



BAB II

TEKNOLOGI INFORMASI

UNIVERSITAS DARMA PERSADA

2025

BAB II

LANDASAN TEORI

2.1 Reduksi Sampah

Reduksi sampah merupakan upaya untuk mengurangi jumlah sampah yang dihasilkan oleh masyarakat melalui berbagai pendekatan dan strategi yang berkelanjutan. Reduksi sampah menjadi bagian penting dalam pengelolaan sampah modern untuk mengurangi dampak lingkungan, memperpanjang umur tempat pembuangan akhir (TPA), dan meningkatkan efisiensi dalam penggunaan sumber daya. Strategi reduksi sampah berfokus pada pendekatan yang dikenal sebagai prinsip 3R (*Reduce, Reuse, Recycle*), yang bertujuan untuk mengurangi produksi sampah di sumbernya, menggunakan kembali barang yang masih dapat dimanfaatkan, dan mendaur ulang material untuk mencegah pembuangan langsung ke TPA.

Menurut penelitian (Kamal, 2024) Tujuan dari penelitian ini adalah mengevaluasi system operasional pengelolaan sampah di TPS Kecamatan Kembangan serta menghitung capaian dan potensi reduksi yang bisa dilakukan di TPS di Kecamatan Kembangan. Terdapat dua metode yang digunakan dalam penelitian ini yaitu analisis Kinerja Operasional Pengelolaan Sampah dan Analisis Mass Balance. Hasil dari penelitian yang dilakukan menunjukkan capaian reduksi sampah yang bisa dilakukan di TPS Kecamatan Kembangan sebesar 7%. Serta menunjukkan bahwa hasil dari evaluasi sistem operasional menghasilkan 2 kategori TPS yaitu Kategori sesuai dan Kurang sesuai, kategori sesuai hanya terdapat di 1 TPS dan 7 TPS lainnya termasuk ke kategori Kurang Sesuai.

Menurut Peraturan Pemerintah Indonesia No. 81 Tahun 2012 tentang Pengelolaan Sampah Rumah Tangga dan Sampah Sejenis Sampah Rumah Tangga, reduksi sampah harus dilakukan dengan mendorong perubahan perilaku masyarakat dalam pengelolaan sampah dari hulu hingga hilir. Hal ini mencakup pemilahan sampah di sumber, pengurangan penggunaan plastik sekali pakai, dan optimalisasi penggunaan kembali bahan organik maupun anorganik.

2.2 Data Mining

Data mining merupakan teknik analisis data yang dapat digunakan untuk mengidentifikasi pola dan tren dalam pengelolaan sampah. Menurut penelitian (Suryanti, 2024), penelitiannya bertujuan untuk menghasilkan alat bantu untuk visualisasi data volume sampah, dengan melakukan visualisasi data dapat diketahui pola atau kecenderungan perilaku masyarakat pada suatu wilayah dalam pengelolaan sampah. Untuk memudahkan dalam pengambilan keputusan tingkat eksekutif berdasarkan hasil pengolahan segmentasi data mining. Manfaat mengetahui visualisasi data tersebut memudahkan dalam identifikasi wilayah yang memiliki volume sampah paling banyak untuk menentukan langkah strategis yang dipilih dalam mengurangi volume sampah di suatu wilayah. Data volume sampah dalam penelitian ini menggunakan data sampah kota surakarta tahun 2017 sampai dengan tahun 2023. Segmentasi pada data volume sampah tersebut bertujuan untuk mengenali pola sumber penghasil sampah volume (ton) terbesar, volume sedang dan volume kecil. Pengetahuan terhadap segmen tersebut dapat digunakan sebagai pendukung dalam pengambilan keputusan strategis pengelolaan sampah Kota Surakarta. Adanya alat bantu ini berguna untuk mengetahui pola atau kecenderungan masyarakat di suatu wilayah akan perilaku kesadaran lingkungan.

Alat bantu ini tidak secara langsung berdampak di lapangan. Namun lebih memudahkan eksekutif dalam mengetahui kondisi sampah disuatu wilayah dan memutuskan tindakan apa yang tepat untuk dilakukan berdasarkan visualisasi data atau hasil pengolahan data dengan data mining.

Data mining terdiri dari beberapa tahapan, antara lain:

1. Pembersihan Data (*Data Cleaning*): Menghilangkan data yang tidak lengkap, berisik, atau tidak relevan.
2. Integrasi Data (*Data Integration*): Menggabungkan data dari berbagai sumber untuk mendapatkan dataset yang lengkap.
3. Seleksi Data (*Data Selection*): Memilih data yang relevan untuk dianalisis lebih lanjut.
4. Transformasi Data (*Data Transformation*): Mengubah data ke dalam format yang lebih sesuai untuk proses *mining*.
5. Proses *Data Mining*: Melakukan analisis dengan menggunakan algoritma tertentu untuk menemukan pola atau informasi.
6. Evaluasi Pola (*Pattern Evaluation*): Menginterpretasikan pola-pola yang ditemukan untuk memastikan relevansinya.
7. Presentasi Pengetahuan (*Knowledge Presentation*): Menyajikan hasil data mining dalam bentuk yang mudah dimengerti oleh pengguna.

Data mining banyak diaplikasikan dalam berbagai bidang, seperti bisnis untuk analisis perilaku konsumen, kesehatan untuk prediksi penyakit, dan keamanan siber untuk mendeteksi anomali. *Data mining* digunakan dalam penelitian pengelolaan sampah untuk mengidentifikasi pola dan tren dari data historis, sehingga memungkinkan pengambilan keputusan yang lebih efektif dalam upaya

pengurangan sampah. *Data mining* digunakan dalam penelitian pengelolaan sampah untuk mengidentifikasi pola dan tren dari data historis, sehingga memungkinkan pengambilan keputusan yang lebih efektif dalam upaya pengurangan sampah. Metode ini memanfaatkan algoritma analitik untuk menemukan informasi tersembunyi dan pola yang relevan dalam dataset pengelolaan sampah (Suryanti *et al.*, 2024). Secara umum, teknik data mining dapat dikelompokkan menjadi dua kategori utama:

1. *Prediksi (Prediction)*: Teknik yang digunakan untuk memperkirakan hasil di masa depan berdasarkan data historis.
2. *Deskriptif (Descriptive)*: Teknik yang digunakan untuk menggambarkan karakteristik atau hubungan data tanpa melakukan prediksi.

Beberapa teknik data mining yang populer antara lain:

1. *K-Means Clustering*: Metode yang membagi data ke dalam kelompok (*cluster*) berdasarkan kemiripan data.
2. *Decision Tree*: Teknik untuk membuat model prediktif dengan struktur pohon yang bercabang.
3. *Association Rule Learning*: Metode yang menemukan aturan hubungan antar item dalam dataset, seperti pada analisis keranjang belanja (*market basket analysis*).

Data mining memegang peranan penting dalam membantu organisasi untuk mengubah data mentah menjadi wawasan yang bermanfaat, sehingga dapat meningkatkan efisiensi, mengoptimalkan strategi bisnis, dan memberikan keunggulan kompetitif.

2.3 Time Series

Time series analysis adalah metode statistik yang digunakan untuk menganalisis data yang dikumpulkan dari waktu ke waktu. Dalam konteks pengelolaan sampah, analisis time series dapat digunakan untuk memprediksi pertumbuhan volume sampah di masa depan. Menurut (Bagwan, 2024) Penelitian ini berfokus pada negara bagian Maharashtra, menyelidiki tren masa lalu dan masa depan menggunakan model *Auto Regressive Integrated Moving Average* (ARIMA), yang biasa digunakan dianalisis dan peramalan deret waktu. Studi ini memperkirakan bahwa rata-rata kapasitas pengolahan limbah elektronik (daur ulang/pembongkaran) dari tahun 2023–2030 akan menjadi 163563,15 MT (metrik ton), dengan nilai perkiraan terus meningkat selama bertahun-tahun, mencapai 248 pendaur ulang pada tahun 2030. Dengan menganalisis laju perubahan dalam pengolahan limbah elektronik kapasitas pada tahun 2023–2030, terdapat peluang sebesar 6,86 % setiap tahunnya. Ini juga menyoroti ruang lingkup kewirausahaan di industri daur ulang limbah elektronik dengan mendeteksi rata-rata peningkatan pendaur ulang sebesar 7,23% per tahun. Studi ini menekankan pentingnya peran kebijakan dan pengambilan keputusan dalam mengelola sampah yang berkembang pesat ini mengalir dari perspektif pengelolaan lingkungan dan ekonomi sirkular.

Secara matematis, model time series dapat direpresentasikan sebagai berikut:

$$Y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-n}) + \varepsilon_t$$

Di mana:

- a. Y_t adalah nilai deret waktu pada waktu t .
- b. $y_{t-1}, y_{t-2}, \dots, y_{t-n}$ adalah nilai deret waktu sebelumnya

- c. F adalah fungsi yang menggambarkan hubungan antara nilai sebelumnya dan nilai saat ini.
- d. ε_t adalah kesalahan atau noise yang tidak dapat dijelaskan oleh model.

Model *ARIMA (AutoRegressive Integrated Moving Average)* adalah salah satu metode yang umum digunakan dalam analisis time series untuk memprediksi nilai masa depan berdasarkan data historis. Menurut Santoso et al. (2023), penerapan model ARIMA dalam *forecasting* pertumbuhan sampah di Indonesia menunjukkan hasil yang akurat dan dapat diandalkan. Penelitian ini menekankan bahwa model ARIMA dapat membantu pengelola sampah dalam merencanakan kapasitas pengelolaan yang diperlukan untuk mengantisipasi peningkatan volume sampah. Buku "*Forecasting Waste Generation with ARIMA*" oleh Patel (2023) juga membahas penerapan model ARIMA dalam konteks pengelolaan sampah, serta memberikan panduan praktis untuk implementasinya.

Di tingkat internasional, penelitian oleh Alavi et al. (2022) menunjukkan bahwa model ARIMA dapat digunakan untuk memprediksi pertumbuhan sampah di berbagai negara dengan hasil yang konsisten. Penelitian ini menyoroti pentingnya pemilihan parameter yang tepat dalam model ARIMA untuk meningkatkan akurasi prediksi. Dengan demikian, penerapan model ARIMA dalam *forecasting* pertumbuhan sampah dapat memberikan wawasan yang berharga bagi pengelola limbah dalam merencanakan strategi pengelolaan yang lebih efektif.

Model ARIMA terdiri dari tiga komponen utama yang dinyatakan dalam bentuk parameter (p, d, q) , yaitu:

1. *Autoregressive (AR) – p*

Komponen ini merepresentasikan hubungan antara nilai saat ini dengan nilai sebelumnya dalam seri waktu berdasarkan regresi linear terhadap nilai-nilai sebelumnya. Model AR(p) menggunakan parameter p , yang menunjukkan jumlah lag atau keterlambatan yang digunakan dalam model. Persamaan matematis untuk model AR adalah:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t$$

di mana:

- Y_t adalah nilai aktual saat ini,
- ϕ adalah koefisien regresi,
- ϵ_t adalah kesalahan acak (*white noise*).

3. *Integrated (I) – d*

Komponen ini digunakan untuk menghilangkan tren atau membuat data menjadi stasioner melalui proses *differencing* (pembedaan) berulang hingga data memiliki nilai rata-rata dan varians yang konstan. Parameter d menunjukkan berapa kali differencing perlu dilakukan agar data menjadi stasioner. Differencing pertama didefinisikan sebagai:

$$Y_t' = Y_t - Y_{t-1}$$

4. *Moving Average (MA) – q*

Komponen MA mengukur hubungan antara nilai saat ini dengan kesalahan residual dari periode sebelumnya. Model MA(q) menggunakan

parameter q , yang menunjukkan jumlah keterlambatan dari kesalahan sebelumnya yang digunakan dalam model. Persamaan matematis untuk model MA adalah:

$$Y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q}$$

di mana:

- μ adalah rata-rata,
- θ adalah koefisien MA,
- ϵ_t adalah kesalahan acak pada waktu t .

2.4 K – Means Clustering

K-Means Clustering adalah teknik analisis data yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok berdasarkan kesamaan. Dalam konteks pengelolaan sampah, metode ini dapat digunakan untuk mengidentifikasi wilayah dengan penyebaran sampah terbanyak, sedang, dan sedikit. Menurut (Magriaty, 2022) Metode penelitian yang digunakan adalah PCA (*principal component analysis*) dan *K-Means cluster* untuk mendapatkan zona pengelolaan sampah. Hasil penelitian menunjukkan bahwa zona pengelolaan sampah dapat diklasifikasikan ke dalam 3 jenis zona. Zona tipe 1 terdiri dari 25 desa yang terletak di daerah perkotaan yang memiliki timbulan sampah dengan kerentanan tertinggi. Zona tipe 2 terdiri dari 36 desa dan terletak relatif dekat dengan daerah pedesaan-perkotaan yang memiliki tingkat pertumbuhan penduduk tertinggi. Zona tipe 3

terdiri dari 36 desa yang terletak sangat jauh dari daerah perkotaan yang memiliki timbunan sampah Secara matematis.

algoritma *K-Means* meminimalkan nilai *Within-Cluster Sum of Squares (WCSS)* yang dinyatakan sebagai:

$$WCSS = \sum_{j=1}^k \sum_{i=1}^{n_j} (x_i - \bar{x}_j)^2$$

Di mana:

- a. k adalah jumlah kluster.
- b. n_j adalah jumlah titik data dalam kluster j .
- c. x_i adalah titik data i dalam kluster j .
- d. \bar{x}_j adalah rata-rata centroid dari kluster j .

Tujuan dari *K-Means* adalah meminimalkan nilai WCSS, sehingga data dalam satu kluster memiliki karakteristik yang lebih mirip dibandingkan dengan data di kluster lain. Algoritma ini sering digunakan untuk segmentasi pelanggan, pengelompokan data gambar, dan analisis pola data lainnya.

WCSS adalah metrik yang digunakan untuk mengukur seberapa baik data dikelompokkan dalam kluster yang diberikan. WCSS menggambarkan total jarak kuadrat antara setiap titik data dan centroid klusternya, yang berarti semakin rendah nilai WCSS, semakin baik kluster yang terbentuk. Penurunan nilai WCSS pada grafik elbow digunakan untuk menentukan jumlah kluster yang optimal.

Rumus perhitungan WCSS dapat dinyatakan sebagai:

$$WCSS = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2$$

Di mana:

- a. k adalah jumlah kluster.
- b. x adalah data dalam kluster C_i .
- c. μ_i adalah centroid kluster C_i .
- d. $\|x - \mu_i\|^2$ adalah jarak Euclidean antara data x dan centroid μ_i .

Dengan meminimalkan nilai WCSS, algoritma *K-Means* berusaha menemukan kelompok yang optimal dan meningkatkan efektivitas dalam segmentasi data.

Adapun untuk melakukan evaluasi terhadap hasil *K-Means Clustering* dilakukan dengan *Sillhoute Score*. *Silhouette Score* mengukur seberapa baik data terkelompok dengan memperhitungkan jarak antar kluster dan dalam kluster. Nilainya berkisar antara -1 hingga 1