 *record* sehingga *Entropy* nilai akar total yang didapat:

$$\begin{aligned} Entropy(S) &= \sum_{i=1}^n -pi * \log_2 pi \\ &= (-103/199 * \log_2 (103/199)) + (-96/199 * \log_2 (96/199)) \\ &= 0,999107262 \end{aligned}$$

2. Menghitung nilai *Entropy* pada atribut fisik barang dengan kategori baik, jumlah kasus ada 127 *record*, yang diketahui status kelayakan barang yang layak ada 101 *record* dan status kelayakan barang yang tidak layak ada 26 *record*.

$$\begin{aligned} Entropy(S) &= \sum_{i=1}^n -pi * \log_2 pi \\ &= (-101/127 * \log_2 (101/127)) + (-26/127 * \log_2 (26/127)) \\ &= 0,731276872 \end{aligned}$$

3. Menghitung nilai *Entropy* pada atribut fisik barang dengan kategori buruk, jumlah kasus ada 72 *record*, yang diketahui status kelayakan barang yang layak ada 1 *record* dan status kelayakan barang yang tidak layak ada 71 *record*.

$$\begin{aligned} Entropy(S) &= \sum_{i=1}^n -pi * \log_2 pi \\ &= (-1/72 * \log_2 (1/72)) + (-71/72 * \log_2 (71/72)) \\ &= 0,105591036 \end{aligned}$$

4. Menghitung nilai *gain* pada atribut fisik barang dengan total jumlah kasus 199 *record* dimana kategori buruk 72 *record* dan kategori baik 127 *record*.

$$\begin{aligned} Gain(S,A) &= Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \\ &= 0,999107262 - (127/199 * 0,731276872) + (72/199 * 0,999107262) \\ &= 0,494209185 \end{aligned}$$

5. Menghitung nilai *Entropy* pada atribut fisik kemasan dengan kategori baik, jumlah kasus ada 142 *record*, yang diketahui status kelayakan barang yang layak ada 101 *record* dan status kelayakan barang yang tidak layak ada 41 *record*.

$$\begin{aligned} Entropy(S) &= \sum_{i=1}^n -pi * \log_2 pi \\ &= (-101/142 * \log_2 (101/142)) + (-41/142 * \log_2 (41/142)) \\ &= 0,867078162 \end{aligned}$$

6. Menghitung nilai *Entropy* pada atribut fisik kemasan dengan kategori rusak, jumlah kasus ada 57 *record*, yang diketahui status kelayakan barang yang layak ada 0 *record* dan status kelayakan barang yang tidak layak ada 57 *record*.

$$\begin{aligned} Entropy(S) &= \sum_{i=1}^n -pi * \log_2 pi \\ &= (-0/57 * \log_2 (0/57)) + (-57/57 * \log_2 (57/57)) \\ &= 0 \end{aligned}$$

7. Menghitung nilai *gain* pada atribut fisik kemasan dengan jumlah kasus 199 *record* dimana kategori baik 142 *record* dan kategori rusak 57 *record*.

$$\begin{aligned} Gain(S,A) &= Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \\ &= 0,999107262 - (142/199 * 0,867078162) + (57/199 * 0) \\ &= 0,380388171 \end{aligned}$$

8. Menghitung nilai *Entropy* pada atribut masa penggunaan barang dengan tiga bulan, jumlah kasus ada 50 *record*, yang diketahui status kelayakan barang yang layak ada 21 *record* dan status kelayakan barang yang tidak layak ada 29 *record*.

$$\begin{aligned} Entropy(S) &= \sum_{i=1}^n -pi * \log_2 pi \\ &= (-21/50 * \log_2 (21/50)) + (-29/50 * \log_2 (29/50)) \\ &= 0,981453895 \end{aligned}$$

9. Menghitung nilai *Entropy* pada atribut masa penggunaan barang dengan enam bulan, jumlah kasus ada 61 *record*, yang diketahui status kelayakan barang yang layak ada 32 *record* dan status kelayakan barang yang tidak layak ada 29 *record*.

$$\begin{aligned} Entropy(S) &= \sum_{i=1}^n -pi * \log_2 pi \\ &= (-32/61 * \log_2 (32/61)) + (-29/61 * \log_2 (29/61)) \\ &= 0,998254569 \end{aligned}$$

10. Menghitung nilai *Entropy* pada atribut masa penggunaan barang dengan tidak terhingga, jumlah kasus ada 88 *record*, yang diketahui status kelayakan barang yang layak ada 49 *record* dan status kelayakan barang yang tidak layak ada 39 *record*.

$$\begin{aligned} Entropy(S) &= \sum_{i=1}^n -pi * \log_2 pi \\ &= (-49/88 * \log_2 (49/88)) + (-39/88 * \log_2 (39/88)) \\ &= 0,990664927 \end{aligned}$$

11. Menghitung nilai *gain* pada atribut fisik kemasan dengan jumlah kasus 199 *record* dimana kategori tiga bulan 50 *record*, kategori enam bulan 61 *record* dan kategori tidak terhingga 88 *record*.

$$\begin{aligned}
 Gain(S,A) &= Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \\
 &= 0,999107262 - (50/199*0,981453895) \\
 &\quad + (61/199*0,998254569) + 88/199*0,990664927) \\
 &= 0,008430191
 \end{aligned}$$

Berdasarkan hasil perhitungan nilai *Entropy* dan nilai *gain* untuk tiap atribut dapat dilihat pada tabel perhitungan *gain Node 1* berikut.

Tabel 2. Perhitungan Gain Node 1

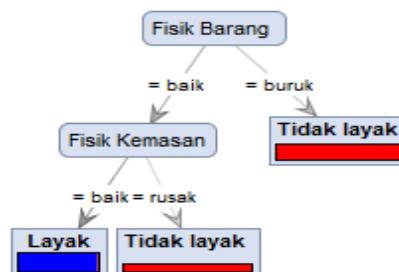
Node 1	Jumlah Kasus	Tidak Layak (S1)	Layak (S2)	Entropy	Gain
Total Fisik Barang	199	96	103	0,999107262	0,494209185
	Baik	127	26	0,731276872	
Fisik Kemasan	Buruk	72	71	0,105591036	0,380388171
	Baik	142	41	0,867078162	
Masa Penggunaan Barang	Rusak	57	57	0	0
	Tiga Bulan	50	29	0,981453895	0,008430191
	Enam Bulan	61	29	0,998254569	
	Tidak Terhingga	88	39	0,990664927	

Berdasarkan hasil perhitungan nilai *Entropy* dan *gain* yang telah dilakukan, diketahui bahwa nilai *gain* terbesar adalah variabel Fisik Barang dengan nilai 0,494209185 dan dijadikan simpul akar untuk mengulangi perhitungan nilai *Entropy* dan nilai *gain*. Dari hasil perhitungan nilai *Entropy* dan nilai *gain* untuk tiap atribut dengan variabel fisik barang yang menjadi simpul akar, dapat dilihat pada tabel perhitungan *gain Node 1.1* berikut.

Tabel 3. Perhitungan Gain Node 1.1.

Node 1.1	Jumlah Kasus	Tidak Layak (S1)	Layak (S2)	Entropy	Gain
Fisik Barang	Baik	127	26	0,731276872	0,667322783
	Buruk	103	1	0,078856014	
Fisik Kemasan	Rusak	24	0	0	0,026851501
	Baik	25	4	0,634309555	
Masa Penggunaan Barang	Enam Bulan	45	14	0,894451885	
	Tiga Bulan	57	8	0,58515699	
	Tidak Terhingga				

Setelah dilakukan hasil perhitungan *Entropy* dan *gain*, maka terbentuk pohon keputusan sebagai berikut:



Gambar 3. Pohon Keputusan yang Terbentuk

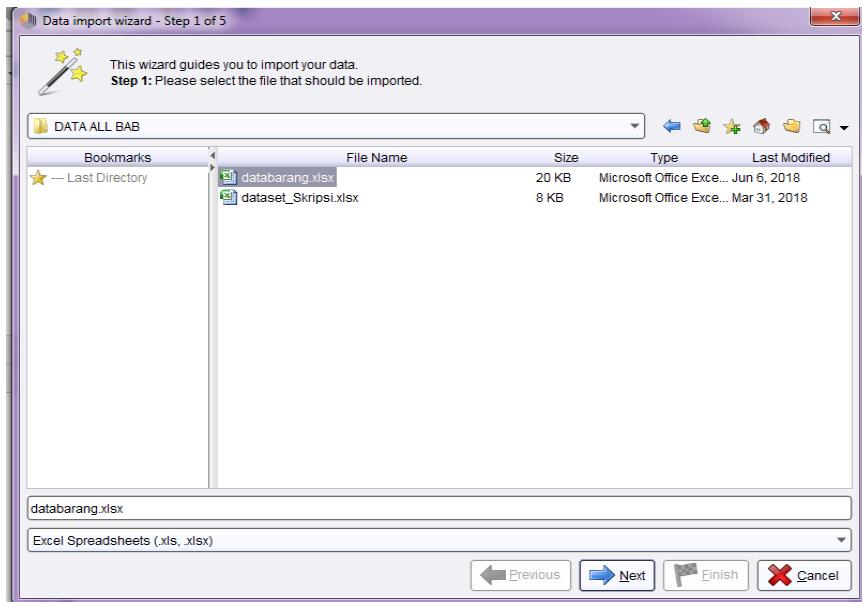
Berdasarkan bentuk pohon keputusan diperoleh *rule* model berupa teks sebagai berikut:

Fisik Barang = baik
| Fisik Kemasan = baik: Layak {Layak=101, Tidak layak =1}

| Fisik Kemasan = rusak: Tidak layak {Layak=0, Tidak layak =25}
 Fisik Barang = buruk: Tidak layak {Layak=1, Tidak layak =71}

3.2. Model Aturan

Dalam menjalankan proses dalam membentuk model aturan, terlebih dahulu *importing* data dari data yang sudah di transformasi terlebih dahulu ke dalam *microsoft excel* dalam format **.csv* agar bisa diakses menggunakan *software Rapid miner*. Atribut yang digunakan sebagai label adalah status kelayakan penggunaan barang menggunakan data yang telah dipilih yaitu data barang pada gudang Hotel Sapadia Pematangsiantar. Dalam melakukan *importing* data, dibutuhkan operator data barang **.csv*. Pilih data yang akan digunakan untuk di *import*. Lakukan *drag and drop* data yang dipilih kedalam *view process*. Berikut alur proses *import* data yang dilakukan.



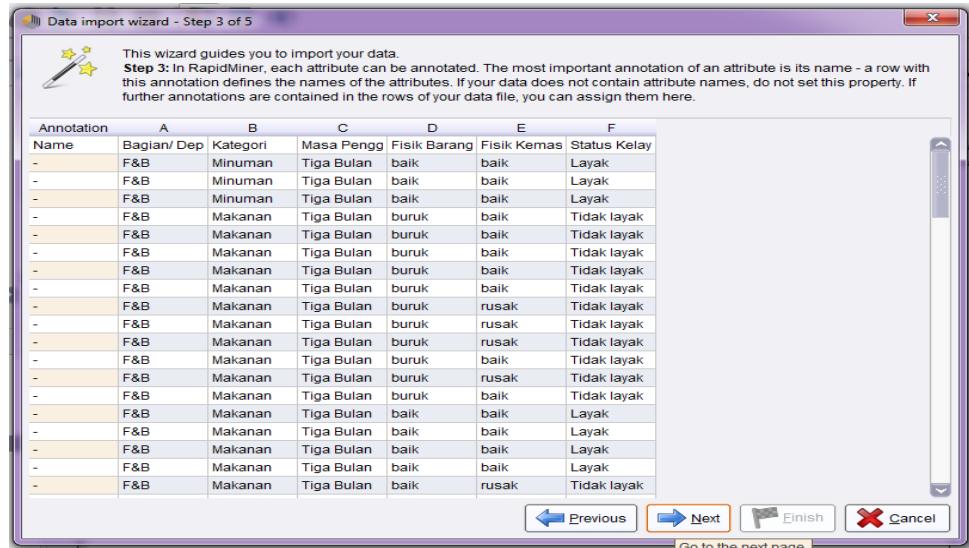
Gambar 4. Alur Proses *Import Data Step 1*

Pada *import* data *wizard step 1*, pilih data berformat **.csv* pada folder yang akan di *import*.

Kode_Baran	Nama_Bara	Bagian/ Dep	Kategori	Masa Pengg	Fisik Barang	Fisik Kemas	Status Kelay
1101001	Sari Wangi	F&B	Minuman	Tiga Bulan	baik	baik	Layak
1101002	Coff Gold SF	F&B	Minuman	Tiga Bulan	baik	baik	Layak
1101003	Kopi Merah	F&B	Minuman	Tiga Bulan	baik	baik	Layak
1101004	tepung Bera	F&B	Makanan	Tiga Bulan	buruk	baik	Tidak layak
1101005	tepung Kanji	F&B	Makanan	Tiga Bulan	buruk	baik	Tidak layak
1101006	tepung Panci	F&B	Makanan	Tiga Bulan	buruk	baik	Tidak layak
1101007	Kecap Asin	F&B	Makanan	Tiga Bulan	buruk	baik	Tidak layak
1101008	Tom Yam Pe	F&B	Makanan	Tiga Bulan	buruk	baik	Tidak layak
1101009	Essen Cokelat	F&B	Makanan	Tiga Bulan	buruk	rusak	Tidak layak
1101010	Vanila Bubuk	F&B	Makanan	Tiga Bulan	buruk	rusak	Tidak layak
1101011	Fermipan	F&B	Makanan	Tiga Bulan	buruk	rusak	Tidak layak
1101012	Tepung pulu	F&B	Makanan	Tiga Bulan	buruk	baik	Tidak layak
1101013	Bicarbonate	F&B	Makanan	Tiga Bulan	buruk	rusak	Tidak layak
1101014	Ajinamoto	F&B	Makanan	Tiga Bulan	buruk	baik	Tidak layak
1101015	Kerupuk udang	F&B	Makanan	Tiga Bulan	baik	baik	Layak
1101016	Beras	F&B	Makanan	Tiga Bulan	baik	baik	Layak
1101017	Agar-agar Si	F&B	Makanan	Tiga Bulan	baik	baik	Layak
1101018	Susu Bende	F&B	Makanan	Tiga Bulan	baik	baik	Layak
1101019	Susu Bende	F&B	Makanan	Tiga Bulan	baik	rusak	Tidak layak

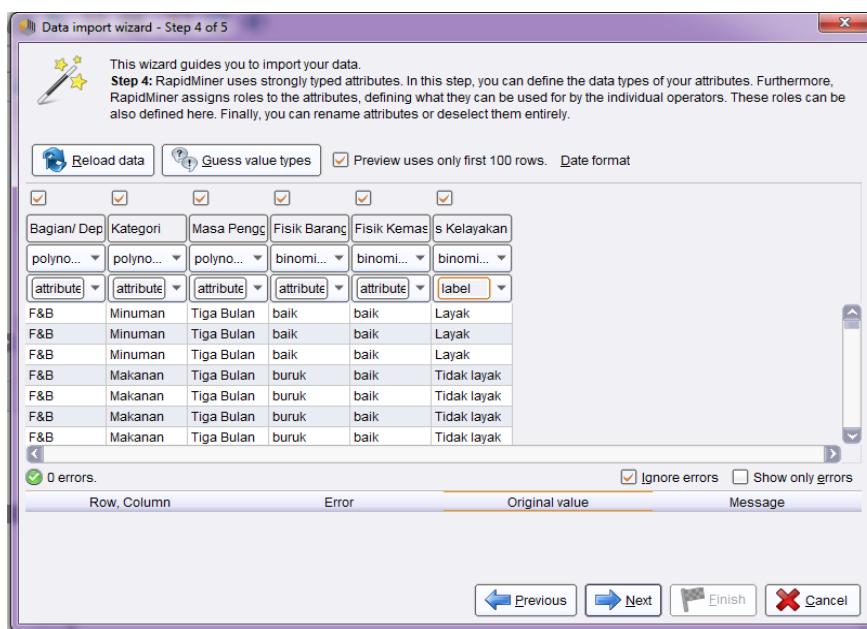
Gambar 5. Alur Proses *Import Data Step 2*

Pada data *import wizard step 2*, data barang **.csv* dengan *import excel* dimasukan. File *excel* dapat berisi beberapa lembar, cukup menandai beberapa sel yang akan dimuat. Pada proses kali ini, peneliti memilih sel 1 untuk dilanjutkan prosesnya kemudian klik *next*.



Gambar 6. Alur Proses Import Data Step 3

Pada *step* ke 3 ini tidak ada dilakukan apapun, setiap atribut dapat di anotasi. Yang paling penting pada *step* 3 ini adalahnya namanya mengandung atribut maka dari itu langsung ke data *import wizard* *step* 4 dengan cara klik *Next* maka akan muncul form data *import wizard* *step* 4.



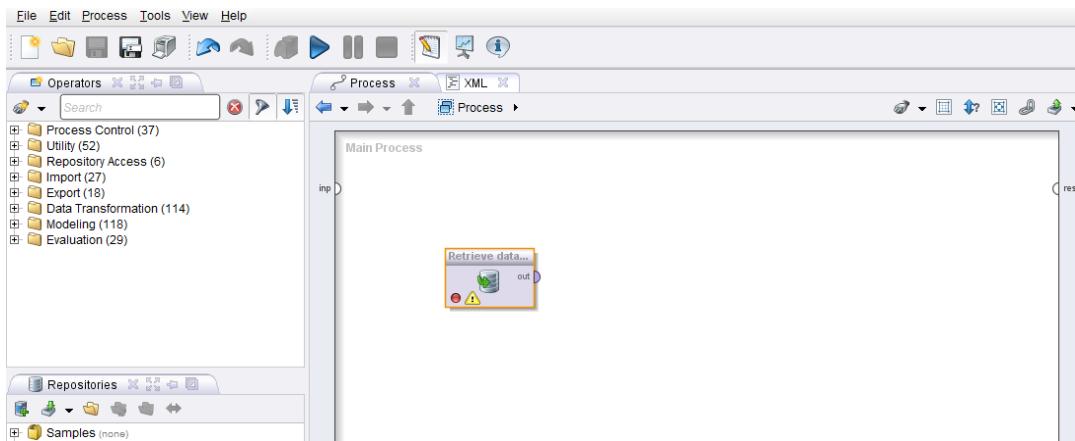
Gambar 7. Alur Proses Import Data Step 4

Pada form data *import wizard* *step* 4 seperti gambar 7 pilih salah satu atribut pada klasifikasi yang menjadi label atau disebut juga atribut predikat. Setelah ditentukan label yang dipilih maka klik *Finish*, tentukan tempat untuk menyimpan data pada *local repository*.



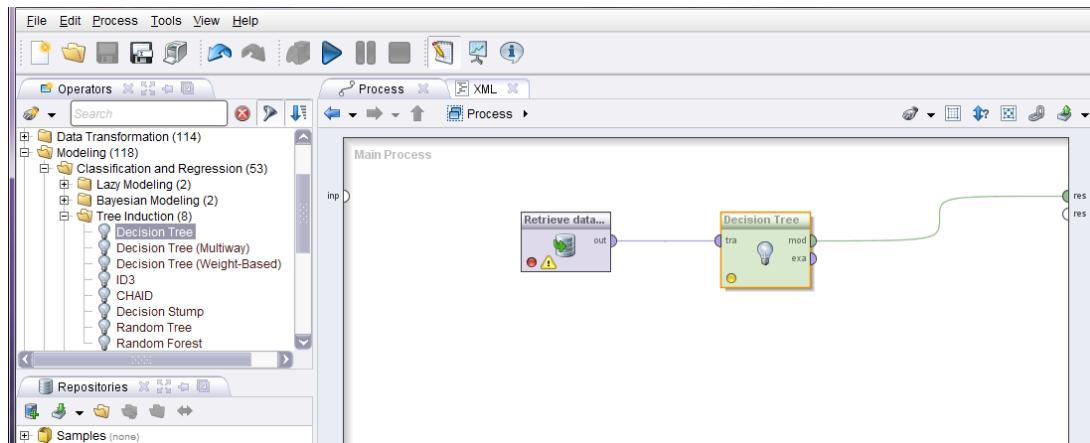
Gambar 8. Alur Proses Import Data Step 5

Pada saat alur *import* data selesai, maka selanjutnya *drag and drop* kedalam *main process*. Didalam *main process* akan memvalidasikan data yang dimodelkan kedalam algoritma dan fungsi *validation* ini adalah memaksimalkan nilai akurasi pengolahan data.



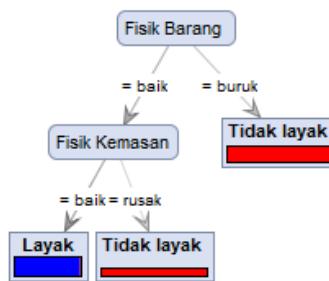
Gambar 9. Drag and Drop Data Barang

Langkah berikutnya *drag* dan *drop* operator *Decision Tree* lalu hubungkan pada *output*.



Gambar 10. Drag and Drop Operator Decesion Tree

Pada saat *tool run* diklik akan muncul pohon keputusan seperti gambar 11 berikut.



Gambar 11. Pohon Keputusan yang Terbentuk

Berdasarkan bentuk pohon keputusan yang terbentuk diperoleh *rule* model berupa teks sebagai berikut:

Fisik Barang = baik

- | Fisik Kemasan = baik: Layak {Layak=101, Tidak layak =1}
- | Fisik Kemasan = rusak: Tidak layak {Layak=0, Tidak layak =25}

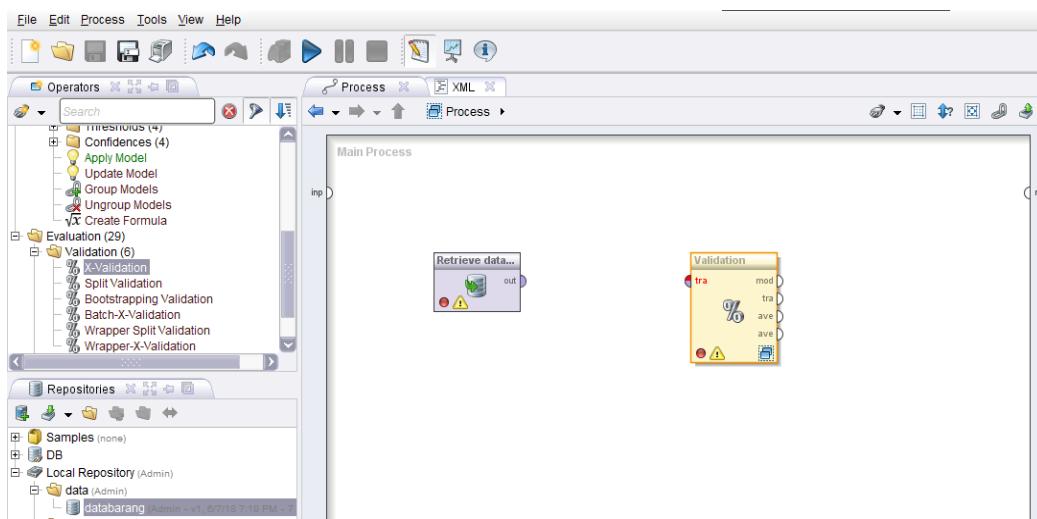
Fisik Barang = buruk: Tidak layak {Layak=1, Tidak layak =71}

Berdasarkan hasil pohon keputusan yang terbentuk diperoleh model aturan dari menentukan kelayakan penggunaan barang pada gudang Hotel Sapadia Pematangsiantar.

3.3. Validasi Performance

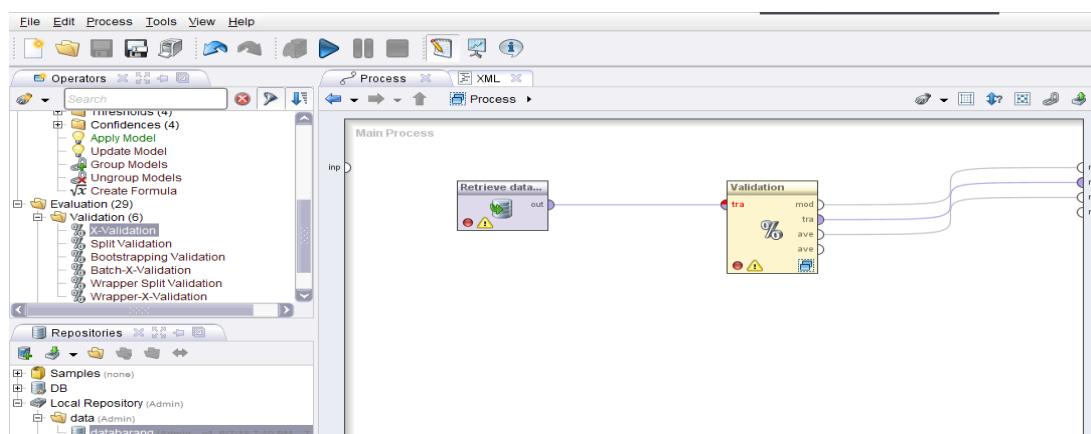
Dalam melakukan validasi *performance* digunakan operator *X-Validation*. Operator *X-validation* melakukan validasi silang untuk memperkirakan kinerja statistik operator pembelajaran (biasanya pada set data yang tak terlihat). Operator ini juga

digunakan untuk memperkirakan seberapa akurat suatu model yang akan tampil dalam praktik. Operator X-Validasi merupakan operator bersarang yang memiliki dua sub-proses: *training sub-process* (sub-proses percobaan) dan *testing sub-process* (sub-proses pengujian). Sub-proses percobaan digunakan untuk melatih sebuah model. Model yang terlatih kemudian diterapkan dalam sub-proses pengujian. Biasanya proses belajar mengoptimalkan parameter model untuk membuat model sesuai dengan data percobaan. Jika kita kemudian mengambil sampel independen dari data pengujian, umumnya model tersebut tidak cocok dengan data percobaan maupun data pengujian. Hal ini disebut dengan istilah '*over-pas*', dan sangat mungkin terjadi ketika ukuran set data training kecil, atau ketika jumlah parameter dalam model besar. Sehingga validasi silang merupakan cara untuk memprediksi kesesuaian model untuk satu set pengujian hipotesis ketika set pengujian eksplisit tidak tersedia. Untuk menemukan operator X-Validation, pilih *Evaluation* pada View Operator. Kemudian untuk melanjutkan proses pengolahan data, *drag and drop x validation* ke dalam main process.



Gambar 12. Operator Databarang *.csv dan X Validation pada Main Process

Pada saat *retrieve* data barang dan *x validation* didalam *main process*, hubungan keduanya dengan menarik garis dari tabel *retrieve* data barang ke operator *x validation* dan menghubungkan lagi ke garis hasil di sisi kanan.

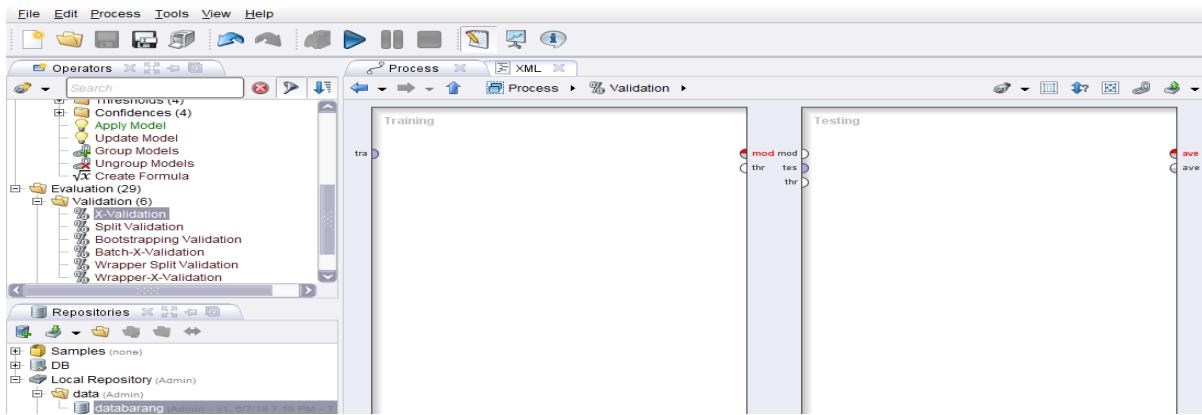


Gambar 13. Menghubungkan Tabel Data barang *.csv dengan Operator Validation

Pada Operator *x validation* memiliki *port input* yaitu, *training example set (tra)* sebagai *port input* memperkirakan, *Example Set* untuk melatih sebuah model (*training data set*). *Example Set* yang sama akan digunakan selama sub-proses pengujian untuk menguji model. Selain itu, operator ini juga memiliki *port output* sebagai berikut:

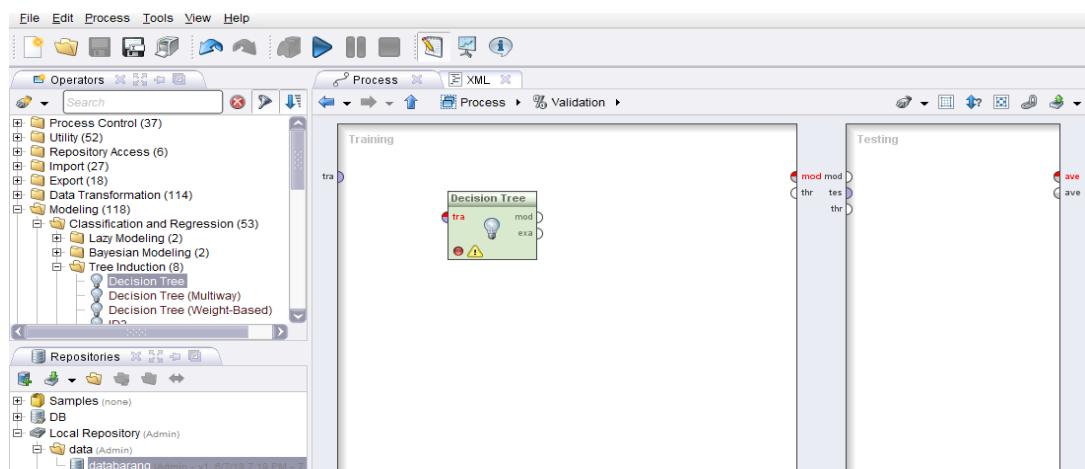
1. *Model (mod)*, pelatihan *sub-process* harus mengembalikan sebuah model yang dilatih pada input *example set* dan model yang dibangun *example set* disampaikan melalui port ini.
2. *Training Example Set (tra)*, the *example set* yang diberikan sebagai masukan pada port input pelatihan dilewaskan tanpa mengubah ke output melalui port ini. Port ini biasa digunakan untuk menggunakan kembali *example set* sama di operator lebih lanjut atau untuk melihat *example set* dalam *workspace result*.
3. *Average able (ave)*, sub-proses pengujian harus mengembalikan *vector kinerja*. Hal ini biasanya dihasilkan dengan menerapkan model dan mengukur kinerjanya. Dua port tersebut diberikan tetapi hanya digunakan jika diperlukan. Kinerja statistic dihitung dengan skema estimasi hanya perkiraan (bukan perhitungan yang tepat) dari kinerja yang

akan dicapai dengan model yang akan dibangun pada set data yang disampaikan secara lengkap. Setelah *table operator *.csv* dan *x validation* telah dihubungkan maka langkah selanjutnya *double klick* pada operator *x validation*



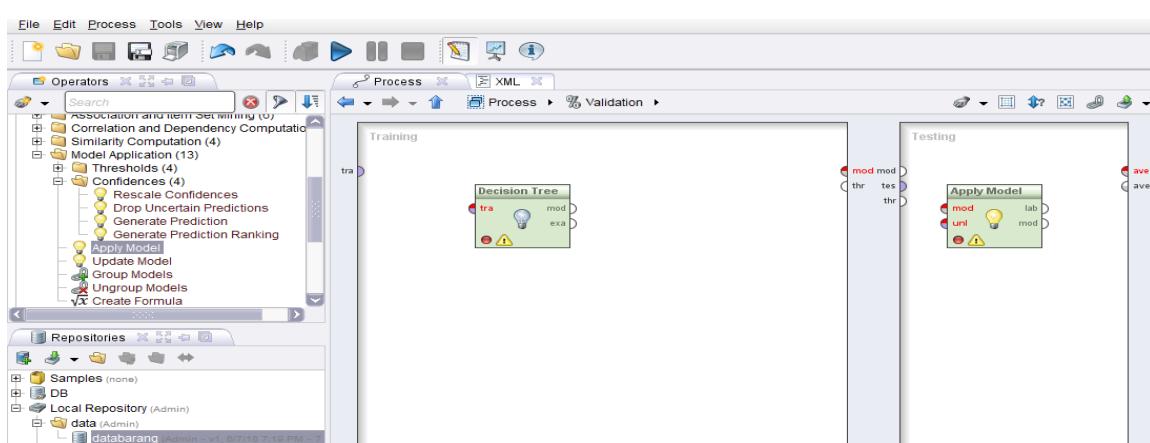
Gambar 14. Tampilan X Validation

Dalam membuat pohon keputusan, pada *x validation* diperlukan operator *Decision Tree*, operator *apply model* dan operator *performance*. Pada *training process*, *drag and drop* operator *Decision Tree* untuk membentuk pohon keputusan. Untuk menemukan operator *Decision Tree* pilih *Modeling* pada *view operator*, lalu pilih *Classification and Regression*, kemudian pilih *Tree Induction* dan pilih *Decision Tree*, kemudian *drag and drop* kedalam *training process*.



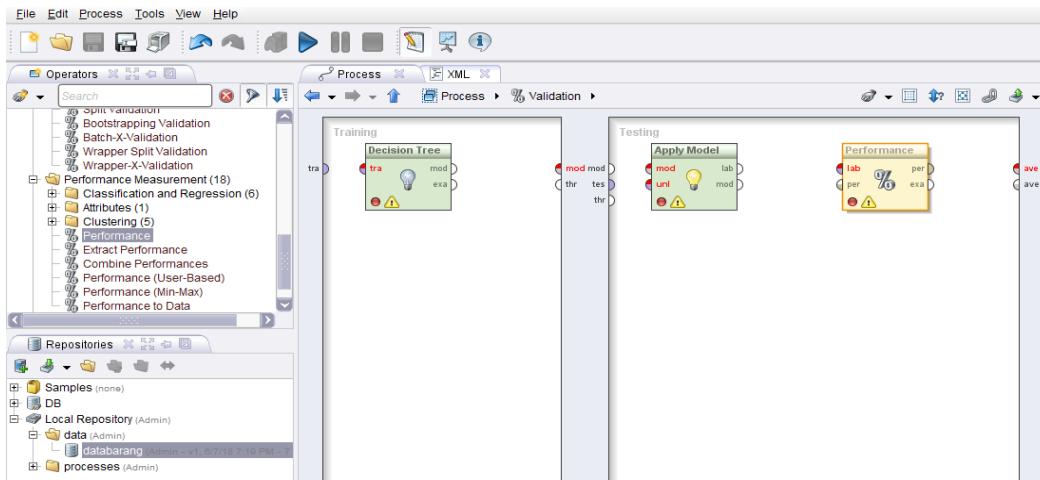
Gambar 15. Drag and Drop Decision Tree pada X Validation

Operator *apply model* digunakan dalam menerapkan suatu model terlatih pada sebuah *Example Set* untuk prediksi. Untuk menemukan operator *Apply Model*, pilih *Modeling* pada *View Operator*, lalu pilih *Model Application*, lalu pilih *Confidence* dan pilih *Apply Model*, kemudian *drag and drop* *apply model* kedalam *testing process*.



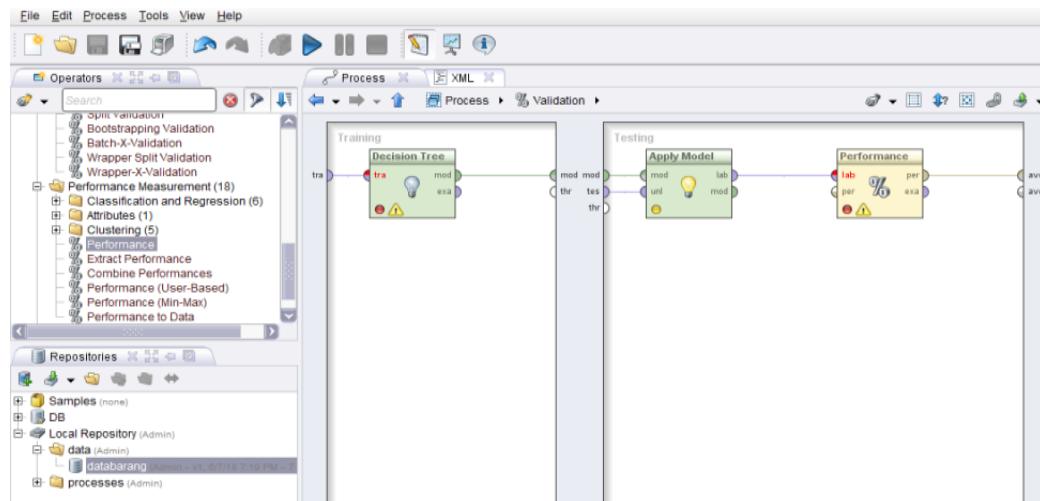
Gambar 16. Drop and Drag Operator ApplyModel kedalam X Validation

Dalam membuat pohon keputusan untuk menentukan apakah barang layak atau tidaknya digunakan, diperlukan juga operator *Performance*. Operator *performance* digunakan untuk evaluasi data barang. Operator ini memberikan daftar nilai kriteria data barang secara otomatis ditentukan agar sesuai dengan jenis kriteria. Untuk menemukan operator *Performance*, pilih *Evaluation* pada *View Operator*, lalu pilih *Performance and Measurement*, lalu pilih *Performance*. Setelah menemukan operator *Performance*, drag operator tersebut lalu *drop* ke dalam *view testing Process*.



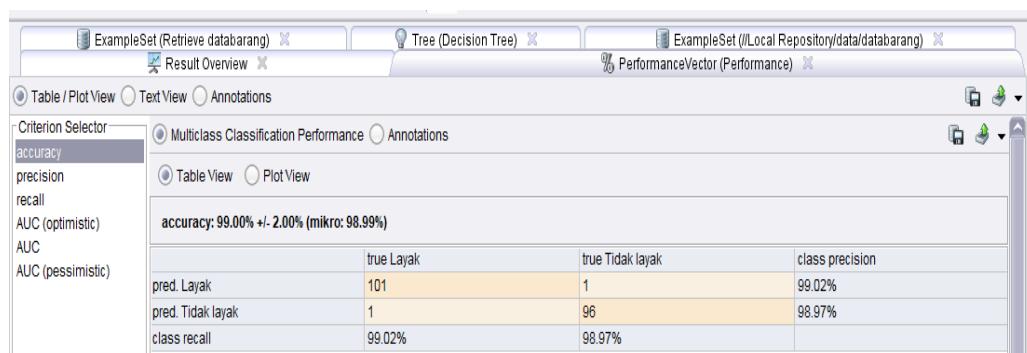
Gambar 17. Drag and Drop operator *Performance* kedalam *view Testing Process*

Setelah operator *Decision Tree*, *apply model* dan *performance* tersusun pada *view process*, kemudian hubungkan garis input dan output pada operator tersebut. Seperti gambar 18 berikut ini.



Gambar 18. Menghubungkan *Decision Tree* ke operator *Apply model* dan *Performance* ke dalam *view testing process*

Hasil pengujian Model Algoritma C4.5 *Classfier* ditunjukkan pada gambar berikut:

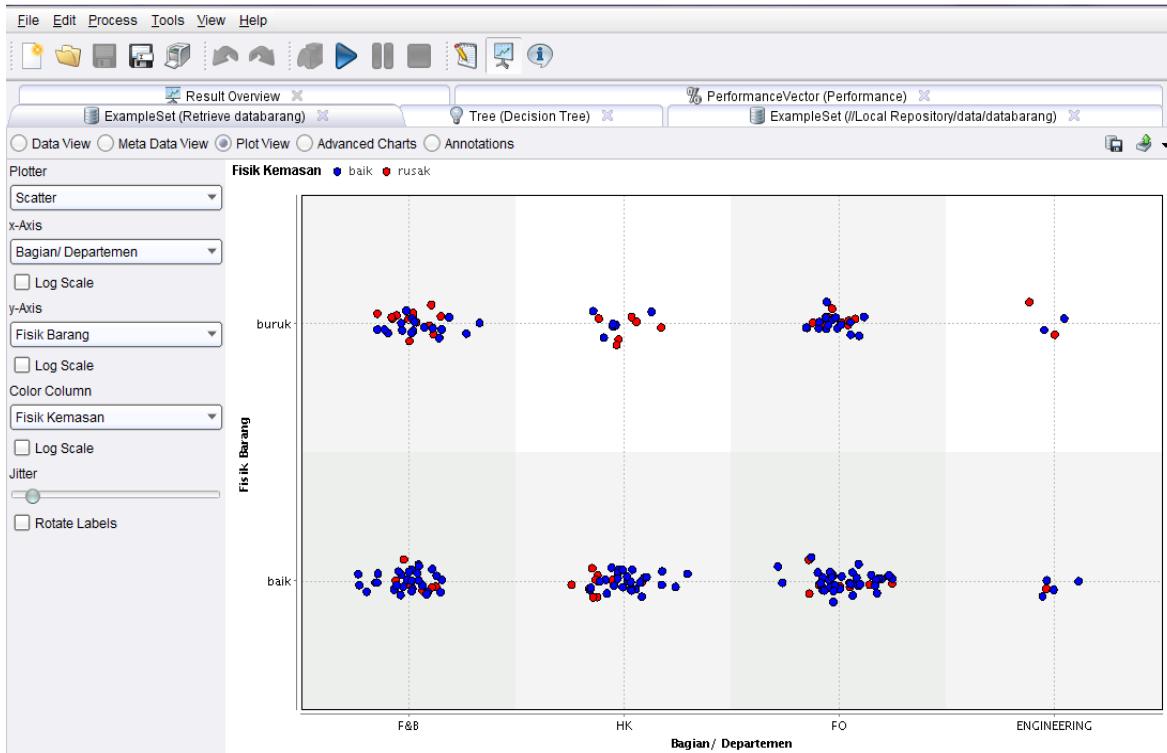


Gambar 19. Nilai *Accuracy Performance*

Keterangan :

1. Jumlah prediksi barang layak dan kenyataannya benar barang layak adalah 101 record (TP)
2. Jumlah prediksi barang layak dan kenyataannya tidak benar barang layak adalah 1 record (TN)
3. Jumlah prediksi barang tidak layak dan kenyataannya benar barang layak adalah 1 record (FP)
4. Jumlah prediksi barang tidak layak dan kenyataannya tidak benar barang layak adalah 96 record (FN)

Pada gambar Nilai *Accuracy Performance* dijelaskan bahwa prediksi layak memiliki nilai 101 dengan *class precision* 99,02%, prediksi tidak layak memiliki nilai 96 dengan *class precision* 98,97%. Dalam melihat ketergantungan hubungan data barang dengan data kuesioner dapat dilihat dari *scatter multiple*, gambar predikat kelayakan barang.



Gambar 20. Profil Kelayakan Barang

Gambar 20 merupakan hasil pengelompokan data yang 2 variabel yang disesuaikan dengan warnanya.

4. KESIMPULAN

Kesimpulan dari penelitian ini diantaranya adalah: Diperoleh suatu model aturan yang dapat memperlihatkan aturan keterhubungan antara atribut fisik kemasan dengan masa penggunaan barang, dari hasil penelitian didapat model aturan kelayakan penggunaan barang berdasarkan fisik barang. Dalam studi kasus menentukan kelayakan penggunaan barang pada Hotel Sapadia Pematangsiantar bahwa sebagian besar barang yang termasuk atribut fisik kemasan dengan kategori baik layak digunakan dan yang termasuk atribut fisik barang dengan kategori buruk tidak layak digunakan. Akurasi yang dihasilkan dari pemodelan Algoritma C4.5 sebesar 99,02%. Dengan jumlah true positif barang layak (TP) sebanyak 101 record, false positif barang layak (FP) sebanyak 1 record, jumlah true negative barang layak (TN) sebanyak 1 record, dan jumlah false negative barang tidak layak (FN) sebanyak 96 record. Pada penelitian ini menunjukkan ada dua dari tiga variabel memiliki korelasi yang sangat signifikan terhadap model aturan keterhubungan data barang untuk menentukan layak atau tidaknya barang untuk digunakan seperti yang diusulkan (fisik barang dan fisik kemasan).

REFERENSI

- [1] A. Dedy dan Y. Alfandi, "Pengaruh kualitas pelayanan dan fasilitas hotel terhadap kepuasan pelanggan di sari ater hot springs resort ciater," *Jurnal Sains Manajemen*, vol. 4, no. 1, hal. 18–25, 2022.
- [2] S. S. S. P. Sitorus dan A. Suseno, "Evaluasi Pengendalian Persediaan Barang Dagang," *Jurnal Serambi Engineering*, vol. 8, no. 2, hal. 5517–5527, 2023.
- [3] S. H. Amira, "Implementation and evaluation between single-used expandable strategy and environmental impact on supply chain and pallets on pt. sumber rezeki palletindo," *COSTING: Journal of Economic, Business and Accounting*, vol. 6, no. 2, hal. 1715–

1732, 2023.

- [4] I. S. Purba *dkk.*, "Accuracy Level of Backpropagation Algorithm to Predict Livestock Population of Simalungun Regency in Indonesia," *Journal of Physics: Conference Series*, vol. 1255, no. 1, hal. 1–6, 2019.
- [5] W. Saputra, J. T. Hardinata, dan A. Wanto, "Resilient method in determining the best architectural model for predicting open unemployment in Indonesia," *IOP Conference Series: Materials Science and Engineering*, vol. 725, no. 1, hal. 1–7, 2020.
- [6] I. A. R. Simbolon, F. Yatussa'ada, dan A. Wanto, "Penerapan Algoritma Backpropagation dalam Memprediksi Persentase Penduduk Buta Huruf di Indonesia," *Jurnal Informatika Upgris*, vol. 4, no. 2, hal. 163–169, 2018.
- [7] A. Wanto dan J. T. Hardinata, "Estimations of Indonesian poor people as poverty reduction efforts facing industrial revolution 4.0," *IOP Conference Series: Materials Science and Engineering*, vol. 725, no. 1, hal. 1–8, 2020.
- [8] R. E. Pranata, S. P. Sinaga, dan A. Wanto, "Estimasi Wisatawan Mancanegara Yang Datang ke Sumatera Utara Menggunakan Jaringan Saraf," *Jurnal semanTIK*, vol. 4, no. 1, hal. 97–102, 2018.
- [9] W. Saputra, J. T. Hardinata, dan A. Wanto, "Implementation of Resilient Methods to Predict Open Unemployment in Indonesia According to Higher Education Completed," *JITE (Journal of Informatics and Telecommunication Engineering)*, vol. 3, no. 1, hal. 163–174, Jul 2019.
- [10] A. Wanto *dkk.*, "Levenberg-Marquardt Algorithm Combined with Bipolar Sigmoid Function to Measure Open Unemployment Rate in Indonesia," in *The 3rd International Conference of Computer, Environment, Agriculture, Social Science, Health Science, Engineering and Technology (ICEST)*, 2021, no. 1, hal. 22–28.
- [11] M. A. Hanafiah dan A. Wanto, "Implementation of Data Mining Algorithms for Grouping Poverty Lines by District/City in North Sumatra," *International Journal of Information System & Technology*, vol. 3, no. 2, hal. 315–322, 2020.
- [12] A. Pradipta, D. Hartama, A. Wanto, S. Saifullah, dan J. Jalaluddin, "The Application of Data Mining in Determining Timely Graduation Using the C45 Algorithm," *IJISTECH (International Journal of Information System & Technology)*, vol. 3, no. 1, hal. 31–36, 2019.
- [13] I. Parlina *dkk.*, "Naive Bayes Algorithm Analysis to Determine the Percentage Level of visitors the Most Dominant Zoo Visit by Age Category," *Journal of Physics: Conference Series*, vol. 1255, no. 1, hal. 012031, 2019.
- [14] N. Arminarahmah, A. D. GS, G. W. Bhawika, M. P. Dewi, dan A. Wanto, "Mapping the Spread of Covid-19 in Asia Using Data Mining X-Means Algorithms," *IOP Conference Series: Materials Science and Engineering*, vol. 1071, no. 1, hal. 1–7, 2021.
- [15] T. H. Sinaga, A. Wanto, I. Gunawan, S. Sumarno, dan Z. M. Nasution, "Implementation of Data Mining Using C4.5 Algorithm on Customer Satisfaction in Tirta Lihou PDAM," *Journal of Computer Networks, Architecture, and High-Performance Computing*, vol. 3, no. 1, hal. 9–20, 2021.
- [16] T. Imandasari, E. Irawan, A. P. Windarto, dan A. Wanto, "Algoritma Naive Bayes Dalam Klasifikasi Lokasi Pembangunan Sumber Air," in *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, 2019, vol. 1, hal. 750–761.