

Klasterisasi Pengendalian Persediaan Aki Menggunakan Metode K-Means

Ricsa Andean¹, Septian Fendy², Aryo Nugroho³

e-mail: ¹ricsa.andean@fik.narotama.ac.id, ²septian.fendyputra@fik.narotama.ac.id,

³aryo.nugroho@narotama.ac.id

Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Narotama Surabaya

Abstract—The process of inventory control is an important part for companies to supply goods. Inventory control is used to meet the amount of goods needed at the time of marketing, but the company has problems in determining the amount of inventory due to the amount of goods needed when marketing changes every time. This research conducts data mining to help companies to be able to measure the amount of inventory before being marketed by extracting data using one of the data mining methods, namely grouping data. In this research grouping distribution report data and sales reports on battery companies, the data is grouped into 3 groups using one of the data grouping algorithms K-Means, from the results of grouping it is used to determine the group of products that are of interest, moderate and less desirable then the group is used as a reference for many, medium and little inventories. The data used are distribution report data and sales reports for the period January - December 2018 with a total of 125 products. After the data is processed using the K-Means algorithm, show that group 1 with the value of low data object is a product group that is less desirable, group 2 with the value of medium data object is a group of interested products in the medium category, while group 3 with the value of high data object is a group of products that are in high demand.

Intisari—Proses pengendalian persediaan merupakan bagian penting bagi perusahaan untuk melakukan penyediaan barang. Pengendalian persediaan dilakukan untuk memenuhi jumlah barang yang dibutuhkan pada saat pemasaran dilakukan, namun perusahaan memiliki permasalahan dalam menentukan jumlah persediaan barang dikarenakan jumlah barang yang dibutuhkan pada saat pemasaran berubah setiap waktu. Penelitian ini melakukan data mining guna membantu perusahaan agar dapat mengukur jumlah persediaan barang sebelum dipasarkan dengan melakukan penggalian data menggunakan salah satu metode data mining yaitu pengelompokan data. Pada penelitian ini melakukan mengelompokkan data laporan distribusi dan laporan penjualan pada perusahaan aki, data tersebut di kelompokkan menjadi 3 kelompok dengan menggunakan salah satu algoritma pengelompokan data yaitu K-Means, dari hasil pengelompokan digunakan untuk mengetahui kelompok produk yang diminati, sedang dan kurang diminati kemudian kelompok tersebut digunakan sebagai

acuan persediaan banyak, sedang dan sedikit. Data yang digunakan merupakan data laporan distribusi dan laporan penjualan periode Januari – Desember 2018 dengan total sebanyak 125 produk. Setelah data tersebut diolah menggunakan algoritma K-Means, menunjukkan bahwa kelompok 1 dengan nilai objek data rendah merupakan kelompok produk yang kurang diminati, kelompok 2 dengan nilai objek data sedang merupakan kelompok produk diminati dalam kategori sedang, sedangkan kelompok 3 dengan nilai objek data tinggi merupakan kelompok produk yang banyak diminati.

Kata kunci—Pengendalian, Persediaan, Data mining, Klasterisasi, K-Means, Produk diminati.

I. PENDAHULUAN

Di era globalisasi, teknologi semakin berkembang dengan cepat, perkembangan teknologi di dunia banyak membantu manusia memudahkan pekerjaannya agar lebih efektif dan efisien di berbagai bidang [1]. Salah satu teknologi yang dapat digunakan pada bidang industri yaitu memudahkan pekerjaan dalam melakukan pengendalian persediaan. Pengendalian persediaan merupakan bagian penting dalam perusahaan, Pengendalian persediaan dilakukan oleh perusahaan untuk menyediakan barang – barang yang dibutuhkan dengan tujuan memenuhi permintaan konsumen [2]. Namun perusahaan mengalami permasalahan dalam melakukan pengendalian persediaan yaitu melakukan persediaan yang tidak efisien dikarenakan persediaan yang berlebih (*over stock*) dan kekurangan persediaan (*out of stock*) menyebabkan kerugian pada perusahaan itu sendiri [3].

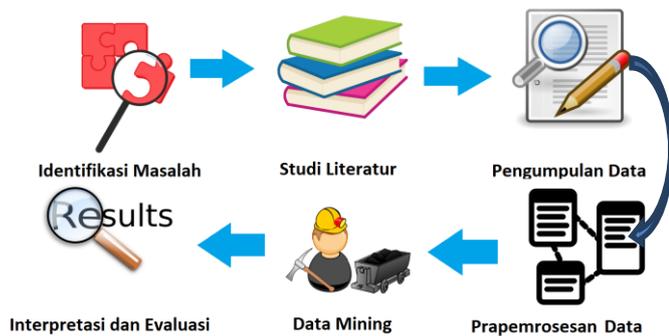
Persediaan yang berlebih mengakibatkan penumpukan barang dan biaya sewa penyimpanan barang, sedangkan kekurangan persediaan mengakibatkan tidak adanya ketersediaan stok barang yang akan dijual [4][5]. Untuk mengatasi kelebihan dan kekurangan persediaan penelitian melakukan penggalian data menggunakan metode klasterisasi yaitu melakukan pengelompokan data ke dalam beberapa kelompok berdasarkan jarak terdekat atau tingkat kemiripan data menggunakan metode algoritma *K-Means*. Pada penelitian ini menggunakan data historis laporan distribusi dan laporan penjualan dengan variabel yang diperoleh yaitu tipe aki (Deskripsi), data jumlah stok awal (SA), pembelian (BL), return penjualan (RB), penjualan aki (JL), return penjualan

(RJ), penyesuaian stok (AD) dan jumlah keseluruhan (Saldo)[6].

Data dibagi menjadi 3 kelompok dengan tujuan untuk mengetahui kelompok produk mana yang diminati, sedang dan kurang diminati kemudian data tersebut digunakan sebagai acuan untuk menentukan persediaan banyak, sedang, dan sedikit [7]. Adapun tujuan penelitian diharapkan dapat membantu perusahaan meminimalkan kelebihan dan kekurangan persediaan dengan mengetahui tingkat produk yang diminati [2].

II. METODE PENELITIAN

Metode penelitian dalam hal ini memaparkan tahapan dan prosedur yang akan dilakukan guna memecahkan permasalahan dalam penelitian ini. Berikut tahapan – tahapan penelitian dapat dilihat pada gambar 1 dibawah ini.



Gambar 1. Tahapan Penelitian

A. Identifikasi Masalah

Pada tahap ini melakukan identifikasi masalah pada perusahaan distributor serta merencanakan solusi untuk menyelesaikan permasalahan tersebut. Permasalahan yang ada yaitu pengendalian persediaan masih belum terorganisir dengan baik oleh karena itu masalah tersebut akan digunakan sebagai sumber permasalahan pada penelitian ini.

B. Studi Literatur

Tahap ini mencari referensi teori yang relevan pada kasus pengendalian persediaan dari jurnal, artikel, dan buku. Tujuan dari referensi tersebut akan digunakan sebagai dasar teori untuk menyelesaikan permasalahan pada penelitian ini.

C. Pengumpulan Data

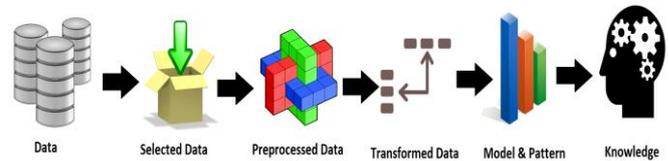
Pengumpulan data merupakan tahapan untuk mendapatkan informasi pada perusahaan sebagai bahan analisis penelitian ini. Data yang diperoleh merupakan data primer yaitu data yang dikumpulkan dan diolah sendiri dari perusahaan distributor aki. Data yang digunakan penelitian ini adalah data laporan penjualan dan laporan distribusi yang berbentuk angka yaitu data pembelian, penjualan dan stok awal persediaan aki.

D. Prapemrosesan Data

Pada tahap ini melakukan perubahan data dikarenakan data tersebut masih belum sempurna seperti data tidak lengkap dan tidak konsisten. Tujuan utama dalam tahap prapemrosesan data adalah melakukan pembersihan, penambahan,

pengurangan dan penyusunan data menjadi terstruktur sesuai kebutuhan pada proses *mining*.

Teknik yang digunakan dalam melakukan mengolah data adalah menggunakan metode *KDD (Knowledge Discovery)* metode ini meliputi beberapa tahapan. Tahapan – tahapan *KDD* sebagai berikut[8].



Gambar 2 Metode KDD

- 1) *Selected Data*: untuk meningkatkan akurasi dan kualitas hasil data mining, data diambil setiap minggu dengan rentang waktu setahun yaitu pada bulan Januari 2018 sampai bulan Desember 2018 dengan jumlah 52 minggu dan banyaknya produk sebanyak 125 produk.
- 2) *Preprocessed Data*: pada tahap ini melakukan pembersihan data karena data yang diperoleh dari perusahaan umumnya masih kotor seperti data yang kurang lengkap dan atribut yang tidak relevan, data yang masih kotor menyebabkan hasil yang kurang akurat dan terjadinya eror saat pemrosesan data sehingga pada proses ini perlu dilakukan pembersihan data. Pada proses ini peneliti menghapus data yang tidak lengkap yaitu produk yang tidak bersinambung dikarenakan produk tersebut sudah tidak di produksi dan variabel serta kolom yang tidak memiliki nilai.
- 3) *Transformation*: tahap ini melakukan perubahan data, data yang didapatkan dari perusahaan masih menggunakan format yang tidak sesuai, karena pada saat pemrosesan data hanya bisa menerima input data kategorikal dan tipe data numerik sehingga data yang digunakan harus berbentuk format excel dengan ekstensi CSV.
- 4) *Model & Patterns*: tahap ini merupakan proses pembentukan pola yaitu mengolah data menjadi kelompok dengan menggunakan metode klusterisasi. Proses ini bertujuan untuk menentukan kategori produk yang diminati dengan menggunakan algoritma *K-Means*[9].

Algoritma *K-Means* merupakan salah satu metode klusterisasi yang berbasis partisi yaitu metode yang digunakan untuk memisah data menjadi kelompok, kelompok yang memiliki kemiripan yang sama di masukan dalam kelompok yang sama dan kelompok yang berbeda di masukan dalam kelompok yang berbeda, untuk memisahkan kelompok tersebut metode *K-Means* membagi data dengan mengelompokkan objek berdasarkan nilai *k* yang telah ditentukan. Tujuan membagi data adalah untuk mengetahui kelompok produk yang diminati dan kurang diminati[10]. Beberapa tahapan metode *K-Means* sebagai berikut dalam bentuk *pseudocode*.

Input: D (Data), K (jumlah klaster)

- Tentukan jumlah kelompok (K)
- Inisialisasi nilai pusat secara acak

Process:

- Hitung setiap titik data dalam D dengan nilai pusat terdekat. Data yang telah dihitung akan menjadi kelompok K.
- Hitung ulang posisi nilai pusat.

Repeat:

Ulangi langkah *process* sampai tidak ada lagi perubahan keanggotaan titik data.

Output:

Data yang sudah dikelompokkan berdasarkan perhitungan jarak minimum. Untuk menghitung setiap objek data dengan nilai pusat menggunakan rumus Euclidean Distance sebagai berikut:

$$d(x, y) = \sqrt{\sum_{j=1}^n (x_j^2 - y_j^2)^2} \quad (1)$$

d = jarak antara x dan y

x = nilai pusat

y = data pada atribut

j = setiap data

n = jumlah data

x_j = nilai pusat ke j

y_j = objek data ke j

5) *Knowledge:* Pada tahap ini mengetahui dari hasil penggalian data. Hasil tersebut digunakan sebagai sumber pengetahuan untuk digunakan pada tujuan penelitian ini dan mempresentasikan pengetahuan dalam bentuk yang mudah dipahami.

E. Data Mining

Pada tahap ini melakukan penggalian data digunakan untuk menemukan pola – pola yang tidak diketahui sebelumnya menjadi informasi. Informasi tersebut dapat dimanfaatkan untuk menyelesaikan masalah yang ada pada penelitian ini.

F. Interpretasi dan Evaluasi

Tahap terakhir adalah memvisualisasikan hasil dari proses *data mining* berupa gambar atau tabel dan mengevaluasi apakah dari hasil *data mining* dapat digunakan untuk menyelesaikan masalah pada penelitian ini.

III. HASIL DAN PEMBAHASAN

Proses *data mining* pada penelitian ini menggunakan program *weka 3.8*, program tersebut digunakan untuk membantu penelitian dalam proses mengelompokkan data ke dalam klaster dengan menggunakan *tool cluster* dan salah satu metode algoritma klasterisasi K-Means. Langkah – langkah proses data mining sebagai berikut.

Pengujian menggunakan data tabel dengan mengubah data awal menjadi penyusunan sebagai berikut:

- Atribut yang digunakan pada penelitian ini berjumlah 5 atribut yaitu atribut deskripsi, minggu, saldo awal (SA), pembelian (BL) dan penjualan (JL).
- Penelitian ini menggunakan data dengan jumlah produk sebanyak 125.

A. Prapemrosesan Data

Tahap awal yang dilakukan adalah melakukan pembersihan data, data yang kurang relevan dan menambahkan atribut yang dibutuhkan. Berikut perubahan dari data awal dan data yang digunakan.

Tabel 1. Data Awal

Deskripsi	Satuan	SA	BL	RB	JL	RJ	AD	Saldo
YUASA 115F51/N120	PCS	13.00	12.00	0.00	0.00	0.00	0.00	25.00
YUASA 115F51/N120	PCS	25.00	0.00	0.00	8.00	0.00	0.00	17.00
YUASA 115F51/N120	PCS	17.00	10.00	0.00	0.00	0.00	0.00	23.00
YUASA 115F51/N120	PCS	23.00	0.00	0.00	0.00	0.00	0.00	20.00
YUASA 115F51/N120	PCS	20.00	0.00	0.00	0.00	0.00	0.00	18.00
...
YUASA YB9-B	PCS	04.00	0.00	0.00	0.00	0.00	0.00	04.00

Tabel 2. Data yang Digunakan

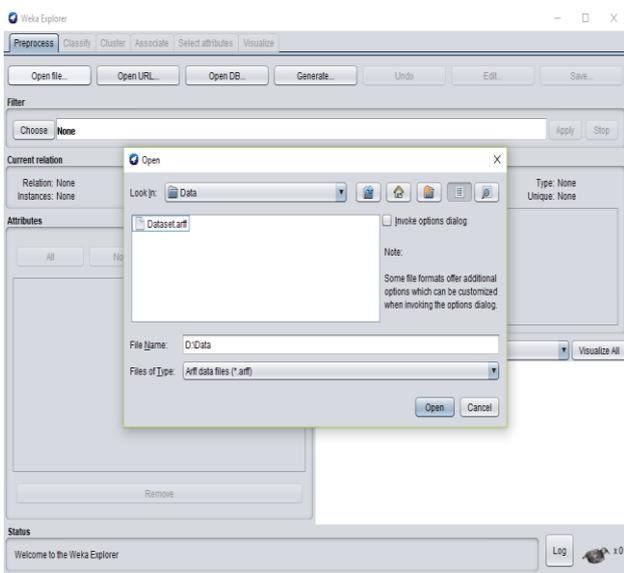
No.	Deskripsi	Minggu	SA	BL	JL
1.	YUASA 115F51/N120	M1	13	12	0
2.	YUASA 115F51/N120	M2	25	0	8
3.	YUASA 115F51/N120	M3	17	10	4
4.	YUASA 115F51/N120	M4	23	0	3
5.	YUASA 115F51/N120	M5	20	0	2
...
6500	YUASA YB9-B	M52	4	0	0

B. Data Mining

Pada tahap ini melakukan penggalian data digunakan untuk menemukan pola – pola yang tidak diketahui sebelumnya menjadi informasi. Untuk membantu melakukan proses penggalian data peneliti menggunakan perangkat lunak *weka 3.8* yang diawali dengan *input* data yang akan digunakan sebagai berikut.

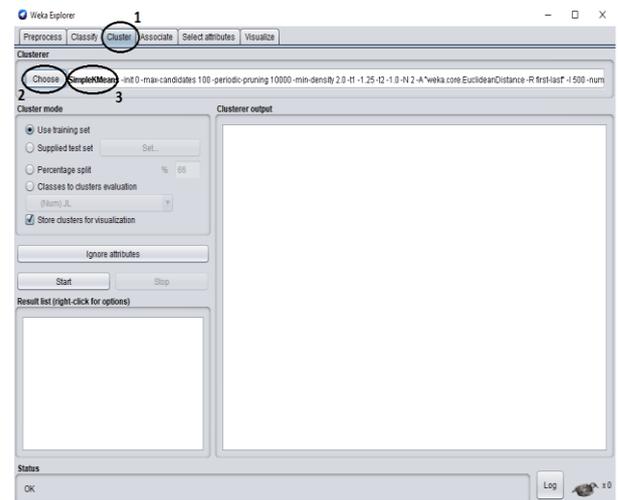


Gambar 3. Halaman Utama Weka



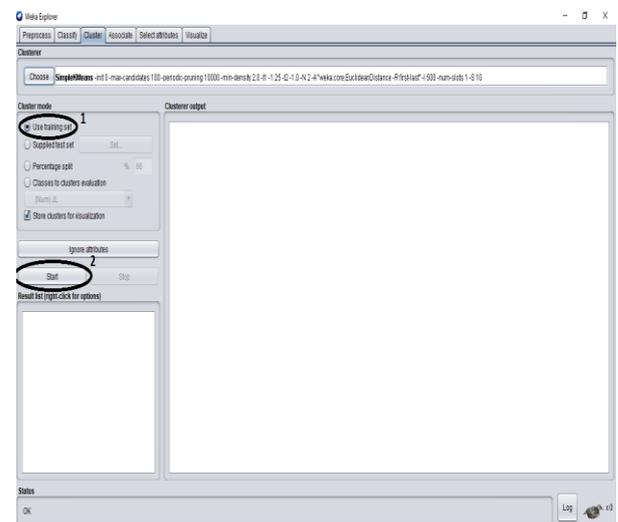
Gambar 4. Input Data

Setelah memilih data yang digunakan kemudian memilih algoritma yang akan digunakan dan menentukan jumlah kelompok yang diinginkan sebagai berikut.



Gambar 5. Pilih Algoritma

Setelah memilih algoritma dan menentukan jumlah kelompok selanjutnya memulai pemrosesan data sebagai berikut.



Gambar 6. Memulai Proses Data

Proses pengelompokan data diawali dengan melakukan inialisasi nilai tengah suatu objek data (centroid), nilai centroid digunakan sebagai perhitungan jarak objek data ke dalam kluster, menentukan centroid diawali dengan menentukan centroid awal dahulu kemudian centroid awal digunakan untuk menemukan centroid akhir, untuk menentukan centroid awal metode K-Means menentukan secara acak yang didapat dari nilai objek data. Hasil penentuan centroid awal yang digunakan dapat dilihat pada tabel berikut.

Tabel 3. Centroid Awal

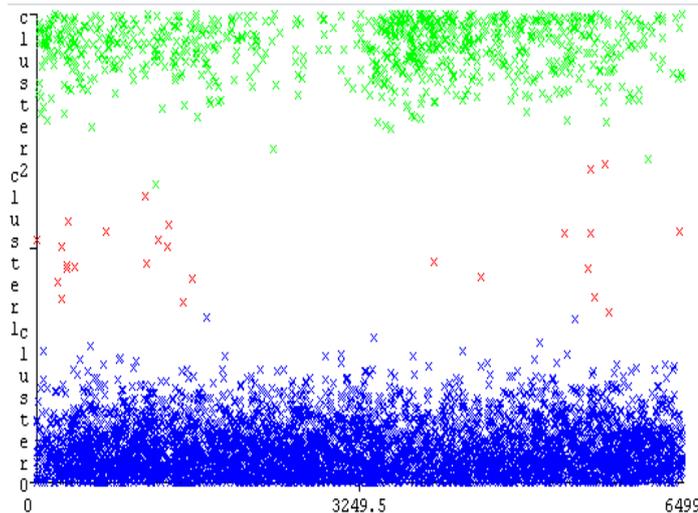
<i>Initial starting points (random):</i>			
Cluster 0	1	0	0
Cluster 1	0	20	5
Cluster 2	6	0	1

Pada tabel 3 di atas menunjukkan penentuan nilai pada objek data yang akan digunakan sebagai centorid awal kemudian nilai tersebut digunakan untuk melakukan perhitungan jarak terdekat (iterasi), perhitungan iterasi menggunakan rumus *euclidean distance* dengan tujuan menemukan nilai centroid akhir yang akan digunakan sebagai nilai pusat kluster. Centroid akhir yang di dapatkan dari hasil perhitungan iterasi dapat dilihat sebagai berikut.

Tabel 4. Centroid Akhir

Attribute	0	1	2
SA	2.7001	21.5385	26.6429
BL	0.4092	47	0.4567
JL	0.2568	38.6923	1.7592

Proses ini menempatkan objek data ke dalam beberapa kelompok (kluster). Data akan di masukkan ke dalam kluster dengan melakukan iterasi ulang berdasarkan jarak minimum dari nilai centroid dengan nilai setiap objek data menggunakan rumus *euclidean distance*. Hasil dari program weka melakukan iterasi sebanyak 11 kali dengan selisih error yang ditunjukkan menggunakan perhitungan *sum of squared error (SSE)* sebesar 19%, hasil dari pengujian data dapat di asumsikan bahwa tingkat keberhasilan mengelompokkan data ke dalam kluster sebesar 81%. Hasil pengelompokkan objek data ke dalam kluster dapat dilihat sebagai berikut.



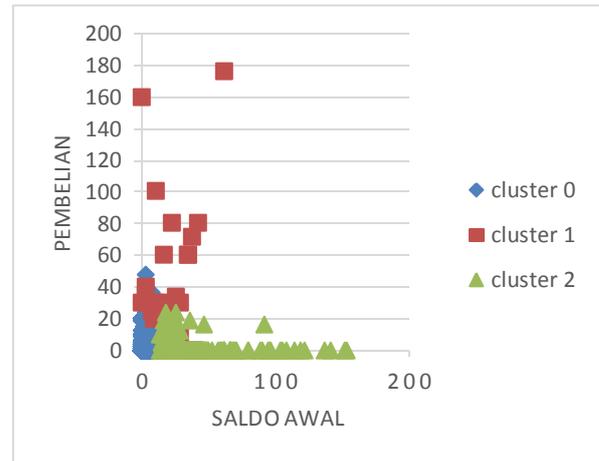
Gambar 7. Grafik Plot Hasil

Pada gambar 7 dapat dilihat hasil penyebaran data ke dalam kluster sesuai dengan anggota yang telah dihitung berdasarkan jarak terdekat dari data ke nilai pusat. Pada gambar 7 di atas menunjukkan bahwa cluster0 berwarna biru berada di posisi bawah merupakan kumpulan objek data yang memiliki nilai rendah selanjutnya cluster1 berwarna merah berada di posisi tengah merupakan kumpulan objek yang memiliki nilai tinggi, sedangkan cluster2 berwarna hijau posisi atas merupakan kumpulan objek data dengan nilai tinggi.

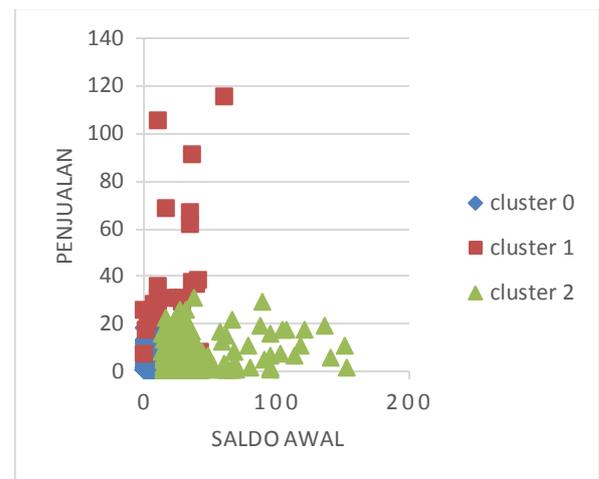
Tabel 5. Hasil Kluster

Kluster	Jumlah	Persen
0	5631	(87%)
1	26	(0%)
2	843	(13%)

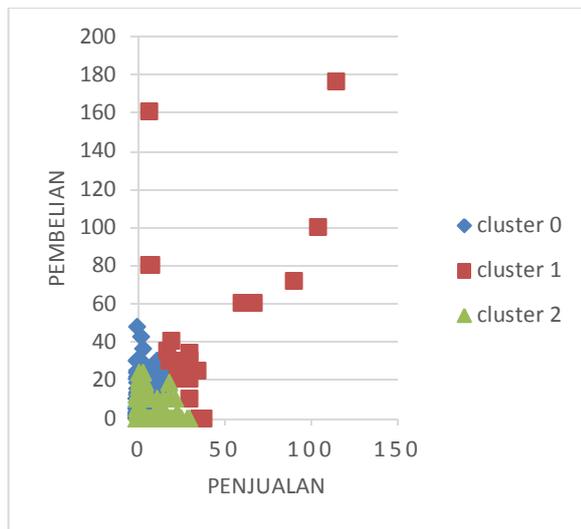
Pada tabel 5 dan gambar 8 di atas menunjukkan jumlah produk yang telah dibagi dan tergabung dalam 3 kelompok yaitu kluster 0, kluster 1, dan kluster 2, kemudian dapat dilihat berapa persen jumlah data yang ada di setiap kluster.



Gambar 8. Grafik Pembelian dan Saldo Awal



Gambar 9. Grafik Penjualan dan Saldo Awal



Gambar 10. Grafik Pembelian dan Penjualan

Tabel 6. Hasil Klasterisasi

Deskripsi	Minggu	SA	BL	JL	Cluster
YUASA 115F51/N120	M1	13	12	0	cluster0
YUASA 115F51/N120	M2	25	0	8	cluster2
YUASA 115F51/N120	M3	17	10	4	cluster2
YUASA 115F51/N120	M4	23	0	3	cluster2
YUASA 115F51/N120	M5	20	0	2	cluster2
YUASA 115F51/N120	M6	18	0	2	cluster2
YUASA 115F51/N120	M7	16	0	4	cluster2
YUASA 115F51/N120	M8	12	0	2	cluster0
YUASA 115F51/N120	M9	10	5	2	cluster0
YUASA 115F51/N120	M10	13	0	0	cluster0
YUASA 115F51/N120	M11	13	0	0	cluster0
YUASA 115F51/N120	M12	13	0	9	cluster0
YUASA 115F51/N120	M13	4	10	5	cluster0
YUASA 115F51/N120	M14	9	0	4	cluster0
YUASA 115F51/N120	M15	5	0	2	cluster0
YUASA 115F51/N120	M16	3	10	1	cluster0
YUASA 115F51/N120	M17	12	0	1	cluster0
YUASA 115F51/N120	M18	11	100	105	cluster1
YUASA 115F51/N120	M19	6	0	0	cluster0

Dari hasil klasterisasi dapat dilihat dari gambar 8, 9 dan 10 gambar tersebut menunjukkan bahwa kelompok produk yang berada didalam cluster 0 melakukan pembelian maksimal 50 barang dan saldo awal tidak sampai 50 barang artinya bahwa kelompok produk cluster 0 tidak menyimpan terlalu banyak dan merupakan produk yang kurang diminati dikarenakan kelompok cluster 0 memiliki nilai terkecil, selanjutnya kelompok produk cluster 1 dengan nilai pembelian dan penjualan diatas 100 barang dapat diartikan bahwa kelompok tersebut merupakan kelompok yang diminati, sedangkan kelompok 2 memiliki nilai sedang dengan pembelian dan penjualan dibawah kelompok cluster 1 dan di atas kelompok cluster 0 dapat diartikan bahwa kelompok tersebut diminati dalam kategori sedang.

Namun dari tabel 6 diatas menunjukkan bahwa ada produk yang masuk kedalam klaster yang berbeda dikarenakan kelompok tersebut disetiap minggunya memiliki nilai yang berbeda seperti pada tabel 6 yaitu produk dengan tipe aki YUASA 1151/N120 di minggu ke 18 melakukan pembelian dan penjualan lebih banyak diantara minggu yang lain.

IV. KESIMPULAN DAN SARAN

Berdasarkan dari hasil penelitian, metode *K-Means* dapat diterapkan untuk melakukan klasterisasi dengan mengubah data menjadi 3 kelompok kemudian kelompok tersebut dapat digunakan untuk mengetahui produk yang diminati, sedang dan kurang diminati. Dari hasil klasterisasi, perusahaan dapat melakukan pengendalian persediaan berdasarkan kelompok produk seperti penjelasan hasil dan pembahasan di atas yaitu produk klaster 0 dapat digunakan sebagai persediaan dengan jumlah sedikit, sedangkan kelompok produk pada klaster 1 digunakan sebagai persediaan dengan jumlah banyak, dan kelompok produk klaster 2 dapat digunakan sebagai persediaan dengan jumlah sedang. Saran untuk pengembangan penelitian selanjutnya dapat dilakukan dengan melakukan perbandingan dengan menggunakan metode klasterisasi lain sebagai perbandingan tingkat akurasi yang lebih akurat.

DAFTAR PUSTAKA

- [1] A. Nugroho, M. N. Al-Azam, A. G. Sooai, and K. Fujisawa, "Benefits Of Enrichment International Publication Programme For Transforming Knowledge and Wisdom Between Researcher," presented at the ASAIHL Conference, Soka University, Hachioji, Tokyo, Japan, 2018, vol. 4, pp. 155–159.
- [2] D. A. Efrilianda and R. R. Isnanto, "Inventory control systems with safety stock and reorder point approach," in *2018 International Conference on Information and Communications Technology (ICOIACT)*, 2018, pp. 844–847.
- [3] M. E. Widiana, K. Hadayati, and A. Nugroho, "Model Enhancement for Batik Artisan Empowerment through Customer Relationship Management Approach based on Information Technology," *IJAME*, 2018.
- [4] P. Ignaciuk, "Discrete-Time Control of Production-Inventory Systems With Deteriorating Stock and Unreliable Supplies," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 45, no. 2, pp. 338–348, Feb. 2015.
- [5] T. Y. Lin, H. H. C. Chuang, and F. Yu, "Tracking Supply Chain Process Variability with Unsupervised Cluster Traversal," in *2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*, 2018, pp. 966–973.
- [6] B. Fang and L. Sun, "Research on inventory collaborative supply for multi-industrial chains based on cloud services platform," in *2017 3rd IEEE*

- International Conference on Computer and Communications (ICCC)*, 2017, pp. 993–999.
- [7] J. Qi, Y. Yu, L. Wang, and J. Liu, “K*-Means: An Effective and Efficient K-Means Clustering Algorithm,” in *2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom) (BDCloud-SocialCom-SustainCom)*, 2016, pp. 242–249.
- [8] M. Danubianu, “StepBy Step Data Preprocessing for Data Mining. A Case Study,” p. 9, 2015.
- [9] J. Yang and J. Wang, “Tag clustering algorithm LMMSK: Improved K-means algorithm based on latent semantic analysis,” *J. Syst. Eng. Electron.*, vol. 28, no. 2, pp. 374–384, Apr. 2017.
- [10] S. Kapil, M. Chawla, and M. D. Ansari, “On K-means data clustering algorithm with genetic algorithm,” in *2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, 2016, pp. 202–206.

