



Jurnal Inovasi Pendidikan Matematika

Volume 1, Nomor 2, Desember 2020

Penentuan Pengelompokan Penjualan dengan Kombinasi K-Means dan Hamming Distance

Retno Tri Vulandari¹, Dwi Remawati², Cindy Yuli Prastya³

Teknologi Informasi, STMIK Sinar Nusantara Surakarta¹²; Jl. K.H Samanhudi No.84-86, Purwosari, Kec. Laweyan, Kota Surakarta, Jawa Tengah 57149; retnotv@sinus.ac.id; dwirema@sinus.ac.id. Informatika, STMIK Sinar Nusantara Surakarta³; Jl. K.H Samanhudi No.84-86, Purwosari, Kec. Laweyan, Kota Surakarta, Jawa Tengah 57149; cindyuliprastya@gmail.com

Abstract

Determination of sales can be done in various ways. Grouping the determination of sales, consisting of items that sell and do not sell. The results of the grouping are used as considerations in producing goods and determining marketing strategies. Calculations performed in this study use the K-Means Clustering and Hamming Distance methods. The attribute data consists of non-numeric data, so the appropriate distance is the hamming distance. Is the item classified as goods that are included in the market or not. The results of this study are to produce sales report data that can report both salable and unsold goods in each branch of the company.

Keywords: *Hamming Distance, K-Means, Clustering*

Abstrak

Penentuan penjualan dapat dilakukan dengan berbagai cara. Pengelompokan penentuan penjualan, terdiri dari barang yang laku dan tidak laku. Hasil pengelompokan digunakan sebagai pertimbangan dalam memproduksi barang dan menentukan strategi pemasaran. Penghitungan yang dilakukan pada penelitian ini menggunakan metode K-Means Clustering dan Hamming Distance. Data atribut terdiri dari data non numerik, sehingga jarak yang sesuai adalah jarak hamming. Apakah barang tersebut tergolong barang yang dimasukkan dalam pasaran atau tidak. Hasil dari penelitian ini adalah menghasilkan data laporan penjualan yang dapat melaporkan barang laku dan tidak laku di setiap cabang perusahaan.

Kata kunci: *Jarak Hamming, K-Means, Pengelompokan*

INFO ARTIKEL

<p><i>ISSN</i> : 2733-0597 <i>e-ISSN</i> : 2733-0600 <i>DOI</i> : http://dx.doi.org/10.30587/postulat.v1i2.2090</p>	<p style="text-align: center;"><i>Jejak Artikel</i></p> <p>Submit Artikel: 2 Juni 2020</p> <p>Submit Revisi: 1 Oktober 2020</p> <p>Upload Artikel: 5 Desember 2020</p>
---	--

PENDAHULUAN

Perkembangan bisnis di Indonesia saat ini mengalami pertumbuhan yang cepat terutama di bidang bisnis retail. PT Cartenz Indonesia adalah salah satu perusahaan yang bergerak di bidang bisnis retail perlengkapan alat gunung. Saat ini sudah banyak produk yang dijual oleh PT Cartenz Indonesia seperti jaket gunung, sandal gunung, celana gunung, tas gunung, dan perlengkapan lainnya. Pada penelitian ini menggunakan metode k-means clustering dan metode hamming distance untuk mengelompokkan data non-numerik berdasarkan data penjualan. Indikator yang digunakan untuk perhitungan adalah jeni, size, warna dan warna kombinasi. data berbentuk non-numerik tersebut akan dikonversikan menjadi data biner terlebih dahulu.

Tujuan dari penelitian ini adalah menentukan barang mana yang laku dan tidak laku. Program yang dibuat dapat menghasilkan data laporan 10 barang laku dan 10 barang tidak laku. Hasil penelitian tersebut dapat menjadi bahan pertimbangan untuk menentukan produksi barang dan menentukan strategi marketing.

METODE PENELITIAN

Metode penelitian terdiri atas pengumpulan data dan pengembangan sistem. **Pengumpulan data** secara langsung baik berupa laporan penjualan ataupun hasil wawancara. Data primer dalam penelitian ini adalah data penjualan seluruh cabang PT Cartenz Indonesia yang berada di area Jawa Tengah meliputi Solo, Magelang, Semarang, Salatiga, Purwokerto, dan Yogyakarta.

Pengumpulan data dengan cara ini adalah cara pengambilan data dengan menggunakan mata tanpa pertolongan alat. Sehingga penulis mengamati berjalannya data di PT Cartenz Indonesia secara langsung. Mulai data masuk dan bagaimana mengolah data tersebut.

Wawancara merupakan cara yang ditujukan untuk memperoleh keterangan dengan tanya jawab dengan pengelola terkait. Dimana penulis melakukan wawancara dengan *manager sales* PT Cartenz Indonesia.

Pengembangan Sistem terdiri atas perumusan masalah, penentuan metode, pemecahan masalah, implementasi program, dan pengujian program. **Perumusan masalah** berupa permasalahan yang dihadapi oleh perusahaan terutama dibidang penjualan. Seperti yang terjadi pada PT Cartenz Indonesia dimana perusahaan tersebut bergerak di bidang produksi dan retail peralatan gunung. Tentunya banyak mengalami masalah salah satunya untuk menentukan strategi pasar dengan melihat barang mana yang laku dan tidak laku dipasaran terutama di area Jawa Tengah.

Dalam menentukan sebuah metode tentunya harus melihat dari masalah yang akan dihadapi terlebih dahulu. Permasalahan yang dihadapi adalah bagaimana menentukan barang laku dan tidak laku dalam suatu area yang terdapat beberapa cabang. Metode yang digunakan dalam penelitian ini adalah metode *K-Means Clustering* dan *Hamming Distance*. Metode *K-Means Clustering* digunakan untuk mengelompokkan data yang laku dan tidak laku. Metode *Hamming Distance* digunakan untuk mengukur jarak antar data menggunakan bilangan biner. Karena data yang digunakan untuk menghitung jarak merupakan data non-numerik.

Pemecahan masalah dilakukan dengan menggunakan metode yang sudah ditentukan sebelumnya. Berdasarkan data yang diperoleh tersebut data non-numerik akan dikonversikan menjadi bilangan biner terlebih dahulu. Kemudian jarak akan dihitung berdasarkan rumus *Hamming Distance*. Dari hasil penghitungan jarak tersebut data akan dikelompokkan menjadi data barang laku dan tidak laku.

Implementasi Program menggunakan pemrograman php dengan database MySQL. Metode tersebut diterapkan dalam sebuah sistem pemrograman php untuk memproses data sesuai dengan penghitungan dan tahapan metode *K-Means Clustering* dan *Hamming Distance*.

Pengujian sistem dilakukan untuk menguji apakah sistem yang diterapkan sudah sesuai. Pada pengujian ini peneliti menggunakan aplikasi Tanagra untuk membandingkan hasilnya. Tanagra merupakan aplikasi penghitungan untuk data mining. Dari hasil perhitungan

pengelompokkan aplikasi hasilnya sama jika dibandingkan dengan sistem yang dibuat. Dari hasil perbandingan tersebut sistem yang dibuat dapat dikatakan sesuai dengan hasil yang diharapkan.

HASIL PENELITIAN DAN PEMBAHASAN

Kriteria Barang Laku dan Tak laku

Barang dapat dikatakan laku apabila penjualannya lebih dari 50% dalam kurun waktu 3 bulan dari total stok yang dimiliki oleh sebuah cabang. Kemungkinan besar barang yang laku akan diproduksi kembali dengan melihat stok bahan baku yang ada. Barang dikatakan tidak laku apabila penjualannya dibawah 30% dalam kurun waktu 3 bulan dari total stok yang dimiliki oleh sebuah cabang. Barang yang tidak laku ini akan mendapatkan perlakuan khusus missal akan dimutasi ke cabang lain atau akan di berikan program diskon untuk meningkatkan produk tersebut.

Tahap Pengelompokan

Tahapan pengelompokkan pada penelitian ini ada beberapa tahap yang harus dilakukan, antara lain :

1. Langkah pertama adalah konversikan indikator yang digunakan untuk penghitungan ke dalam bentuk biner.

Tabel 1 Konversi Warna

No	Nama	Kode
1	Hitam	0000001
2	Merah	0000010
3	Biru	0000100
4	Hijau	0001000
5	Cokelat	0010000
6	Orange	0100000
7	Abu	1000000

Pada Tabel 1 menunjukkan daftar warna dasar yang digunakan untuk memproduksi barang. Pada tabel ini juga menampilkan konversi biner untuk digunakan pada perhitungan *hamming distance*.

Tabel 2. Konversi Warna Kombinasi

No	Nama	Kode
1	Hitam	0000000001

2	Biru	0000000010
3	Merah	0000000100
4	Kuning	0000001000
5	Orange	0000010000
6	Cokelat	0000100000
7	Hijau	0001000000
8	Ungu	0010000000
9	Abu	0100000000
10	Putih	1000000000

Tabel 2 warna kombinasi yang digunakan ada 10 macam. Sehingga type data binernya ada 10 digit angka sesuai dengan jumlah warna kombinasi yang digunakan.

Tabel 3. Konversi Ukuran

No	Nama	Kode
1	M	001
2	L	010
3	XL	100

Tabel 3 hanya terdapat 3 data. Data tersebut merupakan ukuran jaket yang digunakan untuk diproduksi. Ada ukuran M, L, dan XL.

Tabel 4. Konversi Jenis Jaket

No	Nama	Kode
1	Reguler	1
2	Extreme	0

Pada Tabel 4 terdapat 2 data yang menunjukkan jenis jaket yang diproduksi. Jaket extreme untuk suhu di bawa 20 derajat celcius dan jaket regular untuk suhu di atas 25 derajat celcius.

2. Kemudian Menentukan centroid awal yaitu centroid 1 dan centroid 2 secara acak. Centroid 1 dan centroid 2 juga dirubah dalam bentuk biner.

Penghitungan jarak dilakukan dengan menggunakan data yang sudah di konversi. Sebelum melakukan penghitungan yang pertama harus dilakukan adalah menentukan *Centroid* Awal. Tentukan Centroid 1 dan centroid 2 secara acak

Centroid 1 : Jaket L Reguler Adrenalin Hitam Merah | Qty 11

Centroid 2 : Jaket L Extreme Alpin Extreme Abu Hijau | Qty 1

Centroid 1	1	0000001	0000000100	010
Centroid 2	0	1000000	0001000000	010

3. Langkah ketiga adalah menghitung masing – masing jarak data ke centroid 1 dan centroid 2. Penghitungan hanya melalui 1 iterasi saja.

Hitung jarak Jacket L Regular Denly Hitam Merah | Qty 2 terhadap Centroid Awal.

- Ukuran L = 010
- Jenis Regular = 1
- Warna Hitam = 0000001
- Warna kombinasi Merah = 0000000100

Dalam pengitungan yang dilakukan menggunakan persamaan *Hamming Distance* yang sudah dijelaskan sebelumnya. Langkah pertama data dibandingkan dengan centroid 1

$$d_{ij} = \frac{\left[\left[\frac{(010 \text{ XOR } 010)}{3} \right] + \left[\frac{(1 \text{ XOR } 1)}{1} \right] + \left[\frac{(0000001 \text{ XOR } 0000001)}{7} \right] + \left[\frac{(0000000100 \text{ XOR } 0000000100)}{10} \right] \right]}{4}$$

$$d_{ij} = \frac{\left[\frac{0}{3} + \frac{0}{1} + \frac{0}{7} + \frac{0}{10} \right]}{4} = 0.00$$

Kemudian membandingkan centroid 1 dengan centroid 2

$$d_{ij} = \frac{\left[\left[\frac{(010 \text{ XOR } 010)}{3} \right] + \left[\frac{(1 \text{ XOR } 0)}{1} \right] + \left[\frac{(0000001 \text{ XOR } 1000000)}{7} \right] + \left[\frac{(0000000100 \text{ XOR } 0001000000)}{10} \right] \right]}{4}$$

$$d_{ij} = \frac{\left[\frac{0}{3} + \frac{1}{1} + \frac{2}{7} + \frac{2}{10} \right]}{4} = 0.37142$$

Ketentuan

- Jika 0 bertemu 1 maka hasilnya 1 begitu juga sebaliknya
 - Jika 1 bertemu 1 maka hasilnya 0
 - Pembagian 6 mewakili 6 jenis warna yang di gunakan
 - Pembagian 3 mewakili jenis ukuran yang digunakan
4. Langkah terakhir adalah mengelompokkan hasil perhitungan data. Data yang menghasilkan angka mendekati 0 masuk pada kelompok barang laku dan data yang nilainya semakin menjauhi 0 masuk pada kelompok barang tidak laku.

Nilai Hitung centroid 1 = 0.00 → Masuk kelompok c1

Nilai Hitung centroid 2 = 0.37142 → Masuk kelompok c2

Dari hasil perbandingan bahwa : nilai urutan 1 adalah nilai yang paling mendekati 0 dan itu artinya barang yang memiliki hasil perhitungan mendekati 0 adalah barang masuk pada centroid tersebut. Barang yang memiliki nilai hasil paling menjauhi angka 0 adalah

barang masuk pada centroid kedua. Proses penghitungan tersebut dilakukan pada semua data sehingga menghasilkan data kelompok. kelompok barang laku dan kelompok barang tidak laku adalah hasil akhirnya.

Dalam sistem yang telah dibuat ini penentuan centroid sudah dilakukan sistem secara acak. User hanya perlu memasukkan nama cabang dan selang waktu yang digunakan untuk perhitungan. Tampilan perhitungannya seperti pada Gambar 1.

The screenshot shows a web dashboard titled 'DASHBOARD ADMIN'. On the left, there is a navigation menu with 'Data Perjualan', 'Produk', and 'Hitung Pengelompokan'. The main content area is titled 'Tentukan Centroid Awal' and contains a form with the following fields: 'Kota' (a dropdown menu with the placeholder text '--- Silahkan Pilih Kota ---'), 'Bulan 1' (a date input field with the format 'mm/dd/yyyy'), and 'Bulan 2' (another date input field with the format 'mm/dd/yyyy'). A blue 'Proses' button is located at the bottom of the form.

Gambar 1. Tampilan Penentuan Centroid Awal

Kemudian setelah di proses, sistem akan menghasilkan tabel data. Data yang ditampilkan adalah data pengelompokan sesuai kelompok masing masing. Hasil proses program seperti pada Gambar 2.

The screenshot displays two side-by-side tables. The left table is titled '10 Produk laku' and the right table is titled '10 Produk Kurang laku'. Both tables have a search bar at the top and a table with the following columns: 'Nama Barang', 'Warna', 'Warna Kombinasi', 'Size', and 'Jenis'. The data in the tables is as follows:

Nama Barang	Warna	Warna Kombinasi	Size	Jenis
Jaket Cozmeed L Extreme Absorbenne Merah	Merah	Merah	L	extreme
Jaket Cozmeed L Extreme Argon Pro Merah	Merah	Merah	L	extreme
Jaket Cozmeed XL Extreme Absorbenne Merah	Merah	Merah	XL	extreme
Jaket Cozmeed XL Extreme Argon Pro Merah	Merah	Merah	XL	extreme
Jaket Cozmeed XL Extreme Kambora	Merah	Merah	XL	extreme
Jaket Cozmeed M Regular Denial Hitam Biru	Hitam	Biru	M	regular
Jaket Cozmeed M Regular Denial BIRU	Biru	Biru	M	regular
Jaket Cozmeed M Regular V8 Daulagat Hitam Hijau	Hitam	Hijau	M	regular
Jaket Cozmeed M Regular V8 Daulagat Hitam Biru	Hitam	Biru	M	regular
Jaket Cozmeed M Regular Aconicagua Biru	Biru	Biru	M	regular

Gambar 2. Hasil Proses Pengelompokan

Pada Gambar 2 ditunjukkan hasil proses pengelompokkan oleh sistem. Dimana terdapat 2 tabel yang menunjukkan 10 barang paling laku dan 10 barang yang tidak laku. dari hasil tersebut tindak lanjut yang dapat ditempuh adalah memproduksi kembali barang yang laku atau menyiapkan strategi marketing untuk mengatasi barang yang tidak laku.

10 Produk laku					10 Produk Kurang laku				
Nama Barang	Warna	Warna Kombinasi	Size	Jenis	Nama Barang	Warna	Warna Kombinasi	Size	Jenis
Jaket Cozmeed L Extreme Alpinextreme Merah	Merah	Merah	L	extreme	Jaket Cozmeed M Regular Denaly Hitam Biru	Hitam	Biru	M	reguler
Jaket Cozmeed L Extreme Argon Pro Merah	Merah	Merah	L	extreme	Jaket Cozmeed M Regular Denaly BIRU	Biru	Biru	M	reguler
Jaket Cozmeed XL Extreme Alpinextreme Merah	Merah	Merah	XL	extreme	Jaket Cozmeed M Regular VII Daslagit Hitam Hijau	Hitam	Hijau	M	reguler
Jaket Cozmeed XL Extreme Argon Pro Merah	Merah	Merah	XL	extreme	Jaket Cozmeed M Regular VII Daslagit Hitam Biru	Hitam	Biru	M	reguler
Jaket Cozmeed XL	Merah	Merah	XL	extreme	Jaket Cozmeed M Regular Acorncaqua	Biru	Biru	M	reguler

Gambar 4. Hasil perhitungan sistem

Pada Gambar 4 menunjukkan bahwa Jaket Extreme Alpinextreme warna merah ukuran L masuk pada kelompok barang laku dan Jaket Regular Denaly warna hitam biru dengan ukuran M masuk pada kelompok barang tidak laku. jika dilihat secara keseluruhan hasil perhitungannya sama namun hanya berbeda tampilan saja.

KESIMPULAN, DISKUSI DAN REKOMENDASI

Berdasarkan hasil penelitian dan pengolahan data yang telah dilakukan dapat diambil kesimpulan bahwa Pengelompokan data untuk menentukan barang laku dan tidak laku dapat dilaporkan sesuai dengan cabang dan rentang waktu yang diinginkan. Selain itu hasil penelitian ini menampilkan 2 tabel yaitu tabel barang laku dan tidak laku. Masing – masing tabel menampilkan 10 barang terlaris dan 10 barang dengan penjualan terendah

DAFTAR PUSTAKA

- Han, J., Kamber, M., & Pei, J. (2012). In *Data Mining Concepts and Techniques*. United State f America: Elsevier Inc.
- Hastuti, N. F. (2013). *Pemanfaatan Metode K-Means Clustering Dalam Pnentuan Penerimaan Beasiswa*. Surakarta: Universitas Sebelas Maret.
- Hitayatullah, P., & Kawistara, J. K. (2014). *Pemrograman Web*. Bandung: Informatika Bandung.
- Kusuma, V. M. (2017, Juni). Implementasi Metode Fuzzy Subtractive Clustering Untuk Pengelompokkan Data Potensi Kebakaran Hutan/Lahan. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 1, 878.
- Lutvi, & Putro, L. S. (2013). Penerapan Kombinasi Algoritma Minhash dan Binary Hamming Distance pada Hybrid Rekomendasi Lagu. *JURNAL ITSMART*, VOL 2 NO.1.

- M. Agustin, F. E., Fitria, A., & Hanifah, A. (2015). *Implementasi Algoritma K-Means Untuk Menentukan Kelompok Penggayaan Materi Mata Pelajaran Ujian Nasional (Studi Kasus : SMP Negeri 101 Jakarta)*.
- Metisen, B. M., & Sari, H. L. (2015). Analisis Clustering Menggunakan Metode K-Means Dalam Pengelompokan Penjualan Produk Pada Swalayan Fadhila. *Jurnal Media Infotama, XI*, 110-118.
- Rochmawati, Y., & Kusumaningrum, R. (2016, April). Studi Perbandingan Algoritma Pencarian String dalam Metode Approximate String Matching untuk Identifikasi Kesalahan Pengetikan Teks. *Jurnal Buana Informatika, 7*, 127.
- Sari, C. A., & D, D. S. (2014). Kluster K-Means Data Mahasiswa Baru Terhadap Program Studi Yang Dipilih.
- Sugiarti, Y. (2013). *Analisis & Perancangan UML (Unified Modeling Language) Generated VB.6*. Yogyakarta: Graha Ilmu.
- Sutrisno. (2014). *Pembuatan Aplikasi Pos (Point Of Sale) Berbasis PHP dan MySQL*. Surakarta.
- Suyatno. (2017). *Data Mining Untuk Klasifikasi dan Klasterisasi Data*. Bandung: Informatika Bandung.
- Yuniarti, K. C. (2012). *Analisis Sistem Pengendalian Internal Siklus Pendapatan (Studi Kasus pada PT Catenz Indonesia Cabang Surakarta, Jawa Tengah)*. Salatiga.