

Penerapan Algoritma K-Means Dalam Penentuan Peminatan Kain

Munawarah¹⁾ Nur Khafidhoh²⁾

¹⁾Fakultas Teknologi Informasi Universitas KH.A. Wahab Hasbullah

¹⁾e-mail : munawarahunwaha@gmail.com

²⁾Fakultas Teknologi Informasi Universitas KH.A. Wahab Hasbullah

²⁾e-mail : nur.khafidhohunwaha@gmail.com



©2019 –EPiC Universitas KH. A. Wahab Hasbullah Jombang ini adalah artikel dengan akses terbuka dibawah lisensi CC BY-NC-4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>).

ABSTRAK

Banyak sektor industri yang semakin berkembang di kondisi pandemik Covid-19 saat ini, seperti sektor industri kecil menengah yang paling merasakan efeknya. Salah satu sentra industri yang memiliki potensi untuk mengembangkan inovasi dalam Perusahaan Tekstil. Dalam ekonomi industri kecil menengah memiliki peranan yang cukup penting dalam membangun perodaan ekonomi negara, yang menjadikan industri-industri tersebut penting dikembangkan. Penelitian ini bertujuan untuk mengetahui bagaimana analisis pemetaan kluster inovasi yang terbentuk pada IKM (Industri Kecil Menengah) di daerah masing-masing berdasarkan inovasi produk, inovasi proses, dan inovasi pasar. Populasi pada penelitian ini adalah pelaku usaha sentra industri produk tekstil. Penelitian ini bertujuan untuk Penyediaan kebutuhan kain yang dibutuhkan dan untuk memudahkan pengelompokan dalam peminatan rendah, sedang, tinggi kain yang dibutuhkan oleh masyarakat. Teknik analisis data dalam penelitian ini adalah deskriptif dan analisis kluster dengan melakukan pengolahan data menggunakan Algoritma K-Means. Hasil dari penelitian ini adalah bahwa sebanyak 244 peminatan rendah telah melakukan inovasi produk, sebanyak 232 peminatan sedang telah melakukan inovasi proses, sebanyak 56 peminatan tinggi telah melakukan inovasi pasar.

Kata kunci : Tekstil kain, Clustering, K-means

PENDAHULUAN

Banyak sektor industri yang semakin berkembang di kondisi pandemik Covid-19 saat ini, seperti sektor industri kecil menengah yang paling merasakan efeknya. Salah satu sentra industri yang memiliki potensi untuk mengembangkan inovasi dalam Perusahaan Tekstil. Dalam ekonomi industri kecil menengah memiliki peranan yang cukup penting dalam membangun perodaan ekonomi negara, yang menjadikan industri-industri tersebut penting dikembangkan. Penelitian ini bertujuan untuk mengetahui bagaimana analisis pemetaan kluster inovasi yang terbentuk pada IKM (Industri Kecil Menengah) di daerah masing-masing berdasarkan inovasi produk, inovasi proses, dan inovasi pasar. Populasi

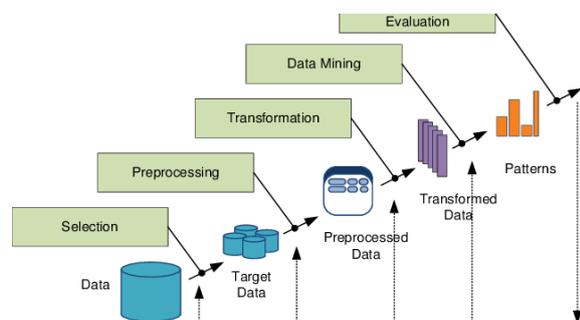
pada penelitian ini adalah pelaku usaha sentra industri produk tekstil. Penelitian ini bertujuan untuk Penyediaan kebutuhan kain yang dibutuhkan dan untuk memudahkan pengelompokan dalam peminatan rendah, sedang, tinggi kain yang dibutuhkan oleh masyarakat. Teknik analisis data dalam penelitian ini adalah deskriptif dan analisis kluster dengan melakukan pengolahan data menggunakan Algoritma K-Means. Pada inovasi proses, dimana untuk melakukan efisiensi dalam pembuatan produk diperlukan mesin canggih yang akan menekan biaya operasional. Namun memerlukan biaya yang besar karena harga mesin yang mahal, ditambah biaya pengiriman dan bea cukai jika impor. Pada inovasi pasar, beberapa IKM hanya

berorientasi pada penjualan atau sales saja, tidak melakukan upaya branding. Sehingga peluang terambil oleh kompetitor, baik yang baru masuk kepasar atau kompetitor lama. (Sumber: Hasil wawancara dengan Bapak Nio, selaku pengelola KTC 2018) Inovasi menjadi perhatian penting karena sangat dibutuhkan, terlebih dalam perkembangan dunia yang terus berubah. Sebagai pengusaha, harus bisa membuat ide atau inovasi yang dapat membedakan produk yang dimiliki dengan competitor. Manfaat dari ketiga inovasi tersebut adalah pada inovasi produk untuk meningkatkan kualitas produk dan pengembangan produknya. Inovasi proses dengan tujuan melakukan penghematan biaya produksi juga efisiensi pembuatan produk. Dan Inovasi pasar untuk menciptakan pasar baru atau memperluas daerah pasar, dalam mengembangkan pasar lebih luas dalam segmen yang lebih tinggi. Pentingnya dilakukan pemetaan klaster inovasi salah satunya adalah untuk menstimulasi pengembangan IKM dalam bersaing dengan produk-produk lain.

METODE

2.1 Knowledge Discovery in Database

KDD adalah metode yang digunakan untuk dapat memperoleh pengetahuan yang berasal dari database yang ada. Hasil pengetahuan yang diperoleh dapat dimanfaatkan untuk basis pengetahuan (*knowledge base*) yang digunakan dalam keperluan mengambil keputusan. Secara lebih detail, proses KDD seperti pada gambar berikut ini yang diadopsi dari :



Gambar 2. Proses KDD

1. *Selection* : Selection digunakan untuk menentukan variabel yang akan diambil agar tidak ada kesamaan dan terjadi perulangan

yang tidak diperlukan dalam pengolahan data mining.

2. *Preprocessing* : Pada preprocessing terdapat dua tahap, yaitu sebagai berikut :
 - a. *Data Cleaning* : Menghilangkan data yang tidak diperlukan seperti menangani missing value, noise data serta menangani data – data yang tidak konsisten dan relevan.
 - b. *Data Integration* : Dilakukan terhadap atribut yang mengidentifikasi entitas yang unik.
3. *Transformation* : Merubah data sesuai format ekstention yang sesuai dalam pengolahan data mining karena beberapa metode pada data mining memerlukan format khusus sebelum dapat diproses pada data mining.
4. *Data mining* : Proses utama pada metode yang diterapkan untuk mendapatkan pengetahuan baru dari data yang diproses. Pada penelitian ini diterapkan teknik clustering yaitu metode K-Means Clustering.
5. *Evaluation/ Interpretation* : Mengidentifikasi pola – polayang menarik kedalam knowdge base yang diidentifikasi. Pada tahap ini, menghasilkan pola – polakhas maupun model prediksi yang dievaluasi untuk menilai kajian yang ada sudah memenuhi target yang diinginkan.
6. *Knowledge Pola* : Pola yang dihasilkan akan dipresentasikan kepada pengguna. Pada tahapan ini pengetahuan baru yang dihasilkan bisa dipahami semua orang yang akan dijadikan acuan pengambilan keputusan.

Clustering atau klasterisasi merupakan metode pengelompokan data. Menurut Tan, 2006 clustering adalah sebuah proses untuk mengelompokan data ke dalam beberapa cluster atau kelompok sehingga data dalam satu cluster memiliki tingkat kemiripan yang maksimum dan data antar cluster memiliki kemiripan yang minimum. Clustering merupakan proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan cluster. Objek yang di dalam cluster memiliki kemiripan karakteristik antar satu sama lainnya dan berbeda dengan cluster yang lain. Partisi tidak dilakukan secara manual melainkan dengan suatu algoritma clustering. Oleh karena itu, clustering sangat berguna dan bisa menemukan group atau kelompok yang tidak dikenal dalam data. Clustering banyak digunakan dalam berbagai

aplikasi seperti misalnya pada business intelligence, pengenalan pola citra, web search, bidang ilmu biologi, dan untuk keamanan (security). Di dalam business intelligence, clustering bisa mengatur banyak customer ke dalam banyaknya kelompok. Contohnya mengelompokkan customer ke dalam beberapa cluster dengan kesamaan karakteristik yang kuat. Clustering juga dikenal sebagai data segmentasi karena clustering mempartisi banyak data set ke dalam banyak group

ISSN : 2355-9357 e-Proceeding of Management : Vol.6, No.2 Agustus 2019 | Page 2023 berdasarkan kesamaannya. Selain itu clustering juga bisa sebagai outlier detection

2.2 Proses Algoritma K-Means

1. Algoritma k-means adalah metode clustering yang paling sederhana dan umum. Hal ini dikarenakan K-means mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien. Dalam penelitian ini dilakukan pengelompokan berdasarkan peminatan yang ada dengan menggunakan langkah :

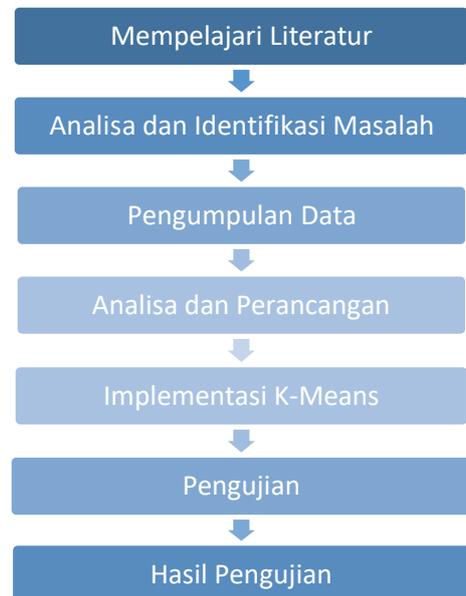
- a. Peminat rendah (C1)
- b. Peminat sedang (C2)
- c. Peminat tinggi (C3)

Langkah selanjutnya adalah menentukan jumlah cluster, adapun cluster yang diinginkan terdiri atas 2 cluster. Kemudian diberikan nilai centroid yang diambil secara acak dari tabel diatas, dengan menggunakan formula : Dimana j adalah nomor index kriteria atau $\{j | 1 \leq j \leq 3, j \in A\}$ dan $j = 1 \dots$

Langkah berikutnya adalah menghitung jarak menggunakan rumus Euclidean index random centroidnya adalah sebagai berikut yang mana untuk C1 diambil dari data yang mempunyai nilai tertinggi dan c2 diambil data yang mempunyai nilai terendah. Menghitung jarak setiap data yang ada terhadap nilai centroid, ditampilkan untuk data iterasi pertama dengan rumus berikut:

$$D_{L_i}(x_2, x_1) = \|x_2 - x_1\| = \sqrt{\sum_{j=1}^p |x_{2j} - x_{1j}|^2}$$

$$D_{L_2}(x_2, x_1) = \|x_2 - x_1\| = \sqrt{\sum_{j=1}^p (x_{2j} - x_{1j})^2}$$



Gambar 2. Kerangka Kerja Penelitian

HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data yang berasal dari Laporan Penjualan Kain. Parameter yang digunakan adalah nama kain, unit, jumlah item kain setiap bulan dari Januari hingga Desember. Proses *cluster* dilakukan untuk mengetahui pola kebutuhan Ukm Sentra Industri Produk Tekstil.

3.1 Data Integration

Pada tahap data integration ini dilakukan penggabungan data-data kain yang terdapat pada laporan setiap bulannya kedalam satu tahun menjadi total peminatan Untuk peminatan kain yang tidak ada jumlah peminatan akan diberi nilai 0, seperti yang terlihat di tabel 2.

Tabel 1. Data Kain

No	Nama Kain	Total Peminatan
1	aurora crepe 58 hijau semen	2505
2	aurora crepe 58 khaki	2338
3	aurora crepe 58 navy	2895
4	baby doll 8537 bunga abu	2059
5	baby doll 8537 bunga mint	2088

6	baby doll 8537 bunga navy	2042
....
20	ct str b1080 onde hitam	4827.5
21	ct str b1080 onde navy	4823
22	ct str b1080 onde pink	4860.5
....
288	KS SCUBA PLS MARON 72	3275.67
289	KS SCUBA PLS MERAH CABE 72	2363.87
290	KS SCUBA PLS NAVY 72	2780.72
291	KS SCUBA POLOS BW 72	3184.29
....
521	yd flanel ktk 301 misty	5874.5
522	yd flanel ktk 302 abu	5803
523	yd flanel ktk 302 biru	5790
524	yd flanel ktk 302 cream	5794.5
521	yd flanel ktk 301 misty	5874.5
522	yd flanel ktk 302 abu	5803
523	yd flanel ktk 302 biru	5790

3.2 Data Cleaning

Pada tahap data cleaning ini yang dilakukan adalah membuang atribut yang tidak relevan atau tidak konsisten. Pada data ini yang digunakan yaitu nama kain, total peminatan, dataset yang telah di cleaning dapat dilihat pada tabel 1.

Tabel 2. Data Integration

Nama Kain	Jan	Feb	Mar	...	Total
aurora crepe 58 hijau semen	0	818	1585	2505
aurora crepe 58 khaki	0	302	2036	2338
aurora crepe 58 navy	0	308	2587	2895
yd flanel ktk 302 merah	782	1376	704	782
yd flanel ktk 303 biru	600	1094	299	600
yd flanel ktk 303 merah	589	1100	300	589
yd flanel ktk 303 putih	600	1090	327	600
yd flanel ktk 303 tosca	610	1100	292	610
yd flanel ktk 923 br	600	706	0	600

yd flanel ktk 923 mr	600	700	0	600
yd flanel ktk 923 ts	600	710	0	600
yd flanel ktk 302 merah	782	1376	704	782
yd flanel ktk 303 biru	600	1094	299	600
....

3.3 Proses Training

1. Iterasi 1 Tentukan K jumlah pusat cluster secara acak (random). Pada percobaan pertama ini ditentukan 3 data secara acak sebagai titik pusat awal untuk perhitungan jarak dari seluruh kelompok cluster yang akan dibentuk. Jumlah cluster = 3 (sedikit, sedang, tinggi) Jumlah data = 532 dengan atribut = 1
2. Hitung jarak tiap data dengan masing-masing cluster pusat dengan menggunakan persamaan (3) yaitu persamaan Euclidean Distance.

Tabel 3. Perhitungan Euclidean Distance Iterasi

nama kain	total pemakaian	cluster
yd flanel ktk 923 mr	2000	c1
rayon calista semangka 83887 navy	3624	c2
ctn oxford pls cream	7978	c3

Anggota dipilih dari terkecil diantara 3 cluster jika terkecil pada bagian C1 maka termasuk sebagai anggota C1 yaitu sebanyak 244 data, jika terkecil pada bagian C2 maka termasuk sebagai anggota C2 yaitu sebanyak 232 data, dan jika terkecil pada bagian C3 maka termasuk sebagai anggota C3 yaitu sebanyak 56 data.

3. Lakukan iterasi ke 2 Tentukan posisi centroid baru dengan cara menghitung rata-rata dari data-data yang ada pada centroid yang sama atau anggota yang sama
4. Jika hasil iterasi ke-2 tidak sama dengan iterasi ke-1 sehingga perlu dilakukan kembali perhitungan ke iterasi ke-3 dan

seterusnya sampai mendapatkan hasil yang sama.

5. Dan jika hasil iterasi ke-3 masih tidak sama dengan iterasi ke-2 sehingga perlu dilakukan kembali perhitungan ke iterasi ke-4 dan seterusnya sampai mendapatkan hasil yang sama.
6. Iterasi ke 4 dengan iterasi ke-3 hasilnya sudah sama maka tidak perlu lagi melanjutkan ke iterasi ke-5 atau cukup berhenti di iterasi ke-4.

Hasil Clusterisasi data kain yang dilakukan dengan algoritma k-means didapatkan setelah melakukan iterasi ke-4 yaitu terdapat kelompok peminatan kain yang peminatan sedikit terdapat pada cluster 1 yang memiliki 244 anggota, kelompok kain yang peminatan sedang terdapat pada cluster 2 yang memiliki 232 anggota, dan kelompok kain yang peminatan tinggi terdapat pada cluster 3 yang memiliki 56 anggota.

Tabel 4. Anggota Cluster 1 (Sedikit)

nama kain	total peminatan
ctn 4032 40s hjtn 488b	3092
ctn oxford kotak 0.8cm cream	3722
ctn oxford kotak 1cm abu	3695
ctn str b1078 bunga coku	3132
rayon calista semangka 83887 hitam	3545
rayon calista semangka 83887 kubus	3580
rayon calista semangka 83887 navy	3624
rayon calista semangka 83887 pink	3265
rayon calista shibori 84146 hitam	3120
rayon camelia hawai 83372 navy	3165
rayon naila patchwork 34976 navy	3435
rayon twill saffron bw	3148
supernova pls 06/merah	3395
supernova pls 23/biru abu	3384
toyobo fd 86	3454
twill rainbow spdx #13	3139
.....

Tabel 5. Anggota Cluster 2 (Sedang)

nama kain	total peminatan
aurora crepe 58 hijau semen	2505
aurora crepe 58 khaki	2338
aurora crepe 58 navy	2895
baby doll 8537 bunga abu	2059
aurora crepe 58 hijau semen	2505
ctn str a5607 denim hitam	2049.5
ctn str a5607 denim navy	2049
ctn str a6495a onde 0.8 ht	2359
ctn str a5607 denim hitam	2049.5
twill rainbow salur 1cm navy	2655
twill rainbow spdx #04	2029
twill rainbow spdx #20	2511
uniqlo twill spdx 01	2418
uniqlo twill spdx 02	2579
twill rainbow salur 1cm navy	2655
.....

Tabel 3.6 Anggota Cluster 3 (Tinggi)

nama kain	total pemakaian
ct str b1080 onde hitam	4827.5
ct str b1080 onde navy	4823
ct str b1080 onde pink	4860.5
ct str b1080 onde tosca	4822
yd flanel ktk 303 biru	5091
yd flanel ktk 303 merah	5089
yd flanel ktk 303 putih	5123
yd flanel ktk 303 tosca	5091
yd flanel ktk 303 biru	5091
yd flanel ktk 303 merah	5089
.....

SIMPULAN DAN SARAN

Penelitian ini bertujuan untuk memetakan klaster inovasi yang terbentuk pada IKM (Industri Kecil Menengah) berdasarkan inovasi. Terdapat tiga kategori inovasi yang digunakan dalam penelitian ini yaitu inovasi produk, inovasi proses, dan inovasi pasar. Berdasarkan hasil analisis dan pengolahan data yang telah dilakukan dalam penelitian, didapatkan kesimpulan bahwa terdapat 244 anggota usaha tekstil yang masuk

kedalam klaster 1, 232 anggota pengusaha tekstil yang masuk kedalam klaster 2 dan 56 anggota klaster 3 adalah usaha-usaha yang telah melakukan inovasi pada produknya. Usaha-usaha yang telah melakukan inovasi pada produk, proses, pasar dan klaster yang belum melakukan inovasi pada produk, proses, pasar.

DAFTAR PUSTAKA

Aji Supriyanto, B. J. (2017). Klasterisasi UMKM dan Potensi Wilayah Berbasis Peta Sebagai Strategi Pengembangan Ekonomi Daerah. *Jurnal Pekommas*, Vol. 2, No. 2, 143-150.

Alfina, T., Santosa, B., & Barakbah, A. R. (2012). Analisa Perbandingan Metode Hierarchical Clustering, Kmeans dan Gabungan Keduanya dalam Cluster Data (Studi kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS). *Jurnal Teknik ITS* Vol. 1, 521-525.

Anggriani, A. D. (2018). Pengaruh Kreativitas Dan Inovasi Terhadap Keberhasilan Usaha Kecil Menengah (UKM) (Studi Pada Cawang ART, Kriya Nusantara). Bandung: Universitas Telkom.

Arikunto, S. (2000). *Manajemen Penelitian*. Jakarta: Rineka Cipta. [5]Arthurs, D., Cassidy, E., H.Davis, C., & Wolfe, D. (2009). Indicators to Support Innovation Cluster Policy. *International Journal of Technology Management* Vol. 46, Nos. 3/4, 263-277.

Baso, T. (2015). Strategi Pengembangan Daya Saing Pada Sentra Industri Tekstil Cigondewah Bandung. Bandung: Universitas Telkom.

Bungin, M. B. (2005). *Metodologi Penelitian Kuantitatif: Komunikasi, ekonomi, dan kebijakan publik serta ilmu-ilmu sosial lainnya*. Jakarta: Prenada Media.

Damanik, E. D. (2013). Analisis Klaster Mengenai Konsumsi Media Televisi Di Kalangan Remaja Kota Bandung (Studi Remaja SMP Negeri dan SMP Swasta). Bandung: TEBS Library & Knowledge Center.

Deepali Virmani ,Shweta Taneja ,Geetika Malhotra, "Normalization based K means Clustering Algorithm" , ISSN: 2349-6495, *International Journal of Advanced Engineering Research and Science (IJAERS)*, [Vol-2, Issue-2, Feb.- 2015].

Dhewanto, W., Lantu, D., & Anggadwita, G. (2015). The innovation cluster of ICT start-up companies in developing countries: Case of Bandung, Indonesia. *International Journal of Learning and Intellectual Capital* Vol. 12, No. 1, 32-44.

Fauziah, D. A. (2016). Pola Pengembangan Sentra Industri Kain Cigondewah Kidul (Studi Kasus Pada Pedagang Sentra Kain di Cigondewah Kidul Kecamatan Bandung Kulon Kota Bandung). Bandung: Universitas Pasundan. [11]H, R. M., Maghfiroh, A., R, A. C., & P, M. R. (2013). *Manajemen Inovasi*. Malang: Universitas Brawijaya.

Hadjar, I. (1996). *Dasar-dasar Metodologi Penelitian Kuantitatif dalam Pendidikan*. Jakarta: Raja Grafindo Persada.

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques 3rd Edition*. Waltham: Elsevier.

Hasri, M. N. (2018). Pengaruh Market Orientation Terhadap Kinerja Perusahaan Dengan Inovasi Sebagai Variabel Intervening (Studi Kasus Industri Ekonomi Kreatif Sub Sektor Fashion Di Social Commerce Kota Bandung) Universitas Telkom. Bandung: Universitas Telkom

Marvin Herry, Anggadwita Grisna (2019). Analisis Pemetaan Klaster Inovasi Pada Sentra Industri Produk Tekstil Cigondewah. Vol.6, No.2, ISSN : 2355-9357

Sirega, M.Hasyim (2018). Klasterisasi Penjualan Alat-Alat Bangunan Menggunakan Metode K-Means (Studi Kasus Di Toko Adi Bangunan). Vol. 1 No. 2, Hal 83-91, ISSN : 2622-1659