Analisis RFM untuk Menentukan Indeks Produk pada Permainan Hay Day dengan Algoritma K-means

Muhammad Ainul Yaqin [1)],Muhammad Adib zamzam [1)], Berlian Gita Cahyani [1)], Silva Ahmad Zaky Zamani [1)]

[1)]Teknik Informatika, Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
yaqinov@ti.uin-malang.ac.id

Abstrak. *Permainan Hay Day merupakan salah satu permainan yang banyak diminati dari berbagai kalangan. Tidak hanya menghadirkan hiburan tapi juga pembelajaran dalam mengelola produk ataupun menjualnya. Hasil penjualan dari produk tersebutlah yang nantinya akan digunakan sebagai modal untuk mengembangkan bisnis kita pada permainan Hay Day. Tujuan penelitian ini yaitu mengelompokkan produk - produk dalam game ke cluster yang dapat dinilai tingkatnya sehingga diketahui produk-produk yang berpotensi bagus. Langkah awal yaitu mengumpulkan data transaksi penjualan produk, selanjutnya data prepocessing dengan memilih data yang layak. Data kemudian menjadi acuan menghitung nilai RFM (Recency, Frequency dan Monetary) dari produk. Kemudian menerapkan algoritma K-Means untuk menghasilkan clustering produk. Dari 75 produk yang menjadi cakupan, hasil yang diperoleh dari analisis dan pembahasan yaitu terdapat 4 cluster dimana cluster 0 Superstar terdiri dari 4 produk, cluster 1 Bronze terdiri dari 18 produk, cluster 2 Gold terdiri dari 23 produk, cluster 3 Silver terdiri dari 12 produk.*

***Kata kunci**: Analisis; RFM; Hay Day; Clustering; K-means;*

1. Pendahuluan

Hayday merupakan salah satu permainan simulasi pertanian, peternakan dan bisnis di platform Mobile (Android / iOS). Dari sisi bisnis pemain bisa mendapatkan keuntungan (gold dan exp) dimana keuntungan tersebut didapat dari pengolahan pertanian, peternakan dan hasil-hasil produksi lain.

Game ini memiliki banyak fitur yang terbagi dalam 5 kategori, yaitu bangunan dan perkebunan utama, bangunan penyimpanan, hasil panen, binatang binatang, termasuk binatang peliharaan, dan bangunan untuk produksi. Setiap kategori memiliki fitur yang beragam yang dapat dimainkan oleh player. Namun, tidak semua fitur dapat dimainkan di saat pertama kali menginstall game ini. Fitur-fitur tersebut akan terbuka seiring dengan naiknya level seorang player. Player harus menaikkan levelnya sampai yang tertinggi agar bisa memainkan semua fitur pada game Hayday dan membuka semua produk yang ada pada game tersebut. Semakin tinggi level sorang player, maka semakin beragam pula permintaan produk dari para pelanggan. Meskipun begitu, banyaknya jenis pesanan tidak selalu dapat memberikan keuntungan yang signifikan. Beberapa produk memiliki nilai jual yang tinggi akan tetapi produk tersebut jarang dipesan. Sedangkan produk yang nilai jualnya murah, seringkali lebih banyak dipesan oleh pelanggan. Oleh karena itu segmentasi produk diperlukan. Selain untuk mengetahui produk mana saja yang laris, segementasi dilakukan untuk mengetahui produk mana saja punya potensi bagus untuk kedepannya.

Penggunaan teknik data mining merupakan salah satu solusi untuk persoalan segmentasi produk. Data mining diharapkan dapat membantu proses pengambilan keputusan yang tepat, memungkinkan pemain untuk mengelola data yang dikumpulkan, data warehouse, atau tempat penyimpanan lainnya menjadi sebuah informasi dan pengetahuan (knowledge) yang baru. Data mining dapat mengekstrak pengetahuan berharga dari sejumlah data yang diatur oleh pemahaman manusia 0 mining penting untuk pengembangan bisnis.

Clustering merupakan teknik data mining yang membagi data dalam suatu himpunan ke dalam beberapa kelompok yang mana kesamaan data dalam suatu kelompok lebih besar dibandingkan kesamaan data tersebut dengan data dalam kelompok lain [2]. Semakin kecil jumlah cluster yang digunakan, maka akan mempermudah pemahaman terhadap struktur data di dalamnya, tetapi pola informasi penting yang ada akan terabaikan. Sebaliknya dengan memberikan jumlah cluster akhir yang

lebih besar, maka kesamaan yang dimiliki antar kelompok akan semakin meningkat pula dan mengabaikan struktur data yang ada [4].

Salah satu metode clustering yang sangat populer dan banyak dipelajari untuk meminimalkan kesalahan clustering untuk titik ruang Euclidean disebut K-Means clustering [5]. Metode K-Means berusaha mengelompokkan data yang ada kedalam beberapa kelompok dimana data dalam satu kelompok memiliki karakteristik yang berbeda antara satu dengan yang lain.

Penelitian lebih mengutamakan pada produk-produk yang ada di level 1-30 yang berjumlah total 75 produk. Data diambil selama 5 jam dengan interval waktu 30 menit.

Model RFM

1. Kriteria R (recency)

Recency berarti skala kekinian sebuah produk diukur ketika produk ditransaksikan atau dipesan.

2. Kriteria F (frequency)

Membutuhkan atribut yang merepresentasikan berapa kali produk dipesan. Kriteria ini ini dapat dilihat dari berapa banyak produk dengan nama yang sama muncul dalam data histori. Atribut yang dibutuhkan adalah atribut jumlah total setiap produk yang dipesan selama periode analisis.

3. Kriteria M (monetary)

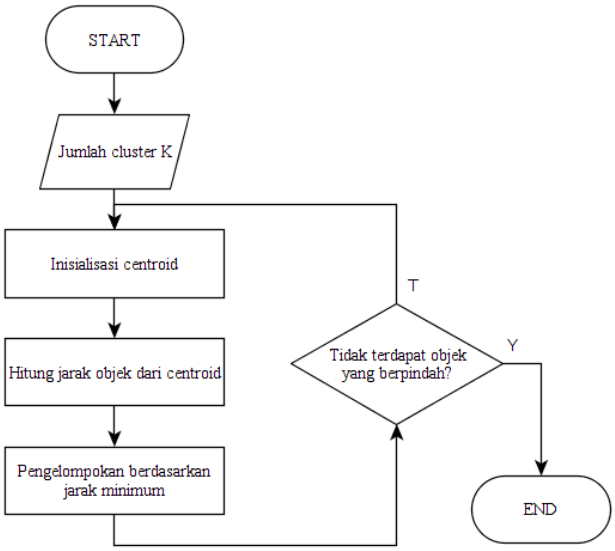
Atribut terakhir dari RFM adalah *monetary* (terlihat pada kolom kelima tabel 3). Atribut ini berhubungan dengan harga atau keuntungan yang didapat bila barang tersebut terjual.

Data Mining

Teknik data mining telah diterapkan pada berbagai domain. Aktivitas transaksi pada sebuah organisasi yang terjadi hampir setiap hari, membuat data transaksi menjadi sangat banyak dan kompleks. Dengan teknik data mining, khususnya teknik clustering, dapat diterapkan untuk membagi seluruh pelanggan / produk menjadi beberapa kelompok (cluster) berdasarkan beberapa kesamaan pada pelanggan / produk.

Algoritma K-Means

Algoritma K-Means adalah algoritma klasik untuk memecahkan masalah clustering [6]. Metode K-Means merupakan metode non-hirarkis. Metode ini merupakan teknik penyekatan (partition) yang membagi atau memisahkan objek k ke daerah yang terpisah [2]. Metode K-Means digunakan untuk mengelompokkan n vector berdasarkan atribut partisi k, dimana k <n, tergantung pada beberapa tindakan. Pusat cluster adalah mean (nilai tengah) semua vector pada cluster tertentu. Algoritma ini dimulai dengan memilih centroid k awal secara acak (randomly), kemudian memberikan nilai vektor ke centroid terdekat dengan Euclidean distance dan menghitung ulang centroid baru. Proses ini berulang sampai vector tidak lagi mengubah cluster antar iterasi [7]. Alur algoritma K-Means ditunjukkan pada gambar 1.



Gambar 1. Alur algoritma K-Means

Persamaan euclidean distance ditunjukkan pada persamaan 1.

$$dist = \sqrt{\sum_{k=1}^{n} (p_k - q_k)^2} \quad (1)$$

Elbow Method

Menurut Kodinariya dan Makwana [2], elbow method digunakan untuk menentukan jumlah cluster dari dataset. Ide dasar dari unsupervised model (misal: K-means clustering) adalah untuk menentukan cluster sehingga total intra-cluster variation (dikenal sebagai total within-cluster variation atau total within-cluster sum of square) dimini- malkan. Metode ini merupakan metode visual. Idenya ada lah dimulai dengan k = 2, dan terus meningkat dalam setiap langkah dengan ditambah 1 pada nilai k. Pada nilai k=3, apabila terjadi perubahan drastis yang berbanding terbalik dengan nilai sebelumnya, maka nilai sebelum terjadinya perubahan tersebut dianggap sebagai jumlah cluster yang paling tepat. Semakin kecil nilai *within sum of squares* maka *cluster* yang terbentuk semakin baik [2]. Rumus dari *within sum of squares* (SSE) pada K-Means adalah seperti berikut :

$$W_k = \sum_{r=1}^{k} \frac{1}{n_r} D_r \quad (2)$$

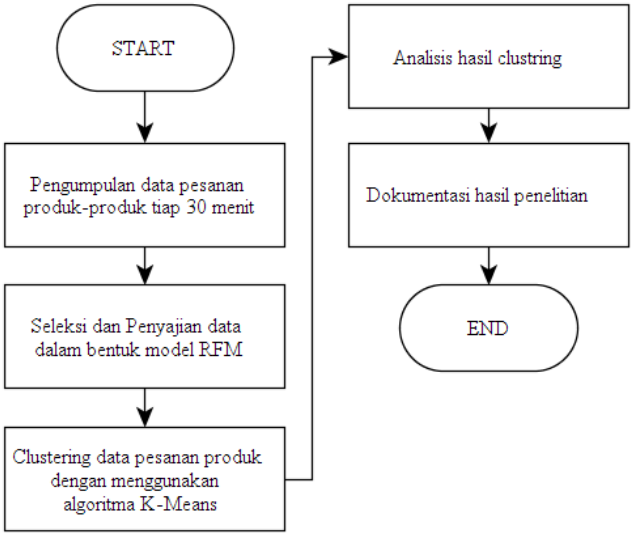
$$D_r = \sum_{i=1}^{n_r - 1} \sum_{j=i}^{n_r} \left((d_i - d_j) \right)^2 \quad (3)$$

Adapun menurut Bholowalia dan Kumar [7] algoritma metode elbow dalam menentukan nilai k untuk K-Means adalah:

1. Inisialisasi awal nilai K
2. Mulai
3. Peningkatan nilai K
4. Mengukur SSE dari tiap nilai K
5. Melihat hasil SSE dari nilai K yang turun secara drastis
6. Tetapkan nilai K yang berbentuk siku
7. Selesai.

Metodologi penelitian

Alur Penelitian ditunjukkan pada gambar 2:



Gambar 2. Alur Penelitian

Penelitian dimulai dengan mengumpulan data pembelian setiap 30 menit dengan total waktu 5 jam, hal ini bertujuan untuk mendapatkan data yang dibutuhkan dalam kurun waktu yang singkat. Selanjutnya data tersebut diseleksi data yang memiliki indeks 0 akan dihapus, dan data yang telah disortir ditampilkan dalam bentuk RFM. Langkah berikutnya Clustering data menggunakan algoritma K-Means sehingga didapat sejumlah cluster dengan jumlah yang optimal. Data cluster kemudian dianalisis untuk mengetahui sebuah cluster termasuk pada label apa. Kemudian data & hasil analisis didokumentasikan.

Penelitian terkait

Clustering dapat digunakan ke berbagai subjek yang memiliki banyak dimensi penilaian, seperti klasterisasi pemain hockey yang cenderung memiliki nilai yang sama di beberapa aspek. Clustering dilakukan dengan mengambil nilai hit, push, tapping serta data Multi level running speed dan Sprint 50 meter. Metode evaluasi cluster yang digunakan menggunakan DBI dan purity. Hasil penelitian menunjukkan bahwa pengelompokan dengan-kmeans menghasilkan cluster yang masih belum maksimal. Hasil ini dapat dipengaruhi berbagai hal diantaranya adalah proses pengukuran data (tolok ukur) yang mengakibatkan nilai diantara pemain hampir sama [7][9].
Penelitian clustering juga mencakup subjek bidang bisnis, seperti pada penelitian Adiana[Adiana]. Penelitian dilakukan dengan melakukan clustering terhadap konsumen tetap (pelanggan) berdasarkan pada aspek Recency (nilai skala kekinian), Frequency (jumlah total transaksi), Monetary (jumlah total uang transaksi) dari masing-masing pelanggan. Hasil penelitiannya yaitu pelanggan terkelompokkan menjadi tiga cluster yang membantu *decision making* [10].

2. Pembahasan

Menurut metodologi penelitian, maka penelitian diawali dengan persiapan labeling per cluster dan pengumpulan data. Labeling yang dibuat :

Tabel 1. Karakteristik Produk

Kategori	R	F	M
Superstar	Tinggi	Tinggi	Tinggi
Gold	Tinggi/Sedang	Tinggi/Sedang	Tinggi/Sedang
Silver	Sedang/Rendah	Sedang/Rendah	Sedang/Rendah
Bronze	Rendah	Rendah	Rendah

Persiapan dan Analisis RFM

Data yang digunakan untuk pemodelan clustering adalah data rincian produk yang tersedia pada saat berada di level 30 pada game HayDay dan data histori produk yang dipesan. Data diambil secara manual yang dituliskan dalam Microsoft Excel yang kemudian dilakukan pemilihan atribut yang disesuaikan dengan kebutuhan kriteria model RFM, yaitu rentang waktu periode analisis, jumlah frekuensi produk yang dipesan selama periode analisis, serta jumlah nominal untuk setiap produk selama periode analisis [11]. Berikut uraian mengenai atribut yang dibutuhkan untuk model RFM :

1. Kriteria R (recency)

   Recency berarti skala kekinian sebuah produk diukur ketika produk ditransaksikan atau dipesan. Di penelitian ini penilaian menggunakan skala 1-4 secara berurutan. Semakin kecil nilai skala R maka produk tersebut semakin usang dari segi pemesanan. Detail penilaian Recency ditampilkan pada tabel 2.

Tabel 2. Karakteristik Produk

R	Rentang waktu
1	>3 jam terakhir
2	3 jam terakhir
3	1 jam terakhir
4	30 menit terakhir

2. Kriteria F (frequency)

Membutuhkan atribut yang merepresentasikan berapa kali produk dipesan. Kriteria ini ini dapat dilihat dari berapa banyak produk dengan nama yang sama muncul dalam data histori. Atribut yang dibutuhkan adalah atribut jumlah total setiap produk yang dipesan selama periode analisis. Misalkan pada contoh (Tabel 3), produk dengan id:1 memiliki nama *Wheat* memiliki frekuensi=11, maka data tersebut berarti produk *Wheat* tersebut telah dipesan sebanyak 11 kali selama periode analisis (150 menit). Begitu juga dengan produk-produk lainnya seperti *Egg*, memiliki F sebanyak 28, *Corn*, memiliki F sebanyak 132, dan seterusnya.

3. Kriteria M (monetary)

Atribut terakhir dari RFM adalah *monetary* (terlihat pada kolom kelima tabel 3). Atribut ini berhubungan dengan harga atau keuntungan yang didapat bila barang tersebut terjual. Sehingga, bila produk *Wheat* terjua, maka poin keuntungan yang kita dapatkan dalam game HayDay tadi adalah 33, untuk produk *Egg* adalah 504, produk *Corn* adalah 924, dan seterusnya. Berikut adalah cuplikan data yang telah melewati tahap pengumpulan data analisis RFM:

TABEL 3. Cuplikan Detail RFM masing-masing produk

ID Produk	Produk	R	F	M
1	Wheat	3	11	33
2	Egg	4	28	504
3	Corn	4	132	924
4	Bread	2	3	63
6	Soybean	1	1	10

ID Produk	Produk	R	F	M
8	Milk	1	1	32
9	Cream	3	3	150
10	Sugarcane	1	2	28
11	Corn Bread	2	3	216
12	Brown Sugar	1	3	96

Pra Proses Data

Setelah data RFM tersaji, langkah berikutnya adalah melakukan Pra Proccessing data yang terdiri dari 3 proses yaitu data cleaning, data reduction, dan data transformation dengan Log10 dan normalisasi Min-Max. Pra proccess dilakukan agar clustering menjadi lebih akurat. Data hasil pra proses data terlihat pada Tabel 4.

Data reduction

Data reduction bertujuan untuk menghasilkan tabel input clustering dengan cara menghilangkan record serta kolom-kolom atribut yang tidak dibutuhkan. Pengurangan kolom dilakukan untuk mengurangi volume data agar tidak terlalu besar dan agar clustering dapat dilakukan dengan mudah. Berikut atribut yang di proses.

Data cleaning

Data cleaning merupakan proses pembersihan data dari item-item dengan nilai yang kosong. Penyesuaian data dilakukan untuk memastikan seluruh data produk aktif sehingga total keuntungannya tidak ada yang bernilai 0.

Data transformation

Atribut yang ada pada data mentah sebagian besar berisi keterangan yang tidak dapat digunakan di dalam proses clustering secara langsung. Oleh karena itu, perlu dilakukan transformasi data yang bertujuan mengubah data mentah ke dalam bentuk yang lebih terukur sehingga dapat digunakan sebagai atribut clustering [11]. Berikut adalah cuplikan hasil pra prosesing data:

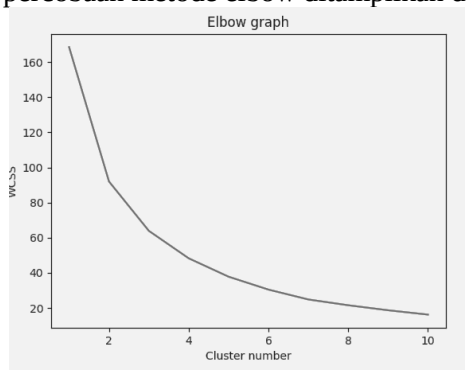
TABEL 4. Hasil *preprocess*

ID Produk	R	F	M
1	0.47712	1.04139	1.518
2	0.60205	1.44715	2.702
3	0.60205	2.12057	2.965
4	0.30102	0.47712	1.799
6	0	0	1

ID Produk	R	F	M
8	0	0	1.5051
9	0.47712	0.47712	2.176
10	0	0.30102	1.447
11	0.30102	0.47712	2.33
12	0	0.47712	1.982

Penentuan Jumlah *Cluster*

Penetuan nilai k menggunakan Metode Elbow. Pada plot tersebut didapatkan titik siku yang terbentuk diantara titik dua dan empat, setelah titik 4 sudah tidak lagi terjadi penurunan yang signifikan secara *intuitive*, sehingga dapat disimpulkan bahwa jumlah *cluster* menurut metode Elbow yaitu sebanyak 4 *cluster*. Grafik percobaan metode elbow ditampilkan di gambar 3.



Gambar 3. Grafik hasil cluster terbaik

K-Means clustering

Langkah-langkah clustering dilakukan dengan contoh perhitungan sebagai berikut :

1. Inisialisasi centroid Terdapat 3 variabel input clustering yaitu R,F dan M dari hasil preprocess maka clustering dilakukan dengan 3 dimensi pengukuran. Jumlah cluster terbaik dari nilai elbow adalah 4 maka sebagai permisalan inisialisasi cluster 0 (c0) [1,1,1], cluster 1 (c1) [2,2,2], cluster 2 (c2) [3,3,3] dan cluster 3 (c3) [4,4,4].
2. Hitung jarak objek dari centroid dengan menggunakan rumus Euclidean distance dilakukan perhitungan jarak nilai tiap produk dengan centroid yang telah diinisialisasi. Sebagai sampel perhitungan diberikan 2 macam produk berikut.

Produk 1

$$c_1 = \sqrt{(0.477-1)^2+(1.041-1)^2+(1.518-1)^2}$$
$$= \sqrt{(-0.522)^2+(0.041)^2+(0.518)^2}$$
$$= \sqrt{0.273+0.001+0.268}$$
$$= 0.543$$

$$c_2 = \sqrt{(0.4771-2)^2+(1.041-2)^2+(1.518-2)^2}$$
$$= \sqrt{(-1.522)^2+(-0.958)^2+(-0.481)^2}$$
$$= \sqrt{2.319+0.918+0.231}$$
$$= 3.469$$

$c_3 = \sqrt{(0.477-3)^2+(1.041-3)^2+(1.518-3)^2}$
$= \sqrt{(-2.522)^2+(-1.958)^2+(-1.481)^2}$
$= \sqrt{6.364+3.836+2.194}$
$=12.3958$

$c_4 = \sqrt{(0.477-4)^2+(1.041-4)^2+(1.518-4)^2}$
$= \sqrt{(-3.522)^2+(-2.958)^2+(-2.4814)^2}$
$= \sqrt{12.410+8.753+6.157}$
$=27.321$

Produk 2

$c_1 = \sqrt{(0.477-1)^2+(1.041-1)^2+(1.518-1)^2}$
$= \sqrt{(-0.397)^2+(0.447)^2+(1.702)^2}$
$= \sqrt{0.273+0.001+0.2688}$
$=3.2565$

$c_2 = \sqrt{(0.602-2)^2+(1.447-2)^2+(2.702-2)^2}$
$= \sqrt{(-1.397)^2+(-0.552)^2+(0.7024)^2}$
$= \sqrt{1.954+0.305+0.493}$
$=2.7532$

$c_3 = \sqrt{(0.602-3)^2+(1.447-3)^2+(2.702-3)^2}$
$= \sqrt{(-0.397)^2+(0.4477)^2+(1.702)^2}$
$= \sqrt{5.7501+2.411+0.088}$
$=8.24998$

$c_4 = \sqrt{(0.602-4)^2+(1.447-4)^2+(2.702-4)^2}$
$= \sqrt{(-3.397)^2+(-2.552)^2+(-1.297)^2}$
$= \sqrt{11.5459+6.5170+1.683}$
$=19.7466$

Himpunan nilai distance sebagai berikut:

Produk1 produk2

$$D^0 = \begin{pmatrix} 0.54 & 3.25 \\ 3.46 & 2.75 \\ 12.39 & 8.24 \\ 27.32 & 19.74 \end{pmatrix} \begin{matrix} \rightarrow c_0 \\ \rightarrow c_1 \\ \rightarrow c_2 \\ \rightarrow c_3 \end{matrix}$$

3. Pengelompokan berdasarkan jarak minimum

Jika hasil distance (jarak) dari nilai produk dan centroid adalah yang paling kecil maka produk tersebut termasuk pada cluster tersebut. Pada perhitungan ini nilai terkecil pada Produk1 adalah 0.54 yaitu pada cluster 0 dan Produk2 adalah 2.75 yaitu pada cluster 1.

$$D^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{matrix} \rightarrow c_0 \\ \rightarrow c_1 \\ \rightarrow c_2 \\ \rightarrow c_3 \end{matrix}$$

Tidak terdapat objek yang berpindah

Objek pada cluster kemudian dibandingkan dengan iterasi sebelumnya, jika masih terdapat objek / produk yang berpindah cluster maka kembali melakukan inisialisasi centroid berdasarkan rata-rata

nilai keseluruhan produk pada cluster yang baru. Jika tidak ada objek yang berubah clusternya maka iterasi berhenti dengan hasil cluster yang terakhir.
Dari hasil akhir clustering dianalisis dan diketahui karakteristik nilai RFM setiap data di tiap cluster, maka hasil clustering dan jumlah produk pada tiap cluster dapat dilabelkan seperti pada tabel 5.

TABEL 5. Hasil Clustering produk dalam Game HayDay

Cluster	Label	Jumlah produk
c0	Superstar	4
c1	Bronze	18
c2	Gold	23
c3	Silver	12

Hasil clustering indeks produk menunjukkan bahwa terdapat 4 produk dengan peringkat tertinggi yaitu pada cluster 0, Superstar. 4 produk tersebut adalah *Egg, Corn, Wol* dan *Gold Bar*. Produk Bronze teratas yaitu Platinum bar, Bacon pie dan Butter. Produk Gold teratas yaitu Cheese, Red berry cake, Silver bar. Produk Silver teratas yaitu Wheat, Cream, Chili pepper.

3. Kesimpulan

Kesimpulan yang diperoleh dari analisis dan pembahasan yaitu terdapat 4 cluster dimana cluster 0,1,2,3 termasuk pada kategori Superstar, Bronze, Gold, Silver secara berurutan. Hasil clustering indeks produk menunjukkan bahwa terdapat 4 produk dengan peringkat tertinggi yaitu pada cluster 0, Superstar. 4 produk tersebut adalah *Egg, Corn, Wol* dan *Gold Bar*. Produk Bronze teratas yaitu Platinum bar, Bacon pie dan Butter. Produk Gold teratas yaitu Cheese, Red berry cake, Silver bar. Produk Silver teratas yaitu Wheat, Cream, Chili pepper.

Ucapan Terima Kasih
Puji syukur kehadirat Allah Subhanallah Wata'alla, karena dengan izin-Nya kami dapat merampungkan penelitian ini dengan lancar. Terima kasih kami ucapkan kepada dosen matakuliah Manajemen Proyek. Dan tidak lupa terima kasih kepada rekan-rekan mahasiswa UIN Maulana Malik Ibrahim khususnya pada jurusan Teknik Informatika yang telah bersama berbagi ilmu dan pengalaman selama penelitian. Semoga hasil penelitian ini dapat bermanfaat bagi penelitian selanjutnya dan bagi kemajuan teknologi.

Daftar Pustaka

[1]. Y. Luo, Q. R. Cai, H. X. Xi, Y. J. Liu, and Z. M. Yu. 2012, *Telecom customer segmentation with K-means clustering,* ICCSE 2012 - Proc.
[2]. Kodinariya, Trupti M. & Makwana, Prashant R.,(2013). *Review on determining number of cluster in K - Means Clustering. International Journal of Advance Research in Computer Science andManagement Studies*, I(6), pp. 90-95
[3]. D. Zheng. 2013. *Application of silence customer segmentation in securities industry based on fuzzy cluster algorithm,* J. Inf. Comput. Sci., vol. 10, no. 13, pp. 4337–4347.
[4]. R. J. Kuo, S. H. Lin, and Z.-Y. Chen 2015. *Integration of Particle Swarm Optimization and Immune Genetic Algorithm-Based Dynamic Clustering for Customer Clustering*, Int. J. Artif. Intell. Tools, vol. 24, no. 5, p. 1550019.
[5]. W. Li. 2008. *Modified K-Means Clustering Algorithm,* 2008 Congr. Image Signal Process., pp. 618–621.

[6]. J. Zhao, W. Zhang, and Y. Liu 2010. *Improved K-Means cluster algorithm in telecommunications enterprises customer segmentation,* Proc. 2010 IEEE Int. Conf. Inf. Theory Inf. Secur. ICITIS 2010, pp. 167–169,
[7]. D. Birant, *Data Mining Using RFM Analysis, in Knowledge-Oriented Applications in Data Mining*, Turkey: KImito Funatso, 2011, pp. 91–108.
[8]. Bholowalia, Purnima & Kumar, Arvind, 2014. *EBK-Means: A Clustering Techiniques based on Elbow Method and K-Means in WSN*. International Journal of Computer
[9]. Muhammad, A.F. 2015."Klasterisasi Proses Seleksi Pemain Menggunakan Algoritma k-Means". Jurusan Teknik Informatika FIK UDINUS. Semarang.
[10]. Adiana, B.E. Soesanti, I. Permanasari, A. E. 2018. "Analisis Segmentasi Pelanggan Menggunakan Kombinasi RFM Model dan Teknik Clustering". Jurusan Teknik Elektro dan Teknologi Informasi, Universitas Gadjah Mada.
[11]. H. P. Daniel, R. I. Retno, P. R. Andi. 2018. "Analisis Pemetaan Pelanggan Potensial Menggunakan Algoritma K-Means dan LRFM Model Untuk Mendukung Strategi Pengelolaan Pelanggan (Studi Pada Maninjau Center Kota Malang)." Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Brawijaya