



Implementation of the K-Means Algorithm for Inventory Data Clustering

Implementasi Algoritma K-Means Untuk Clustering Data Inventori

Amelia Rahmadani¹, Nursyahira²

¹Program Studi Sistem Informasi, Fakultas Sains dan Teknologi UIN Sultan Syarif Kasim

²Program Studi Teknik Informatika, Sekolah Tinggi Teknologi (STT) Dumai

E-Mail: 112250320334@students.uin-suska.ac.id, nursyahira608@gmail.com

*Makalah: Diterima 13 Januari 2025; Diperbaiki 25 Januari 2025; Disetujui 5 Februari 2025
Corresponding Author: Amelia Rahmadani*

Abstrak

Manajemen persediaan perangkat teknologi pada sebuah bagian fungsi Shared Service Center Information Communication Technology (SSC ITC) menjadi kunci dalam memastikan ketersediaan barang guna mendukung sistem informasi dan komunikasi perusahaan. Ketidakseimbangan antara stok dan permintaan yang dinamis menjadi tantangan utama dalam manajemen inventori perangkat teknologi. Oleh karena itu, Shared Service Center Information Communication Technology (SSC ICT) berperan penting dalam memastikan ketersediaan barang guna mendukung sistem informasi dan komunikasi perusahaan. Sehingga pengelompokan data perlu dilakukan untuk memahami pola dan hubungan antar data dalam inventori. Penelitian ini bertujuan untuk menerapkan algoritma K-Means dalam mengelompokkan data inventori berdasarkan stok dan durasi penyimpanan perangkat. Hasil analisis menunjukkan bahwa berdasarkan uji coba terhadap *cluster* 2, 3, 4, 5, 6, 7, 8, 9, dan 10, pemilihan K=6 didasarkan pada hasil nilai Davies-Bouldin Index (DBI) terkecil dibandingkan dengan nilai DBI pada K lainnya, yaitu sebesar 0,693, yang mengindikasikan kualitas kluster yang optimal. Dengan demikian, metode ini dapat membantu dalam mengoptimalkan strategi pengelolaan inventori.

Keyword: inventori, k-means, cluster, davies bouldin index (DBI), stok barang

Abstract

Inventory management of technological devices in a part of the Shared Service Center Information Communication Technology (SSC ITC) function is key in ensuring the availability of goods to support the company's information and communication systems. The imbalance between stock and dynamic demand is a major challenge in technology device inventory management. Therefore, the Shared Service Center Information Communication Technology (SSC ICT) plays an important role in ensuring the availability of goods to support the company's information and communication systems. So that data clustering needs to be done to understand the patterns and relationships between data in inventory. This research aims to apply the K-Means algorithm in clustering inventory data based on stock and duration of device storage. The results of the analysis show that based on the trial of clusters 2, 3, 4, 5, 6, 7, 8, 9, and 10, the selection of K=6 is based on the results of the smallest Davies-Bouldin Index (DBI) value compared to the DBI value at other K, which is 0.693, which indicates the optimal cluster quality. Thus, this method can help in optimizing inventory management strategies.

Keyword: inventory, k-means, cluster, davies bouldin index (DBI), stock items

1. PENDAHULUAN

PT Kilang Pertamina Internasional (KPI) RU II, salah satu dari tujuh unit pengolahan di Indonesia, memiliki kapasitas 170.000 barrel per hari, menjadikannya kilang terbesar di Sumatera. Kilang Putri Tujuh Dumai - Sungai Pakning memproduksi BBM dan Non BBM untuk distribusi nasional, mendukung kebutuhan energi dan kemajuan daerah. Untuk mendukung operasional, KPI RU II mendirikan SSC ICT (Shared Service Center Information Communication Technology) untuk meningkatkan efektivitas teknologi informasi dalam proses bisnis. SSC ICT mengelola inventori perangkat teknologi, seperti PC, laptop, printer, scanner, CCTV, kabel jaringan, dan lainnya, dengan total 691 item dari April 2021 hingga Agustus 2024. Meski menggunakan perangkat lunak, tantangan tetap ada dalam menyesuaikan stok dengan kebutuhan dinamis. Tantangan utama dalam pengelolaan inventori adalah ketidakseimbangan antara stok perangkat dan kebutuhan operasional yang fluktuatif, yang dapat mempengaruhi kinerja layanan. Untuk mengatasi hal tersebut, diperlukan teknik

pengelompokan data guna memahami pola distribusi dan hubungan antar data dalam inventori. Dengan mengelompokkan data inventori berdasarkan karakteristik tertentu, perusahaan dapat membuat keputusan yang lebih akurat dalam pengadaan dan penyimpanan barang.

Salah satu metode yang dapat digunakan untuk menghadapi tantangan ini adalah mengelompokkan inventori berdasarkan pola penggunaan dan durasi penyimpanan. Dalam hal ini, cabang kecerdasan buatan dalam bidang ilmu komputer mampu menangani masalah tersebut secara kompleks, salah satunya melalui data mining. Data mining merupakan proses pengolahan data untuk mengidentifikasi pola tertentu dari sekumpulan data [1]. Data yang dianalisis menggunakan data mining akan menghasilkan wawasan baru yang dapat digunakan sebagai acuan dalam membuat keputusan. [2]. Salah satu metode dalam teknik data mining untuk mengurangi permasalahan yang digunakan adalah *Clustering* [3]. Melalui *clustering*, barang-barang dapat dikelompokkan berdasarkan karakteristik penggunaan dan frekuensi durasi penyimpanan, sehingga perusahaan dapat menetapkan stok minimum dengan lebih efektif [4]. Konsep dasar *clustering* adalah mengelompokkan sejumlah objek ke dalam *cluster*, di mana *cluster* yang ideal memiliki tingkat kesamaan yang tinggi di antara objek-objek yang tergabung di dalamnya [5]

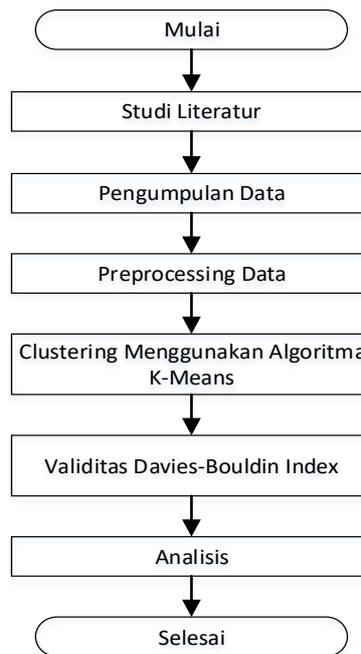
Algoritma *K-Means* adalah salah satu metode yang sering diterapkan dalam proses *clustering*, dikenal sebagai *K-Means Clustering*. Algoritma ini menggunakan pendekatan berbasis jarak dalam data mining untuk melakukan pengelompokan data [6]. Metode *K-Means* membagi data ke dalam beberapa kelompok berdasarkan kesamaan karakteristik, di mana data dengan karakteristik serupa dikelompokkan dalam satu cluster, sedangkan data dengan karakteristik yang berbeda dimasukkan ke *cluster* yang lain [7].

Algoritma *K-Means* dipilih karena menawarkan tingkat akurasi yang tinggi, mudah digunakan, dan efisien dalam menangani data berukuran besar. Keunggulan lainnya adalah fleksibilitasnya, yang memungkinkan pengguna untuk menentukan jumlah *cluster* sesuai dengan kebutuhan [5]. Algoritma *K-Means* telah banyak digunakan untuk mengatasi permasalahan inventori persediaan barang, seperti penelitian yang dilakukan oleh F. Nurdiyansyah mengenai implementasi *K-Means* di Poultry Shop [3]. Penelitian ini memilih *K-Means* untuk mengelompokkan persediaan barang yang ada di CV. Muria PS dengan tiga kategori, yaitu laris, cukup laris dan kurang laris. Tiga *cluster* yang terbentuk terdiri dari *cluster* 1 dengan dua item, *cluster* 2 dengan sembilan item, dan *cluster* 3 dengan 25 item, hasil dari *Clustering* tersebut dapat dijadikan referensi bagi pihak lain untuk mengatur stok persediaan barang. Meskipun begitu, *K-Means* sensitif terhadap outlier, di mana nilai ekstrem dapat memengaruhi hasil pengelompokan [8] dan hasil pengelompokan bersifat tidak selalu unik [1]. Dalam penelitian ini, perhitungan *K-Means* dilakukan menggunakan perangkat lunak RapidMiner, dengan *Davies-Bouldin Index* sebagai tolok ukur untuk mengevaluasi kualitas *cluster* yang terbentuk.

Berlandaskan latar belakang diatas yang maka diterapkan penggunaan data mining dengan menggunakan metode *K-Means Clustering* dengan memanfaatkan data yang langsung diambil dari bagian SSC ICT PT. Kilang Pertamina Internasional (KPI) RU II. Dengan menerapkan analisis *Clustering*, PT. Kilang Pertamina Internasional RU II dapat mengoptimalkan strategi pengadaan barang, memastikan persediaan barang yang sesuai dengan kebutuhan produksi, serta mendukung efisiensi operasional yang lebih baik. Dengan demikian, penelitian ini bertujuan untuk menerapkan algoritma *K-Means* dalam pengelompokan inventaris gudang agar pengelolaan dan penyimpanan barang-barang dapat dikelompokkan dengan baik dan menjadi lebih efisien.

2. METODOLOGI PENELITIAN

Terdapat enam tahapan utama yang dilakukan dalam penelitian ini, yakni tahap perencanaan dan studi literatur, pengumpulan data inventori perangkat teknologi dari SSC ICT PT KPI RU II, dataset yang diperoleh akan melalui proses preprocessing untuk memastikan kualitas data, data yang telah siap kemudian diproses menggunakan algoritma *K-Means* untuk membentuk *cluster* yang diolah menggunakan *tools RapidMiner*, validitas *cluster* yang terbentuk diuji menggunakan *Davies-Bouldin Index (DBI)*, dan hasil analisis klaster serta validitas DBI akan digunakan untuk memberikan rekomendasi pengelolaan inventori yang lebih optimal. Metodologi penelitian dapat dilihat pada Gambar 1.



Gambar 1. Metodologi Penelitian

2.1 Data Mining

Data mining mengolah data mentah yang tersimpan di dalam basis data sehingga menghasilkan informasi yang berguna. Data mentah merupakan data yang disimpan hanya sebagai dokumentasi [3]. Data mining akan mengestrak atau menambang informasi yang berguna dari data yang besar dan kompleks. Ini mencakup teknik statistik, algoritma, dan proses yang digunakan untuk menemukan pola dan hubungan dalam data [9]. Data mining sering disebut dengan istilah Knowledge discovery in database (KDD). Knowledge discovery in database (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar [1].

2.2 Clustering

Clustering adalah metode pembelajaran mesin yang digunakan untuk mengelompokkan data menjadi beberapa kelompok atau klaster yang sesuai. Tujuan dari *Clustering* adalah untuk menemukan struktur di data yang tidak diketahui sebelumnya [9]. *Clustering* berfungsi mengelompokkan beberapa data sehingga menghasilkan satu data yang mirip dengan satu data baru. *Clustering* adalah teknik penambahan data yang banyak digunakan, yang memisahkan sekumpulan titik data menjadi beberapa kelompok atau cluster [10]. Titik data yang terdapat kemiripan karakteristik yang satu sama dengan yang lain berada dalam satu klaster dan tentunya berbeda dengan data yang ada pada klaster [11]. Dengan melakukan *Clustering*, kita dapat menemukan pola yang tidak diketahui yang ada pada data [12]

2.3 K-Means

Pengelompokan *K-Means* adalah jenis pembelajaran tanpa pengawasan (*unsupervised learning*), yang digunakan ketika memiliki data yang tidak berlabel yang dikelompokkan ke dalam grup k atau kluster k , di mana k mewakili jumlah klaster. Dalam pengelompokan *K-Means*, setiap klaster diwakili oleh titik pusat (*centroid*) atau rata-rata yang dan dihitung sebagai nilai rata-rata data dalam klaster tersebut [11]. *K-Means* termasuk kedalam analisis klaster (*cluster analysis*) non hirarki yang digunakan untuk mengelompokkan sejumlah data berdasarkan suatu variabel (*feature*). Tujuan dari *K-Means Clustering* yaitu untuk mendapatkan kelompok data dengan cara memaksimalkan aspek kesamaan karakteristik data dalam suatu klaster dan memaksimalkan perbedaan antar klaster data [13]

Langkah-langkah metode *K-Means Clustering* diantaranya [1]:

- Tentukan K sebagai jumlah *cluster* yang akan dibentuk.
- Inisialisasi K centroid awal (titik pusat *cluster*) bisa dilakukan dengan berbagai cara. Cara yang sering dilakukan adalah dengan random atau acak.
- Alokasikan semua data/objek ke *cluster* terdekat. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat *cluster*. Untuk menghitung jarak semua data ke setiap titik pusat *cluster* dapat menggunakan teori jarak Euclidean dengan persamaan :

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

Dimana $D(i, j)$ = jarak data ke pusat *cluster j*.

X_{ki} = Data ke *i* pada atribut data ke *k*.

X_{kj} = Data ke *j* pada atribut ke *k*.

- d) Kembali ke step 3, apabila masih ada data yang berpindah cluster atau apabila perubahan nilai *centroid*.

2.4 Davies-Bouldin Index (DBI)

Davies-Bouldin Index (DBI) adalah salah satu metode yang digunakan untuk mengukur validitas cluster dalam metode clustering. Jika jarak antar klaster maksimal, itu berarti bahwa kemiripan karakteristik antara masing-masing cluster kecil sehingga perbedaan antar antar cluster tampak lebih jelas. Jika jarak antar-cluster minimum berarti setiap objek dalam cluster memiliki tingkat kemiripan karakteristik yang tinggi [14].

Davies-Bouldin Index (DBI) bekerja dengan menghitung rata-rata dari nilai yang dihitung untuk setiap item dalam kumpulan data. Nilai setiap titik dihitung sebagai jumlah nilai kompak yang dibagi dengan jarak antara dua titik pusat (*centroid*) dari kelompok-kelompok sebagai ukuran pemisahan (*separation*) [15]. Jika DBI bernilai lebih kecil, itu menunjukkan bahwa kita memiliki jumlah cluster yang lebih baik. Nilai DBI yang kecil menandakan bahwa cluster yang dihasilkan sangat kompak di dalamnya dan cukup terpisah dengan cluster yang lain, yang berarti pengelompokan tersebut dilakukan dengan baik [16].

3. ANALISA DAN HASIL

3.1 Pengumpulan Data

Data yang digunakan pada penelitian adalah data inventori bagian SSC ICT PT.Kilang Pertamina Internasional RU II tercatat dari April 2021 s.d Agustus 2024. Data yang diperoleh pada pengumpulan data ini berjumlah 691 record. Adapun atribut dan data yang digunakan dapat dilihat pada Tabel 1 berikut:

Tabel 1. Data Inventori

Id	Nama Barang	Qty	Stok	Satuan	Kategori	Jenis	Tanggal Penyimpanan	Durasi Penyimpanan - Agustus 2024 (Hari)
1	_SPK_CCTV_K UARTA	30		Unit	NET > CCTV & Surveillance Devices	NonConsumable	2021-04-19	1220
2	Telephone Digital Aastra DBC-4223			Unit	NET > Telephone & Tablets	NonConsumable	2021-04-24	1215
3	PHP Application	1			ICT > Web Applications	NonConsumable	2021-04-24	1215
4	ASUS NoteBook K6501Z				COMP > PCs & Notebooks	NonConsumable	2021-06-05	1173
5	_SWITCH_SEW A_GSK	30	1	Unit	NET > Network & Control Devices	NonConsumable	2021-06-05	1173
...
691	Bracket TV Wall Mount 22"-43" Lynk 2205, Ximus	3	2	Pcs	ICT > Generic Materials & Tools	NonConsumable	2024-07-25	27

3.2 Preprocessing Data

Preprocessing data merupakan salah satu tahapan proses yang harus dilakukan untuk memproses data pada proses mining. Kondisi data yang dihasilkan tidak selalu dalam kondisi yang siap untuk diproses. Data tersebut seringkali terdapat berbagai permasalahan yang bisa mengganggu hasil dari proses mining itu sendiri, seperti data yang missing value, outliers, noise, data redundant, ataupun format data yang tidak sesuai dengan sistem [17]. Oleh sebab itu tahap Preprocessing ini dibutuhkan untuk mengatasi permasalahan tersebut.

3.2.1 Data Selection

Seleksi data adalah proses pemilihan data dari kumpulan data operasional yang telah didapatkan dari pemilihan data awal, sehingga dataset tersebut dapat diolah menggunakan *RapidMiner*.

3.2.2 Cleaning Data

Data cleaning merupakan proses untuk dapat mengatasi nilai yang hilang, noise dan data yang tidak konsisten [13]. Pada tahap ini, duplikat, kesalahan dalam data, kekosongan data, dan atribut yang tidak berkorelasi akan dihapus.

Caranya adalah dengan mendeteksi adanya error atau corrupt pada data, kemudian memperbaiki atau menghapus data jika memang diperlukan. Serta menghapus atribut yang tidak memberikan kontribusi yang relevan terhadap analisis klusterisasi. Atribut yang tidak relevan jika dipertahankan dapat meningkatkan kompleksitas analisis tanpa memberikan wawasan tambahan yang bermanfaat. Situasi seperti ini juga memerlukan data cleaning agar tidak muncul masalah yang lebih rumit dalam pengolahan data kedepannya. Pada tahap ini atribut Satuan dan Tanggal Penyimpanan dihapus karena tidak relevan dengan algoritma K-Means. Data yang digunakan setelah dilakukannya proses cleaning ini menjadi 675 record. Data hasil cleaning dapat dilihat pada Tabel 2.

Tabel 2. Hasil Cleaning

Id	Nama Barang	Qty	Stok	Kategori	Jenis	Durasi Penyimpanan - Agustus 2024 (Hari)
1	_SPK_CCTV_KUA RTA	30	0	NET > CCTV & Surveillance Devices	NonConsumable	1221
2	RADIO_TRUNKIN G_NBN	863	787	NET > Radio Communications	NonConsumable	1216
3	HP_IS_BOC	88	5	NET > Telephones & Tablets	NonConsumable	1216
4	Hp Pagewide 586	56	16	COMP > Printers & Scanners	NonConsumable	1174
5	Hp Lasrjet M880 A3	12	7	COMP > Printers & Scanners	NonConsumable	1174
...
675	Switching Power Supply Running Text 200W PS YHY200	0	0	COMP > Sound System & Multimedia Devices	Consumable	405

3.2.3 Transformasi Data

Agar data mudah untuk diolah menggunakan metode algoritma K-Means, maka data inventori yang akan diolah dilakukan dulu proses transformasi data. Tahap ini bertujuan untuk memudahkan proses clustering. Karena algoritma K-Means memerlukan tipe numerik, konversi data pada fase ini digunakan untuk menginisialisasi data yang bernilai normal menjadi angka [18]. Pada tahap ini atribut Nama Barang dirubah special role menjadi id untuk mempertahankan isi data dari atributnya dan atribut Kategori dan Bulan di transformasikan menjadi nilai numerik. Data setelah dilakukan normalisasi dapat dilihat pada Tabel 3.

Tabel 3. Hasil Transformasi

Id	Nama Barang	Qty	Stok	Kategori	Jenis	Durasi Penyimpanan - Agustus 2024 (Hari)
1	_SPK_CCTV_KUARTA	30	0	0	0	1221
2	RADIO_TRUNKING_NBN	863	787	1	0	1216

3	HP_IS_BOC	88	5	2	0	1216
4	Hp Pagewide 586	56	16	3	0	1174
5	Hp Lasrjet M880 A3	12	7	3	0	1174
...
675	Switching Power Supply Running Text 200W PS YHY200	0	0	7	1	405

3.2.4 Normalisasi Data

Normalisasi data merupakan proses atau tahapan dalam mengurangi redundancy data atau meminimalkan anomali yang ada pada data utama sehingga didapatkan data yang lebih sederhana dan lebih akurat untuk digunakan dalam berbagai keperluan yang spesifik, misal: dipakai pada implementasi clustering data [19]. Tujuan dari normalisasi data dalam dataset adalah untuk membentuk data dalam posisi nilai dengan rentang yang sama. Karena algoritma K-Means sedikit sensitif dengan adanya data outlier (pencilan), sehingga dengan dilakukan normalisasi deviasi dari outlier menjadi distribusi data normal [20]. Hasil Normalisasi dapat dilihat pada Tabel 4.

Tabel 4. Hasil Normalisasi

Id	Nama Barang	Qty	Stok	Kategori	Jenis	Durasi Penyimpanan - Agustus 2024 (Hari)
1	Battery 1,5 Volt Size "AAA"	0.055	-0.137	-3.672	- 1.219	0.930
2	Facsimile Continues Roll Paper (Kertas Fac)	6.513	10.471	-2.985	- 1.219	0.916
3	PC Dekstop G4	0.505	-0.070	-2.297	- 1.219	0.916
4	Notebook G4 dan G5	0.257	0.078	-1.069	- 1.219	0.796
5	Barrel Connector Coupler RJ45	-0.084	-0.043	-1.069	- 1.219	0.796
...
675	Switching Power Supply Running Text 200W PS YHY200	-0.177	-0.137	1.141	0.819	-1.394

3.2.5 Outlier Data

Deteksi outlier dilakukan menggunakan *RapidMiner* untuk memastikan bahwa data yang digunakan dalam penelitian ini bebas dari anomali yang dapat mengganggu analisis. Dengan menggunakan operator "*Outlier Detection*" data yang secara signifikan berbeda dari pola umum diidentifikasi dan dihapus. Hasilnya, dari dataset awal, diperoleh 655 data yang telah dibersihkan dari outlier, sehingga lebih representatif dan siap untuk analisis lebih lanjut.

3.3 Klasterisasi Algoritma K-Means

Selanjutnya proses klasterisasi menggunakan software *RapidMiner* dengan menggunakan algoritma *K-Means* menggunakan data yang sudah dilakukan preprocessing data.

3.3.1 Hasil Klasterisasi

Setelah data dikelompokkan menggunakan K-Means, setiap titik data dalam *cluster* memiliki nilai *Euclidean* yang merupakan jarak antara titik tersebut dengan pusat *cluster* (*centroid*). Jarak ini menunjukkan seberapa dekat atau jauh data poin tersebut dengan pusat cluster. Semakin kecil jarak *Euclidean*, semakin dekat data poin dengan centroidnya, yang berarti data tersebut lebih "representatif" dari cluster tersebut. Pada proses klasterisasi menggunakan percobaan *cluster* 2, 3, 4, 5, 6, 7, 8, 9, dan 10 berdasarkan penelitian [8] dapat dilihat pada Tabel 5 berikut:

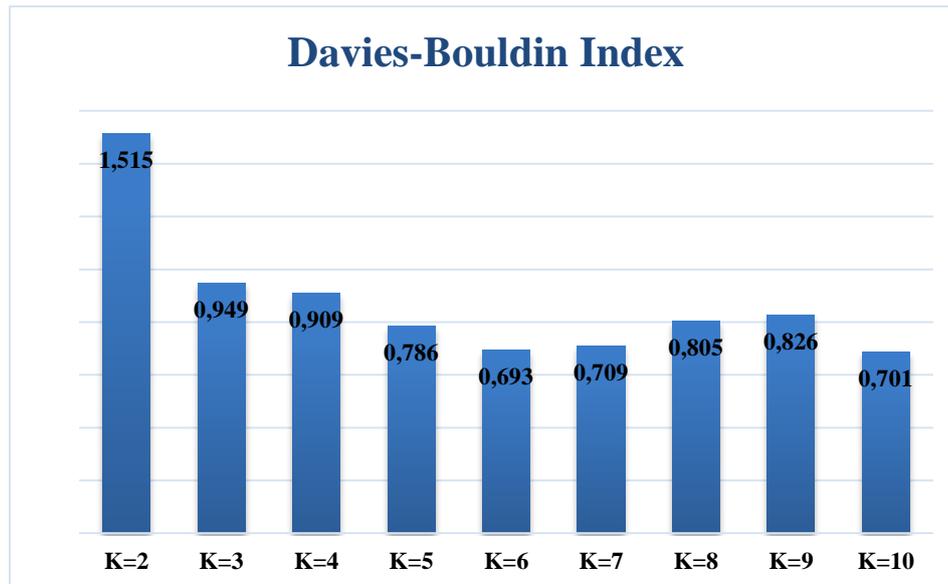
Tabel 5. Hasil Klasterisasi Menggunakan Algoritma K-Means

Percobaan	K0	K1	K2	K3	K4	K5	K6	K7	K8	K9
Nilai K=2	2.100	2.775								
Nilai K=3	1.053	1.238	2.065							

Nilai K=4	0.853	1.259	0.533	1.969						
Nilai K=5	0.890	0.481	0.384	1.128	1.238					
Nilai K=6	0.420	0.481	0.890	1.160	0.668	0.384				
Nilai K=7	0.890	0.271	1.052	0.420	0.378	0.668	0.348			
Nilai K=8	0.420	0.412	1.347	0.348	0.554	0.890	0.201	0.520		
Nilai K=9	0.147	0.554	0.801	0.348	0.420	1.347	0.520	0.373	0.873	
Nilai K=10	0.384	0.857	0.594	0.270	0.668	0.420	0.334	0.707	0.349	0.076

3.3.2 Nilai Validitas Cluster

Selanjutnya dilakukan uji validitas cluster dengan menggunakan teknik DBI. Nilai DBI diambil dari masing-masing hasil Clustering data menggunakan algoritma K-Means, tertera pada Gambar 2 berikut:



Gambar 2. Nilai BDI K-Means

Berdasarkan nilai DBI terbaik, cluster terbaik pada algoritma K-Means terletak pada percobaan K=6 dengan nilai DBI yang diperoleh sebesar 0,693 yang membagi data menjadi 6 Cluster. Nilai DBI yang rendah mencerminkan bahwa cluster-cluster tersebut saling terpisah dengan baik dan memiliki kesamaan tinggi di dalamnya. Setelah melakukan proses analisis data dengan menerapkan metode K-Means Clustering maka didapatkan hasil akhir dari pengelompokan data persediaan barang yang ada di SSC ICT PT.KPI RU II pada tabel dibawah ini:

Tabel 6. Hasil Klusterisasi untuk *cluster_0*

Cluster	Id	Nama Barang	Qty	Stok	Jenis
Cluster_0	1	Krisbow Hand Tool Kit Mechanical set 66pcs,	4	0	NonConsumable
cluster_0	2	Bracket TV Emmy Mount DF 90-T, 65"-90",	2	0	NonConsumable
cluster_0	3	Ausdom Web Camera HD1080P, AF640,	1	1	NonConsumable
cluster_0	4	Puneng Network Multi Purpose, Lines Communication Scanning and Testing Device (Tone Cecker)	6	4	NonConsumable
cluster_0	5	Flike Network IntelliTone Pro 200 Toner and Proble Kit MI8200-60-Kit (Tone Cecker)	3	3	NonConsumable
cluster_0	6	Televisi TCL Google TV 65 Inch, 4KHDRTV	10	4	NonConsumable
...
cluster_0	70	Focusrite Scarlett 2i2, Generation 2 in, 2 out USB Audio Interface soundcard	4	3	NonConsumable

Berdasarkan data hasil klusterisasi yang dapat dilihat pada Tabel 6 menunjukkan bahwa kelompok stok barang yang berada pada *cluster_0* merupakan kategori produk Nonconsumable yang memiliki tingkat penggunaan moderat, dengan persediaan yang tidak memerlukan kapasitas penyimpanan dalam jumlah besar.

Tabel 7. Hasil Klusterisasi untuk *cluster_1*

Cluster	Id	Nama Barang	Qty	Stok	Jenis
cluster_1	1	MPS_HP586/556_981Y_Black	12	0	Consumable
Cluster_1	2	MPS_HP586/556_981Y_Cyan	7	0	Consumable
cluster_1	3	Albion Cable Duct 16x20x300 White	21	0	Consumable
cluster_1	4	Blank Compact Disc Writeable (CD-R)	0	0	Consumable
cluster_1	5	Blank Digital Video Disc Writeable (DVD-R)	0	0	Consumable
cluster_1	6	Clipsal Clamp Conduit @100 Pcs 20mm White	0	0	Consumable
...
cluster_1	328	LED TV 43 Inch	3	0	NonConsumable

Berdasarkan data Tabel 7 menunjukkan bahwa kelompok stok barang yang berada pada *cluster_1* didominasi oleh produk Consumable, yang menunjukkan frekuensi penggunaan tinggi atau tingkat konsumsi yang cepat. Akibatnya, persediaan dalam cluster ini cenderung lebih rendah dibandingkan dengan jumlah awal yang tersedia.

Tabel 8. Hasil Klusterisasi untuk *cluster_2*

Cluster	Id	Nama Barang	Qty	Stok	Jenis
cluster_2	1	ToolBag Kenmaster	30	30	Consumable
cluster_2	2	Connector Port RJ-45 Cat 3, Commscope	95	95	Consumable
cluster_2	3	Dual Port BS Faceplate Kit, Commscope	46	46	Consumable
cluster_2	4	Connector Port Cable RJ-45, Commscope	20	20	Consumable
cluster_2	5	Contact Cleaner Flammable (Mudah terbakar), CRC	7	4	Consumable
cluster_2	6	AMP TE Connectivity SL Series Rj-45 Jack Cat 6, Krone	165	129	Consumable
...
cluster_2	65	Kabel PVC @500M, Supreme	7	1	Consumable

Berdasarkan data Tabel 8 menunjukkan bahwa kelompok stok barang pada *cluster_2* merupakan kategori produk Consumable yang menunjukkan keseimbangan antara jumlah yang tersedia dan tingkat penggunaannya, mencerminkan efisiensi dalam manajemen stok.

Tabel 9. Hasil Klusterisasi untuk *cluster_3*

Cluster	Id	Nama Barang	Qty	Stok	Jenis
cluster_3	1	Hp Pagewide 556	5	4	NonConsumable
cluster_3	2	Scanjet Pro 4500	5	5	NonConsumable
cluster_3	3	_SWITCH_SEWA_GSK	30	1	NonConsumable
cluster_3	4	Hp Pagewide 586	56	16	NonConsumable
cluster_3	5	Ploter A0	2	2	NonConsumable
cluster_3	6	Card Module TMU/12	0	0	NonConsumable
...
cluster_3	60	Telepon Digital DBC 222, Mitel /Aastra	5	1	NonConsumable

Berdasarkan data Tabel 9 menunjukkan bahwa kelompok stok barang pada *cluster_3* merupakan kategori produk Nonconsumable yang disimpan dalam jumlah besar namun dengan tingkat penggunaan yang relatif rendah, sehingga stoknya sering kali berkurang dengan lambat.

Tabel 10. Hasil Klusterisasi untuk *cluster_4*

Cluster	Id	Nama Barang	Qty	Stok	Jenis
cluster_4	1	Toner Cartridge HP LJ 1536 dnf (CE276A)	5	2	NonConsumable
cluster_4	2	ZTE Wifi Model F670L 12 V, 1.5 A	4	0	NonConsumable
cluster_4	3	JOINWIT Optical Power Meter JW3208	4	4	NonConsumable
cluster_4	4	OTB Rackmount 8 Core	20	20	NonConsumable
cluster_4	5	Wall Mounted Cabinet Indorack 5512S	3	1	NonConsumable
cluster_4	6	Fiber Media Converter 10/100Mbps FL-8110 SB-11-20A, Flexreme	3	0	NonConsumable
...
cluster_4	47	Telephone Digital Aastra DBC-4223	5	5	NonConsumable

Berdasarkan data Tabel 10 menunjukkan bahwa kelompok stok barang pada *cluster_4* merupakan kategori produk Nonconsumable yang menunjukkan stabilitas dalam pola penggunaan dan penyimpanan, dengan rasio antara jumlah dan stok yang konsisten.

Tabel 11. Hasil Klusterisasi untuk *cluster_5*

Cluster	Id	Nama Barang	Qty	Stok	Jenis
cluster_5	1	Honeyweel JT-ENCIP Single Panel Enclosure	0	0	Consumable
cluster_5	2	AMP Wiring Management 1U	0	0	Consumable
cluster_5	3	ICT-Digital Satelit Finder	0	0	Consumable
cluster_5	4	TOA M15E Speaker Explotion Proof	4	0	Consumable
cluster_5	5	TOA ZR2015S Megaphone 15 W,	3	2	Consumable
cluster_5	6	Jabra Speak 810 UC Speakerphone for Converncing	2	2	Consumable
...
cluster_5	95	ComfortFit Stereo Headset, A4	13	11	NonConsumable

Berdasarkan data Tabel 11 menunjukkan bahwa kelompok stok barang pada *cluster_5* didominasi oleh produk Nonconsumable, di mana rasio antara jumlah yang tersedia dan stok yang dimiliki tidak menunjukkan variasi yang signifikan, sehingga mencerminkan distribusi dan penggunaan yang seimbang.

4. KESIMPULAN

Dalam penelitian ini, telah dilakukan proses klasterisasi terhadap data inventori bagian SSC ICT pada PT. KPI RU II menggunakan algoritma *K-Means*. Berdasarkan analisis *Davies-Bouldin Index* (DBI), *cluster* terbaik diperoleh pada percobaan $K=6$ dengan nilai DBI sebesar 0,693. Proses klasterisasi ini berhasil mengelompokkan data inventori menjadi 6 cluster, dengan *cluster_0* berjumlah 70 item, *cluster_1* berjumlah 328 item, *cluster_2* berjumlah 65 item, *cluster_3* berjumlah 60 item, *cluster_4* berjumlah 47, dan *cluster_5* berjumlah 95 item. Berdasarkan hasil analisis dan pengolahan data menggunakan algoritma *K-Means*, mencapai kesimpulan sebagai berikut:

1. Algoritma *K-Means* mampu mengelompokkan data inventori dengan efektif, berdasarkan atribut stok dan durasi penyimpanan. Hasil pengelompokan ini memberikan informasi yang terstruktur terkait karakteristik barang, sehingga tujuan utama dalam penelitian ini telah tercapai.
2. Pola yang terbentuk dari hasil klasterisasi menunjukkan kecenderungan barang dengan stok tinggi dan durasi penyimpanan panjang berada dalam kelompok yang berbeda dengan barang dengan stok rendah dan durasi penyimpanan yang lebih pendek. Hal ini mempermudah perusahaan dalam memahami pola distribusi barang dan memperbaiki efisiensi pengelolaan stok.

3. Evaluasi menggunakan Davies-Bouldin Index menunjukkan bahwa nilai Davies-Bouldin Index berada pada tingkat yang rendah maka mengindikasikan kluster yang terbentuk memiliki kualitas yang baik.

Kesimpulan ini mengindikasikan bahwa algoritma K-Means dengan K=6 adalah pilihan terbaik untuk mengelompokkan data inventori di KPI RU II bagian SSC ICT. Temuan ini diharapkan dapat mendukung pengambilan keputusan yang lebih tepat dalam manajemen inventori di KPI RU II bagian SSC ICT.

REFERENSI

- [1] I. Nasution, A. P. Windarto, and M. Fauzan, "Penerapan Algoritma K-Means Dalam Pengelompokan Data Penduduk Miskin Menurut Provinsi," *Building of Informatics, Technology and Science (BITS)*, vol. 2, no. 2, pp. 76–83, Dec. 2020, doi: 10.47065/bits.v2i2.492.
- [2] C. A. Sugianto, A. H. Rahayu, and A. Gusman, "Algoritma K-Means untuk Pengelompokan Penyakit Pasien pada Puskesmas Cigugur Tengah," *Journal of Information Technology*, vol. 2, no. 2, pp. 39–44, Aug. 2020, doi: 10.47292/joint.v2i2.30.
- [3] F. Nurdiansyah and I. Akbar, "Implementasi Algoritma K-Means untuk Menentukan Persediaan Barang pada Poultry Shop," *Jurnal Teknologi dan Manajemen Informatika*, vol. 7, no. 2, pp. 86–94, Dec. 2021, doi: 10.26905/jtmi.v7i2.6377.
- [4] Z. Sitorus, I. Syahputra, C. Indra Angkat, and D. Sartika, "Implementation of K-Means Clustering for Inventory Projection," *International Journal of Science, Technology & Management*, vol. 5, no. 3, pp. 673–678, May 2024, doi: 10.46729/ijstm.v5i3.856.
- [5] R. A. Farissa, R. Mayasari, and Y. Umaidah, "Perbandingan Algoritma K-Means dan K-Medoids Untuk Pengelompokan Data Obat dengan Silhouette Coefficient di Puskesmas Karangsambung," *Journal of Applied Informatics and Computing*, vol. 5, no. 2, pp. 109–116, Oct. 2021, doi: 10.30871/jaic.v5i1.3237.
- [6] Sekar Setyaningtyas, B. Indarmawan Nugroho, and Z. Arif, "TINJAUAN PUSTAKA SISTEMATIS: PENERAPAN DATA MINING TEKNIK CLUSTERING ALGORITMA K-MEANS," *Jurnal Teknoif Teknik Informatika Institut Teknologi Padang*, vol. 10, no. 2, pp. 52–61, Oct. 2022, doi: 10.21063/jtif.2022.v10.2.52-61.
- [7] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan," *Jurnal Tekno Kompak*, vol. 15, no. 2, p. 25, Aug. 2021, doi: 10.33365/jtk.v15i2.1162.
- [8] I. Kamila, U. Khairunnisa, and M. Mustakim, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau," *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 5, no. 1, p. 119, Feb. 2019, doi: 10.24014/rmsi.v5i1.7381.
- [9] M. Rafi Nahjan, Nono Heryana, and Apriade Voutama, "IMPLEMENTASI RAPIDMINER DENGAN METODE CLUSTERING K-MEANS UNTUK ANALISA PENJUALAN PADA TOKO OJ CELL," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 1, pp. 101–104, Jan. 2023, doi: 10.36040/jati.v7i1.6094.
- [10] Md. T. Islam, P. Kumar Basak, P. Bhowmik, and M. Khan, "Data Clustering Using Hybrid Genetic Algorithm with k-Means and k-Medoids Algorithms," in *2019 23rd International Computer Science and Engineering Conference (ICSEC)*, IEEE, Oct. 2019, pp. 123–128. doi: 10.1109/ICSEC47112.2019.8974797.
- [11] A. Supriyadi, A. Triayudi, and I. D. Sholihati, "PERBANDINGAN ALGORITMA K-MEANS DENGAN K-MEDOIDS PADA PENGELOMPOKAN ARMADA KENDARAAN TRUK BERDASARKAN PRODUKTIVITAS," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 6, no. 2, pp. 229–240, Nov. 2021, doi: 10.29100/jupi.v6i2.2008.
- [12] Prihandoko, Bertalya, and M. I. Ramadhan, "An analysis of natural disaster data by using K-means and K-medoids algorithm of data mining techniques," in *2017 15th International Conference on Quality in Research (QiR) : International Symposium on Electrical and Computer Engineering*, IEEE, Jul. 2017, pp. 221–225. doi: 10.1109/QIR.2017.8168485.
- [13] F. P. Ferdy Pangestu, N. Y. Nur Yasin, R. C. Ronald Chistover Hasugian, and Y. Yunita, "Penerapan Algoritma K-Means Untuk Mengklasifikasi Data Obat," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 12, no. 1, pp. 53–62, Mar. 2023, doi: 10.32736/sisfokom.v12i1.1461.
- [14] M. Mughnyanti, S. Efendi, and M. Zarlis, "Analysis of determining centroid clustering x-means algorithm with davies-bouldin index evaluation," *IOP Conf Ser Mater Sci Eng*, vol. 725, no. 1, p. 012128, Jan. 2020, doi: 10.1088/1757-899X/725/1/012128.
- [15] Y. Arie Wijaya, D. Achmad Kurniady, E. Setyanto, W. Sanur Tarihoran, D. Rusmana, and R. Rahim, "Davies Bouldin Index Algorithm for Optimizing Clustering Case Studies Mapping School Facilities," *TEM Journal*, pp. 1099–1103, Aug. 2021, doi: 10.18421/TEM103-13.

- [16] G. Ghufron, B. Surarso, and R. Gernowo, "The Implementations of K-medoids Clustering for Higher Education Accreditation by Evaluation of Davies Bouldin Index Clustering," *Jurnal Ilmiah Kursor*, vol. 10, no. 3, Nov. 2020, doi: 10.21107/kursor.v10i3.232.
- [17] H. Syah and A. Witanti, "ANALISIS SENTIMEN MASYARAKAT TERHADAP VAKSINASI COVID-19 PADA MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)," *Jurnal Sistem Informasi dan Informatika (Simika)*, vol. 5, no. 1, Feb. 2022, doi: 10.47080/simika.v5i1.1411.
- [18] D. Putri Adilah Asih, B. Irawan, and A. Bahtiar, "PENGELOMPOKAN DATA TRANSAKSI DALAM MENENTUKAN STRATEGI PENJUALAN MENGGUNAKAN ALGORITMA K-MEANS," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 1, pp. 141–147, Feb. 2024, doi: 10.36040/jati.v8i1.8321.
- [19] A. N. Sihananto, A. Puspita Sari, H. Khariono, R. Akhmad Fernanda, and D. Cakra Mudra Wijaya, "Implementasi Metode K-Means Untuk Pengelompokan Kasus Covid-19 Tingkat Provinsi Di Indonesia," *Jurnal Informatika dan Sistem Informasi*, vol. 3, no. 1, pp. 76–85, Apr. 2022, doi: 10.33005/jifosi.v3i1.472.
- [20] Ahmad Harmain, P. Paiman, H. Kurniawan, K. Kusriani, and Dina Maulina, "NORMALISASI DATA UNTUK EFISIENSI K-MEANS PADA PENGELOMPOKAN WILAYAH BERPOTENSI KEBAKARAN HUTAN DAN LAHAN BERDASARKAN SEBARAN TITIK PANAS," *TEKNIMEDIA: Teknologi Informasi dan Multimedia*, vol. 2, no. 2, pp. 83–89, Jan. 2022, doi: 10.46764/teknimedia.v2i2.49.