

Klasterisasi Data Supplier Menggunakan Algoritma Fuzzy C-Means (Studi Kasus: PT. Jogja Rekayasa Engineering)

M. Lukman Prayoghi¹, Arief Setyanto², Eko Pramono^{3*}

^{1,3} Teknik Informatika Fakultas Ilmu Komputer Universitas Amikom Yogyakarta

Jl. Ring Road Utara, Condong Catur, Sleman, Yogyakarta, INDONESIA

¹prayoghi.indo@gmail.com, ²arief_s@amikom.ac.id, ³eko.p@amikom.ac.id

INTISARI

Penelitian dengan studi kasus pada PT Jogja Rekayasa Engineering ini adalah sebuah pengembangan dari aplikasi pencatatan keuangan, yang termasuk di dalamnya, kontrak, rencana anggaran proyek, rencana anggaran belanja, beserta data vendor, juga pekerjaan dan barang yang di supply. menjadi suatu tolok ukur pengambilan keputusan pemilihan vendor supplier yang tepat bagi perusahaan PT. Jogja Rekayasa Engineering yang sebelumnya sudah di implementasikan dengan aplikasi berbasis web. Proses penelitian memanfaatkan metode Fuzzy C-Means untuk klasterisasi pada data yang di ambil dari database aplikasi, penelitian ini di harapkan agar perusahaan dapat memilih serta menentukan supplier mana yang terbaik untuk dapat menjalin kerja sama, dengan memanfaatkan beberapa variabel pengukuran yang tersedia dari banyaknya data yang sudah di klaster. Penelitian sendiri akan berupa simulasi data yang besar untuk proses clustering itu sendiri. Ketika proses cluster dengan algoritma Fuzzy C-Means sudah berjalan seperti yang di harapkan, kemudian akan ditanamkan pada fitur aplikasi perusahaan untuk menentukan keputusan pemilihan vendor, yang akan di kerjakan oleh tim web developer mereka.

Kata Kunci — SPK, Clustering, FCM, Fuzzy, Supplier, Vendor.

ABSTRACT

This research with case study at PT Jogja Engineering Engineering is a development of financial recording application, including contract, project budget plan, budget plan, beside the vendor data, also work and goods in supply. become a benchmark decision making vendor selection appropriate supplier for company PT. Jogja Engineering Engineering which previously been implemented with web-based applications. The research process utilizes the Fuzzy C-Means method for clustering the data taken from the application database, this research is expected to enable the company to choose and determine which supplier is best to be able to work together, by utilizing several measurement variables available from the number of data already in the cluster. The research itself will be a large data simulation for the cluster-ing process itself. When the cluster process with the Fuzzy C-Means algorithm is already running as expected, then it will be embedded in the enterprise application features to determine the vendor selection decision, which will be done by their web developer team.

Keyword – DSS, Clustering, FCM, Fuzzy, Supplier, Vendor..

I. PENDAHULUAN

PT. JRE yang menjadi studi kasus penelitian ini sendiri merupakan sebuah perusahaan mesin yang memproduksi mesin te-pat guna dengan kandungan lokal. Dalam pengerjaan bagian-bagian mesinnya, PT JRE bekerja sama dengan perusahaan penyedia material dalam negeri lainnya. Untuk satu mesinnya saja perusahaan ini bekerja sama dengan hampir 20 perusahaan di Indonesia untuk mendapatkan partisi mesin terbaik. Pada suatu kasus, pernah terjadi kesulitan dalam menentukan vendor mana yang akan di pilih untuk menjadi supplier karena

berbagai aspek yang sulit di putuskan secara tepat, seperti banyaknya penawaran yang masuk dari beberapa supplier dengan harga, lokasi, dan kualitas yang juga berbeda-beda satu dan lainnya. Dari beberapa data perusahaan yang telah bekerja sama, diperlukan klasterisasi data serta mengelompokkan vendor mana yang layak untuk kembali diajak bekerja sama pada proyek berikutnya, dengan memperhatikan hasil kerja, efisiensi biaya dan lokasi dari vendor perusahaan yang sudah tersedia saat ini, tidak menutup kemungkinan data akan semakin banyak untuk beberapa tahun kedepan.

Dengan memanfaatkan konsep data mining untuk pengambilan keputusan adalah solusi yang tepat pada era perkembangan digital saat ini, lebih tepatnya merancang algoritma untuk pengembangan aplikasi penunjang keputusan yang akan dikerjakan oleh IT developer PT JRE

II. TINJAUAN PUSTAKA

Klasterisasi nilai citra RGB pada buah papaya madu berdasarkan mutu buah menggunakan algoritma FCM dilakukan oleh Sila Abdullah S, Mulyadi. (2017) Penelitian ini mengklaster beberapa data buah berdasarkan warna dan menghasilkan tiga kondisi sesuai data real papaya dengan sortir menggunakan panca indra yakni kondisi matang mentah $R < GB$ $R=9.5825$, $GB=17.8833$, Kondisi matang mengkal $R=GB$ $R=17.6735$, $GB=17.2993$. Kondisi matang matang $R > GB$ $R=24.6697$, $GB=16.6703$

Kemudian Esmael Mehdizadeh (2016) mengklaster serta membuktikan efektifitas aturan kelistrikan menggunakan algoritma FCM. Isidorus Cahyo Adi (2016) menganalisa data penjualan untuk meprediksi pendistribusian produk helm XYZ juga dengan algoritma FCM. Penelitiannya menunjukkan nilai sum of square terendah dalam percobaan dan menghasilkan 8 cluster sebagai output penelitiannya.

Irma Irandha, Arna Fariza (2017) mengklaster data keluarga miskin di Kota Surabaya, kecamatan Wonocolo dengan menggunakan metode Fuzzy C-Means, kemudian di implemetasikan dengan menggunakan web based app dan menghasilkan output label yang telah di klaster, setelah sebelumnya algoritma sudah ditanamkan pada aplikasi. Kesimpulan dari penelitian ini adalah klasterisasi menggunakan Fuzzy C-Means dapat dilakukan dan menghasilkan data yang seperti diharapkan untuk di Analisa lebih lanjut.

Prisma Wahyu Wulan (2017) melakukan penelitian dengan merancang aplikasi komputerisasi pencatatan sampah masuk ke TPA Klotok dengan menerapkan metode K-Means. Guna mengetahui daerah-daerah mana saja yang memasok sampah terbesar di wilayah kota Kediri dengan mendata berat sampah, volume sampah, nomor polisi kendaraan, pengangkut sampah dan daerah sampah berasal.

Sandi Fajar Rodiansyah (2013) melakukan pengklasifikasian data yang di crawling dari twitter, kemudian menerapkan algoritma naïve Bayesian. Dalam jurnal ini membantu peneliti agar tetap dalam benang merah bahwasanya teknik klaster dan klasifikasi adalah dua hal yang berbeda walau sebenarnya masih dalam satu koridor dalam data mining.

III. METODOLOGI PENELITIAN

Tahap awal dalam penelitian ini sendiri adalah tahap membangun data dengan aplikasi web based yang sedang dikembangkan oleh PT. JRE, dengan mengikuti aturan input dari perusahaan akan menghasilkan aplikasi yang terkoneksi dengan database internal dengan data berupa informasi supplier yang telah dan akan bekerja sama dengan PT. JRE, material apa saja yang di supplai, harga, lokasi dan ketepatan waktu dalam pengerjaan akan menjadi tolak ukur dalam tahapan klasterisasi.

Kemudian dalam penelitian ini penulis menggunakan pendekatan kuantitatif dengan memanfaatkan Diagnosing Action, dan Taking Action.

Dalam tahapan pengumpulan data sendiri. Penulis melakukan wawancara serta konsultasi dengan CEO PT. JRE untuk membangun sebuah aplikasi sistem informasi yang akan menjadi sebuah basis data yang diperlukan untuk klasterisasi data supplier nantinya.

IV. LANDASAN TEORI

Definisi data dalam Webster New World tertulis bahwasanya “something known or assumed”. Artinya, data merupakan suatu yang diketahui/ dianggap. Dengan demikian, data dapat memberi gambaran tentang suatu keadaan atau persoalan. Sedangkan, data menurut kamus Oxford Dictionary adalah The Facts. Jadi, dapat disimpulkan bahwa sesuatu yang bersifat real juga diketahui untuk keperluan suatu Analisa, diskusi, presentasi ilmiah atau yang lainnya. Menurut Ong (2013) data dibagi menjadi 4 bagian, yaitu: Jenis data berdasarkan Sifatnya, jenis data menurut sumbernya, jenis data menurut cara memperolehnya, jenis data menurut waktu pengumpulannya.

Clustering merupakan proses pengelompokan record, pengamatan atau memperhatikan dan membentuk kelas objek-

objek yang memiliki kemiripan. Cluster adalah kumpulan record yang memiliki kemiripan antara satu dan lainnya, juga memiliki ketidakmiripan dengan record dalam cluster yang lain.

Clustering berbeda dengan klasifikasi, dimana tidak adanya variable target dalam clustering. Proses ini tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variable target. Akan tetapi, algoritma clustering mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (Homogen), yang dimana kemiripan pada record di dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan record dalam kelompok lain akan bernilai minimal. Clustering pada data sendiri untuk memberikan label kelas pada data dan merupakan salah satu metode data mining. Pada penelitian yang dilakukan oleh Jimmy, Sherwin[1] tentang segmentasi citra spot dengan menggunakan pendekatan Fuzzy C-means menyimpulkan bahwa metode logika samar (Fuzzy C-Means Clustering) memiliki tingkat kestabilan output/ hasil yang lebih baik daripada pendekatan metode konvensional (K-Means Clustering).

Fuzzy C-Means (FCM) dalam Clustering didasarkan pada teori logika fuzzy. Teori ini pertama kali diperkenalkan oleh Lotfi Zadeh (1965) dengan nama himpunan fuzzy (fuzzy set). Dalam teori fuzzy, keanggotaan sebuah data tidak diberikan nilai secara tegas dengan nilai 1 dan 0 melainkan dengan suatu nilai derajat keanggotaan yang jangkauan nilainya 0 sampai 1. Nilai keanggotaan menjadi 0 ketika sama sekali tidak menjadi anggota, dan menjadi 1 ketika menjadi anggota secara penuh dalam suatu himpunan. Semakin tinggi nilai keanggotaannya maka semakin tinggi derajat keanggotaannya dan semakin kecil maka semakin rendah derajat keanggotaannya.

V. HASIL DAN PEMBAHASAN

Memanfaatkan algoritma Fuzzy C-Means untuk klusterisasi dengan menentukan pusat cluster yang akan menandai lokasi rata-rata untuk setiap cluster. Dengan cara memperbaiki pusat cluster dan derajat keanggotaan setiap titik data secara berulang. Data yang akan di proses berupa data contoh, dengan menentukan beberapa variable pendukung yang akan di temukan saat proses observasi di lapangan. Sebagai contoh seperti

pada tabel 3, dataset vendor mempunyai 3 variabel, jarak dalam satuan kilometer, harga dalam nilai random, dan kelayakan hanya mempunyai 2 nilai, yaitu 1 = tidak layak dan 2 = layak..

TABEL I.
CONTOH DATA SET

No	Vendor	Jarak	Harga
1	CV PRIMA BAJA TEKNIK	321.6	85
2	CV SARANA LOGISTIK	321.6	80
3	CV SURYA DOTA TEKNIK	122.9	83
4	CV TALENTA HARAPAN	93.3	70
5	TOKO BESI T	122.9	68
6	BENGKEL LUHUR YOGYAKARTA	15.8	65
7	CV GLOBAL MEDIATECH	9.4	64
8	CV BENGKEL IDEK	3.2	72
9	PT BENGKEL API REFRACTORINDO	256.5	69
10	PT LIANG CHI INDONESIA	122.9	75
11	WIRATAMA ANUGRAH	321.6	75
12	CV MITRA ABADI TEKNIK	23.2	72
13	PT LINICO INDONESIA	417.5	81
14	PT SOLUSINDO TIRTATEK	417.5	71
15	CV BANGKIT	60.2	80
16	PT SHAHS DAYA TEKNIK	417.5	83
17	PT ELECTIC CONSULTING	538.2	70
18	PT EVERALD PRIMA PERKASA	563.7	68
19	PT PRIMA UNTUNG BERSAMA	321.6	65
20	PT MITSUBISHI ELECTRIC INDONESIA	538.2	64

Output dari Fuzzy C-Means sendiri bukanlah merupakan Fuzzy Inference System, namun merupakan deretan pusat cluster dan beberapa derajat keanggotaan untuk tiap titik data. Kemudian menentukan matriks x data yang akan di cluster, berukuran k x j, dengan k = jumlah data yang akan di cluster. Dan j = jumlah variable/atribut.

$$\begin{bmatrix} X_{11} & X_{12} & \dots & X_{1j} \\ X_{21} & X_{22} & \dots & X_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ X_{k1} & X_{k2} & \dots & X_{kj} \end{bmatrix}$$

Kemudian menentukan jumlah cluster yang akan dibentuk, pembobot, maksimum iterasi (max n), kriteria penghentian, serta menentukan fungsi obyektif awal. Setelah itu membentuk matriks partisi awal U, guna mencari derajat keanggotaan dengan ukuran k x i, k = jumlah data dan i = jumlah cluster.

$$\begin{bmatrix} U_{11} & U_{12} & \dots & U_{1j} \\ U_{21} & U_{22} & \dots & U_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ U_{k1} & U_{k2} & \dots & U_{kj} \end{bmatrix}$$

Misalkan matriks partisi awal U yang terbentuk secara random sebagai berikut:

TABEL II.
CONTOH PSEUDO-PARTITAL

	c1	c2	c3
1	0.3	0.3	0.4
2	0.3	0.5	0.2
3	0.8	0.1	0.1
4	0.5	0.2	0.3
5	0.5	0.1	0.4
6	0.2	0.1	0.7
7	0.3	0.4	0.3
8	0.6	0.2	0.2
9	0.5	0.2	0.3
10	0.5	0.1	0.4
11	0.2	0.1	0.7
12	0.3	0.4	0.3
13	0.3	0.5	0.2
14	0.8	0.1	0.1
15	0.5	0.2	0.3
16	0.5	0.1	0.4
17	0.5	0.2	0.3
18	0.5	0.1	0.4
19	0.2	0.1	0.7
20	0.3	0.4	0.3

Menghitung pusat cluster dengan rumus:

$$V_{ij} = \frac{\sum_{k=1}^n (\mu_{ik})^w \cdot x_{kj}}{\sum_{k=1}^n (\mu_{ik})}$$

Keterangan :

V_{ij} = pusat cluster pada cluster ke-i dan atribut ke-j.+

μ_{ik} = data partisi (pada matriks U) pada cluster ke-i dan data ke-k.

X_{kj} = data (pada matriks U) pada atribut ke-j dan data ke-k.

w = pembobot

Dengan rumus diatas, total dari cluster 1 yang di iterasi di kalkulasikan dengan total kuadrat dari cluster yang sama. Maka menghasilkan Pusat Cluster sebagai berikut:

TABEL III.
PUSAT CLUSTER

C1	C2	C3
0.5	0.15814	0.353488
123.7213	37.04302	0.755814
42.71709	11.72907	0.239535
86.33628	25.11163	0.513953

Kemudian menghitung nilai obyektif (P n) dengan rumus :

$$P_n = \sum_{k=1}^n \sum_{i=1}^n (\mu_{ik})^w (d_{ik})^2$$

Keterangan :

μ_{ik} = data partisi (pada matriks U) pada cluster ke-i dan data ke-k.

d_{ik} = fungsi ukuran jarak untuk jarak Euclidean pada pusat cluster ke-i dan data ke-k.

w = pembobot.

P_n = nilai obyektif pada iterasi ke-n.

Pada tahapan perbaikan derajat keanggotaan setiap data cluster akan dihitung menggunakan matriks partisi. Dengan rumus sebagai berikut:

$$\begin{aligned} \mu_{ik} &= \left[\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{w-1}} \right]^{-1} d_{ik} \\ &= d(x_k - v_i) \\ &= \left[\sum_{j=1}^m (x_{kj} - v_{ij}) \right]^{1/2} \end{aligned}$$

Keterangan:

μ_{ik} = data partisi (pada matriks U) pada pusat cluster ke-i dan data ke-k.

d_{ik} = fungsi ukuran jarak untuk jarak Euclidean pada pusat cluster ke-i dan data ke-k.

d_{jk} = fungsi ukuran jarak untuk jarak Euclidean pada pusat cluster ke-j dan data ke-k.

w = pembobot.

X_{kj} = data (pada matriks U) pada atribut ke-j dan data ke-k.

TABEL IV.
MIU KUADRAT

0.09	0.1	0.16
0.09	0.3	0.04
0.64	0	0.01
0.25	0	0.09
0.25	0	0.16
0.04	0	0.49
0.09	0.2	0.09
0.36	0	0.04
0.25	0	0.09
0.25	0	0.16
0.04	0	0.49
0.09	0.2	0.09
0.09	0.3	0.04
0.64	0	0.01
0.25	0	0.09
0.25	0	0.16
0.25	0	0.09
0.25	0	0.16
0.04	0	0.49
0.09	0.2	0.09

4.3	1.4	3.04
-----	-----	------

TABEL VI.
X V

110303.8	41456.03	83144.88	58935.87
109480.4	41001.46	82437.18	58361.98
21844.95	2112.78	11509.42	4688.207
13492.44	2013.149	5957.231	2065.67
19586.99	960.5579	9598.816	3178.527
4438.975	12428.6	3562.9	6566.69
4157.704	13797.52	3845.378	7433.71
5168.961	15747.36	5194.764	9110.403
70275.62	18651.57	48983.67	30882.11
20585.78	1442.956	10435.61	3827.965
108709.2	40598.38	81781.99	57840.07
5676.961	11326.52	4014.08	6184.951
180427.1	88239.87	145263.8	112795.1
178910.3	87460.73	143978.4	111777.3
9939.231	5880.281	4967.15	3696.079
180752.2	88418.21	145542.4	113020.7
294001.9	172880.6	248901.9	206198
321799.5	194541.4	274592.7	229717.7
107310.1	39937.75	80614.05	56940.33
293199.8	172521.1	248238.7	205695.3

TABEL V.
X1 X2 X3

28.944	7.65	0.09	28.944	7.65	0.09	51.466	13.8	0.16
28.944	7.2	0.09	80.4	20	0.25	12.864	3.2	0.04
78.866	53.12	0.64	1.229	0.83	0.01	1.229	0.83	0.01
23.325	17.5	0.5	3.732	2.8	0.08	8.397	6.3	0.18
30.725	17	0.5	1.229	0.88	0.02	19.664	10.88	0.32
0.632	2.6	0.04	0.158	0.85	0.01	7.742	31.85	0.49
0.846	5.76	0.18	1.504	10.24	0.32	0.846	5.76	0.18
1.162	25.92	0.36	0.128	2.88	0.04	0.128	2.88	0.04
64.125	17.25	0.25	10.26	2.76	0.04	23.085	6.21	0.09
30.725	18.75	0.5	1.229	0.75	0.02	19.664	12	0.32
12.864	3	0.08	3.216	0.75	0.02	157.584	38.75	0.98
2.088	6.48	0.09	3.712	11.52	0.16	2.088	6.48	0.08
37.575	7.29	0.18	104.375	20.25	0.5	16.7	3.24	0.08
287.2	45.44	1.28	4.175	0.71	0.02	4.175	0.71	0.02
15.05	20	0.25	2.408	3.2	0.04	5.418	7.2	0.09
104.38	20.75	0.25	4.175	0.83	0.01	66.8	13.28	0.16
134.55	17.5	0.5	21.528	2.8	0.08	48.438	6.3	0.18
140.93	17	0.5	5.637	0.88	0.02	90.192	10.88	0.32
12.864	2.6	0.04	3.216	0.85	0.01	157.584	31.85	0.49
48.438	5.76	0.18	86.112	10.24	0.32	48.438	5.76	0.18

1084	318.57	8.5	387.387	100.87	2.08	742.462	216	4.42
------	--------	-----	---------	--------	------	---------	-----	------

TABEL VII.
L

9927.339	3731.043	13303.18	26961.56
9853.232	10250.37	3297.487	23401.08
13980.77	21.1278	115.0942	14116.99
3373.109	80.52596	536.1508	3989.786
4896.747	9.605579	1535.811	6442.163
177.559	124.286	1745.821	2047.666
374.1934	2207.603	346.084	2927.88
1860.826	629.8945	207.7905	2698.511
17568.9	746.0629	4408.53	22723.5
5146.444	14.42956	1669.697	6830.571
4348.369	405.9838	40073.17	44827.53
510.9265	1812.244	361.2672	2684.437
16238.44	22059.97	5810.552	44108.96
114502.6	874.6073	1439.784	116817
2484.808	235.2112	447.0435	3167.062
45188.05	884.1821	23286.78	69359.01
73500.47	6915.222	22401.17	102816.9
80449.87	1945.414	43934.83	126330.1
4292.404	399.3775	39500.88	44192.66
26387.98	27603.37	22341.48	76332.83

Fungsi Objektif & Selisih Fungsi Objektif = **742776.2**

Dengan metode analisis data menggunakan algoritma Fuzzy C-Means ini sendiri sampai rumus yang dijabarkan terakhir perlu beberapa langkah penelitian mengingat data yang didapatkan di lapangan akan seperti apa. Setelah menghentikan iterasi jika pusat cluster V tidak berubah. Cluster akan dipilih berdasarkan nilai matriks partisi terbesar, tentu dengan melalui beberapa iterasi terlebih dahulu.

IV. KESIMPULAN

Dari percobaan iterasi pertama dengan variable yang ada, belum sampai pada hasil akhir yang memuaskan, maka perlu di lakukan beberapa iterasi lagi dengan variable pendukung yang ditemukan di tempat penelitian.

REFERENSI

- [1] Suyanto, 2017, Data Mining Untuk Klasifikasi dan Klasterisasi Data, Informatika Bandung, Bandung.
- [2] Efraim Turban, Jay E. Aronson, Ting-Peng Liang, 2007, Decision Support System and Intelligent Systems, Seventh Edition, Prentice-Hall, New Delhi.
- [3] Eko Prasetyo, 2014, Data Mining Mengolah Data Menjadi Informasi menggunakan Matlab, Andi Publishing, Yogyakarta.
- [4] Kusriani, 2007, Konsep dan Aplikasi Sistem Pendukung Keputusan, First Edition, Andi Publishing, Yogyakarta.
- [5] Irma Irandha, Arna Fariza, 2017, Analisa Keluarga Miskin dengan Menggunakan Metode Fuzzy C-Means Clustering, TECHCSI Surabaya.
- [6] Sandi Fajar, 2013, Klasifikasi Posting Twitter Kemacetan Lalu Lintas Kota Bandung Menggunakan Naïve Bayesian Classification, IJCCS, Bandung
- [7] Esmael Mehdizadeh., 2016, Electrical fuzzy C-Means: A new heuristic fuzzy clustering algorithm, Cogent Engineering, Qazvin, Iran.
- [8] Sylvia Jane, 2007, Fuzzy C-Means Clustering Model Data Mining For Recognzing Stock Data Sampling Pattern. IJCCS, Medan, Indonesia.
- [9] Isidorus, C. A., 2016, Implementasi Data Mining menggunakan Metode Clustering Untuk Prediksi Penjualan di PT. XYZ, Skripsi, Teknik Informatika, Universitas Sanata Dharma, Yogyakarta.
- [10] Prisma Wahyu WD., 2017, Klasterisasi Daerah Pemasok Sampah Terbesar ke TPA Klotok ke TPA Klotok di Kota Kediri dengan Metode K-Means Method, Skripsi, Teknik Informatika, UNP, Kediri.