



KLASIFIKASI KEBUTUHAN PRODUK ECOPRINT MENGUNAKAN METODE CLUSTERING K-MEANS

Sakti Sudirman¹, Lestari Yusuf^{2(*)}

Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Nusa Mandiri
Jl. Jatiwaringin No. 2, Cipinang Melayu, Makasar, Jakarta Timur, Telp ((021) 28534471)
saktisudirmanbbm@gmail.com¹, lestari.lyf@nusamandiri.ac.id²

Abstrak-- Mencari tahu tingkat minat tertinggi konsumen produk ecoprint berdasarkan data daerah, data usia, data jenis kelamin, dan data jenis produk. Data jenis produk antara lain sepatu sport, sepatu boots, sepatu flat pria, sepatu flat wanita, dan tas. Dalam mencari tahu tingkat peminat tertinggi dari pembelian produk, penelitian ini menerapkan data mining dengan metode algoritma Clustering K- Means dan implementasi data menggunakan aplikasi Rapidminer. Pada hasil akhir penelitian dengan menerapkan metode algoritma Clustering K-Means, cluster dikelompokkan menjadi 3 yaitu cluster 1 berjumlah 73 data, cluster 2 berjumlah 15 data, dan cluster 3 berjumlah 12 data. Dari cluster yang didapat menghasilkan 3 kategori peminatan yaitu sangat diminati, diminati, kurang diminati. Produk yang sangat diminati konsumen terbanyak adalah sepatu sport dan sepatu flat wanita kemudian produk tas dan sepatu flat pria dan yang terendah adalah produk sepatu boots. Produk ini sangat diminati oleh kalangan perempuan dan kurang diminati oleh kalangan laki-laki. Dengan rentan usia konsumen yang sangat berminat yaitu kalangan usia 25 sampai 50 tahun, rentan usia yang kurang berminat kalangan usia 18-25 dan 50 keatas. Asal daerah konsumen yang sangat berminat yaitu daerah Pulau Jawa, Kota Jakarta, dan sekitarnya, sedangkan asal daerah konsumen yang kurang berminat yaitu Sumatera

Kata kunci: *produk ecoprint, minat, clustering k-means*

Abstract - Finding out the highest level of consumer interest in ecoprint products based on regional data, age data, gender data, and product type data. Data on product types include sports shoes, boots, men's flats, women's flats, and bags. In finding out the highest level of interest from product purchases, this study applies data mining with the K-Means Clustering algorithm and data implementation using the Rapidminer application. At the end of the research by applying the Clustering K-Means algorithm method, the clusters are grouped into 3, namely cluster 1 with 73 data, cluster 2 with 15 data, and cluster 3 with 12 data. From the clusters obtained, there are 3 categories of specialization are very attractive, desirable, and less attractive. The products that are in high demand by consumers are sport shoes and women's flat shoes, then bags and men's flat shoes and the lowest product is boots. This product is in great demand by women and less desirable by men. With a vulnerable age of very interested consumer, namely the age group of 25 to 50 years, the less interested age group is the age group of 18-25 and 50 and above. The area of origin consumers who are very interested is the area of Java Island, Jakarta City, and its surroundings, while the area of origin less interested consumer is Sumatra.

Keywords: *ecoprint product, interest, clustering k-means*

I. PENDAHULUAN

Dalam dunia bisnis yang selalu dinamis dan penuh persaingan, para pelakunya harus senantiasa memikirkan cara-cara untuk terus *survive* dan jika mungkin mengembangkan skala bisnis mereka [1].

Karena mengikuti pasar bisnis modern, gaya tekstil dan variasi yang berkembang pada produk lebih banyak bersifat *free/bebas* [2]. Salah satu industri tekstil tradisional yang mengikuti perkembangan zaman yaitu *ecoprint*. Sesuai dengan namanya, kata *eco* berasal dari kata ekosistem yang berarti alam dan

*Korepondensi.

Alamat E-mail : jurnal.larik@bsi.ac.id.

Diterima 28 Agustus 2022; Direvisi 20 Oktober 2022; Diterima 21 Desember 2022

© 2022 Jurnal Larik.

kata *print* yang berarti mencetak. Teknik pewarnaan *ecoprint* ini dipelopori pertama kali oleh *India Flint* [3]. *Ecoprint* diartikan sebagai proses mencetak warna dan bentuk daun ke kain, kulit, kertas dan keramik melalui kontak secara langsung [4]. Keunikan *ecoprint* yaitu walaupun sumber yang dipakai berasal dari jenis tumbuhan yang serupa, namun warna yang dikeluarkan belum tentu akan menghasilkan warna yang serupa [5]. Awalnya teknik *ecoprint* muncul memang dilakukan pada media kain, setelah melalui banyak eksplorasi teknik *ecoprint* mengalami banyak kemajuan yang cukup pesat sampai ketahap menggunakan kulit domba sebagai medianya. Pemilihan kulit domba sebagai dasar pembuatan *ecoprint* karena kulit domba adalah salah satu jenis kulit binatang yang mempunyai tekstur yang lembut, tipis, dan lentur sehingga sangat sering digunakan untuk produk-produk busana dan produk *fashion* seperti produk *ekoprint* [6].

Dalam rangka menghadapi persaingan bisnis di dunia *fashion* dan meningkatkan pendapatan perusahaan sekaligus meminimalisir pengeluaran, pihak terkait didalam perusahaan tersebut dituntut harus dapat mengambil keputusan yang baik dalam menentukan strategi pemasaran produk *ecoprint* yang akan dipasarkan. Dalam penelitian ini peneliti menentukan minat konsumen dilakukan berdasarkan data kota, jenis produk yang diminati, jenis kelamin, dan usia dengan menerapkan *Data Mining* pada metode *Clustering K-Means* dan diakhiri dengan mengimplementasikan data menggunakan aplikasi *Rapidminer* agar data yang dihasilkan akurat. *Clustering* merupakan proses mengelompokkan sebuah data ke dalam *cluster* berdasarkan parameter tertentu sehingga obyek-obyek dalam *cluster* memiliki tingkat kemiripan yang tinggi satu dengan yang lain dan pada objek yang lain sangat tidak mirip pada *cluster* yang berbeda [7]. Penerapan *Data Mining* dalam bisnis di perusahaan sering dilakukan untuk membuat strategi bisnis yang baik dan mendapatkan profit yang maksimal [8]. *Data mining* merupakan proses pencarian informasi/pola menarik dalam sebuah data terpilih dengan memakai metode dan teknik tertentu [9].

Proses KDD sangat berkaitan dalam teknik integrasi serta interpretasi, penemuan ilmiah serta visualisasi pola pada himpunan data-data [8]. *Knowledge Discovery in Database* adalah penemuan sebuah pengetahuan di dalam bagian integral database yang berproses pada keseluruhan untuk mengubah data mentah menjadi pengetahuan yang berguna dalam mengambil keputusan [10]. Dalam proses KDD (*knowledge discovery in database*) terdapat beberapa langkah-langkah diantaranya

selection, preprocessing, transformation, data mining, dan Interpretation/evaluation [11].

Dalam penelitian ini penggunaan aplikasi *Rapidminer* juga sangat berperan penting untuk menguji hasil yang didapat dari pengolahan dari menggunakan metode *Clustering K-Means*. *Rapidminer* merupakan suatu perangkat lunak untuk sains data yang dibentuk oleh *Rapidminer, Inc* yang menyajikan suatu *text mining*, terintegrasi data, machine learning, dan analisis prediktif, serta membantu langkah dalam suatu proses pembelajaran mesin seperti hasil optimasi, visualisasi, persiapan data, dan model tervalidasi [12]. *Rapidminer* sangat *compatible* dengan beberapa sumber data, lingkungan dan memungkinkan pengguna untuk menerapkan model yang mereka rancang pada perangkat keras perusahaan apapun [13].

Penelitian ini bertujuan untuk menyelesaikan persoalan yang ada di perusahaan yaitu untuk mengetahui tingkat minat belanja konsumen terhadap produk *ecoprint* agar perusahaan bisa mengambil keputusan untuk menentukan strategi terbaik dalam produksi dan pemasarannya. Dalam produksi perusahaan dapat memperbanyak produksi produk yang paling diminati para konsumen dan mengurangi produksi produk yang sedikit diminati para konsumen tujuannya untuk menekan biaya produksi agar lebih terarah dan meminimalisir pengeluaran yang sia-sia. Dalam pemasaran perusahaan dapat mengetahui target pasar dari segi wilayah, umur, dan jenis kelamin tujuannya untuk memfokuskan pemasaran agar lebih terarah dan mendapatkan omset lebih baik.

II. METODE PENELITIAN

A. Pengumpulan Data

1. Wawancara

Wawancara dilakukan untuk mendapatkan permasalahan yang terjadi. Dikarenakan adanya kelebihan biaya produksi yang akan berdampak kepada biaya produksi, produk yang terlalu lama disimpan, kebutuhan ruang penyimpanan yang meningkat dikarenakan tidak memperhatikan permintaan pasar sehingga ditemukan permasalahan kelebihan memproduksi suatu barang yang menimbulkan kerugian pada pemilik *Ecoprint*. Maka wawancara dilakukan kepada pemilik untuk mengetahui keinginan dan kebutuhan pemilik.

2. Observasi

Pengamatan yang dilakukan pada proses pengklasifikasian produk ini langsung pada perusahaan yaitu proses pemesanan produk sampai

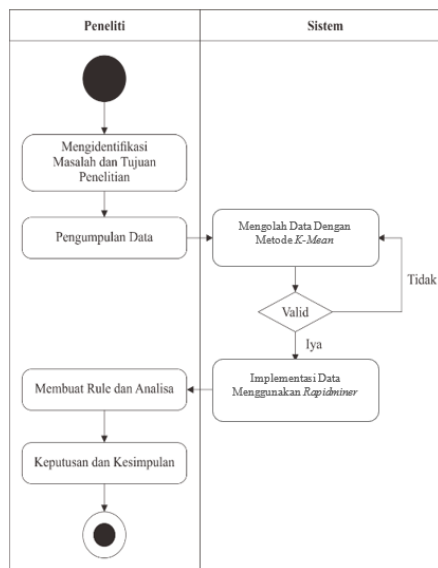
transaksi. Dan mendapatkan sumber data yaitu data penjualan produk Ecoprint Agustus 2020-Mei 2022.

3. Studi pustaka

Dalam metode ini dilakukan dalam media internet dalam pencarian bahan-bahan penelitian yang berada di Internet, buku-buku dan juga jurnal sebagai referensi dalam penulisan ini.

B. Metode penelitian

Alur kerja penelitian yang penulis lakukan dapat dilihat pada gambar 1 dibawah ini:



Gambar 1. Alur kerja penelitian

Tahapan penelitian pertama dimulai dari mengidentifikasi masalah pada perusahaan melalui proses wawancara. Dalam penelitian ini penulis melakukan pencarian data secara langsung melalui tanya jawab kepada narasumber yaitu owner. Adapun hal yang ditanyakan yaitu berkaitan dengan masalah yang terjadi pada perusahaan selama ini. Setelah selesai wawancara penulis menyimpulkan permasalahan pada perusahaann yaitu pemborosan produksi secara berlebihan.

Tahap kedua yaitu tahap pengumpulan data yang diambil langsung dari perusahaan Ecoprint. Data yang digunakan dalam penelitian ini merupakan data yang diambil dari beberapa proses pengumpulan data konsumen dan data penjualan produk ecoprint dari Agustus 2020 sampai dengan Mei 2022 yang berjumlah 500 data berupa format .xlsx. Data yang dikumpulkan antara lain data daerah konsumen yang diambil berdasarkan data pengiriman produk, data usia dan jenis kelamin konsumen diambil berdasarkan perkiraan melalui wawancara terhadap owner, dan data jenis-jenis

produk yang terjual diambil dari data penjualan perusahaan. Dalam penelitian ini penulis juga mengumpulkan data dengan memanfaatkan media internet dalam rangka mencari data-data pendukung. Buku dan jurnal adalah contoh data yang dikumpulkan melalui media internet yang sangat berguna untuk refrensi penulis pada penelitian ini.

Tahap ketiga yaitu transformasi data, data yang berjenis alfabet seperti jenis kelamin, kota konsumen harus dilakukan proses inisialisasi data terlebih dahulu kedalam bentuk angka/numerik karena data tidak dibentuk *numerik*. Data harus ditransformasikan terlebih dahulu dengan cara melakukan frekuensi pada data yang terbanyak muncul dengan mengurutkan frekuensi tertinggi ke terendah dan melakukan inisial data.

Tahap keempat yaitu menerapkan metode *Clustering K-means* untuk mengelompokkan data yang sudah dikumpulkan. Pertama-tama menentukan jumlah *cluster* terbaik yang akan digunakan sebagai perhitungan data dengan metode *Clustering K-means*. Selanjutnya menentukan *centroid* secara *random* atau acak. Menghitung jarak dari *centroid*, caranya menghitung jarak dari data ke-1 terhadap pusat *cluster* dengan rumus *euclidian distance* [10]. Selanjutnya alokasikan masing-masing objek ke *centroid* terdekat, untuk melakukan alokasi objek ke dalam masing-masing *cluster* dengan cara mengelompokkan berdasarkan jarak minimum objek ke pusat *cluster* dengan memberikan kode "1" jika jarak terpendek di *cluster* 1, memberikan kode "2" jika jarak terpendek di *cluster* 2, dan memberikan kode "3" jika jarak terpendek di *cluster* 3. Perhitungan dilakukan terus sampai ke-500 data terhadap pusat *cluster*. Langkah-langkah ini akan diulang ke iterasi selanjutnya sampai data tidak mengalami perubahan dan setelah data tidak mengalami perubahan perulangan akan diberhentikan karena sudah mendapatkan hasil yang dicari.

Tahap kelima yaitu hasil dari perhitungan tersebut akan diuji atau di implementasikan kedalam aplikasi *Rapidminer*. Setelah data berhasil melewati sistem uji di *Rapidminer* penulis akan membuat rule dan menganalisanya. Selanjutnya penulis akan membuat rule dan analisa. Tahap terakhir penulis akan memberikan keputusan dan kesimpulan dari hasil penelitian.

III. HASIL DAN PEMBAHASAN

Tabel 1. Data Minat Belanja Konsumen pada Produk Ecoprint

| No. | Kode | L/P | Usia | Daerah Konsumen | Jenis Produk Yang Diminati | | | | |
|------|-------|------|-------|----------------------|----------------------------|--------------|------------------|--------------------|------|
| | | | | | Sepatu Sport | Sepatu Boots | Sepatu Flat Pria | Sepatu Flat Wanita | Tas |
| 1. | K-001 | P | 18-25 | Jakarta & sekitarnya | 1 | 0 | 0 | 0 | 0 |
| 2. | K-002 | L | 35-50 | Jakarta & sekitarnya | 1 | 0 | 0 | 0 | 0 |
| 3. | K-003 | L | 35-50 | Pulau Jawa | 0 | 0 | 1 | 0 | 0 |
| 4. | K-004 | P | 35-50 | Pulau Jawa | 0 | 0 | 0 | 1 | 0 |
| 5. | K-005 | L | 25-35 | Pulau Sumatera | 1 | 0 | 0 | 0 | 0 |
| 6. | K-006 | P | 18-25 | Pulau Sumatera | 1 | 0 | 0 | 0 | 0 |
| 7. | K-007 | L | 35-50 | Pulau Jawa | 1 | 0 | 0 | 0 | 1 |
| 8. | K-008 | P | 35-50 | Pulau Jawa | 1 | 0 | 0 | 1 | 0 |
| 9. | K-009 | P | 25-35 | Jakarta & sekitarnya | 0 | 0 | 0 | 1 | 0 |
| 10. | K-010 | P | 35-50 | Pulau Jawa | 0 | 1 | 0 | 0 | 0 |
| | | | | | | | | | |
| 497. | K-497 | L | 25-35 | Jakarta & sekitarnya | 1 | 1 | 0 | 1 | 0 |
| 498. | K-498 | P | 35-50 | Pulau Jawa | 1 | 0 | 1 | 0 | 0 |
| 499. | K-499 | P | 35-50 | Pulau Sumatera | 1 | 0 | 0 | 1 | 0 |
| 500. | K-500 | P | 18-25 | Jakarta & sekitarnya | 1 | 1 | 0 | 1 | 0 |
| | | | | Jumlah | 326 | 68 | 78 | 205 | 114 |

Tabel 1 merupakan data sample yang digunakan untuk menentukan tingkat kebutuhan konsumen terhadap produk Ecoprint yang telah dikumpulkan.

Langkah pertama data akan ditransformasikan karena ada beberapa data tidak bernilai *numeric*. Caranya adalah melakukan frekuensi dari data yang terbanyak ke terendah dan lakukan inisialisasi.

Tabel 1. Inisialisasi Jenis Kelamin

| P/L | Frekuensi | Inisial |
|---------------|-----------|---------|
| Perempuan (P) | 359 | 1 |
| Laki-Laki (L) | 141 | 2 |

Tabel 2 menjelaskan inisialisasi jenis kelamin konsumen. Untuk jenis kelamin perempuan diberikan inisial 1 dan laki-laki 2.

Tabel 3. Inisialisasi Usia

| Usia | Frekuensi | Inisial |
|-------------|-----------|---------|
| 35-50 tahun | 212 | 1 |
| 25-35 tahun | 169 | 2 |
| 18-25 tahun | 80 | 3 |
| 50+ tahun | 39 | 4 |

Tabel 3 merupakan inisialisasi Usia konsumen dengan usia paling muda mulai dari 18 tahun sampai usia 50 tahun keatas.

Tabel 2. Inisialisasi Daerah Konsumen

| Daerah Konsumen | Frekuensi | Inisial |
|----------------------|-----------|---------|
| Wil. Jawa | 230 | 1 |
| Jakarta & Sekitarnya | 178 | 2 |
| Wil. Sumatera | 92 | 3 |

Tabel 4 berisikan inisiasi daerah para konsumen yang memesan produk-produk Ecoprint. Dilihat dari dokumen penjualan daerah penjualan pulau jawa diberi inisialisasi no 1 dikarenakan jumlah frekuensi penjualannya. Dilanjutkan jakarta dan sekitarnya yang di inisialisasi no 2 dan pulau Sumatera nomer 3.

Tabel 3. Inisialisasi Jenis Produk

| Jenis Produk | Frekuensi | Inisial |
|--------------------|-----------|---------|
| Sepatu Sport | 326 | 1 |
| Sepatu Flat Wanita | 205 | 2 |
| Tas | 114 | 3 |
| Sepatu Flat Pria | 78 | 4 |
| Sepatu Boots | 68 | 5 |

Pada tabel 5 merupakan penjelasan inisialisasi jenis produk dengan frekuensi yang sering diberi nilai 1 dan berlanjut ke nomer berikutnya sampai ke nomer 5 yang diberikan untuk mewakili sepatu boots.

Langkah kedua setelah semua data ditransformasikan ke dalam bentuk angka, maka data-data tersebut telah dapat dikelompokkan dengan menggunakan algoritma *Clustering K-Means* sampai mendapatkan *cluster* terbaik. Pertama-tama menentukan jumlah *cluster* yang diinginkan. Dalam penelitian ini data-data yang akan digunakan sebanyak 3 *cluster* ($k=3$). Setelah itu menentukan titik pusat awal dari setiap *cluster* atau *centroid*. Dalam penelitian ini centroid ditentukan secara *random*, dimana nilai *cluster* 1 diambil dari baris ke-14, nilai *cluster* 2 pada baris ke-26, dan *cluster* 3 pada baris ke-28. Keterangan centroid di jelaskan pada tabel 6.

Tabel 4. Titik Pusat Awal Cluster

| No. | L/P | Usia | Daerah Konsumen | Jenis Produk Yang Diminati | | | | |
|-----|-----|------|-----------------|----------------------------|--------------|------------------|--------------------|-----|
| | | | | Sepatu Sport | Sepatu Boots | Sepatu Flat Pria | Sepatu Flat Wanita | Tas |
| 14. | 1 | 2 | 1 | 1 | 0 | 0 | 2 | 0 |
| 26. | 2 | 1 | 3 | 1 | 0 | 4 | 0 | 0 |
| 28. | 1 | 1 | 2 | 0 | 5 | 0 | 0 | 3 |

Langkah ketiga menghitung jarak *centroid* dengan cara menghitung jarak dari data ke-1 sampai

data ke-500 terhadap 3 pusat *cluster*. Perhitungan ini dilakukan dengan rumus *euclidian distance*.

Rumus *Euclidian Distance* :

$$dist(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Perhitungan jarak data pertama dengan pusat *cluster* pertama adalah :

$$\sqrt{(1-1)^2 + (3-2)^2 + (2-1)^2 + (1-1)^2 + (0-0)^2 + (0-0)^2 + (0-2)^2 + (0-0)^2} = 2,45$$

Perhitungan jarak data pertama dengan pusat *cluster* kedua adalah :

$$\sqrt{(1-1)^2 + (3-2)^2 + (2-1)^2 + (1-1)^2 + (0-0)^2 + (0-0)^2 + (0-2)^2 + (0-0)^2} = 4,69$$

Perhitungan jarak data pertama dengan pusat *cluster* ketiga adalah :

$$\sqrt{(1-1)^2 + (3-2)^2 + (2-1)^2 + (1-1)^2 + (0-0)^2 + (0-0)^2 + (0-2)^2 + (0-0)^2} = 6,24$$

Perhitungan diatas menghasilkan tabel cluster baru. Selanjutnya peneliti akan menampilkan data pengelompokan *group* iterasi ke-1, dengan penjelasan jika nilai 1 berarti data tersebut berada didalam kelompok *cluster* tersebut seperti yang digambarkan pada tabel 7 dibawah ini:

Table 7. Pengkelompokan Group Iterasi ke-1

| C1 | C2 | C3 |
|----|----|----|
| 1 | | |
| 1 | | |
| | 1 | |
| 1 | | |
| 1 | | |
| 1 | | |
| 1 | | |
| 1 | | |
| 1 | | |
| 1 | | |
| | | 1 |
| 1 | | |
| 1 | | |
| 1 | | |
| 1 | | |
| | 1 | |
| 1 | | |

| | | |
|---|---|---|
| 1 | | |
| 1 | | |
| 1 | | |
| 1 | | |
| 1 | | |
| 1 | | |
| | | 1 |
| 1 | | |
| | 1 | |
| 1 | | |
| | | 1 |

Langkah keempat alokasikan masing-masing objek ke *centroid* terdekat, untuk melakukan alokasi objek ke dalam masing-masing *cluster* dengan cara mengelompokkan berdasarkan jarak minimum objek ke pusat *cluster* dengan memberikan kode "1" jika jarak terpendek di *cluster* 1, memberikan kode "2" jika jarak terpendek di *cluster* 2, dan memberikan kode "3" jika jarak terpendek di *cluster* 3. Perhitungan dilakukan terus sampai ke-500 data terhadap pusat *cluster*. Tetapi disini hanya dilampirkan 28 data saja secara urut.

Tabel 8. Data Iterasi ke-1

| Data ke-i | c1 | c2 | c3 | Jarak Terpendek |
|-----------|------|------|------|-----------------|
| 1 | 2,45 | 4,69 | 6,24 | 1 |
| 2 | 2,65 | 4,12 | 6 | 1 |
| 3 | 4,8 | 2,24 | 7,21 | 2 |
| 4 | 1,41 | 5,1 | 6,24 | 1 |
| 5 | 3 | 4,12 | 6,16 | 1 |
| 6 | 3 | 4,58 | 6,32 | 1 |
| 7 | 3,87 | 5,39 | 5,29 | 1 |
| 8 | 1 | 5 | 6,32 | 1 |
| 9 | 1,41 | 4,9 | 6,24 | 1 |
| 10 | 5,57 | 6,86 | 3,16 | 3 |
| 11 | 2,24 | 4,36 | 6 | 1 |
| 12 | 2,24 | 5 | 6,32 | 1 |
| 13 | 3,74 | 5,66 | 5,57 | 1 |
| 14 | 0 | 5,1 | 6,4 | 1 |
| 15 | 4,9 | 2,45 | 7,42 | 2 |
| 16 | 3,87 | 5,39 | 5,29 | 1 |
| 17 | 2,45 | 4,69 | 6,24 | 1 |

| | | | | |
|----|------|------|------|---|
| 18 | 2,45 | 4,24 | 5,92 | 1 |
| 19 | 2,83 | 4,24 | 6,08 | 1 |
| 20 | 3,74 | 5,83 | 5,57 | 1 |
| 21 | 3,32 | 6,24 | 5,83 | 1 |
| 22 | 2,45 | 4,24 | 5,92 | 1 |
| 23 | 2,65 | 4,12 | 6 | 1 |
| 24 | 5,2 | 7,42 | 4,24 | 3 |
| 25 | 2,45 | 4,9 | 6,4 | 1 |
| 26 | 5,1 | 0 | 7,28 | 2 |
| 27 | 3 | 4,12 | 6 | 1 |
| 28 | 6,4 | 7,28 | 0 | 3 |

Jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat *cluster* terdekat. Dengan cara membandingkan hasil *cluster* dan diambil yang paling kecil. Setelah itu menghitung *cluster* baru untuk mendapatkan *cluster* ke-2 dengan rumus jumlahkan semua anggota masing-masing *cluster* lalu bagi dengan jumlah anggota kelompok.

Tabel 9. Hasil Iterasi ke-2

| Data ke-i | c1 | c2 | c3 | cluster |
|-----------|------|------|------|---------|
| 1 | 3,79 | 4,15 | 6,3 | 1 |
| 2 | 3,79 | 4,17 | 6,28 | 1 |
| 3 | 5,64 | 1,73 | 7,52 | 2 |
| 4 | 3,37 | 4,8 | 5,99 | 1 |
| 5 | 3,89 | 4,2 | 6,38 | 1 |
| 6 | 3,99 | 4,31 | 6,44 | 1 |
| 7 | 2,42 | 5,2 | 5,53 | 1 |
| 8 | 3,22 | 4,7 | 5,91 | 1 |

| | | | | |
|----|------|------|------|---|
| 9 | 3,19 | 4,62 | 5,94 | 1 |
| 10 | 6,35 | 6,64 | 3,85 | 3 |
| 11 | 3,63 | 4,04 | 6,18 | 1 |
| 12 | 3,85 | 4,22 | 6,31 | 1 |
| 13 | 1,64 | 5,63 | 5,3 | 1 |
| 14 | 3,1 | 4,58 | 5,87 | 1 |
| 15 | 5,65 | 1,46 | 7,58 | 2 |
| 16 | 2,42 | 5,2 | 5,53 | 1 |
| 17 | 3,79 | 4,15 | 6,3 | 1 |
| 18 | 3,73 | 4,18 | 6,22 | 1 |
| 19 | 3,83 | 4,21 | 6,32 | 1 |
| 20 | 2,42 | 5,18 | 5,55 | 1 |
| 21 | 1,69 | 5,64 | 5,28 | 1 |
| 22 | 3,73 | 4,18 | 6,22 | 1 |
| 23 | 3,79 | 4,17 | 6,28 | 1 |
| 24 | 6,07 | 6,91 | 3,45 | 3 |
| 25 | 3,92 | 4,21 | 6,37 | 1 |
| 26 | 5,65 | 1,67 | 7,57 | 2 |
| 27 | 3,93 | 4,34 | 6,36 | 1 |
| 28 | 5,56 | 7,24 | 2,38 | 3 |

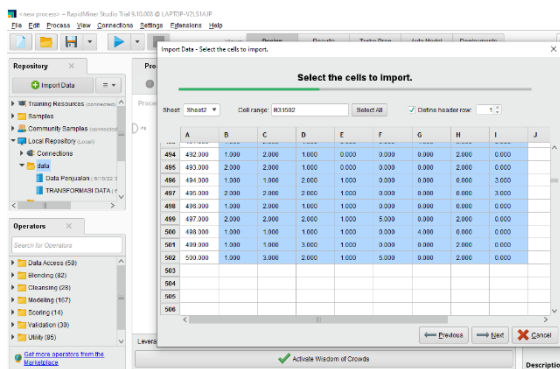
Pada perhitungan ini iterasi berhenti pada iterasi ke-2 karena kelompok data sudah sama atau tidak mengalami perubahan. Perhitungan data menggunakan metode *Clustering K-Means* pun dinyatakan selesai dari 500 data mendapatkan hasil *cluster* 1 berjumlah 354 data, *cluster* 2 berjumlah 78 data, dan *cluster* 3 berjumlah 68 data. Hasil akhir clustering dijelaskan pada tabel 10.

Tabel 10. Hasil Akhir Cluster

| Cluster | Data Ke-i | L/P | Usia | Daerah | Jenis Produk Yang Diminati | | | | |
|---------|-----------|-----|------|--------|----------------------------|--------------|------------------|--------------------|-----|
| | | | | | Sepatu Sport | Sepatu Boots | Sepatu Flat Pria | Sepatu Flat Wanita | Tas |
| C1 | 1 | 1 | 3 | 2 | 1 | 0 | 0 | 0 | 0 |
| C1 | 2 | 2 | 1 | 2 | 1 | 0 | 0 | 0 | 0 |
| C1 | 4 | 1 | 1 | 1 | 0 | 0 | 0 | 2 | 0 |
| C1 | 5 | 2 | 2 | 3 | 1 | 0 | 0 | 0 | 0 |
| C1 | 6 | 1 | 3 | 3 | 1 | 0 | 0 | 0 | 0 |
| C1 | 7 | 2 | 1 | 1 | 1 | 0 | 0 | 0 | 3 |
| C1 | 8 | 1 | 1 | 1 | 1 | 0 | 0 | 2 | 0 |
| C1 | 9 | 1 | 2 | 2 | 0 | 0 | 0 | 2 | 0 |
| C1 | 11 | 1 | 2 | 2 | 1 | 0 | 0 | 0 | 0 |

| | | | | | | | | | |
|----|----|---|---|---|---|---|---|---|---|
| C1 | 12 | 1 | 3 | 1 | 1 | 0 | 0 | 0 | 0 |
| C1 | 13 | 1 | 2 | 3 | 0 | 0 | 0 | 2 | 3 |
| C1 | 14 | 1 | 2 | 1 | 1 | 0 | 0 | 2 | 0 |
| C1 | 16 | 2 | 1 | 1 | 1 | 0 | 0 | 0 | 3 |
| C1 | 17 | 1 | 3 | 2 | 1 | 0 | 0 | 0 | 0 |
| C1 | 18 | 1 | 1 | 2 | 1 | 0 | 0 | 0 | 0 |
| C1 | 19 | 1 | 2 | 3 | 1 | 0 | 0 | 0 | 0 |
| C1 | 20 | 1 | 3 | 1 | 1 | 0 | 0 | 0 | 3 |
| C1 | 21 | 1 | 3 | 1 | 0 | 0 | 0 | 2 | 3 |
| C1 | 22 | 1 | 1 | 2 | 1 | 0 | 0 | 0 | 0 |
| C1 | 23 | 2 | 1 | 2 | 1 | 0 | 0 | 0 | 0 |
| C1 | 25 | 2 | 3 | 1 | 1 | 0 | 0 | 0 | 0 |
| C1 | 27 | 1 | 1 | 3 | 1 | 0 | 0 | 0 | 0 |
| C2 | 3 | 2 | 1 | 1 | 0 | 0 | 4 | 0 | 0 |
| C2 | 15 | 2 | 3 | 2 | 0 | 0 | 4 | 0 | 0 |
| C2 | 26 | 2 | 1 | 3 | 1 | 0 | 4 | 0 | 0 |
| C3 | 10 | 1 | 1 | 1 | 0 | 5 | 0 | 0 | 0 |
| C3 | 24 | 1 | 3 | 1 | 0 | 5 | 0 | 2 | 0 |
| C3 | 28 | 1 | 1 | 2 | 0 | 5 | 0 | 0 | 3 |

Setelah data berhasil diolah menggunakan metode *Clustering K-Means* dengan cara manual dan dengan bantuan Microsoft Excel. Langkah terakhir adalah untuk menguji hasil data agar lebih akurat dilakukanlah implementasi data pada aplikasi *Rapidminer* versi 9.10 menggunakan metode *Clustering K-Means* yang hasilnya akan membantu perusahaan untuk memudahkan dalam proses pengambilan keputusan sebagai penentu minat konsumen terhadap produk *ecoprint*. Dalam penelitian ini *rapidminer* yang digunakan adalah versi 9.0, setelah membuka aplikasi *rapidminer*, import data dalam format *.xlsx* yang berjumlah sebanyak 500 data seperti pada gambar 2.



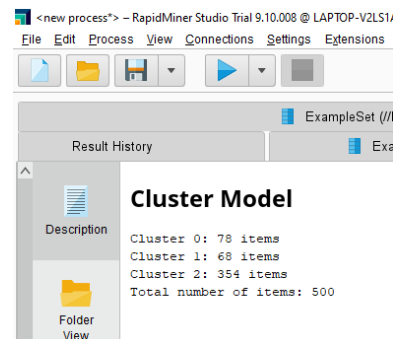
Gambar 1. Penginputan data pada rapidminer 9.0

Hasil centroid dijelaskan pada gambar 3 .

| Attribute | cluster_0 | cluster_1 | cluster_2 |
|-------------|-----------|-----------|-----------|
| LIP | 1.526 | 1.119 | 1.260 |
| Usia | 2.064 | 1.750 | 1.881 |
| Daerah | 1.808 | 1.588 | 1.732 |
| Sport | 0.487 | 0.412 | 0.734 |
| Ebooks | 0 | 5 | 0 |
| Flat Pria | 4 | 0 | 0 |
| Flat Wanita | 0 | 1.191 | 0.927 |
| Tias | 0 | 1.015 | 0.171 |

Gambar 3. Hasil Centroid Table pada Cluster Model

Dari data hasil *Cluster Model (Clustering)* dapat diketahui proses perhitungan manual dan implementasi dengan *Rapidminer* maka didapatkan hasil yang sama. Hasil akurasi dari proses Implementasi *RapidMiner* menggunakan algoritma *Clustering K-Means*.



Gambar 4. Tampilan Description Pada Cluster Model

Pada *Cluster Model (Clustering)* ini dapat dilihat keseluruhan data berjumlah 500 data yang dikelompokkan menjadi 3 *cluster* yaitu *cluster* 0 sebanyak 78 *items*, *cluster* 1 sebanyak 68 *items* dan *cluster* 2 sebanyak 354 *items*. Jumlah *cluster* yang didapat dari pengolahan data menggunakan aplikasi *Rapidminer* sama seperti pengolahan data dengan cara manual.

IV. SIMPULAN DAN SARAN

Dari permasalahan tentang data penjualan produk *ecoprint* pada minat konsumen berhasil diselesaikan dengan metode Algoritma *Clustering K-Means* dan aplikasi *Rapidminer* dilakukan dengan mengelompokkan data dalam jumlah yang banyak. Dilakukan dengan cara menghitung jarak dari data ke-1 sampai data ke-500 terhadap pusat *cluster*. Setelah itu dilakukan perbandingan dan memilih jarak terdekat antara data dengan pusat *cluster*, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat *cluster* terdekat. Tahap terakhir data akan diimplementasikan pada aplikasi *Rapidminer* untuk mengetahui bahwa hasil *cluster* yang didapat benar-bener akurat. Hasil *cluster* inilah yang sangat berpengaruh pada keputusan yang akan diambil pada perusahaan ini.

Pada hasil akhir penelitian dengan menerapkan metode algoritma *Clustering K-Means*, *cluster* dikelompokkan menjadi 3 yaitu *cluster* 1 berjumlah 354 data, *cluster* 2 berjumlah 78 data, dan *cluster* 3 berjumlah 68 data. Dari *cluster* yang didapat menghasilkan 3 kategori peminatan yaitu sangat diminati, diminati, kurang diminati. produk yang sangat diminati konsumen terbanyak adalah sepatu sport dan sepatu flat wanita kemudian produk tas dan sepatu flat pria dan yang terendah adalah produk sepatu boots. Produk ini sangat diminati oleh kalangan perempuan dan kurang diminati oleh kalangan laki-laki. Dengan rentan usia konsumen yang sangat berminat yaitu kalangan usia 25 sampai 50 tahun, rentan usia yang kurang berminat kalangan usia 18-25 dan 50 keatas. Asal daerah konsumen yang sangat berminat yaitu daerah Jawa, Kota Jakarta, dan sekitarnya, sedangkan asal daerah konsumen yang kurang berminat yaitu Sumatera.

Dalam produksi, perusahaan dapat memperbanyak produk yang paling diminati konsumen yaitu sepatu sport dan sepatu flat wanita dan mengurangi produksi yang sedikit diminati konsumen seperti sepatu boots tujuannya untuk menekan biaya produksi agar lebih terarah dan meminimalisir pengeluaran yang sia-sia. Dalam pemasaran perusahaan peneliti mendapatkan target pasar terbaik yaitu dari segi wilayah Pulau Jawa, Jakarta dan sekitarnya, rentan umur usia 25-50

tahun, dan jenis kelamin perempuan. Tujuannya untuk memfokuskan pemasaran agar lebih terarah dan mendapatkan omset yang maksimal.

V. REFERENSI

- [1] R. Sari, "Enur Irdiansyah PENERAPAN DATA MINING PADA PENJUALAN PRODUK MINUMAN DI PT . PEPSI COLA INDOBEVERAGES MENGGUNAKAN ..."
- [2] A. M. Dixit *et al.*, "No 主観的健康感を中心とした在宅高齢者における健康関連指標に関する共分散構造分析Title," *Anal. Biochem.*, vol. 11, no. 1, pp. 1–5, 2018.
- [3] E. Saptutyingsih and B. P. Kamiel, "Pemanfaatan Bahan Alami untuk Pengembangan Ecoprint dalam Mendukung Ekonomi Kreatif," in *Prosiding Seminar Nasional Unimus*, 2019, vol. 2.
- [4] S. Ristiani, "EKSPLOKASI TEKNIK ECOPRINT PADA MEDIA KULIT / EXPLORATION OF ECOPRINT TEQNIQUES ON LEATHER MEDIA," no. 7.
- [5] F. Yanti and E. Br, "Febry yanti elmida br. giting 161201035," 2021.
- [6] A. Budiawan,S.E, *The Secret of Ecoprint on Leather by Ecoprint Bali*. 2021.
- [7] L. Iryani, "Penerapan Data Mining Menentukan Minat Baca Mahasiswa Di Perpustakaan Universitas Bina Darma Palembang Menggunakan Metode Clustering," *INTECOMS J. Inf. Technol. Comput. Sci.*, vol. 3, no. 1, pp. 82–89, 2020.
- [8] A. Nazaruddin, *Data Mining dan Penerapan Algoritma*. 2021.
- [9] A. Asroni, H. Fitri, and E. Prasetyo, "Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokan Data Calon Mahasiswa Baru di Universitas Muhammadiyah Yogyakarta (Studi Kasus: Fakultas Kedokteran dan Ilmu Kesehatan, dan Fakultas Ilmu Sosial dan Ilmu Politik)," *Semesta Tek.*, vol. 21, no. 1, pp. 60–64, 2018, doi: 10.18196/st.211211.
- [10] S. Anita; Muharmi, Yulya; Tria, "Penentuan Tingkat Minat Belanja Online Melalui Media Sosial Menggunakan Metode Clustering K-Means," vol. I, 2018.
- [11] Alfira, "PENERAPAN MARKET

- BASKET ANALYSIS MENGGUNAKAN PROSES KDD (KNOWLEDGE DISCOVERY IN DATABASE) SEBAGAI STRATEGI PENJUALAN PRODUK SWALAYAN (STUDI KASUS: SWALAYAN X),” pp. 509–516, 2018.
- [12] M. Hasanah, *Implementasi Barcode Dan Algoritma Regresi Linear Untuk Memprediksi Data Persediaan Barang*. 2020.
- [13] S. Kumari, *Data Science and Analytics*. Emerald Publishing Limited, 2020