**Literature Review**

**Introduction**

The growing populations and consumption patterns in cities worldwide are increasing urban waste generation, posing a significant challenge for sustainable city planning and environmental management. Accurate prediction of urban waste generation is crucial for developing sustainable waste management strategies. In this section, the existing literature will be reviewed to find a limitation in the current body of knowledge. This project tries to build a bridge the gap between urban waste generation and sustainable waste management by employing advanced machine learning algorithms.

**Predicting urban waste generation**

Coskuner et al. (2020) employed artificial neural networks (ANNs) to predict domestic, commercial and construction wastes. The ANN model was trained on a dataset of historical waste generation data and various socioeconomic factors. The results indicated that the ANN model was able to accurately predict waste generation, demonstrating the potential of machine learning techniques in waste management.

Verma et al. (2019) utilized multiple linear regression (MLR) to forecast municipal waste generation. The MLR model included various independent variables, such as population, income, and household size. The results demonstrated that MLR could effectively predict waste generation, providing a simple and interpretable model for waste management planning.

Adeleke et al. (2021) investigated the performance of ANNs in predicting municipal solid waste generation. The result of the study verifies that waste composition prediction can be done in a single hidden-layer satisfactorily.

**Clustering urban waste generation**

Du et al. (2022) employed K-means clustering to classify municipal solid waste (MSW) prediction in mainland china. they used six economic, social, and climatic indicators: population, per capita GDP (PCGDP), environmental sanitation investment (ESI), average temperature, average precipitation, and average humidity.The results show that the new city classification system can be used to predict MSW generation and formulate appropriate regional policies.

Guleryuz (2020) evaluated the waste management performance of 39 districts in Istanbul was assessed using K-means clustering. The comparison took into account domestic waste, medical waste, population, municipal budget, and mechanical sweeping area. The analysis revealed that there are significant variations among the districts in the clusters formed by these variables.

Various machine learning techniques, including artificial neural networks, support vector regression, and multiple linear regression, have been employed for forecasting municipal solid waste generation. Clustering urban waste generation patterns helps identify areas with similar waste characteristics, enabling targeted waste management interventions. Studies have explored various clustering algorithms, including k-means, hierarchical clustering, and self-organizing maps, to group urban areas based on waste generation patterns.

**Data Description**

**Introduction**

Waste management is a critical challenge facing urban centers worldwide, with the ever-increasing volume of waste posing a significant threat to the environment and public health. Accurately predicting and clustering urban waste generation patterns are essential steps towards developing effective and sustainable waste management strategies. The research project tries to address the following question:

1. What are the key factors that influence urban waste generation patterns?
2. Can we accurately predict urban waste generation using data-driven approaches?
3. Can we identify distinct clusters of urban areas based on waste generation characteristics?

**Dataset**

Data on waste generation for 39 districts of Istanbul in 2019 was obtained from an academic article published by Guleryuz in 2019. The dataset includes five variables: domestic waste amount, population, municipal budget, medical waste amount, and mechanical sweeping area. Table 1 presents the descriptive statistics of the normalized dataset.

Table 1: Descriptive Statistics of the normalized dataset

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variables | DW | PO | MB | MW | MS |
| Mean | -1.00E-16 | 6.83E-17 | 3.20E-16 | 2.03E-16 | -7.70E-17 |
| Std | 1 | 1 | 1 | 1 | 1 |
| Min | -1.97 | -1.89 | -2.28 | -1.08 | -1.45 |
| Max | 2.8 | 2.76 | 2.77 | 2.71 | 2.8 |
| Skewness | 0.62 | 0.56 | 0.06 | 1.17 | 0.95 |
| Kurtosis | 1.03 | 0.51 | 0.85 | 0.76 | 1.08 |

**Technology Review**

**Introduction**

The increasing volume of urban waste poses a significant challenge to municipalities worldwide, demanding the development of effective and sustainable waste management strategies. Technological advancements have emerged as a promising avenue to address this challenge, offering tools that can predict waste generation patterns, optimize waste collection routes, and streamline waste management processes. This technology review examines the capabilities of machine learning and data analytics tools in predicting and clustering urban waste generation.

**Technology Review**

Machine learning and data analytics tools have revolutionized the way we analyze and interpret data, offering powerful algorithms that can uncover hidden patterns and extract meaningful insights. In the context of urban waste management, these tools can be employed to analyze waste generation data, socioeconomic factors, and other relevant variables to predict future waste generation trends. This predictive capability enables municipalities to proactively allocate resources, optimize collection schedules, and plan waste disposal strategies. Moreover, data analytics tools can be utilized to cluster urban areas based on their waste generation patterns, identifying areas with similar waste characteristics.

The integration of machine learning and data analytics tools into waste management systems offers several benefits that align perfectly with the goals of this project. By leveraging these technologies, we can:

1. Enhance the accuracy of waste generation predictions, enabling more efficient resource allocation and cost savings.
2. Identify distinct waste generation patterns across urban areas, facilitating targeted waste management interventions.

Various machine learning and data analytics tools are available for predicting and clustering urban waste generation. In this project the following machine learning algorithms will be employed:

1. **Regression Analysis**: Regression analysis is a statistical technique that establishes a relationship between a dependent variable. It provides a simple and interpretable model for predicting waste generation, but it may not capture complex nonlinear relationships between dependent variable and independent variables.
2. **Artificial neural networks (ANNs):** ANNs are powerful machine learning algorithms that can learn from complex data patterns and make accurate predictions. It can capture complex nonlinear relationships between dependent variable and independent variables, but they also require more computational power.
3. **K-means clustering:** K-means clustering is a simple yet effective algorithm for grouping data points into a predefined number of clusters, enabling the identification of distinct waste generation patterns. K-means clustering, while simple and effective in identifying distinct waste generation patterns, is prone to limitations. Its greedy nature can lead to entrapment in local minima, and its sensitivity to initial centroid placement.

Table 2: summary of the strengths and weaknesses of each method

|  |  |  |
| --- | --- | --- |
| **Method** | **Strengths** | **Weaknesses** |
| **Regression Analysis** | Simple, interpretable, can handle missing data | May not capture complex nonlinear relationships |
| **ANNs** | High accuracy, can capture complex nonlinear relationships | Requires large amounts of data, computationally expensive |
| **K-means clustering** | Simple, efficient, easy to interpret | Sensitive to initial centroids, may stuck in local minima |
|  |  |  |

Despite the significant progress made in applying machine learning algorithms to urban waste management applications, further research is needed in combining predictive and clustering algorithms to develop more sophisticated machine learning techniques to capture the waste generation patterns in complex urban environment.

**References**

Adeleke, O., Akinlabi, S. A., Jen, T.-C., & Dunmade, I. (2021). Application of artificial neural networks for predicting the physical composition of municipal solid waste: An assessment of the impact of seasonal variation. *Waste Management & Research*, *39*(8), 1058–1068. https://doi.org/10.1177/0734242X21991642

Coskuner, G., Jassim, M. S., Zontul, M., & Karateke, S. (2020). Application of artificial intelligence neural network modeling to predict the generation of domestic, commercial and construction wastes. *Waste Management & Research*, *39*(3), 499–507. https://doi.org/10.1177/0734242X20935181

Du, X., Niu, D., Chen, Y., Wang, X., & Bi, Z. (2022). City classification for municipal solid waste prediction in mainland China based on K-means clustering. *Waste Management*, *144*, 445–453. https://doi.org/https://doi.org/10.1016/j.wasman.2022.04.024

Guleryuz, D. (2020). Evaluation of waste management using clustering algorithm in megacity Istanbul. *Environmental Research and Technology*, *3*(3), 102–112. https://doi.org/10.35208/ert.764363

Verma, A., Singh, N. B., & Kumar, A. (2019). Application of Multi Linear Model for Forecasting Municipal Solid Waste Generation in Lucknow City: A Case Study. *Current World Environment*, *14*. https://doi.org/10.12944/CWE.14.3.10