

Evaluation of Product Sales Data Using Clustering Method and Hierarchical Divisive Clustering at PT.AYN

Ahir Yugo Nugroho Harahap^{1*}, Ria Eka Sari², Heri Gunawan³, Adnan Buyung Nasution⁴

^{1,2,3}Universitas Potensi Utama Medan, Indonesia

⁴Universitas Borobudur

Corresponding Author: Ahir Yugo Nugroho Harahap
ahiryugo.potensi@gmail.com

ARTICLE INFO

Keywords : Data Mining, Sales, Hierarchical Divisive Clustering, Purchasing Trends

Received : 05 June

Revised : 20 June

Accepted: 23 July

©2024 Harahap, Sari, Gunawan, Nasution: This is an open-access article distributed under the terms of the [Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



ABSTRACT

Data Mining, focusing on the Hierarchical Divisive algorithm, can provide a solution for PT.AYN in overcoming the problem of unifying and evaluating sales data, a company that sells various types of disposable tissues. This study aims to identify products that are in demand and less in demand and to group sales data based on company and product type. The results of this study provide valuable insights for evaluating sales data, understanding distributor purchasing trends, and supporting more effective stock planning, shipping, and marketing strategies. Through the application of the clustering method and the Hierarchical Divisive algorithm, this study offers an effective solution to optimize the use of sales data at PT.AYN, which has the potential to be a valuable asset in formulating long-term business strategies. The conclusion of this study is that the Clustering method and the Hierarchical Divisive algorithm can be used to solve problems in grouping sales data based on various factors, including region and product type, and can assist in developing better marketing strategies.

Evaluasi Data Penjualan Produk Menggunakan Metode Clustering Dan Hierarki Divisive Clustering Di PT.AYN

Ahir Yugo Nugroho Harahap^{1*}, Ria Eka Sari², Heri Gunawan³, Adnan Buyung Nasution⁴

^{1,2,3}Universitas Potensi Utama Medan, Indonesia

⁴Universitas Borobudur

Corresponding Author: Ahir Yugo Nugroho Harahap

ahiryugo.potensi@gmail.com

ARTICLE INFO

Kata Kunci: Data Mining, Penjualan, Hierarki Divisive Clustering, Tren Pembelian

Received : 05 Juni

Revised : 20 Juni

Accepted: 23 Juli

©2024 Harahap, Sari, Gunawan, Nasution: This is an open-access article distributed under the terms of the [Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



ABSTRAK

Data Mining, dengan fokus pada algoritma Hirarki Divisive, dapat memberikan solusi bagi PT.AYN dalam mengatasi masalah penyatuan dan evaluasi data penjualan, sebuah perusahaan yang menjual berbagai jenis Tissue sekali pakai. Penelitian ini bertujuan untuk mengidentifikasi produk yang diminati dan kurang diminati serta mengelompokkan data penjualan berdasarkan perusahaan dan jenis produk. Hasil penelitian ini memberikan wawasan berharga untuk evaluasi data penjualan, pemahaman tren pembelian distributor, serta mendukung perencanaan stok, pengiriman, dan strategi pemasaran yang lebih efektif. Melalui penerapan metode clustering dan algoritma Hierarchical Divisive, penelitian ini menawarkan solusi yang efektif untuk mengoptimalkan penggunaan data penjualan di PT.AYN, yang berpotensi menjadi aset berharga dalam perumusan strategi bisnis jangka panjang. Kesimpulan dari penelitian ini adalah bahwa metode Clustering dan algoritma Hierarchical Divisive dapat digunakan untuk menyelesaikan permasalahan dalam pengelompokkan data penjualan berdasarkan berbagai faktor, termasuk daerah dan jenis produk, serta dapat membantu dalam pengembangan strategi pemasaran yang lebih baik.

PENDAHULUAN

Dalam era digital saat ini, perusahaan di berbagai industri menghadapi tumpukan data yang terus tumbuh. Data penjualan adalah aset berharga yang dapat memberikan wawasan penting untuk pengambilan keputusan strategis. Namun, banyak perusahaan masih menghadapi tantangan dalam pengelolaan dan evaluasi data penjualan mereka. Hal ini juga berlaku untuk PT.AYN, perusahaan yang telah beroperasi sejak September 2005, dan berfokus pada penjualan Tissue sekali pakai dengan berbagai tipe di Medan.

Saat ini, data penjualan di PT.AYN dikelola dengan cara yang kurang efisien, menggunakan Microsoft Excel, dan seringkali data tersebut hanya dilihat sebagai berkas-berkas yang tidak terpakai. Kondisi ini menyebabkan informasi berharga yang terkandung dalam data penjualan tidak dimaksimalkan. Selain itu, pengelolaan data yang belum terintegrasi juga mempersulit manajemen perusahaan dalam mengakses dan menganalisis data penjualan secara komprehensif. Untuk mengatasi permasalahan ini, penelitian ini akan menjelaskan penerapan teknik Data Mining, dengan fokus pada metode Clustering menggunakan algoritma Hirarki Divisive. Melalui penggunaan algoritma ini, data penjualan akan diklasifikasikan dan dikelompokkan berdasarkan karakteristik tertentu, sehingga perusahaan dapat mengidentifikasi produk yang diminati dan yang kurang diminati. Data yang terintegrasi ini akan memberikan pemahaman yang lebih baik tentang tren pembelian distributor tertentu terhadap produk tertentu, dan akan menjadi landasan untuk perencanaan stok, pengiriman, dan strategi pemasaran di masa depan.

Pengelompokan, juga dikenal sebagai pengelompokan data, melibatkan pengelompokan data dengan nilai-nilai yang mirip (homogen). Data yang dapat dikelompokkan meliputi hasil pengamatan, catatan data, atau kelas dan objek dengan kesamaan. Dalam pengelompokan, berbeda dengan klasifikasi karena tidak menggunakan variabel keputusan/target, seperti tinggi, sedang, dan rendah. Sebagai contoh, pengelompokan keluarga yang layak dan tidak layak untuk Program Keluarga Harapan (PKH) berdasarkan pendapatan, jumlah tanggungan, aset, pekerjaan, kepemilikan, dan kondisi rumah [2]. Sementara itu, Analisis Hirarki Divisive adalah metode pengelompokan hierarki yang memecah data dengan strategi top-down. Ia dimulai dengan menempatkan semua objek dalam satu kelompok besar, yang berfungsi sebagai akar hierarki. Kemudian, ia membagi kelompok akar menjadi kelompok sub yang lebih kecil, dan proses ini diulangi, memecah kelompok menjadi yang lebih kecil. Proses pemartisian berlanjut hingga setiap kelompok pada tingkat terendah hanya berisi satu objek, atau objek dalam sebuah kelompok cukup mirip satu sama lain, membentuk yang bisa disebut sebagai kelompok-kelompok yang lebih kecil [3].

Dengan demikian, penelitian ini bertujuan untuk memberikan solusi dalam mengoptimalkan penggunaan data penjualan di PT.AYN, serta membantu perusahaan dalam memahami perilaku konsumen dan tren pembelian, yang pada gilirannya dapat meningkatkan efisiensi operasional dan keunggulan kompetitif perusahaan di pasar yang semakin kompetitif.

TINJAUAN PUSTAKA

A. Tantangan dalam Data Mining

Meskipun memiliki banyak manfaat, data mining juga menghadapi berbagai tantangan. Salah satu tantangan utama adalah kualitas data. Data yang tidak lengkap, tidak konsisten, atau bias dapat menghasilkan pola yang tidak akurat (Han, Kamber, & Pei, 2011). Selain itu, masalah privasi dan etika juga menjadi perhatian, terutama ketika data mining digunakan untuk analisis data pribadi. [4]. Perkembangan teknologi dan metodologi baru terus mendorong batas-batas data mining. Deep learning dan big data analytics adalah dua tren utama yang saat ini mendominasi penelitian dan aplikasi data mining (Zhang, Yang, & Qin, 2018). Penggunaan algoritma deep learning dalam data mining memungkinkan pengenalan pola yang lebih kompleks dan akurat dari data yang sangat besar [5].

B. Penerapan Metode Clustering

Metode clustering diterapkan dalam berbagai bidang. Misalnya, dalam analisis pasar, clustering digunakan untuk segmentasi pelanggan berdasarkan perilaku belanja mereka (Berry & Linoff, 2004). Dalam pengenalan pola, clustering digunakan untuk pengelompokan gambar berdasarkan fitur visual (Duda, Hart, & Stork, 2001) [6]. Meskipun metode clustering sangat berguna, ada beberapa tantangan yang dihadapi dalam penerapannya. Salah satu tantangan utama adalah menentukan jumlah cluster yang optimal. Selain itu, kualitas clustering dapat dipengaruhi oleh pemilihan metrik jarak dan parameter algoritma (Jain, 2010) [7].

C. Penerapan Hierarchical Divisive Clustering

HDC banyak digunakan dalam berbagai aplikasi seperti biologi komputasi, pengelompokan teks, dan analisis pasar. Dalam biologi, HDC digunakan untuk pengelompokan gen berdasarkan ekspresi mereka (Eisen et al., 1998). Dalam pengelompokan teks, HDC membantu dalam klasifikasi dokumen dan analisis sentimen [8]. Tren terbaru dalam HDC melibatkan kombinasi dengan teknik machine learning lainnya seperti pembelajaran mendalam untuk meningkatkan akurasi dan efisiensi clustering (Zhao et al., 2018). Selain itu, pendekatan berbasis graf dan teori jaringan juga digunakan untuk meningkatkan pemahaman tentang struktur data yang kompleks [9]. Kelebihan HDC termasuk kemampuan untuk menangani data dengan struktur hierarkis yang kompleks dan fleksibilitas dalam menentukan jumlah cluster akhir. Namun, kelemahan utama HDC adalah kompleksitas komputasi yang tinggi dan kepekaan terhadap outlier dan data noise (Jain, Murty, & Flynn, 1999) [10].

METODOLOGI

a. Desain Penelitian

Penelitian ini mengadopsi pendekatan pengembangan sistem, dengan fokus pada desain aplikasi dan penerapan metode algoritma dalam mengevaluasi data penjualan. Desain penelitian ini terbagi menjadi tiga

tahap utama: (a) pengumpulan data, (b) perancangan sistem dan (c) penerapan metode.

b. Pengumpulan Sumber Data

Tahap pertama adalah wawancara mandiri dengan Manager di PT.AYN. Peneliti menanyakan kepada Manager mengenai masalah pada evaluasi data penjualan yang ada di perusahaan tersebut. Setelah bertanya Peneliti Mencoba Untuk mencari masalah yang nantinya akan diselesaikan dengan menggunakan aplikasi. selanjutnya yaitu menganalisis masalah, menentukan masalah apa saja yang ada serta metode clustering yang telah di teliti. Hal ini berguna untuk memecahkan masalah yang ada agar menjadi lebih teratur.

c. Perancangan Sistem

Perancangan sistem dibuat dengan menggunakan UML yang terdiri dari usecase diagram, activity diagram, sequence diagram & class diagram. Desain interface dibuat dengan menggunakan Visio untuk menjelaskan template dari masing-masing halaman yang dibuat.

d. Penerapan Metode

Tahap terakhir adalah mengaplikasikan permasalahan tersebut pada metode Clustering untuk mengetahui dan menyelesaikan masalah yang ada pada perusahaan dengan menentukan berbagai klaster yang ada pada data yang terdapat di perusahaan tersebut seperti, nama distributor, nama produk, serta jumlah datanya. Hal yang dihasilkan nantinya dapat berupa tren penjualan yang dapat membantu perusahaan dalam menggali informasi lebih dalam tentang penjualan melalui data penjualan.

e. Uji Coba dan Evaluasi

Pengujian dilakukan untuk mengevaluasi kinerja sistem dari berbagai aspek, yaitu dengan model pengujian blackbox testing. Uji coba juga melibatkan penggunaan umum (user) dan fitur, sistem, seperti kegagalan setiap aktivitas atau adanya bug dari semua proses.

f. Analisis Data

Data yang diperoleh dari pengujian sistem akan dianalisis secara kualitatif dan kuantitatif. Evaluasi kinerja sistem akan dibandingkan dengan kriteria keberhasilan yang ditetapkan sebelumnya. Temuan analisis akan digunakan untuk mendukung kesimpulan dan saran dalam bagian akhir penelitian. Dengan mengikuti metodologi ini, diharapkan dapat mencapai pengembangan sistem Data Mining Menggunakan Metode Clustering Dan Hierarki Divisive Clustering Dalam Evaluasi Data Penjualan.

HASIL PENELITIAN DAN PEMBAHASAN

A. Penerapan Metode

Untuk mengetahui Tren produk (paling diminati), PT.AYN telah menggunakan beberapa data pola penjualan, data tersebut digunakan dalam perhitungan Metode Clustering dengan transaksi penjualan selama 1 tahun pada tahun 2022. Adapun penerapan tahapan-tahapan algoritma apriori yang digunakan dalam penelitian ini adalah sebagai berikut :

1. Langkah Pertama : Menentukan data awal penjualan

| Kode | Produk | Kategori | Ellara Karya CV - Banda Aceh | Bintang Mutiara Cemerlang - Medan | Everbright - Medan | Catur Sentosa Amugrah - Pekanbaru | Everbright - Padang | Teratai Kristal - Bukittinggi | Indogaben Sukses Perkasa - Batam | Indo Global Bestari - Kepulauan Riau | Viva Central Tembilahan | Niaga Sukses Mandiri CV - Riau |
|------|---|----------|------------------------------|-----------------------------------|--------------------|-----------------------------------|---------------------|-------------------------------|----------------------------------|--------------------------------------|-------------------------|--------------------------------|
| A001 | Facial Soft Pack 250 sheets | Facial | 4932 | 2416 | 3640 | 2060 | 2030 | 1927 | 2858 | 2244 | 1340 | 2395 |
| A002 | Facial Travel Pack 50 sheets | Facial | 7389 | 6112 | 9488 | 11355 | 7669 | 5877 | 5413 | 5780 | 8004 | 4427 |
| A003 | Facial Soft Pack 200 sheets | Facial | 4002 | 2877 | 2894 | 3952 | 2973 | 2843 | 2810 | 3221 | 2386 | 3415 |
| A004 | Tessa Toilet Roll 110 gram | Toilet | 12000 | 9611 | 10356 | 5591 | 4832 | 9715 | 2823 | 3882 | 5733 | 7438 |
| A005 | Tessa Soft Bathroom 6 Rolls | Toilet | 1909 | 1793 | 1776 | 1470 | 1807 | 4106 | 6988 | 2031 | 3859 | 6743 |
| A006 | Tessa Soft Bathroom 12 rolls | Toilet | 7545 | 2425 | 1243 | 7108 | 737 | 2047 | 840 | 1394 | 5496 | 7365 |
| A007 | Tessa Nature Towel Roll 100 sheets | Nature | 6224 | 8632 | 9847 | 7447 | 7023 | 11798 | 5279 | 8507 | 10369 | 4785 |
| A008 | Nature Eucalyptus Facial Travel Pack 60 sheet | Nature | 1531 | 2182 | 2892 | 1333 | 1891 | 1833 | 1169 | 1753 | 1863 | 1899 |
| A009 | Tessa Nature Kitchen Towel 70 sheets | Nature | 1784 | 854 | 860 | 716 | 683 | 888 | 843 | 960 | 975 | 609 |
| A010 | Tessa Nature Kitchen Towel 150 sheets | Nature | 4148 | 2319 | 1870 | 1118 | 819 | 1221 | 1290 | 835 | 957 | 1200 |

Gambar 1. Tabel Penjualan Tahun 2022

2. Langkah Kedua : Data di konversi menjadi matriks

Untuk penerapan algoritma analisis divisive clustering, akan diberikan hitungan sebagai berikut dimana data tersebut telah di konversi menjadi sebuah matriks jarak Euclid dengan menggunakan rumus Euclidean, dengan model perhitungan sebagai berikut :

Data dijumlahkan menggunakan rumus Euclidean yaitu:

$$\begin{aligned}
 &= |7,389-4,932|^2 + |6,112-2,416|^2 + |9,488-3,640|^2 + |11,355-2,060|^2 + |7,669-2,030|^2 + |5,877-1,927|^2 + |5,413-2,858|^2 + |5,780-2,244|^2 + |8,004-1,340|^2 + |4,427-2,395|^2 \\
 &= 2,4572+3,6962+5,8482+9,2952+5,6392+3,9502+2,5552+3,5362+6,6642+2,0322 \\
 &=6036849+136604166+34199104+86397025+31798321+15602500+6528025+12503296+44408896+4129024 \\
 &= 255263456 \\
 &= \sqrt{255263456} \\
 &= 15976.966
 \end{aligned}$$

Lanjut ke perhitungan selanjutnya

$$\begin{aligned}
 &= |4,002-4,932|^2 + |2,877-2,416|^2 + |2,894-3,640|^2 + |3,952-2,060|^2 + |2,973-2,030|^2 + |2,843-1,927|^2 + |2,810-2,858|^2 + |3,221-2,244|^2 + |2,386-1,340|^2 + |3,415-2,395|^2 \\
 &= (-930)^2+4612+(-746)^2+1,8922+9432+9162+(-48)^2+9772+1,0462+1,0202 \\
 &=864900+212521+556516+3579664+889249+839056+839056+954529+1094116+1040400 \\
 &= 10033255
 \end{aligned}$$

$$= \sqrt{10033255}$$

$$= 3167.5314$$

Perhitungan dilakukan secara terus menerus hingga data yang ada habis dijumlahkan hingga nantinya berbentuk matriks seperti pada gambar dibawah ini :

Table 1. Tabel Hasil Matriks

| | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| X1 | 0 | 15,976 | 3,167 | 16,589 | 7,778 | 9,301 | 19,040 | 4,043 | 5,814 | 3,553 |
| X2 | 15,976 | 0 | 14,049 | 10,714 | 16,806 | 14,745 | 8,471 | 17,938 | 20,699 | 18,916 |
| X3 | 3,167 | 14,049 | 0 | 15,690 | 6,930 | 7,973 | 17,171 | 4,801 | 7,178 | 5,624 |
| X4 | 16,589 | 10,714 | 15,690 | 0 | 17,858 | 15,575 | 10,154 | 19,035 | 21,722 | 19,195 |
| X5 | 7,778 | 16,806 | 6,930 | 17,858 | 0 | 10,530 | 18,539 | 8,253 | 9,938 | 9,368 |
| X6 | 9,301 | 14,745 | 7,973 | 15,575 | 10,530 | 0 | 18,700 | 10,816 | 12,011 | 10,374 |
| X7 | 19,040 | 8,471 | 17,171 | 10,154 | 18,539 | 18,700 | 0 | 20,450 | 23,405 | 21,781 |
| X8 | 4,043 | 17,938 | 4,801 | 19,035 | 8,253 | 10,816 | 20,450 | 0 | 3,446 | 3,413 |
| X9 | 5,814 | 20,699 | 7,178 | 21,722 | 9,938 | 12,011 | 23,405 | 3,446 | 0 | 3,100 |
| X10 | 3,553 | 18,916 | 5,624 | 19,195 | 9,368 | 10,374 | 21,781 | 3,413 | 3,100 | 0 |

Data yang ada di dalam matriks memiliki nilai yang berbeda antara satu objek dengan objek lainnya, maka di lakukan penjumlahan untuk menentukan nilai objek seperti pada berikut :

Tabel 2. Perhitungan I

| Objek | Avarage dissimilarity to the other objects |
|-------|---|
| X1 | $(15,976+3,167+16,589+7,778+9,301+19,040+4,043+5,814+3,553)/9 = 9473.44444$ |
| X2 | $(15,976+14,049+10,714+16,806+14,745+8,471+17,938+20,699+18,916)/9 = 15368.2222$ |
| X3 | $(3,167+14,049+15,690+6,930+7,973+17,171+4,801+7,178+5,624)/9 = 9175.88889$ |
| X4 | $(16,589+10,714+15,690+17,858+15,575+10,154+19,035+21,722+19,195)/9 = 16281.3333$ |
| X5 | $(7,778+16,806+6,930+17,858+10,530+18,539+8,253+9,938+9,368)/9 = 11777.7778$ |
| X6 | $(9,301+14,745+7,973+15,575+10,530+18,700+10,816+12,011+10,374)/9 = 12225$ |
| X7 | $(19,040+8,471+17,171+10,154+18,539+18,700+20,450+23,405+21,781)/9 = 17523.4444$ |
| X8 | $(4,043+17,938+4,801+19,035+8,253+10,816+20,450+3,446+3,413)/9 = 10243.8889$ |
| X9 | $(5,814+20,699+7,178+21,722+9,938+12,011+23,405+3,446+3,100)/9 = 11923.6667$ |
| X10 | $(3,553+18,916+5,624+19,195+9,938+10,374+21,781+3,413+3,100)/9 = 10591.5556$ |

Objek X7 merupakan splinter Group. Untuk setiap objek dari grup yang besar harus dihitung average dissimilarity dengan objek yang tersisa, dan

membandingkan itu dengan average dissimilarity dengan objek dari splinter group seperti tabel dibawah:

Tabel 3. Perhitungan I

| Objek | Average dissimilarity in remain objects | Average dissimilarity to objects of splinter group | Difference |
|-------|---|--|------------|
| X1 | $(15,976+3,167+16,589+7,778+9,301+4,043+5,814+3,553)/8 = 8277.625$ | 19040 | -10762.375 |
| X2 | $(15,976+14,049+10,714+16,806+14,745+17,938+20,699+18,916)/8 = 16230.375$ | 8471 | 7759.375 |
| X3 | $(3,167+14,049+15,690+6,930+7,973+4,801+7,178+5,624)/8 = 8176.5$ | 17171 | -8994.5 |
| X4 | $(16,589+10,714+15,690+17,858+15,575+19,035+21,722+19,195)/8 = 17047.25$ | 10154 | 6893.25 |
| X5 | $(7,778+16,806+6,930+17,858+10,530+8,253+9,938+9,368)/8 = 10932.625$ | 18539 | -7606.375 |
| X6 | $(9,301+14,745+7,973+15,575+10,530+10,816+12,011+10,374)/8 = 11415.625$ | 18700 | -7284.375 |
| X8 | $(4,043+17,938+4,801+19,035+8,253+10,816+3,446+3,413)/9 = 10243.8889$ | 20450 | -11481.875 |
| X9 | $(5,814+20,699+7,178+21,722+9,938+12,011+3,446+3,100)/9 = 11923.6667$ | 23405 | -12916.5 |
| X10 | $(3,553+18,916+5,624+19,195+9,938+10,374+3,413+3,100)/9 = 10591.5556$ | 21781 | -12588.125 |

Terlihat bahwa perbedaan terbesar terdapat pada objek X2. Oleh sebab itu, objek X2 berpindah bagian kepada splinter group, kemudian dihitung kembali sampai dengan data habis.

Tabel 4. Perhitungan I

| Objek | Average dissimilarity in remain objects | Average dissimilarity to objects of splinter group | Difference |
|-------|--|--|------------|
| X1 | $(3,167+16,589+7,778+9,301+4,043+5,814+3,553)/7 = 7177.8571$ | 2606.8125 | 4571.0446 |
| X3 | $(3,167+15,690+6,930+7,973+4,801+7,178+5,624)/7 = 7337.5714$ | 2527.25 | 4810.3214 |
| X4 | $(16,589+15,690+17,858+15,575+19,035+21,722+19,195)/7 = 17952$ | 8803.625 | 9148.375 |

| | | | |
|-----|---|-----------|-----------|
| X5 | $(7,778+6,930+17,858+10,530+8,253+9,938+9,368)/7=10093.571$ | 4599.8125 | 5493.7589 |
| X6 | $(9,301+7,973+15,575+10,530+10,816+12,011+10,374)/7 = 10940$ | 3730.3125 | 7209.6875 |
| X8 | $(4,043+4,801+19,035+8,253+10,816+3,446+3,413)/7 = 7686.7143$ | 3228.0625 | 4458.6518 |
| X9 | $(5,814+7,178+21,722+9,938+12,011+3,446+3,100)/7 = 9029.8571$ | 3891.25 | 5138.6071 |
| X10 | $(3,553+5,624+19,195+9,938+10,374+3,413+3,100)/7 = 7803.8571$ | 3163.9375 | 4639.9196 |

Terlihat bahwa perbedaan terbesar terdapat pada objek X4. Oleh sebab itu, objek X4 berpindah bagian kepada splinter group

Tabel 5. Perhitungan I

| Objek | Average dissimilarity in remain objects | Average dissimilarity to objects of splinter group | Difference |
|-------|---|--|------------|
| X1 | $(3,167+7,778+9,301+4,043+5,814+3,553)/6=5609.3333$ | 9597.9063 | -3988.573 |
| X3 | $(3,167+6,930+7,973+4,801+7,178+5,624)/6 = 5945.5$ | 9108.625 | -3163.125 |
| X5 | $(7,778+6,930+10,530+8,253+9,938+9,368)/6 = 8799.5$ | 11228.906 | -2429.406 |
| X6 | $(9,301+7,97+15,575+10,530+10,816+12,011+10,374)/6 = 10167.5$ | 9652.6563 | 514.84375 |
| X8 | $(4,043+4,80+19,035+8,253+10,816+3,446+3,413)/6 = 5795.3333$ | 11131.531 | -5336.198 |
| X9 | $(5,814+7,178+9,938+12,011+3,446+3,100)/6 = 6914.5$ | 12806.625 | -5892.125 |
| X10 | $(3,553+5,624+9,938+10,374+3,413+3,100)/6 = 5905.3333$ | 11179.469 | -5274.135 |

Terlihat bahwa perbedaan terbesar terdapat pada objek X4. Oleh sebab itu, objek X6 berpindah bagian kepada splinter group

Tabel 6. Perhitungan I

| Objek | Average dissimilarity in remain objects | Average dissimilarity to objects of splinter group | Difference |
|-------|--|--|------------|
| X1 | $(3,167+7,778+4,043+5,814+3,553)/5= 4871$ | 9449.4531 | -4578.4531 |
| X3 | $(3,167+6,930+4,801+7,178+5,624)/5 = 5540$ | 8540.8125 | -3000.8125 |
| X5 | $(7,778+6,930+8,253+9,938+9,368)/5 = 8453.4$ | 10879.453 | -2426.0531 |
| X8 | $(4,043+4,80+19,035+8,253+3,446+3,413)/5 = 4791.2$ | 10973.766 | -6182.5656 |
| X9 | $(5,814+7,178+9,938+3,446+3,100)/5 = 5895.2$ | 12408.813 | -6513.6125 |
| X10 | $(3,553+5,624+9,938+3,100)/5 = 5011.6$ | 10776.734 | -5765.1344 |

Pada bagian ini terlihat, untuk semua difference bernilai negatif. Oleh sebab itu tidak terjadi perpindahan. Dengan ini maka secara otomatis X6 masuk ke cluster 2. Dalam langkah berikutnya akan dilakukan pembagian cluster pembagian cluster dilakukan pada cluster yang memiliki rata-rata terbesar, dengan cara yang sama seperti sebelumnya sampai dengan data habis dengan menggunakan matriks sebagai berikut.

Tabel 7. Perhitungan Matriks lanjutan

| | X1 | X3 | X5 | X6 | X8 | X10 |
|-----|------|------|------|------|------|------|
| X1 | 0 | 3167 | 7778 | 4043 | 5814 | 3553 |
| X3 | 3167 | 0 | 6930 | 4801 | 7178 | 5624 |
| X5 | 7778 | 6930 | 0 | 8253 | 9938 | 9368 |
| X6 | 4043 | 4801 | 8253 | 0 | 3446 | 3413 |
| X8 | 5814 | 7178 | 9938 | 3446 | 0 | 3100 |
| X10 | 3553 | 5624 | 9368 | 3413 | 3100 | 0 |

Tabel 8. Perhitungan 2

| Objek | Average dissimilarity in remain objects |
|-------|--|
| X1 | $(3,167+7,778+4,043+5,814+3,553)/5= 4871$ |
| X3 | $(3,167+6,930+4,801+7,178+5,624)/5 = 5540$ |
| X5 | $(7,778+6,930+8,253+9,938+9,368)/5 = 8453.4$ |
| X8 | $(4,043+4,80+19,035+8,253+3,446+3,413)/5 = 4791.2$ |
| X9 | $(5,814+7,178+9,938+3,446+3,100)/5 = 5895.2$ |
| X10 | $(3,553+5,624+9,938+3,100)/5 = 5011.6$ |

Terlihat bahwa perbedaan terbesar terdapat pada objek X5. Oleh sebab itu, objek X5 berpindah bagian kepada splinter group

Tabel 9. Perhitungan 2

| Objek | Average dissimilarity in remain objects | Average dissimilarity to objects of splinter group | Difference |
|-------|---|--|------------|
| X1 | $(3,167+4,043+5,814+3,553)/4 = 4144.25$ | 7778 | -3633.75 |
| X3 | $(3,167+4,801+7,178+5,624)/4 = 5192.5$ | 6930 | -1737.5 |
| X8 | $(4,043+4,80+19,035+3,446+3,413)/4 = 3925.75$ | 8253 | -4327.25 |
| X9 | $(5,814+7,178+3,446+3,100)/4 = 4884.5$ | 9938 | -5053.5 |
| X10 | $(3,553+5,624+3,413+3,100)/4 = 3922.5$ | 9368 | -5445.5 |

Proses dihentikan karena kluster telah selesai dibagi, dengan pembagian kluster yang menghasilkan nilai negative di setiap pembagiannya. Pada tahap ini kita dapat menyimpulkan bahwa :

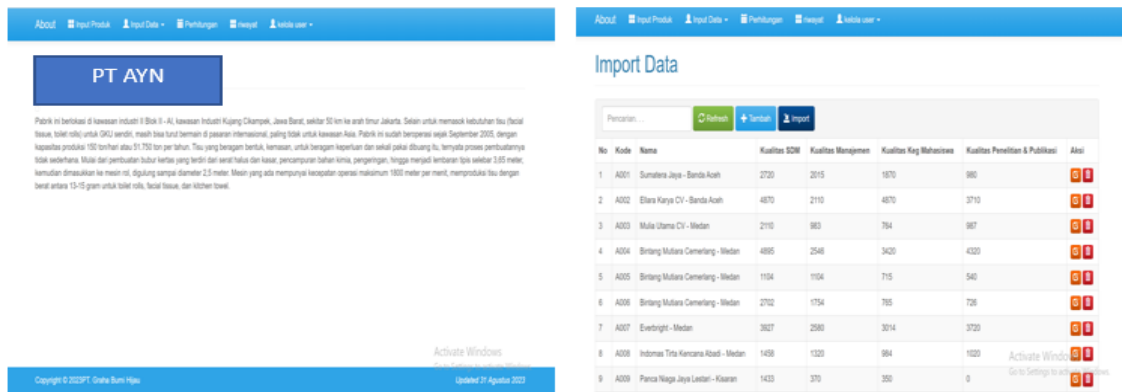
- A. Kluster 1 terdiri dari 3 produk (X7, X2, dan X4) yaitu Tessa Nature towel roll 100 Sheets, Facial Travel Pack 50 sheets, Tessa Toilet Roll 110 dan menjadikan Tessa Nature towel roll 100 Sheets dengan penjualan terbaik dan berada di posisi pertama.
- B. Kluster 2 terdiri dari 1 produk (X6) yaitu Tessa Soft Bathroom 12 rolls.
- C. Kluster 3 terdiri dari 5 produk (X1, X3, X5, X8, X9, dan X10) yaitu Facial Soft Pack 250 sheets, Facial Soft Pack 200 sheets, Nature Eucalyptus Facial Travel Pack 60 sheet, Tessa Nature Kitchen Towel 70 sheets, Tessa Soft Bathroom 6 Rolls, dan Tessa Nature Kitchen Towel 150 sheets, yang membuat produk yang berada di cluster 3 menjadi

Dengan begitu, terlihat bahwa dari perbandingan diatas dapat dilihat bahwa Perusahaan Indogaben Sukses Perkasa - Batam memiliki pembelian paling baik, dibandingkan dengan Niaga Sukses Mandiri CV - Riau. Oleh karena itu, Hal ini dapat dijadikan acuan dalam evaluasi penjualan pada periode berikutnya.

B. Implementasi Sistem

1. Menu utama

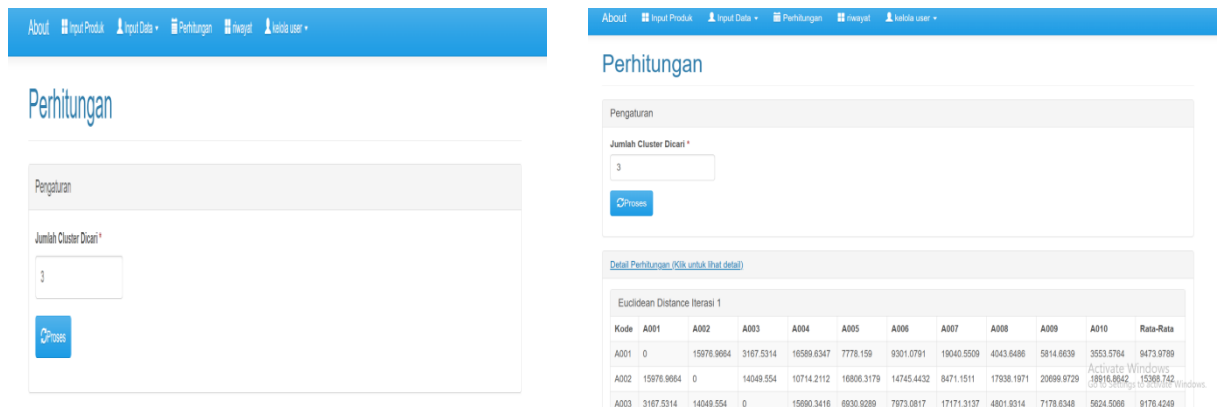
Form ini berfungsi untuk menampilkan menu bagian utama, menu input data, kelola user, Input Data Riwayat dan proses hitung seperti terlihat pada gambar berikut.



Gambar 3. Menu Utama dan Impor Data

2. Menu Perhitungan

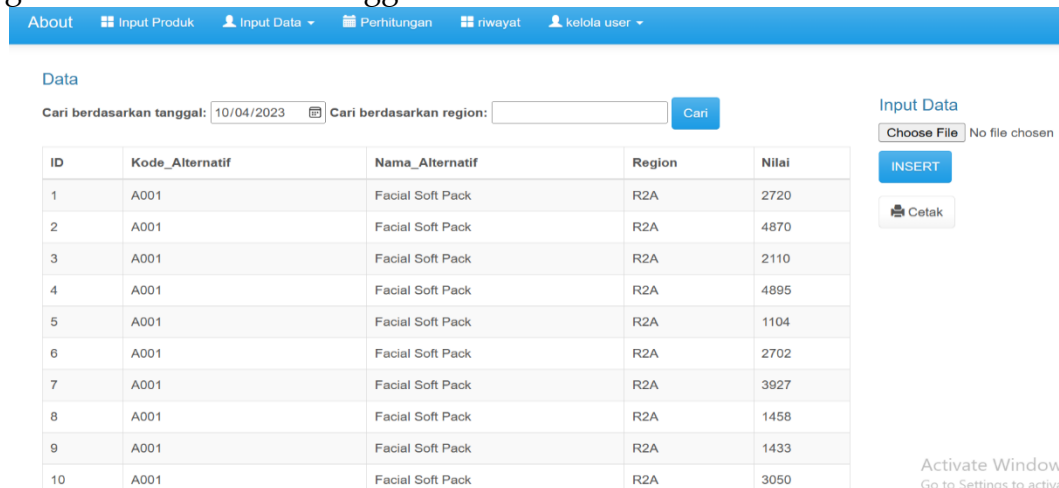
Pada form ini akan ditampilkan form proses untuk menghitung dan juga dapat di setting ingin terbagi dalam berapa cluster :



Gambar 4. Perhitungan

3. Menu Riwayat

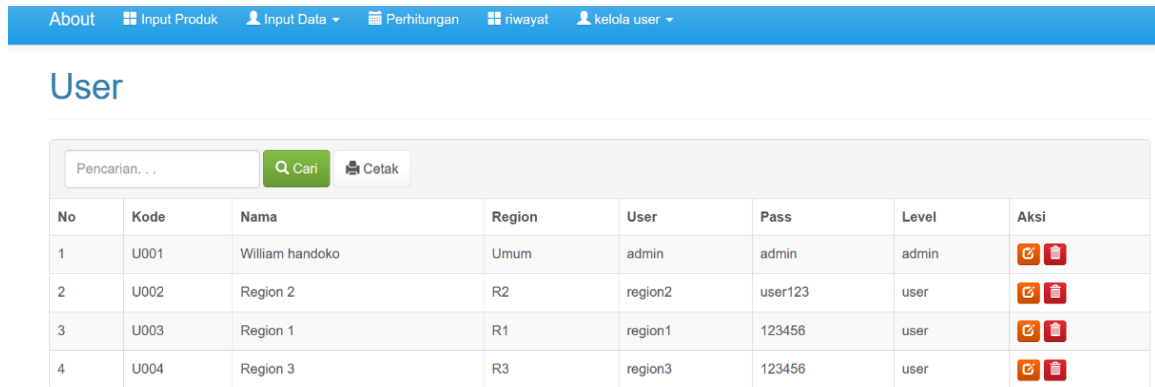
Dalam form ini akan ditampilkan riwayat data dimana jika kita masukkan data pada form data maka data akan tersimpan di system, dan jika ingin dicari kembali berdasarkan tanggal, user ataupun admin dapat menyetel hal tersebut di bagian cari berdasarkan tanggal.











Gambar 5. Menu Riwayat

4. Menu Kelola User

Pada bagian ini hanya admin yang dapat melihat dan mengontrol form tersebut.



The screenshot shows a web application interface for user management. At the top, there is a navigation bar with links: About, Input Produk, Input Data, Perhitungan, riwayat, and kelola user. Below the navigation bar, the title 'User' is displayed. A search bar with the placeholder 'Pencarian...' and a 'Cari' button is present. Below the search bar is a table with the following data:

| No | Kode | Nama | Region | User | Pass | Level | Aksi |
|----|------|-----------------|--------|---------|---------|-------|---|
| 1 | U001 | William handoko | Umum | admin | admin | admin |   |
| 2 | U002 | Region 2 | R2 | region2 | user123 | user |   |
| 3 | U003 | Region 1 | R1 | region1 | 123456 | user |   |
| 4 | U004 | Region 3 | R3 | region3 | 123456 | user |   |

Gambar 6. Kelola User

KESIMPULAN DAN REKOMENDASI

Dari penulisan penelitian yang berjudul Implementasi Data Mining Untuk Evaluasi Data Penjualan Menggunakan Metode Clustering Dan Algoritma Hirarki Divisive Pada PT.AYN. Kesimpulan yang di dapat sebagai berikut:

1. Metode Clustering dan algoritma Hierarki Divisive dapat digunakan untuk menyelesaikan permasalahan perhitungan serta pengelompokkan region dan produk berdasarkan data yang ada pada PT.AYN. Setiap daerah memiliki pola dan kecenderungan yang berbeda-beda sesuai dengan data yang di masukkan. Hal ini ditunjukkan oleh perbedaan hasil yang dihasilkan dalam proses perhitungan.
2. Data penjualan dapat digolongkan sesuai dengan model dari algoritma yang digunakan pada metode clustering, sehingga data tersebut dapat diprediksi dan dicari polanya dengan menggunakan metode ini.
3. Algoritma hirarki divisive dapat membantu mengembangkan strategi pemasaran dengan memberikan saran kepada konsumen.
4. Pengaplikasian algoritma hirarki divisive menggunakan tools database sql dan sublime dimulai dengan cara mengimport data penjualan per bulan dengan menggunakan Microsoft Excel. Semakin banyak data, maka semakin banyak cluster yang tercipta.

PENELITIAN LANJUTAN

Pengembangan penelitian masih perlu dilakukan dalam penelitian ini yaitu dengan mencoba metode-metode lain yang berhubungan dengan data mining atau juga mengembangkan menjadi data warehouse dengan melihatkan sub bagian lain selain penjualan. Dalam hal aplikasi juga perlu mengembangkan dalam hal otomasi ataupun fitur-fitur dan segi kemanan data perlu menjadi tantangan kedepan guna menjadikan penelitian ini menjadi lebih baik lagi.

REFERENSI

- Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern Classification*. Wiley
- Eisen, M. B., Spellman, P. T., Brown, P. O., & Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences*, 95(25), 14863-14868.
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques*. Elsevier.
- Hartomi, Z. H., Sabna E., Yulanda, Rinaldi A, M., & Sanjaya R. (2022). Penerapan Data Mining Menggunakan Metode Clustering Evaluasi Data Penjualan PT. *Aspacindo Kedaton Motor*, E-ISSN : 2579 - 3918 | P-ISSN : 2302 - 710X.
- Irawan, y. (2019). *Implementation Of Data Mining For Determining Majors Using K-Means Algorithm In Students Of SMA Negeri 1 Pangkalan Kerinci*. *Journal of Applied Engineering and Technological Science*. Vol 1(1) 2019 : 17-29.
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651-666
- Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys (CSUR)*, 31(3), 264-323
- Mardi, Y. (2023). *Data Mining : Klasifikasi Menggunakan Algoritma C4.5*. ISSN : 2407-0491 E-ISSN : 2541-3716.
- Zhang, C., Yang, J., & Qin, Q. (2018). Data Mining with Deep Learning: A Review. *Journal of Data Mining and Knowledge Discovery*, 2(1), 1-18.
- Zhao, Y., Li, H., Yang, L., & Yang, J. (2018). A Hierarchical Divisive Clustering Algorithm Based on Kernel Density Estimation. *Information Sciences*, 432, 82-99..