ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

Background Noise Reduction using FFBPNN-LM Network and Adaptive Filter

G. Deepika[1], M. Vidya[2]

PG Scholar, Department of ECE, SVCE, Sriperumbudur, India [1]

Assistant Professor, Department of ECE, SVCE, Sriperumbudur, India[2]

ABSTRACT: Speech is the most efficient and popular means of communication. Speech is usually contaminated with background noises from environment. These noises usually degrade the performance of the speech signal. The main aim of this project is to filter the background noises from the speech signal. A signal classifier is constructed to classify the noise affected input speech signal into various classes based on the type of noise. The classification can be done based on certain features such as energy, entropy, variance, covariance, standard deviation, bandwidth, pitch, periodicity and frequency regarding their ability to distinguish the different speech signals. The input speech signals passing through the network should be such that it gets grouped under any one of the four output classes. The neural network which has been chosen is the "Feed forward Back Propagation Neural Network". The algorithm used to train this network is "Levenberg Marquardt Algorithm". The output of the Feed Forward Back Propagation Neural Network is fed to an adaptive filter where the noise gets removed to some extent in the speech signal. MATLAB version 13 is used for simulation.

KEYWORDS: Feedforward Backpropagation Neural Network, Levenberg Marquardt, Mean Square Error.

I. INTRODUCTION

Generally, we all live in an environment flooded with noise. The recovery of the original signal from the noise speech signal is a very common goal in the design of signal processing system. These noises usually degrade the performance of the speech signal, thereby making it difficult for the people living in the environment. In this project, the noises which are supposed to have lower frequency ranges have been dealt.

Speech enhancement in a noisy environment has been a challenging problem for decades. Noise gets added to speech signal almost in an uncontrolled manner. Speech processing system (e.g. speech coding, speech recognition, speaker verification) encounters different types and levels of background acoustical noise (e.g. traffic noise, car noise, office noise etc.). These systems pick up those noisy signals along with speech. The noise signals result in performance degradation of those systems.Therefore, many techniques have been used to reduce or completely remove those noisy signals.The main technique usually followed everywhere is the adaptive noise cancellation technique where the noise gets removed and a distortion less speech is obtained.The algorithms which have been used in these techniques are Least Mean Square Algorithm, Recursive Least Square Algorithm etc.

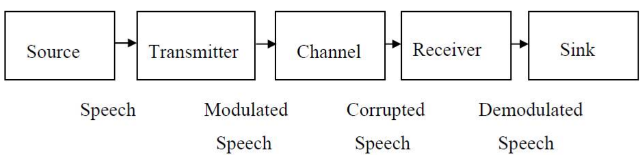

Fig 1. Speech Processing System

The source generates the speech signal. The transmitter is a system designed in such a waythat the original signal can be transmitted through thechannel and reaches itsdestination. If suppose the signal is weak, it can't be transmitted as such, it must be modulated with a carrier wave first. Channel is the medium through which the signal is transmitted. One cannot avoid the channel though the signal being transmitted will be corrupted by noise, distorted and attenuated.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

As environmental noises vary in texture and dynamics, using one coding scheme has proven to be inadequate for many common types of noises. Therefore,multicoding scheme can be adopted. The receiver would serve as the inverse system to the transmitter. However, because of the channel, the receiver must do its best to reproduce the original signal. Finally, the received speech is sent to the sink which somehow makes use of the speech.
Noise classification is also one of the major parts in many of the intelligent systems. Noise classification is used to reduce the effect of environmental noises on speech processing tasks. It is used to design natural-quality, multi-mode noise coding algorithms. Similarly, multi-mode comfort noise generators are designed to remedy the noise contrast problem, reported in discontinuous transmission-based cellular systems (GSM). Noise classification is also used in many other applications such as programmable hearing-aid devices and noise monitoring systems. In programmable hearing-aid devices, a classification algorithm automatically matches a program mode with the listening environment of the user. In noise monitoring systems, classification of environmental noises is done to control noise pollution.
Many efforts were made to design automatic systems forcontrolling noise. In pattern classification problems, neural network has been found to provide successful results. The neural network can be trained to discern the criteria used to classify the inputs into groups, and it can do so in a generalized manner allowing successful classification of new inputs under various noisy environments.
The system which has been designed can classify any incoming noise affected speech signal. The noises usually include the low frequency noises such as dogs barking noise, classroom noise, fan noise, traffic noise, rain, military noise etc. By doing such a classification, the system will be able to recognize any lowfrequency noise affected speech signal entering the system.The features which have been chosen include entropy, energy, standard deviation, pitch, periodicity, variance, co-variance, bandwidth and frequency. Once all the features of the corresponding noise affected speech signal have been trained in the feed forward back propagation neural network, then testing can be done. Finally, they have been send to an adaptive filter where the noise gets filtered.

II. PROPOSED SYSTEM

The input noise affected speech signal is first fed to the to the feed forward back propagation neural network [13] . The output of forward back propagation neural network is the classification of the noise affected speech signal based on the detectable background noise. This noise affected speech signal is then sent to the adaptive filter where the noise from the speech signal gets removed. Finally, a clean speech signal up to a certain extent has been obtained. The figure 2 depicts a general block diagram of the proposed system.

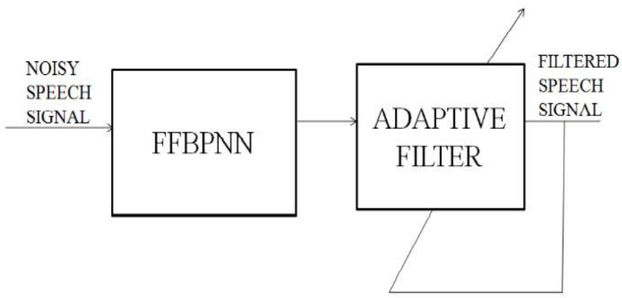

Fig 2. General Block Diagram of the proposed system

Initially, noise affected speech signals from the NOIZEOUS database are first taken.Then the features of these signals are extracted. Feedforward backpropagation neural network is then trained with these by a suitable algorithm called Levenberg Marquardt algorithm [12]. After the training gets finished, the testing phase is done where the signal is fed to the network which is classified under a particular output class. The figure 3 depicts a detailed block diagram of the proposed system.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

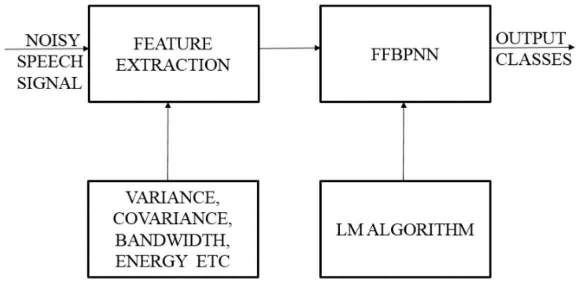

Fig 3. Detailed Block Diagram of the proposed system

III. FEATURE EXTRACTION

When an audio signal gets classified under a given class, the features in that audio signal must be extracted before the classification process. These features usually decide the class to which the audio signal belongs. Feature extractioninvolves the analysis of the input ofthe audio signal. Thefeature extraction techniques can be classified as temporal analysis and spectral analysis technique. In Temporal analysis,the waveform of the audio signal itself is used for analysis. Spectral analysis utilizes spectral representation of the audio signal for analysis. All audio features are usually extracted by breaking the input signal into a succession of analysis windows or frames, each of around 10-40-ms length, and computing one feature value for each of the windows. An approach which has been followed is to take the values of all features for a given analysis window to form the feature vectors for the classification decision, so that the class assignments can be obtained in real time.Thus, feature extraction plays an important role in classification of an audio signal. There are different types of features such as pitch, periodicity etc that are explained below.
A major step in the design of a signal classification system is the selection of a good set of features such that theycan separate the signals in the feature space. Features that capture the temporal and spectral structure of the input signal are commonly used.The feature extraction is needed to reduce the dimensionality of the data passed on to the neural network. The neural network determines the best way to process the data to arrive at a classification. Some of the features considered for classification are variance, covariance, bandwidth, energy etc.

A.Pitch

The sound that comes through vocal tract starts from the larynx where vocal cords are controlled by nerves from brain. The vibration in the vocal cords and the shape of the vocal tract are controlled by nerves from brain. The sound, which we produce, could be categorized into voiced and unvoiced sounds. When unvoiced sounds are produced, the vocal cords do not vibrate and they stay open whereas during voiced sounds they vibrate and producewhat is known as glottal pulse. A pulse is a summation of a sinusoidal wave offundamental frequency and its harmonics (Amplitude decreases as frequency increases). The fundamental frequency of glottal pulse is known as the pitch.

B.Periodicity

Periodicity is a measure of the periodic structure of speech. The largest peak in the normalized function is chosen as the estimate of the pitch period and the value of the peak becomes the periodicity measure.This voicing function is a measure of how strong the speech frame is.

$$\text{Periodicity} = \max_{m} \{Ri(m)/\ Ri\ (0)\},\ 20 \le m \le 120$$

whereRi(m) and Ri (0) are the autocorrelation values and m is the delay.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

C.Entropy

Entropy means a process in which order deteriorates with the passage of time.The entropy of a system is maximal when the system has reached the thermodynamic equilibrium.

$$\mathrm{H}(x) = -\sum_{i=0}^{N} \mathrm{P}(x_i) log_{10} p(x_i)$$

where x = {x1, x2, . . ., xN} is a set of random phenomena, and p (xi) is a probability of a random phenomenon xi.

D. Energy

The amplitude of unvoiced segments is noticeably lower than that of the voiced segments. The short- time energy of speech signals reflects the amplitude variation. In a typical speech signal we can see that its certain properties considerably change with time.The energy function can also be used to locate approximately the time at which voiced speech become unvoiced speech and vice versa, and for high quality speech (high signal to noise ratio) the energy can be used to distinguish speech from silence.
The short time energy of the speech signal provides a convenient representation that reflects the amplitude variation and can be defined as

$$\mathrm{E_N} = \sum_{m=-\infty}^{\infty} [x(m)W(n-m)]^2$$

wherex (m) represents the data sequence and W (n-m) represents a limited time window sequence.

E. Variance

Variance is a measurement of the spread between numbers in a data set. The variance measures how far each number in the set is from the mean. Variance is calculated by taking the differences between each number in the set and the mean

$$\sigma^2 = \frac{\sum(X-\mu)^2}{\mathrm{N}}$$

X: individual data point
µ: mean of data points
N: total no of data points

F. Covariance

Co-variance is a measure of the degree to which returns on two risky assets move in tandem.Covariance is calculated by analyzing at return surprises (standard deviations from expected return), or by multiplying the correlation between the two variables by the standard deviation of each variable.

Covariance can be calculated using the following formula

$$Cov(X,Y) = \frac{\sum_{i=1}^{n}(X-\bar{X})(Y-\bar{Y})}{\mathrm{N}-1}$$

$\bar{X}$ = Individual value of X
X = Mean of X
$\bar{Y}$ = Individual value of Y
Y = Mean of Y

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

n = the number of data points

G. Standard Deviation

Standard deviation is a measure of the dispersion of a set of data from its mean. If the data points are further from the mean, there is higher deviation within the data set. Standard deviation is calculated as the square root of variance by determining the variation between each data point relative to the mean.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{n}(x_i - \mu)^2}$$

xi = individual data point
µ = mean of data points
N = total no of data points

H. Bandwidth

Bandwidth is defined as a band containing all frequencies between upper cut-off and lower cut-off frequencies.

$$fb = fu - fl$$

fb = bandwidth
fu = upper cut off frequency
fl = lower cut off frequency

I. Frequency

Frequency is the number of complete cycles per second.

$$f = 1/T$$

f is the frequency in hertz
T is the Time period in msec

IV. FEEDFORWARD BACKPROPAGATION NEURAL NETWORK

Neural networks are one among the family of computational architectures which are inspired by biological brains. Such architectures are referred to as connectionist systems.These systems constitute interconnected and interacting components called nodes or neurons. Neural networks are usually characterized by absence of explicit representation of knowledge.It is also said that there are no symbols or values that directly correspond to classes of interest. Rather, knowledge is implicitly represented in the patterns of interactions between network components. A graphical depiction of a typical feed forward neural network is given in Figure 4. The term “feed forward” indicates that the network has a flow that extends in only one direction. Except during training, there is no backward flow in a feed forward network; all links proceed from input nodes toward output nodes.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

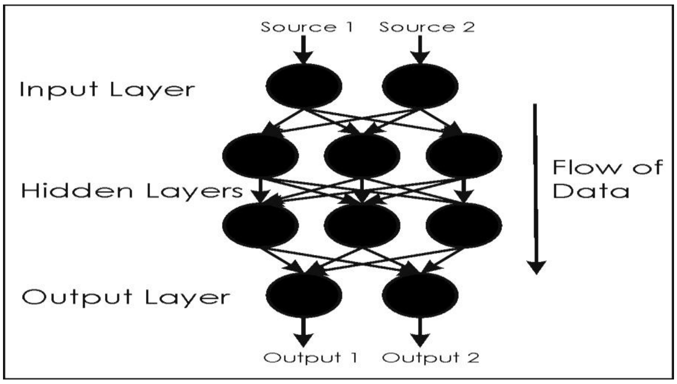

Fig 4. A Typical Feedforward Neural Network

Individual nodes in a neural network resemble biological neurons by taking input data and performing simple operations on those data by selectively passing the results on to the next layer which also constitutes the neurons. The output which is obtained from each node is called "activation". For each vector and node in the network, weight values are associated with these values vector impel the input data to flow through the network. Weights associated with individual nodes arealso known as biases. Weight values are determined by the consecutive flow oftraining data through the network (i.e., weight values are established during a training period in which the network learns how to identify particular classes by their typical input data characteristics).
After the training phase, the neural network has been said to focus on the classification of new data. Classifications are performed by trained networks through

- The activation which is obtained from the network input nodes by relevant data sources (these data sources must directly match those used in the training of the network).
- The forward flow of this data through the network.
- The ultimate activation of the output nodes.

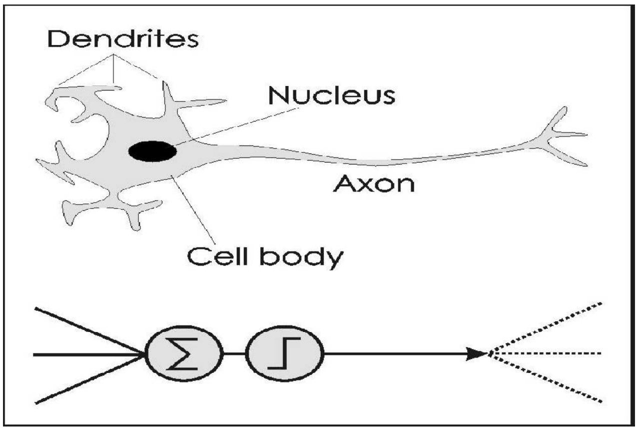

Fig 5. Comparison between biological neuron and an artificial neuron

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

A. Properties of Neural Network

- The NNs are capable of displaying mapping capabilities, i.e.input patterns are mapped to their associated output patterns.
- The NNs have the capacity to learn by examples. Therefore, NN architectures can be trained with known examples of problems before they are tested for their capability on unknown instances of the problem. Therefore, they can identify new objects previously untrained.
- The NNs possess the ability to make a problem more general. Thus, they can predict new outcomes from the past trends.
- The NNs are robust systems and are fault tolerant. They can, therefore, evoke full patterns from incomplete, partial or noisy patterns.
- The NNs have the capability to process information in parallel, at high speed, and in a distributed manner.

B. Levenberg Marquardt Algorithm

The Levenberg-Marquardt algorithm was designed to appeal second-order training speed without having the need to compute the Hessianmatrix. When the performancefunction resembles the form of a sum of squares (as is typical in training feed forward networks), then the Hessian matrix can be approximated as

$$H = J^T J$$

and the gradient can be computed as

$$g = J^T e$$

where J is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and back propagation technique that is much less complex than biases, and e corresponds to vector of network errors. The Jacobian matrix can be computed through a standard computing the Hessian matrix.

The Levenberg-Marquardt algorithm uses an approximation to the Hessian matrix in the subsequent Newton-like update:

$$x^{k+1} = x^k - [J^T J + \mu I]^{-1} J^T e$$

When the value of μ is zero, this is just Newton's method, using the approximate Hessian matrix. When the value of μ is large, it becomes gradient descent with a small step size. It is also said that Newton's method is faster and more accurate near an error minimum, so the aim is to shift towards Newton's method as quickly as possible. Thus, μ is decreased after each successive step (reduction in performance function) and is said to increaseonly when a tentative step is said to increase the performance function. In this way, the performance function will always be reduced at each iteration of the algorithm.

C. Multiclass Classification

Multiclass or Multinomial classification is termed as the problem of classifying instances into one of the more than two classes. Each training point belongs to one of N different classes. The goal is to construct a function which when given a new data point, will correctly predict the class to which the new data belongs.

In this project, four output classes have been defined for the Feed Forward Back Propagation Neural Network.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

- Airport noise class
- Car noise class
- Restaurant noise class
- Train noise class

V. SOFTWARE IMPLEMENTATION AND RESULTS

A. Neural Network Toolbox

The Neural Network Toolbox is a tool which is present in a directory called nnet. In order to list the help topics,type help nnet. A number of demonstrations areincluded in the toolbox.Neural network demonstration and application scripts can be found by typing help nndemos. To install the Neural Network Toolbox, instructions found in the following two MATLAB documents: The Installation Guide for PC or the Installation guide for UNIX can be followed. A few of the new features and improvements introduced with this version of the Neural Network Toolbox are discussed below

Control System Applications
• Network model predictive control
• Model reference adaptive control
• Feedback linearization controller

Graphical User Interface
A graphical user interface has been present in this toolbox. This interface allows you to:
• Create networks using the functions in toolbox.
• Enter necessary data into the GUI
• Initialize, train, and simulate networks
• The training results are exported from the GUI to the command line workspace
• The data can be imported from the command line workspace to the GUI
To open the Network window type nntool.

New Training Functions
The toolbox now has four training algorithms that apply weight and bias learning rules. One algorithm applies the learning rules in batch mode. Three algorithms apply learning rules in three different incremental modes:
• trainb - Batch training function
• trainc - Cyclical order incremental training function
• trainr - Random order incremental training function
• trains - Sequential order incremental training function
All four functions present the whole training set in each epoch (pass through the entire input set).

B. Database

A noisy speech corpus (NOIZEUS) has been present to facilitate comparison of speech enhancement algorithms among research groups. The noisy database contains 30 IEEE sentenceswhich are produced by three male and three female speakers corrupted by eight different real-world noises at different SNRs. The noise was taken from the AURORA database and includes suburban train noise, babble, car, exhibition hall, restaurant, street, airport and train-station noise. This corpus is available to the researchers in free of charge.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

TABLE I

FEATURES OF TRAINING DATA SET

Features	Airport noise signal	Car noise signal	Restaurant noise signal	Train noise signal
Variance	0.003	0.003	0.004	0.004
Entropy (J)	3.235	3.442	3.379	3.353
Standard deviation	0.049	0.063	0.067	0.062
Energy (J)	56	84	72	81
Covariance	0.002	0.003	0.004	0.003
Bandwidth (Hz)	0.002	0.004	0.003	0.003
Pitch (Hz)	888	888	888	666
Periodicity (Hz)	2252	2111	1762	1826
Frequency (Hz)	297	823	377	383

C. Simulation Results

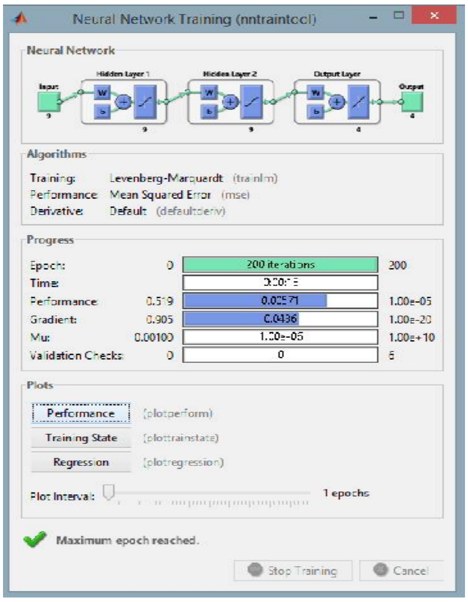

Fig 6.NN Training Tool

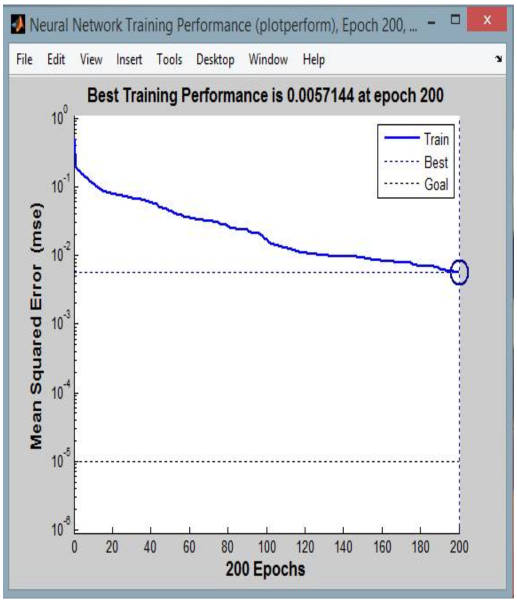

Fig 7. Performance Plot

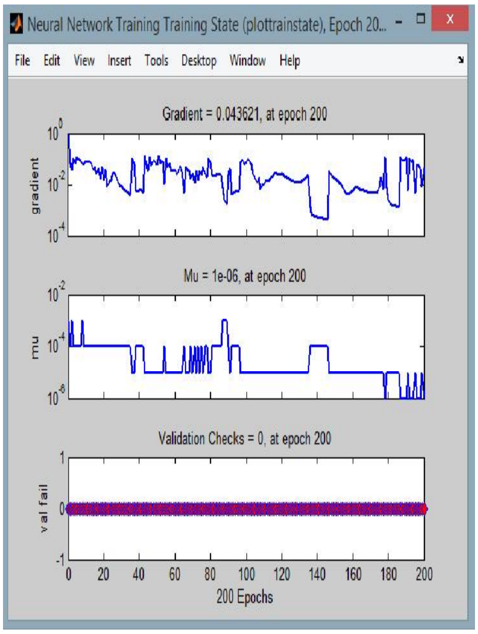

Fig 8. Training State Plot

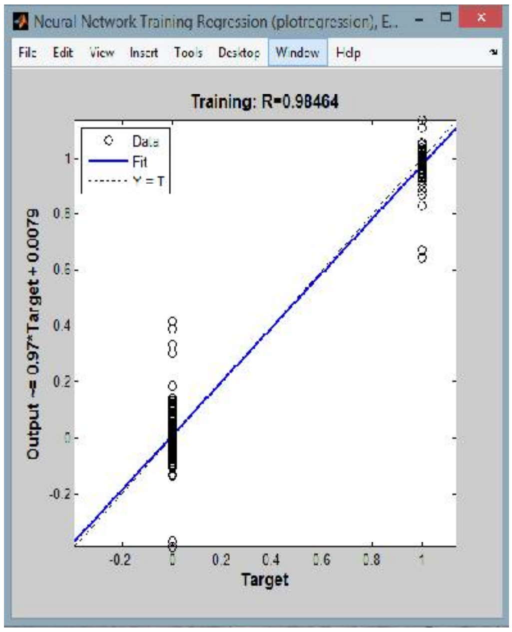

Fig 9. Regression Plot

VI. CONCLUSION

Thus an ANN is constructed and it can recognize few types of noise affected speech signal. Feed forward Back propagation Neural Network, trained with a training algorithm Levenberg Marquardt works efficiently by classifying the corresponding speech signals into a particular class. When a speech signal comes under a particular class, the corresponding class is represented by `1`.

A. Future Scope

- In phase 1, noise affected speech signal is fed to a FFBP-NN and the noisy speech signal gets classified in any one of the output classes based on the detectable background noise. Number of classes can also be increased further.
- In phase 2, the output of the FFBP-NN is given to an adaptive filter to reduce the background noise.

REFERENCES

[1] Michael A.Carlin,MounyaElhilali,"A Framework for Speech Activity Detection Using Adaptive Auditory Receptive Fields",IEEE /ACM Transactions On Audio, Speech, And Language Processing, Vol. 23, No. 12, December 2016.
[2] Francisco Ortega-Zamorano, Jose M. Jerez, Daniel UrdaMunoz, Rafael M. Luque-Baena, and Leonardo Franco, "Efficient Implementation of the Back propagation Algorithm in FPGAs and Microcontrollers ", IEEE Transactions On Neural Networks And Learning Systems, Vol. 27, No. 9, September 2016.
[3] Ofer Schwartz, Sharon Gannot and Emanuel A. P. Habets, "An Expectation-Maximization Algorithm for Multimicrophone Speech Dereverberation and Noise
Reduction With Coherence Matrix Estimation", IEEE /ACM Transactions On Audio, Speech, And Language Processing, Vol. 24, No. 9, September 2016.
[4] Xiao-Lei Zhang, DeLiang Wang, "Boosting Contextual Information for Deep Neural Network Based Voice Activity Detection", IEEE Transactions On Audio, Speech, And Language Processing, Vol. 24, No. 2, February 2016.
[5] MahfoudHamidia, AbderrahmaneAmrouche," Improved variable step size NLMS Adaptive filtering algorithm for acoustic ", ELSEIVER Publications, November 2015.
[6] Toyota Fujioka, Yoshifumi Nagata, and Masato "High-Precision Harmonic Distortion Level Measurement of a Loudspeaker Using Adaptive Filters in a Noisy Environment", IEEE /ACM Transactions On Audio, Speech, And Language Processing, Vol. 23, No. 10, October 2015.
[7] MichielHermans, Joni Dambre, and Peter Bienstman, "Optoelectronic Systems Trained With Back propagation Through Time", IEEE Transactions On Neural Networks And Learning Systems, Vol. 26, No. 7, July 2015.

[8] Zongsheng Zheng, HaiquanZhao,"Affine projection M-estimate subband adaptive filters for robust adaptive filtering in impulsive noise", ELSEIVER Publications, December 2014.
[9] Ian McLoughlin, Haomin Zhang, ZhipengXie, Yan Song, and Wei Xiao," Robust Sound Event Classification using Deep Neural Networks", IEEE Transactions On Audio, Speech, And Language Processing, Vol. 25, No. 10, May 2014.
[10] Xiaojia Zhao, Yuxuan Wang, and DeLiang Wang," Robust Speaker Identification in Noisy and Reverberant Conditions", IEEE Transactions On Audio, Speech, And Language Processing, Vol. 22, No. 4, April 2014. 33
[11] Danilo Comminiello, Michele Scarpiniti, Raffaele Parisi, Aurelio Uncini "Combined adaptive beamforming schemes for nonstationary interfering noise reduction ", ELSEIVER Publications, October 2013.
[12] Jose de Jesus Rubio, PlamenAngelov, Jaime Pacheco ,"Uniformly Stable Back propagation Algorithm to Train a Feed forward Neural Network", IEEE Transactions On Neural Networks , Vol. 22, No. 3, March 2011.
[13] SamanRazavi and Bryan A. Tolson, "A New Formulation for Feedforward Neural Networks", IEEE Transactions On Neural Networks , Vol. 22, No. 10, October 2011.
[14] A. Krueger, E. Warsitz, R. Haeb-Umbach," Speech enhancement with a GSC-like structure employing eigenvector-based transfer function ratios estimation", IEEE Transactions On Audio, Speech, And Language Processing (2011) 206-219.
[15] D. Panagiotopoulos, C. Orovas, and D. Syndoukas, "A heuristically enhanced gradient approximation (HEGA) algorithm for training neural networks", Neurocomputing, March 2010.