IMPLEMENTASI DATA MINING SEBAGAI INFORMASI STRATEGIS PENJUALAN BATIK (STUDI KASUS BATIK MAHKOTA LAWEYAN)

Yusuf Sulistyo Nugroho[1], Fatah Yasin Al Irsyadi[1]
[1]Fakultas Komunikasi dan Informatika, Universitas Muhammadiyah Surakarta
Jl. A. Yani Tromol Pos I Pabelan, Kartasura, Surakarta, Indonesia
[*]E-mail : yusuf.nugroho@ums.ac.id

Abstrak

Sebuah data warehouse yang telah dimiliki oleh Batik Mahkota Laweyan bisa dijadikan sebagai sumber informasi bagi manajemen perusahaan terkait dengan rencana strategis perusahaan di masa yang akan datang yang berkaitan dengan tren kategori batik berdasarkan nama pola batik dan wilayah pemasarannya dari waktu ke waktu. Namun, hal ini belum didukung dengan metode yang baik serta kemudahan dalam hal akses data untuk menentukan rencana strategis perusahaan. Penelitian ini dimaksudkan untuk melakukan data mining guna menemukan informasi strategis terkait penjualan batik bagi manajemen Batik Mahkota Laweyan. Metode dalam data mining yang digunakan untuk menggali informasi dari data warehouse Batik Mahkota Laweyan adalah Decision Tree dengan algoritma penentuan kriteria atributnya menggunakan information gain. Variabel yang dicari informasi strategisnya adalah nama kategori batik berdasarkan variabel nama pola, propinsi wilayah pemasaran, dan jenis kelamin pelanggan. Hasil penelitian menunjukkan bahwa atribut yang memiliki pengaruh paling tinggi untuk menentukan klasifikasi nama kategori batik adalah nama pola (X1). Hal ini ditunjukkan dengan variabel X1 yang menempati sebagai simpul akar (root node) dalam diagram decision tree.

***Kata kunci**: batik mahkota, data mining, decision tree, information gain, klasifikasi*

1. PENDAHULUAN

Kemajuan teknologi informasi telah menyebabkan banyak orang dapat memperoleh data dengan mudah bahkan cenderung berlebihan. Data tersebut semakin lama semakin banyak dan terakumulasi, akibatnya pemanfaatan data yang terakumulasi tersebut menjadi tidak optimal. Sebagai contoh perusahaan retail yang akan memberikan brosur penawaran barang-barang yang dijual ke pelanggan sesuai basis data pelanggan yang mereka punya. Jika perusahaan retail tersebut mempunyai satu juta data pelanggan dan masing-masing pelanggan tersebut dikirimkan sebuah brosur penawaran dimana biaya pengiriman brosur tersebut adalah dua ribu rupiah, maka biaya yang akan dikeluarkan oleh perusahaan tersebut adalah dua juta rupiah per bulan. Dari penggunaan dana tersebut mungkin hanya sepertiganya atau bahkan 8% saja yang secara efektif membeli penawaran tersebut (Buhlman, 2002).

Batik Mahkota Laweyan merupakan perusahaan batik yang distribusinya sudah menyebar di berbagai daerah di Indonesia. Sebuah data warehouse yang telah dibangun pada penelitian sebelumnya bisa dijadikan sebagai sumber informasi bagi manajemen Batik Mahkota Laweyan terkait dengan rencana strategis perusahaan di masa yang akan datang yang berkaitan dengan tren jenis kategori batik berdasarkan nama pola barang dan propinsi wilayah pemasarannya dari waktu ke waktu. Namun, hal ini belum didukung dengan pembuatan laporan akhir dan kemudahan mengakses data untuk menentukan rencana strategis perusahaan.

Al Irsyadi (2014) menyatakan bahwa data warehouse Batik Mahkota Laweyan memiliki empat tabel dimensi (dimensi Produk, dimensi Wilayah, dimensi Waktu dan dimensi Pelanggan), empat tabel sub dimensi (dimensi Kategori, dimensi Sub_Kategori, dimensi Pola dan dimensi Jenis Kelamin) dan satu tabel Fakta, yaitu Fakta Penjualan. Proses ekstraksi dilakukan agar data transaksi sesuai dengan format data warehouse. Proses ekstraksi menghasilkan tabel-tabel dimensi (dimensi Produk, dimensi Wilayah, dimensi Waktu dan dimensi Pelanggan) dan tabel-tabel sub dimensi

(dimensi Kategori, dimensi Sub_Kategori, dimensi Pola dan dimensi Jenis Kelamin). Semua monitoring terhadap data-data penjualan produk Batik Mahkota Laweyan dilakukan menggunakan cube browser, sehingga masing-masing dimensi dapat dimasukkan atau dikeluarkan dari tabel untuk melihat data sesuai dengan kebutuhan analisis dengan cara *drag and drop*. Informasi yang ditampilkan oleh setiap dimensi dapat dilihat secara lebih rinci dengan proses *drill down* atau *roll up* sesuai dengan aturan hirarki field setiap dimensi.

Di sisi lain, data transaksi dari waktu ke waktu dalam data warehouse yang telah dimiliki jika dibiarkan menumpuk, hanya akan menjadi data sampah dan akan membebani basis data. Selain itu, data tersebut tidak memiliki manfaat bagi perusahaan jika tidak diolah dengan baik. Data mining merupakan salah satu metode yang dapat digunakan untuk mengatasi permasalahan penumpukan data. Metode ini dapat dimanfaatkan untuk menggali informasi strategis bagi perusahaan terhadap data-data yang menumpuk. Salah satu cara mengatasi masalah ini adalah dengan menggunakan teknik *Data Mining* (DM) yang bisa digunakan untuk pengolahan data menjadi sumber informasi strategis dengan metode klasifikasi dan klastering. *Data mining* dapat membantu sebuah organisasi yang memiliki data melimpah untuk memberikan informasi yang dapat mendukung pengambilan keputusan (Kiron et al, 2012). Dalam bidang analisis perusahaan dan manajemen resiko, *data mining* digunakan untuk merencanakan keuangan dan evaluasi aset, merencanakan sumber daya (*resources planning*) dan memonitor persaingan (Anggraini, 2009).

Berdasarkan latar belakang tersebut, penelitian ini dimaksudkan untuk melakukan data mining guna menemukan informasi strategis terkait penjualan batik dari data warehouse yang dimiliki sebagai sumber informasi bagi manajemen Batik Mahkota Laweyan berkaitan dengan tren kategori batik berdasarkan nama pola barang dan wilayah pemasarannya dari waktu ke waktu.

2. METODOLOGI

2.1. Jenis Penelitian dan Penentuan Variabel Data Mining

Penelitian ini termasuk jenis *research and development* karena akan meneliti dan mengembangkan sistem informasi dengan cara memonitor produksi, tren batik di suatu wilayah untuk menghasilkan rencana strategis pemasaran produk di masa yang akan datang.

Variabel untuk klasifikasi ini terdiri dari 2 jenis, yaitu:

a) Variabel dependen (Y)

Menurut Nugroho (2015), variabel Y adalah variabel yang akan dicari nilainya berdasarkan perhitungan-perhitungan menggunakan variabel lainnya. Variabel Y dalam proses klasifikasi ini adalah Nama kategori batik.

b) Variabel independen (X)

Variabel X merupakan variabel yang nilainya sebagai data perhitungan-perhitungan untuk menentukan nilai variabel Y.

Proses *data mining* ini, variabel X yang digunakan terdiri dari:

1) Nama pola batik, sebagai X1
2) Jenis kelamin pelanggan, sebagai X2
3) Propinsi distribusi penjualan, sebagai X3

2.2. Klasifikasi Kategori Batik Menggunakan *Decision Tree*

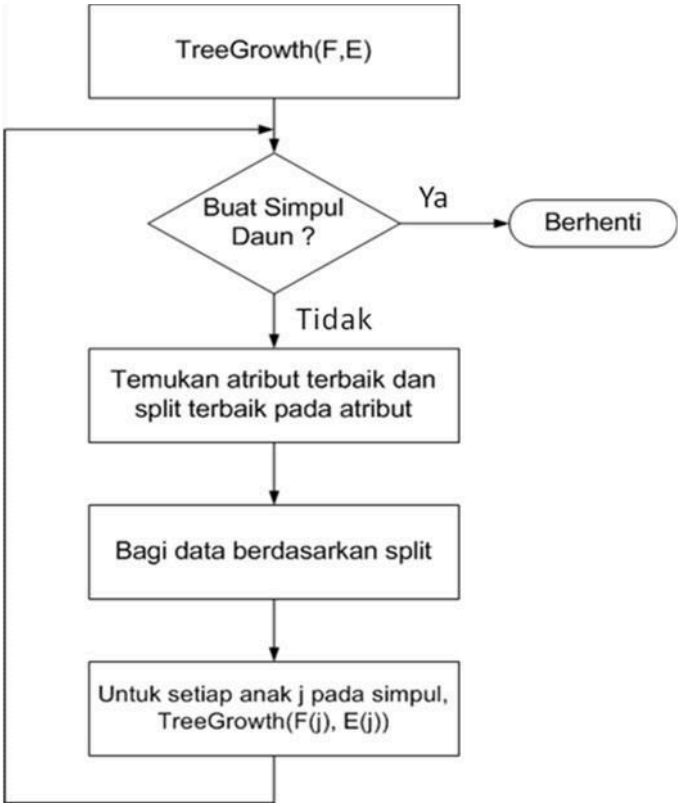
Tahapan analisis dilakukan untuk menentukan klasifikasi nama kategori batik berdasarkan nama nama pola, jenis kelamin pelanggan dan propinsi distribusi penjualannya. Klasifikasi nama kategori batik dilakukan dengan metode *Decision Tree* menggunakan algoritma *hunt* seperti pada gambar 1.

Adapun klasifikasi nama kategori batik dilakukan dengan metode *Decision Tree* menggunakan algoritma *hunt* serta penentuan atributnya menggunakan *information gain* berdasarkan nilai entropi dari masing-masing atribut yang telah ditentukan seperti terlihat pada persamaan (1) dan (2).

$$Entropi(y) = -p_1 log_2 p_1 - p_2 log_2 p_2 - \cdots - p_n log_2 p_n \quad \text{..... (1)}$$

$$gain(y, A) = entropi(y) - \sum_{c \in \overline{nilai}\,(A)} \frac{y_c}{y} entropi(y_c) \quad \text{..... (2)}$$

Penentuan simpul akar maupun simpul-simpul internal dalam pohon keputusan ditentukan berdasarkan nilai *information gain*. Atribut yang memiliki nilai *information gain* terbesar untuk setiap anak cabang, maka atribut tersebut ditentukan sebagai atribut yang menempati suatu simpul akar atau simpul internal. Proses pencarian simpul-simpul dilakukan hingga membentuk simpul daun untuk semua cabang dalam pohon keputusan. Jika suatu cabang sudah membentuk simpul daun, maka proses penghitungan nilai *information gain* dihentikan dan dilanjutkan pada cabang lain yang belum menemukan simpul daun.



Gambar 1. Algoritma Induksi Pohon Keputusan

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data dan *Preprocessing*

Penelitian ini menggunakan semua data dari *data warehouse* Batik Mahkota Laweyan yaitu sebanyak 1201 data penjualan batik sebagai data pelatihan untuk proses klasifikasi. Kelas data yang digunakan untuk *data mining* bersifat *binominal* atau *polynominal* sesuai aturan yang telah dibuat berdasarkan nilai datanya.

Tabel 1 menunjukkan contoh potongan hasil pengumpulan field-field dan data penjualan batik yang digunakan sebagai data pelatihan dalam proses data mining. Tabel 2 merupakan pembagian variabel dan kelas data yang digunakan dalam klasifikasi.

Tabel 1. Potongan Data Penjualan Batik

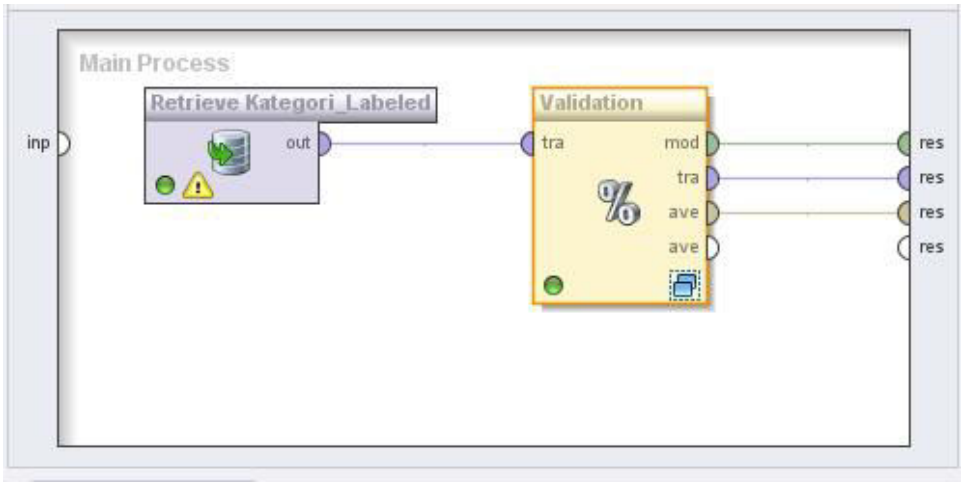
No	Nama Kategori	Nama Pola	Jenis Kelamin	Propinsi
1	Sarimbit	Tulis	Wanita	Jawa Timur
2	Longdress	Tulis	Wanita	Dki Jakarta
3	Kaos	Print	Pria	Jawa Timur
4	Kaos	Print	Pria	Jawa Timur
5	Kaos	Print	Pria	Jawa Timur
6	Bolero	Cap	Pria	Kepulauan Riau
7	Bolero	Cap	Pria	Kepulauan Riau
8	Sarimbit	Tulis	Wanita	Jawa Tengah
9	Selendang	Cap	Wanita	Jawa Barat
10	Selendang	Cap	Wanita	Jawa Barat
11	Bahan	Print	Wanita	Jawa Tengah
12	Bahan	Print	Wanita	Jawa Tengah
13	Bahan	Print	Wanita	Jawa Tengah
14	Celana	Cap	Wanita	Jawa Barat
15	Celana	Cap	Wanita	Jawa Barat

Tabel 2. Pembagian Variabel dan Kelas Data

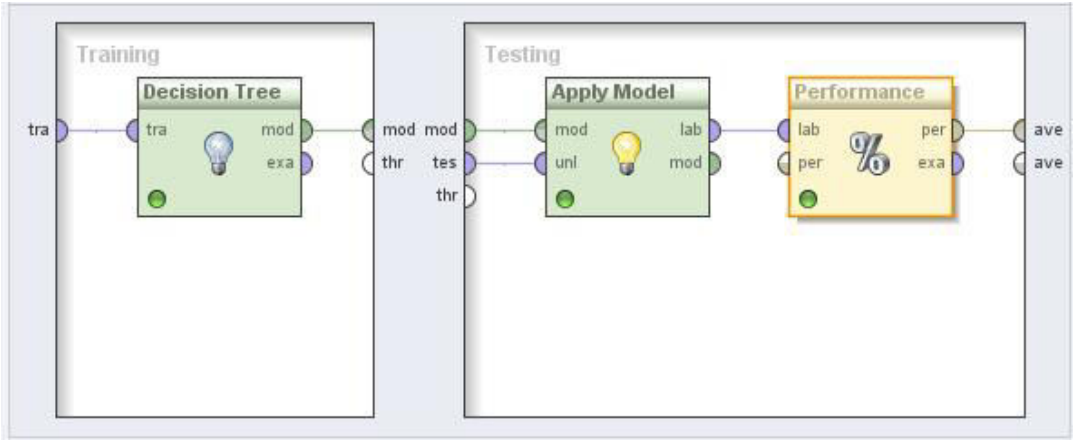
Variabel	Nama Field	Jenis Kelas Data	Kelas data yang digunakan
Y	Nama Kategori	*Binominal*	Abaya, Babydol, Blazer, Bahan, Batik, Blus, Bolero, Celana, Daster, Dress, Gantungan, Hem, Jaket, Jam Dinding, Jarik, Jas, Jilbab, Kaos, Kemben, Kipas, Koko, Longdress, Mukena, Rok, Sackdress, Sajadah, Sandal, Sarimbit, Sarung, Selendang, Setelan, Sprei, Syal, Tas.
X1	Nama Pola	*Polynominal*	Cap, Print, Tulis dan Lainnya.
X2	Jenis Kelamin	*Binominal*	Pria, Wanita
X3	Propinsi	*Binominal*	Bali, DKI Jakarta, Jawa Barat, Jawa Tengah, Jawa Timur, dan Kepulauan Riau.

3.2. Klasifikasi Data dengan Pohon Keputusan

Rancangan proses klasifikasi nama kategori batik dengan metode *Decision Tree* menggunakan aplikasi *Rapid Miner 5* ditunjukkan pada gambar 2 dan gambar 3. Rancangan proses tersebut kemudian dieksekusi untuk menghasilkan sebuah skema pohon keputusan untuk mengetahui tren nama kategori batik berdasarkan variabel-variabel bebas yang diajukan. Hasil eksekusi dapat diperoleh sebuah skema pohon keputusan yang ditunjukkan pada gambar 4.



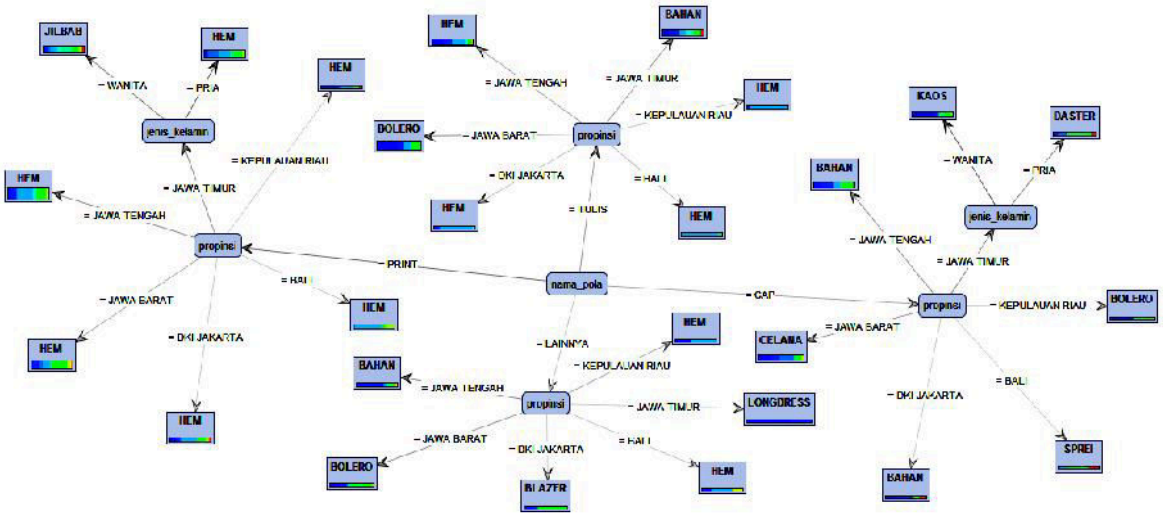
Gambar 2. Rancangan Proses Klasifikasi Nama Kategori Batik



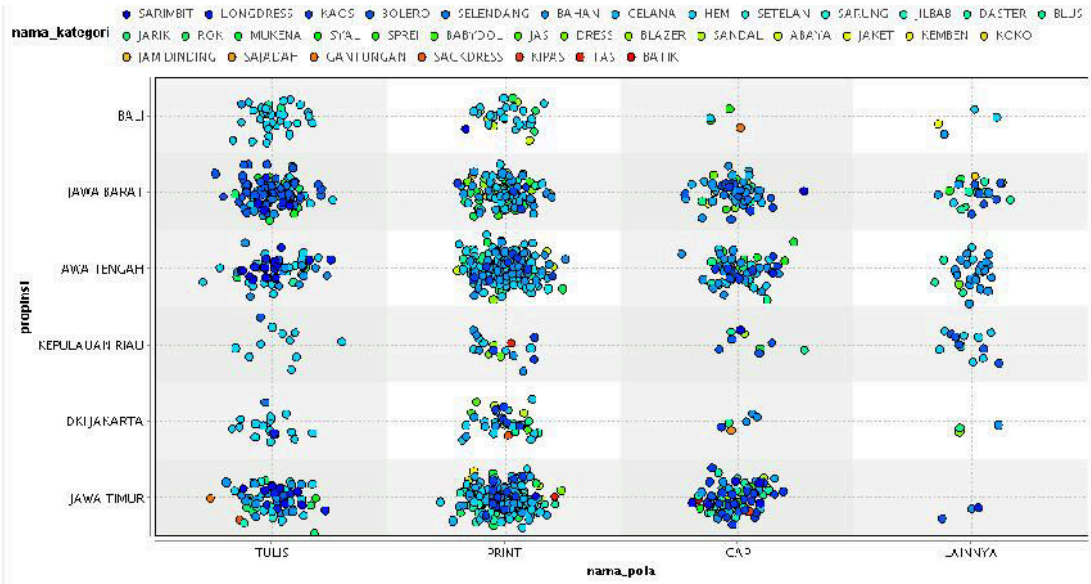
Gambar 3. Rancangan Proses Penerapan Model *Decision Tree* untuk Klasifikasi Nama Kategori Batik

Berdasarkan hasil pohon keputusan pada gambar 4, dapat dilihat bahwa atribut yang memiliki pengaruh paling tinggi untuk menentukan klasifikasi nama kategori batik adalah nama pola (X1). Hal ini ditunjukkan dengan variabel X1 yang menempati sebagai simpul akar (*root node*).

Contoh penerapan klasifikasi pohon keputusan pada sebuah data sebagai berikut, jika batik yang dicetak memiliki nama pola Print dan penjualannya akan didistribusikan di propinsi Jawa Tengah, maka batik tersebut diklasifikasikan dalam kelompok nama kategori Hem.



Gambar 4. Skema Decision Tree untuk Klasifikasi Nama Kategori Batik



Gambar 5. Distribusi Nama Kategori Batik terhadap Nama Pola dan Propinsi Menggunakan *Decision Tree*

Selain klasifikasi nama kategori batik dapat dilihat menggunakan pohon keputusan, contoh pola distribusi antar atribut terhadap wilayah dapat disajikan menggunakan grafik scatter plot pada gambar 5. Berdasarkan gambar 5 dapat dilihat bahwa setiap pola batik yang diproduksi oleh perusahaan memiliki wilayah distribusi yang paling banyak tingkat keminatannya. Sebagai contoh untuk pola batik tulis, wilayah yang paling tinggi tingkat keminatannya adalah Jawa Barat. Sementara itu, untuk pola batik print paling banyak diminati oleh wilayah Jawa Tengah, sedangkan untuk pola batik cap lebih banyak diminati oleh wilayah Jawa Timur. Namun secara keseluruhan, pola batik yang paling banyak terjual di berbagai wilayah adalah pola batik print.

Berdasarkan hal tersebut diatas, keputusan yang paling baik diambil oleh perusahaan terkait dengan tren pola batik yang akan diproduksi di masa-masa mendatang adalah dengan mempertimbangkan wilayah distribusinya. Selain itu, wilayah distribusi juga dapat digunakan sebagai pendukung keputusan untuk menentukan kategori produk yang akan dibuat oleh perusahaan batik Mahkota Laweyan.

Dengan demikian, perusahaan dapat memprediksi kebutuhan produksi batik yang akan dibuat terkait dengan kategori dan wilayah pemasarannya berdasarkan nama pola batik tersebut.

Hal ini tentunya dapat memberikan keuntungan yang cukup besar bagi perusahaan antara lain efisiensi biaya produksi dan waktu untuk membuat batik.

3.3. Rencana Strategis Batik Mahkota Laweyan

Berdasarkan interpretasi hasil penelitian, berikut beberapa kriteria batik yang bisa diterapkan sebagai sebuah kebijakan strategis bagi Batik Mahkota Laweyan:

1. Kriteria Nama Pola CAP.
   a. Jika Propinsi adalah Jawa Barat, maka nama kategori batik yang dijual adalah Celana.
   b. Jika Propinsi adalah Jawa Tengah dan DKI Jakarta, maka nama kategori batik yang dijual adalah Bahan.
   c. Jika Propinsi adalah Bali, maka nama kategori batik yang dijual adalah Sprei.
   d. Jika Propinsi adalah Kepulauan Riau, maka nama kategori batik yang dijual adalah Bolero.
   e. Jika Propinsi adalah Jawa Timur, maka nama kategori batik yang dijual tergantung dari Jenis Kelamin pelanggan, yaitu:
      1) Jika Wanita, maka nama kategori batik adalah Kaos.
      2) Jika Pria, maka nama kategori batik adalah Daster.
2. Kriteria Nama Pola PRINT.
   a. Jika Propinsi adalah Jawa Barat, Jawa Tengah, DKI Jakarta, Bali, dan Kepulauan Riau, maka nama kategori batik yang dijual adalah Hem.
   b. Jika Propinsi adalah Jawa Timur, maka nama kategori batik yang dijual tergantung dari Jenis Kelamin pelanggan, yaitu:
      1) Jika Pria, maka nama kategori batik adalah Hem.
      2) Jika Wanita, maka nama kategori batik adalah Jilbab.
3. Kriteria Nama Pola TULIS.
   a. Jika Propinsi adalah Jawa Barat, maka nama kategori batik yang dijual adalah Bolero.
   b. Jika Propinsi adalah Jawa Tengah, DKI Jakarta, Kepulauan Riau, dan Bali, maka nama kategori batik yang dijual adalah Hem.
   c. Jika Propinsi adalah Jawa Timur, maka nama kategori batik yang dijual adalah Bahan.
4. Kriteria Nama Pola LAINNYA.
   a. Jika Propinsi adalah Jawa Barat, maka nama kategori batik yang dijual adalah Bolero.
   b. Jika Propinsi adalah Jawa Tengah, maka nama kategori batik yang dijual adalah Bahan.
   c. Jika Propinsi adalah Bali dan Kepulauan Riau, maka nama kategori batik yang dijual adalah Hem.
   d. Jika Propinsi adalah DKI Jakarta, maka nama kategori batik yang dijual adalah Blazer.
   e. Jika Propinsi adalah Jawa Timur, maka nama kategori batik yang dijual adalah Longdress.

4. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan maka dapat disimpulkan bahwa:

(1) Proses *data mining* dengan menggunakan metode *decision tree* telah berhasil dilakukan untuk menggali informasi strategis bagi Batik Mahkota Laweyan. Hasil proses ini menunjukkan bahwa atribut yang memiliki pengaruh paling tinggi untuk menentukan klasifikasi nama kategori batik adalah nama pola (X1). Hal ini ditunjukkan dengan variabel X1 yang menempati sebagai simpul akar (*root node*) dalam diagram pohon keputusan.

(2) Propinsi (X3) sebagai variabel wilayah pemasaran menempati sebagai faktor kedua yang mempengaruhi kategori batik. Hal ini dapat dilihat dalam diagram pohon keputusan bahwa variabel propinsi terletak dalam simpul cabang di bawah nama pola (X1).

(3) Perusahaan Batik Mahkota Laweyan perlu memperhatikan nama pola dan wilayah pemasaran untuk membuat produk batik dengan nama kategori tertentu.

DAFTAR PUSTAKA

Al Irsyadi, Fatah Yasin. (2014). Implementasi Data Warehouse dan Data Mining untuk Penentuan Rencana Strategis Penjualan Batik. Jurnal KomuniTi Fakultas Komunikasi dan Informatika Universitas Muhammadiyah Surakarta. Volume 6 No. 1, Maret 2014.

Anggraini, Dyah. (2009). Analisis Perubahan Kelompok Berdasarkan Perubahan Nilai Jual Pada Bloomberg Market Data dengan Menggunakan Formal Concept Analysis. Available from :

http://www.gunadarma.ac.id-/Akuntansi/Artikel_92106032.pdf. [Diakses pada tanggal 30 Januari 2014]

Bühlman, P. dan Yu, B. (2002). Analyzing Bagging, The Annals of Statistics. Vol. 30 no. 4, hal 927-961.

Kiron, D., Shockley, R., Kruschwitz, N., Finch, G., & Haydock, M. (2012). Analytics: The Widening Divide. MIT Sloan Management Review, 53(2), 1-22.

Nugroho, Yusuf Sulistyo. (2015). Klasifikasi dan Klastering Mahasiswa Informatika Universitas Muhammadiyah Surakarta. Prosiding Bidang Teknik dan Rekayasa. The 1st University Research Colloquium 2015. ISSN: 2407-9189.