



PENGELOMPOKAN INDUSTRI MIKRO DI INDONESIA DENGAN METODE *K-MEANS* CLUSTERING

Aswan S. Sunge¹, Heri Hermawanto²

Program Studi Teknik Informatika Fakultas Teknik Universitas Pelita Bangsa

¹aswan.sunge@pelitabangsa.ac.id

Abstraksi

Industri merupakan bagian dari proses produksi yang tidak mengambil bahan-bahan tersebut langsung dari alam untuk di konsumsi tetapi bahan-bahan di proses dan akhirnya menjadi komoditas yang berharga kepada masyarakat. Industri mikro adalah perusahaan industri yang tenaga kerjanya terdiri dari 5 – 19 orang. Tujuan dari penelitian ini adalah untuk mengelompokan industri mikro yang ada di Indonesia berdasarkan 3 klaster yaitu klaster rendah, klaster sedang dan klaster tinggi dengan menggunakan metode *K-Means Clustering*. Data yang digunakan dalam penelitian ini adalah data jumlah perusahaan industri mikro di indonesia berdasarkan provinsi pada tahun 2013 – 2015. Berdasarkan data tersebut diperoleh daerah dengan jumlah perusahaan industri mikro dengan klaster rendah (C0) sebanyak 31 provinsi, klaster sedang (C1) 2 provinsi dan klaster tinggi (C2) 1 provinsi.

Kata kunci : *Industri, Klastering, Algoritma K-means*

Abstract

Industry is part of the production sector that does not take these materials directly from nature for consumption but the materials are processed and finally become valuable and valuable to society. Micro industry is an industrial company whose workforce consists of 5-19 people. The purpose of this research is to classify the micro industry in Indonesia based on 3 clusters, namely the low cluster, medium cluster and high cluster using the K-Means Clustering method. The data used in this study is data on the number of micro industrial companies in Indonesia by province in 2013 - 2015. Based on this data, it was obtained that the regions with the number of micro industrial companies with low clusters (C0) were 31 provinces, medium cluster (C1) 2 provinces and clusters. height (C2) 1 province.

Keywords : *Industry, Clastering, Algoritma K-means*

1. Pendahuluan

Usaha Mikro, Kecil dan Menengah merupakan salah satu unit usaha yang memiliki

peran penting dalam perkembangan dan pertumbuhan perekonomian Indonesia. Dengan adanya sektor UMKM, pengangguran akibat angkatan kerja yang tidak terserap dalam dunia kerja menjadi berkurang. Tumbuhnya usaha mikro menjadikannya sebagai sumber pertumbuhan kesempatan kerja dan pendapatan. Dengan banyak menyerap tenaga kerja berarti UMKM mempunyai peran strategis dalam upaya pemerintah dalam memerangi kemiskinan dan pengangguran. Usaha Mikro, Kecil dan Menengah (UMKM) mempunyai peran yang strategis dalam pembangunan ekonomi nasional, oleh karena selain berperan dalam pertumbuhan ekonomi dan penyerapan tenaga kerja juga berperan dalam pendistribusian hasil-hasil pembangunan.

Dalam Krisis ekonomi yang terjadi di negara Indonesia pada tahun 1997 hingga 1998 dimana banyak usaha berskala besar yang mengalami stagnasi bahkan berhenti aktifitasnya, sektor UMKM terbukti lebih tangguh dalam menghadapi krisis tersebut. Pasca krisis ekonomi tahun 1997-1998 jumlah UMKM tidak berkurang, justru meningkat terus, bahkan mampu menyerap 85 juta hingga 107 juta tenaga kerja sampai tahun 2012. Pada tahun 2012,

jumlah pengusaha di Indonesia sebanyak 56.539.560 unit. Dari jumlah tersebut, Usaha Mikro Kecil dan Menengah (UMKM) sebanyak 56.534.591 unit atau 99,99% [1]. Namun disisi lain ditengah pentingnya sektor UMKM bagi Indonesia, sampai saat ini masih belum adanya sistem informasi yang mampu mengelompokkan jumlah industri mikro yang ada di Indonesia, mengingat begitu pentingnya sektor UMKM ini bagi negara Indonesia.

2. Landasan Pemikiran

2.1 Industri

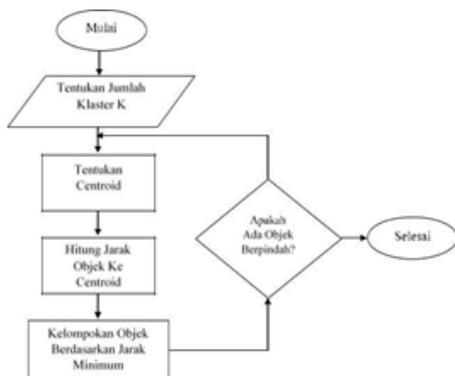
Pengertian industri menurut Undang – Undang No. 3 Tahun 2014 adalah seluruh bentuk kegiatan ekonomi yang mengolah bahan baku dan memanfaatkan sumber daya industri sehingga menghasilkan barang yang mempunyai nilai tambah atau manfaat lebih tinggi, termasuk jasa industri [2]. Industri kecil merupakan industri yang jumlah karyawan/tenaga kerja berjumlah antara 5–19 orang. Modal relatif kecil karena modal disediakan oleh seorang pemilik atau sekelompok kecil pemilik modal, tenaga kerjanya berasal dari lingkungan sekitar atau masih ada hubungan saudara [2].

2.2 Clustering

Clustering adalah proses mengelompokkan objek berdasarkan informasi yang diperoleh dari data yang menjelaskan hubungan antar objek dengan prinsip untuk memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas-nya. Penerapan metode clustering menghasilkan pengetahuan berupa penentuan beberapa cluster catatan data yang memiliki kemiripan atribut [3].

2.3 Algoritma K-means

Algoritma K-means merupakan salah satu algoritma dengan partitional, karena K-means didasarkan pada penentuan jumlah awal kelompok dengan mendefinisikan nilai centroid awalnya. K-means clustering merupakan salah satu metode data clustering non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih cluster/kelompok.



Gambar 1. K-means Method

Proses pengelompokkan data ke dalam suatu cluster dapat dilakukan dengan cara menghitung jarak terdekat dari suatu data ke sebuah titik centroid. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai persamaan berikut :

$$D(i,j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{ik} - x_{jk})^2}$$

Dimana :

(i,j) = Jarak data ke-i ke pusat cluster j

x_{ki} = Data ke-i pada atribut data ke-k

x_{kj} = Titik pusat ke-j pada atribut ke-k

jarak pusat cluster dihitung kembali dengan keanggotaan cluster yang sekarang. Pusat cluster adalah rata-rata dari semua data/objek dalam cluster tertentu. Jika dikehendaki bisa juga menggunakan median dari cluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai. Setiap objek kemudian ditugaskan kembali memakai pusat cluster yang baru [4].

3. Metodologi Penelitian

Dalam penelitian ini, lingkup objek penelitian yang diterapkan peneliti sesuai dengan permasalahan yang akan diteliti yaitu mengelompokkan industri mikro di Indonesia dengan metode k-means clustering, adapun objek pada penelitian ini yaitu data jumlah industri mikro yang ada di Indonesia yang didapatkan dari situs Badan Pusat Statistik Indonesia.



Gambar 2. Kerangka Berpikir

3.1 Dataset

Pada penelitian ini dataset yang digunakan sebanyak 34 data yang merupakan jumlah provinsi yang ada di Indonesia. Pengambilan data dilakukan pada tanggal 23 Juni 2020 berdasarkan jumlah industri mikro yang ada di Indonesia yang didapatkan dari situs

Badan Pusat Statistik Indonesia sebagai data riset untuk bahan penelitian yang kemudian akan diubah kedalam sebuah file excel. Data yang diambil berdasarkan data jumlah industri mikro dari tahun 2013 hingga 2015. Kriteria data yang digunakan terdiri dari nama provinsi dan tahun. Adapun dataset yang dimaksud adalah sebagai berikut :

Tabel 1. Data Industri Mikro Tahun 2013 -2015

Kep. Riau	13706	14638	7231
DKI Jakarta	20738	15110	28378
Jawa Barat	382899	437985	421881
Jawa Tengah	650115	766782	934814
Yogyakarta	67454	73266	52907
Jawa Timur	539320	608774	771185
Banten	71736	75760	108235
Bali	84149	107434	95282
NTB	93694	93645	79764
NTT	100761	109266	71768
Kalimantan Barat	35892	36311	53867
Kalimantan Tengah	17456	18936	11884
Kalimantan Selatan	64235	67674	55564
Kalimantan Timur	20689	15866	11084
Kalimantan Utara	0	0	1180
Sulawesi Utara	37091	35527	39431
Sulawesi Tengah	30247	38511	20745
Sulawesi Selatan	94537	100526	112896
Sulawesi Tenggara	57180	68711	46084
Gorontalo	20934	22610	12458
Sulawesi Barat	26028	27888	11123
Maluku	35208	36422	19312
Maluku Utara	8328	7851	6939
Papua Barat	2730	2353	1442
Papua	9292	9413	6973

Tabel 2. Centroid data awal

Provinsi	2013	2014	2015
Aceh	74880	69316	64009
Sumatera Utara	64034	76227	94979
Sumatera Barat	57987	71413	63409
Riau	15241	14355	16791
Jambi	22590	25441	22415
Sumatera Selatan	57495	58751	47516
Bengkulu	10595	11310	11663
Lampung	90051	94739	76728
Kep. Bangka Belitung	9723	7752	5914

4. Pembahasan

Dalam penerapan algoritma K-means dihasilkan nilai centroid dari data yang didapat dengan ketentuan klasterisasi yang diinginkan adalah 3. Penentuan cluster dibagi atas 3 klaster

yaitu cluster tingkat rendah (C0), cluster tingkat sedang (C1) dan cluster tingkat tinggi (C2). Pada penelitian ini penentuan titik cluster dilakukan dengan mengambil nilai secara acak pada tabel diatas.

C2	10	13706	14638	7231
C3	20	35892	36311	53867
27	94537	100526	112896	2
28	57180	68711	46084	2
29	20934	22610	12458	0
30	26028	27888	11123	0
31	35208	36422	19312	0
32	8328	7851	6939	1
33	2730	2353	1442	1
34	9292	9413	6973	1

4.1 Klastering Data

Dengan menggunakan centroid yang telah diperoleh, data di klaster menjadi 3 klaster.

Porses klaster data dilakukan dengan mengambil jarak terdekat dari setiap data yang di oleh. Hasil pengelompokan data dapat dilihat pada table 3.

Tabel 3. Penentuan jarak pusat cluster pada iterasi 1

Prov	2013	2014	2015	Cluster
1	74880	69316	64009	2
2	64034	76227	94979	2
3	57987	71413	63409	2
4	15241	14355	16791	1
5	22590	25441	22415	0
6	57495	58751	47516	2
7	10595	11310	11663	1
8	90051	94739	76728	2
9	9723	7752	5914	1
10	13706	14638	7231	1
11	20738	15110	28378	0
12	382899	437985	421881	2
13	650115	766782	934814	2
14	67454	73266	52907	2
15	539320	608774	771185	2
16	71736	75760	108235	2
17	84149	107434	95282	2
18	93694	93645	79764	2
19	100761	109266	71768	2
20	35892	36311	53867	2
21	17456	18936	11884	1
22	64235	67674	55564	2
23	20689	15866	11084	1
24	0	0	1180	1
25	37091	35527	39431	2
26	30247	38511	20745	0

Pada iterasi 1 diperoleh klaster data jumlah perusahaan mikro berdasarkan provinsi yang dapat dilihat pada tabel 3. Proses klastering tersebut

berhenti pada iterasi ke 7. pada iterasi 7 menentukan nilai titik tengah atau centroid dapat diketahui pada Tabel 4.

Tabel 4. Centroid iterasi 7

c1	20.77777778	18693.78	18905.22	16046.11111
c2	13.92307692	75245.62	82056	74549.30769

c3	13.33333333	524111.3	604513.7	709293.3333
----	-------------	----------	----------	-------------

Setelah mendapatkan centroid data, dilakukan perhitungan dengan cara yang sama seperti pada iterasi pertama. Dengan mencari jarak terdekat

Pada penelitian ini, peneliti menggunakan bantuan software rapidminer untuk mendukung ke akuratan hasil perhitungan. Berikut adalah hasil penggunaan software rapidminer.



Gambar 3. Cluster Model



Gambar 4. Cluster Model Visualization

Pada iterasi 7, pengelompokan data yang dilakukan terhadap 3 klaster memiliki hasil yang sama dengan iterasi 6. Sehingga proses perhitungan berhenti pada iterasi ke-7.

5. Penutup

Berdasarkan penelitian yang telah

dari setiap data.

dilakukan menggunakan *Rapidminer Studio* maka dapat disimpulkan metode *K-means Clustering* pada pengelompokan industri mikro berdasarkan provinsi dapat diterapkan. Data diolah menjadi 3 klaster yaitu klaster rendah, sedang dan tinggi. Dimana hasil penelitian menyimpulkan dari 34 provinsi yang ada di Indonesia dan 3 klaster yang ada diperoleh hanya 1 provinsi di Indonesia dengan tingkat klaster tinggi (C2) yaitu Jawa Barat. 2 provinsi di Indonesia dengan tingkat klaster sedang (C1) yaitu Jawa Tengah dan Jawa Timur. Dan 30 provinsi lainnya dengan cluster rendah (C0).

Daftar Pustaka

- [1] D. N. Batubara, A. P. Windarto, A. Wanto, D. Hartama, and E. Irawan, "Penerapan Datamining Klastering Pada Perusahaan Industri Mikro di Indonesia," pp. 330–335, 2020.
- [2] Y. R. Suci, S. Tinggi, and I. Ekonomi, "Perkembangan UMKM (Usaha Mikro Kecil Menengah) di Indonesia," *J. Ilm. Fak. Ekon.*, vol. 6, no. 1, pp. 51–58, 2017.
- [3] N. Hikmah, D. Ariyanti, and M. Sugesti, "Penerapan Teknik Data Mining untuk Clustering Armada pada PT. Siaga Transport Indonesia Menggunakan Metode k- Means," *Explore*, vol. 9, no. 1, p. 8, 2019, doi: 10.35200/explore.v9i1.116.
- [4] L. Magdalena and R. Fahrudin, "Penerapan Data Mining Untuk Koperasi Se-Jawa Barat Menggunakan Metode Clustering pada Kementerian Koperasi dan UKM," *J. Digit*, vol. 9, no. 2, pp. 190–201, 2019.