International Journal of Research and Reviews in Artificial Intelligence
Vol. 1, No. 1, March 2011

www.sciacademypublisher.com

Performance Comparison of the Artificial Neural Network and the K-Nearest Neighbor Classifiers in Classroom Speech Intelligibility Prediction Application

M. Ridhwan Tamjis , M. Naufal Mansor, Paulraj M P, Sazali Yaacob, A. Nazri Abdullah, Raymond B. W. Heng

Intelligent Signal Processing Research Cluster, Universiti Malaysia Perlis, Kangar, Malaysia

Correspondence should be addressed to M. Ridhwan Tamjis, mystril_nd@yahoo.com

Abstract – Classroom speech intelligibility has become one of the major concerns in education nowadays. In any classrooms and educational facilities, an optimal speech intelligibility level is required to ensure that the listeners at any location in the classroom have a good perception of the message that is conveyed by the speaker. Classrooms need to be designed carefully in order to give optimal classroom speech intelligibility level. In this paper, two different types of artificial intelligence methods are proposed to implement the prediction application: Artificial Neural Network (ANN) and the k-Nearest Neighbor (k-NN). Both classifiers are trained separately using the previously acquired datasets which consist of acoustical parameters and the speech intelligibility of actual classrooms. Results show that the ANN performs better on imbalanced datasets compared to the k-NN.

1. Introduction

Speech intelligibility is a term that is widely used in acoustic, architectural and audiology studies. It is a measure of the quality of the perceptions of the listeners towards the speaker. For any classroom or teaching facilities, it is necessary to have an excellent level of speech intelligibility to make sure that sound is distributed sufficiently to all listeners. Optimum speech intelligibility is considered achieved when a listener hears correctly the words uttered by a speaker and the word is not mistaken with other words. For example, a person might speak with an accent but is still understood by a listener who is face-to-face with the talker. The accent might be a distortion in communication, but as long as the listener understands the message, speech intelligibility is established. Several researchers reported that classrooms with plausible level of speech intelligibility yields high performance students [1]-[3].

Several methods have been introduced to increase the speech intelligibility in the classroom [4]. However, most of the methods presented are only cost effective when implemented at the stage of classroom design. Renovation of an established classroom requires a large investment. Moreover, the acoustical requirements such as the length, height, and width of the classroom must be computed carefully to obtain classroom with optimal speech intelligibility.

To ease the work of classrooms architect, we are proposing an artificial intelligence based classroom speech intelligibility prediction system. This system will predict the classroom speech intelligibility based on the given acoustical parameters. This research will utilize the sample-based learning method to train the prediction system.

For any artificial intelligence application, it is crucial to have ample amount of data to train and test the classifiers. This is on the basis that the classifier needs to be fed with actual input and output to classify something. This concept imitates human requirements to recognize objects. For example is on recognizing a screwdriver. What are the features that differentiate screwdriver from other objects? We know that basically screwdrivers have long and edgy iron tips, have gripping handle that is made of plastics, and rod shaped. These features made our brain classify what we see as a screwdriver. The same concept is applied on artificial intelligence.

The objective of this paper is to compare the performance of the Artificial Neural Network (ANN) and k-Nearest Neighbor (k-NN) in the classroom speech intelligibility prediction application. This paper will also investigate the effect of imbalanced datasets toward the performance of both artificial intelligence classifiers.

2. Methodology

2.1. The Acoustical Measurement

Acoustical measurement requires time-consuming and expensive field-measurement. This is because this field demands strict adherence to the measurement and test standards. Furthermore, the measurement activity requires sophisticated and sensitive measuring devices.

For classroom speech intelligibility prediction application, measurement needs to be conducted in the actual classrooms conditions with acoustical instruments that comply with the acoustical standards.

2.1.1. *The Measuring Instruments*

For the measurement, we have employed several acoustical measuring equipments and devices. These include the unidirectional loudspeakers, power amplifiers, and mixers. But most important is the equipment to measure the acoustical criteria in the classroom. For that reason, a sound level meter that is able to record and store the sound spectrum both in time and frequency domain is used.

2.1.2. *Classroom Condition and Experiment Set up*

According to established standard [4], the classroom to be measured needs to be unoccupied at all time except for the person conducting the measurement. The adjacent classrooms need to be empty as well to reduce the noise disturbances. The furniture available in the classroom is left in their original positions to maintain the actual classroom condition. The door and the window are tightly shut at all time during the measurement.

Figure 1 gives the basic idea on the experiment set up in the classroom.

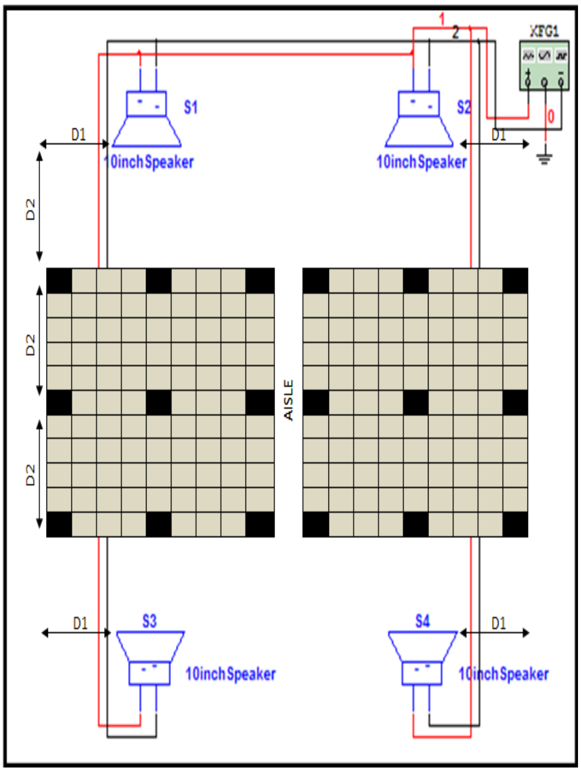

Figure 1. Basic Experiment Setup

From Figure 1, a total of four unidirectional speakers are used to accommodate the sound distribution requirements in university classrooms. These speakers are placed facing each other in a balanced distance.

2.1.3. *Measurement Method*

The measurement was conducted based on the requirements stated in the standards and by previous researchers. Figure 2 depicts the work flow of the measurement activity.

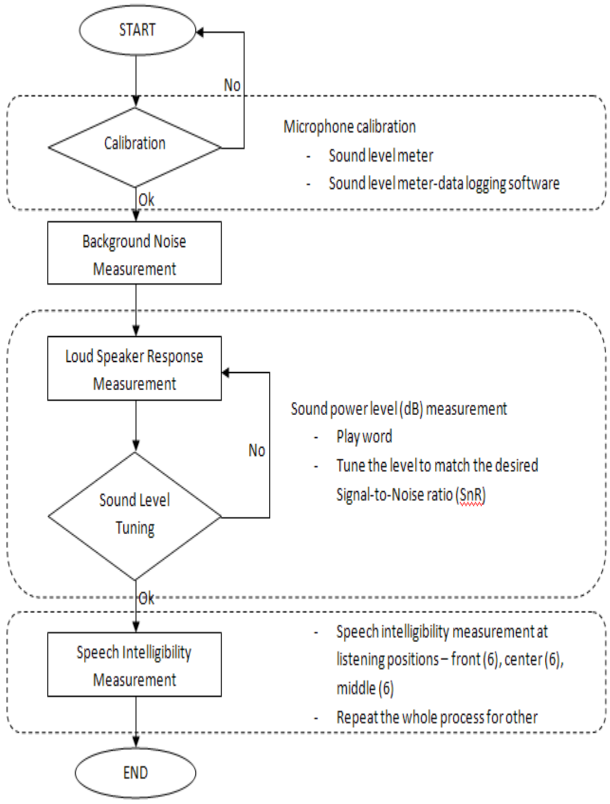

Figure 2. Measurement Workflow

Further elaboration on the measurement workflow is discussed in [5].

For the purpose of developing the classroom speech intelligibility prediction system, we used a total of 3280 datasets. The datasets consist of eight different inputs, namely X-position, Y-Position, classroom's length, height and width, background noise level, word's sound power level, and signal-to-noise ratio. The output is the Speech Transmission Index (STI). The datasets are summarized in Table 1.

Table 1. Data Summary

	Min	Max
X	1	6
Y	1	6
RT	1.51	2.42
Width	6.4	10.06
Length	12.802	18.9
Height	3.03	3.56
Word	55	72
SNR	0	20
STI	0.1	0.78

It is important to note here that the outputs of the classifiers (STI) are separated into five different classes as shown in Table 2.

Table 2. STI Classes

STI	0-0.3	0.3-0.45	0.45-0.6	0.60-0.75	0.75-1.0
Specification	Worse	Poor	Normal	Good	Excellent
Class	1	2	3	4	5

2.2. Artificial Intelligence Based Prediction

For the purpose of developing the classroom's speech intelligibility prediction system, two different types of artificial intelligence classifiers are proposed which are (1) ANN and (2) k-NN.

2.2.1. Artificial Neural Network

ANN is form of computation that is motivated right from its inception by the recognition that the human brain computes in an entirely different way from the digital computer [6]. It has been used in many classification applications for its robustness in classifying the desired output. ANN comes in many forms, but the most interesting method is the feed forward network due to its learning and generalization properties.

In this research, a simple feed forward back propagation neural network was developed to predict the classroom's speech intelligibility. The network consists of two hidden layers, with eight input neurons in the first layer which represents all eight acoustic parameters. Each hidden layer is set to have 22 hidden neurons. The output of each layer is computed using the Log-Sigmoid transfer function.

2.2.2. K-Nearest Neighbor

The k-NN is a method for classifying objects based on the closest training examples in the feature space. It is a type of instance-based learning, or lazy learning where the function in only approximated and all computation is deferred until classification. The k-NN is one of the fundamental and one of the simplest classifying methods in artificial intelligence. It classify an object based on the majority voting of its neighbors, with the object being assigned to the class is most common amongst its k nearest neighbor [7].

In this work, for each test parameter (to be predicted), minimum distance from the test parameters to the training set was calculated to locate the k-NN category of the training data set. The Euclidean Distance measure was used to calculate how close each member of the training set was to the test class that was being examined. The Euclidean distance was measured using the equation as in (1).

$$d_E(x, y) = \sum_{i=1}^{N} \sqrt{x_i^2 - y_i^2} \tag{1}$$

From this k-NN category, class label of the test parameters was determined by applying majority voting.

3. Results and Discussion

The performances of both classifiers were measured by calculating the error rate between the original and the predicted output. The results of both classifiers are shown in Table 2.

Table 3. Classification Results

Classifier	Neural Network	K-NN
Train Samples	2644	2644
Test Samples	1133	1133
Average Performance	83.2%	65.8%

As mentioned earlier, the goal of this paper is to compare the performance of the ANN and the k-NN. For that reason, both networks were trained separately using the same preprocessing and post-processing method. However, the method used will not be discussed in this paper.

It can be clearly seen in Table 3 that the neural network yields the best overall classification percentage when compared to the k-NN. The overall performances were calculated based on the ability of the network to classify each class with minimal error. The reason for the K-NN having low performance when compared to the neural network might be due to the imbalanced of the datasets.

For any artificial intelligence method, ample amounts of input and output data are required to fulfill the sample-based training requirement. However, the processes of obtaining the data are not cheap and might consume more than half of the research time, especially in acoustic application research. Moreover, the measurement might not yield the results that meet our expectation due to several environmental factors. For example in this research, it is almost impossible to obtain an excellent intelligibility reading (0.75-1.0) due to the condition of the classroom itself. A classroom with an excellent acoustical condition is hard to find unless acoustical renovation is done to the current classroom for the research to get good data sample.

Due to the unavailability of a classroom with excellent acoustical condition, we have obtained imbalanced datasets. In other words, the obtained data do not sufficiently cover the whole classes. Thorough analysis in Table 4 shows that both classifiers classify correctly on the classes that have large amount of samples. On the other hand, the classifiers produces low classification rate for classes with small number of samples. The k-NN yields the worse classification rate when compared to ANN.

Table 4. Classification Rate of Each Class

		Classification Rate (%)	
Class	**∑ of Samples**	**ANN**	**k-NN**
1	1133	95.76	97.63
2	1133	100	96.35
3	1133	98.85	96.9
4	374	76.2	26.74
5	4	50	0

4. Conclusion

From the findings, it can be concluded that neural network performs better compared to the K-NN, especially for imbalanced datasets. To develop an excellent and accurate classroom speech intelligibility prediction system, it is necessary to have a balanced dataset that covers the whole classes of STI.

Acknowledgments

The authors would like to thanks the administration of the Universiti Malaysia Perlis and the Ministry of Higher Education, Malaysia for providing the necessary equipments and measurement spaces to conduct this research.

References

[1] [1] F. s. Berg, J. C. Blair, and P. V. Benson, "Classroom acoustics: The problem, impact, and solution," Language, Speech, and Hearing Services in School, vol. 27, pp. 16-20, 1996.
[2] [2] J. S. Lukas, R. B. DuPree, and J. W. Swing, "Effect of noise on academic achievements and classroom behaviour," Berkeley,1981.
[3] [3] J. S. Bradley, "Speech intelligibility studies in classrooms," Journal Acoustical Society of America, vol. 80, pp. 846-854, 1986.
[4] [4] Acoustical performance, criteria, design requirements, and guidelines for school, ANSI Standard-2002.
[5] [5] M. R. Tamjis, M. P. Paulraj, S. Yaacob, and A. N. Abdullah, "Measuring the speech intelligibility of the sound reinforced classroom," in Proceedings of the International Postgraduate Conference on Engineering (IPCE 2010), Perlis, Malaysia, 2010.
[6] [6] Simon Haykin, Neural Networks: A Comprehensive Foundation. New Jersey: Prentice Hall Upper Saddle River, 1999, pp 1-15.
[7] [7] Dudani, S.A. The distance-weighted k-nearest-neighbor rule. IEEE Trans. Syst. Man Cybern., SMC-6:325–327, 1976.

Mohammad Ridhwan Tamjis received his Bachelor Degree (Hons.) in Mechatronic Engineering from Universiti Malaysia Perlis, Malaysia in 2009. He is currently working as a Research Assistant while pursuing Master of Science in Mechatronic Engineering at the university where he obtained his bachelor degree. His research interest includes artificial intelligence, signal processing, machine vision, automation and aerospace technology.

Muhammad Naufal Mansor received his Diploma in Mechanical Engineering from Politeknik Sultan Abd Halim Mu'adzam Shah and a Bachelors degree in Mechatronic Engineering from University, Malaysia Perlis, Malaysia. He is currently doing Master in Mechatronic Engineering at Universiti Malaysia Perlis, Malaysia. His research interests include signal processing, image processing, vision, mechanical system, fuzzy system and artificial neural network.

Paulraj Murugesa Pandian is currently working as an Associate Professor at Universiti Malaysia Perlis, Malaysia. His research interests include artificial intelligence, Fuzzy systems, Speech processing, acoustic engineering and biosignal processing applications. He has published more than 100 papers in referred journals and conferences. He has authored a book titled Introduction to Artificial Neural Networks. He is a member of the Institution of Engineers, India and MISTE,Indi

Sazali Yaacob is currently working as a Professor at Universiti Malaysia Perlis, Malaysia. He has published more than 150 papers in Journals and Conference Proceedings. His research interests are in Artificial Intelligence applications in the fields of acoustics, vision and robotics. He received Charted Engineer status by the Engineering Council, United Kingdom in 2005 and also a member of the IET (UK).

Ahmad Nazri Abdullah is currently working as an Associate Professor at Universiti Malaysia Perlis, Malaysia. With a background in technical communication, he plays an important role in literature based acoustic researches such as vowel and phoneme classification in speech. His research interests includes acoustic, speech processing and artificial intelligence.

Raymond BW Heng is currently working as Professor at Universiti Malaysia Perlis. He has published over 100 papers in Acoustics, Noise and Vibrations in Journals and Conference Proceedings. He is a Fellow of the Institute of Acoustics, UK, and Chartered Engineer and Professional Engineer, a member of the Institution of Mechanical Engineers, UK, Acoustical Society of America, British Society of Audiology, and Environmental Engineering Society.