Exploring Machine Learning and Deep Learning Approaches for Multi-Step Forecasting in Municipal Solid Waste Generation

Oshan Mudannayake[a,1], Disni Rathnayake[a,1], Jerome Dinal Herath, Dinuni K Fernando[a,*], MGNAS Fernando[a]

[a] *University of Colombo, School of Computing, 35, Reid Avenue, Colombo 00700, Sri Lanka.*

Abstract

Municipal Solid Waste(MSW) management enact a significant role in protecting public health and the environment. The main objective of this paper is to explore the utility of using state-of-the-art machine learning and deep learning-based models for predicting future variations in MSW generation for a given geographical region, considering its past waste generation pattern. We consider nine different machine learning and deep-learning models to examine and evaluate their capability in forecasting the daily generated waste amount. In order to have a comprehensive evaluation, we explore the utility of two training and prediction paradigms, a single model approach and a multi-model ensemble approach. Three Sri Lankan datasets from; Boralesgamuwa, Dehiwala, and Moratuwa, and open-source daily waste datasets from the city of Austin and Ballarat, are considered in this study. In addition, we provide an in-depth discussion on important considerations to make when choosing a model for predicting MSW generation.

Keywords: Solid waste generation, Prediction, Deep learning, Machine learning

1. Introduction

Daily human activities are directly and indirectly linked to solid waste generation. Globally, there is around 2.01 billion tons of Municipal Solid Waste (MSW) generated per year, of which at least 33% is not managed in an environmentally safe manner (Salem et al. (2021)).

[*]Corresponding author at: University of Colombo, School of Computing, 35, Reid Avenue, Colombo 00700, Sri Lanka.
E-mail address: dkf@ucsc.cmb.ac.lk (D. Fernando).

[1]Oshan Mudannayake and Disni Rathnayake contributed equally to this work, and should be considered co-first authors.

Poor management and unsafe disposal of solid waste pose a threat to both the environment and human health. The management of MSW faces various challenges related to urbanization, climate change (Raab et al. (2021)) and population growth, which adds complexity and dynamics to the problem. The attentiveness of urban waste requires suitable disposal facilities, infrastructure and transport (Ulgiati and Zucaro (2019)). Further, main tasks and fundamental causes of managing solid waste are lack of waste sorting, poor waste collection mechanisms and absence of public engagement in waste management. Moreover, managing solid waste is a crucial phenomena in both developing and developed countries as it directly impact health and hygiene related issues. Thus, proper tracking and waste collection mechanisms are needed to quantify and predict waste generation.

In particular, the ability to forecast the quantity of waste generated in future would alleviate the burden of managing solid waste, where authorities could factor in future predicted variations in solid waste generation into decision making at the present time. Thereby effectively utilizing resources for waste collection, sorting and other waste management practices. Overall, the ability to accurately estimate future waste generation rates can help motivate gap analysis in existing waste management and pave the way for better strategic planning.

Against this backdrop, in this work we aim to explore the utility of using state-of-the-art machine learning and deep learning-based models for the propose of predicting future variations in solid waste generation for a given geographical region, considering its past waste generation pattern. The work in (Guo et al. (2021); Nguyen et al. (2021)) have already provided a study on comparing machine learning and deep learning models for MSW-prediction, however unlike in the study of Guo et al. (2021) we consider nine different machine learning and deep-learning models from the simplest Linear Regression model to state-of-the-art deep learning models like Transformers (Wolf et al. (2019)) in order to evaluate the suitability of each model with daily solid waste prediction. Additionally, unlike in the study of Nguyen et al. (2021) we consider daily solid waste data from different geographical areas; Sri Lanka (Boralesgamuwa, Dehiwala, Moratuwa), City of Austin in Texas in USA, and City of Ballarat in Australia. Furthermore, we consider the weekly seasonal patterns especially in Ballarat and Austin datasets and explore a multi-model ensemble approach which specifically gives additional focus to the weekly seasonal patterns that exist within the waste generation of each day of the week.

In this work we evaluated the predictive power of nine forecasting models, five machine learning-based models–Linear regression(Stanton (2001)), Auto ARIMA(Matsila and Bokoro (2018)), Light GBM(Ke et al. (2017)), Random Forest(Kane et al. (2014)), Prophet(Taylor and Letham (2018)) and four deep learning-based models–Long short-term memory(LSTM)(Yu et al. (2019)), Temporal Convolutional Network(TCN)(Hewage et al. (2020)), Transformer (Wolf et al. (2019)), and N-Beats(Oreshkin et al. (2019)). We considered these models because they are used in many different time series forecasting studies (Samal et al. (2019),Satrio et al. (2021),Lara-Benítez et al. (2020), Hewage et al. (2020); Lin et al. (2021); Cirstea et al. (2021), Ke et al. (2017), Sagheer and Kotb (2019); Cao et al. (2019); Chimmula and Zhang (2020); Elsworth and Güttel (2020)). We explored two trains of through with respect to training models, a single model approach where a single predictive model is trained to predict solid waste generation similar to a typical time series forecasting task, and a multi-model ensemble approach where seven different models of the same type were trained and used separately for each day of the week. We explored these two options due to the seasonal pattern observed in solid waste generation in Austin and Ballarat datasets, and to identify if there would be any increment in predictive power or decrease in resource utilization in terms of training smaller deep learning-based models for the ensemble.

In this work, we compared the prediction ability of these models mentioned above to forecast daily waste amounts for datasets chosen from three geopolitically diverse locations (i.e., Australia, USA and Sri Lanka). We consider five datasets across these regions, a dataset from Ballarat, which is the third largest city in Victoria, Australia, another dataset from Austin, capital of U.S. state of Texas and also three datasets collected from different Municipal and Urban Authorities in Sri Lanka–Dehiwala, Boralesgamuwa, and Moratuwa. We applied both single-model and multi-model approaches to all nine models we used in this study. The models were evaluated based on the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) values.

In a nutshell our contribution can be summarized as follows. We explore the utility of five machine learning-based predictive models and four state-of-the-art deep learning-based forecasting models for the purpose of predicting solid waste generation. We compare the predictive capability of these models extensively, across five datasets. Sri Lankan datasets are from three local authorities in Colombo; Boralesgamuwa, Dehiwala, and Moratuwa. Additionally, two open source datasets from Ballarat, Australia and Austin, Texas. We explore the utility of two training and prediction paradigms of using these models, a single model approach and a multi-model ensemble approach. Finally we provide extensive evaluation by discussing three important points that may be useful for employing models for predicting solid waste generation–1) the seasonality of data 2) choosing between a single-model or multi-model approach and 3) choosing between a machine learning or deep learning-based model.

Organization: The rest of the paper is organized as follows. Section 2 provides related work of solid waste prediction and time series forecasting using machine learning techniques. In Section 3, we explore the problem statement. Section 4 describes the datasets and data pre-processing steps carried out in this work. In Section 5, we discussed the methodologies of the study. Section 6 presents our evaluation for forecasting solid waste generation in this study including experimental setup, experimental results and discussion. In Section 7, we presented the conclusions of the study.

2. Related Work

This section discusses background details of solid waste prediction and time series forecasting using machine learning techniques.

2.1. Solid waste prediction

MSW generation is becoming one of the crucial issues with the rapid development around the world (Johnson et al. (2017)). Presently, the global waste generation of 3.3 million tonnes per day is becoming unmanageable, and this amount is expected to rise up to 11 million tonnes per day by 2100 (Hoornweg et al. (2013)). Accurate forecasting and prediction of waste are very important because the best strategies for waste management and planning are highly dependent on waste quantification (Tonjes and Greene (2012); Perera and Fernando (2020)).

According to various studies, MSW forecasting methods can be mainly classified into five categories (Abbasi and El Hanandeh (2016)). They are statistical analysis (Sha'Ato et al. (2007)); regression analysis (Wei et al. (2013); material flow analysis (Zhang et al. (2012)); time series analysis (Xu et al. (2013)); and artificial intelligence (Abbasi et al. (2013); Yusoff et al. (2018); Ali and Ahmad (2019); Soni et al. (2019)). However, each and every model or method have their own merits and demerits comparatively. Among them, the artificial intelligence model has been gaining popularity in the forecasting of the generation of MSW due to its high flexibility and proven prediction capabilities compared to the other conventional methods, like regression analysis, time series analysis, (Abbasi and El Hanandeh (2016); Ali and Ahmad (2019); Soni et al. (2019); Dissanayaka and Vasanthapriyan (2019); Juba (2017); Patel and Meka (2013)) etc.

2.2. Machine learning for time series forecasting

Statistical time series modeling is widely used in many prediction and forecasting tasks (Cerqueira et al. (2019); Taylor and Letham (2018)). Autoregressive Moving Average(ARMA) and Autoregressive Integrated Moving Average(ARIMA, which is a generalization of ARMA) models are widely used to fit to the time series data either to better recognize the data or to forecast data in the series. Mwenda et al. (2014) has analyzed, compared and selected the best time series model to forecast solid waste generation for the next years in the city of Arusha in Tanzania among these two models ARMA and ARIMA, and Exponential Smoothing models. Chauhan and Singh (2017) developed a suitable ARIMA model, on the basis of different statistical parameters, in order to forecast healthcare waste quantity from the hospitals of Garhwal region of Uttarakhand, India. Rimaitytė et al. (2012) aimed in selecting and evaluating several methods like regression(Life Cycle Assessment of Integrated Waste Management (LCA-IWM) (Beigl et al. (2003), available at http://www.lca-iwm.net)) and time series modeling methods(ARIMA and Seasonal Exponential Smoothing(SES)) for MSW forecasting in a medium-scaled Eastern European city of Kaunas, Lithuania, with respect to affluence-related and seasonal impacts in the study. For the time series analysis, the combination of ARIMA and SES techniques were found to be the most accurate.

In several recent studies, Artificial Neural Network(ANN) was trained and tested to model waste generation. Azadi and Karimi-Jashni (2016) predicted solid waste generation rates using ANN and Multiple Linear Regression(MLR) in Fars region of Iran. Tiwari et al. (2012) used ANN model to predict industrial solid waste generation and then compared the value with the results obtained from an ANFIS (Adaptive Neuro-Fuzzy Inference System) model. Soni et al. (2019) compared six ANN and ANFIS based models to evaluate and determine the effectiveness in MSW forecasting. According to the results obtained, GA-ANN(i.e. Soni et al. (2019) used genetic algorithm techniques to determine the optimal biases and the weights of the ANN, instead of using the back-propagation optimization.) was found to be the most accurate model among the six models. Juba (2017) analyzed and compared ANN and ARMA to predict the weekly amounts of solid waste generated by individuals in fourteen households in the residential area of Kator in Juba city. According to the literature, many studies have successfully applied ANNs in the time series analysis and forecasting of solid waste.

There are many other machine learning models were used for time series forecasting. LSTM is a common candidate in time series forecasts, in many recent studies (Sagheer and Kotb (2019); Cao et al. (2019); Chimmula and Zhang (2020); Elsworth and Güttel (2020)). The LSTM model is the elegant recurrent neural network variant, which uses the purpose-built LSTM memory cells to represent the long-term dependencies in time series data (Greff et al. (2016)).Niu et al. (2021) aimed at the temporal variation of MSW generation in their study, and a LSTM neural network consisting of LSTM layers and a dropout layer was established and optimized for forecasting MSW generation. To have better illustrate of the accuracy and reliability of the LSTM neural network, an ARIMA model and a conventional ANN model was used to forecast MSW. Results proved that LSTM neural network's superior capability in forecasting solid waste.

Ke et al. (2017) showed in their experiments on multiple public datasets, that LightGBM speeds up the training process of conventional gradient boosting decision tree by up to over 20 times while achieving almost the same accuracy. Some recent studies of time series forecasting used state-of-the-art machine learning models like Facebook Prophet, TCN which have outperformed statistical time series modeling methods like SARIMA, ARIMA, etc.(Samal et al. (2019),Satrio et al. (2021),Lara-Benítez et al. (2020),Hewage et al. (2020); Lin et al. (2021); Cirstea et al. (2021))

3. Problem Statement

In this work, we explore the feasibility of state-of-the-art machine learning and deep learning-based predictive models for the purpose of accurately predicting the daily amount of solid waste generated within a designated geographical authority. The ability to accurately predict solid waste generation multiple days into the future would be helpful for authorities to maximize landfill diversion and better utilize resources to help manage proper disposal and logistics (Zhang et al. (2012); Yermal and Balasubramanian (2017)), thereby effectively reducing waste management costs and increasing operational efficiency.

We model this problem as a uni-variate time series forecasting task, where the objective at time T is to predict the daily amount of solid waste for k days into the future (i.e., $\hat{Y}_T = [\hat{y}_{T+1}, \hat{y}_{T+2}, \ldots, \hat{y}_{T+(k-1)}, \hat{y}_{T+k}]$) based on the amount of daily solid waste generated in the past n days (i.e., $X_T = [x_{T-n}, x_{T-(n-1)}, \ldots, x_{T-1}, x_T]$). In our formulation we denote $\hat{Y}_T$ as the predictions made by the model and $Y_T = [y_{T+1}, y_{T+2}, \ldots, y_{T+(k-1)}, y_{T+k}]$ as the actual solid waste amounts generated between the T^{th} day and $(T+k)^{th}$ day.

4. Datasets and Data preprocessing

In this section, we describe the datasets and data preprocessing steps carried out in this work.

4.1. Datasets

We utilized five different datasets of daily waste collected in different cities (shown in Table A.3). This includes two open source datasets from Ballarat, Australia and Austin, Texas. Ballarat is primarily a residential area, along with significant industrial, commercial and rural areas. It is a city in the Central Highlands of Victoria, Australia. Austin is the most sub urban major metro in Texas, United States with a strong economy. Additionally, we also utilized datasets from three distinct local authorities in Sri Lanka, which is a developing/emerging country with a lower-middle income economy.

Sri Lankan datasets are from three local authorities in Colombo, Sri Lanka. Colombo is the commercial capital and the largest city of Sri Lanka in terms of population. The urban area of Colombo extends well beyond the boundaries of a single local authority, encompassing other municipal and urban councils. In this study, we used the daily collected waste amounts from the Boralesgamuwa Urban Council, Dehiwala Mount Lavinia Municipal Council and Moratuwa Urban Council. We found that the Sri Lankan datasets contained many missing values due to the irregular waste collection and reporting of the data by the waste collection authorities.

4.1.1. Ballarat, Australia dataset

Ballarat, Australia MSW dataset (Ballarat) contains the daily statistics of garbage collection in the City of Ballarat. It includes date(July 2000 - March 2015), number of garbage bins collected, tonnes of waste collected, and area of collection. For our study, we have extracted the tonnes of waste collected per day.

4.1.2. Austin, Texas dataset

Austin, Texas MSW dataset (Austin (2019)) contains waste collection information based on several variables. It includes the daily MSW amount(in tonnes) from January 2003 to July 2021 in city of Austin, Texas. In this study we have extracted only the Report Date and Load Weight variables. Then we calculated the summation of Load Weight per Report Date.

4.1.3. Boralesgamuwa Urban Council dataset

Boralesgamuwa Urban Council's MSW is the first Sri Lankan region dataset consisting of 2541 data points. It contains the daily MSW amount(in kg) from January 2012 to December 2018. MSW amount in the Boralesgamuwa Urban Council area varies from 3741.3kg to 89580kg from the year 2012 to 2018.

4.1.4. Dehiwala Mount Lavinia Municipal Council dataset

Dehiwala Mount Lavinia Municipal is the second Sri Lankan MSW dataset consisting of 2534 data points. It includes the daily MSW amount(in kg) from January 2012 to December 2018. MSW amount in the Dehiwala Mount Lavinia Municipal Council area varies from 3600kg to 39487kg from the year 2012 to 2018.

4.1.5. Moratuwa Urban Council dataset

Moratuwa Urban Council MSW is the third Sri Lankan dataset which contains 1376 data points. It includes the daily MSW amount(in kg) from January 2012 to December 2018. MSW amount varies from 590kg to 51544kg from the year 2014 to 2018 in the Moratuwa Urban Council area.

4.2. Data Preprocessing

This section present the data preprocessing steps. We utilized a machine learning pipeline consisting of three main preprocessing steps. First, we removed the outliers. Then we completed the datasets for a specific period by imputing the missing values. We carry out data imputation by filling in missing values with estimated values based on available data (Barnard and Meng (1999); Acock (2005); Lall (2016)). Finally, we split the data taking 70% as the training data and the remaining 30% as the testing data. In the following section we discuss the data imputation steps in more detail.

4.2.1. Identifying and dealing with the missing values

We found that training a machine learning model on existing data is the best way to impute missing values in this study Jerez et al. (2010); Saad et al. (2020). We opted for a supervised learning approach with lag features to use the available data from the entire dataset to train a model and impute the missing values instead of utilizing a time series model such as ARIMA which could only use the previous values to impute a particular missing value. We believe that the supervised learning approach is much superior to the alternative for datasets like Moratuwa where more than 10% of the values had to be imputed. We selected the XGBoost model Zhang et al. (2019) for this task, where the predictions made by the model for a corresponding missing data instance was used to fill in that sequential position in the time series. The XGBoost algorithm can identify a best way to combine the individual variable context information with those about variables efficiently. We chose a grid search to determine the parameters and the number of lag features of the model that could best fit the existing data. The range of each parameter of the grid search is depicted in Table A.1. We also tried a initial attempt for imputing values through an ARIMA model by treating each dataset as a series of values. However, this made imputing missing values that appeared early in the series difficult as only the data before the missing value could be used to train the model.

After the grid search, we chose two models for each dataset based on the RMSE of the models in order to satisfy the following conditions.

***Model I*:** The model with the lowest RMSE value was chosen as the main imputation model when sufficient data for the lag features preceding the current missing point in the series is present, this would result in better imputation due to the presence of optimal number of lag features.

***Model II*:** There may be cases where the first missing value in the dataset does not have sufficient data preceding it to create the lag features of *Model I*. Then a second model (i.e., *Model II*) was chosen for these scenarios which requires fewer lag features than the number of available values used for *Model I*.

5. Methodology

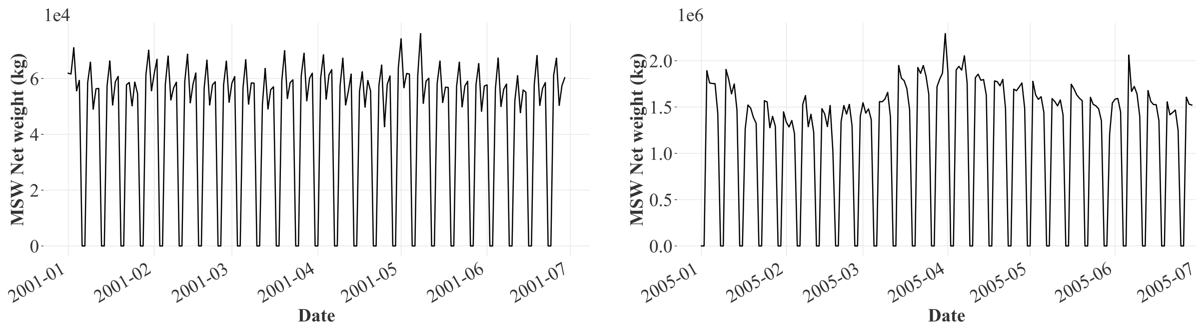
In this work we explore two machine learning-based prediction paradigms, a single-model approach and a multi-model (Huang et al. (2021); Doblas-Reyes et al. (2005)). Additionally, we explore the utility of five machine learning-based time series prediction models and three state-of-the-art deep learning-based time series forecasting models. In this section, we explain our rationale behind exploring these two approaches and an overview of the models utilized in this work.

5.1. Implementation Approaches

We consider two different approaches in utilizing machine learning and deep learning-based models for this predictive task, a single-model approach and a multi-model approach.

Single-model approach: A single predictive model is trained to predict solid waste generation–similar to a typical time series predictive task. The entire dataset is split into two sets, a *train set* and a *test set*, where both sets comprise of a continues stream of data. A given model is trained on the *train set* and performance is compared using the *test set*.

Multi-model ensemble approach: During our data analysis, we found clear weekly seasonal patterns as shown in examples for the Ballarat and Austin shown in Figures 1a and



(a) Ballarat.

(b) Austin.

Figure 1: Weekly pattern of the daily waste data.

1b respectively. As shown in these figures in both these datasets, we observe comparatively lower values for the weekends where as values for weekdays follow a tentative weekly pattern.

This observation prompted us to explore a multi-model approach for this predictive task. Instead of predicting future solid waste generation using a single model, we trained seven models of similar architecture to predict the waste generation of each day in the week. The distinction is that we consider each separate day in a week as a different time series by grouping past solid waste generation values for a given day into its own series.

Here, we first extract the data belonging to each day in the original dataset as a separate series and split each of the series according to the original 70% : 30% ratio. At the end of the prediction task, all the predicted series of different days of the week were combined together to form one single prediction. The main purpose of this approach was to investigate whether it was possible to achieve better performance through modeling each day separately in contrast to using a single model to encompass all the data in a dataset.

5.2. Machine Learning and Deep Learning Models

This section presents a brief overview of all the machine learning and deep learning models explored in this study. In total, we consider five machine learning models (i.e., Linear Regression (Stanton (2001)), Auto ARIMA (Matsila and Bokoro (2018)), Light GBM (Ke et al. (2017)), Random Forest (Kane et al. (2014)) and Prophet (Taylor and Letham (2018)) and four state-of-the-art deep learning-based time series prediction models–LSTM (Yu et al. (2019)),TCN (Hewage et al. (2020)) Transformers (Wolf et al. (2019)) and N-Beats (Oreshkin et al. (2019)) in this work.

5.2.1. Linear Regression

Linear regression takes a linear approach to model the relationship between a dependent variable and one or more independent variables. Linear regression attempts to estimate a straight line that best fits the given data and the equation of that line gives the regression equation. Using one explanatory variable for regression is called simple linear regression, which we use as a baseline model to forecast solid waste generation. Simple linear regression is commonly used in time series forecasting and also in financial analysis. Multiple Linear Regression (MLR) is when several explanatory variables are used for the regression. In this work, we consider a multiple linear regression model as a forecasting model with some of the target series' lag features which are variables in regression that contains data from earlier time steps. We have empirically chosen these lag values after tuning the linear regression model specifically for each dataset.

5.2.2. Auto ARIMA

ARIMA (Ariyo et al. (2014)) (Autoregressive Integrated Moving Average) is a time series forecasting model that operates with three parameters, ARIMA (p, d, q), where; p is the number of autoregressive terms which refers to past values used to predict the next value, d is the number of nonseasonal differences to eliminate the seasonality of time series data, and q is the number of lagged forecast errors in the prediction equation used to define the number of past forecast errors used to predict future values. When training an ARIMA model, statistical techniques are used to generate these p, d, and q values by performing the differencing to eliminate the non-stationary nature of data and plotting the autocorrelation function and the partial autocorrelation function graphs. In Auto ARIMA, the model itself generates the optimal p, d, and q values that would fit the dataset in order to provide the best predictions. In this study Auto ARIMA model is considered and implemented as a thin wrapper around *pmdarima* library.

5.2.3. Light GBM

Light Gradient Boosting Machine (Ke et al. (2017)) also known as Light GBM is a gradient boosting framework that uses tree-based learning algorithms. Light GBM shows leaf-wise tree growth. Since it is based on decision tree algorithms, it divides the tree by leaf with the best fit, while other boosting algorithms divide the tree by depth or level rather than by leaf. The leaf-wise algorithm can reduce more losses than the level-wise algorithm used in other gradient boosting methods and therefore gives much better precision which can rarely be achieved by any of the existing boosting algorithms. In addition, Light GBM is very fast in training. In our work, we consider a LightGBM implementation of the Gradient Boosted Trees algorithm as a univariate forecasting model with lag features.

5.2.4. Random Forest

Random Forest is a type of ensemble machine learning algorithm. It can be used for both classification and regression problems while playing as an extension of bootstrap aggregation of decision trees. However, they can also be used for time series forecasting, although this requires that the time series first be turned into a supervised learning problem. It also requires evaluating the model using walk-forward validation, as evaluating the model using k-fold cross validation would result in optimistically biased results. In this study, we use random forest regression as a forecasting model for prediction solid waste generation. It also uses lag features in order to obtain a forecast. Our Random Forest implementation is a wrapper around the *RandomForestRegressor* in *sklearn* as (Breiman (2001); Geurts et al. (2006)).

5.2.5. Prophet

Prophet is an open source time series forecasting framework based on the idea of using decomposable models, developed by Facebook (Taylor and Letham (2018)). Unlike the previous models, Prophet supports the inclusion of the impact of custom seasonality and holidays. Prophet works with decomposable time series containing three components; trend, seasonality and holidays (Navratil and Kolkova (2019)). We use a wrapper around Prophet implementation in our experiments. We have only added the optional argument of holidays for the datasets of Ballarat and Austin, as the Sri Lankan calendar has not yet been made available in the library.

5.2.6. LSTM

The Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber (1997)) is an improved Recurrent Neural Network (RNN) based model that has shown promising results with respect to learning long and short-term relationships in time series data. LSTMs overcome two major obstacles that RNN's have had to deal with, which are vanishing gradients and exploding gradients. LSTM fixes this by having a gated structure. LSTMs allow RNNs to remember input over a long time period. This is because LSTMs hold information in a memory. In addition to handling long term dependencies, LSTM retain short term information. The LSTM able to read, write and delete information from its memory via this gated mechanism.

5.2.7. Temporal Convolutional Networks

Temporal Convolutional Network (TCN) (Hewage et al. (2020)) is a specialized deep learning architecture designed for time series tasks. TCN is able to extract long-term patterns using dilated causal convolutions and residual blocks, which may also be more computationally efficient. This convolution increases the receptive field of the neural network without resorting to pooling operations, so there is no loss of resolution (Yu and Koltun (2016)). TCN satisfies two main principles: the network outlet has the same length as the input sequence, similar to LSTM networks; and they prevent information leakage from future to the past using causal convolutions (Bai et al. (2018a)). The TCN architecture used in this study is an implementation of a dilated TCN used for forecasting, inspired from the experiments done in the work of Bai et al. (2018b).

5.2.8. Transformers

Transformers are the state-of-the-art deep learning model that is commonly used for natural language processing (NLP) tasks. Transformers can also be used for time series forecasting tasks as well. Transformers are an encoder-decoder architecture. Its main feature is known as a multi-head attention mechanism, which is able to establish intra-dependencies in the input vector and in the output vector known as auto-attention, as well as inter-dependencies between input and output vectors known as encoder-decoder attention. The multi-head attention mechanism is highly parallelizable when used with GPUs. In our study, we used an implementation of Transformers architecture based on the study of Vaswani et al. (2017).

5.2.9. N - Beats

Oreshkin et al. (2019) proposed N-BEATS: Neural Basis Expansion Analysis, a deep neural architecture designed to solve the univariate times series point forecasting problem using deep learning. N-BEATS is known as a pure deep learning architecture in time series forecasting. This model is constructed using backward and forward residual links and a deep stack of fully-connected layers for interpretable time series forecasting. In our study, we used the univariate version of the implementation of N-BEATS architecture, as outlined in the study of Oreshkin et al. (2019).

6. Evaluation

In this section, we present our evaluation for forecasting solid waste generation by machine learning and deep learning models. Additionally, we provide a discussion on important considerations on choosing an appropriate forecasting model.

6.1. Experimental Setup

In our experiments we explored the prediction power of eight forecasting models across two prediction paradigms, a single-model approach and a multi-model approach. All experiments in both single-model and multi-model ensemble approach considers a multi-step prediction of the last 30% values in each dataset. All tests were run on a 12-core Ryzen 3900 machine with a base clock speed of 3.1GHz and 32GB RAM. The models were trained on a Nvidia RTX 2070 Super GPU with 8 GB GDDR6 VRAM. The implementations for the machine learning and deep learning models were carried out using Python and Darts Herzen et al. (2021), a machine learning library for Python with a focus on time series forecasting.

Machine learning models were tuned with an exhaustive grid search heuristic. Deep learning models were manually tuned to the best of our ability. The parameters of the best models for the single-model approach and multi-model approach are in Table A.4 and Table A5, respectively. We haven't included the parameters for the Prophet and Auto ARIMA in these tables as they were chosen automatically by the respective algorithms.

Evaluation Metrics: We used three metrics for the evaluation of the models during this study– RMSE, MAE, and MAPE. Let y_{ij} be the ith test sample for the jth prediction step where $i \in [1, k]$, and $\hat{y}_{ij}$ be the predicted value of y$_{ij}$ and k is the number of test samples. The RMSE, MAE and MAPE are given the by Equations 1, 2 and 3 respectively.

$$RMSE_j = \sqrt{\sum_{i=1}^{k} \frac{(\hat{y}_{ij} - y_{ij})^2}{k}} \tag{1}$$

$$MAE_j = \sum_{i=1}^{k} \frac{|\hat{y}_{ij} - y_{ij}|}{k} \tag{2}$$

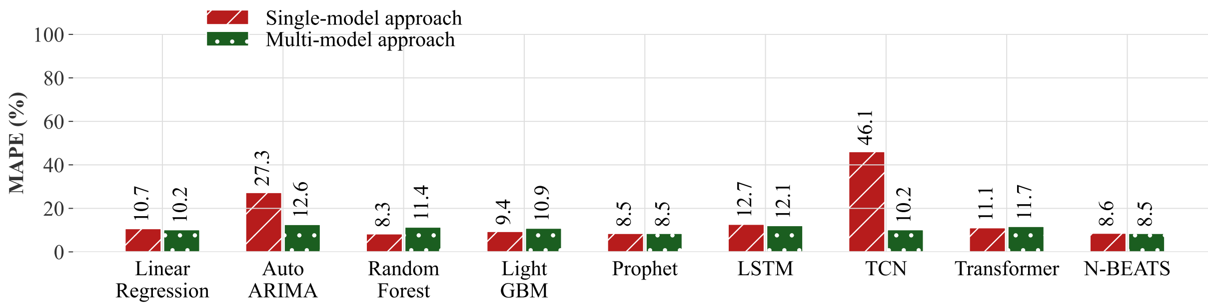
$$MAPE_j = \frac{1}{k} \sum_{i=1}^{k} \frac{|\hat{y}_{ij} - y_{ij}|}{y_{ij}} \tag{3}$$

6.2. Experimental Results

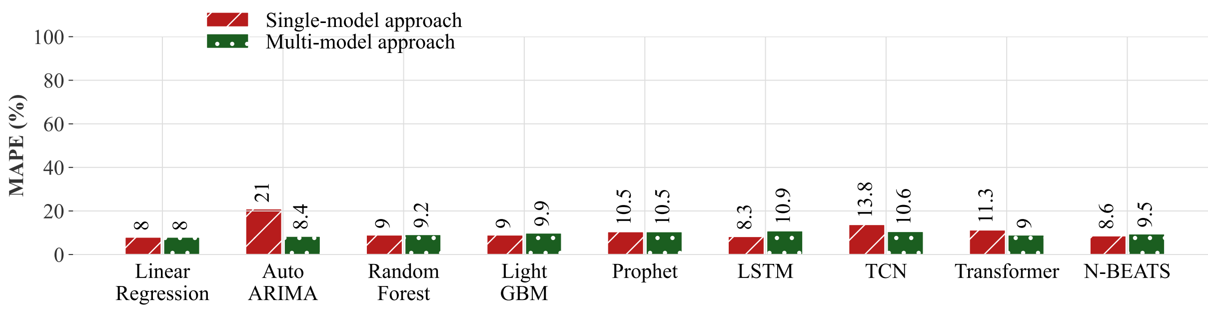
This section will describe the results of all the experiments conducted during our study. Table A.6, Table A.7, Table A.8, Table A.9 and Table A.10 contains the performance of each model on Ballarat, Austin and Sri Lankan datasets respectively. The tables contain the RMSE, MAE and MAPE values of the best model in each experiment. Figures 2a, 2b, and 2c correspond to the MAPE values for the best models for Ballarat, Austin and Sri Lankan datasets respectively.

Figure 2a and Table A.6 shows the results for Ballarat dataset for machine learning and deep learning models trained using the single-model approach and the multi-model ensemble approach, respectively. As shown in Figure X, the Random Forest model and the N-BEATS model show the best performance for single-model and multi-model training approaches for the Ballarat dataset as 8.3% and 8.47%. Overall, the single-model Random Forest is the most successful model in predicting the waste generation patterns in the Ballarat dataset with an average improvement of 3.85% against the rest of the machine learning models and an average improvement of 6.82% against the deep learning-based models.

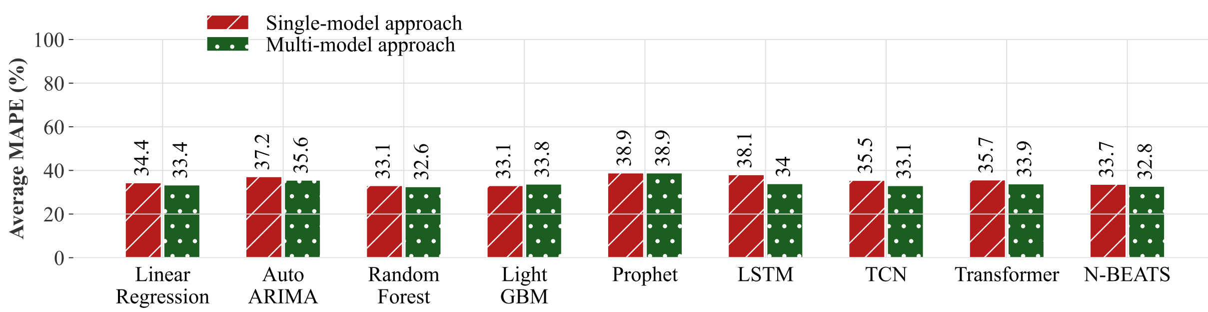
Prophet also shows strong results for the Ballarat dataset (i.e., 8.47% in MAPE). For the Ballarat dataset, Auto ARIMA and TCN seems to have shown a significant improvement of over 15% reduction of MAPE with the multi-model approach. Performance of single-model



(a) Ballarat dataset



(b) Austin dataset



(c) Sri Lankan datasets (average)

Figure 2: MAPE of the chosen models with different datasets.

and multi-model approaches average around MAPE of 15.85% and 10.67% across all model types. Therefore the multi-model approach has worked better for the Ballarat dataset.

We see similar variations in MAPE results for Austin dataset (shown in Table A.7 and Figure 2b) where Linear Regression has obtained the best performance for the Austin dataset in all scenarios. The error for the results in this dataset are comparable to the Ballarat dataset. Here, on average the best Linear regression model outperforms all other machine learning models by 4.37% and other deep learning models by 2.46% when considering the single model approach. As for the multi-model approach Linear regression is 1.49% greater than machine learning models and 1.96% greater than deep learning models. Auto ARIMA's results have significantly improved by over 12% with the multi-model approach. LSTM and N-BEATS has shown a decrease in performance in multi-model approach in comparison to the single-model training. The single-model approach shows an average MAPE of 11.07% while the multi-model approach performed slightly better at an average MAPE of 9.56%.

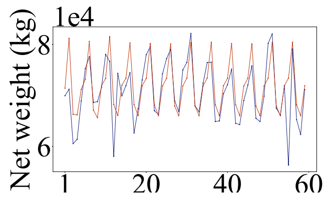
Random Forest model shows the best performance for all Sri Lankan datasets with an average of 32.86% in MAPE except for the single-model training mode in Boralesgamuwa dataset. Light GBM shows the best single-model performance for the Boralesgamuwa dataset (i.e., 28.84% in MAPE). It is apparent that the Random Forest has been most successful in capturing the patterns in the Sri Lankan datasets, which didn't have visible seasonal patterns.

In most cases, training according to the multi-model approach shows and improvement in predictive performance. On average, deep learning models haven't shown a significant improvement in prediction for the chosen datasets over machine learning models. Random forests have been able to achieve comparable results as the best deep learning model after a grid search.

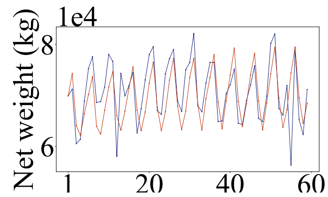
6.3. Discussion

In this section we further discuss three points to consider when choosing a predictive model for forecasting solid waste generation, namely 1) seasonality in data 2) the modeling approach: single-model vs multi-model and 3) choosing between a machine learning vs a deep learning model.

Effect of seasonality in data: Overall we observe a lower average predictive error for Ballarat and Austin of 13.26% and 10.31% across all models, irrespective of the approach used (i.e., single or multi-model). In contrast there is higher error for the Sri Lankan regional datasets (i.e., average of 25.64%). We attribute this disparity in predictive performance to the seasonal patterns observed in Austin and Ballarat.

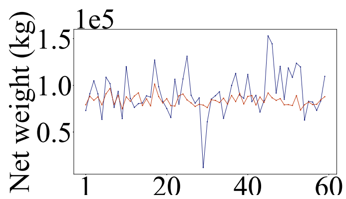


(a) Best ML model

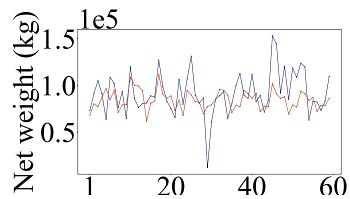


(b) Best DL model

Figure 3: Actual vs Predictions for Ballarat.



(a) Best ML model



(b) Best DL model

Figure 4: Actual vs Predictions for Moratuwa-Sri Lanka

Figures 3 shows actual vs predictions made for Ballarat dataset for both best machine learning and deep learning models. Similarly, Figure 4 shows the actual vs predictions made for the Moratuwa-Sri Lankan dataset. Both the figures show the predictions for best performing machine learning and deep learning model for each dataset considering both single-model or the multi-model approach. Based on Figure 3 it is clear that predictions made by the models tend to capture variations more accurately in the presence of seasonality. In contrast as shown in Figure 4, in a dataset that doesn't have seasonality, both the machine learning (Fig 3.a) and deep learning (Fig 3.b) models have a harder time in learning temporal patterns. This is a common trend we see in all other non-seasonal data (i.e., Boralesgamuwa and Dehiwala-Sri Lankan data).

Single-model or multi-model approach: The seasonality in the Ballarat and the Austin datasets prompted us to explore the utility of exploring a single-model and a multi-model approach, where the multi-model approach specifically gives additional focus to the weekly seasonal patterns that exist within the waste generation of each day of the week. Our initial assumption was that the multi-model approach would perform better against data with clear seasonal patterns. Table 1 shows the average RMSE, MAE and MAPE error considering all the datasets against a single-model and multi-model approach. Overall, we observe that the multi-model approach reported slightly better performance than the single-model approach (Table 1). However, while the predictive performance itself shows slight differences, the biggest difference in choosing to use either of these approaches comes with the effort and resources in training these models.

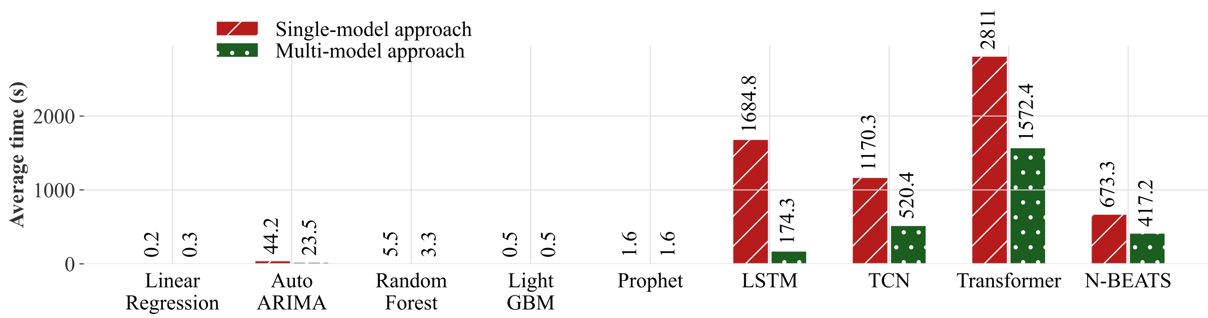


Figure 5: Average training time for the models across all the datasets.

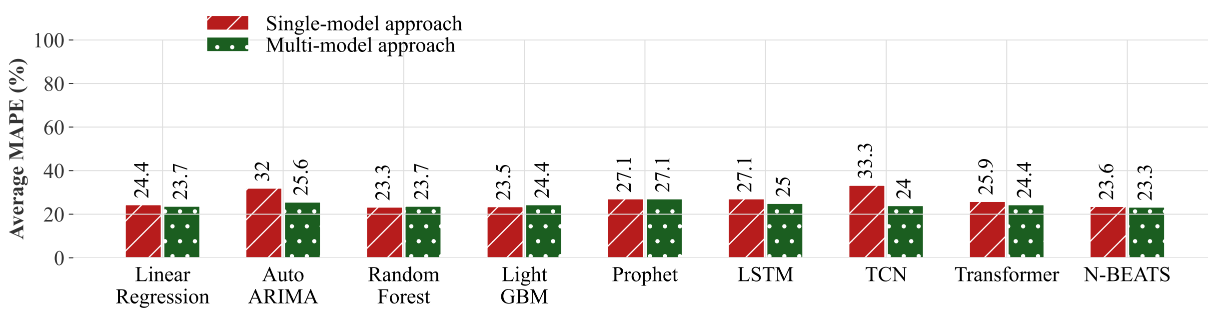


Figure 6: Average MAPE of the models across all the datasets.

Figure 5 shows the average training time for each of the best performing models across all datasets. The average training time of single-model and multi-model approaches were was 710.15 seconds and 301.5 seconds. And the average MAPE values varied as 26.7% and 24.59%. Therefore the multi-model approach has taken less time to train the models with better performance. It takes significantly less time to train the deep learning models using the multi-model approach than the single-model approach. This is because a set of smaller models trained according to the multi-model approach were able to capture the pattern better than one large model that was trained using the conventional single-model approach.

Table 1: Average performance of single-model vs multi-model training methods

Training method	RMSE	MAE	MAPE	Training time (s)	Predicting time (s)
Single-model	71747.34	56118.70	26.70%	710.15	1.82
Multi-model	62217.67	48348.99	24.59%	301.50	24.13

Table 2: Average performance of machine learning vs deep learning models

Training paradigm	RMSE	MAE	MAPE	Training time (s)	Predicting time (s)
Machine learning	67574.97	52457.47	25.44%	23.23	15.23
Deep learning	66051.49	51882.44	25.96%	1264.18	9.44

This directly benefits in reducing the computational cost of the training phase. Therefore, for situations with higher constraints in time, we believe a multi-model approach may be more suitable as it obtains similar predictive performance while requiring less training time, specifically for instances where deep learning models are used.

Machine learning or deep learning: Our experiments considered both machine learning and deep learning models in order to determine comparatively what types of models may be more suitable. Figure 6 shows the average MAPE of each of the models considering all the datasets. While the deep learning models do shows a slightly better reduction in error, it is marginal accounting to less than a decrease of 5% in MAPE. Meaning that there seems to be less utility in using deep neural network architectures for this forecasting task (i.e., summarized in Table 2).

In addition, the deep learning models take a significantly longer time in training than the machine learning models. As shown in Table 2 on average, the deep learning models have taken 54× times more time to train than the machine learning models. All the machine learning models were trained on the CPU, where as the deep learning models were trained on a dedicated GPU. The machine learning models were trained with relatively less computational cost than the deep learning models, adding to the utility of using machine learning-based models.

We consider datasets spanning different time periods. The longest spanning dataset

Austin contains the records of 14 years. The smallest dataset is the Moratuwa dataset which spans only for 3 years. For both these datasets, the performance of the best machine learning model and the best deep learning model varies by less than 2% MAPE. Therefore we can conclude that the length of the dataset is not a significant factor contributing to the design choice of selecting a machine learning or deep learning model. It seems like both model types work equally well for long-term (i.e., more than 10 years) or shorter-term (i.e., around 2-3 years) time periods of data. While it can be argued that a deep learning model might be able to predict the waste data generation of a much larger dataset with better accuracy, it may not be realistic to assume that the same waste generation patterns could exist for a time period beyond 15 years due changing in urban populations, waste management policies that could be implemented within such a long period of time.

We have used a grid search method to train the machine learning models. The deep learning models were trained manually with extreme care. Therefore the machine learning models have presented a greater advantage in being able to be trained in environments with less supervision and skilled personal than deep learning models.

While machine learning models on average have slightly outperformed the deep learning models, deep learning models have outperformed machine learning models on several occasions. However the performance improvement of the deep learning models have come at a much greater computational cost. Therefore it is apparent that the machine learning models are specifically well suited for developing regions such as Sri Lanka where there exist limitations in modeling waste data such as lack of computational power and skilled personals.

7. Conclusion

In this paper, we investigated how well the machine learning and the state-of-the-art deep learning models are able to forecast daily waste amount of five different geographical areas. We compared the performance of nine different machine learning and deep learning models across all the five datasets. In our study, we observe comparable results in both machine learning and deep learning models, while machine learning models on average have slightly outperformed the deep learning models. However deep learning models have taken more computational power during the training phase. Therefore we can conclude that machine learning models are sufficient for forecasting municipal solid waste in a given geographical location. Furthermore the training time has been reduced by using the multi-model training paradigm. Also results shows that It also contributed to a slight increase of performance as well.

References

Abbasi, M., Abduli, M., Omidvar, B., Baghvand, A., 2013. Forecasting municipal solid waste generation by hybrid support vector machine and partial least square model .

Abbasi, M., El Hanandeh, A., 2016. Forecasting municipal solid waste generation using artificial intelligence modelling approaches. Waste management 56, 13–22.

Acock, A.C., 2005. Working with missing values. Journal of Marriage and family 67, 1012–1028.

Ali, S.A., Ahmad, A., 2019. Forecasting msw generation using artificial neural network time series model: a study from metropolitan city. SN Applied Sciences 1, 1–16.

Ariyo, A.A., Adewumi, A.O., Ayo, C.K., 2014. Stock price prediction using the arima model, in: 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, IEEE. pp. 106–112.

Austin, T.d., 2019. Austin, texas dataset. `https://data.world/cityofaustin/mbnu-4wq9`.

Azadi, S., Karimi-Jashni, A., 2016. Verifying the performance of artificial neural network and multiple linear regression in predicting the mean seasonal municipal solid waste generation rate: A case study of fars province, iran. Waste management 48, 14–23.

Bai, S., Kolter, J.Z., Koltun, V., 2018a. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv preprint arXiv:1803.01271 .

Bai, S., Kolter, J.Z., Koltun, V., 2018b. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv:1803.01271 .

Ballarat, A.d., . Ballarat, australia dataset. `https://data.gov.au/data/dataset/ballarat-garbage-co`

Barnard, J., Meng, X.L., 1999. Applications of multiple imputation in medical studies: from aids to nhanes. Statistical methods in medical research 8, 17–36.

Breiman, L., 2001. Random forests. Machine learning 45, 5–32.

Cao, J., Li, Z., Li, J., 2019. Financial time series forecasting model based on ceemdan and lstm. Physica A: Statistical Mechanics and its Applications 519, 127–139.

Cerqueira, V., Torgo, L., Soares, C., 2019. Machine learning vs statistical methods for time series forecasting: Size matters. arXiv preprint arXiv:1909.13316 .

Chauhan, A., Singh, A., 2017. An arima model for the forecasting of healthcare waste generation in the garhwal region of uttarakhand, india. International Journal of Services Operations and Informatics 8, 352–366.

Chimmula, V.K.R., Zhang, L., 2020. Time series forecasting of covid-19 transmission in canada using lstm networks. Chaos, Solitons & Fractals 135, 109864.

Cirstea, R.G., Kieu, T., Guo, C., Yang, B., Pan, S.J., 2021. Enhancenet: Plugin neural networks for enhancing correlated time series forecasting, in: 2021 IEEE 37th International Conference on Data Engineering (ICDE), IEEE. pp. 1739–1750.

Dissanayaka, D., Vasanthapriyan, S., 2019. Forecast municipal solid waste generation in sri lanka, in: 2019 International Conference on Advancements in Computing (ICAC), IEEE. pp. 210–215.

Doblas-Reyes, F.J., Hagedorn, R., Palmer, T., 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting — II. Calibration and combination. Tellus A: Dynamic Meteorology and Oceanography 57, 234–252. doi:10.3402/tellusa.v57i3.14658.

Elsworth, S., Güttel, S., 2020. Time series forecasting using lstm networks: a symbolic approach. arXiv preprint arXiv:2003.05672 .

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Machine learning 63, 3–42.

Greff, K., Srivastava, R.K., Koutník, J., Steunebrink, B.R., Schmidhuber, J., 2016. Lstm: A search space odyssey. IEEE transactions on neural networks and learning systems 28, 2222–2232.

Guo, H.n., Wu, S.b., Tian, Y.j., Zhang, J., Liu, H.t., 2021. Application of machine learning methods for the prediction of organic solid waste treatment and recycling processes: A review. Bioresource technology 319, 124114.

Herzen, J., Lässig, F., Piazzetta, S.G., Neuer, T., Tafti, L., Raille, G., Van Pottelbergh, T., Pasieka, M., Skrodzki, A., Huguenin, N., et al., 2021. Darts: User-friendly modern machine learning for time series. arXiv preprint arXiv:2110.03224 .

Hewage, P., Behera, A., Trovati, M., Pereira, E., Ghahremani, M., Palmieri, F., Liu, Y., 2020. Temporal convolutional neural (tcn) network for an effective weather forecasting using time-series data from the local weather station. Soft Computing 24, 16453–16482.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural computation 9, 1735–1780.

Hoornweg, D., Bhada-Tata, P., Kennedy, C., 2013. Environment: Waste production must peak this century. Nature News 502, 615.

Huang, Y., Du, C., Xue, Z., Chen, X., Zhao, H., Huang, L., 2021. What Makes Multi-modal Learning Better than Single (Provably) , 1–13URL: `http://arxiv.org/abs/2106.04538`, `arXiv:2106.04538`.

Jerez, J.M., Molina, I., García-Laencina, P.J., Alba, E., Ribelles, N., Martín, M., Franco, L., 2010. Missing data imputation using statistical and machine learning methods in a real breast cancer problem. Artificial intelligence in medicine 50, 105–115.

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.

Juba, P., 2017. Forecasting solid waste generation in juba town, south sudan using artificial neural networks (anns) and autoregressive moving averages (arma). Journal of Environment and Waste Management 4, 211–223.

Kane, M.J., Price, N., Scotch, M., Rabinowitz, P., 2014. Comparison of arima and random forest time series models for prediction of avian influenza h5n1 outbreaks. BMC bioinformatics 15, 1–9.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.Y., 2017. Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems 30, 3146–3154.

Lall, R., 2016. How multiple imputation makes a difference. Political Analysis 24, 414–433.

Lara-Benítez, P., Carranza-García, M., Luna-Romera, J.M., Riquelme, J.C., 2020. Temporal convolutional networks applied to energy-related time series forecasting. applied sciences 10, 2322.

Lin, W.H., Wang, P., Chao, K.M., Lin, H.C., Yang, Z.Y., Lai, Y.H., 2021. Wind power forecasting with deep learning networks: Time-series forecasting. Applied Sciences 11, 10335.

Matsila, H., Bokoro, P., 2018. Load forecasting using statistical time series model in a medium voltage distribution network, in: IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society, IEEE. pp. 4974–4979.

Mwenda, A., Kuznetsov, D., Mirau, S., 2014. Time series forecasting of solid waste generation in arusha city-tanzania. Mathematical Theory and Modeling 4, 29–39.

Navratil, M., Kolkova, A., 2019. Decomposition and forecasting time series in the business economy using prophet forecasting model. Central European Business Review 8, 26.

Nguyen, X.C., Nguyen, T.T.H., La, D.D., Kumar, G., Rene, E.R., Nguyen, D.D., Chang, S.W., Chung, W.J., Nguyen, X.H., Nguyen, V.K., 2021. Development of machine learning-based models to forecast solid waste generation in residential areas: A case study from vietnam. Resources, Conservation and Recycling 167, 105381.

Niu, D., Wu, F., Dai, S., He, S., Wu, B., 2021. Detection of long-term effect in forecasting municipal solid waste using a long short-term memory neural network. Journal of Cleaner Production 290, 125187.

Oreshkin, B.N., Carpov, D., Chapados, N., Bengio, Y., 2019. N-beats: Neural basis expansion analysis for interpretable time series forecasting. arXiv preprint arXiv:1905.10437 .

Patel, V., Meka, S., 2013. Forecasting of municipal solid waste generation for medium scale towns located in the state of gujarat, india. International Journal of Innovative Research and Science 2, 4707–4716.

Perera, C., Fernando, M., 2020. Comparison of multiple linear regression and artificial neural network models for the prediction of solid waste generation in sri lanka. Trans. Mach. Learn. Data Min. 13, 3–25.

Raab, K., Tolotti, G., Wagner, R., 2021. Challenges in solid waste management: Insights into the disposal behavior of suburban consumers in guatemala city .

Rimaitytė, I., Ruzgas, T., Denafas, G., Račys, V., Martuzevicius, D., 2012. Application and evaluation of forecasting methods for municipal solid waste generation in an eastern-european city. Waste Management & Research 30, 89–98.

Saad, M., Chaudhary, M., Karray, F., Gaudet, V., 2020. Machine learning based approaches for imputation in time series data and their impact on forecasting, in: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE. pp. 2621–2627.

Sagheer, A., Kotb, M., 2019. Time series forecasting of petroleum production using deep lstm recurrent networks. Neurocomputing 323, 203–213.

Salem, H.S., Yihdego, Y., Pudza, M.Y., 2021. Scientific and technical perspectives of waste-to-energy conversion. Journal of Nature Science and Sustainable Technology 15, 193–217.

Samal, K.K.R., Babu, K.S., Das, S.K., Acharaya, A., 2019. Time series based air pollution forecasting using sarima and prophet model, in: proceedings of the 2019 international conference on information technology and computer communications, pp. 80–85.

Satrio, C.B.A., Darmawan, W., Nadia, B.U., Hanafiah, N., 2021. Time series analysis and forecasting of coronavirus disease in indonesia using arima model and prophet. Procedia Computer Science 179, 524–532.

Sha'Ato, R., Aboho, S., Oketunde, F., Eneji, I., Unazi, G., Agwa, S., 2007. Survey of solid waste generation and composition in a rapidly growing urban area in central nigeria. Waste management 27, 352–358.

Soni, U., Roy, A., Verma, A., Jain, V., 2019. Forecasting municipal solid waste generation using artificial intelligence models—a case study in india. SN Applied Sciences 1, 162.

Stanton, J.M., 2001. Galton, pearson, and the peas: A brief history of linear regression for statistics instructors. Journal of Statistics Education 9.

Taylor, S.J., Letham, B., 2018. Forecasting at scale. The American Statistician 72, 37–45.

Tiwari, M.K., Bajpai, S., Dewangan, U., 2012. Prediction of industrial solid waste with anfis model and its comparison with ann model-a case study of durg-bhilai twin city india. International Journal of Engineering and Innovative Technology (IJEIT) 6, 192–201.

Tonjes, D.J., Greene, K.L., 2012. A review of national municipal solid waste generation assessments in the usa. Waste Management & Research 30, 758–771.

Ulgiati, S., Zucaro, A., 2019. Challenges in urban metabolism: sustainability and well-being in cities. Frontiers in Sustainable Cities 1, 1.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. Advances in neural information processing systems 30.

Wei, Y., Xue, Y., Yin, J., Ni, W., 2013. Prediction of municipal solid waste generation in china by multiple linear regression method. International Journal of Computers and Applications 35, 136–140.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., et al., 2019. Huggingface's transformers: State-of-the-art natural language processing. arXiv preprint arXiv:1910.03771 .

Xu, L., Gao, P., Cui, S., Liu, C., 2013. A hybrid procedure for msw generation forecasting at multiple time scales in xiamen city, china. Waste management 33, 1324–1331.

Yermal, L., Balasubramanian, P., 2017. Application of auto arima model for forecasting returns on minute wise amalgamated data in nse, in: 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), IEEE. pp. 1–5.

Yu, F., Koltun, V., 2016. Multi-scale context aggregation by dilated convolutions international conference on learning representations (iclr) 2016.

Yu, Y., Si, X., Hu, C., Zhang, J., 2019. A review of recurrent neural networks: Lstm cells and network architectures. Neural computation 31, 1235–1270.

Yusoff, S.H., Din, U.N.K.A., Mansor, H., Midi, N.S., Zaini, S.A., 2018. Neural network prediction for efficient waste management in malaysia. Indonesian Journal of Electrical Engineering and Computer Science (IJEECS) 12, 738–747.

Zhang, L., Yuan, Z., Bi, J., Huang, L., 2012. Estimating future generation of obsolete household appliances in china. Waste management & research 30, 1160–1168.

Zhang, X., Yan, C., Gao, C., Malin, B., Chen, Y., 2019. Xgboost imputation for time series data, in: 2019 IEEE International Conference on Healthcare Informatics (ICHI), IEEE. pp. 1–3.