Received: February 24, 2024. Revised: May 17, 2024.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

An Enhanced GRU model with Optimized Update Gate: A Novel Approach in Municipal Solid Waste Prediction

Tuba Batool[1] **Rozaida Ghazali**[1]* **Nureize Binti Arbaiy**[1] **Lokman Hakim Ismail**[1]
Irfan Javid[2]

[1]*Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Malaysia*
[2]*Department of Computer Science and Information Technology, University of Poonch, Rawalakot, AJK, Pakistan*
* Corresponding author's Email: hi210012@student.uthm.edu.my

Abstract: The management of Municipal Solid Waste (MSW) establishes a decisive facet in the development of sustainable and vigorous societies that requires an adequate framework in order to prevent a risk of environmental pollution. Therefore, it's imperative to establish a robust framework to handle the generated MSW. In this regard, early predictions of waste generation with high accuracy emerges as a pivotal factor in serving the municipal authorities to formulate an effective MSW management system. Various researchers have significantly contributed to the field of early waste predictions by developing different deep learning models to attain the highly accurate predicted results. However, it is important to note that each model has its strengths and limitations. This study has predominantly focused on the enhancement of the Gated Recurrent Unit (GRU) model. The primary limitation in the standard GRU model lies in both of its gates (reset and update) processing the identical data. This redundancy has contributed to a prolonged training time and a diminished convergence rate. Therefore, this study has proposed an enhanced GRU model with optimized update gate (EGRU-OU) to address the redundancy issue that lies between the two gates. The EGRU-OU model will provide the filtered data specifically to the update gate which is instrumental in significantly reducing the redundant information between the two gates. Moreover, this study has employed two different datasets to conduct a comprehensive analysis of the results. One dataset has been collected for 16 different developed countries, whereas, the other dataset has been collected for Multan City, Pakistan. These datasets have been segmented into three subdivisions: training data, constituting 70% of the dataset for model training; testing data, comprising 15% for model evaluation; and validation data, representing the remaining 15% for additional model validation. In addition, The EGRU-OU model has been compared with other benchmark models including standard GRU model, Long Short Term Memory (LSTM) and Artificial Neural Network (ANN) model based on three distinct error metrics: Mean Absolute Error (*MAE*), Root Mean Square Error (*RMSE*), and the coefficient of determination (R^2). The outcomes have clearly demonstrated a superior performance of the EGRU-OU approach as compared to other models with the least error matrices values of *MAE* being 0.036 and *RMSE* being 0.0684.

Keywords: Municipal solid waste, ANN, GRU, LSTM.

1. Introduction

Municipal solid waste (MSW) is referred to the leftover which is originated from the homes, businesses and various non-industrial sources which encompasses the variety of items utilized and castoff by individuals on daily basis [1]. As stated by a report [2] from Global Market Insights, the magnitude of the MSW market surged to $117 billion in 2022 and is anticipated to experience a yearly growth rate of around 3.3% from 2023 to 2032. A substantial surge in MSW generation has now become a paramount factor contributing to the degradation of quality of life, therefore, it is essential to develop a robust framework to facilitate the municipal authorities in MSW management. The prediction results with high accuracy of MSW can certainly play a pivotal role in the development of an effective system [3]. Additionally, Different studies [4, 5] have been conducted for the projection of MSW quantity by leveraging pertinent data gathered from past years. Direct projection of MSW prediction is quite challenging and is contingent upon the specific socio-demographic, and economic factors inherent to a particular region. [6]. However, the developing countries grapple with uncertain and insufficient data, [5] necessitating the implementation of appropriate approaches to predict MSW generation. Different researchers have used different models, like regression techniques and time series methodology [7, 8, 9]. A study [9] has revealed that the time series methodologies have exhibited superior performance as compared to the regression methods. The accuracy of the time series models have been augmented by using a hierarchical artificial neural network (ANN) for the prediction of waste generation [10]. Moreover, to enhance the performance of ANN model, the researchers have implemented different deep learning methodologies [11, 12] comprising gated recurrent unit (GRU) models and long short-term memory (LSTM).

Kyunghyun Cho et al. introduced the GRU model in 2014 [13] and it has shown exceptional performance in the MSW generation prediction [14]. However, the standard GRU model still has some shortcomings. In the conventional GRU model, there are two different gates; one is a reset gate and the other is update gate. The primary purpose of a reset gate is to regulate the flow of less critical information towards the subsequent state, while the update gate preserves the most crucial information for the ensuing hidden state. The primary limitation of the conventional GRU model is the data redundancy that lies between its two gates as the same input is taken by both gates. However, there is no necessity for the update gate to handle the data that has already been processed and ignored by the reset gate. The redundant information leads to the low convergence rate and high training time. This issue needs to be resolved aptly to improve the model’s performance.

The major purpose of this study is to propose EGRU-OU model to address the data redundancy issue that lies between the two gates in the standard GRU model and compare the results with other baseline models which includes standard GRU, LSTM and ANN model based on the prediction accuracy of MSW generation. The proposed model aims to enhance performance by achieving higher accuracy along with reduced training time through the utilization of filtered data.

2. Literature review

2.1 Deep learning approaches in MSW prediction

Deep learning models, which differ from machine learning models in their architecture, are characterized by neural networks with three layers or more and do not necessitate many complex data pre-processing. [15]. However, different deep learning models demonstrate complex details with different neural networks to address different problems [16]. Deep learning models commonly include Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). [16]. Convolutional Neural Networks (CNNs), often referred to as ConvNets, are layered neural networks predominantly employed for tasks such as object detection and image processing. And, RNNs are typically considered to perform exceptionally well when dealing with natural language processing and time series data [17]. The two common RNNs include Long Short Term Memory (LSTM), and Gated Recurrent Unit (GRU).

Moreover, many researchers have conducted a range of studies employing different deep learning approaches to forecast the generation rate of MSW, some of these studies have been elucidated in this section. A study has been conducted in Mainland, China where ANNs were employed to predict the MSW generation rate and the results have shown the outstanding performance of the model with a high regression value being 0.96 and low Root Mean Square Error (*RMSE*) value being 17.6 [18]. In addition, a study was carried out to forecast the generation of solid waste in Sousse, Tunisia which compared Bidirectional Long Short-Term Memory (BLSTM) with other alternative models, including LSTM. The outcomes have revealed that BLSTM outperformed LSTM with a mean squared error (*MSE*) value being 0.15 and a mean absolute error (*MAE*) value being 0.21 [19]. Moreover, recurrent neural networks with long short-term memory (RNN-LSTM) were used in a research for forecasting the generation rate of MSW and the outcomes have illustrated the acceptable performance of the model, as reflected by the regression values ranging from 0.70 to 0.86 [20]. A further study [6] aimed to compare LSTM and GRU, integrating them with gray relational analysis (GRA) based on MSW prediction and the results have revealed the superior performance of GRA-GRU with a low *MAE* value being 18.801 and a low *RMSE* value being 21.830.

Furthermore, a study has been carried out to predict the urban waste generation by employing the

Table 1. A summary of deep learning approaches for predicting solid waste

Model	Data	MAE	MSE	RMSE	R^2	Reference
ANN	National Bureau of Statistics of China (NBSC, 1980–2018) Yearly data	-	-	17.6	0.968	[18]
BLSTM	734 monthly records collected over a one-year period, for Tunisia from historical records.	0.21	0.15	-	-	[19]
RNN-LSTM	8.8 years (January 2013 – October 2021) of daily mixed waste disposal data, obtained from the landfill historical records at the landfill scale.	-	-	72 - 95	0.70 - 0.86	[20]
GRA-GRU	Historical yearly data on MSW generation from 1979 to 2019 (BMBS, 1980–2020) in Beijing	18.801	-	21.830	-	[21]
RNGRU	2000 records collected from Organization for Economic Co-operation and Development (OECD)	0.014	0.0010	-	-	[22]
Multisite-LSTM	Dataset collected from the municipality of Herning, Denmark, (weekly observations from 1,000 households in the period between 2011 and 2018).	0.41	-	0.50	-	[23]

Regularized Noise-based Gated Recurrent Unit (RNGRU) and the findings have illustrated the superior performance of RNGRU with a low *MAE* value of 0.0147 and a low *MSE* value of 0.0010 [21]. A different research [22] has contributed in predicting the generation rate of MSW through the utilization of multisite Long Short-Term Memory (LSTM) and the outcomes have presented a satisfactory performance of multisite LSTM, as in a low MAE value of 0.41 and an RMSE value of 0.50. A summary of deep learning models in solid waste predictions has been given in Table 1.

In addition, we could observe that GRU model has performed better than LSTM and ANN model in solid waste prediction. Therefore, this study has opted the GRU model as a baseline for further enhancement, aiming to achieve superior accuracy while reducing training time.

3. Methodology

This section demarcates the comprehensive methodology that has been employed in this study, encompassing various aspects such as data collection sources, the development of models including GRU, LSTM, ANN, and EGRU-OU, and the evaluation metrics employed for analysing the performance of these models. The inclusive workflow of this study is illustrated in Fig. 1.

3.1 Data collection

Two distinct datasets have been utilized in this research to have a thorough analysis of the results.

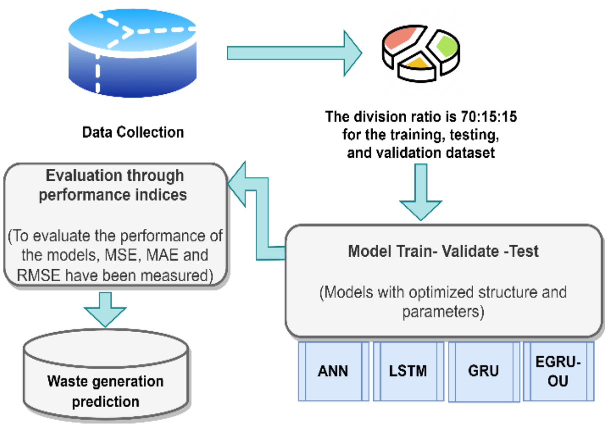


Figure. 1 General workflow of this study

One data has been collected for 16 different developed countries through the Organization for Economic Co-operation and Development (OECD) via their website link https://data.oecd.org/.

The dataset contains 15 different variables (as shown in Table 2) for the past 25 years (1996 - 2020) with MSW as target variable. Other dataset has been collected from the Multan Waste Management Company (MWMC) for Multan City, Pakistan. This data has been collected month wise for five years (2017-2022). These datasets have been segmented into three subdivisions: training data, constituting 70% of the dataset for model training; testing data, comprising 15% for model evaluation; and validation data, representing the remaining 15% for additional model validation.

Table 2. Summarized dataset of 16 different countries for the year 2019

Sr#	Country	Year	Population (MLN_PER)	Upper Secondary (%)	GWH	PC_ WKGPOP	Disposable income (USD_CAP)	Municipal Waste (K Tons)
1	Australia	2019	8.877637	51.79036	71101.13	73.525	38607.52	5220.28
2	Belgium	2019	11.46202	38.01605	89913.4	65.3	36617.33	4799.862
3	Chile	2019	19.10722	40.96816	82100.15	64.09287	18216.63	8207.56
4	Denmark	2019	5.814461	41.23565	28689.56	75	34951.08	4907.466
5	Finland	2019	5.521605	43.33691	66053	72	34864.08	3122.705
6	Poland	2019	38.38648	60.61755	151363	67.525	23519.16	12752.778
7	France	2019	67.35605	42.5336	547042.8	66.375	36033.75	37397.05
8	Germany	2019	83.09296	54.85022	575864	75.65	41021.82	50611.789
9	United Kingdom	2019	66.79681	32.25974	310171.3	76.175	34697.54	30677.985
10	Italy	2019	59.72908	42.58225	283950.1	59.05	31656.89	30023
11	Japan	2019	126.1669	43.38796	998494.2	78.11421	30908.73	42737
12	Korea	2019	51.76482	38.64747	559099.5	66.82268	25964.24	21155.91
13	Norway	2019	5.347893	37.20805	134665.8	75.3	40062.63	4150.795
14	Spain	2019	47.10536	23.17828	134665.8	63.3	28503.36	22261.69
15	Switzerland	2019	8.57528	44.00263	71739.21	80.475	41057.11	6079
16	Turkey	2019	82.57945	20.44486	289135.8	50.3	22463.28	35017.392

3.2 Model development

3.2.1. Standard ANN model

Artificial Neural Networks are the computational methods which are capable of being trained to discern intricate relationships among two or more independent variables or datasets and uses a matrix programming environment to deal with the mathematical challenges [23].The architecture of the ANN model follows a specific format, characterized by neurons arranged in a complex and nonlinear configuration [24]. Typically a simple ANN model consists of three different layers; an input layer that gets the inputs and transfer in the next layer, a hidden layer to process the input according to the problem, and lastly an output layer that provides the final outcomes. As the complexity of the problem escalates both the number of layers and the intricacy of the ANN model also proliferate [25]. This study has utilized the ANN approach for the prediction of MSW generation where various inputs influencing MSW generation are fed into the input layer, processed through the hidden layer, and ultimately, the predicted MSW output is generated by the output layer as shown in Fig. 2.

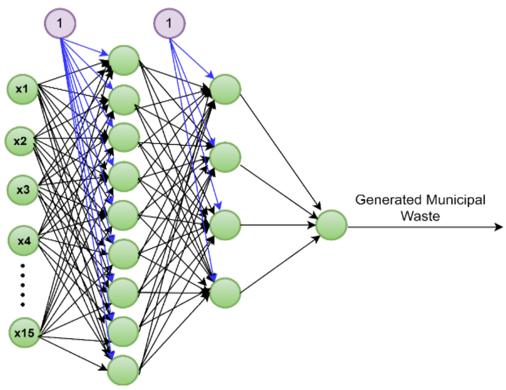


Figure. 2 Internal architecture of ANN model

3.2.2. Standard LSTM model

The LSTMs and ANNs share the similar basic architecture encompasses input, output and hidden layer. The principal distinction between the LSTM and the conventional ANN lies in the composition of their hidden layers. Unlike ANN, the hidden layers of LSTM incorporate LSTM layers and dropout layers, marking a characteristic feature of the LSTM architecture [11]. The LSTM algorithm demonstrates outstanding performance in forecasting time series data including MSW predictions [26]. An LSTM cell consists of three different gates; input, forget and output gate. The mathematical equations for an LSTM cell are as follows.

$$N_t = \sigma\,(Xa_t + Xh_{t-1} + b) \tag{1}$$

$$G_t = \sigma\,(Xa_t + Xh_{t-1} + b) \tag{2}$$

$$C_t = tanh\,(Xa_t + Xh_{t-1} + b) \tag{3}$$

$$C_{ct} = N_t \odot C_{t-1} + I_t * C_t \tag{4}$$

$$Q_t = \sigma\,(Xa_t + Xh_{t-1} + b) \tag{5}$$

$$H_t = Q_t \odot tanh\,(C_{ct}) \tag{6}$$

Where X and b represents weights and biases
at represents input and ht represents hidden state
Nt represents input value given at time t
Gt has represented the forget gate given at time t
Ct has represented the update cell state
Cct has represented the cell state
Qt has represented the output gate
Ht has the representation of hidden state and tanh and σ have served as the activation functions.

3.2.3. Standard GRU model

The GRU model shares the gated architecture with LSTM, however, GRU does not have an output gate rather it has two gates; a reset and an update gate [13]. A GRU architecture has the least number of gates and hence it has fewer parameters than an LSTM. In this architecture, the reset gate discerns less pertinent information from the prior state, excluding it from the subsequent state. Whereas, the update gate preserves crucial information from the previous state for incorporation into the subsequent state. The mathematical formulations governing the data processing in a GRU model are delineated as follows:

$$R_t = \sigma\,(Xa_t + Xh_{t-1} + b) \tag{7}$$

$$Z_t = \sigma\,(Xa_t + Xh_{t-1} + b) \tag{8}$$

$$H_{ht} = tanh\,(Xa_t + R_t \odot Xh_{t-1} + b) \tag{9}$$

$$H_t = (1 - Z_t) \odot h_{t-1} + Z_t \odot H_{ht} \tag{10}$$

where X and b represents weights and biases
at represents input and ht represents hidden state
Rt has represented the reset gate at given time t
Zt has represented update gate at time t
Hht has the representation of the candidate hidden state
Ht represents hidden state and tanh and σ have served as the activation functions

3.2.4. Enhanced GRU with optimized update gate

Although, GRU model has top-notch performance for resolving various issues including

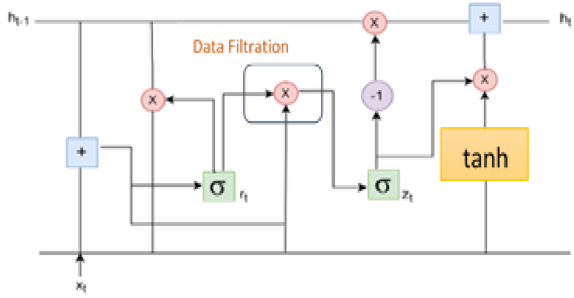


Figure. 3 Architecture of EGRU-OU model

solid waste predictions, yet it has its own pros and cons like every other model. In the standard GRU model, the reset gate is responsible for controlling least important information from heading towards the next state whereas, the update gate is responsible for upholding the most important information for the next hidden state. The paramount flaw in the standard GRU model is the same data is processed by the both gates, however there is absolutely no need for the update gate to process the data which has already been processed and ignored by the reset gate as update gate will also ignore that data. This redundancy that lies between its two gates eventually leads to the enhanced training time and low convergence rate. Therefore, this study has developed the Enhanced GRU with optimized update gate model to resolve the redundancy between the two gates by providing the filtered data to the update gate. This filtration of the data is responsible to shrink the redundant information to a greater extend between the two gates. Furthermore, the training time of the EGRU-OU model will also reduce by providing a filtered data to the update gate. The graphical representation of the EGRU-OU model has been presented in Fig. 3.

In the EGRU-OU approach the input of the update gate has been altered by multiplying the actual input with the reset gate's output. In this way, the update gate will get the already filtered data that would not contain the least important information and hence, update gate will not have to reprocess the data which is already been ignored by the reset gate and eventually data redundancy issue would be addressed. The mathematical equations for Filtered GRU model are given below.

$$R_t = \sigma\,(Xf_t + Xh_{t-1} + b) \tag{11}$$

$$Z_t = \sigma\,(Xf_t + Xh_{t-1} + b) \odot R_t \tag{12}$$

$$H_{ht} = tanh\,(Xf_t + R_t \odot Xh_{t-1} + b) \tag{13}$$

$$H_t = (1 - Z_t) \odot h_{t-1} + Z_t \odot H_{ht} \quad (14)$$

Where X and b represents weights and biases
ft represents input and ht represents hidden state
Rt has represented reset gate at time t
Zt has represented update gate at time t
Hht represents the candidate hidden state
Ht represents hidden state and tanh and σ have served as the activation functions.
A step by step flow of work in EGRU-OU model has been given in Algorithm 1.

Algorithm I: Enhanced Gated Recurrent Unit with Optimized Update Gate (EGRU-OU)

Parameters initialization
EGRU-OU Factors
-Set of Input features: $[y_1, y_2, y_3, \dots \dots \dots \dots, y_n]$
-Weight matrix: $[m_1, m_2, m_3, \dots \dots \dots \dots, m_n]$
- Update gate Bias: b_z
-Reset gate Bias: b_r
-Bias of candidate state: $b_{\hat{h}}$
Step 1: Create EGRU-OU approach
Allocate $[f_1, f_2, f_3, \dots \dots \dots \dots, t_n]$ to $[\bar{f}_1, \bar{f}_2, \bar{f}_3, \dots \dots \dots \dots, \bar{f}_n]$
-Sum up the 1st layer of the GRU (l_1 units) with the activation function (*Sigmoid*) where the dropout is d_1 and the recurrent dropout is r_1.
-Sum up the 2nd layer of the GRU (l_2 units) with the activation function (*Sigmoid*) where the dropout is d_2 and the recurrent dropout is r_2.
Step 2: Train and validate the model
-Calculate the reset gate r_t: by utilizing the Equation (10)
-Compute update gate z_t: using Equation (12)
- Calculate the candidate key $\bar{h}_t$: by utilizing the Equation (12)
-Utilize the Equation (13) to compute the output h_t
while stopping criteria not met **do**
while all instances training **do**
-Compute categorical cross-entropy loss function.
-Compute the grad g^i_{SC} of MSE_i.
end while
Update rule
-Update the weights $X: X = X + X_n$ while X_n= Weight update factor
-Update the Biases $b_i: b = b + b_{in}$ while b_{in}= Input units bias update factor
-Update the Biases $b: b_j = b_j + b_{jn}$ while b_{jn}= Hidden units bias update factor
While (All layers have been trained)
End while
Step 3: Test the Model
-Test fine-tuned hyperparameters with the test data subset.
return Compute the final outcomes in the test data subset.

Moreover, 100 epochs have been employed to train the ANN, LSTM, GRU and EGRU-OU models. During the training phase random weights have been used in each iteration. The model's efficiency has been improved through the application of an optimizer called "Adam." In addition to that a dropout value of 0.02 has been configured within the layers to address concerns related to over-fitting. Furthermore, early stopping criteria have been instigated, where training halts with no further improvement in the validation loss within a span of 20 epochs. The evaluation of the developed model's performance has been conducted using error matrices.

3.3 Performance evaluation criteria

In the current study, three different error matrices RMSE, MAE, and coefficient of determination (R2) have been employed for the performance evaluation of the ANN, LSTM, GRU and EGRU-OU model for the prediction of MSW generation.

A model is deemed to exhibit higher performance when the R² attains a high value and could be computed by the statistical formula given below.

$$R2 = 1 - \frac{SSR}{SST} = 1 - \frac{\sum_{i=1}^{T}(Ai - Bi)2}{\sum_{t=1}^{T}(\bar{Ai} - \bar{Bi})2} \quad (15)$$

Where SSR = squared sum of residuals,
SST = total sum of squares,
T = total values,
Ai and Āi represents actual value and its mean at ith point and
Bi and B̄i represents forecasted value and its mean respectively at ith point.

The additional error metric employed in this study is RMSE which can be computed by determining the division of the difference between real and forecasted values to the total number of values, as illustrated in Eq. 16.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{X}(Ai - Bi)2}{X}} \quad (16)$$

Where X = total observations,
Ai = observed value given at ith point,
Bi = forecasted value given at ith point.
Moreover, for a comprehensive analysis, MAE error metrics has also been employed in this study which is computed by evaluating the division of the total number of values to the difference between the forecasted and actual values, as depicted in Eq. 17.

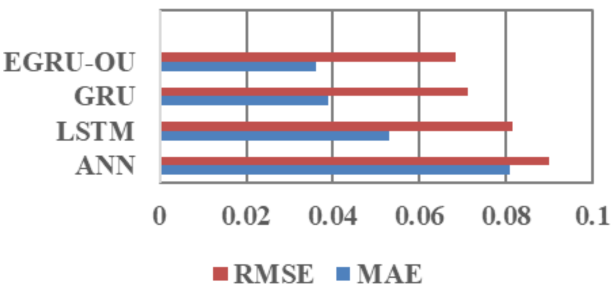
$$MAE = \frac{\sum_{i=1}^{X}|Ai - Bi|}{X} \quad (17)$$

Where X = total values,
Ai = actual value given at ith point,
Bi = forecasted value given at ith point.

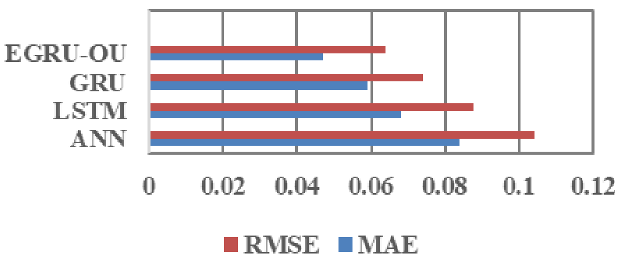
4. Results

The conducted study has employed two hidden layers in all models encompassing standard GRU, LSTM, ANN, and EGRU-OU model. Two different datasets (OECD dataset and Multan dataset) have been utilized in order to train the models and the results have indicated the superior performance of the EGRU-OU model compared to GRU, LSTM and ANN. According to the results, the error matrices have the least values for EGRU-OU model with *MAE* value being 0.036 and *RMSE* value being 0.0684 for Multan dataset. Moreover, the GRU model has demonstrated the second-highest level of performance with *MAE* value being 0.039 and *RMSE* value being 0.0713. However, the LSTM has also shown the acceptable performance with *MAE* value of 0.053 and *RMSE* value of 0.0815 and the least performance has been depicted by the ANN model with the highest values of *MAE* being 0.081 and *RMSE* being 0.090. The graphical representation of the error matrix values for Multan dataset has been illustrated in Fig. 4 (a).

Additionally, the error matrix values for OECD dataset has also underscored the superior performance of the EGRU-OU model exhibiting the lowest *MAE* value being 0.047 and *RMSE* value being 0.0638. The GRU model has followed as the second-highest performer coming up with a low *MAE* value of 0.059 and *RMSE* value of 0.0741. Meanwhile, the LSTM performance remained with a *MAE* value of 0.068 and an *RMSE* value of 0.0874 and again the least performance has been demonstrated by the ANN model, recording the high *MAE* value of 0.084 and *RMSE* value of 0.104. Fig. 4 (b) visually depicts the error matrix values for the OECD dataset.



(a)



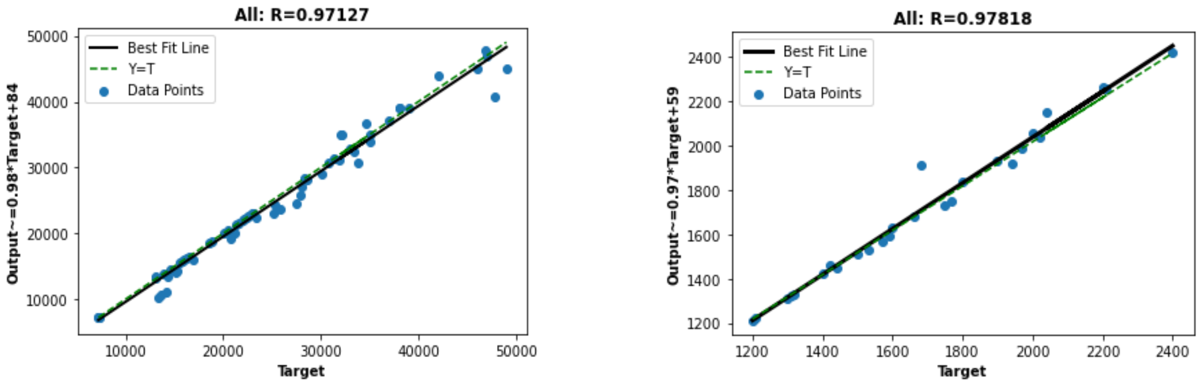
(b)

Figure. 4 Comparative analysis of ANN, LSTM, GRU and EGRU-OU models using: (a) Multan data and (b) OECD dataset based on *MAE* and *RMSE* values

Moreover, to conduct an in-depth analysis of the models' performances, coefficient of determination values have been systematically assessed for all methodologies including EGRU-OU, GRU, LSTM and ANN models. The quantitative values of R^2 serve in gauging the accuracy of the model's prediction. The high value of *R* corresponds to the high accuracy of the model's predictions. The results of *R* values have demonstrated that the EGRU-OU model has performed better than the standard GRU, LSTM and ANN model with the highest values for *R* being 0.978 for Multan dataset and 0.971 for OECD dataset as illustrated in Fig. 5. The values of *R* for the standard GRU model has shown the satisfactory performance of the GRU model with *R* being 0.964 for Multan dataset and 0.957 for OECD dataset as shown in Fig. 6. Furthermore, the values of *R* for LSTM are 0.959 for Multan dataset and 0.948 for OECD dataset as shown in Fig. 7. The ANN model exhibits the lowest *R* value, demonstrating its least predicting performance with *R* being 0.917 for Multan dataset and 0.906 for OECD dataset as depicted in Fig. 8.

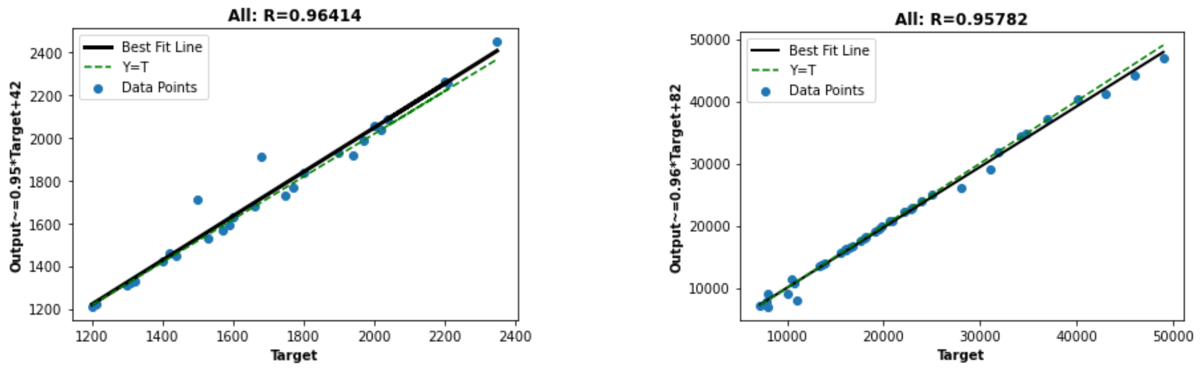
In addition, the overall disparity between the real and the forecasted values of the EGRU-OU model for both OECD and Multan dataset has been shown in Fig. 9 where superior performance of the EGRU-OU model with high accuracy of predicted results can be noticed. Additionally, the standard GRU model demonstrates a high level of accuracy in its forecasted results following the EGRU-OU model. Fig. 10 has elucidated the overall discrepancy between the real and forecasted values of the standard GRU model, highlighting the commendable performance of the GRU approach. Furthermore, the general variance between the real and forecasted values of the LSTM model for both the OECD and Multan datasets is depicted in Fig. 11. The figure has highlighted the commendable performance of the LSTM model, showcasing satisfactory accuracy in its predicted results. And, the least prediction accuracy results have been demonstrated by the ANN model as illustrated in Fig. 12.

Received: February 24, 2024. Revised: May 17, 2024.



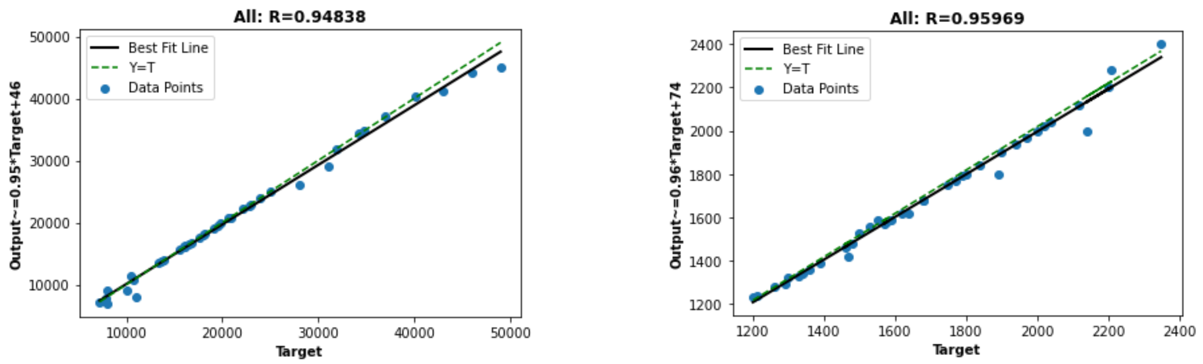
(a) (b)

Figure. 5 The value of *R* in EGRU-OU for: (a) OECD dataset and (b) Multan dataset



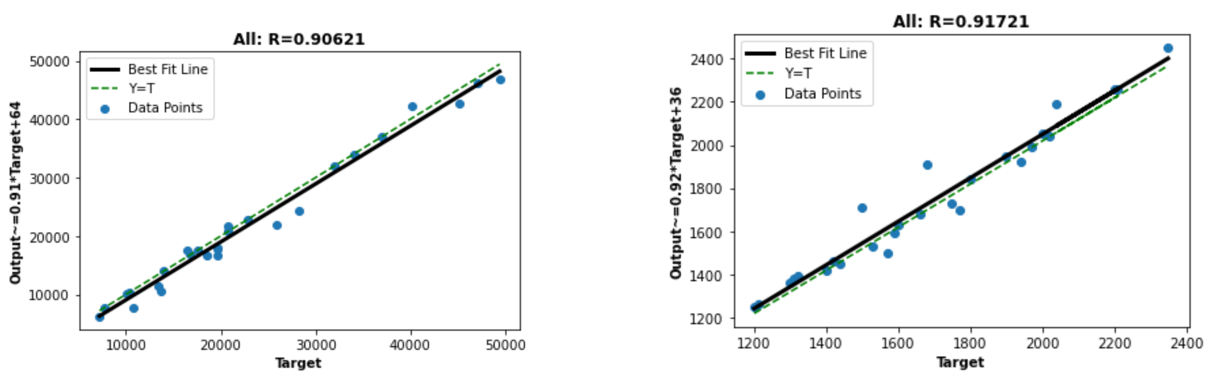
(a) (b)

Figure. 6 The value of *R* in GRU model for: (a) OECD dataset and (b) Multan dataset



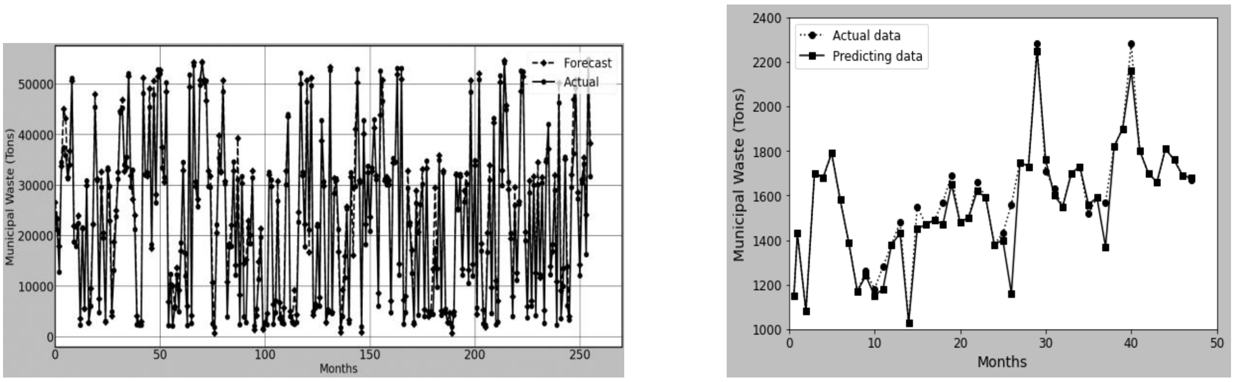
(a) (b)

Figure. 7 The value of *R* in LSTM model for: (a) OECD dataset and (b) Multan dataset



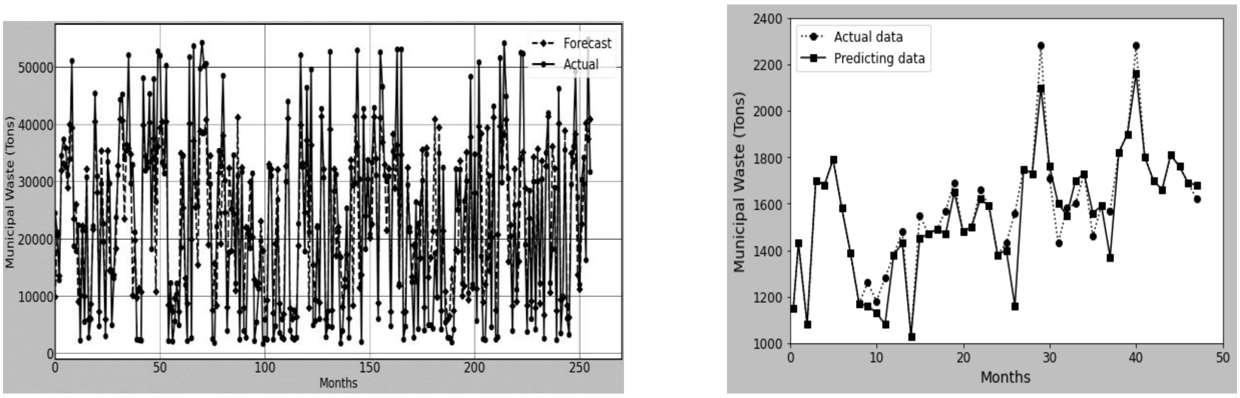
(a) (b)

Figure. 8 The value of *R* in ANN model for: (a) OECD dataset and (b) Multan dataset



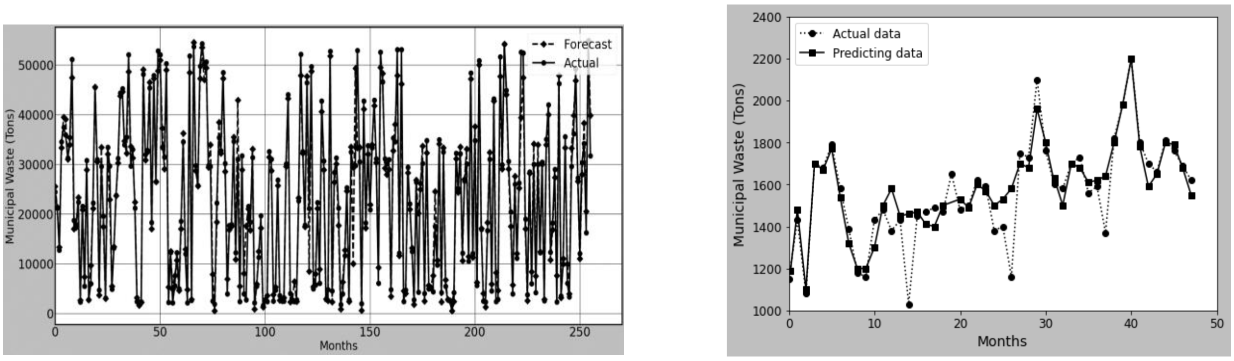
(a) (b)

Figure. 9 MSW prediction accuracy of the EGRU-OU model for: (a) OECD and (b) Multan dataset



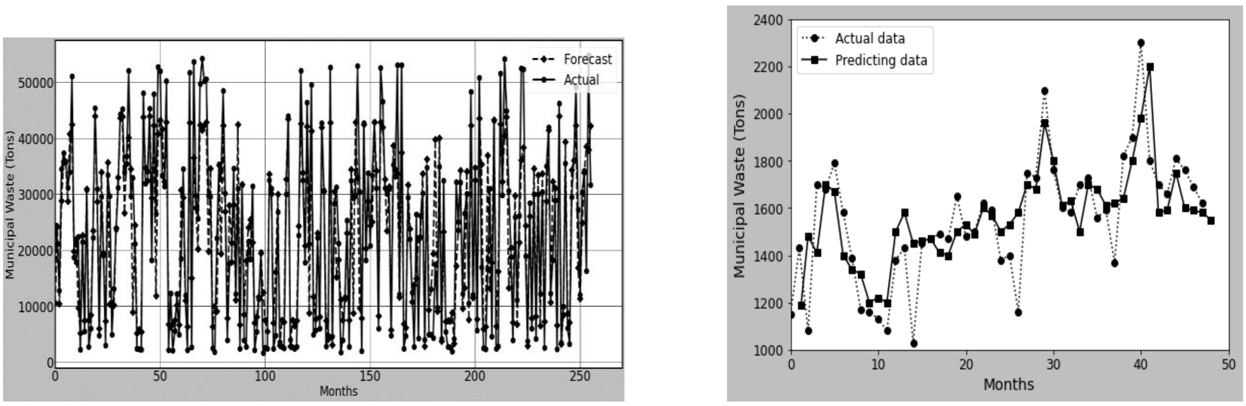
(a) (b)

Figure. 10 MSW prediction accuracy of the standard GRU for: (a) OECD and (b) Multan dataset



(a) (b)

Figure. 11 MSW prediction accuracy of the LSTM model for: (a) OECD and (b) Multan dataset



(a) (b)

Figure. 12 MSW prediction accuracy of the ANN model for: (a) OECD and (b) Multan dataset

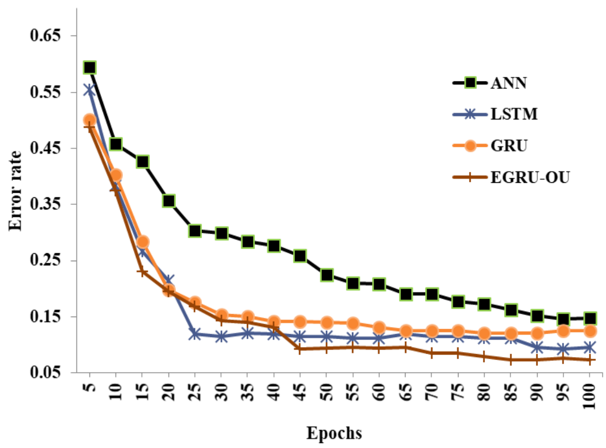


Figure. 13 Comparative analysis of EGRU-OU, GRU, LSTM and ANN based on error rate

5. Discussion

The conducted study has proposed the EGRU-OU approach in order to forecast the MSW generation rate by using two different datasets and compared the results with other benchmark models including standard GRU, LSTM and ANN model. The results have clearly demonstrated the superior performance of the EGRU-OU approach as compared to other state-of-art approaches as shown in Fig. 13 where least error rate could be noticed for the EGRU-OU model.

Furthermore the proposed model could be easily integrated with the solid waste management applications that would help the municipal authorities to have more accurate prediction results regarding the waste generation in future. The more accurate will be the results, the more effective would be the planning to handle the waste. However, the EGRU-OU model still have some limitations, primarily it may struggle when dealing with the very long sequences of the data. This limitation could be addressed in the future studies.

6. Conclusion

The early predictions with high accuracy of the MSW generation grasps the potential to empower the municipal authorities in articulating a robust framework for the management of generated waste. This brings in the necessity of the development of models capable of predicting MSW generation rate with a highest degree of accuracy. The literature review conducted in this study has unveiled the different contributions of researchers in the prediction of waste generation. Several models have been developed to achieve the high accuracy of prediction results. However, it is important to note that each model has its strengths and limitations. In this study, the GRU model has been predominantly emphasized for enhancement purposes. The primary limitation in the standard GRU model arises from its both gates (reset and update) processing the same data. Therefore the major contribution of the conducted study is to propose the EGRU-OU model in order to resolve the redundancy issue. In the EGRU-OU model the filtered data has been provided to the update gate that would shrink the redundant information to a greater extent. And finally EGRU-OU model has been compared with other state of art models including standard GRU, LSTM and ANN model where three different error matrices *RMSE*, *MAE*, and R^2 were employed for the performance evaluation of the models in forecasting the MSW generation. The results have clearly depicted a superior performance of the EGRU-OU approach as compared to other models with the least error matrices values of *MAE* being 0.036, *RMSE* being 0.0684, and highest regression value of 0.978.

Conflicts of Interest

The authors of this research paper declare that they have no conflicts of interest related to the research presented in this manuscript.

Author Contributions

Each author contributed to this research in different ways. The specific contributions are as follows: "Conceptualization, methodology, data analysis literature review, and data collection by Tuba Batool; software, validation, formal analysis by Tuba Batool, Irfan Javid and Nureize Binti Arbaiy; writing—original draft preparation by Tuba Batool, and Lokman Hakim Ismail; visualization, supervision, project administration, and funding acquisition by Dr. Rozaida Ghazali.

Acknowledgments

This research was supported by the ministry of Science, Technology and Innovation (MOSTI) through the Technology Development Fund 1(TDF06211409) and Universiti Tun Hussein Onn Malaysia.

References

[1] https://archive.epa.gov/epawaste/nonhaz/municipal/web/html/ (Accessed 19 December 2023).

[2] https://www.gminsights.com/industryanalysis/municipal-solid-waste-management-market (accessed 01 June 2023).

[3] T. Batool, R. Ghazali, I. Javid, and N. B. Arbaiy, "Gated Recurrent Unit for Landfill Area Estimation Based on Municipal Solid Waste Prediction", *International Journal of Intelligent Engineering & Systems*, Vol. 16, No. 5, 2023, doi: 10.22266/ijies2023.1031.58.

[4] N. Sun, and S. Chungpaibulpatana, "Development of an appropriate model for forecasting municipal solid waste generation in Bangkok", *Energy Procedia*, Vol. 138, pp. 907-912, 2017

[5] S. Azadi, and A. K. Jashni, "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*, Vol. 48, pp. 14-23, 2016, doi: 10.1016/j. wasman.2015.09.034.

[6] L. Bingchun, L. Zhang, and Q. Wang, "Demand gap analysis of municipal solid waste landfill in Beijing: Based on the municipal solid waste generation", *Waste Management*, Vol. 134, pp. 42-51, 2021.

[7] C. E. Ossandon, A. M. Nieto, and N. Harsch, "Using a fuzzy TOPSIS-based scenario analysis to improve municipal solid waste planning and forecasting: a case study of Canary archipelago (1999–2030)", *J. Cleaner Prod*, Vol. 176, pp. 1198-1212, 2018, 10.1016/j.jclepro.2017.10.324

[8] K. Kolekar, T. Hazra, and S. Chakrabarty, "A Review on Prediction of Municipal Solid Waste Generation Models", *Procedia Environmental Sciences*, Vol. 35, pp. 238-244, 2016.

[9] C. Ghinea, E. Drăgoi, E. Comăniţă, M. Gavrilescu, T. Câmpean, S. Curteanu, and M. Gavrilescu, "Forecasting municipal solid waste generation using prognostic tools and regression analysis", *J Environ Manage*, Vol. 1, No. 182, pp. 80-93, 2016, doi: 10.1016/j.jenvman. 2016.07.026.

[10] S. A. Ali and A. Ahmad, "Forecasting MSW generation using artificial neural network time series model: a study from metropolitan city", *SN Applied Sciences*, Vol. 1, No. 11, pp. 1-16, 2019, 10.1007/s42452-019-1382-7.

[11] N. Dongjie, W. Fan, D. Shijin, H. Sheng, and W. Boran, "Detection of long-term effect in forecasting municipal solid waste using a long short-term memory neural network", *Journal of Cleaner Production*, Vol. 290, p. 125187, 2021.

[12] U. Anıl, and S. K. Kaya, "Deep Learning Based a Comprehensive Analysis for Waste Prediction", *Operational Research in Engineering Sciences: Theory and Applications*, Vol. 5, No. 2, pp. 176-189, 2022.

[13] I. Javid, A. K. Z. Alsaedi, R. Ghazali, Y. M. M. Hassim, and M. Zulqarnain, "Optimally organized GRU-deep learning model with Chi 2 feature selection for heart disease prediction", *Journal of Intelligent & Fuzzy Systems*, Vol. 42, No. 4, pp. 4083-4094, 2022.

[14] S. Kujawa, J. Mazurkiewicz, and W. Czekała, "Using convolutional neural networks to classify the maturity of compost based on sewage sludge and rapeseed straw", *J. Clean. Prod.*, Vol. 258, 2020.

[15] C. Bhatt, I. Kumar, V. Vijayakumar, K. U. Singh and A. Kumar, "The state of the art of deep learning models in medical science and their challenges", *Multimedia Systems*, Vol. 27, No. 4, pp. 599-613, 2021.

[16] https://www.ibm.com/topics/deeplearning#:~:text=the%20next%20step-,What%20is%20deep%20learning%3F,from%20large%20amounts%20of%20data. (Accessed on 21 December 2023).

[17] https://www.simplilearn.com/tutorials/deeplearning-tutorial/deep-learning-algorithm.

[18] W. Fan, D. Niu, S. Dai, and B. Wu, "New insights into regional differences of the predictions of municipal solid waste generation rates using artificial neural networks", *Waste Management*, Vol. 107, pp. 182-190, 2020.

[19] H. Jammeli, R. Ksantini, F. B. Abdelaziz, H. Masri, "Sequential artificial intelligence models to forecast urban solid waste in the city of Sousse, Tunisia", *IEEE Transactions on Engineering Management*, 2021.

[20] K. K. Adusei, K. T. W. Ng, T. S. Mahmud, N. Karimi, C. Lakhan, "Exploring the use of astronomical seasons in municipal solid waste disposal rates modeling", *Sustainable Cities and Society*, Vol. 86, p. 104115, 2022.

[21] G. Rashmi, and S. S. Kumar, "Regularized Noise based GRU Model to Forecast Solid Waste Generation in the Urban Region", *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, Vol. 12, No. 10, pp. 5449-5458, 2021.

[22] M. Cubillos, "Multi-site household waste generation forecasting using a deep learning approach", *Waste Management*, Vol. 115, pp. 8-14, 2020.

[23] M. M. Hoque, and M. T. U. Rahman, "Landfill area estimation based on solid waste collection prediction using ANN model and final waste disposal options", *Journal of Cleaner Production*, Vol. 256, p. 120387, 2020.

Received: February 24, 2024. Revised: May 17, 2024.

[24] A. Malekian, and N. Chitsaz, "Concepts, procedures, and applications of artificial neural network models in streamflow forecasting", *Advances in Streamflow Forecasting*, pp. 115-147, 2021.

[25] https://www.analyticsvidhya.com/blog/2021/07/understanding-the-basics-of-artificial-neural-network-ann/ (Accessed on 23 December 2023)

[26] K. Lin, Y. Zhao, and J. H. Kuo, "Deep learning hybrid predictions for the amount of municipal solid waste: a case study in Shanghai", *Chemosphere*, Vol. 307, p. 136119, 2022.