Food Waste Analysis with Machine Learning

Nur Hidayah binti Ahmad Shafii

1. INTRODUCTION

In 2016, the United Nations Sustainable Development Group introduced 17 global goals as a universal call to action to end poverty, protect the planet, and ensure prosperity for all by 2030. Each goal has specific targets that address various global challenges, with SDG 12 focusing on responsible consumption and production (Nik Mahdi et al., 2023). Understanding the trends and factors contributing to food waste is crucial to achieving SDG 12, which aims to ensure sustainable consumption and production patterns.

Food waste is irrational economic management where food is wasted regardless of expiration date, left uneaten, or spoiled (Radzymińska et al., 2016). The Food and Agriculture Organization estimates that one-third of global food production is for human consumption, while 1.3 billion tons are lost or wasted. Food waste in Kuala Lumpur averaged 2.1 tonnes per day in March 2020, despite Malaysia's strong solid waste management awareness (Wafi & Tumiran, 2022) Inefficient food management and overproduction lead to significant waste, contributing to greenhouse gas emissions, landfills, and water pollution(Masuk Shopnil et al., 2023) In Malaysia, several types of waste management practices are employed, including recycling, composting, incineration, inert landfill, sanitary landfill, and other disposal sites (Wahidah & Ghafar, 2017). However, landfill disposal remains the most common method due to its cost-effectiveness and simplicity, even though many landfills have reached their capacity (Lim et al., 2016).

Machine learning offers a powerful toolset for analyzing and understanding the complex dynamics of food waste. By leveraging data-driven insights, researchers can identify patterns and predictive factors, thereby contributing to more sustainable food systems and improved resource efficiency (Panda & Dwivedi, 2020). This study focuses on using machine learning to analyze the relationship between food waste and population growth in a Southeast Asian country, providing valuable insights to support SDG 12. Institutions can achieve a more sustainable and efficient future by recognizing and addressing the variables that contribute to food waste. This proactive strategy not only supports global sustainability objectives but also encourages responsible management of the environment and economic strength.

2. PROBLEM STATEMENTS

Food waste is a pressing issue throughout the globe caused by various socio-economic, cultural, and infrastructural factors. Most of the food waste is disposed of in landfills or incinerated, leading to numerous issues for the environment, economy and society (Amiruddin et al., 2023). Recent studies have underscored the need for detailed investigations into food waste patterns in Southeast Asia to inform policy and practice (Jereme et al., 2016; Tamasiga et al., 2022; Wafi & Tumiran, 2022). The research conducted by Radzymińska et al. (2016) revealed that consumers did not practice food management although fully aware of the social and ecological effects of food waste. They suggested a mass educational campaign against food waste to combat negative attitudes and stimulate model behavior toward food waste. Therefore, this study aims to fill this critical gap by providing insights into the current situation and underlying causes of food waste in Southeast Asia, thereby contributing to more informed and effective waste management strategies.

The advancements in machine learning offer new opportunities to enhance predictive accuracy by analyzing a wide range of features (Masuk Shopnil et al., 2023). Previous research has demonstrated the potential of machine learning models to forecast food waste and identify key influencing factors (Masuk Shopnil et al., 2023; Panda & Dwivedi, 2020) However, lack of food waste prediction using machine learning algorithms in Malaysia (Ramlan et al., 2023). This study seeks to apply these advanced techniques to investigate the correlation between food waste and population to provide valuable insights and effective waste reduction strategies.

3. OBJECTIVE

The objectives of this study are as follows:

i. To provide insight into the food waste trends in Southeast Asia.
ii. To investigate the relationship between food waste against population for a Southeast Asia country using machine learning.

4. METHODOLOGY

In this research, Cross-Industry Standard Processes for Data Mining (CRISP-DM) as shown in Fig xx are used to solve business issues using analytics.

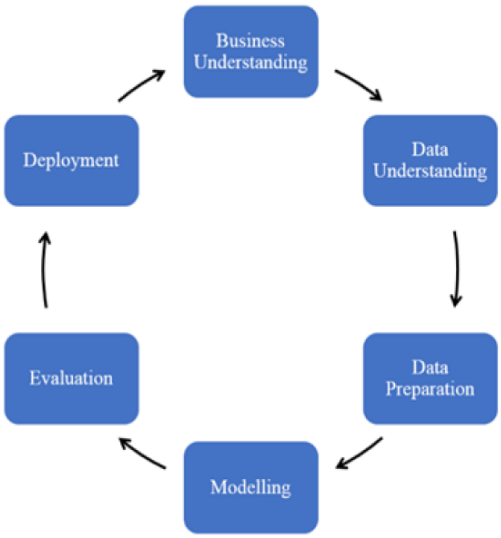

Figure 4.1 CRISP-DM Methodology

Business understanding is the initial phase where the researcher converts the research objectives and requirements into a data mining problem definition. This step ensures that

the goals of the research are well understood and aligned with the data analysis tasks. In the data understanding process, it focuses on data collection and identifying data quality issues. Researchers gather relevant data and conduct an initial analysis to assess its quality, completeness, and relevance to the problem at hand. Next, data preprocessing and cleaning are involved in the data preparation phase. handling missing values, ensuring data consistency, and preparing the dataset for analysis. The data is transformed into a suitable format for modeling. The preprocessed dataset undergoes EDA to uncover patterns and insights. Visualizations such as charts showing the highest combined per-capita food waste, food waste by source, and choropleth maps are created to illustrate key findings and understand data distribution. In the modeling phase, various machine learning models are applied to the dataset to explore the relationship between the dependent variable (e.g., total food waste) and independent variables (e.g., population size). The models used include linear regression, random forest and support vector regression. The models are evaluated and validated in the evaluation phase to ensure their quality and performance. Performance metrics such as Mean Squared Error (MSE) and R-squared are used to compare the models. The model that achieves the highest performance and meets the research criteria is selected as the final model. Lastly, The selected model is then deployed in real-world situations. This involves integrating the model into practical applications where it can be used to predict food waste and assist in decision-making processes to achieve sustainable consumption and production patterns.

5. DATASETS

The data for this study was obtained from a reliable source which was the United Nations Environment Programme webpage. There were 46 variables in the dataset such as country, region, category, and so on. Some of the attributes were not included in the experimental process since the attributes did not have any significant impact. In this study, the data was limited to the year 2021. Hence, it limits the ability to analyze trends over time. Future studies incorporating data from multiple years would be necessary to provide a more comprehensive understanding of the subject matter and to track changes and trends over time.

6. PROGRAMS/TOOLS TO BE USED

In this study, Python 3 was chosen as the programming language due to its extensive collection of libraries. For instance, the Pandas library provides tools for analysing and manipulating data, the NumPy library gives a wide range of mathematical operations, and the Matplotlib package provides tools for visualising data. Furthermore, Google Colab was selected as the primary platform to support the Python code.

7. REFERENCES

Amiruddin, N. N. H. N., Rozamri, N. A., Baharudin, F., & Mohamad, I. N. (2023). A Study on Household Food Waste Management and Composting Practice. *IOP Conference Series: Earth and Environmental Science*, *1205*(1). https://doi.org/10.1088/1755-1315/1205/1/012018

Jereme, I. A., Siwar, C., Begum, R. A., & Abdul Talib, B. (2016). Addressing the problems of food waste generation in Malaysia. *International Journal of ADVANCED AND APPLIED SCIENCES*, *3*(8), 68–77. https://doi.org/10.21833/ijaas.2016.08.012

Lim, W. J., Chin, *, Yusof, Yahya, & Tee, T. P. (2016). Food waste handling in Malaysia and comparison with other Asian countries Abstract. In *International Food Research Journal* (Vol. 23).

Masuk Shopnil, M. M., Husna, A., Sultana, S., & Islam, M. N. (2023). An Ensemble ML Model to Predict the Wastage of Food: Towards Achieving the Food Sustainability. *2023 International Conference on Next-Generation Computing, IoT and Machine Learning, NCIM 2023*. https://doi.org/10.1109/NCIM59001.2023.10212669

Nik Mahdi, N. A., Fernando, Y., & Abdalla, Y. A. (2023). Understanding The Sustainable Development Goals Concept: Malaysia Report and Trend. *Journal of Governance and Integrity*, *5*(3), 317–327. https://doi.org/10.15282/jgi.5.3.2022.8938

Panda, S. K., & Dwivedi, M. (2020). Minimizing Food Wastage Using Machine Learning: A Novel Approach. *Smart Innovation, Systems and Technologies*, *159*, 465–473. https://doi.org/10.1007/978-981-13-9282-5_44

Radzymińska, M., Jakubowska, D., & Staniewska, K. (2016). Consumer Attitude and Behaviour towards Food Waste. *Journal of Agribusiness and Rural Development*, *10*(1). https://doi.org/10.17306/JARD.2016.20

Tamasiga, P., Miri, T., Onyeaka, H., & Hart, A. (2022). Food Waste and Circular Economy: Challenges and Opportunities. In *Sustainability (Switzerland)* (Vol. 14, Issue 16). MDPI. https://doi.org/10.3390/su14169896

Wafi, M. A. A., & Tumiran, M. A. (2022). Food Waste Management in Malaysia: A Brief Introduction. In *Selected Proceedings from the 1st International Conference on Contemporary Islamic Studies (ICIS 2021)* (pp. 275–282). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-2390-6_26

Wahidah, S., & Ghafar, A. (2017). *Food Waste in Malaysia: Trends, Current Practices and Key Challenges*. http://ap.fftc.agnet.org/ap_db.php?id=774&print=1