Prediction of Classroom Reverberation Time using Neural Network

Fathin Liyana Zainudin¹,  Abd Kadir Mahamad¹,², Sharifah Saon¹,² and Musli Nizam Yahya³

¹Faculty of Electrical & Electronic Engineering, Universiti Tun Hussein Onn Malaysia
²Internet of Things Focus Group, Faculty of Electrical & Electronic Engineering, Universiti Tun Hussein Onn Malaysia
³Faculty of Mechanical & Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia

Abstract. In this paper, an alternative method for predicting the reverberation time (RT) using neural network (NN) for classroom was designed and explored. Classroom models were created using Google SketchUp software. The NN applied training dataset from the classroom models with RT values that were computed from ODEON 12.10 software. The NN was conducted separately for 500Hz, 1000Hz, and 2000Hz as absorption coefficient that is one of the prominent input variable is frequency dependent. Mean squared error (MSE) and regression (R) values were obtained to examine the NN efficiency. Overall, the NN shows a good result with MSE < 0.005 and R > 0.9. The NN also managed to achieve a percentage of accuracy of 92.53% for 500Hz, 93.66% for 1000Hz, and 93.18% for 2000Hz and thus displays a good and efficient performance. Nevertheless, the optimum RT value is range between 0.75 – 0.9 seconds.

1. Introduction
In initial room architecture design, one of the most important parameter in order to ensure a pleasant room acoustic is the reverberation time (RT). RT is the time for the sound to decay 60dB from the initial value as seen in figure 1.

Achieving the optimal RT in a classroom is important since a bad acoustic can generate a poor working condition to a lesson due to conflicting sounds between the source and its reflecting noises [1]. Optimum RT can ensure good speech intelligibility as it was proven that students showed better results in speech intelligibility tests conducted in classrooms with less noise [2]. This paper is seeking an optional method in estimating RT value that can be up to par with other existing methods.

2. Literature Review
In 1890, Sabine had discovered an easy way to estimate RT value by incorporating the volume of the room (V), the absorption coefficient (α), and the surface areas within the room (S) [3] as seen in equation 1.

\[ RT_{60} = 0.05V \left( \sum S\alpha \right)^{-1} \] (1)

RT₆₀ = 0.05V \left( \sum S\alpha \right)^{-1}
Eyring had come out with a formula in 1930 that compromised for non-diffuse and non-uniform surface absorption coefficient surroundings \[4\]. Still, improvements have been made throughout the years in order to overcome the shortcomings \[5-6\].

Apart from formulaic calculations, other techniques and methods were also being presented such as the computer simulations ray-tracing techniques (ODEON) \[7\] and finite element models (FEM) \[8\]. Since computer simulation and FEM required a lot of time in the initial model designing phase, NN has been another option in estimating the RT value \[9\].

Researches on NN for RT prediction were executed since 1999 and pioneered by Nannarielo and Fricke \[10\] where they examined NN using dataset from actual building to predict RT in large halls. This research has proven that NN is capable in predicting room RT. In 2010, Yahya et al. \[11\] continued the research by executing experiments to predict RT for classrooms using dataset gathered from FEM computation. Aliabadi et al. in 2014 \[12\] then developed an empirical acoustic model in predicting RT for industrial workrooms using NN where they were able to strengthen the potential of NN method in minimizing the uncertainties in acoustics' modeling. This research conducted in designing a classroom RT prediction method using NN with dataset that were computed from ODEON software.

2.1. Neural Network Application

NN was founded by McCulloch and Pitts in 1943 that loosely imitates the work of the axons that transmit and translate information in human brain \[13\]. In supervised learning method, the NN need to learn and memorize certain input patterns and finally it should be able to translate and compute the inputs provided. Thus, the training dataset (that pairs the inputs with its output) is essential in order for the NN to familiarize with the input patterns and capable in distinguishing the final output.

A multilayer perceptron (MLP) is a feed forward neural network (FFNN) that computes a set of inputs to its appropriate output \[14\]. MLP consists of 3 main layers: input, hidden nodes (can be more than 1), and output that are connected to one another. Figure 2 shows the architecture of a MLP with a single hidden layer. MLP is a fully connected network where each node in each layer connects with a particular weight to every node in the following layer.

**Figure 1**: Reverberation Time.

**Figure 2**: MLP model architecture with a single hidden layer.

3. Data Collection

Classrooms models with different heights, widths, and lengths were created using Google SketchUp software. Windows were also added to some of the models. For this research, only shoebox shaped room were designed. Figure 3 shows the sample model that was created with Google SketchUp software. ODEON 12.10 software was used to compute the output dataset. Each room design was assigned with a different combination of material types for the room surface making a total of 600 data in the dataset. Ray tracing method is applied by ODEON software to calculate the room RT as seen in figure 4.
4. Neural Network Training

In this research, the NN system was trained and analyzed on 500Hz, 1000Hz and 2000Hz frequency. A total of 13 variables that were applied as the NN system inputs, which were the room volume \((V)\), length \((L)\), width \((W)\), height \((H)\), equivalent absorption coefficient of wall1 area \((S_{\alpha w1})\), equivalent absorption coefficient of wall2 area \((S_{\alpha w2})\), equivalent absorption coefficient of floor area \((S_{\alpha f})\), equivalent absorption coefficient of door area \((S_{\alpha d})\), equivalent absorption coefficient of ceiling area \((S_{\alpha c})\), equivalent absorption coefficient of window area \((S_{\alpha w})\), and sound source position \((x/L, y/W, z/H)\).

Figure 5 shows a standard shoebox model that displays the \(L, W, H\), as well as \(x, y, z\) starting point for sound source placement.
Figure 5. Length ($L$), Width ($W$), Height ($H$), as well as $x$, $y$, RT starting point for sound source placement in a classroom.

Table 1. Room geometrical characteristics for NN training dataset.

|               | Max     | Min     | Mean   | Standard Deviation |
|---------------|---------|---------|--------|--------------------|
| $V$ ($m^3$)   | 1307.32 | 209.17  | 369.60 | 255.72             |
| $H$ ($m$)     | 4.20    | 2.60    | 3      | 0.31               |
| $L$ ($m$)     | 18.44   | 9.50    | 11.40  | 2.78               |
| $W$ ($m$)     | 23.39   | 7.16    | 10     | 4.63               |
| $S_{w1}(500)$ | 123.56  | 0       | 8.73   | 20.76              |
| $S_{w2}(500)$ | 135     | 0       | 16.98  | 25.41              |
| $S_{g}(500)$  | 216.24  | 2.39    | 18.19  | 41.83              |
| $S_{door}(1000)$ | 10.94   | 0.18    | 1.08   | 2.82               |
| $S_{ceil}(1000)$ | 212.91  | 31.93   | 71.89  | 44.81              |
| $S_{win}(1000)$ | 18.87   | 0       | 2.84   | 4.28               |
| $S_{ceil}(1000)$ | 94.93   | 0       | 11.12  | 17.44              |
| $S_{door}(1000)$ | 88.99   | 0       | 10.68  | 14.90              |
| $S_{win}(1000)$ | 232.87  | 2.39    | 33.19  | 45.12              |
| $S_{ceil}(1000)$ | 8.81    | 0.18    | 0.83   | 2.08               |
| $S_{win}(1000)$ | 249.51  | 27.94   | 85.12  | 58.16              |
| $x/L$         | 0.892   | 0.179   | 0.488  | 0.212              |
| $y/W$         | 0.819   | 0.036   | 0.559  | 0.207              |
| $z/H$         | 0.803   | 0.122   | 0.47   | 0.17               |

Table 1 shows the input dataset geometrical characteristics for the NN system development. All data in database should be normalized before being implemented to the NN to minimize the data redundancy and dependency without losing information. For this paper, all data were normalized to the range of 0.1 to 0.9 using equation 2,

$$X' = 0.1 + 0.8 \left( X - X_{\text{min}} \right) \left( X_{\text{max}} - X_{\text{min}} \right)^{-1}$$  \hspace{1cm} (2)

where; $X$ is the data of the original database, $X'$ is the data to the transformed database, $X_{\text{min}}$ and $X_{\text{max}}$ are the minimum and maximum of the original database, respectively.

The datasets were then arranged randomly and then divided into 3 sets; 60% for training, 20% for validation, and the other 20% for testing. Validation data is important to make sure overfitting did not occur. Testing data used only for testing the final solution in order to confirm the predictive power of
the network. Training data must be more than validation and testing data to ensure that each type of texture is trained. Matlab NN toolbox was used to train the collected data. In this research, the NN system for 500Hz, 1000Hz, and 2000Hz were run and trained separately as absorption coefficients are frequency dependent. The mean squared error (MSE) and the regression ($R$) value between the obtained output and the desired output were observed. The MSE is the parameter that is used to estimates and measures the error between the obtained and the desired output while $R$ is the parameter that measures the correlation between both outputs. The obtained output is considered to be good fit if $R > 0.9$ [15]. Equation (3), (4), and (5) are the formula for MSE, $R$, and percentage of accuracy with $e_i$ is the error, $t_i$ is the desired value, $y_i$ is the predicted value, $t_{mean}$ and $y_{mean}$ are the mean values, and $N$ is a number of data.

$$\text{MSE} = \sum_{i=1}^{N} (e_i)^2 = \sum_{i=1}^{N} (t_i - y_i)^2$$  \hspace{1cm} (3)

$$R = \left( \frac{\sum_{i=1}^{N} (t_i - t_{mean})(y_i - y_{mean})}{\sqrt{\sum_{i=1}^{N} (t_i - t_{mean})^2 \sum_{i=1}^{N} (y_i - y_{mean})^2}} \right)^{1/2}$$  \hspace{1cm} (4)

$$\text{acc (\%)} = \frac{\text{Image correctly classified}}{\text{Total image}} \times 100\%$$  \hspace{1cm} (5)

5. Results and Discussion

Table 2 shows the MSE for data training, validation, and testing. The Matlab FFNN training performances can be seen in figure 6, which observes the MSE values for each epoch. NN will continue to train and updates the weight values until the MSE value for the validation data increases in comparison to the MSE value from the previous epoch in order to prevent overfitting occurs. Overfitting occurs when the NN is able to learn the complicated relationship between inputs and outputs but limited to the training data only and resulting to NN unfamiliarity when new data is introduced.

Table 3 shows the $R$ values for data training, validation, and testing. The $R$ values obtained indicate that the training, validation, and testing data for all networks having a good fit as $R > 0.9$. The regression plots for testing data for 500Hz, 1000Hz and 2000Hz, respectively can be seen in figure 7.

| Frequency | Train | Validation | Testing |
|-----------|-------|------------|---------|
| 500Hz     | 0.0014| 0.0036     | 0.0016  |
| 1000Hz    | 0.0005| 0.0011     | 0.0012  |
| 2000Hz    | 0.0010| 0.0022     | 0.0024  |

| Frequency | Train | Validation | Testing |
|-----------|-------|------------|---------|
| 500Hz     | 0.96351| 0.90915    | 0.95497 |
| 1000Hz    | 0.98259| 0.95889    | 0.96393 |
| 2000Hz    | 0.97245| 0.93408    | 0.94281 |
Figure 6. Matlab FFNN Training performances for (a) 500Hz, (b) 1000Hz, and (c) 2000Hz.

Figure 7. Regression plot scatter for (a) 500Hz, (b) 1000Hz, and (c) 2000Hz.

Table 4 shows the testing data percentage accuracy compared to the ODEON 12.10 computation. Overall, a good percentage of accuracy has been achieved with 92.53% for 500Hz, 93.66% for 1000Hz, and 93.18% for 2000Hz. This result shows the accuracy for testing data percentage is good since achieving more than 90%. While Table 5 shows the validation of average reverberation time in seconds (s) from the actual measurements and ODEON. The result from table 5 had confirmed that the value of
classroom RT is acceptable since the differences between the actual RT through measurement and through simulation using ODEON software is small, which is within 0.01 – 0.05.

Table 4. Testing data percentage accuracy

| Frequency | % accuracy |
|-----------|------------|
| 500Hz     | 92.53      |
| 1000Hz    | 93.66      |
| 2000Hz    | 93.18      |

Table 5. Reverberation Time from Actual Measurement and ODEON

| Frequency | Actual RT (s) | ODEON (s) |
|-----------|--------------|-----------|
| 500Hz     | 0.77         | 0.75      |
| 1000Hz    | 0.87         | 0.90      |
| 2000Hz    | 0.90         | 0.89      |

6. Conclusion

The NN system shows a good result with 92.53%, 93.66%, and 93.18% accuracy for 500Hz, 1000Hz, and 2000Hz respectively. A low MSE values were also achieved with 0.0016 for 500Hz, 0.0012 for 1000Hz, and 0.0024 for 2000Hz as well as R value > 0.9 for all that indicates a good fit. While, the RT value for a classroom is range between 0.75 – 0.9 seconds, which consider an optimum RT value. In conclusion, the NN system is proven to be a good alternate method for classroom RT prediction.

In the future, validations with actual classroom measurement need to be inspected in order to check the NN system accuracy and efficiency in the real environment.

References

[1] Rasmussen B, Brunskog J, and Hoffmeyer D 2012 Reverberation time in class rooms – Comparison of regulations and classification criteria in the Nordic countries Proceedings of Joint Baltic-Nordic Acoustics Meeting (Nordic Acoustic Association)

[2] Rabelo A T V, Santos J N, Oliveira R C and Magalhães M D C 2014 Effect of classroom acoustics on the speech intelligibility of students CoDAS vol 26 no 5 (Sociedade Brasileira de Fonoaudiologia) pp 360-366

[3] Sabine W C Collected Papers on Acoustics (Peninsula, Los Altos CA) pp 1900-1915

[4] Eyring C F 1930 Reverberation time in “dead” rooms J. Acoust. Soc. Am. vol 1(2A) (Acoustical Society of America) pp 217-241

[5] Fitzroy D 1959 Reverberation formulae which seems to be more accurate with non-uniform distribution of absorption J. Acoust. Soc. Am. vol 31 (Acoustical Society of America) pp 893-897

[6] Neubauer R O 2001 Estimation of reverberation time in rectangular rooms with non-uniformly distributed absorption using a modified Fitzroy equation Building Acoustics vol 8 no 2 (SAGE Journals) pp 115-137

[7] Hamidon M S, Zainudin F L, Mahamad A K, and Saon 2016 Simulation of acoustically demanding space by using Odeon software ARPN J. Eng. And App. Sciences (Asian Research Publishing Network) pp 5102

[8] Easwaran V and Craggs A 1996 An application of acoustic finite element models to finding the reverberation times of irregular rooms Acta Acustica united with Acustica vol 82 no 1 (European Acoustics Association) pp 54-64

[9] Cox T J, Li F and Darlington P 2001 Extracting room reverberation time from speech using artificial neural networks Journal of the Audio Engineering Society vol 49 no 4 (Audio Engineering Society) pp 219-230
[10] Nannariello J and Fricke F 1999 The prediction of reverberation time using neural network analysis Applied Acoustics vol 58 no 3 (Elsevier) pp 305-325

[11] Yahya M N, Otsuru T, Tomiku R and Okozono T 2010 Investigation the capability of neural network in predicting reverberation time on classroom International Journal of Sustainable Construction Engineering and Technology vol 1 no 1 (UTHM) pp 1-13

[12] Aliabadi M Golmohammadi R, Ohadi A, Mansoorizadeh M, Khotanlou H and Sarrafa Zadeh M S 2014 Development of an empirical acoustic model for predicting reverberation time in typical industrial workrooms using artificial Neural Networks Acta Acustica united with Acustica vol 100 no 6 (European Acoustics Association) pp 1090-1097

[13] Funahashi K I 1989 On the approximate realization of continuous mappings by neural networks Neural networks vol 2 no 3 (Elsevier) pp 183-192

[14] Rumelhart D E, Hinton G E and Williams R J 1986 Learning internal representations by error propagation (No. ICS-8506) (California Univ San Diego La Jolla Inst for Cognitive Science)

[15] Sivanandam S N and Sumathi S 2005 Introduction to Neural Networks Using MATLAB 6.0 (Tata McGraw-Hill) ISBN: 978-0-07-059112-7

Acknowledgments
The financial support received from the Office for Research, Innovation, Commercialization and Consultancy Management (ORICC), University Tun Hussein Onn Malaysia (UTHM) is gratefully acknowledged.