Investigation of Speech Intelligibility Using Artificial Neural Network Model

Dina Harith Shaker*,M.Sc(Asst.Lecturer)

Abstract

A classroom acoustic is an important and difficult part of university classroom design. Good design is achieved more on the basis of acoustics expertise than on pure engineering design. In this paper, the Artificial Neural Network (ANN) model is used for predicting speech intelligibility in classroom. There are several classroom properties such as diminution of the class, signal to noise ratio (SNR), the location of the student and teacher , background noise where collected from the classroom. A set of word is complied and a speech signal data base was created. The sound pressure levels are then measured using sound pressure meter at different classroom positions. A datasheet was obtained from the measurement and then used to provide as training database into learning process of (ANN) to predict the speech intelligibility at various listeners' position of classroom. This method improve high accuracy, efficiency and economic of calculation intelligibility in classrooms. Therefore it reduces the error by using the classic methods.

Keywords: Acoustics, Noise, sound pressure level, speech intelligibility, and Artificial Neural Network (ANN) model.

* University of Technology

Introduction

The investigation of good teaching in the classroom is a part and content in the administration of teaching in universities and school classrooms. It is recognized as the quality of teaching will affect on the learning process and training. it is an important tasks to find how teachers evaluating' teaching quality, scientifically and objectively, how to reduced the difficult in the educational process[1].

There are many difficulties in the teaching quality evaluation on account to that the teaching quality in the classroom is depends on the various factors [2].

Such as intelligibility of speech, the level of background noise, no need to speak in a raised voice, the comfort teaching and learning [3].

Therefore the universities always have been attempt to solve the problem and find to how correct the reflect factors for quality of teaching in the classrooms and provide the targeted by teachers in the classrooms to developed the methods of teaching quality in the classrooms.. Paulraj M P [4] present a simple system for classrooms speech intelligibility prediction, many classrooms characteristics such as size, signal- to – noise ratio and speech transmission index were collected from classrooms in university of Malaysia Perlis, Elman network was selected to develop the system robustness in prediction application for speech intelligibility and the network performances was dependent to the normalization method.

Musli Nizam [5] develop an Artificial Neural Network (ANN) model for predicting reverberation times in classrooms. The simulation system is developed at Oita University, to achieve that Finite Element Method (FEM) is used to simulation database and fed into Artificial Neural Network (ANN) for learning process.

Rabelo ATV, Santos JN, Oliveira RC, Magalhaes MC [6] analyzed the acoustic performance of 18 classrooms and measured the following parameters: pressure level, reverberation time, speech transmission index and speech intelligibility. The results of speech intelligibility test were compared to the values of the acoustic parameters with the help of students test. In the conclusion of this study shows that the acoustic parameters have a direct effect on the speech intelligibility of students.

To answer on many questions in the acoustical design of classrooms, together with practice for quality of teaching evaluation, were consider in this work them by means of including the evaluation system by using artificial neural network (ANN). To achieve that, it was used measurement method to measure sound pressure level in classroom and used it to simulated a database and fed into (ANN) for learning process. The use of artificial neural networks obviates the need to specify, calculate and measure acoustic quantities. The acoustics of classroom depends on several parameters such as the size of the class, the shape and surface finishes in it, therefore to subjective acoustic ratings have been obtained, then can be used the artificial neural networks to predict how well a new hall will be perceived.

Speech intelligibility

In university classrooms, the major acoustical concern is that of verbal communication. Inadequate acoustical condition, resulting in poor verbal communication, causes two main problems. First, they lead to reduced learning efficiency. Second, they can lead to fatigue, health and stress problems, amongst lecturers, were there are forced to reparation for poor conditions of acoustic and by raising their voices, for example.

The good quality of communication can be quantified by speech intelligibility (SI). This quantity is the percentage of speech material which is correctly understood by the average listener. It has been suggested that, in the case of normal – hearing adults working in their first language, the speech intelligibility should exceed 97% [7]. In the case of acoustically-challenged people, such as hard of hearing students, students working in second language and children, the requirements are undoubtedly more stringent; Bradley suggests aiming for 100%[8].

The speech intelligibility can be determined by several methods such as Signal- to Noise Ratio (SNR), Direct to Reverberant ratio (DRR).In this research, Were used (SNR) to infer the amount of speech intelligibility in classroom. Where, (SNR) is equal to the level of speech (SL) in dB minus the level of background noise (BGN) in dB both in listener position. The level of speech depends on the speakers' voice level, the distance between the listener and the speaker in the classroom. The (BGN) results from the ventilation system, projectors, in class student activity and the noise comes from outside the classroom [9].

Signal-to-Noise Ratio (SNR) measurement

The method is based on the Signal-to-Noise Ratio (SNR) on a specific location, usually the location of the student with the hearing impairment.

The signal is the speech of lecturer and therefore the measurement can be performed during classroom hours. To identify the noise level the sound pressure level before and after the lecturer talking is used. In this way a SNR is obtained that is actually present during the speech signal. This is an advantage over using an average noise level. The identify when the lecturer is talking a sound pressure meter is placed at the student place. The sound pressure level will be much higher when the lecturer is talking compared to silent period. The measurements need to be performed for at least twenty minutes to get enough samples to make an accurate calculation of Signal-to-Noise Ratio (SNR).[10,11]
The difference in the sound level at recovers' location between lecturers is a quite (noise) and lecturer talk (signal) determines the Signal-to-Noise Ratio (SNR) and it is given by Equations below[12]:

$$SNR_{being} = p_g - p_{r1} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$
$$SNR_{end} = p_g - p_{r2} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$
$$SNR - {}^{1}/_{2}\ [2p_g - p_{r1} - p_{r2}] \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Where:

- p_g is the average sound level during the signal.
- p_{r1} is the average sound level before the signal.
- p_{r2} is the average sound level after the signal.

Artificial Neural Network (ANN) model

The Artificial Neural Network (ANN) is a model worked by using multilayer perception input layer (MLP), many number of hidden layer and one output layer, where each of input layer is called neuron, were these layer connected line between each neuron with neurons in the up layer of, the same neurons layer is not connected, according to function x to mutual influence between neurons. The activation function is selected as (S) functions [13]. it is expressed in Equations (4 and 5).

$$f(x) = 1/1 + \exp(-ax) \qquad \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

Also has other function, such as:

$$f(x) = \tan\left(\frac{x}{2}\right) = [\,1 - \exp(-ax)]/[1 + \exp(-ax)] \ldots\ldots\ldots\ldots\ldots\ldots (5)$$

the network structure is represented by Schematic diagram shown in Figure (1)[14].

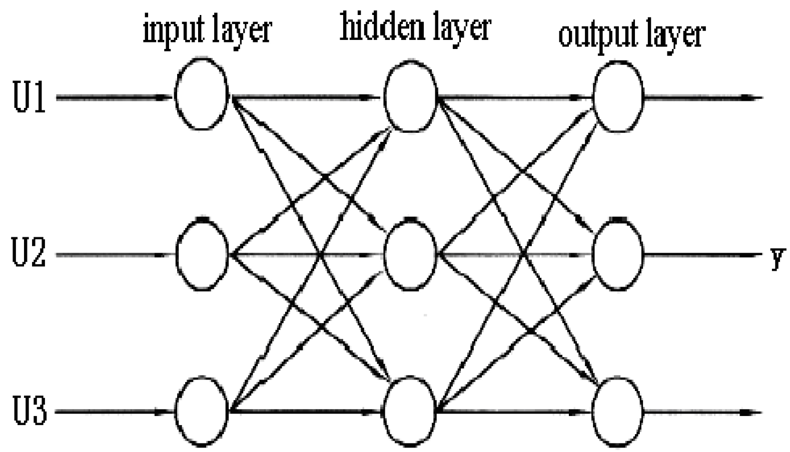


Figure (1) : Artificial Neural Network Model construction**.**

Proposed model

The proposed model divided into three stages :

- Measurement the sound pressure level in select location in the classroom.
- Calculate the Speech Intelligibility (SI) in the classroom by using Signal-to-Noise Ratio (SNR) equation.
- The intelligent stage.

The three stage operate together to identifies weak of the (SI) location in classroom.The width, length and height of classroom are measured and the listener positions in the classroom are identified. The distance between any two adjacent listener positions in the same row is 0.5 m (D_1) and the distance between each row is 1m (D_2). The classroom basic plan is shown in Figure (2).

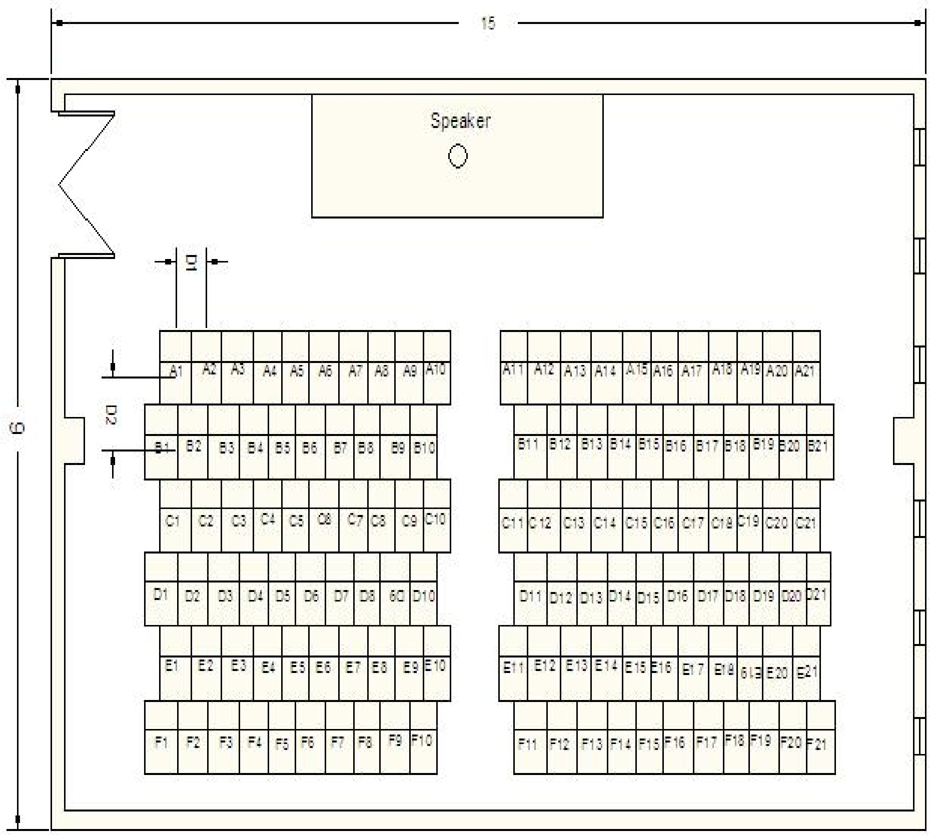


Figure (2): The basic plan of classroom**.**

The Omni directional source is placed at the source position representing a lecturer at a high of 1.5m from the flower level. The source is placed at 1m from the floor level at each listeners seating position representing the human ear. Various pre-recorded words formats are played at various speech and noise levels.

Arabic and English word are recorded at different range of speech levels (from 47dB and 72dB) and range of noise levels is between (54dB to 70dB) are played back by using Omni directional speaker. The sound pressure levels are measured by using sound pressure meter shown in Figure (3) in each selected location.

Figure (3): Sound pressure level meter**.**

The experimented measurement sound pressure levels and calculated SNR are shown in Table (1) which is represented the speech intelligibility of sound in classroom.
The network was learned by a training set represent the sound pressure level and the position of the listener.

The learning rate and momentum are the most important parameters in the Multilayer Perceptron (MLP) algorithms. The learning rate varied during the learning process while the momentum value adjusted at (0.9).

The input layer consist of (264) neurons. The hidden layer contains (33) neurons were selected by try and error procedure. The try and error procedure stopped when the network learned to satisfy the desire output as shown in Figure (4) where it represent the Sun Square Network error (SSN) for Artificial Neural Network (ANN) network . The output layer consists of (1).

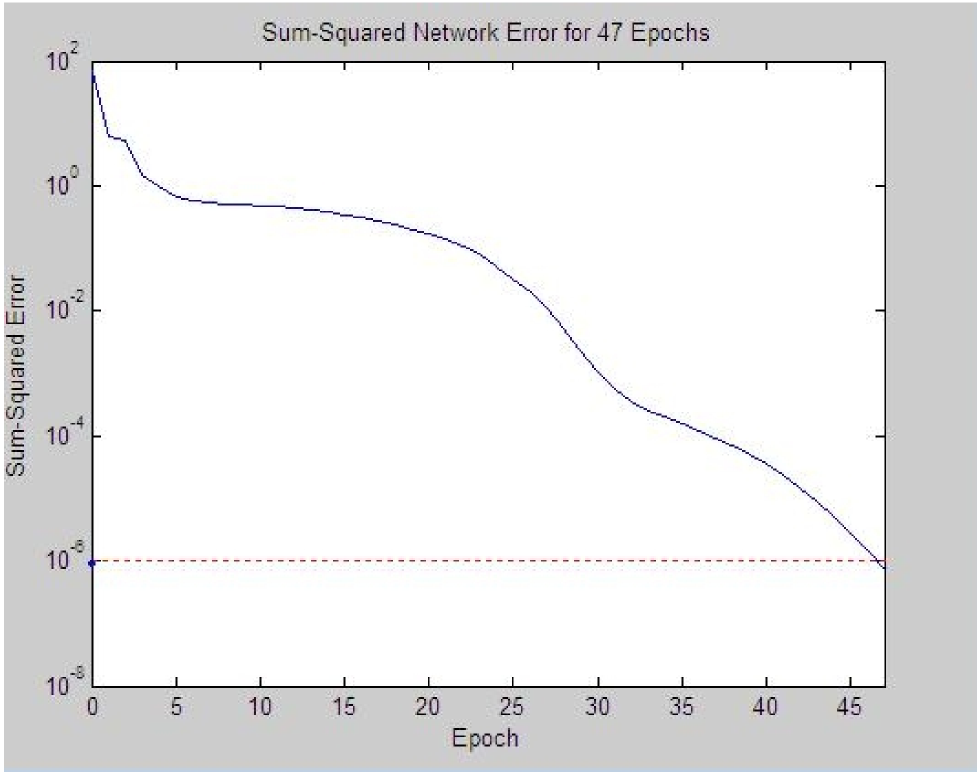


Figure (4): The Sum Square Network error (SSN) for (ANN) network.

All data (input and output matrix used by neural network algorithms were normalized to values between (0-1) by using Equation (6) for more accuracy and fast convergence.

$$Normalized\,value = \frac{y - y_{min}}{y_{max} - y_{min}} \quad \dots\dots\dots\dots\dots\dots\dots (6)$$

Using MATLAB 7 to build the proposed model. The model flowchart is shown in Figure (5).

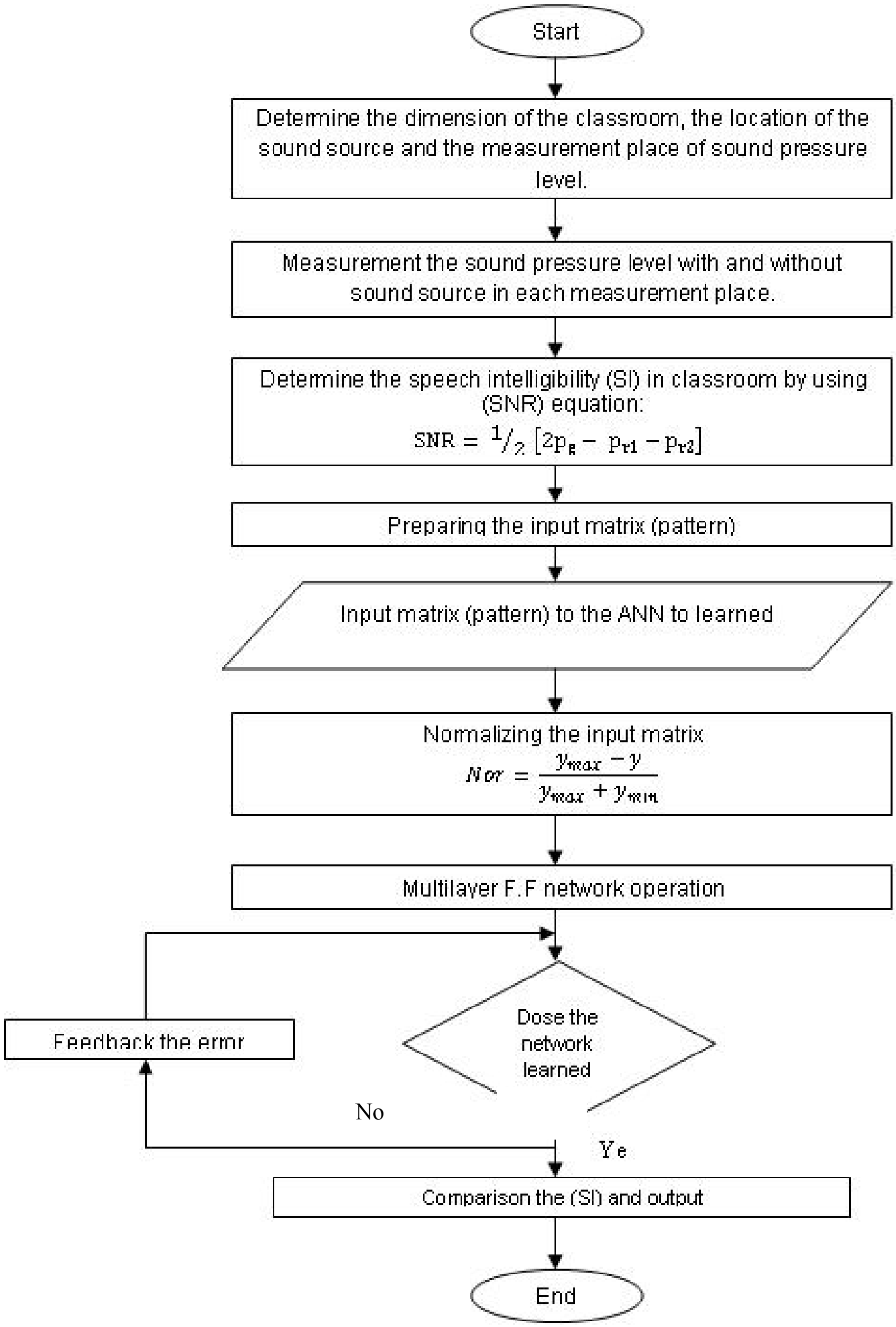


Figure (5): The flow chart of the proposed model.

Results and Discussions

The speech intelligibility level at various listeners' positions in the classroom is determined using the artificial neural network model. The resulting speech intelligibility level at different location is represented in Figures (6), (7) and (8). From this figures it can be inferred that as a Signal- to- Noise Ratio (SNR) increased, the speech intelligibility increased.

Figure (6) represent the practical result of speech intelligibility in each row of seats in the classroom. Where the speech intelligibility be better in the front seats than that of the rear seats because its proximity to the direct sound as noise goes down with increasing distance, Where a good Signal- to- Noise Ratio (SNR) should be not less than +15dB, that is to say the signal strength should be at least 15dB above the background noise level in order to achieve good speech intelligibility by the listener [13].

The better speech intelligibility is in seats away from windows (away from noise source) as well as to the presence of reflective walls that provide early reflections which serve to strength the direct sound signal and lost it cause weakness in the speech intelligibility. Figure (7) show the speech intelligibility distribution in all raw of classroom.

Figure (8) represent the contour model of classroom speech intelligibility simulating results. That as attempt to detect the week and strength of speech intelligibility place in classroom.

Where a dark area represents good speech intelligibility, while the pastel area represents a bad speech intelligibility. It can be observed that the back side of the classroom shows the poor speech intelligibility this is because they fare from the sound source and the seats in the far side from the windows is better speech intelligibility than the other side because it far from the noise source.

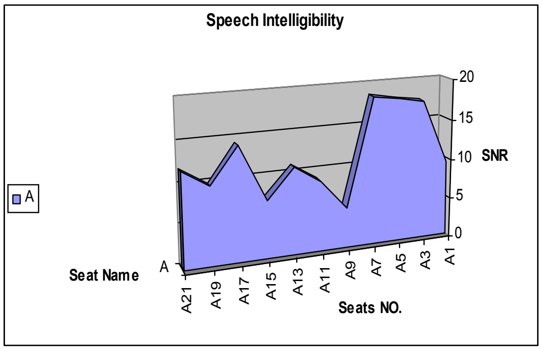
Conclusion

This paper shows that a simple system to predict the speech intelligibility in classroom has been developed using Artificial Neural Network (AAN) model to represent classroom speech intelligibility it is provide a high accuracy method, simple process calculation, it make the work more efficiency and economic benefits. In the same time it reduces the error when uses the manual calculation caused.

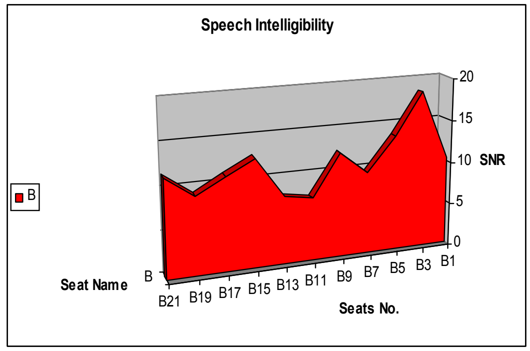
Table (1): The measurement sound pressure level and Signal- to Noise Ratio

The sound pressure level measurement				
SNR	**pr2**	**pr1**	**pg**	**location set**
9.45	65	54.1	69	A1
17.45	68	51.7	77.3	A3
17.95	67.2	63.7	83.4	A5
18.5	67.7	64.5	84.6	A7
5.1	72.8	56.6	69.8	A9
8.8	70.1	54.5	71.1	A11
10.9	68.1	60.1	75	A13
7.25	72.5	63.2	75.1	A15
14.15	64.1	59	75.7	A17
9.75	70.1	60.6	75.1	A19
11.8	65.6	60.2	74.7	A21
10.5	66.6	64.8	76.2	B1
18.5	68.4	50.8	78.1	B3
13.65	75.1	50.4	76.4	B5
9.65	72.6	66.9	79.4	B7
12.25	63.2	61.3	74.5	B9
7.35	70.6	65.5	75.4	B11
7.95	68.9	61.6	73.2	B13
12.7	55	52	66.2	B15
11.05	65.9	62.6	75.3	B17
9.2	70.6	57.2	73.1	B19
11.6	66	60.2	74.7	B21
8.05	73.6	61.5	75.6	C1
15.45	72.6	52.1	78.3	C3
15.85	72.7	53.4	78.9	C5
12.95	67.7	50	71.8	C7
13.85	68.9	65.2	80.9	C9
13.2	71.2	59.2	78.4	C11
9.35	73.8	64.3	78.4	C13
7.1	69.4	57.8	70.7	C15
10.35	69.3	61	75.5	C17
11.75	71.3	61.4	78.1	C19
10.65	69.7	57.6	74.3	C21
6.15	62.3	58.4	66.5	D1
18.25	62.5	63.2	81.1	D3
21.85	63.5	53.4	80.3	D5
13.45	68.7	60	77.8	D7

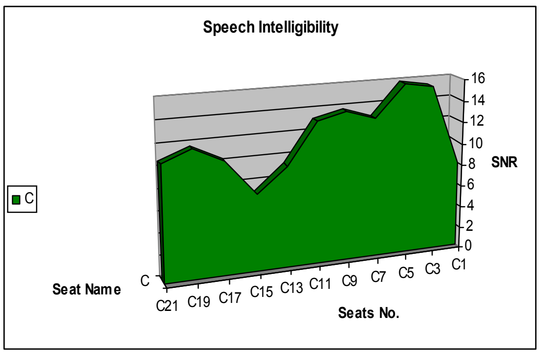
The sound pressure level measurement				
SNR	**pr2**	**pr1**	**pg**	**location set**
18.75	64.6	55.9	79	D9
6.1	69.8	66.4	74.9	D11
9.2	64.7	54.5	69.2	D13
7.05	69.8	59.1	71.5	D15
17.7	65.2	57.4	79	D17
7.65	70.7	61.4	73.7	D19
7.3	64.3	64.9	71.9	D21
15.75	62.3	57.2	74.5	E1
13.8	61.2	60.8	74.8	E3
13.6	63.5	62.3	76.5	E5
17.45	62.7	62	79.8	E7
6.85	67.3	62.4	71.7	E9
6.7	63.4	56.6	66.7	E11
14.2	68.3	55.1	75.9	E13
6.15	61.2	53.9	63.7	E15
15.3	71.2	53.4	77.6	E17
8.5	67.9	57	75.2	E19
10.25	64.1	55.6	70.1	E21
9.2	63.1	63.7	72.6	F1
16.65	62.1	53.8	74.6	F3
20.3	63.1	53.1	78.4	F5
14.7	66.7	53.5	74.8	F7
5.3	68.4	65.8	72.4	F9
5.15	61.4	60.1	65.9	F11
3.25	55.9	57.8	60.1	F13
13.35	62.1	62	75.4	F15
8.7	64.3	57.9	69.8	F17
15.55	66.9	58.4	78.2	F19
14.75	69.1	52.4	75.5	F21



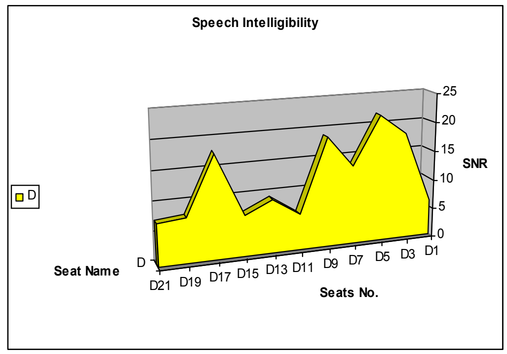
(A) speech intelligibility in the raw A



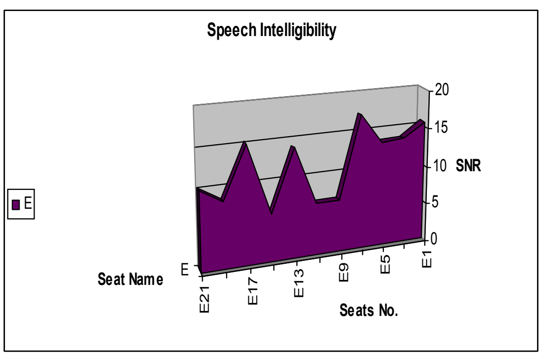
(B) speech intelligibility in the raw B



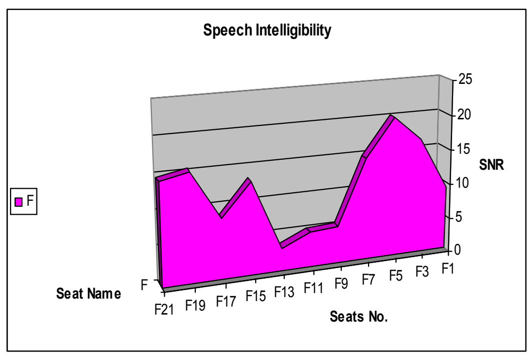
(C) speech intelligibility in the raw C



(D) speech intelligibility in the raw D



(E) speech intelligibility in the raw E



(F) speech intelligibility in the raw F

Figure (6): The practical result of speech intelligibility in each row of seats in the classroom.

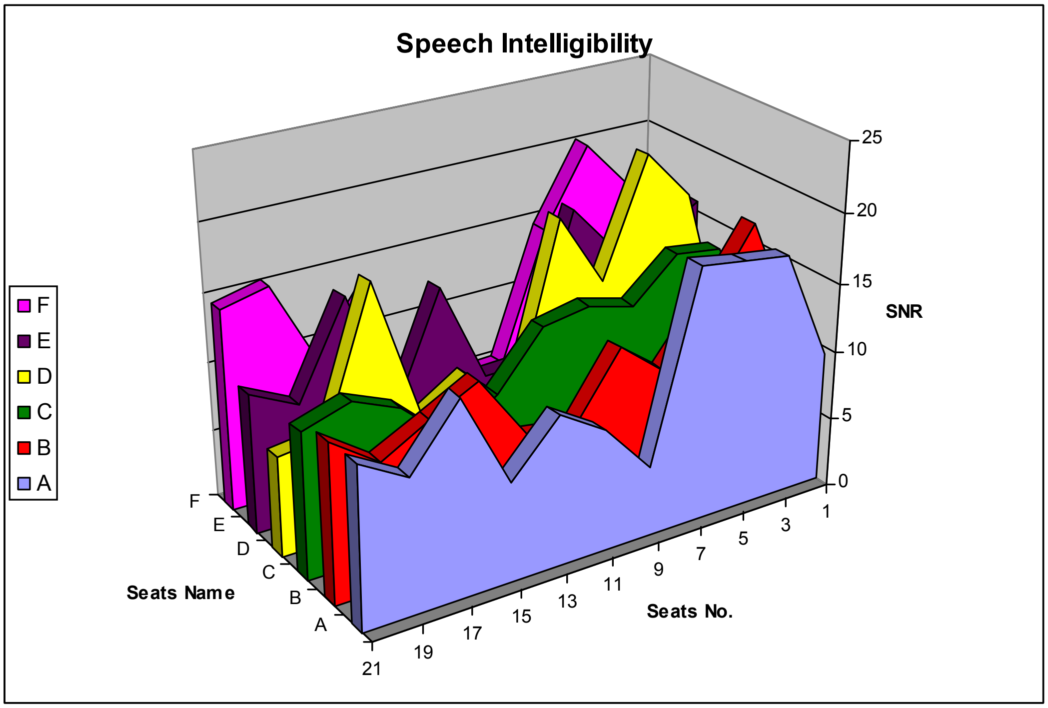
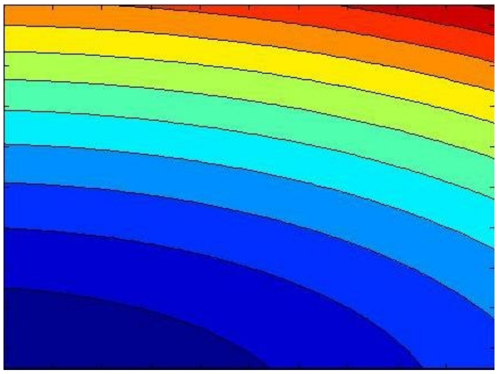
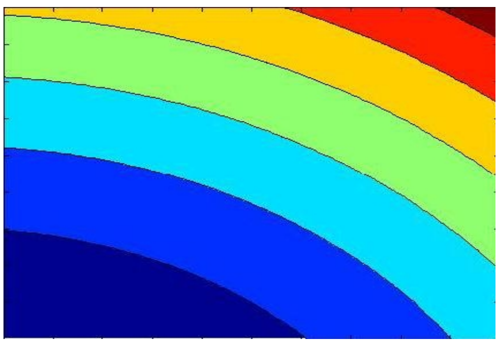


Figure (7): The speech intelligibility distribution in all raw of classroom.

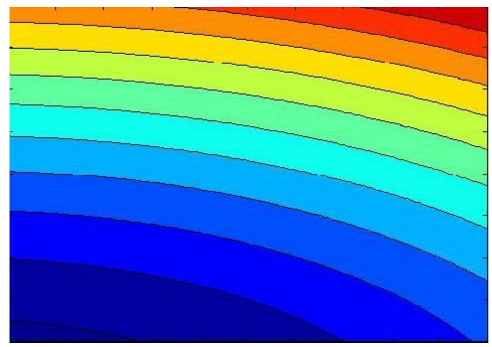


(A) speech intelligibility in the raw A

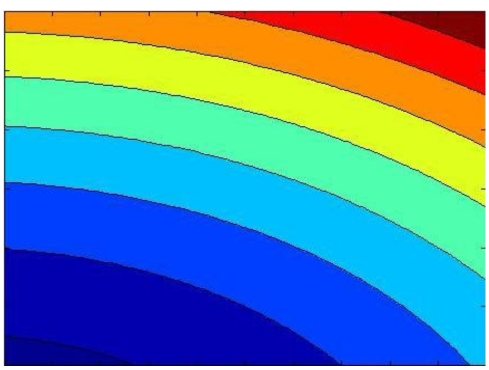
(B) speech intelligibility in the raw B



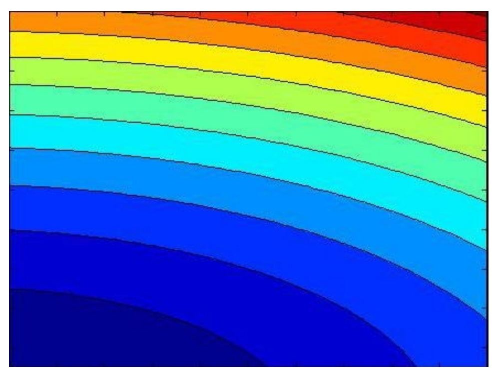
(C) speech intelligibility in the raw C



(D) speech intelligibility in the raw D



(E) speech intelligibility in the raw E



(F) speech intelligibility in the raw F

Figure (8): The contour model of classroom speech intelligibility simulating results.

References:

[1] Qi Yan-ming, Wang Chang-long, Yang Kai," Evaluation of classroom teaching quality in universities based on artificial neural network", IILTA International Conference on Control, Automation and Engineering. IEEE,2009.

[2] Jan Radosz," Global Index of the Acoustic Quality of Classroom, "Archives of Acoustics, Vol.38, No. 2, PP. 159-168, 2013.

[3] Fergus Fricke, Chan Hoon Haan, "Predicting the Acoustics of Concert Halls Using an Artificial Neural Network" Acoustic Australia, Vol. 23, No. 3-87,1995.

[4] Paulraj M P,Sazali Yaacob, A. Nazri Abdullah, M. Thagirarani, M. Ridhwan Tamjis "Classroom Speech Intelligibility Prediction Using Elman Neural Network".6th International Colloquium on Signal processing and Applications (CSPA),IEEE,2010.

[5] Musli Nizam Bin Yahya, Tour Otsuru, Reiji Tomiku, Takeshi Okuzuno and Toshiki Naqqsawa," Reverberation times Prediction on Classroom Using Neural Networks Model", Proceeding of 20th International Congress on Acoustics, ICA, Sydney, Australia, 2010.

[6] Rabelo ATV, Santos JN, Oliveira RC, Magalhaes MC,"Effect of Classroom Acoustics on the Speech Intelligibility of Students",Original Article,CoDAS;26(5):360-6, 2014.

[7] H.G.Latham "The Signal – to – Noise Ratio for Speech Intelligibility- an Auditorium Acoustics Design Index", Applied Acoustics 12, 253-320 (1979).

[8] J.S. Bradley "Speech Intelligibility Studies in Classrooms", J. Acouct. SOC. AM. ,80(3), 846-854(1986).

[9] Murray Hodgson "UBC-Classroom Acoustical Aurvey " Canadian acoustics- acoustique canadienne , 22(4) 3-10 (1994).

[10] Hongisto, V.,"AModel Predicting the Effect of Speech of Varying Intelligibility on Work Performance", Blackwell Mnuksgaard, 2005.

[11] AnsI, "Acoustical Performance criteria, Design, Requirement and Guidelines for Schools", SI 2.60, 2002.

[12]G.J. Zeilstra,"Speech Intelligibility in Classroom",Facalty of applied sciences,2009.

[13] Hagen," Neural Network Design", kui dai, translated, Beijing Machinery Industry Press, 2005.

[14] De Giuli V, Da Pos O, De Carli M , " Indoor Environmental Quality and Pupil Perception in Italian Primary Schools", Building and Environment, 2012.

التحقيق في مفهوميه الكلام عن طريق نموذج الشبكات العصبية الاصطناعية

م.م. دينا حارث شاكر*

المستخلص

إن الخواص الصوتية للقاعات الدراسية هو جزء مهم في التصميم الهندسي ولكن يصعب تحقيقه في هذه القاعات وخاصة القاعات الجامعية. ويتحقق التصميم الجيد باستخدام الخبرة في التصميم الصوتي مع التصميم الهندسي النقي. في هذه الورقة تستخدم الشبكات العصبية الاصطناعية (ANN) للتنبؤ بمفهومية الكلام في الفصول الدراسية ولتشخيص مواقع الضعف والخلل من الناحية الصوتية , وهنالك العديد من الخصائص الصوتية في القاعات الدراسية التي يجب أخذها بنظر الاعتبار عند التصميم منها أبعاد القاعة, النسبة بين الإشارة الصوتية إلى الضوضاء(SNR), موقع كل من الطالب والأستاذ, الضوضاء الخلفية قد تم اختيار مجموعة من النصوص وإنشاء خطاب قاعدة البيانات كإشارة ومن ثم قياس منسوب الضغط الصوتي باستخدام جهاز قياس منسوب الضغط الصوتي في مواقع مختلفة داخل القاعة. وتم الحصول على ورقة البيانات من هذا القياس واستخدامها كقاعدة بيانات لتدريب الشبكات العصبية الاصطناعية (ANN) التي تعمل على التنبؤ بمفهومية الكلام داخل القاعة ولمواضع مختلفة للمستمعين. وهذا كله يؤدي إلى تقليل الأخطاء الصوتية في القاعات الدراسية وتحسين كفاءة العملية التعليمية.

الكلمات المفتاحية: الصوتيات , الضوضاء, منسوب الضغط الصوتي,مفهومية الكلام, الشيكات العصبية الاصطناعية.

*الجامعة التكنولوجية