PREDICTING MEDICAL WASTE GENERATION BASED ON DEMOGRAPHIC AND SOCIO-ECONOMIC PARAMETERS THROUGH MACHINE LEARNING APPROACH IN KHULNA CITY

Tanjina Mannan Risha*[1] and Islam Md Rafizul[2]

[1]*Undergraduate student, Department of Civil Engineering, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh, e-mail: tanjina.risha@gmail.com,*
[2]*Professor, Department of Civil Engineering, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh, e-mail: imrafizul@ce.kuet.ac.bd*

***Corresponding Author**

ABSTRACT

Proper medical waste management is essential for environmental sustainability and public health preservation. The goal of this study is to predict medical waste generation in Khulna city using machine learning techniques taking into account socioeconomic and demographic factors. Healthcare centers are known to produce significant volumes of several types of waste such as hazardous waste, infectious waste, toxic waste, sharp waste, etc. This study aimed to create a predictive algorithm that can predict medical waste generation based on selective parameters. Machine Learning algorithms are trained over examples through which they learn from past experiences and also analyze historical data. Predicting waste generation helped with the implementation, improvement, and optimization of waste management operations. In addition, a Python script is developed utilizing sci-kit-learn to train a machine-learning model. Demographic and socio-economic data related to medical waste (number of beds, average number of patients daily, bed occupancy, service charge, etc.) across various healthcare centers in Khulna City were collected. This dataset was then utilized to train and validate the machine-learning model. Results indicated that among the variables considered, the number of patients had the strongest relationship with the amount of medical waste generated. This analysis also revealed that bed occupancy has the weakest relationship with medical waste generation. This is because the total number of patients is considered, which includes both outdoor patients and indoor patients. Some hospital has fewer beds. but they get more outdoor patient because of their good lab quality, and this results in the weakest relationship between generated waste and bed occupancy. After considering all factors, the prediction has a Confidence score of 0.95 and a mean absolute error of 18.68. It indicates that as the number of parameters increases, the accuracy of the model increases.

Keywords: *Machine Learning, Python, Scikit learn, Mean absolute error, Confidence score.*

1. INTRODUCTION

Due to increasing urbanization and population expansion, the issue of disposing of toxic hazardous wastes, including hospital waste, is becoming more and more of a global concern. Hospital waste or clinical waste, which poses serious threats to environmental health, requires specialized treatment and management before its final disposal (PRISM Bangladesh 2005).
All wastes resulting from both human and animal activity are included in the term "waste." Medical waste is one of the riskiest types of waste. Because it is contagious. Medical waste is a unique kind of trash that is extremely harmful shown in Figure 1. After nuclear waste, this form of waste is the second most dangerous. The quantitative and qualitative characteristics of healthcare service wastes have evolved and become more harmful due to population growth and technological advancements.. Medical wastes include extremely toxic compounds and can spread disease. The use of disposable medical items has increased along with the medical sector's global expansion, which has significantly increased the creation of medical waste. Medical wastes include highly toxic chemicals, toxic metals, pathogenic viruses, and bacteria that can lead to pathological dysfunction in the human body (Mazloomi et al. 2019, Zamparas et al. 2019). For arbitrary usage and management, medical waste presents a significant risk to patients, hospital visitors, doctors, nurses, technicians, sweepers, and others. The Medical Waste Tracking Act of 1988 in the United States defines medical waste as "any solid waste that is created in the diagnosis, treatment, or immunization of human beings or animals, in study pertaining thereto, or in the production or testing of biologicals." In contrast, the European Union defines medical waste items as those related to chapter 18 of the European Waste Catalogue, which is defined as the wastes from human or animal health care and/or related investigation. Medical trash in Turkey is described as pathogenic, contagious, and sharp items that come from the medical field. The way that different nations define medical waste can have a significant impact.

Fig-1: Medical waste

Disposing of medical waste simply into trash cans, drains, and canals, or its ultimate disposal outside of cities, presents a major risk to public health. This disrespect for maintaining public health results from a lack of knowledge, expertise, and infrastructure among those involved in the creation and removal of hospital waste, as well as from a dearth of treatment facilities and systems in the community. The growing number of clinics, hospitals, and diagnostic labs in the city is making the issue worse. Because of this, researching clinical waste management is essential to finding a workable solution to this issue.

Current disposal technique: Transport and off-site management is done by the NGO PRODIPAN, a leading NGO that deals with the medical waste of Khulna. They collect medical waste from 1 public hospital, 8 private hospitals, 60 clinics, and 38 diagnostic centers. They divide the hospital waste into 4 categories to manage it easily. They have a proposal for 4 different color drums. Black drums are for compostable food or kitchen wastes, and green drums are for recyclable wastes like plastics. Red drums for sharp wastes, yellow drums for infectious medical wastes. They have two covered vans for waste transport. The van has a capacity of 1.5 tons of waste carrying per trip. The mode of transport is primary disposal. They spend about 1 lakh 65 thousand tk per month on medical waste management. The vehicle transports the waste to the Rajbadh, Khulna. Where they have a place for segregation, treatment, and disposal.

Not all hospitals separate waste according to the rules during generation. That's why after the collection of waste from the hospital, waste is separated according to type such as hazardous, sharp, infectious, etc. At first, they segregate the mixed wastes into four categories. They compost the kitchen or food waste from black drums and deliver it to the land. The sharp wastes from red drums are stored underground in the septic tank. They have shredding arrangements for recyclable wastes like plastics and paper from green drums. These are sold after shredding. The medical wastes like gauges and bandages from yellow drums are wetted in water then chlorine is added and dried. After that, the dry wastes are burned in the incinerator which has a high chimney. Here fig-3,fig-3, and fig-5 show waste collection, waste separation, and incineration respectively.

Fig-3: Waste collection van

Fig-4: Waste separation according to type

Fig-5: Incinerator

2. METHODOLOGY

2.1 Study Area

Khulna is one of the largest cities in Bangladesh faced with growing waste management problems. Global concerns for urban environmental pollution are increasing day by day. The effects of environmental pollution are already visible in Bangladesh. However, the concern about pollution has created awareness of the need to control the environment. Community organization is an essential element in ensuring the effectiveness of solid waste management in Khulna city and increasing the likelihood of its sustainability. The major activities of the community organization process will be participatory appraisal, awareness and motivational campaigns, and the establishment of waste management committees. Clinical waste management (CWM) today is considered to be one of the most immediate and serious problems of the environment, confronting urban local governments in developing countries. Long experience states the failure and inefficiencies of local governments by hundreds of years in CWM. Khulna is the third largest metropolitan city in Bangladesh, The significance of the city is increasing day by day for its socioeconomic role, which helps to the overall development of Bangladesh. Here fig-6 shows all the medical centres situated in Khulna City corporation.

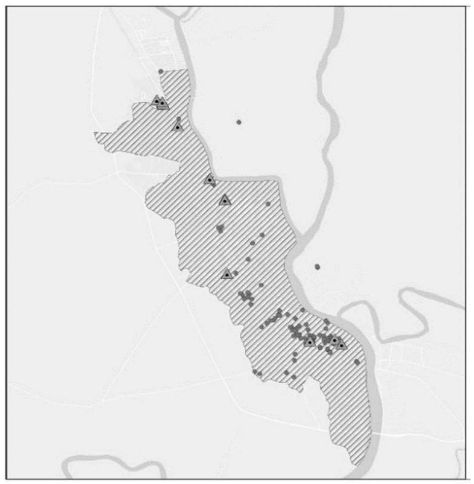

Fig-6: Map showing medical centres in Khulna City Corporation

2.1.1 Dataset

A total of 30 hospitals have been used in this study as the dataset. All the hospitals are situated in Khulna City Corporation. All the hospitals have been divided into various types such as hospitals, clinics, diagnostic centers, etc. Those hospitals are also categorized as Category A, Category B, and Category C. Here Category A means it is a large type of diagnostic center, and almost all types of tests are done here. Category B means it is a medium-type diagnostic center, and Category C means it is small in size only a few tests are done here. A questionnaire survey form is prepared. Data related to medical waste (no of beds, number of patients, service charge, daily generated waste, etc.) were collected by field visit. Questions for collecting required data were asked to several medical personnel of selected medical centers. A questionnaire survey form is prepared. Data related to medical waste (no of beds, number of patients, service charge, daily generated waste, etc.) were collected by field visit. Questions for collecting required data were asked to several medical personnel of selected medical centers. Then the data was prepared for use in machine learning algorithms.

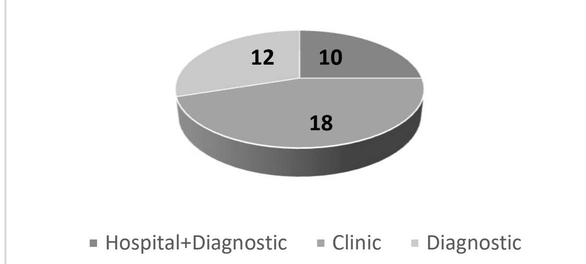

Fig-7: Pie chart based on the type of surveyed hospital

2.2 Model development

Data wrangling and pre-processing were done to build a suitable training data set for Machine Learning model development. Machine learning algorithms are trained on examples, analyzing historical data, and learning from prior events. A Python script is developed utilizing Scikit Learn to train the model. Python is a general-purpose, high-level programming language. Its design ethos places a high value on code readability and emphasizes indentation. A Python package for statistical modeling and machine learning models is called Scikit-Learn, or sk-learn. We can create several machine learning models for regression, classification, and clustering using sci-kit-learn, and we may use statistical tools to analyze these models. Lastly, Mean Absolute Error was evaluated and prediction was done on the test data set. Fig-8 shows the flow diagram of the machine learning process.

Input data → Analyze Data → Find Patterns → Prediction → Stores feedback

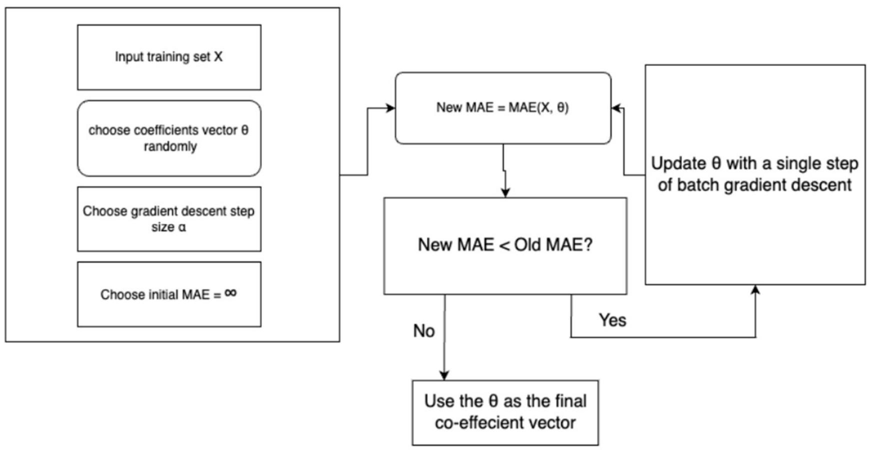

Fig-8: Flow diagram of model

2.3 Multiple linear regression analysis

Multiple linear regression is a powerful statistical tool for modeling the relationship between a dependent variable and a set of independent variables. It assumes a linear relationship between the variables and can be used to describe, predict, and test hypotheses. The method is a generalization of simple linear regression, allowing for the inclusion of multiple explanatory variables. It is particularly useful for assessing the associations of several variables with a continuous outcome and can be extended to include interactions, polynomial terms, and categorical variables. The linear regression technique is popularly employed for estimating waste generation due to its success in this regard (Hoang et al. 2017). Multiple linear regression (MLR) is a frequently used method by diverse disciplines to examine the case where a variable is influenced by more than one independent variable. The form of the general equation is shown in Eq. (1)

$$y = a + b \cdot x1 + c \cdot x2 + \cdots + z \cdot xn \qquad (1)$$

In the above equation, y is a dependent variable, which changes by alterations of independent variables. Independent variables are shown as x1, x2,…, xn. The a value is a constant and the b, c,…, z values are coefficients. The relationship between dependent and independent variables is controlled by coefficients. The coefficients' sign, which can be either positive or negative, indicates how the independent factors influence the dependent variable.

3. RESULTS AND DISCUSSION

Collected data from field visits to different medical centers, including the number of beds, number of patients, bed patient occupancy, and service charge are used for predicting waste. This study aimed to determine the factors that have the strongest influence on medical waste generation and assess the accuracy of our predictive model. To evaluate the performance of the models three statistical indices are used: the mean absolute error (MAE), the root-mean-square error (RMSE), and correlation coefficient (R2) values. Here fig-9, fig-10, fig-11, and fig-12 show the graphical relation between hospital waste and the number of beds, number of patients, service charge, and bed occupancy respectively.

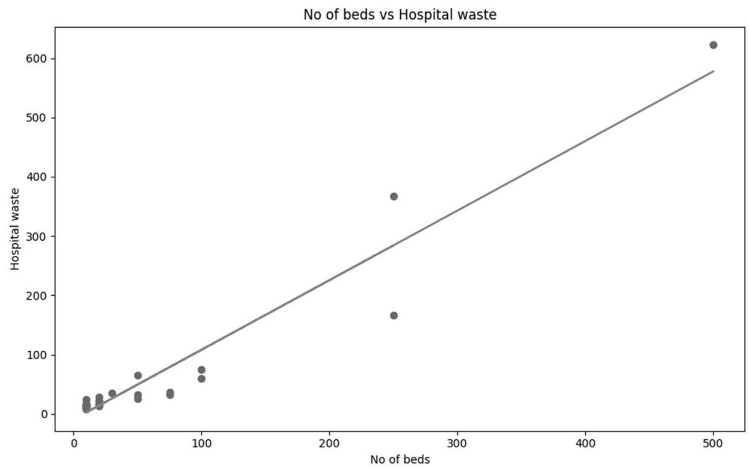

Fig-9: Regression using only total beds

Regression using only total beds
Confidence score: 0.9283518439050974
Coefficients (weights): [1.17383864]
Intercept: -9.2723721059495
Mean root squared error:33.24792402556723

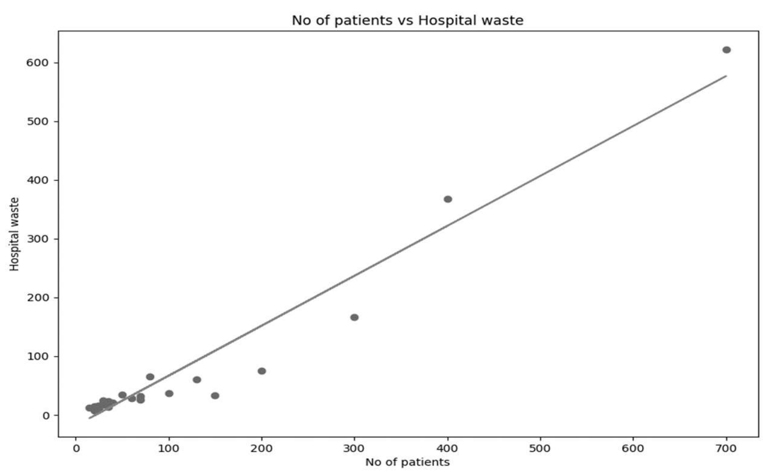

Fig-10: Regression using only no of patients

Confidence score: 0.9444983687114304
Coefficients (weights): [0.84971523]
Intercept: -18.313497789852264
mean absolute error:20.878964702890958
mean root squared error:29.26273392369493

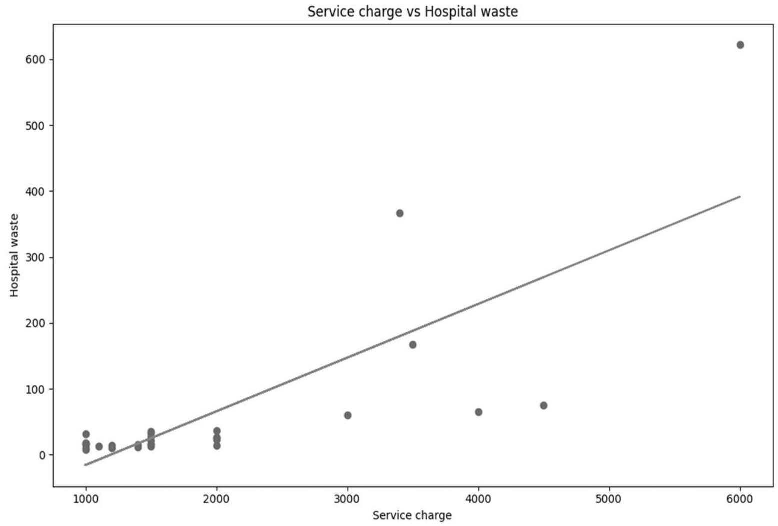

Fig-11: Regression using only the service charge

Confidence score: 0.6150917246402248
Coefficients (weights): [0.08145227]
Intercept: -97.25074571215511
mean absolute error:47.288379318916235
mean root squared error:77.06206609495212

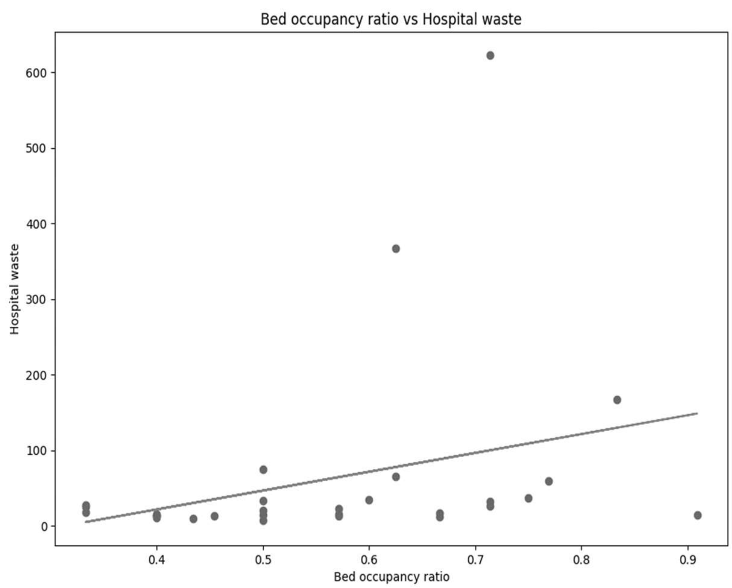

Fig-12: Regression using only bed occupancy ratio

Confidence score: 0.09487857683709466
Coefficients (weights): [249.12182303]
Intercept: -77.84022537275898
mean absolute error:62.13013501335164
mean root squared error:118.17223421711235

4. CONCLUSION

Results indicated that among the variables considered, the number of patients had the strongest relationship with the amount of medical waste generated. This finding suggests that as the number of patients increases, so does the amount of medical waste generated. This can be attributed to the fact that more patients typically result in increased utilization of medical resources, such as disposable supplies and medications, which contribute to the overall waste generation in healthcare settings. Furthermore, our analysis revealed that bed occupancy had the weakest relationship with medical waste generation. This implies that the utilization of beds, as measured by the occupancy rate, does not significantly impact the amount of waste generated. When the waste is predicted using all the factors, the Confidence score is 0.95, the mean absolute error is 18.68, and the mean root squared error is 27.86. This indicates that as the number of factors increases the accuracy of the model increases. In conclusion, our results are promising and may play a useful role in establishing proper medical waste management on the healthcare waste sources, generation, collection, transportation, treatment, and disposal.

5. RECOMMENDATION

After the study, it is recommended that Color-coded bins should be provided at each source of generation of medical waste. Hospitals should arrange training and awareness programs among the staff and ward boys periodically to provide them with knowledge for the proper collection process. Each hospital should have some workers at the storage room or transfer station to segregate the mixed wastes. Hospitals should spend more budget to ensure the proper treatment of their waste. Collection vehicles should have some facilities to convey the waste separately. A special treatment plant is needed for only treating medical waste.

ACKNOWLEDGEMENT

We want to express our special appreciation to all experts involved in this research resulting in such useful data. A special thanks to the manager of PRODIPON Md Kamal.

REFERENCES

Ahsan, A., Alamgir, M., Imteaz, M., Shams, S., Rowshon, K., Aziz, M. G., & Idrus, S. (2015). MUNICIPAL SOLID WASTE GENERATION, COMPOSITION AND MANAGEMENT: ISSUES AND CHALLENGES. A CASE STUDY. *Environment Protection Engineering*, *41*(3). https://doi.org/10.5277/epe150304

Alam Chowdhury, M. (2002). *A participatory approach to hospital waste management*.

Al-Shayea, Q., & El-Refea, G. (2013). *Predicting the Effects of Medical Waste in the Environment Using Artificial Neural Networks: A Case Study*. www.IJCSI.org

Altin, F. G., Budak, İ., & Özcan, F. (2023). Predicting the amount of medical waste using kernel-based SVM and deep learning methods for a private hospital in Turkey. In *Sustainable Chemistry and Pharmacy* (Vol. 33). Elsevier B.V. https://doi.org/10.1016/j.scp.2023.101060

Biswas, A., Amanullah, A., & Santra, S. C. (n.d.). *Medical Waste Management in the Tertiary Hospitals of Bangladesh: An Empirical Enquiry* (Vol. 5, Issue 2).

Çetinkaya, A. Y., Kuzu, S. L., & Demir, A. (2020). Medical waste management in a mid-populated Turkish city and development of medical waste prediction model. *Environment, Development and Sustainability*, *22*(7), 6233–6244. https://doi.org/10.1007/s10668-019-00474-6

Cheng, Y. W., Sung, F. C., Yang, Y., Lo, Y. H., Chung, Y. T., & Li, K. C. (2009). Medical waste production at hospitals and associated factors. *Waste Management*, *29*(1), 440–444. https://doi.org/10.1016/j.wasman.2008.01.014

Cubillos, M. (2020). Multi-site household waste generation forecasting using a deep learning approach. *Waste Management*, *115*, 8–14. https://doi.org/10.1016/j.wasman.2020.06.046

Elshaboury, N., Abdelkader, E. M., Alfalah, G., & Al-Sakkaf, A. (2021). Predictive analysis of municipal solid waste generation using an optimized neural network model. *Processes*, *9*(11). https://doi.org/10.3390/pr9112045

Erdebilli, B., & Devrim-İçtenbaş, B. (2022). Ensemble Voting Regression Based on Machine Learning for Predicting Medical Waste: A Case from Turkey. *Mathematics*, *10*(14). https://doi.org/10.3390/math10142466

Icrc. (n.d.). *MEDICAL WASTE MANAGEMENT*. www.icrc.org

Jahandideh, S., Jahandideh, S., Asadabadi, E. B., Askarian, M., Movahedi, M. M., Hosseini, S., & Jahandideh, M. (2009). The use of artificial neural networks and multiple linear regression to predict rate of medical waste generation. *Waste Management*, *29*(11), 2874–2879. https://doi.org/10.1016/j.wasman.2009.06.027

Jeba, J. S., & Rahman, Md. M. (2020a). Medical waste management in Khulna City Corporation, Bangladesh. *National Geographical Journal of India*, *66*(4), 306–319. https://doi.org/10.48008/ngji.1750

Jeba, J. S., & Rahman, Md. M. (2020b). Medical waste management in Khulna City Corporation, Bangladesh. *National Geographical Journal of India*, *66*(4), 306–319. https://doi.org/10.48008/ngji.1750

Johnson, N. E., Ianiuk, O., Cazap, D., Liu, L., Starobin, D., Dobler, G., & Ghandehari, M. (2017). Patterns of waste generation: A gradient boosting model for short-term waste prediction in New York City. *Waste Management*, *62*, 3–11. https://doi.org/10.1016/j.wasman.2017.01.037

Mohan, J., Gupta, A., Jaypee Institute of Information Technology University, Institute of Electrical and Electronics Engineers. Uttar Pradesh Section. SP/CS Joint Chapter, & Institute of Electrical and Electronics Engineers. (n.d.). *2019 International Conference on Signal Processing and Communication (ICSC) : 07-09 March 2019, Jaypee Institute of Information Technology, NOIDA*.

Patwary, M. A., O'Hare, W. T., & Sarker, M. H. (2011). Assessment of occupational and environmental safety associated with medical waste disposal in developing countries: A qualitative approach. *Safety Science*, *49*(8–9), 1200–1207. https://doi.org/10.1016/j.ssci.2011.04.001

Professor Rejuan Hossain Bhuiyan, S. (2008). *MEDICAL WASTE MANAGEMENT SYSTEM IN DHAKA CITY CORPORATION: SOUTH, BANGLADESH*.

Rathod, T., Hudnurkar, M., & Ambekar, S. (2020). USE OF MACHINE LEARNING IN PREDICTING THE GENERATION OF SOLID WASTE. In *PJAEE* (Vol. 17, Issue 6).

Sayed Mohammad Akid, A., Mahbub Hassan, K., Shahrear Apu, O., Arifur Rahman, S., Mahbub Hassan, K., & Azizur Rahman, M. (2014). Evaluating the existing condition of clinical waste management in Khulna city and its possible health hazards One Dimensional Non-Linear Consolidation Analysis of Dredged Soil : An Experimental and Theoretical study View project Recycled Aggregate Concrete View project EVALUATING THE EXISTING CONDITION OF CLINICAL WASTE MANAGEMENT IN KHULNA CITY AND ITS POSSIBLE HEALTH HAZARDS. In *KUET*. https://www.researchgate.net/publication/336305247

Zhou, H., Yu, X., Alhaskawi, A., Dong, Y., Wang, Z., Jin, Q., Hu, X., Liu, Z., Kota, V. G., Abdulla, M. H. A. H., Ezzi, S. H. A., Qi, B., Li, J., Wang, B., Fang, J., & Lu, H. (2022). A deep learning approach for medical waste classification. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-06146-2