7th European Conference on Industrial Engineering and Operations Management
Augsburg, Germany, July 16-18, 2024
Publisher: IEOM Society International, USA DOI: 10.46254/EU07.20240256
Published: July 16, 2024

Machine Learning Approach for the Prediction of Pandemic-Induced Medical Waste in the United Arab Emirates: Covid-19 Case Study

Reem Belal Irshaid, Noora Mohamed Abdulla Zayed, Mariam Eisa Mohamed Ismaeel Alhosani, Alia Saeed Mohammed Shareef Alkhoori, Mira Omar Mohamed Hassan Abdulla and Saed T. Amer
Health, Safety and Environment Engineering
Department of Management Science and Engineering
Khalifa University
Abu Dhabi, United Arab Emirates
100064411@ku.ac.ae, noora.zayed@ku.ac.ae, 100049684@ku.ac.ae, 100044344@ku.ac.ae, 100038235@ku.ac.ae, saed.amer@ku.ac.ae

Abstract

The COVID-19 pandemic triggered the urge to utilize machine learning to build models to predict biomedical waste for futuristic waste management. In this study, the machine learning approach (Random Forest Regressor) is used to build a model that predicts biomedical waste weight generated based on the COVID-19 pandemic case. Data gathered during the period of COVID-19, which is specific to the United Arab Emirates and key features of the study were involved in the machine learning algorithm (PCR tests, vaccine doses, Masks, and Death Wraps). The generated model showed 92.67% accuracy, reflecting the overall predictive accuracy relative to the average biomedical waste observed. Pareto chart is used to identify which feature affects the overall generated biomedical waste; in this case, the model's initial prediction for a hypothetical scenario showed that around 2015.53 kg of the total biomedical waste is generated from PCR tests, whereas the least contributor was due to death wraps around 41.13 kg. The model can predict a COVID-19 case like a pandemic that generates a huge amount of waste, with futuristic model improvements to include other vital features to the model's algorithms. The model has shown high accuracy which can be used in future pandemics to manage waste produced, especially medical waste.

Keywords
Medical waste prediction, Machine learning, Random Forest Regressor, COVID-19, Waste management

1. Introduction
The outbreak of the COVID-19 pandemic has not only posed unprecedented challenges to global healthcare systems but has also significantly impacted waste management practices, particularly in the field of medical waste. **Figure 1** shows the number of cases in GCC countries, vaccine doses, and deaths caused by the COVID-19 pandemic (Mathieu 2020). The rapid growth in infectious medical waste, including personal protective equipment (PPE), masks, gloves, and other disposable materials, has strained existing waste management infrastructures worldwide. The United Arab Emirates has been no exception to this challenge with its strong healthcare sector and dynamic economy. As the UAE continues to combat the COVID-19 pandemic, effective management of pandemic-induced medical waste has become a critical priority for ensuring public health and environmental sustainability.

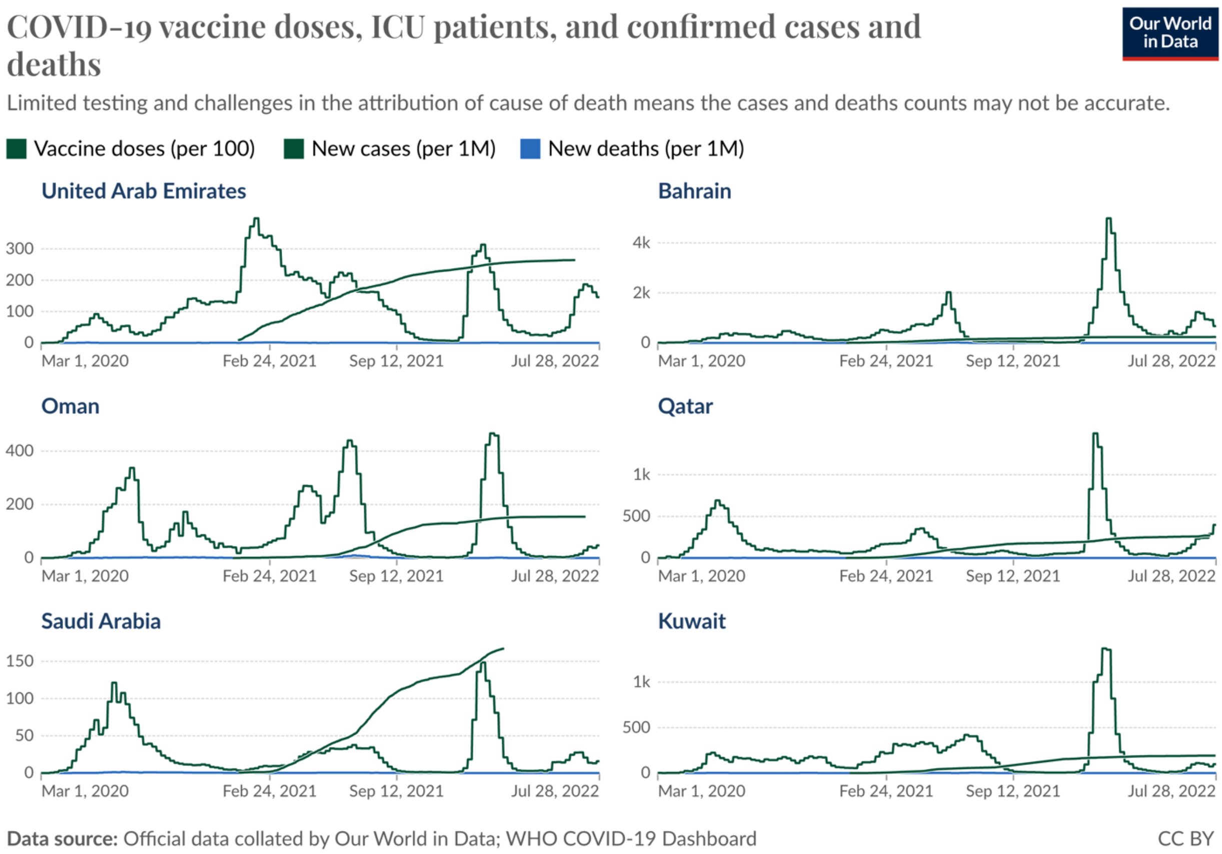

Figure 1. Covid-19 cases, vaccines, and deaths in the GCC counties. The graphs show the rapid spread of the disease between March 2020 and July 2022. These numbers led to the generation of medical waste across the countries.

In response to the pressing need for innovative solutions, this paper proposes a machine-learning approach for predicting pandemic-induced medical waste in the UAE, focusing on the COVID-19 case analysis. Machine learning, a subset of artificial intelligence, has shown great potential in various domains, including healthcare and environmental science. By manipulating data on medical waste generation, epidemiological trends, population dynamics, and other relevant factors, machine learning models can be trained to forecast the volume and composition of medical waste generated during the COVID-19 pandemic in the UAE (Ngoc 2022).

Assessing and evaluating the COVID-19 waste management plan is crucial for developing a comprehensive framework for future pandemics. Also, exploring existing waste management models globally provides insights into creating an integrated framework for effective pandemic waste management. Moreover, proposing an adaptable waste management system based on the UAE approach sets the stage for a comprehensive and sustainable approach to handling pandemic medical waste (Dhaiban 2023).

By implementing the power of machine learning, this research seeks to provide valuable insights and decision-support tools to policymakers, healthcare administrators, waste management authorities, and other stakeholders involved in pandemic response efforts in the UAE. Through proactive planning and evidence-based decision-making, the proposed approach aims to enhance the resilience and effectiveness of the UAE's medical waste management infrastructure in the face of the ongoing COVID-19 pandemic and future public health crises.

2. Literature Review

Healthcare facilities serve as the primary sources of healthcare waste. Healthcare waste, also known as medical waste and biomedical waste, consists of various materials generated within healthcare facilities, ranging from harmless to hazardous, infectious to non-infectious, and chemical in nature. Approximately 1-2% of general waste production is attributed to healthcare waste, 85% of wastes are non-hazardous and 15% are classified as hazardous, including radioactive, infectious, or toxic materials. Hospitals, medical labs, clinics, research facilities, and blood banks are among the primary generators of healthcare waste. The production of healthcare waste varies among countries globally and is influenced by several factors. These factors include waste management practices, the diversity of healthcare facilities and their specialties, the availability of reusable equipment, and the volume of daily patient care. Nevertheless, documented healthcare waste production tends to be lower in developing nations compared to developed ones (Janik-Karpinska 2023).

Since March 2020, the global focus has been on combating the COVID-19 pandemic, prompting concerns about whether insufficient waste management could contribute to the spread of the virus. Studies have revealed varying survival rates of the virus on different surfaces, ranging from 4 hours on copper to up to 3 days on plastic surfaces and stainless steel. The stream in waste generation during the pandemic, along with the disposal of contaminated disposable personal protective equipment (PPE), has strained waste management systems. Consequently, ensuring efficient, safe management of COVID-19 medical waste has become a crucial aspect of pandemic control efforts. Alongside established standards covering identification, collection, storage, transportation, processing, and disposal, factors such as disinfection and training have emerged as integral components of effective healthcare waste management. There are no substantial differences between the overall management of COVID-19-related healthcare waste and standard medical waste management practices before the pandemic (Janik-Karpinska 2023).

The COVID-19 pandemic has resulted in a surge in household waste production, leading governments to postpone waste recycling initiatives. Consequently, there has been an increase in the shipping of unrecycled packaging materials, exacerbating environmental degradation. Moreover, the management and disposal of waste during the pandemic pose significant environmental challenges. The widespread use of PCR tests for diagnosing the virus has generated over 15,000 tons of plastic waste globally, raising concerns about its proper disposal and control. Around 97% of plastic residues from these tests are incinerated, emitting toxic chemicals into the atmosphere during the process. Additionally, the excessive use and disposal of personal protective equipment (PPE) has led to marine environment contamination in various regions worldwide. The United Arab Emirates ranks ninth among the top ten countries globally in terms of plastic waste generation. The improper disposal of face masks has contributed to environmental degradation, particularly in areas with insufficient waste collection infrastructure. The increased use of facemasks in the UAE has led to the contamination of marine ecosystems (Alalawi 2022).

Globally, the frequent occurrences of the virus and the waste made it incredibly challenging for academic research. This research primarily delves into the root causes, consequences, and management strategies of the crisis. Additionally, scholars have examined external factors contributing to the medical waste disposal crisis, including deficiencies in regional management systems and the government's failure to fulfill its public health role. As for the manifestations of the medical waste crisis, the widespread improper disposal of medical waste poses significant threats to society, the economy, and the environment. These impacts include severe deterioration of air quality, threats to public hygiene, risks to public health, destruction of aquatic ecosystems, and substantial disruptions to global trade and investment (Ma 2022). The proposed study involves the construction of a machine-learning model that predicts pandemic-induced waste so that it can be handled and disposed of correctly.

Used machine learning models solve practical problems that include image classification and analysis. These include convolutional neural networks and deep neural networks. These models help in reducing the manual image classification for medical waste. However, the main problem is finding an algorithm that will help manage these wastes (Zhou 2022). The random forest model, a prominent machine learning approach, has become a significant focus in artificial intelligence research due to its robust adaptive learning capabilities and nonlinear mapping proficiency. This makes it well-suited for simulating wastewater treatment processes, which are characterized by significant lag, non-linearity, and multiple variables (Cheng 2023).

A key application of the random forest algorithm is random forest regression, based on the statistical learning theory developed by Breiman. Random forest regression uses Bootstrap resampling to draw multiple samples from the original dataset and build decision trees for each sample. These decision trees are then aggregated to make predictions, with the final prediction being the average of the outputs from all the trees. The core of the random forest regression algorithm lies in using multiple decision trees to make predictions, combining their outputs to enhance accuracy. This approach offers high prediction precision, strong generalization capabilities, and quick convergence, and requires minimal parameter adjustments, effectively mitigating the risk of overfitting. It is well-suited for handling diverse datasets and robustly extracting variables from ultra-high-dimensional vector spaces. Random forest regression has widespread application across various fields, including medicine, management, and agriculture. By leveraging limited samples to construct multiple decision tree models, random forest regression increases the diversity of the decision trees and enhances the accuracy of the final integrated model (Cheng 2023).

3. Methods

3.1. Prediction of Biomedical Waste Quantities

The prediction of biomedical waste was coded in Python, Jupyter's Notebook, as a Random Forest Regressor model, where the machine learning algorithm is versatile, robust, and can handle complex data relationships. The Random Forest Regressor combines ensemble learning and decision trees to make predictions for continuous outcomes.

Data Extraction, Preprocessing, and Feature Engineering

The extracted data from the literature was firstly inputted into the algorithm, which consists of columns of the number of cases, PCR weight (kg), vaccine weight (kg), masks weight (kg), population, death wrap weight (kg) and total biomedical waste (kg) over two years; the climax of the pandemic (2020-2022). The selected columns were converted to float time as part of the data preprocessing to ensure that they were numerical for further calculations. Consequently, feature engineering was performed as new features were calculated based on the existent features in the dataset. The features included 'ratio of PCR weight to the number of cases', 'ratio of vaccine weight to the population', 'ratio of mask weight to the population,' and lastly, 'ratio of death wrap weight to the population'. These new features can provide additional information for predicting biomedical waste quantities.

Data Splitting and Feature Scaling
The data was then split into training and testing sets. This allows for model training on one portion of the data (80%) and evaluation on another portion (20%) to assess its performance. The features were also standardized to have a mean of 0 and a standard deviation of 1, which is crucial to ensure fair treatment of features for improved model performance, faster convergence, and better interpretability and robustness to outliers.

Model Training and Model Evaluation
The training data were initialized and trained as a Random Forest Regressor model; the model learns to predict the total biomedical waste based on the features provided. The model was then evaluated through various regression metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Square Error (RMSE), and R-squared (R2) score to evaluate the performance of the trained model on the test data.

Feature Importance Visualization
The importance of each feature was visualized in biomedical waste prediction using a bar plot. This gives a visual representation of the most significant influence on the model's predictions.

3.2. Prediction of Total Biomedical Waste for a Hypothetical Next Pandemic
A hypothetical scenario was stimulated where predictions are made for the features and total biomedical waste for a 'next pandemic' based on a trained Random Forest Regressor model.

Sample Data Creation, Feature Selection, and Feature Scaling
The next pandemic data was defined as a sample dataset that represents the number of cases and population for the next pandemic scenario. The target columns' population' and 'number of cases' were dropped from the next pandemic data frame to create another Data Frame that contains only the input features required for prediction. The features' population' and 'number of cases' were selected to train a model for predicting other features. And the features were scaled.

Model Training for Feature Prediction
The Random Forest Regressor model is initialized for reproducibility and the scaled training data is used as well as the features to be predicted to train the model; during the training, the model learns the underlying patterns and relationships between the input features population and number of cases and the target features to be predicted 'total biomedical waste (kg) and possibly other features. The Random Forest Regressor algorithm constructs an ensemble of decision trees based on the training data where each decision tree in the ensemble is trained on a random subset of the training data and features, introducing randomness to reduce overfitting. The trees in the ensemble collectively make predictions, and the final prediction is often an average or a weighted average of the individual tree predictions.

Feature Prediction for the Next Pandemic
The trained model is used to predict the values of features other than 'population' and 'number of cases' for the next pandemic scenario; these predicted features represent various aspects related to the pandemic, such as healthcare resource utilization, disease spread dynamics, or public health interventions. The predicted feature values are stored in a separate Data Frame where each column corresponds to a predicted feature, such as hospital bed demand, medical supply requirements, or healthcare personnel needs. The trained model predicts total biomedical waste based on the predicted features and the next pandemic numbers. This prediction provides insights into the potential impact of the next pandemic, estimating the total biomedical waste, better preparation, and resource allocation to manage waste disposal, and ensuring environmental and public health safety during the pandemic.

Prediction Outcomes
The model outcomes comprise printed predicted total biomedical waste for the next pandemic and the predicted features for the next pandemic.

4. Data Collection
In this section, we provide detailed information about the process of data collection involving the COVID-19 cases

level and quintessential biomedical wastes during the spread of the virus in the United Arab Emirates (UAE). The collected data comprise the number of diagnosed cases, daily PCR tests, vaccinations, masks, and deaths. These datasets play a particularly important role in understanding the pandemic's dynamics and the waste of medical things it generates.

The data was obtained from reputable sources to help guarantee authenticity and boost the information's reliability. Specifically, the following sources were utilized: Information on diagnosed cases, daily PCR tests, and deaths were sourced directly from the official COVID-19 dashboard of the UAE government (National Emergency Crisis and Disaster Management, 2023). To get a representative sample, 68 dates were randomly chosen from March 2020 to June 2022, which makes the sample remarkably diverse because of the variety of case scenarios that were selected. Data on the coronavirus vaccine distributed and administered in the UAE were retrieved from Our World in Data, one of the most competitive databases of statistics and trends in the field (*Daily COVID-19 vaccine doses administered*, 2024). In this respect, the dates that were picked articulate with the cases diagnosed, the conducted PCR tests, and the deaths that occurred on those specific dates for temporal consistency in the analysis. Data relating to the use of face masks in the UAE were obtained from a scholarly article (Al-Omran et al., 2023). This source offers insights into mask utilization and supports the estimation of mask weights in the process of biomedical waste calculation.

In finding out the specific weights of components in biomedical wastes, the specific weights were based on existing standards and relevant literature. The daily number of vaccines was calculated by multiplying the number of vaccines administered by 0.00824kg, which is equal to the weight of a single vaccine dose (Al-Omran et al., 2021). Likewise, the weight of PCR tests per day was calculated through the multiplication of the number of tests carried out by 0.01252 kg, which signifies the weight of one PCR test (Al-Omran et al., 2021). The weight of death wraps, used to cover deceased patients within the infirmaries was projected at 3.5 kg per wrap. Furthermore, the weight of masks for each day was computed and determined by the multiplication of daily mask count by 0.00285 kg, which is the amount of weight a single mask would have.

The total biomedical waste per day was, therefore, computed as the sum of the weights of PCR tests, vaccines, masks, and death wraps. These computations will help form a complete overview and comprehension of the amount and material composition of the medical waste that occurred during the COVID-19 pandemic in the UAE. Overall, the data collection process was carried out in a very meticulous way, which gave the collected information a high level of accuracy and reliability, thus allowing the analysis of the pandemic-induced medical waste management in the UAE to be strong and comprehensive.

References were sought to provide evidence to back up the assigned weights of one PCR test, one mask, one vaccine dose, and one death wrap. These values are the key to determining the whole biomedical waste produced during the COVID-19 epidemic. As an example, comprehending one PCR test's weight lets one calculate the total weight of tests done in a day, and the same goes for vaccine doses, which enables the correct understanding of the total weight of vaccines injected. The same is true for the weight of one death wrap, as it is vital for estimating the overall weight of a deceased patient's wraps. Also, one mask's weight holds the same value for measuring the mass of all daily worn masks. References obtained serve to validate and substantiate these allotted weights, guaranteeing the accuracy and dependability of the waste estimation process.

5. Results and Discussion

5.1 Numerical Results

Data Preparation and Feature Engineering

The study was initiated by transforming all relevant variables into floating-point numbers to maintain consistency across the data manipulations. The analysis focused on several key variables, and the dataset was tested through feature engineering to include ratios such as 'PCR to cases ratio,' 'Vaccine to population ratio,' 'Masks to population ratio,' and 'Death wrap rate'. These ratios were anticipated to provide deeper insights into the factors influencing biomedical waste generation during the pandemic. This step is critical in developing AI models for predicting biomedical waste during pandemics. Proper data preparation ensures consistency and reduces errors. This standardization is crucial for enhancing model compatibility with machine learning algorithms and improving the overall predictive accuracy and reliability of the model.

For instance, ratios such as "PCR to cases ratio," "Vaccine to population ratio," and "Masks to population ratio" provide a perception of the scale of medical interventions relative to the population size and infection rates. These features help the model capture non-linear relationships and improve the interpretation of the model's outputs, enabling policymakers to understand which aspects of pandemic management most heavily influence waste production.

Model Training and Evaluation

The features were standardized to reduce variance and improve the performance of the machine learning algorithm. A random forest regressor was employed due to its robustness and effectiveness in handling outliers and non-linear data. The model's performance was evaluated using various metrics:

- Mean Absolute Error (MAE): 183.30 kg, indicating the average error in predictions.
- Mean Squared Error (MSE): 48,480.18 kg^2, suggesting variability in the prediction errors.
- Root Mean Squared Error (RMSE): 220.18 kg, providing a measure of error magnitude.
- R^2 Score: 0.978, representing a high degree of variance explained by the model.
- Accuracy Percentage: 92.62%, reflecting the overall predictive accuracy relative to the average biomedical waste observed.

These metrics collectively suggest that the Random Forest model achieved a high level of predictive accuracy, making it a reliable tool for forecasting biomedical waste outputs based on the modeled predictors. The effectiveness of the model, as demonstrated in the provided results, underscores its importance in shaping AI-driven strategies for future pandemic responses, specifically in the management of biomedical waste. The random forest regressor, chosen for its robustness and reliability, provided a high R^2 score of 0.978, indicating that the model could explain approximately 98% of the variability in biomedical waste production during the COVID-19 pandemic. This high level of accuracy demonstrates the model's ability to capture complex patterns in pandemic-related data, making it a powerful tool for predicting waste outputs under similar circumstances. Additionally, the relatively low RMSE of 220.18 kg highlights the model's precision in quantitative forecasts, which is crucial for planning waste management logistics and capacities.

The model's evaluation metrics MAE, MSE, RMSE, and a 92.62% accuracy percentage serve multiple functions. They validate the effectiveness of the feature engineering and data preparation strategies. Also, these metrics guide iterative improvements in the model by pinpointing areas where predictions could be enhanced. For instance, by analyzing instances where the model underperforms, the features can be refined or adjusted to improve accuracy.

Prediction of Total Biomedical Waste for a Hypothetical Next Pandemic

Initial predictions focused on deriving expected feature values for a hypothetical scenario. The features predicted include the weights of PCR tests, vaccines, masks, and death wraps, along with corresponding ratios that relate these values to the number of cases and population size. These predictions are instrumental in understanding each factor's relative impact on biomedical waste.

- Predicted Total Biomedical Waste (kg) for the next pandemic: 6684.94 kg
- Predicted PCR weight (kg): 2015.53, suggesting a significant reliance on PCR testing in the modeled scenario.
- Predicted vaccine weight (kg): 1101.43, indicating substantial vaccine distribution.
- Predicted mask weight (kg): 226.36, reflecting protective measures in response to the pandemic.
- Predicted death wrap weight (kg): 41.13, which relates to the management of fatalities due to the pandemic.
- Ratios: These include PCR-to-cases ratio (0.623), vaccine-to-population ratio (0.000115), masks-to-population ratio (0.000024), and the death rate (0.012649), providing insights into the scale and response intensity of medical interventions compared to population and infection rates.

This prediction is vital as it quantifies the potential waste challenges that might arise, allowing for better resource planning and management. The use of predictive modeling to estimate future pandemic-induced biomedical waste is a proactive approach that aids in the preparation and effective management of healthcare resources. The ability to forecast the scale of required medical interventions (such as testing and vaccination) and the resultant waste can significantly influence strategic planning and operational efficiencies in pandemic responses. These predictions also facilitate the development of targeted waste reduction strategies, ensuring environmental sustainability and public health safety are maintained during pandemic outbreaks. These results, derived from a robust AI model, demonstrate the potential of machine learning in enhancing predictive capabilities and preparedness for future health crises, thereby mitigating their impact on both healthcare systems and the environment.

5.2 Graphical Results

The Random Forest Regressor model was subjected to a feature importance analysis to identify the primary factors contributing to biomedical waste generation. Figure 2 illustrates the relative importance of each feature in the predictive model, highlighting which aspects of the pandemic data (e.g., PCR tests, vaccine distribution, mask usage) are most predictive of biomedical waste volumes. As shown in Figure 2, the most significant factor influencing biomedical waste production is the weight of PCR tests (PCR Weight (kg)), which has an importance value of approximately 0.7. This indicates that the frequency and scale of PCR testing play a crucial role in determining the volume of biomedical waste during a pandemic.

Other notable features include the weight of masks (masks weight (kg)) and the vaccine-to-population ratio, which also contribute to the model's predictions but to a lesser extent. The number of cases, the population size, and the death rate have relatively low importance values, suggesting that they are less influential in predicting biomedical waste volumes. These results underscore the importance of PCR testing as a significant driver of biomedical waste. Therefore, managing and optimizing PCR test usage can be critical for controlling biomedical waste during future pandemics. Understanding the importance of these features helps policymakers and healthcare managers prioritize resources and develop targeted waste management strategies to mitigate environmental impacts.

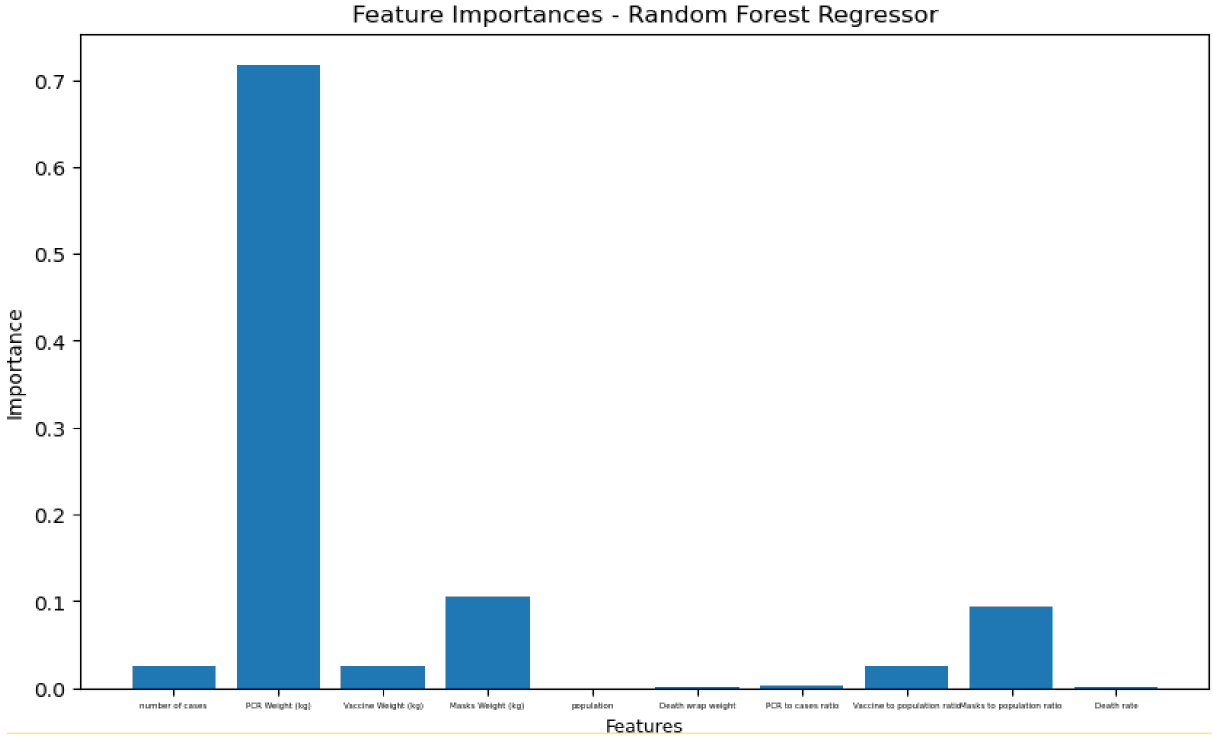

Figure 2. The relative importance of features in predicting biomedical waste generation during a pandemic is derived from the Random Forest model.

This detailed analysis of features of importance provides valuable insights for planning and managing biomedical waste during pandemics. By identifying the key drivers of waste production, the model helps inform strategic decisions and optimize resource allocation to address the most impactful factors.

Model Scalability and Applicability

The model's scalability and applicability to other contexts beyond the UAE can be applied and considered by the utilized data. The data preparation techniques and feature engineering strategies used in this study are universally applicable, allowing the model to be adapted to different geographic and demographic contexts. For instance, global data on PCR tests, vaccine distribution, mask usage, and death rates during the COVID-19 pandemic have been widely collected, making it feasible to apply this model to other countries with similar data availability.

To apply this model in other countries, the following steps are recommended:

1. **Data Collection:** Gather data on PCR testing, vaccine distribution, mask usage, and death rates specific to the country of interest. Ensure the data is comprehensive and covers the entire pandemic period.
2. **Feature Engineering:** Apply the same feature engineering techniques to create ratios such as 'PCR to cases ratio,' 'Vaccine to population ratio,' 'Masks to population ratio,' and 'Death wrap rate.' This ensures consistency and comparability across different regions.
3. **Model Training:** Use the collected data to train the random forest regressor, ensuring the model is tailored to the country's specific context and demographic characteristics.
4. **Model Evaluation:** Evaluate the model's performance using metrics such as MAE, MSE, RMSE, R2 Score, and Accuracy Percentage. Compare these metrics to the original study to assess the model's reliability and accuracy in the new context.
5. **Prediction and Analysis:** Use the trained model to predict biomedical waste outputs for future pandemic scenarios. Analyze the predictions to inform strategic planning and resource allocation.

These global data points demonstrate the model's potential to be scaled and adapted for use in diverse settings, providing valuable insights and enabling effective pandemic management strategies worldwide. This adaptability ensures that the findings and methodologies from this study can contribute to global efforts to predict and manage biomedical waste during future pandemics.

5.3 Proposed Improvements

1. **Data acquisition:** data collection was challenging since we had to find comprehensive and high-quality data specific to COVID-19 cases. In this work, the primary data inventory is selected from the UAE government (National Emergency Crisis and Disaster Management, 2023); however, for future work international data around biomedical wastes generated during the COVID-19 period can be incorporated to have an international model. In addition, data normalizing can be implemented during data reprocessing to enhance the model's interpretability and accuracy.

2. **Features Selection:** this work included the most utilized variables during the identification and treatment of the COVID-19 pandemic (PCR tests, Vaccine doses, Masks, and Death Wraps). However, many features can be considered too. Recursive Feature Elimination analysis can be used to identify the most relevant features for the case study.

3. **Model parameters fine-tuning:** looking at the Mean Absolute Error (MAE), which is around 183.30 kg that corresponds to the average error in prediction, the learning rate, regularization strength, and model complexity can be adjusted to optimize the performance of the model.

4. **Data validation:** techniques such as K-Fold cross-validation can be used to validate various subsets of data. Nonetheless, the accuracy of the model generated in this work reached 92.62% accuracy of the overall predictive accuracy. By applying these improvements, the accuracy and performance of the machine learning model will enhance in addressing futuristic biomedical wastes, hence supporting decision-making for biomedical waste management strategies.

5.4 Validation

To identify the most significant biomedical wastes that contribute to total biomedical wastes generated during COVID-19, a Pareto chart is produced in Figure 3 Bars represent the categories of biomedical wastes or features included in this study (the highest bar is the PCR test) impact the overall waste presented, whereas the line (in green) corresponds to the cumulative total, as it gives an insight into how does the effects of each feature on the other one (steep rise from PCR test to vaccine doses). The chart in Figure 3 shows that PCR tests and Vaccine doses are both potentially the most important variables in the total generated biomedical waste, in which the "70/30" rule is applied where 70% of the effect comes from 30% of the causes. Pareto charts can be used in future pandemics to help prioritize resources and, hence, decide on the appropriate treatment methods and their relative disposal pattern to reduce the overall biomedical waste.

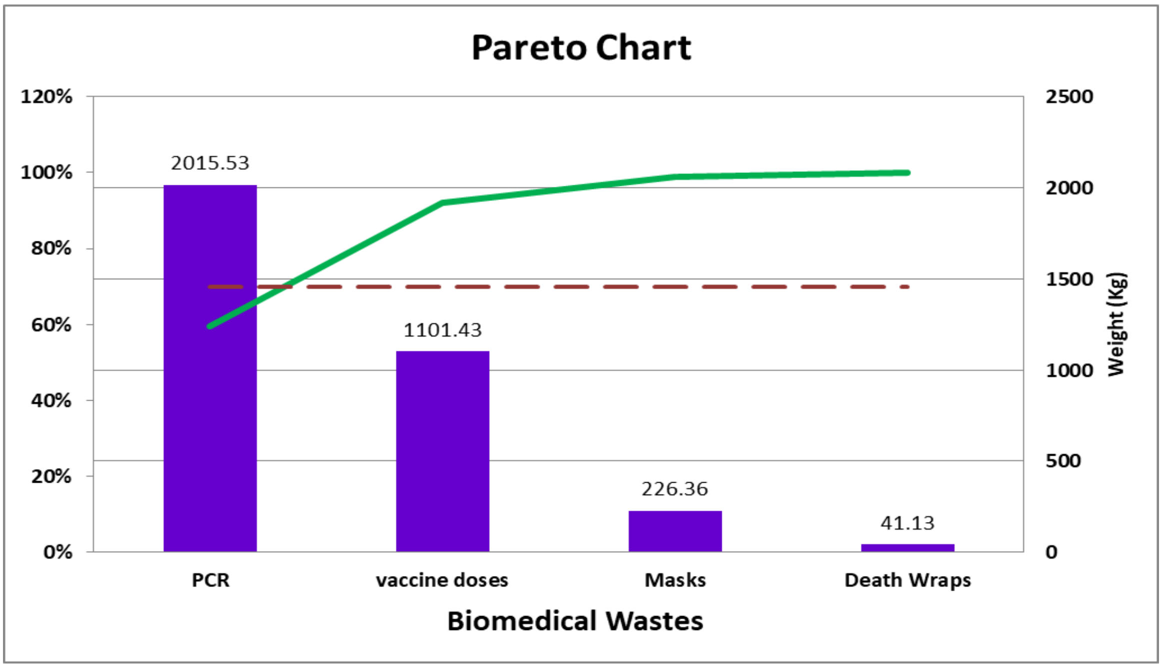

Figure 3. Pareto Chart of Biomedical Wastes "70/30", the chart displays the distribution of biomedical waste types by weight in kilograms, with the left vertical axis showing the percentage contribution, and the right vertical axis showing the total weight. PCR waste and vaccine doses together contribute to over 70% of the total biomedical waste, while masks and death wraps make up a smaller fraction according to the paretro principle.

6. Conclusion

During the COVID-19 pandemic, there was an increase in medical waste generated due to the heightened demand for Personal Protective Equipment (PPE), testing supplies, and disposable medical equipment. Efforts were made to ensure the safe and efficient disposal of COVID-19-related medical waste. Therefore, many countries had to adapt to the increased volume of medical waste by implementing enhanced disposal practices. Machine learning can be used to build models that predict pandemics by analyzing various data sources such as previous pandemic data like patient's healthcare information, and the machine learning algorithm can identify patterns and trends that may indicate the emergence of a pandemic. In this work, a Random Forest Regressor model is used to model the COVID-19 pandemic, in which data is collected from the official COVID-19 dashboard of the UAE government (National Emergency Crisis and Disaster Management, 2023). Then data is reprocessed for machine learning training that involves collecting recent data and scaling these numerical data. The features selection process included the weights of PCR tests, vaccines, masks, and death wraps, which were concluded to be the most used during the COVID-19 period. The importance Vs. The feature plot showed that the PCR test weight (kg) scored the highest in importance, corresponding to the availability of these tests during the COVID-19 period. The model also predicted that 6684.94 kg of biomedical waste would be generated in a COVID-19 case-like scenario while highlighting the contribution of PCR test weight in kg during the COVID-19 pandemic. The generated model can predict a similar case/pandemic in which the main variables/features are similar to the study in this paper, which was the main goal of this study, in which most medical treatments use masks and tests, and in later stages, vaccines will be developed. To conclude, the COVID-19 pandemic encouraged the utilization of machine learning when it comes to waste management by providing a predictive model for proper waste handling and disposal.

References

Mathieu, E., Ritchie, H., Rodés-Guirao, L., Appel, C., Giattino, C., Hasell, J., … Roser, M. Coronavirus (COVID-19) deaths. *Retrieved from* https://ourworldindata.org/covid-deaths, 2020.

Ngoc, S.-M. V., Nguyen, M.-A., Nguyen, T.-L., Thi, H. V., Dao, T. L., Bui, T. M., and Chu, D.-T., COVID-19, and environmental health: A systematic analysis for the global burden of biomedical waste by this epidemic, *Case Studies in Chemical and Environmental Engineering*, vol. 6, 2022.

Janik-Karpinska, E., Brancaleoni, R., Niemcewicz, M., Wojtas, W., Foco, M., Podogrocki, M., and Bijak, M., Healthcare waste—a serious problem for Global Health, Healthcare, vol. 11, no. 2, 2023.

Alalawi, S., Issa, S. T., Takshe, A. A., and ElBarazi, I., A review of the environmental implications of the COVID-19 pandemic in the United Arab Emirates, *Environmental Challenges*, vol. 8, 2022.

Ma, G., Ding, J., and Lv, Y., Seir evolutionary game model applied to the evolution and control of the medical waste disposal crisis in China during the COVID-19 Outbreak, *Sustainability*, vol. 14, no. 18, 2022.

Zhou, H., Yu, X., Alhaskawi, A., Dong, Y., Wang, Z., Jin, Q. Lu, H.. A deep learning approach for medical waste classification*. Scientific Reports*, vol 12, 2022.

Cheng, Q., Chunhong, Z., and Qianglin, L. Development and application of random forest regression soft sensor model for treating domestic wastewater in a sequencing batch reactor. *Scientific Reports*, vol 13, 2023.

Al-Omran, K., Khan, E., and Perna, S., Medical waste generation during COVID-19 pandemic in selected member countries of Arabian Gulf region; Google Trend Analysis, *Arab Journal of Basic and Applied Sciences*, vol. 30, no. 1, pp. 79–91, 2023.

Al-Omran, K., Khan, E., Ali, N., and Bilal, M., Estimation of COVID-19 generated medical waste in the Kingdom of Bahrain, *Science of The Total Environment*, vol. 801, 2021.

Daily COVID-19 vaccine doses administered, Our World in Data, Available: https://ourworldindata.org/grapher/daily-covid-19-vaccination-doses?tab=chart&time=2021-04-07..2021-10-12&country=~ARE, May 2024.

Dhaiban, A. A., Amro, A. Bin, Chulliyil, H. M., Ahmed, S. E., and Amer, S. T., Biological Waste Management for COVID-19 and Future Pandemics: A Holistic Framework, *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Jul. 18, 2023.

National Emergency Crisis and Disaster Management, Home: UAE coronavirus (COVID-19) updates, NCEMA, Available: https://covid19.ncema.gov.ae/en, May 2023.

Biographies

Reem Belal Irshaid is currently a master's student at Khalifa University, pursuing a degree in Biomedical Engineering and Biotechnology. Her academic journey began with a bachelor's degree in Applied Physics and Astronomy, where she developed a solid foundation in scientific principles and analytical thinking. During her undergraduate studies, she cultivated a deep interest in the intersection of biology and technology, which inspired her to further my education in this dynamic field. At Khalifa University, she is actively engaged in cutting-edge research projects that aim to innovate and improve technologies. Her commitment to advancing in technology drives her academic and professional endeavors. With a strong background in physics and a passion for biomedical engineering, she aspires to contribute to groundbreaking advancements in Triboelectric nanogenerator technology.

Noora Mohamed Abdulla Zayed is pursuing a master's degree in biomedical engineering at Khalifa University of Science and Technology. After earning her bachelor's from the same major and university, she interned at the Tissue Engineering Research Group at the Royal College of Surgeons where she published her first paper as a second author titled “Functionalizing Collagen-Based Scaffolds with Platelet Rich Plasma for Enhanced Skin Wound Healing Potential”. After that, she joined the workforce for two years during the pandemic to become a research assistant at the water intelligence testing center that uncovered COVID-19 cases before they were made public through wastewater testing. In that, she published her second paper titled “Long Term Study on Wastewater Sars-Cov-2 Surveillance across the United Arab Emirates”. Her research interests lie in Biotechnology and Tissue Engineering and she hopes for a research position sometime in the future.

Mariam Eisa Mohamed Ismaeel Alhosani is a master’s student in biomedical engineering and biotechnology and the chair of advocacy and outreach for the BMED student council. She is driven by a profound passion for biomedical sciences and human health enhancements and believes in the transformative power of science and technology to address global challenges and create meaningful societal change. During her studies, Ms. Alhosani has explored various aspects of biomedical engineering, including biomaterials, biomechanics, biosensors, and forensic applications, gaining practical skills and a deep understanding of the field's complexities. Her notable work includes a research project on bioelectronics and wearables focusing on TENG technology. Looking forward, she aims to delve deeper into bioelectronics and biosensors to develop innovative diagnostic tools and medical devices. She is also committed to addressing healthcare disparities by making healthcare more accessible and affordable, particularly in underserved communities.

Alia Saeed Mohammed Shareef Alkhoori is a master's student in System Engineering and Management and holds a bachelor's degree in civil engineering. With a strong foundation in engineering principles and a keen interest in leveraging technology for societal benefits, Alia has focused her research efforts on applying machine learning techniques to address contemporary challenges. Her work on predicting pandemic-induced medical waste in the United Arab Emirates exemplifies her commitment to interdisciplinary solutions that blend engineering, management, and data science. Alia is dedicated to pioneering sustainable and efficient healthcare practices, driven by her passion for innovative research and transformative practical applications in the field.

Mira Omar Mohamed Hassan Abdulla is a master’s student in Materials science and engineering at Khalifa University, United Arab Emirates. Mira Completed a bachelor's degree in the mechanical and nuclear engineering field. During her bachelor’s studies, Ms. Mira published her first research paper at the Materials Research Society and proudly presented it at the MRS conference (2017), in the United States of America. In her master’s degree, she is investigating the effects of using carbon material, namely graphene as a support for CO2 methanation catalysts. She worked at NAWAH energy company, United Arab Emirates, in Barakah Nuclear Power Plant as a nuclear reactor operator. Apart from that, she enjoys oil painting as she participates in international exhibitions such as the Louvre Museum, in Paris. She is also interested in linguistic studies; besides being a native Arabic speaker and an excellent English speaker, she is building up her Russian language skills too.

Saed T. Amer holds a Ph.D. in Computer and Information Systems Engineering obtained in August 2012 from Tennessee State University in the United States. He is currently a distinguished faculty member in the Department of Management Science and Engineering at Khalifa University. Dr. Amer's research is centered on Health, Safety, and Environment (HSE) engineering, with a strong emphasis on training and education. His extensive expertise encompasses fire protection, industrial hygiene, system safety, quality health, safety, and environment (HSE), as well as human factors related to HSE.