Design of Smart Waste Bin and Prediction Algorithm for Waste Management in Household Area

Siti Hajar Yusoff, Ummi Nur Kamilah Abdullah Din, Hasmah Mansor, Nur Shahida Midi, Syasya Azra Zaini
International Islamic University MalaysiaJalan Gombak,
53100 Kuala Lumpur, Malaysia, Phone: (+603) 6196 4000, Fax: (+603) 6196 4053

Article Info

ABSTRACT

Maintaining current municipal solid waste management (MSWM) for the next ten years would not be efficient anymore as it has brought many environmental issues such as air pollution. This project has proposed Artificial Neural Network (ANN) based prediction algorithm that can forecast Solid Waste Generation (SWG) based on household size factor. Kulliyyah of Engineering (KOE) in International Islamic University Malaysia (IIUM) has been chosen as the sample size for household size factor. A smart waste bin has been developed that can measure the weight, detect the emptiness level of the waste bin, stores information and have direct communication between waste bin and collector crews. This study uses the information obtained from the smart waste bin for the waste weight while the sample size of KOE has been obtained through KOE's department. All data will be normalized in the pre-processing stage before proceeding to the prediction using Visual Gene Developer. This project evaluated the performance using R2 value. Two hidden layers with five and ten nodes were used respectively. The result portrayed that the average rate of increment of waste weight is 2.05 percent from week one until week twenty. The limitation to this study is that the amount of smart waste bin should be replicated more so that all data for waste weight is directly collected from the smart waste bin.

Corresponding Author:

Siti Hajar Yusoff,
International Islamic University MalaysiaJalan Gombak,
53100 Kuala Lumpur, Malaysia, Phone: (+603) 6196 4000, Fax: (+603) 6196 4053
Email: sitiyusoff@iium.edu.my

1. INTRODUCTION

As years passed by, MSWM will remain to be one of the crucial problems faced in most of the developing countries including Malaysia. This problem is mainly resulted from the variation in the household size [1]. On 1st July 2016, a new smartphone application has been officially launched by Selangor government [2]. The application, namely "iClean Selangor" is assumed to cater the problem of solid waste management in household area. This technique is meant for the people in Selangor to lodge a report to Kumpulan Darul Ehsan Berhad (KDEB) Waste Management if there is any waste related problem arise in Local Council of Klang, Selayang, Ampang Jaya and Kuala Langat area. KDEB Waste Management is the new garbage collector operated under the Selangor state government [3]. Nevertheless, this system is considered impractical because of some possibilities that might occur as not all residents have access to the internet especially in the remote area thus making them harder to lodge the report. Moreover, it has been almost two years since the launching of this system in 2016, yet the record of the downloads is only 500 which is very little. It shows the impracticality and unawareness about the system itself.

Nonetheless, public waste bins will be filling up faster than ever. Inevitably, many of the bins end up overflowing before collection time, causing not only cluttered streets and bad odors, but also negative health and environmental impacts. Consequently, negative impacts that bring harm to the environment will slowly be the alarming issues to the society [4]. Consequences that could happen due to poor handling of solid waste management is the spread of bacteria and insects from waste can increase the possibility of people facing salmonella bacteria which will lead to food poisoning, fever and other severe illnesses. It could also cause overflow of garbage prompt the air pollution index from good to bad drastically. It will eventually disrupt the human respiratory processes and affect our health condition as the bad gases are absorbed to the lung and spread to other parts of the body. Finally, garbage can contaminate surface waters, which can affect all ecosystems.

There are two types of managing municipal solid waste in Malaysia. First is the management of waste as a whole, whereby the future amount of waste generated is needed in order to plan for the landfill and second is the management of waste collection at household area to avoid public bin's overflow. Currently, MSWM in Malaysia do not have the exact statistic of how much waste is generated and how many times the waste bins get full per day. Without these statistics, it is very hard for the government to provide ample spaces of the compost sites and to plan for the garbage pick-up schedule for the future.

On the other hand, maintaining current way of curb side collection for different household size would be very inefficient [5]. The waste collector crews are only following their own schedule to collect the garbage at the respected scheduled area. Unfortunately, this will cause inefficiency in term of fuel consumption and time management. Sometimes, due to the scheduled collection routine, the bins can get overflow way before the collection time. This will further cause diseases including typhoid fever, dengue and other major illnesses.

Therefore, it is very important to predict the amount of waste generated to ease the process in managing future MSWM. Recently, there are many researches on forecasting the SWG based on prediction models [6]-[8]. Prediction models can give information about the future SWG based on many performance's criterion such as Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE) and R^2.

There are few researches that have been conducted to predict future estimation based on ANN [9]-[11]. A study by Sun & Chungpaibulpatana used MLP under ANN model and Pearson Correlation to predict SWG in Bangkok [12]. At the beginning of the research, few modelling techniques have been explored based on few influences such as population growth and household income. Also, interpolation technique has been applied during data collection stage due to some missing values. Neural fitting tool has been used to select, create and train data of the network based on MSE and regression analysis. For MLP, one neuron hidden layer has been applied that results in the acceptable fitting value R^2 of 0.96. During the evaluation stage, the performance for both techniques has been compared. The results managed to illustrate that ANN model is much more accurate compared to Principal Component Analysis-Regression (PCA-Regression) by 10% based on R^2 value. However, the values of MSE for both PCA-Regression and ANN model were very high which are 221805.2 and 63929 respectively.

According to a research, the data regarding garbage levels are very useful to provide optimized route for the waste collector [13]. This data can ultimately help in reducing the cost associated with the fuel consumption used by truck. The longer the route taken by the truck's driver more fuel are used. On top of that, Mahajan et al. said that if the waste collector keeps attending to the garbage bin even though it is not full yet, it will also provide a waste of time and inefficient working hours. The research also mentioned about the importance of detecting the moisture inside the bins. The moisture sensors have been used to provide data of waste segregation inside the bin. This study used three types of sensors which are; the level sensor, humidity sensor and the load cell to detect the waste level, moisture state and the weight of the waste. These sensors have been integrated by a microcontroller named Raspberry-pi 3. However, this study didn't provide data storage. Thus, data such as waste level and weight of waste collected are not stored. This make this project insignificant because these important data are missing. Without these data, future amount of waste generated can't be estimated.

The main objective of this research is to design smart waste bin and to design an efficient prediction algorithm for waste management to predict the generation of waste based on household size in Kulliyyah of Engineering in International Islamic University Malaysia. The remaining section of this paper comprises of few main parts. Section 2 will be explaining about the smart waste bin while section 3 will lay out data acquisition process. In Section 4, pre-processing stage will be discussed. On the other hand, Section 5 will explain about the evaluation while Section 6 lays out the results and analysis of the experimentation. As a conclusion, the last section will conclude the paper.

2. SMART WASTE BIN

This section will discuss about the proposed methodology that is used to design the smart waste bin. In this section, there will be three subsections; smart waste bin overview, software part and hardware part. Subsection 2.1 will explain on how the smart waste bin is working. Whilst, Subsection 2.2 and Subsection 2.3 describes the software and hardware configuration of the smart waste bin respectively.

2.1. Smart Waste Bin Overview

The prototype's design of the waste bin involves both software and hardware part. Further details about the software and hardware to design this smart waste bin will be discussed in Subsection 2.2 and Subsection 2.3 respectively. On the other hand, the flowchart of the system operation is laid out as in Figure 1. Initially, when the project installed and powered on the system, the GSM module will be initialized. It is important to make sure that the GSM has been initialized as if it is not, the system would not able to send notification to the waste collector crews later. Then, the sensors built in the system will measure the acquired data; level and weight of the waste. This information will be sent to SD card and will be displayed at the LCD. If the level has exceeded 90 percent of the capacity of the bin, a notification in the form of SMS will be sent to the waste collector crews to notify for the pickup time. On the other hand, an algorithm has also been developed to push the LCD display to display "Collection time! Do not put any waste" when the level has exceeded 90 percent of the capacity. So, the passerby knows that no more waste can be put into the bin and that the bin has exceeded the capacity of the waste that can be put into it. This is to prevent littering to occur.
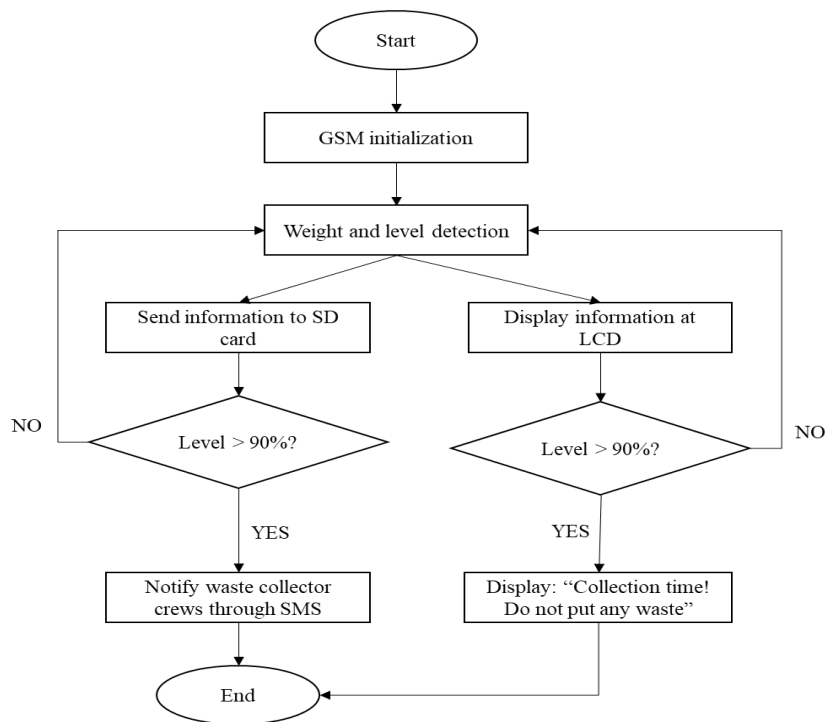

Figure 1. System operation's flowchart

2.2. Software Part

For the software part, the project uses Sublime Text 3 for code's editing; shown in Figure 2, Arduino Integrated Development Environment (IDE) for building, compiling and uploading the code into Arduino MEGA controller; shown in Figure 3, while Fritzing is used to sketch the schematic diagram of the hardware prototype; shown in Figure 4.
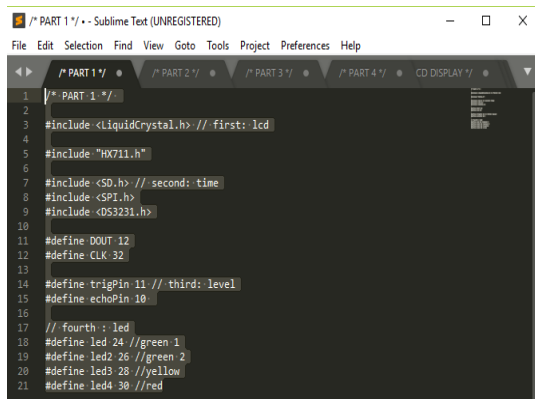
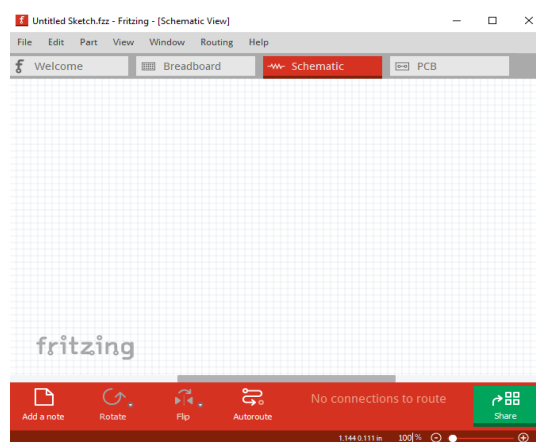Figure 2. Sublime text 3 software

Figure 3. Arduino IDE software

Figure 4. Fritzing software

This project has divided the software's pseudocode into five parts. These five parts are;
a. Pins define and setup of library
b. LCD configuration, load cell scale setup, SD card initialization
c. General setup
d. Calculation of sensor's formula and getting information
e. SD storing and GSM configuration

In the first part; pins define and setup of library, selected pins used in Arduino MEGA have been defined. Also, library used such as "SoftwareSerial.h" and "LiquidCrystal.h" are called. Whilst, during the second part, the pins used for LCD is highlighted. Also, a calibration factor for the load cell is set from several sets of trials and errors. On top of that, a file namely "myFile" is created to enable to store the information regarding the waste weight and level. Part three is where most of the setup is done. Two most important features in part three's algorithm are the initialization of the SD card and GSM. Initializing the GSM is very important as it is to make sure that the network is connected and that the notification can be sent to the waste collector crews later on.

Moving on to part four, this part is where the system collect the data from the waste bin. With the use of Real-Time Clock, this bin is able to provide precise data at the precise time so the data can be very useful to be analyzed later on. Also in this part, the calculation for the level sensor is made. The level sensor used in this project was the Ultrasonic Sensor (model: HC-SR04). The algorithm for the calculation is shown in Equation 1.

$$distance = (duration / 2) * 0.0344 - 0.5$$ (1)

Where, 'distance' is the level of waste filled in the waste bin and 'duration' is the time taken for the echo pulse to go back and forth from the mouth of the bin and top of the waste inside the bin. Finally, the last

part portrayed the algorithm of how the information is stored into the SD card. Also in this step, an SMS will be sent to the waste collector crews if the bin has reached the maximum capacity.

2.3. Hardware Part

The final prototype's breadboard and schematic diagram are shown in Figure 5 and Figure 6 respectively. As for the hardware part shown in Figure 5, Arduino MEGA, ultrasonic sensor module, load cell amplifier, sim800L GSM module, memory card, SD card module, liquid-crystal display (LCD), four light-emitting diode (LED) and real-time clock equipped with temperature sensor are used. As shown in Figure 5 and Figure 6, the microcontroller used in this project was Arduino MEGA 2560. Arduino MEGA 2560 has 72 pins altogether. In this project, not all pins in Arduino MEGA will be used. In general, there were seven components which have been connected to the Arduino MEGA. However, for LCD, not all pins are needed to be used. Hence, these pins will not be connected to the Arduino MEGA.

The general working system design shown in Figure 7 illustrates that there are three bins installed in the KOE's building and each bin has different sim cards of different registered numbers. The reason why different bins have different registered mobile number is because each bin will send its location and bin's status update to the waste collector crews.
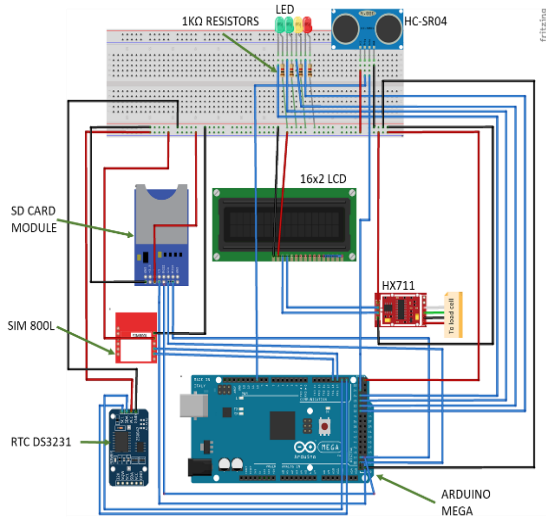

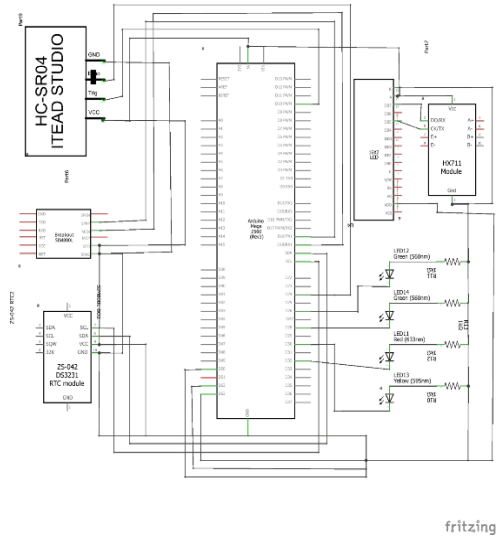Figure 5. Prototype's breadboard diagram

Figure 6. Prototype's schematic diagram

Figure 7. Working system's design for 3 different bins at 3 different locations around KOE's building

3.   DATA ACQUISITION

The data for household size and the amount of waste generated in KOE buildings have been collected via two mediums, from office of Department of Academic Affairs in KOE and through a smart waste bin. In this project, four blocks which are E0, E1, E2, E3 in KOE buildings represent different household size. Table 1 represents the amount of household for each building and the concurrent amount of waste bins installed. The amount of waste bin installed has been obtained through observation by checking each block.

Table 1. Current Waste Bin's Distribution in KOE

Block	Level of study	Number of students	Number of bins
E0	Second year	618	28
E1	First year	679	30
E2	Third year	667	24
E3	Fourth year	437	16

Due to certain circumstances, only three smart waste bins were installed and the rest of the data were acquired via weighing scale. During the experimentation, when the bin has reached 90 percent of the maximum allocation of the size, a message has been sent via GSM that notify the collection time. The message received is shown in Figure 8. This Figure 8 also shows the location of the bin and the amount of waste collected when the waste bin is full. On the other hand, Figure 9 represents the amount of waste generated for week 1 until week 5. As seen in Figure 9, the amount of waste generated in block E1 are highest compared to the other three blocks. Whilst, block E3 generated the least amount of waste. The highest amount of waste collected in one week during the observation was in block E1 in week 5 where the amount was 205kg per week. In next section, this project will continue with the pre-processing stage.

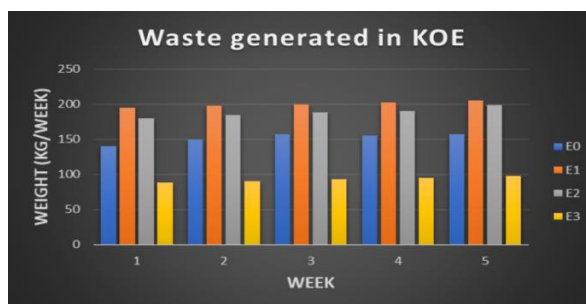Figure 8. Notification sent by the smart bin

Figure 9. Amount of waste generated during five weeks of experimentation

4.   PRE-PROCESSING

Data for waste weight in Figure 9 has undergone the first step in the pre-processing stage which is obtaining the trend line. Using cftool in MATLAB, the R^2 value for each trend lines of each block were obtained. There are many trend lines that have been applied such as linear and polynomial. The results have been laid out in Table 2. As shown in Table 2, all R^2 values are more than 90 percent which in this case, the values go beyond 0.9. Whilst, the highest combination for all trend lines of all blocks are maximum when a polynomial of degree 2 was applied.

Table 2. Trend Lines Via Curve Fitting Tool for Household Size

Block	Combination	Trend line	Fitting options	R² value
E0	1	Linear	-	0.9392
	2	Polynomial	Degree: 2	0.9914
E1	1	Linear	-	0.9952
	2	Polynomial	Degree: 2	0.9964
E2	1	Linear	-	0.9376
	2	Polynomial	Degree: 2	0.9954
E3	1	Linear	-	0.9865
	2	Polynomial	Degree: 2	0.9977

The next step in this pre-processing stage is to forecast the generation of waste using ANN via Visual Gene Developer. All data is needed to be normalized before undergoing the prediction. The normalization step was conducted for each household size variables using Visual Gene Developer. Figure 10 until Figure 13 show the process and the value for normalization for each household size (E0 until E3). This 'max number' is taken as xmax. This normalization done in this software. Different combination will be tested and the resulted R² value will be depicted in Table 3 in Section 5.
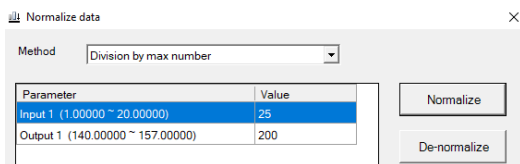
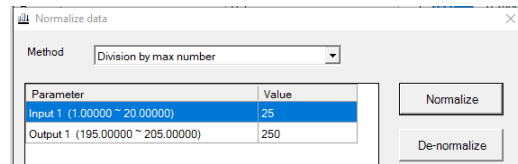Figure 10. Data normalization for E0
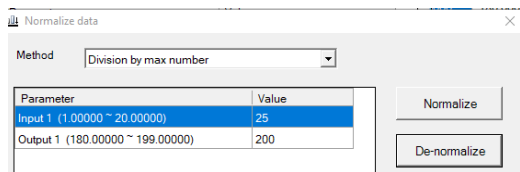
Figure 11. Data normalization for E1
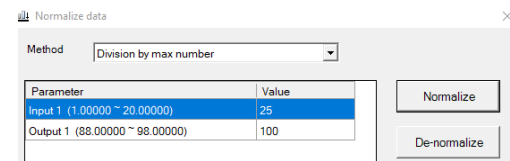
Figure 12. Data normalization for E2

Figure 13. Data normalization for E3

5. EVALUATION

In this section, different sets of regressions obtained in previous Section 4 will be compared and evaluated based on different household sizes. The value of R² obtained are summarized in Table 3. As depicted in Table 3, the highest R² value for all household size obtained when two number of hidden layers have been applied with the number of nodes in layer 1 is 5 and the number of nodes in layer 2 is 10. The highlighted column 2 shown in Table 3 depicted this result.

Table 3. Summary of Performance of Prediction Model based on Different Hidden Layers and Number of Nodes for Household Size Factor

Hidden layer	1st node	2nd node	E0	E1	E2	E3
1	10	-	0.7673237	0.9319049	0.9178493	0.961199
2	5	10	0.7926933	0.9960697	0.9886416	0.983776
2	10	5	0.7597373	0.9919297	0.9309337	0.785193

For household size 1 (E0), the highest R² value obtained was 0.792693, household size 2 (E1), the highest R² value obtained was 0.9960697, household size 3 (E2), the highest R² value obtained was 0.9886416 and for household size 4 (E3), the highest R² value obtained was 0.983776. One of the regression lines for household sizes in Table 3 which is the regression line for household size 1 is shown in Figure 14. The regression line shown in Figure 14 and the highlighted column in Table 3 illustrates the accuracy of the algorithm in reference to the household size factor. As being shown in Figure 14, the blue line is very close to

the threshold line (grey line). This shows the accuracy of the algorithm with reference to the obtained R^2 value.

From Table 3, the project has deducted that the combination of the number of hidden layers and the number of nodes which gives the highest R^2 value is the second combination. Next, this project will use this combination as the prediction algorithm to predict the amount of waste generated for household size factor. Later, this predicted amount of waste will be compared with the observed weight for week 1 until week 5. The result of the comparison for one of the sample size, which is household 1(E0) is depicted in Figure 15. Figure 15 shows that the observed and predicted line will be in the same projection for the upcoming week. As shown in Figure 16, the value of the sum of error for this algorithm is only 0.005932442.
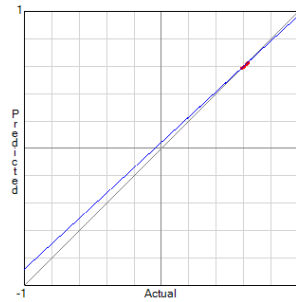

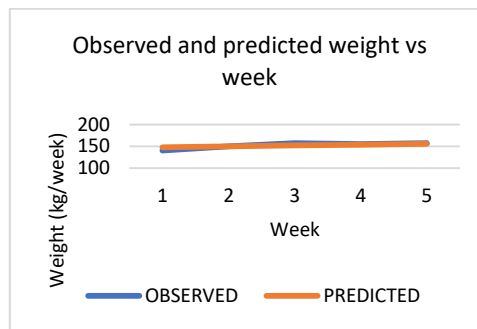Figure 14. Regression line of household size 1 (E0)

Figure 15. Comparison of the observed and predicted weight for household size 1
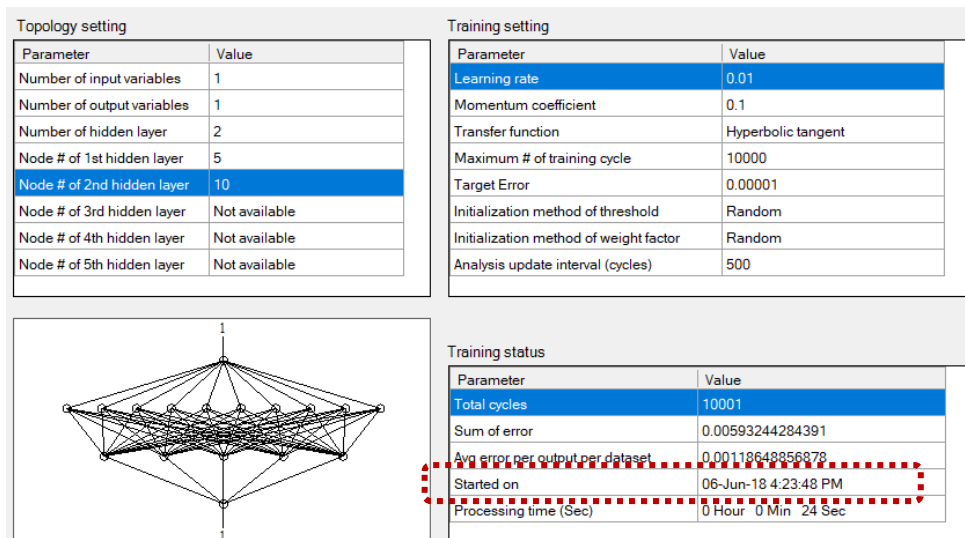

Figure 16. Visual gene developer experimentation layout for household size

6.   RESULTS AND ANALYSIS

Finally, the combination of the hidden layers and the number of nodes in previous Section 5 will be used to forecast the amount of waste generated for the four household sizes (E0-E3) until week 20. The predicted SWG in four buildings (E0-E3) in KOE, IIUM from week 1 until week 20 is shown in Figure 17 until Figure 20.

The criterion of the prediction algorithm deducted from the previous stage are:

    a.   ANN
    b.   Two hidden layers
        i.   First hidden layer: 5 nodes
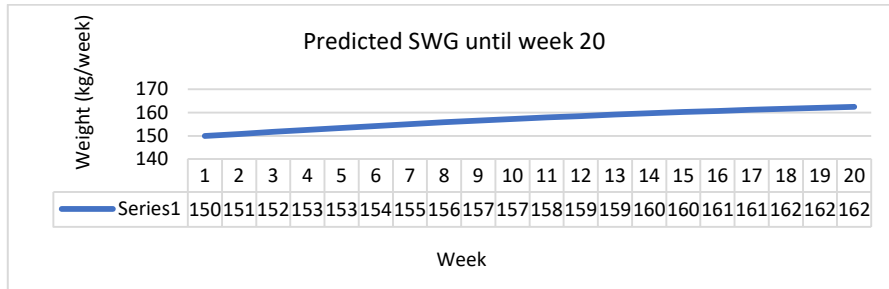   ii.   Second hidden layer: 10 nodes

Figure 17. Prediction of SWG in E0 for twenty weeks

As shown in Figure 17, the number of waste that will be generated in E0 in week 20 is 162.43kg per week, compared to week 1 where the amount of waste was 149.94kg per week. There is only an increase of 8 percent of weight of waste from week 1 until week 20.
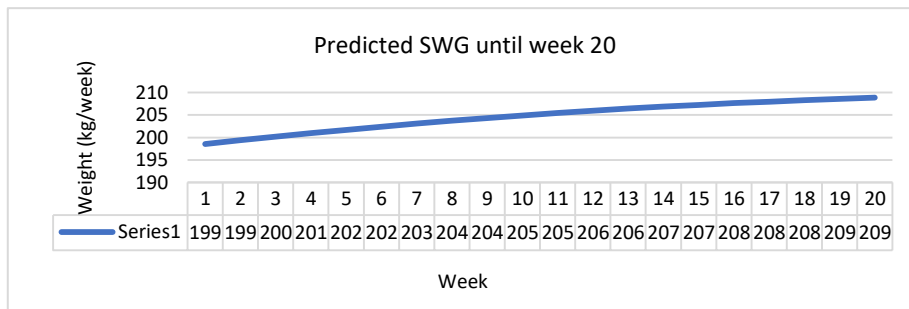

Figure 18. Prediction of SWG in E1 for twenty weeks

As shown in Figure 18, the number of waste that will be generated in E1 in week 20 is 208.87kg per week, compared to week 1 where the amount of waste was 198.54kg per week. There is only an increase of 4.7 percent of weight of waste from week 1 until week 20.
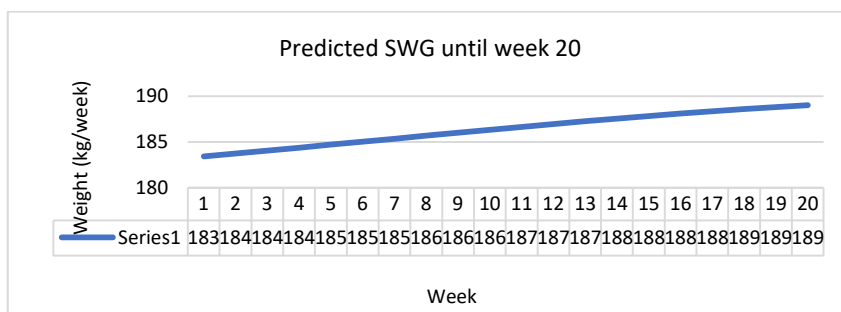

Figure 19. Prediction of SWG in E2 for twenty weeks

As shown in Figure 19, the number of waste that will be generated in E2 in week 20 is 189.05kg per week, compared to week 1 where the amount of waste was 183.42kg per week. There is only an increase of 2.97 percent of weight of waste from week 1 until week 20.
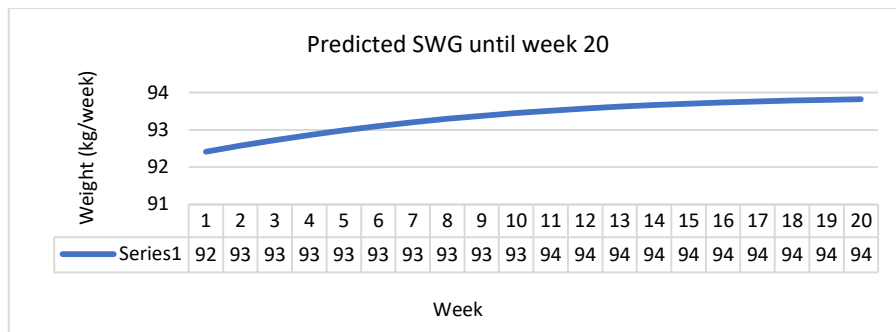

Figure 20. Prediction of SWG in E3 for twenty weeks

As shown in Figure 20, the number of waste that will be generated in E3 in week 20 is 93.82kg per week, compared to week 1 where the amount of waste was 92.42kg per week. There is only an increase of 0.44 percent of weight of waste from week 1 until week 20. The average rate of increment for all household size is 2.05 percent.

On the other hand, Figure 21 will compare different amount of waste produced for different household. Figure 21 shows that different household will produce different amount of waste.
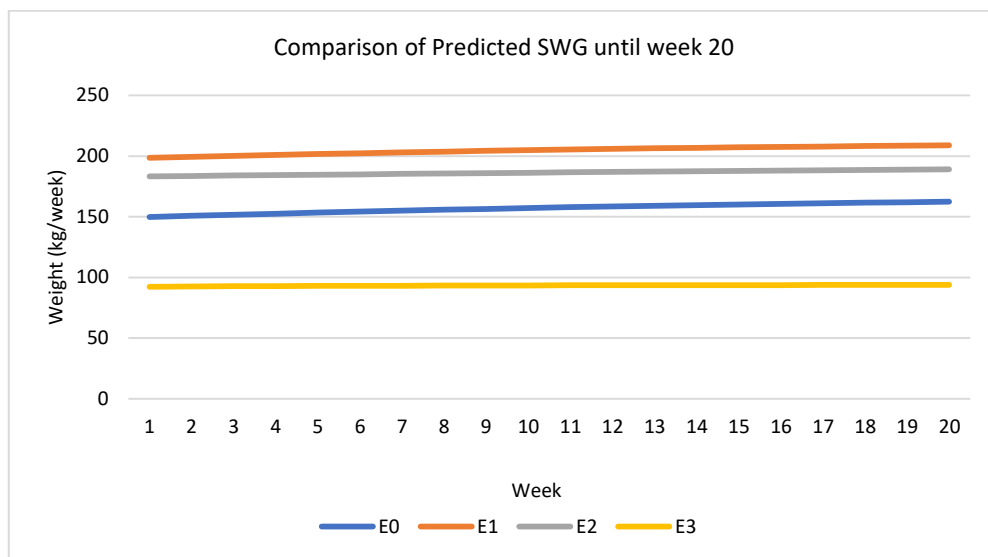

Figure 21. Comparison of prediction of SWG in in four household sizes (E0-E3) for twenty weeks

In relation to the above predicted waste for all household size until week 20, this study deducted that current system of management and collection SWG in KOE, IIUM are needed to be revised time by time. As mentioned in Section 1, this prediction of SWG will help the authorities to help providing an estimation of the prioritized building in KOE. As clearly pictured in Figure 21, the priority of the building will go to E1 which generated highest amount of waste compared to E0, E2 and E3.

7.   CONCLUSION

Poor MSWM will lead to many environmental and health issues such as excessive amount of methane gas production and malaria. Therefore, in this project, prediction algorithms are proposed to provide the forecasted SWG based on household size factor. Prediction algorithm plays a very important role not only to MSWM authorities, but also in handling the waste. This algorithm will provide the management personnel

to have the estimation and how to handle the SWG in the future. Based on the experimentation results, it shows that the objectives of this project have been achieved. The smart waste bin design in Chapter 3 has helped this project in the collection of waste weight in KOE, IIUM. Also, the notification sent by the smart waste bin has helped the waste collector crews to know where the location of the bin that is needed to be emptied. This process has helped the management personnel by not to attend to all bins to check whether it has been full or not. It can save a lot of time and manpower too. Also, if a notification is sent right before the bin is full, this can prevent the waste bin to get overflow before the pick-up schedule and waste bin overflowing's issues can be overcome.

In addition, the result in Section 6 indicated that the prediction of SWG based on household size factor is best suited when ANN is used with two hidden layers where the number of nodes for the first layer is five and the second layer is ten. The result also shows that the prediction algorithm has predicted the average rate of increment of household size factor which is 2.05 percent for the next twenty weeks.

However, the limitation to this study is that the amount of smart waste bin should be replicated more so that all data for waste weight is directly collected from the smart waste bin. On the other hand, room for improvement can always be proposed in any project. Further suggestions can be considered to achieve better results. There are two recommendations that can be considered to improve future work for this project such as the use of other prediction algorithm such as Adaptive Neuro-Fuzzy Inference System (ANFIS) or Nonlinear Autoregressive Network with Exogenous Inputs (NARX) and to consider more SWG factors such as population growth and household income.

REFERENCES
[1]   Saini R, Ahuja N, Bahukhandi K. "Futuristic Projection of Solid Waste Generation in Dehradun City of Uttarakhand using Supervised Artificial Neural Network-Non-Linear Autoregressive Neural Network (NARnet)". *International Journal of ChemTech Research [Internet]*. 2017 [cited 6 July 2018];10(13):283-299. Available from: http://www.sphinxsai.com/2017/ch_vol10_no13/3/(283-299)V10N13CT.pdf
[2]   Pek Mei W. "New app for rubbish collection introduced in Ampang Jaya - Metro News | The Star Online [Internet]". Thestar.com.my. 2017 [cited 10 July 2018]. Available from: https://www.thestar.com.my/metro/community/2017/06/22/new-app-for-rubbish-collection-introduced-in-ampang-jaya/
[3]   Anand R. "Selangor launches app to report uncollected garbage | Malay Mail [Internet]". Themalaymailonline.com. 2016 [cited 10 July 2018]. Available from: http://www.themalaymailonline.com/malaysia/article/selangor-launches-app-to-report-uncollected-garbage.
[4]   Juutinen M. "Overflowing garbage bins: 5 impacts on health and environment, and how to prevent | Ecube Labs [Internet]". *Ecube Labs*. 2016 [cited 10 July 2018]. Available from: http://ecubelabs.com/blog/overflowing-garbage-bins-5-impacts-on-health-and-environment-and-how-to-prevent/
[5]   Chappells H, Shove E. "Bins and the history of waste relations". [Internet]. 2016 [cited 10 July 2018];. Available from: http://www.lancaster.ac.uk/fass/projects/esf/bins.htm
[6]   Marandi F, Ghomi S. "Time series forecasting and analysis of municipal solid waste generation in Tehran city". *2016 12th International Conference on Industrial Engineering (ICIE)*. 2016.
[7]   Younes M, Nopiah Z, Basri N, Basri H, Abushammala M, Maulud K. "Prediction of municipal solid waste generation using nonlinear autoregressive network". *Environmental Monitoring and Assessment*. 2015;187(12).
[8]   Ali Abdoli M, Falah Nezhad M, Salehi Sede R, Behboudian S. "Longterm forecasting of solid waste generation by the artificial neural networks". *Environmental Progress & Sustainable Energy*. 2011;31(4):628-636.
[9]   Hoseini S, Pourahmadi-Nakhli M, Soltani A. "Neural Network Model of Estimation of Body Mass Index Based on Indirect Input Factors". *Indonesian Journal of Electrical Engineering and Computer Science*. 2013;2(3).
[10] Li S. "Forecast Model of Water Quantity Based on Back Propagation Artificial Neural Network". *Indonesian Journal of Electrical Engineering and Computer Science*. 2014;12(6).
[11] Huang H. "A Hybrid Neural Network Prediction Model of Air Ticket Sales". *Indonesian Journal of Electrical Engineering and Computer Science*. 2013;11(11).
[12] Sun N, Chungpaibulpatana S. "Development of an Appropriate Model for Forecasting Municipal Solid Waste Generation in Bangkok". *Energy Procedia*. 2017; 138:907-912.
[13] Mahajan P, Kokane A, Shewale A, Shinde M, Ingale S. "Smart Waste Management System using IoT". *International Journal of Advanced Engineering Research and Science*. 2017;4(4):93-95.