

Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing Studi Kasus: Hoyweapstore

Deni Triyansyah¹ , Devi Fitriana²

Teknik Informatika, Universitas Mercu Buana, Jakarta
41514010005@student.mercubuana.ac.id¹ , devi.fitriana@mercubuana.ac.id²

Abstrak

Penelitian ini mencari strategi marketing yang mampu bersaing dengan kompetitor lain, karena pada era saat ini setiap perusahaan saling belomba dalam menjual produknya dengan ujung tombaknya adalah strategi marketing, masalah yang diangkat pada penelitian ini adalah menumpuknya stok sepatu yang belum terjual dan kurang diperhatikannya reseller yang memiliki potensial dalam penjualan. penelitian ini menggunakan metode K-Means, adalah metode clustering yang membagi data kedalam satu atau lebih kluster, data yang karakteristik sama dikelompokkan satu kluster yang sama dan karakteristik yang berbeda dikelompokkan kedalam kelompok yang lain. K-Means adalah metode yang cocok digunakan dengan satu tipe data dan ditunjang dengan validitas kluster yang menggunakan Davies-Bouldin Index untuk mengetahui seberapa baik proses kinerja pengklusteran. Hasil pengklusteran data untuk mendapatkan informasi reseller yang potensial, dengan menggunakan dua kluster, sedangkan untuk mendapatkan informasi tipe sepatu yang laris adalah dengan menggunakan tiga cluster. Dan pengklusteran ini menghasilkan 1 reseller yang potensial yaitu reseller Hoy, sedangkan proses mencari tipe sepatu yang laris mendapatkan 3 tipe sepatu yang terlaris yaitu Neo Adventage, Cloudfoam Swiftt, dan Bennasi. Hasil dari nilai Davies Bouldin Index yang diperoleh pada clustering pencarian reseller yang potensial adalah 0.271 dan 0.196 pada klustering pencarian tipe sepatu yang laris.

Keywords : Data Mining, strategi marketing, klustering, Metode K- Means, Rapid Miner, Online Shop

Received Juli 2018

Accepted for Publication September 2018

DOI: 10.22441/incomtech.v8i2.4174

1. PENDAHULUAN

Di era saat ini, banyak orang – orang yang melakukan kegiatan wirausaha, dan menjadikannya pekerjaan utamanya dikarenakan jika berwirausaha akan mendapatkan hasil yang tidak terbatas dan mendapatkan kebebasan tanpa adanya peraturan dari pimpinan Serta semakin banyak yang melakukan kegiatan wirausaha, salah satu bentuk wirausaha pada era saat ini adalah Online Shop.

Online Shop adalah suatu proses pembelian barang atau jasa dari mereka yang menjual melalui internet, atau layanan jual-beli secara online tanpa harus bertatap muka dengan penjual atau pihak pembeli secara langsung[1]. Hoyweapstore adalah salah satu online shop di Indonesia yang bergerak dibidang penjualan sepatu sneakers yang berdiri sejak tanggal 9 November 2014 yang berlokasi di Jl.H.Saba Komplek DPR 3 No.1 Meruya Selatan Kembangan Jakarta Barat.

Dalam bisnisnya Hoyweapstore berkembang cukup pesat sejak 3 tahun silang, akan tetapi tidak lepas dari masalah yang ada seperti banyaknya stok tipe sepatu yang belum terjual dan beberapa reseller yang potensial tetapi tidak diperhatikan. Dalam masalah ini pihak hoyweapstore masih melakukan atau mencoba beberapa stategi marketing untuk meningkatkan penjualan dan meminimalisir masalah yang ada.

Berdasarkan masalah tersebut, perlu adanya solusi yang tepat untuk permasalahan yang sedang dihadapi dalam menentukan perencanaan strategi marketing yang tepat. solusi yang digunakan untuk mewujudkannya adalah data mining. Data mining adalah suatu proses kegiatan menganalisa data guna untuk menemukan suatu pola dari sebuah kumpulan data[2], dan metode yang digunakan adalah metode clustering. Metode klustering adalah suatu proses pengelompokan objek data yang mirip satu sama lain kedalam kluster yang sama dan berbeda dengan objek yang ada di kluster lain[3]. Algoritma yang digunakan dalam penelitian ini untuk menentukan stategi marketing adalah algoritma K-Means.

Algoritma K-Means merupakan algoritma yang menjadi salah satu algoritma yang paling penting dalam bidang data mining, Karena Memiliki kelebihan sebagai algoritma yang mudah diimplementasikan, relatif cepat ditinjau dari waktu komputasi dan telah digunakan secara luas untuk menyelesaikan berbagai persoalan komputasi dan Algoritma K-Means bisa dibidang adalah metode pengelompokan yang paling populer[4], dan jika menggunakan metode k-means, harus melakukan transmisi data agar dapat diolah, maka data yang berjenis data nominal seperti tipe sepatu, warna, dan reseller harus diinisialisasikan terlebih dahulu kedalam bentuk angka[5]. dan jika data sudah dalam bentuk angka, nantinya akan diproses menggunakan tools Rapidminer.

Dengan dilakukannya penelitiannya ini diharapkan dapat membantu pihak Hoyweapstore dalam menentukan strategi marketing yang tepat untuk meningkatkan laba penjualan dan mampu bersaing dengan kompetitor.

2. STUDI LITERATUR

Pada langkah ini penulis melakukan studi literatur guna untuk menambah teori-teori dalam melakukan penelitian.

strategi adalah sebuah rencana(Plan) yang disatukan, menyeluruh, dan terpadu yang menghubungkan strategi perusahaan dengan tantangan lingkungan dan yang dirancang untuk memastikan bahwa tujuan utama perusahaan dapat dicapai melalui pelaksanaan yang tepat dalam perusahaan, Willy Pratama Widharta dan Sugiono Sugiharto(2013)[6]. Menurut Dimas Hendika, Zainul Arifin, dan Sunarti. Strategi marketing adalah suatu cara untuk mendapatkan kemenangan keunggulan bersaing yang berkesinambungan, baik perusahaan yang memproduksi barang ataupun jasa.[7]

Menurut Aditya Wardhana dan Mahir Pradana, Online shop adalah pembelian suatu barang yang dilakukan melalui internet sebagai media pemasaran dengan penggunaan situs (website) sebagai katalog[8]. dan menurut Asmai Ishak, Aktivitas Online shop merupakan cara baru dalam melakukan transaksi, dalam aktivitas ini pembeli dan penjual tidak bertemu secara langsung seperti halnya proses penjualan biasa, Efisiensi waktu dan juga efektifitas dalam proses transaksi merupakan hal yang membuat konsumen tertarik untuk melakukan belanja online[9]. Elly Muningsih dan Sri Kiswati mengatakan, Lembaga atau perusahaan yang mengaplikasikan E-Commerce dalam kegiatan penjualan atau perdagangan, biasa dikenal dengan nama Toko Online atau Online Shop[10].

Menurut Asmita Yadav dan Sandeep kumar Singh, dan data yang digunakan adalah data Bank. dalam penelitian, mereka mengurangi kompleksitas komputasi algoritma dengan menggunakan kembali data iterasi sebelumnya dalam iterasi saat ini atau berikutnya untuk mengelompokkan data-objek. Dalam algoritma ini, centroid juga diperlakukan sebagai data objek dalam kluster masing-masing, yang memberikan bantuan untuk menghindari masalah kluster kosong dan memberikan solusi yang optimal dari masalah karena semua data-objek disimpan dalam cluster terdekatnya, tujuan dari penelitiannya adalah peningkatan akurasi, tingkat presisi dan efisiensi algoritma[11].

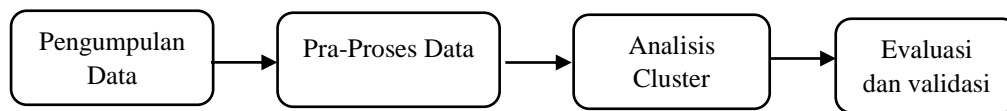
Totok Suprawoto, dalam penelitian ini , peneliti menggunakan data mahasiswa angkatan 2014/2015. Peneliti menggunakan metode k-means untuk untuk mengelompokkan data kategorikal hingga menghasilkan kluster yang lebih stabil, selanjutnya peneliti menganalisis lebih mendalam terhadap hasil klusterisasi tersebut untuk membantu bagian admisi dalam menentukan kebijakan mengenai promosi program studi di lingkungan STMIK AKAKOM[12].

Dalam penelitian Elly Muningsih dan Sri Kiswati mereka menggunakan data produk dari online shop. menurut penelit, K-Means merupakan salah satu metode terbaik dan paling populer dalam algoritma clustering dimana K-Means mencari partisi yang optimal dari data dengan meminimalkan kriteria jumlah kesalahan kuadrat dengan prosedur iterasi yang optimal[10].

Menurut Putri Elfa Mas'udia mengatakan, Data mining adalah proses menemukan hubungan dalam data yang tidak diketahui oleh pengguna dan menyajikannya dengan cara yang dapat dipahami sehingga hubungan tersebut dapat menjadi dasar pengambilan keputusan[13]. menurut Elly Muningsih dan Sri Kiswati, data mining Data mining merupakan sebuah inti dari proses atau Knowledge Discovery in Database(KDD), meliputi dugaan algoritma yang mengeksplor data, membangun model dan menemukan pola yang belum diketahui, penelitian ini menggunakan algoritma k-Means Clustering dengan data yang digunakan yaitu data produk dari online shop. peneliti melakukan pengolahan data dengan algoritma k-Means bertujuan untuk mengelompokkan produk yang dijual pada online shop menjadi beberapa kluster untuk mengetahui produk mana yang paling diminati sehingga jumlah stok harus banyak[10].

3. METODE

Metode Penelitian yang digunakan dalam penerapan algoritma K-Means untuk menentukan Strategi marketing pada online shop Hoyweapstore, menggunakan rancangan penelitian yang ditunjukkan pada Gambar 1.



Gambar 1 Rancangan Penelitian

a. Data

Data yang digunakan dalam penelitian akhir ini adalah data transaksi penjualan yang diperoleh dari online shop Hoyweapstore, Data-data tersebut berisi data transaksi penjualan sepatu seperti Tanggal, Tipe Sepatu, Warna dan Reseller. Jumlah data yang digunakan dalam penelitian ini, berjumlah 1548 data transaksi dari bulan Juli 2017 sampai Februari 2018.

Dalam penelitian ini metode pengumpulan data yang penulis gunakan adalah sebagai berikut :

1. Wawancara, metode pengumpulan data dengan melakukan tanya jawab langsung dengan tatap muka dengan pemilik online shop Hoyweapstore
2. Dokumentasi, metode dengan melakukan pengumpulan data atau dokumen, sampai data terkumpul menjadi satu dan siap untuk kepentingan penelitian.
3. Studi literature, adalah mempelajari dan memahami karya ilmiah ataupun sumber ilmiah yang lainnya yang sesuai dengan penelitian dan memiliki hubungan penelitian yang serupa.

Tabel 1. Sampel tabel transaksi penjualan online shop Hoyweapstore.

TANGGAL	TIPE SEPATU	WARNA	RESELLER
30/06/2017	VANS SLIP ON	PUTIH	HOY
30/06/2017	VANS OLD SKOOL	HITAM PUTIH	HOY
30/06/2017	VANS OLD SKOOL	HITAM PUTIH	1HOY
30/06/2017	AIR MAX ZERO MIDNIGHT	HITAM PUTIH	7DEADLYSN EAKERS
30/06/2017	VANS AUTHENTIC	NAVY	HOY
30/06/2017	VANS AUTHENTIC	NAVY	HOY
30/06/2017	VANS AUTHENTIC	NAVY	DISKON STORE
30/06/2017	VANS OLD SKOOL	HITAM PUTIH	HOY
30/06/2017	NEO ADVANTAGE	ABU-ABU	SOLUP CITY
30/06/2017	NMD	HITAM PUTIH	HOY

b. Pra-Proses Data

Pra-Proses Data adalah salah satu langkah paling penting dalam proses data mining yang berkaitan dengan persiapan dan transformasi dataset awal.

- Transformasi Data

Transformasi Data adalah proses perubahan data, menjadi data yang dapat diolah menggunakan algoritma yang akan dipakai apakah dalam bentuk numerik, klasifikasi dan lain-lain[14]. Agar data di atas dapat diolah dengan menggunakan algoritma K-Means, maka data yang berjenis data nominal seperti nama, tipe sepatu, Warna, dan reseller diinisialisasikan terlebih dahulu dalam bentuk angka.

Tabel 2.Sampel tabel Pra-Proses data tipe sepatu

Tipe Sepatu	Transformasi
VANS SLIP ON	1
VANS OLD SKOOL	2
VANS OLD SKOOL	2
VANS OLD SKOOL	2
AIR MAX ZERO MIDNIGHT	3
AIR MAX ZERO MIDNIGHT	3
VANS SLIP ON	3
VANS OLD SKOOL	2
NEO ADVANTAGE	4
NMD	2

Tabel 3.Sampel tabel Pra-Proses data warna

Warna	Transformasi
HITAM PUTIH	1
HITAM PUTIH	1
PUTIH	2
HITAM PUTIH	1
HITAM PUTIH	1
HITAM PUTIH	1
HITAM PUTIH	1
PUTIH	2
HITAM PUTIH	1
HITAM PUTIH	1

Tabel 4.Sampel tabel Pra-Proses data reseller

Reseller	Transformasi
HOY	1
HOY	1
HOY	1
7DEADLYSNEAKERS	2
HOY	1
HOY	1
DISKON STORE	3
HOY	1
SOLUP CITY	4
HOY	1

Tabel 5.Sampel hasil tabel Pra-Proses data transaksi penjualan online shop Hoyweapstore

TANGGAL	TIPE SEPATU	WARNA	RESELLER
30/06/2017	1	1	1
30/06/2017	2	1	1
30/06/2017	2	2	1
30/06/2017	2	1	2
30/06/2017	3	1	1
30/06/2017	3	1	1
30/06/2017	3	1	3
30/06/2017	2	2	1
30/06/2017	4	1	4
30/06/2017	2	1	1

c. Analisis Kluster (Algoritma K-Means)

K-Means merupakan salah satu metode data clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster/kelompok. Metode ini mempartisi data ke dalam cluster/kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain[15].

Langkah-langkah dalam melakukan clustering dengan metode K-Means[15], adalah sebagai berikut:

1. Tentukan Jumlah Kluster k.
2. Menentukan Titik pusat awal dari setiap kluster
3. Alokasikan semua data/ objek ke kluster terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke kluster tertentu ditentukan jarak antara data dengan pusat cluster. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat kluster. Jarak paling antara satu data dengan satu kluster tertentu akan menentukan suatu data masuk dalam kluster mana. Untuk menghitung jarak semua data ke setiap titik pusat kluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut:

$$D(x,y) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2}$$

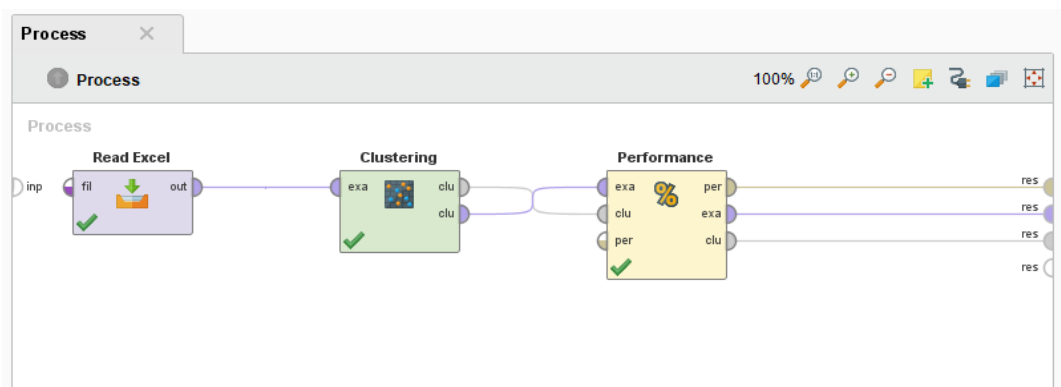
4. Hitung kembali pusat kluster dengan keanggotaan cluster yang sekarang. Pusat cluster adalah rata-rata dari semua data/ objek dalam cluster tertentu. Jika dikehendaki bisa juga menggunakan median dari kluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.
5. Tugaskan lagi setiap objek memakai pusat kluster yang baru. Jika pusat cluster tidak berubah lagi maka proses klustering selesai. Atau, kembali ke langkah nomor 3 sampai pusat kluster tidak berubah lagi.

Penerapan Algoritma K-Means

Penerapan proses algoritma K-Means menggunakan aplikasi RapidMiner studio versi 8.2.001 yang berguna dalam membantu melakukan kegiatan seperti data mining, Klustering, klasifikasi, dan regresi[16]. dilihat pada gambar 2.

- Read Excel

adalah operator ini membaca ExampleSet dari file Excel yang ditentukan.



Gambar 2. Proses clustering di aplikasi rapidminer

- **Klustering**
adalah operator yang melakukan pengelompokan menggunakan algoritma k-means.
- **Performance**
adalah operator yang digunakan untuk evaluasi kinerja metode pengelompokan berbasis centroid. Operator ini memberikan daftar nilai kriteria kinerja berdasarkan centroid kluster.
- **Proses Pengolahan Data**
Pada langkah ini penulis melakukan dua kali proses pengklusteran, proses yang pertama dibagi menjadi tiga cluster dengan menggunakan 3 jenis data yaitu tanggal, tipe sepatu dan warna. dan proses yang kedua menjadi 2 kluster dengan jenis data tanggal, tipe sepatu dan reseller. Setelah semua data telah ditransformasi kedalam bentuk angka, setelah itu data-data tersebut telah dapat dikelompokkan dengan menggunakan algoritma K-Means Clustering.

d. Validitas Kluster

Validitas Kluster, adalah sebuah proses evaluasi pengujian yang memberikan daftar nilai kriteria kinerja berdasarkan centroid kluster. yang berguna untuk mengetahui seberapa baik kinerja dari proses pengklusteran. Pada penelitian ini, evaluasi kluster menggunakan Davies Bouldin Index pada perangkat Rapidminer.

Hasil akhir yang didapat pada proses evaluasi Kluster adalah sebuah nilai, yang dimana jika nilai yang dihasilkan oleh Davies Bouldin Index semakin kecil maka akan semakin baik evaluasi cluster tersebut, dan sebaliknya jika Davies Boludin Index menghasilkan nilai yang semakin tinggi maka proses evaluasi tesebut semakin buruk atau kurang baik.

4. HASIL DAN PEMBAHASAN

- **Hasil proses klustering untuk mendapatkan informasi reseller yang memiliki potensial dalam poenjualan**

Selanjutnya pengolahan data untuk mendapatkan informasi reseller yang memiliki potensial dalam penjualan. proses kali ini hanya menggunakan dua kluster dan atribut data yang digunakan adalah tanggal, tipe data dan reseller. berikut pada tabel 6 adalah titik pusat awal kluster.

- Pertama tentukan jumlah cluster.
- Tentukan titik centroid pada setiap kluster

Tabel 6. titik pusat awal kluster

	Tanggal	Tipe Sepatu	Reseller
Kluster 1	30/06/2017	256	1
Kluster 2	02/10/2017	1	1

- Selanjutnya adalah menghitung jarak terdekat, dikarenakan metode K-Means mengalokasikan setiap data kedalam suatu kluster dari data yang memiliki jarak paling dekat dengan titik pusat dari setiap kluster, maka dari itu perlu dihitung jarak setiap data dengan menggunakan rumus Encludian Distance, sebagai contoh akan dihitung jarak data pertama ke pusat kluster pertama:

$$D(1,1) \sqrt{(1 - 256)^2 + (1 - 1)^2} = 65.025$$

Dari hasil perhitungan di mendapatkan hasil bahwa jarak data pertama dengan pusat cluster pertama adalah 65.025

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

Dari hasil perhitungan di mendapatkan hasil bahwa jarak data pertama dengan pusat cluster kedua adalah 0

Berdasarkan hasil kedua perhitungan di atas bahwa jarak data kedua yang paling dekat adalah dengan cluster 2, sehingga data pertama dimasukkan ke dalam Kluster 2, Sampel Hasil perhitungan selengkapnya untuk 10 data pertama dapat di lihat pada tabel 7.

Tabel 7. Sampel hasil pengelompokan iterasi pertama

No	Tanggal	Tipe Sepatu	Reseller	K1	K2	Kluster
1	30/06/2017	1	1	65025	0	2
2	30/06/2017	2	1	64516	1	2
3	30/06/2017	2	1	64516	1	2
4	30/06/2017	2	1	64516	1	2
5	30/06/2017	3	1	64009	4	2
6	30/06/2017	3	1	64009	4	2
7	30/06/2017	3	1	64009	4	2
8	30/06/2017	2	1	64516	1	2
9	30/06/2017	4	1	63504	9	2
10	30/06/2017	2	1	64516	1	2

Tabel 8. Sample jumlah item dari iterasi pertama dari setiap kluster

KLUSTER 1	KLUSTER 2
288 item	1260 Item

- Hitung kembali pusat kluster dengan keanggotaan Kluster yang sekarang. Pusat Kluster adalah rata-rata dari semua data/ objek dalam Kluster tertentu.

Tabel 9. Sample titik pusat awal kluster yang baru

Kluster	Tipe Sepatu	Warna
K1	190,4166667	5,826388889
K2	42,05555556	3,647619048

- Selanjutnya melakukan proses iterasi yang kedua, pada tabel 10.

Tabel 10. Sampel hasil pengelompokkan iterasi kedua

No	Tanggal	Tipe Sepatu	Reseller	K1	K2	Kluster
1	30/06/2017	1	1	35901,96764	1692,568529	2
2	30/06/2017	2	1	35524,13431	1611,457417	2
3	30/06/2017	2	1	35524,13431	1611,457417	2
4	30/06/2017	2	1	35524,13431	1611,457417	2
5	30/06/2017	3	1	35148,30097	1532,346306	2
6	30/06/2017	3	1	35148,30097	1532,346306	2
7	30/06/2017	3	1	35148,30097	1532,346306	2
8	30/06/2017	2	1	35524,13431	1611,457417	2
9	30/06/2017	4	1	34774,46764	1455,235195	2
10	30/06/2017	2	1	35524,13431	1611,457417	2

Tabel 11. Sample titik pusat awal kluster yang baru pada iterasi kedua

Kluster	Tipe Sepatu	Reseller
K1	183,646875	5,678125
K2	39,95358306	3,629478827

- Tugaskan kembali setiap objek memakai pusat kluster yang baru. Jika pusat cluster tidak berubah lagi maka proses clustering selesai. Atau, kembali ke langkah nomor 3 sampai pusat kluster tidak berubah lagi, Setelah melakukan proses iterasi 2,3,4,5,6,7 dan 8. dapatlah pada iterasi ke 8 pada tabel 12 :

Tabel 12. Sampel hasil pengelompokkan iterasi ke 8

No	Tanggal	Tipe Sepatu	Reseller	K1	K2	Kluster
1	30/06/2017	1	1	31154,82024	1388,26826	2
2	30/06/2017	2	1	30802,91739	1314,937709	2
3	30/06/2017	2	1	30802,91739	1314,937709	2
4	30/06/2017	2	1	30802,91739	1314,937709	2
5	30/06/2017	3	1	30453,0	1243,60	2

	017			1453	7158	
6	30/06/2017	3	1	30453,0 1453	1243,60 7158	2
7	30/06/2017	3	1	30453,0 1453	1243,60 7158	2
8	30/06/2017	2	1	30802,9 1739	1314,93 7709	2
9	30/06/2017	4	1	30105,1 1167	1174,27 6607	2
10	30/06/2017	2	1	30802,9 1739	1314,93 7709	2

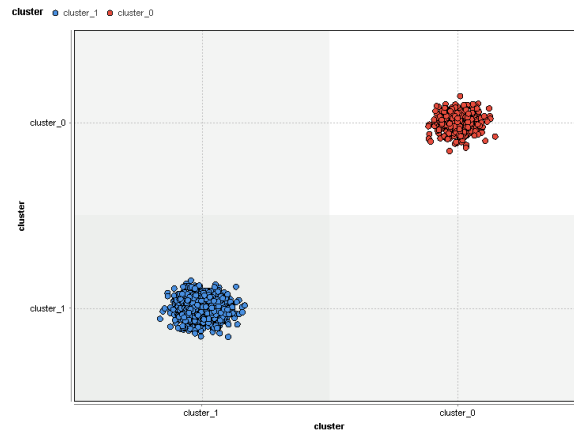
- Dalam penelitian ini terjadi 8 kali iterasi, pada proses iterasi dari 7 ke 8 tidak adanya anggota yang berpindah dari kluster satu kekluster yang lain atau pusat kluster tidak berubah, sehingga proses iterasi selesai. dan hasil dari proses iterasi kedelapan dengan menggunakan atribut data tanggal, tipe sepatu, dan reseller menghasilkan sekelompok data dari setiap masing-masing kluster, dengan kluster 1 menghasilkan sekelompok data yang berjumlah 350 item dan kluster 2 berjumlah 1198 item. Sampel hasil pengelompokkan dapat dilihat ditabel 13.

Tabel 13. Sampel hasil pengelompokkan iterasi ke 8

Kluster 1	Kluster 2
<p>Kluster 1 berjumlah 350 item, yang terdiri dari :</p> <p>Hoy = 309 Dope Homestore = 9 Masrqces = 4 Jt = 2 Creackstore = 2</p> <p>Bagian proses kluster 1 ini, Hoy adalah reseller yang memiliki potensial dalam penjualan, sedangkan Dope Homestore dengan reseller yang lainnya tidak memiliki potensial dalam penjualan. Karena bisa dilihat pada list diatas, lebih dari 50% Hoy melakukan melakukan transaksi penjualan.</p>	<p>Kluster 2 berjumlah 1198 item, yang terdiri dari :</p> <p>Hoy = 1100 7deadlysneakers = 9 Marqces = 8 Creackstore = 5 VGASV = 4</p> <p>Untuk proses kluster 2, Reseller yang memiliki potensial dalam penjualan adalah reseller dengan nama Hoy, sedangkan 7deadlysneakers dengan reseller yang lainnya tidak memiliki potensial dalam penjualan. bisa dilihat pada list diatas, lebih dari 50% Hoy melakukan penjualan.</p>

- Hasil klustering pengolahan data pertama dengan Rapidminer dilihat dari Plot view di gambar 3.

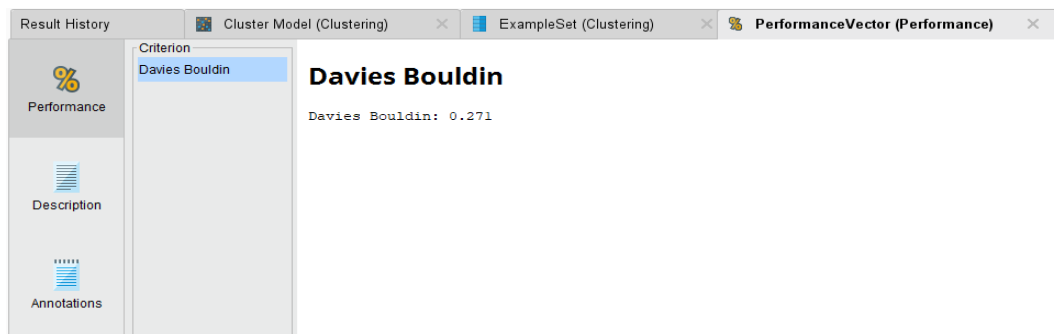
Ini adalah sheet untuk menampilkan representasi dari database yang telah diolah secara keseluruhan lengkap dengan klusternya. Urutan dari clusternya yaitu cluster 0=1 dengan warna merah yang memiliki 350 item dalam kluster, dan dalam kluster ini item dengan nama reseller Hoy adalah item yang memiliki potensial dalam penjualan, bias dilihat pada tabel 13 bahwa transaksi penjualan Hoy lebih dari 50%.



Gambar 3. Tampilan ExampleSet Plot View

Sedangkan cluster 1=2 dengan warna biru adalah representasi dari cluster 2 yang berjumlah 1198 item, dalam cluster ini item dengan nama Hoy adalah item yang melakukan transaksi penjualan lebih dari 50%, yang artinya reseller Hoy ini memiliki potensial dalam penjualan. dan bisa dilihat pada tabel 13 hasil dari klustering, dimana reseller yang potensial dikluster satu dan cluster dua, adalah reseller dengan nama Hoy.

- Hasil kinerja cluster menggunakan Davies Bouldin Index, pada gambar 4.



Gambar 4. Tampilan nilai Davies Bouldin menggunakan Rapidminer

Pada proses ini validitas cluster menggunakan Davies Bouldin menghasilkan nilai 0.271, yang dimana nilai ini dihasilkan dari kinerja kedua cluster. yang dimana jika dari evaluasi tersebut akan menghasilkan sebuah parameter berupa nilai, semakin kecil nilai tersebut maka semakin baik cluster tersebut.

- **Proses pengolahan data untuk mendapatkan informasi tipe sepatu yang laris**

Pada proses pengolahan data kali ini, peneliti menggunakan tiga cluster dan dengan menggunakan atribut data tanggal, tipe sepatu, dan warna. berikut pada table 14 adalah titik pusat awal cluster.

Tabel 14. titik pusat awal cluster

	Tanggal	Tipe Sepatu	Warna
Kluster 1	30/06/2017	1	1
Kluster 2	02/10/2017	256	2
Kluster 3	17/072017	128	7

- Selanjutnya adalah menghitung jarak terdekat, sebagai contoh akan dihitung jarak data pertama ke pusat kluster pertama:

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

Dari hasil perhitungan di mendapatkan hasil bahwa jarak data pertama dengan pusat kluster pertama adalah 0

$$D(1,2) = \sqrt{(1-256)^2 + (1-2)^2} = 65.026$$

Dari hasil perhitungan di mendapatkan hasil bahwa jarak data pertama dengan pusat kluster kedua adalah 665.026

$$D(1,3) = \sqrt{(1-128)^2 + (1-7)^2} = 16.165$$

Dari hasil perhitungan di mendapatkan hasil bahwa jarak data pertama dengan pusat kluster ketiga adalah 16.165

Berdasarkan hasil ketiga perhitungan di atas bahwa jarak data pertama yang paling dekat adalah dengan kluster 1, sehingga data pertama dimasukkan ke dalam kluster 1, Sampel Hasil perhitungan selengkapnya untuk 10 data pertama dapat di lihat pada tabel 15.

Tabel 15. Sampel hasil pengelompokan iterasi pertama

No	Tanggal	Tipe Sepatu	Warna	K1	K2	K3	Kluster
1	30/06/2017	1	1	0	65026	16165	1
2	30/06/2017	2	1	1	64517	15912	1
3	30/06/2017	2	2	2	64516	15901	1
4	30/06/2017	2	1	1	64517	15912	1
5	30/06/2017	3	1	4	64010	15661	1
6	30/06/2017	3	1	4	64010	15661	1
7	30/06/2017	3	1	4	64010	15661	1
8	30/06/2017	2	2	2	64516	15901	1
9	30/06/2017	4	1	9	63505	15412	1
10	30/06/2017	2	1	1	64517	15912	1

Tabel 16. Sample jumlah item dari iterasi pertama dari setiap kluster

KLUSTER 1	KLUSTER 2	KLUSTER 3
846 item	145 Item	557 Item

- Hitung kembali pusat kluster dengan keanggotaan cluster yang sekarang. Pusat cluster adalah rata-rata dari semua data/ objek dalam cluster tertentu

Tabel 17. Sample titik pusat awal kluster yang baru

Kluster	Tipe Sepatu	Warna
K1	16,25886525	5,828605201
K2	222,8482759	7,282758621
K3	110,8833034	6,635547576

- Selanjutnya melakukan proses iterasi yang kedua, pada tabel 18.

Tabel 18. Sampel hasil pengelompokan iterasi kedua

No	Tanggal	Tipe Sepatu	Warna	K1	K2	K3	Kluster
1	30/06/2017	1	1	256,1483969	49256,13056	12106,09977	1
2	30/06/2017	2	1	226,6306664	48813,43401	11887,33316	1
3	30/06/2017	2	2	217,973456	48801,86849	11877,06206	1
4	30/06/2017	2	1	226,6306664	48813,43401	11887,33316	1
5	30/06/2017	3	1	199,1129359	48372,73746	11670,56655	1
6	30/06/2017	3	1	199,1129359	48372,73746	11670,56655	1
7	30/06/2017	3	1	199,1129359	48372,73746	11670,56655	1
8	30/06/2017	2	2	217,973456	48801,86849	11877,06206	1
9	30/06/2017	4	1	173,5952054	47934,0409	11455,79994	1
10	30/06/2017	2	1	226,6306664	48813,43401	11887,33316	1

Tabel 19. Sample titik pusat awal kluster yang baru pada iterasi kedua

Kluster	Tipe Sepatu	Warna
K1	16,25886525	5,828605201
K2	212,9528796	6,848167539
K3	104,5029354	6,739726027

- Tugaskan kembali setiap objek memakai pusat kluster yang baru. Jika pusat cluster tidak berubah lagi maka proses clustering selesai. Atau, kembali ke langkah nomor 3 sampai pusat kluster tidak berubah lagi, Setelah melakukan proses iterasi 2,3,4,5,6,7. dapatlah pada iterasi ke 7 pada tabel 20 :

Tabel 20. Sampel hasil pengelompokan iterasi ke 7

No	Tanggal	Tipe Sepatu	Warna	K1	K2	K3	Kluster
1	30/06/2017	1	1	213,3086122	42452,17272	9765,07622	1
2	30/06/2017	2	1	186,7310786	42041,26574	9568,764936	1
3	30/06/2017	2	2	178,102263	42030,37737	9558,426414	1
4	30/06/2017	2	1	186,7310786	42041,26574	9568,764936	1
5	30/06/2017	3	1	162,153545	41632,35877	9374,453652	1
6	30/06/2017	3	1	162,153545	41632,35877	9374,453652	1
7	30/06/2017	3	1	162,153545	41632,35877	9374,453652	1
8	30/06/2017	2	2	178,102263	42030,37737	9558,426414	1
9	30/06/2017	4	1	139,5760114	41225,45179	9182,142368	1
10	30/06/2017	2	1	186,7310786	42041,26574	9568,764936	1

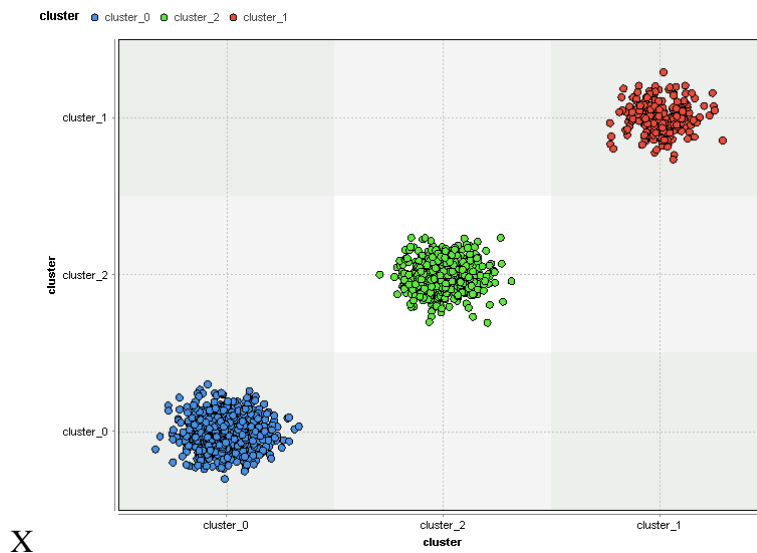
- Pada proses ini terjadi 7 kali iterasi, dimana dari proses iterasi 6 keiterasi 7 tidak adanya anggota yang berpindah dari satu kluster ke kluster yang lain atau pusat kluster tidak berubah, sehingga proses iterasi selesai. dan dari proses ini dengan menggunakan 3 kluster menghasilkan kelompok dari masing-masing kluster dengan menggunakan atribut tanggal, tipe sepatu, dan warna. Dan dapat dilihat pada tabel 21, kluster 1 berjumlah 819 item dengan tipe sepatu Neo Advantage yang memiliki jumlah terbesar dengan warna putih dan didominasi warna hitam, dan kluster 2 berjumlah 215 item dengan tipe yang memiliki jumlah terbesar adalah Cloudfoam Swiftt dengan warna hitam, sedangkan kluster 3 berjumlah 514 item dengan tipe sepatu Bennasi warna hitam didominasi warna putih.

Tabel 21. Sampel hasil pengelompokan iterasi ke 7

Hasil Kluster 1	Hasil Kluster 2	Hasil Kluster 3
Kluster 1 berjumlah 819 item, yang terdiri dari : Neo Advantage = 180 NMD = 119 Vans Old Skool = 86 Adidas Superstar = 69 Vans Authentic = 44 Yeezy = 38 Vans Slip On = 30 Flyknit Racer = 30 Ultraboost 3.0 = 29 Air Jordan Retro = 26 Pada kluster ini Neo advantage	Kluster 2 berjumlah 215 item, yang terdiri dari : Cloudfoam Swiftt = 20 Neo Coderby = 13 Cloudfoam Super Flex = 9 Cosmic 2 = 9 Neo Baseline = 7 Cosmic = 6 Asics Reightning Champ = 5 Converse Ct Hi = 3 Puma Rihanna = 2 Air Max Command = 1 Pada kluster 2 ini Cloudfoam Swiftt dengan warna Hitam adalah tipe sepatu yang paling laris.	Kluster 3 berjumlah 514 item, yang terdiri dari : Bennasi = 177 Neo Cloudfoam = 29 Yeezy Pirate Black = 16 HUARACHE = 16 Yeezy kids t-dove = 11 Neo Baseline = 11 Ultraboost = 11 Yeezty Kids Pirate = 8 Asics Mooncrater = 7 Puma Bow = 6 Dan pada kluster 3 sepatu yang terlaris adalah sepatu

<p>adalah sepatu yang terlaris dengan warna putih dan dominasi warna hitam.</p>		<p>dengan tipe Bennisasi warna hitam dengan dominasi warna putih</p>
---	--	--

- Hasil klustering pengolahan data pertama dengan Rapidminer dilihat dari Plot View di gambar 5.



Gambar 5. Tampilan ExampleSet Plot View

Ini adalah sheet untuk menampilkan representasi dari database yang telah diolah secara keseluruhan lengkap dengan klusternya. Urutan dari klusternya yaitu cluster 0=1 dengan warna biru yang memiliki 819 item dalam kluster, dan dalam kluster ini item dengan nama tipe sepatu neo advantage adalah tipe sepatu yang paling laris dengan warna putih dan dominasi warna hitam.

Sedangkan cluster 1=2 dengan warna merah adalah representasi dari kluster 2 yang berjumlah 215 item, dalam kluster ini tipe sepatu yang laris adalah tipe sepatu dengan nama cloudfoam swiftt dengan warna Hitam.

Sedangkan cluster 2=3 dengan warna hijau adalah representasi dari kluster 3 yang berjumlah 514 item, dengan tipe sepatu yang laris adalah bennisasi dengan warna hitam dengan dominasi warna putih. dari hasil klustering diatas dapat dilihat di tabel 21, bahwa sepatu yang sering terjual atau terlaris pada kluster satu adalah tipe sepatu Neo Advantage, dan kluster kedua adalah tipe sepatu

Cloudfoam Swiftt, sedangkan pada kluster ketiga tipe sepatu yang sering terjual adalah tipe Bennisasi.

- Hasil kinerja kluster menggunakan Davies Bouldin Index, pada gambar 6.



Gambar 6. Tampilan nilai Davies Bouldin menggunakan Rapidminer

Davies bouldin ini bekerja untuk mengevaluasi kinerja dari proses pengklusteran, yang dimana jika dari evaluasi tersebut akan menghasilkan sebuah parameter berupa nilai, semakin kecil nilai tersebut maka semakin baik kluster tersebut, dari hasil pengklusteran diatas menghasilkan nilai Davies Bouldin sebesar 0.196.

5. IMPLIKASI

1. Dari hasil poin satu menghasilkan strategi marketing, yaitu dengan melakukan pengelompokan atau paket penjualan dan diskon, misalnya sepatu A dan b adalah sepatu yang paling diminati dan sepatu c dan d kurang diminati, maka memberlakukan sistem paket dengan membeli sepatu A dan c akan mendapatkan diskon, begitupun sebaliknya dengan sepatu b dan d.
2. Selanjutnya ketika melakukan pemotretan produk. potret produk yang laku dengan yang kurang laris secara bersamaan, fungsinya untuk produk tersebut cepat dilihat oleh konsumen. Agar konsumen dapat mempertimbangkan dalam membeli produk yang mana yang akan dibeli.
3. Dalam Penelitian ini menghasilkan informasi reseller yang memiliki potensial dalam penjualan, Hal yang harus dilakukan adalah mengelola reseller dengan melakukan edukasi penjualan melalui internet dan terus diberi arahan dalam berjualan.
4. Berikan reseller harga special untuk setiap produk yang dibeli.
5. Dengan adanya penelitian ini, semoga pihak online shop hoyweapstore dapat terbantu dalam meningkatkan omset penjualan dan mampu bersaing oleh competitor lain.

6. KESIMPULAN DAN STUDI LANJUTAN

• Kesimpulan

Berdasarkan penelitian yang dilakukan mendapatkan sebuah hasil dalam menentukan strategi marketing pada online shop Hoyweapstore, maka dapat disimpulkan :

1. Berdasarkan pengolahan data dengan menggunakan atribut tanggal, tipe sepatu, warna, dan reseller, menghasilkan 3 tipe sepatu yang diminati oleh konsumen dengan tipe sepatu Neo Adventage, Floodfoam Swiftt, dan Bennasi. serta menghasilkan 1 reseller yang potensial dalam penjualan yaitu.
2. Berdasarkan penelitian klustering menggunakan algoritma K-Means dan dibantu dengan perangkat lunak Rapidminer, yang didapat pada proses klustering untuk mendapatkan informasi reseller yang potensial dengan menggunakan 2 kluster menghasilkan 8 kali iterasi dengan tidak berubahnya titik pusat kluster atau berpindahnya anggota dari satu kluster ke kluster yg lain serta menghasilkan nilai validitas dengan Davies Bouldin Index sebesar 0.271. sedangkan proses klustering untuk mendapatkan informasi tipe sepatu yang laris dengan digunakanya 3 kluster menghasilkan 7 kali iterasi dengan ditandainya tidak ada berubahnya titik pusat kluster atau berpindahnya anggota dari satu kluster ke kluster yg lain dan proses ini menghasilkan nilai Davies Bouldin Index sebesar 0.196.

• Studi Lanjutan

Keterbatasan yang ada pada studi ini mengakibatkan diperlukannya studi lanjutan untuk mendukung serta menyempurnakan temuan studi yang didapatkan pada studi ini.

Studi lanjutan yang disarankan untuk dilakukan adalah :

1. Studi klustering terbaik dengan mengkomparasi algoritma K-Means dengan algoritma Hybrid K-Means KNN.
2. Studi mendapatkan informasi terbaik dan terupdate dengan menambah jumlah data dengan menggunakan data terbaru.

REFERENCES

- [1] C. A. Sari, "Perilaku Berbelanja Online Di Kalangan Mahasiswi Antropologi Universitas Airlangga," *AntroUnairdotNet*, vol. 4, no. 2, pp. 205–216, 2015.
- [2] G. Gunadi and D. I. Sensuse, "Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth) :," *Telematika*, vol. 4, no. 1, pp. 118–132, 2012.
- [3] U. R. Raval and C. Jani, "Implementing & Improvisation of K-means Clustering Algorithm," *Int. J. Comput. Sci. Mob. Comput.*, vol. 55, no. 5, pp. 191–203, 2016.
- [4] T. M. Kodinariya and P. R. Makwana, "Review on determining number of Cluster in K-Means Clustering," *Int. J. Adv. Res. Comput. Sci. Manag. Stud.*, vol. 1, no. 6, pp. 2321–7782, 2013.
- [5] J. O. Ong, "Implementasi Algoritma K-means clustering untuk menentukan strategi marketing president university," *J. Ilm. Tek. Ind.*, vol. vol.12, no. no. juni, pp. 10–20, 2013.
- [6] W. P. Widharta and S. Sugiharto, "Penyusunan strategi dan sistem penjualan dalam rangka meningkatkan penjualan toko damai," *J. Manaj. Pemasar. Petra*, vol. 2, no. 1, pp. 1–15, 2013.
- [7] D. H. Wibowo, Z. Arifin, and Sunarti, "Analisis Strategi Pemasaran Untuk Meningkatkan Daya Saing UMKM (Studi pada Batik Di Jeng Solo)," *J. Adm. Bisnis*, vol. 29, no. 1, pp. 59–66, 2015.
- [8] A. Wardhana, "Viral Marketing Determinants of Top Online Shop Brands in Indonesia," *Mimbar*, vol. 32, no. 1, pp. 25–30, 2016.
- [9] A. Ishak, "Analisis Kepuasan Pelanggan Dalam Belanja Online : Sebuah Studi Tentang Penyebab (Antecedents) Dan Konsekuensi (Consequents)," *J. Siasat Bisnis*, vol. 16, no. 2, pp. 141–154, 2012.
- [10] E. Muningsih and S. Kiswati, "Penerapan Metode K-Means Untuk Clustering Produk Online Shop Dalam Penentuan Stok Barang," *J. Bianglala Inform.*, vol. 3, no. 1, pp. 10–17, 2015.
- [11] J. Wang, "An improved K-Means clustering algorithm," *Technology*, vol. 5, no. m, pp. 4–6, 2011.
- [12] T. Suprawoto, "Klasifikasi Data Mahasiswa Menggunakan Metode K-Means untuk Menunjang Pemilihan Strategi Pemasaran," *J. Inform. dan Komput.*, vol. 1, no. 1, pp. 12–18, 2016.
- [13] P. N. Malang, "Klasifikasi Tugas Akhir Untuk Menentukan Dosen Pembimbing Menggunakan Naïve Bayes Classifier (Nbc)," vol. 7, pp. 135–140, 2015.
- [14] D. Mining and K. Clustering, "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Diare Di Kabupaten Langkat," *Cogito Smart J.*, vol. 2, no. 2, pp. 108–119, 2016.
- [15] Y. Agusta, S. Bali, and B. Denpasar, "K-Means – Penerapan, Permasalahan dan Metode Terkait," *J. Sist. dan Inform.*, vol. 3, no. Pebruari, pp. 47–60, 2007.
- [16] A. Chisholm, *Exploring Data with RapidMiner*. 2013.

