



Clustering Sampah yang Dihasilkan Oleh Masyarakat Kota Surabaya dengan Menggunakan Metode K-Means

Tasya Febrinda Ardika Putri¹, Hilyatun Nuha², dan Sigit Ananda Murwato³

^{1,2,3}Jurusan Teknik Industri, Fakultas Teknik, Universitas 17 Agustus 1945 Surabaya, Jl Semolowaru No. 45 Surabaya, Indonesia

INFORMASI ARTIKEL

Halaman:

38 – 46

Tanggal penyerahan:

2 Juni 2023

Tanggal diterima:

19 September 2023

Tanggal terbit:

30 September 2023

ABSTRACT

The Surabaya environmental service said that around 1600 tons of waste enters the Surabaya landfill per day. The unstable increase means that the pile of rubbish in Surabaya is increasingly experiencing fluctuating tonnage, this is because the higher human growth, the more rubbish in Surabaya will increase. Continuing social life will create many new types of waste. According to Law No. 18 of 2018 concerning waste management, population growth and changes in people's consumption patterns have resulted in an increase in the volume, types and characteristics of increasingly diverse waste. Waste processing in the city of Surabaya can be done by grouping or clustering waste based on the type of waste in each waste bank to make it easier for people to process waste. The results of the cluster model from manual data processing with rapidminer show that from the total number of items of 647 there are 11 areas with the highest clusters in the city of Surabaya. The results of manual data validation tests using the Davies Bouldin Index (DBI) show 2 clusters with the strongest min-max data normalization because they are close to 0 with a cluster value of $k = 2$ resulting in 0.268.

Keywords: Clustering, K-Means, Garbage, Surabaya, Technology

EMAIL

¹tasyafebrinda@surel.untag-sby.ac.id

²hilyatun_n@untag-sby.ac.id

³sigitananda@untag-sby.ac.id

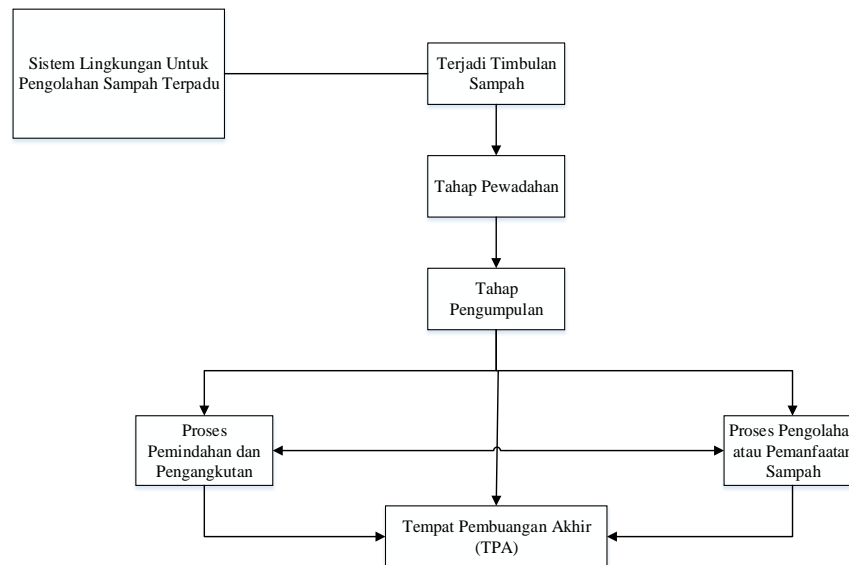
ABSTRAK

Dinas lingkungan hidup Surabaya mengatakan bahwa sampah yang masuk di TPA Surabaya sekitar 1600 ton per hari. Peningkatan yang tidak stabil membuat tumpukan sampah yang ada di Surabaya semakin mengalami tonase yang berubah-ubah, hal ini dikarenakan semakin tinggi pertumbuhan manusia akan meningkat juga sampah yang ada di Surabaya. Kehidupan masyarakat yang terus berlanjut akan membuat banyak jenis sampah yang baru. Menurut UU No 18 Tahun 2018 tentang pengelolaan sampah bahwa penambahan penduduk dan perubahan pola konsumsi masyarakat menimbulkan bertambahnya volume, jenis, dan karakteristik sampah yang semakin beragam. Pengolahan sampah yang ada di kota Surabaya ini dapat dilakukan dengan mengelompokan atau clustering sampah berdasarkan jenis sampah nya pada masing – masing bank sampah untuk mempermudah masyarakat mengolah sampah. Hasil *cluster model* dari pengolahan data manual dengan rapidminer menunjukkan bahwa dari total number of items sebesar 647 terdapat 11 wilayah dengan *cluster* tertinggi di kota Surabaya. Hasil uji validasi data manual menggunakan *Davies Bouldin Index* (DBI) menunjukkan 2 *cluster* dengan normalisasi data min-max paling kuat karena mendekati angka 0 dengan nilai *cluster* $k = 2$ hasil yang menjadi 0.268.

Kata kunci: Clustering, K-Means, Surabaya, Sampah, Teknologi

PENDAHULUAN

Sampah dapat diartikan sebagai sesuatu yang terjadi karena perilaku makhluk hidup, dengan hal ini dapat berbentuk sebuah padatan, cairan, dan gas atau berupa zat organik dan anorganik bersifat dapat terurai atau tidak dapat diuraikan kembali dan akan dianggap sudah tidak dapat digunakan kembali lalu dibuang oleh manusia ke lingkungan. Menurut UU No 18 Tahun 2018 tentang pengelolaan sampah menyebutkan jika pertumbuhan penduduk dan perubahan pola konsumsi masyarakat dapat menimbulkan bertambahnya volume, jenis, dan karakteristik sampah yang semakin beragam dan bermacam – macam.



Gambar 1 Sistem Lingkungan Pengolahan Sampah

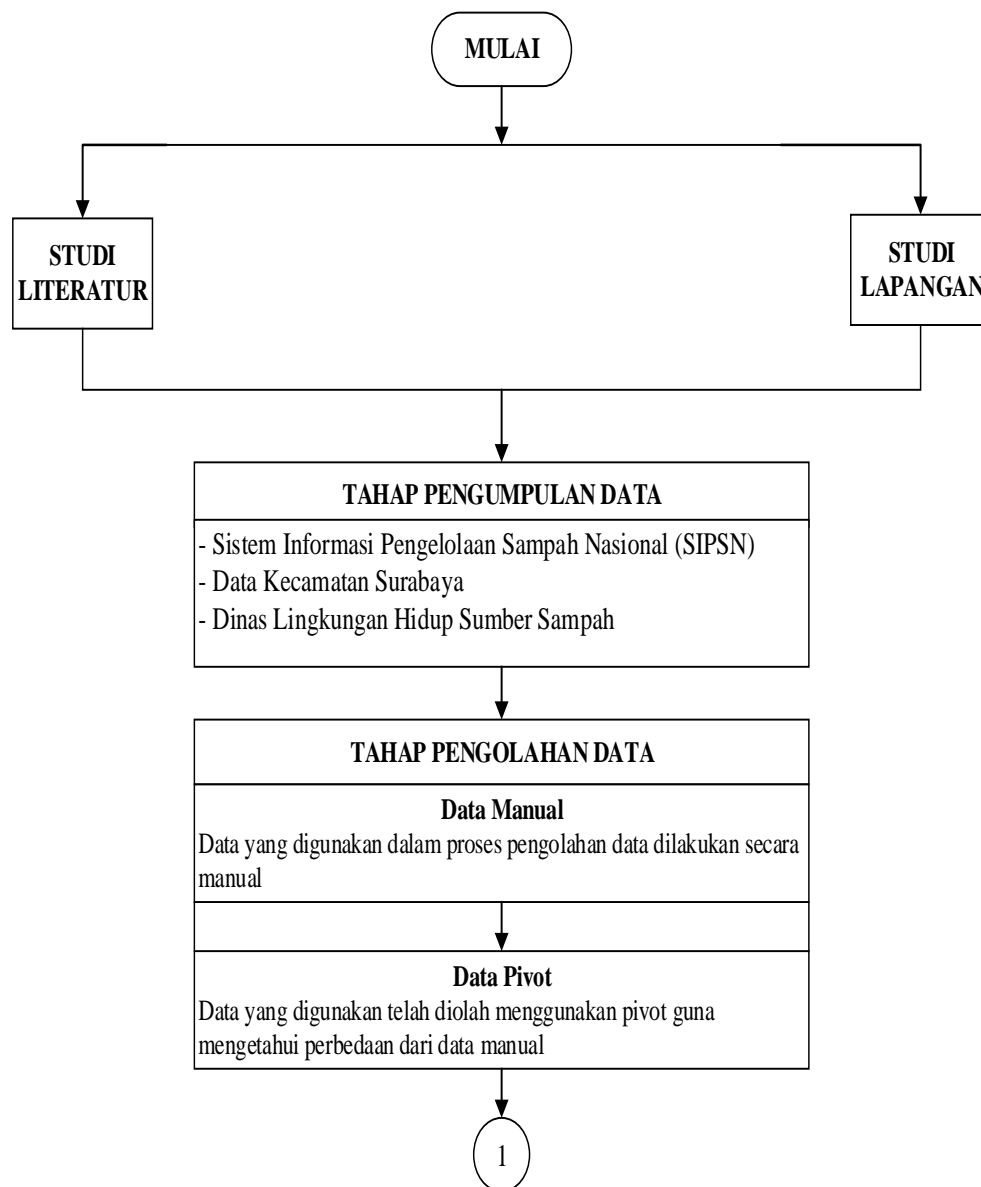
Proses dalam mengolah sampah jika dilihat dalam sistem lingkungan terjadi seperti aliran di atas dengan proses awal yang berasal dari timbulan sampah lalu di arahkan menuju proses selanjutnya yaitu pewadahan dan pengumpulan dimana sampah yang telah ditimbun dikumpulkan menjadi satu guna proses selanjutnya, setelah proses pengumpulan sampah selanjutnya adalah proses pengolahan atau pemanfaatan dari sampah yang telah dikumpulkan, pada proses pengolahan sampah ini tidak hanya diolah tetapi juga langsung dilakukan proses pemindahan agar tidak menumpuk di satu tempat setelah dilakukan proses pengolahan. Proses dari pengolahan tidak hanya dilakukan satu kali tetapi juga akan dilakukan beberapa proses dilihat dari volume sampah yang dihasilkan seberapa besar, semakin besar volume sampah yang dihasilkan maka semakin membutuhkan waktu yang lama pula untuk pengolahan sampah tersebut [1].

Selain perkembangan industri *manufacturing* di kota Surabaya yang semakin meningkat, peran mahasiswa atau lingkungan kampus juga memiliki pengaruh yang cukup tinggi untuk peningkatan sampah yang ada di kota Surabaya, pada beberapa kampus telah memiliki program untuk pengurangan sampah pada kampus tersebut dengan membuat pengelolaan sampah plastik botol, Surabaya telah memiliki program khusus bagi lingkungan kampus untuk membuat strategi supaya sampah yang dihasilkan wilayah kampus tidak terlalu tinggi dan bisa dikelola dengan baik di wilayah kampus yaitu dengan program *eco Campus*. *Eco Campus* memiliki pencapaian terhadap pengelolaan sumber daya alam seperti limbah air serta pencegahan lingkungan sekitar kampus, dengan adanya program khusus di kampus maka dapat mengurangi sampah yang ada di pengelolaan sampah kota Surabaya [2]

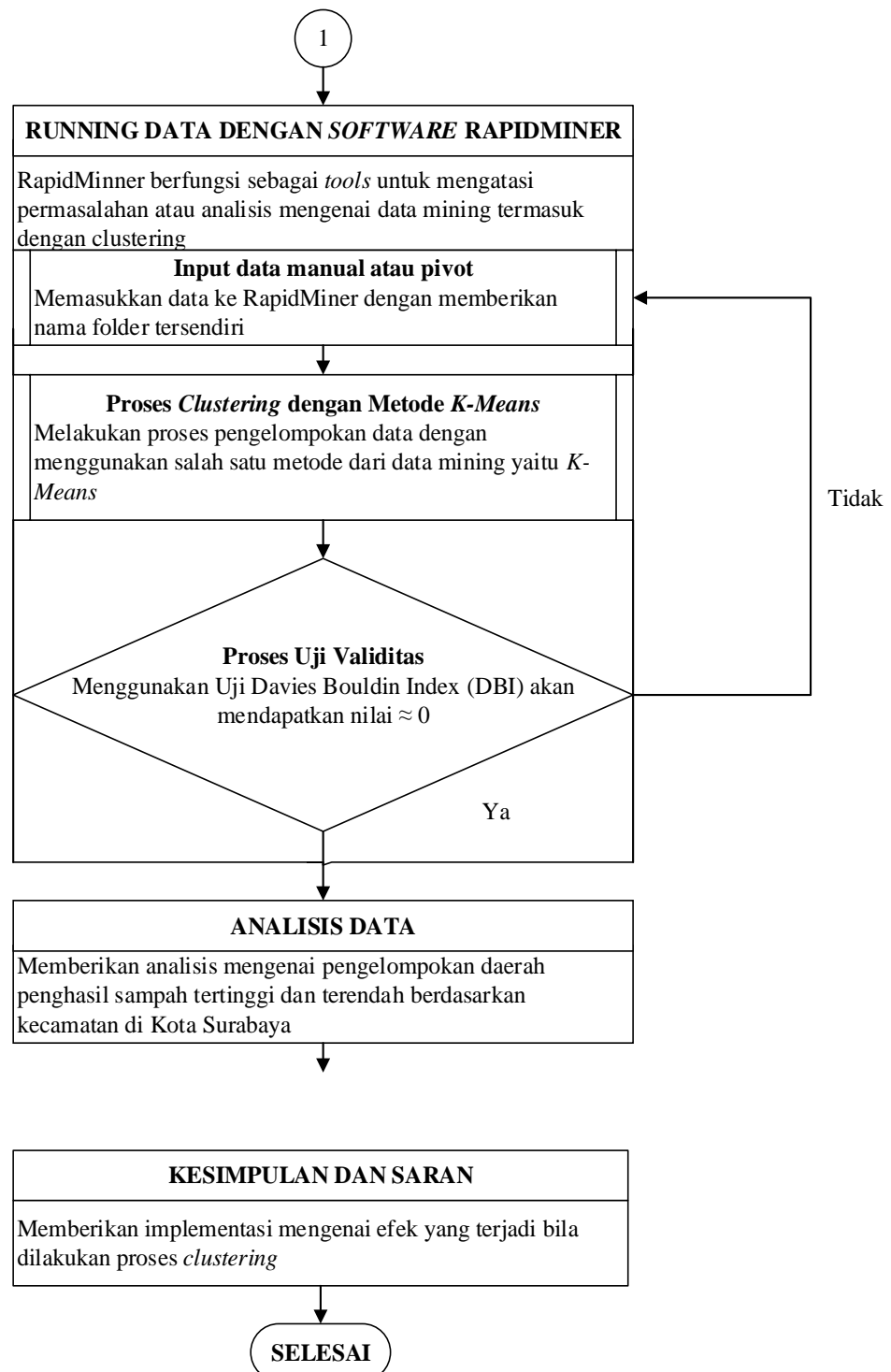
Pengoptimalan untuk wilayah Surabaya penghasil sampah tertinggi harus memiliki penanganan yang lebih baik tetapi tepat pada sasaran, oleh karena itu perlu adanya teknik pengelompokan yang terdapat pada data mining dengan beberapa metode pengolahan, yaitu dengan menggunakan metode clustering *k-means*. *Clustering k-means* merupakan metode dengan mengelompokkan variabel – variabel penelitian ke dalam *clustering* atau *polynomial* yang ada pada *clustering*, dengan adanya proses *clustering* menggunakan metode *k-means* maka nantinya akan diketahui wilayah Surabaya yang mengalami kriteria penanganan lebih guna mengoptimalkan sampah [3]. Pengelompokan menggunakan metode *k-means* diharapkan dapat membantu pemerintah kota Surabaya guna lebih memperhatikan wilayah yang mengalami penanganan secara khusus, sehingga dapat memberikan saran penanganan atau inovasi setelah mengetahui proses pengolahan data dari clustering wilayah penghasil sampah tertinggi di kota Surabaya. Karena jika tidak diatasi mulai dari sekarang maka akan semakin tinggi tingkat ketidaksadaran manusia terhadap lingkungan sekitar, jika terdapat bukti tingkat pengklusteran pada daerah kota Surabaya dengan wilayah tertinggi untuk penghasil sampah maka masyarakat sekitar dapat menjadi penyalur yang baik jika nantinya mendapat bantuan inovasi maupun pengolahan terhadap sampah.

Tabel 1 Penelitian Terdahulu

Peneliti	Tahun	Judul	Hasil Penelitian
Yulia Darmi, Agus Setiawan	2016	Penerapan Metode <i>Clustering K-Means</i> Dalam Pengelompokan Penjualan Produk	Pengklasteran dengan menggunakan metode <i>K-Means</i> didapatkan untuk produk yang laku dan tidak laku
Benri Melpa Metisen, Herlina Latipa Sari	2015	Analisis <i>Clustering</i> Menggunakan Metode <i>K-Means</i> Dalam Pengelompokan Penjualan Produk Pada Swalayan Fadhlila	Proses <i>cluster</i> secara hirarki dengan menggunakan metode <i>K-Means</i> menghasilkan sebuah informasi gambaran penjualan terkluster atau terkelompok.
Tutik Khotimah	2014	Pengelompokan Surat Dalam Al-Quran Menggunakan Algoritma <i>K-Means</i>	Telah dilakukan pengelompokan surat-surat dalam Al-Quran dengan menggunakan algoritma <i>K-Means</i>
Winda Pradnya Dhuhiha	2015	<i>Clustering K-Means</i> Menggunakan Metode Menentukan Status Gizi Balita	Telah membandingkan hasil pengelompokan menggunakan table Growth Chart dan Algoritma <i>K-Means</i> didapat 17 data yang memiliki kelompok sama
Herman Suryadi		Pengklasteran Area Pembangunan Bak Sampah Daun Menggunakan Metode <i>K-Means</i>	Hasil dari penelitian tersebut adalah untuk rekomendasi hasil <i>clustering</i> yang dipilih yaitu dengan 3 <i>cluster</i> untuk tumpukan sampah karena memiliki nilai tertinggi pada <i>Silhouette Coefficient</i>

METODE

Gambar 2 Metode penelitian



Gambar 3 Metode penelitian lanjutan

1. Studi Literatur

Pada bagian ini menjelaskan mengenai bagaimana untuk menyelesaikan permasalahan pada proposal tugas akhir ini dengan membaca beberapa permasalahan yang sama dengan topik yang sedang di teliti dengan mengulas kembali literatur terdahulu, studi literatur juga dapat ditemukan pada buku serta referensi – referensi lain yang terkait dengan topik atau judul yang sedang di teliti.

2. Studi Lapangan

Penelitian ini akan mengacu pada kriteria sampah yang ada di suraabaya dengan berdasarkan jenis, sumber, berat sampah yang ada pada wilayah Surabaya serta mencari informasi mengenai

pengelolaan sampah di Surabaya, pada penelitian ini juga akan dibutuhkan jumlah bank sampah guna mengetahui lokasi untuk masyarakat dapat mengumpulkan sampah.

3. Tahap Pengumpulan Data

Pada tahap pengumpulan data untuk dapat menyelesaikan permasalahan mengenai sampah dengan nantinya akan menggunakan metode pengelompokan *k-means* maka data yang dibutuhkan yaitu Sistem Informasi Pengelolaan Sampah Nasional (SIPSN) data berupa sampah masuk, sampah terkelola, sampah daur ulang serta sampah up-cycle, data kecamatan di Surabaya untuk melihat wilayah yang ada di Surabaya, data dinas lingkungan hidup yaitu mengenai sumber sampah serta jenis sampah yang ada di kota Surabaya.

4. Tahap Pengolahan Data

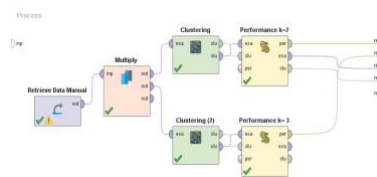
Pada tahapan ini dapat dilakukan sebuah perhitungan dengan menggunakan 2 data yaitu dari data perhitungan dari rata – rata dan melalui data yang telah dipivot dan ditambahkan data jumlah penduduk serta luas wilayah per kecamatan, dalam proses pengelolaan data ini menggunakan *tools excel* guna mempermudah dan mempercepat pengolahan data awal setelah membagi data tersebut maka data – data tersebut diolah dengan menggunakan *tools* atau *software* rapidminer.

HASIL DAN PEMBAHASAN

Pembahasan Data Manual

Dalam perhitungan mengenai pengelompokan wilayah penghasil sampah tertinggi dan terendah maka dilakukan sebuah *clustering* guna mempertimbangkan solusi yang tepat pada wilayah tersebut supaya diberikan penanganan khusus oleh pemerintah. Untuk memberikan proses perhitungan maka terdapat data awal yang telah didapatkan untuk mendukung proses perhitungan. Terdapat beberapa proses setelah mendapatkan data mentah atau awal dengan menggunakan *software* rapidminer.

Design RapidMiner



Gambar 4 *Design* Data Manual di RapidMiner

Proses penggunaan *software* rapidminer tahapan pertama yang dilakukan setelah input data mentah yaitu dengan adanya *design* proses *clustering* dengan metode *k-means* dan *david bouldin index*. Pada proses *design* ini nantinya akan menunjukkan hasil atau implementasi dalam pengelompokan wilayah yang dilihat dari kecamatan Surabaya.

Cluster Model

Proses dalam menentukan *clustering* akan secara langsung terlihat jumlah yang masuk pada 3 cluster beserta dengan total dari data yang di *cluster* yaitu sebagai berikut:

Cluster Model

```
Cluster 0: 636 items
Cluster 1: 11 items
Total number of items: 647
```

Gambar 5 *Cluster* Model RapidMiner

Dapat dilihat dari gambar 4.2 mendapatkan total *number of items* 647 dengan pembagian hasil *cluster* model yaitu *cluster* 0 dengan implementasi 636 alamat yang terdeteksi menggunakan

wilayah kecamatan kota Surabaya merupakan hasil sampah terendah sedangkan *cluster* 1 dengan implementasi 11 alamat terdeteksi menggunakan wilayah kecamatan kota Surabaya mendapatkan nilai penghasil sampah tertinggi.

Centroid Table

Centroid *table* merupakan hasil dari rata – rata data yang masuk di 3 *attribute*, yakni *attribute* sampah masuk, sampah daur ulang dan sampah *up-cycle*.

Tabel 2 Centroid *Clustering*

No	Clustering	Jumlah	Kecamatan
1	0	636	Kenjeran, Asem Rowo, Tegalsari, Bubutan, Krembangan, Simokerto, Benowo, Tandes, Rungkut, Karangpilang dan lainnya.
2	1	11	Sukolilo, Sambikerep, Pakal, Jambangan, Mulyorejo, Sukomanunggal, Lakarsantri, Wiyung, Semampir dan lainnya.

Data Hasil Cluster

Tabel 3 Hasil Data *Cluster* Berdasarkan Kecamatan

Attribute	Cluster_0	Cluster_1
Sampah Masuk	6966.521	26052731.091
Sampah Daur Ulang	3396.458	26052731.091
Sampah Up-Cycle	141985.583	1637980.455

Dengan hasil tabel diatas dapat di jelaskan bahwa 11 kecamatan di kota Surabaya yang termasuk *cluster* 1 menjadi *cluster* dengan penghasil sampah yang tertinggi, sedangkan untuk hasil *cluster* 0 dengan jumlah anggota sebanyak 636 merupakan *cluster* dengan penghasil sampah terendah.

Hasil Davies Bouldin Index

Tabel 4 Uji Validitas *Davies Bouldin Index*

No	Jumlah Cluster	Nilai Uji Davies Bouldin Index
1	K = 2	0.268
2	K = 3	0.402

Dalam hal ini nilai uji yang dipakai adalah nilai $k \approx 0$ maka hasil yang tepat pada tabel diatas adalah berada di nilai $k = 2$ dengan hasil 0.268 dibanding dengan nilai $k = 3$ dengan hasil 0.402.

Pembahasan Data Pivot

Dalam data pivot ini data yang digunakan sama dengan data manual tetapi dilakukan peringkasan terhadap data guna membandingkan pola pada *clustering* jika menggunakan data mentah. Data pivot dilakukan karena terdapat beberapa data yang sama pada kecamatan

dikarenakan letak alamat yang ada di kota Surabaya terekam pada data manual, dengan hal ini maka diperlukan teknik pivot untuk menyederhanakan data.

Design RapidMiner



Gambar 6 *Design Data Pivot* di RapidMiner

Perbedaan dari *design* pengolahan data manual dengan data pivot memang tidak terjadi perbedaan yang signifikan pada bentuk fisiknya hanya berbeda pada penulisan data input awal tetapi nantinya hasil yang dikeluarkan akan berbeda dengan hasil pengolahan data manual.

Cluster Model

Cluster model dari data pivot menghasilkan data output yang berbeda dengan output data manual. Perbedaan itu terjadi dikarenakan jumlah data yang lebih sedikit dibanding dengan data manual

Cluster Model

```
Cluster 0: 30 items
Cluster 1: 3 items
Total number of items: 33
```

Gambar 7 *Cluster Model* RapidMiner

Centroid Table

Proses *running* yang salah satunya menghasilkan *centroid table* dari pengolahan data pivot maka membuat hasil *centroid table* di data manual dengan data pivot juga berbeda.

Tabel 5 Centroid *Clustering*

Attribut	Cluster_0	Cluster_1
Sampah Masuk	5275402	254741162
	0457358	899999680
	54.0	000000.0
Sampah Daur Ulang	7596490	255080973
	1375990.	333333.340
	560	
Sampah Up-Cycle	2019495	520254127
	6018161.	4673.333
	766	

Data Hasil Cluster

Tabel 6 Hasil Data *Cluster* Berdasarkan Kecamatan

No	Clustering	Jumlah	Kecamatan
1	0	30	Asem Rowo, Benowo, Bubutan, Dukuh Pakis, Gunung Anyar, Gubeng, Lakarsantri, Mulyorejo, Rungkut dan lainnya
2	1	3	Pakal, Semampir dan Wiyung.

Jumlah anggota *cluster* sebanyak 30 pada *clustering* 0 dengan pengelompokan berdasarkan kecamatan di kota Surabaya, sedangkan *cluster* yang kedua dengan jumlah anggota sebanyak 3 pada *clustering* 1 dengan pengelompokan juga berdasarkan kecamatan di wilayah kota Surabaya.

Hasil Davies Bouldin Index

Tabel 7 Uji Validitas *Davies Bouldin Index*

No	Jumlah Cluster	Nilai Uji <i>Davies Bouldin Index</i>
1	K = 2	0.326
2	K = 3	0.465

Dalam hal ini nilai uji yang dipakai adalah nilai $k \approx 0$ maka hasil yang tepat pada tabel diatas adalah berada di nilai $k = 2$ dengan hasil 0.326 dibanding dengan nilai $k = 3$ dengan hasil 0.465.

KESIMPULAN

Setelah dilakukan proses perhitungan dan pengolahan data maka hasil dari rumusan masalah dan tujuan dari penelitian mengenai pengelompokan sampah yang dihasilkan masyarakat kota Surabaya dengan menggunakan metode *K-Means* disimpulkan sebagai berikut:

1. Hasil *cluster model* dari pengolahan data manual dengan rapidminer menunjukkan bahwa dari total number of items sebesar 647 terdapat 11 wilayah dengan *cluster* tertinggi di kota Surabaya. Hasil uji validasi data manual menggunakan *Davies Bouldin Index* (DBI) menunjukkan 2 *cluster* dengan normalisasi data min-max paling kuat karena mendekati angka 0 dengan nilai *cluster* $k = 2$ hasil yang terjadi 0.268.
2. Hasil *cluster model* dari pengolahan data pivot dengan rapidminer menunjukkan bahwa dari total number of items sebesar 33 terdapat 3 wilayah dengan *cluster* tertinggi di kota Surabaya. Hasil uji validasi data pivot menggunakan *Davies Bouldin Index* (DBI) menunjukkan 2 *cluster* dengan normalisasi data min-max paling kuat karena mendekati angka 0 dengan nilai *cluster* $k = 2$ hasil yang terjadi 0.326

DAFTAR PUSTAKA

- [1] W. C, Solusi Pengelolaan Sampah Kota, 2021.
- [2] A. Kurniawan, I. D. Mumpuni dan M. As'ad, "Pengklasteran Bank Sampah Menggunakan Metode K-Means Pada Dinas Lingkungan Hidup Kabupaten Pasuruan," *SENASIF*, pp. 687-698, 2017.
- [3] W. Widiasih dan H. Nuha, "Usulan Strategi Sustainable Lifestyle Dalam Menunjang Eco Campus Di Universitas ABC Surabaya," *Symposium Nasional*, pp. 141-147, 2019.
- [4] W. Mega, "Clustering Menggunakan Metode K-Means Menentukan Status Gizi Balita, 2015.
- [5] K. Tutik, "Pengelompokan Surat Dalam Al-Quran Menggunakan Algoritma K-Means, 2014.