STOCHASTIC SOLUTIONS FOR THE CHALLENGES IN PROPOSAL ENGINEERING

A Dissertation Report

Submitted in the Partial fulfillment of the

Requirement for Award of the Degree

of

MASTER OF THECHNOLOGY

In STRUCTURAL ENGINEERING

By

Sathish M

(11604451)

Under the guidance of
Mr. Manoharan Rajalingam
(Asso. Professor and Head of Structural Engineering Discipline)

School of Civil Engineering

Lovely Professional University

Phagwara, Punjab (India)-1444110

DECLARATION

I hereby declare that the dissertation entitled, "**Stochastic Solutions for the Challenges in Proposal Engineering"** submitted for the M.Tech degree is entirely my original work and all ideas and references have been duly acknowledge. It does not contain any work for the award of any other degree or diploma.

Date: Sathish M

11604451

CERTIFICATE

Certified that this project report entitled "**Stochastic Solutions for the Challenges in Proposal Engineering"** submitted independent by student of School of Civil Engineering, Lovely Professional University, Phagwara are carrying out the work under the direction of me for the Award of Degree. The report has not been submitted to a university or institution for the award of any degree.

Signature of Supervisor

Mr. Manoharan Rajalingam

Asso. Professor (Head of Structural Engineering Discipline)

ACKNOWLEDGEMENT

Any accomplishment requires the efforts of people and, this is not different. So it has been also complete with cooperation of many persons. I am thankful to my guide **"Mr. Manoharan Rajalingam"** who gave me the golden opportunity to do this wonderful research on the topic **"Stochastic Solutions for the Challenges in Proposal Engineering"**, which also helped me in doing a lot of Research and I came to know about so many new things.

Secondly I would also like to thank my parents and friends who helped me a lot in finishing this project within the limited time.

Mr. Sathish M

Regd. No. 11604451

M. Tech

(Structural Engineering)

ABSTRACT

Present study is on challenges faced by the construction industries in proposal engineering stage and the techniques used to overcome the difficulties faced in planning and construction stages. Many methods are proposed by the researchers for prediction of problems and quantification of optimized materials. In this study we are more concerned on steel structure quantification. An attempt is being made to improve the accuracy of quantification of steel so that the estimated cost will be reliable as compared to other techniques used in the field. This will result in obtaining better outputs in estimation and costing. The factors affecting the tender are analyzed and higher weightage is given to variable which has higher influence on the project cost. This prediction technique will act as a gateway for the contracting companies to acquire the project since the accuracy of estimation is improved to a greater extent.

TABLE AND CONTENT

LIST OF FIGURES

CHAPTER-1

1. INTRODUCTION:

The main aim of quantity surveyor and cost engineer is to provide a sustainable cost through estimation. The Estimator will always try to make a most likely estimate which is the reflection of reality. Estimating the cost of the project during planning and designing stages is called Pre-Tender cost estimation. Cost estimation prepared at pre-tendering stage is subjected to biases because they are often prepared within limited time frame and without knowing the complete scope of the project. Underestimation in pre-tender cost estimation may lead to non-feasible project if accepted by owner may lead to project failure. Overestimation may lead to rejection of the quoted cost for the tender. The estimated cost of a construction project is influenced by several factors such as the structural, architectural and engineering systems. The structural cost depends on area of the construction, amount of material used etc. The architectural cost depends upon the design and quality of materials used etc. The engineering system cost includes sanitary, electrical, air-conditioning and elevator system.

One of the methods for increasing the accuracy of pre-tender cost estimate is to add contingency reserve by taking in account of the unforeseen cost. There are many methods for estimating contingency reserve namely Percentage allowance method, Monte Carlo, ANN (Artificial Neural Network), Fuzzy logic and Regression.

Percentage allowance method is most widely used method and this method is subjected to the cost engineers experience and judgment. A study by Arafa and Algedra(2011) revealed 7 key parameters influencing the structural cost of the building, namely: ground floor area, typical floor area, number of storeys, number of columns, type of footing, number of elevators and number of rooms. In linear regression model the input given has linear relationship between them which may not always be appropriate. Another disadvantage is they are not suitable to interact among large number of variables. Regression assumes that relationship between input and the output are linear whereas in construction projects the relationships between them are non-linear and sometimes unknown. Multiple linear regressions were developed and are very useful tool for analyzing the contribution of potential new item to the overall estimate. The disadvantage of Fuzzy approach is that the relationship between output and input variable are decided based on the qualitative

information of the project. Artificial Neural Network is non-parametric prediction tool. It is an information processing paradigm built like a biological nervous system. A single neuron is not intelligent. A collection of those neurons is made intelligent by making corporate actions. Collection as a network creates a pattern of inputs to a neural network and processed as a pattern and results as a pattern. The artificial neuron has been modeled mimicking biological neuron similar way working together to produce remarkable results. So Artificial Neural Networks (ANN) is a mathematical model that was developed based on the phenomenon of error minimization.

1.1 Artificial Neural Networks:

ANN is an information process tool inspired by the working of biological nervous system. ANN does not assume about the underlying distribution so it is a non-parametric prediction tool. ANN can accept non-linear variable and mathematically defines cost as a function of variable.

The important element of ANN is their structure. It has large number of interconnected elements called neurons which works simultaneously to solve a specific problem. ANN can accept non-linear variable and mathematically defines cost as a function of variable. ANN can form a relationship between the cost and the cost influencing variables even if the nature of relationships is unknown. Kim et al (2004) discovered that ANN are viable and are better approach for estimating construction cost. However application of ANNs in construction is a relatively new research area (Kim et al., 2004). Several researches demonstrated the potential use of ANN in construction (Moselhi et al, 1992; Hagazy and Ayed, 1998) and their superior performance over traditional regression analysis (De la Garaza and Rouhana, 1995; Creese and Li, 1995; Emsley et al 2002; Setyawati et al). ANN model can reduce the contingency of unforeseen costs between the early estimate and the projected completion time of the project (Aibinu et.al, 2011). Commonly used ANN models in various engineering fields are BPNN(Back-Propagation Neural Network), PNN(Probabilistic Neural Network), GRNN(General Regression Neural Network). The method which is widely used in construction estimation is BPNN. Development of high quality BPNN is difficult. The process of developing BPNN requires experimentation. It requires periodic refinement of network parameters, network redesign and problem reformation. GRNN when compared to BPNN can learn faster with data sets having limited information. GRNN output result increases as number of sample increases. PNN can also learn fast and it can also be further improved when needed.

This study is important because large number of project results in cost overrun. If the inaccuracies in pre-tender estimate are predicted, then the cost advisors would be able to prepare better cost estimate and the project owner can be assured of the costs of their projects during the planning and designing stage.
The patterns in which the neurons are collected are called architecture. Three fundamental types of Neural Network architecture are

1. Single Layered Feedforward Networks: It has the simplest architecture. The input can proceed to the output. The information flow is possible only in one direction. Data will not flow in reverse direction.

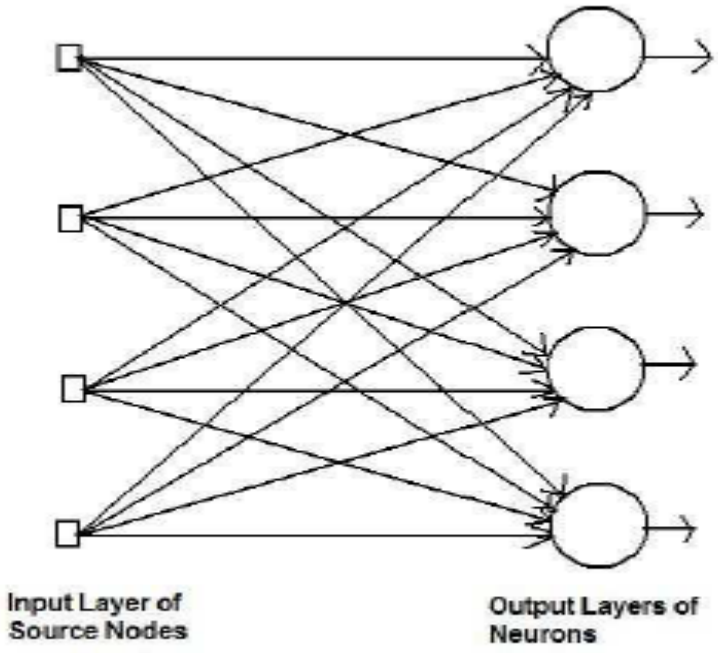

Figure 1 Single Layer FeedForward Network

2. Multilayer FeedForward Network: It has one or more hidden layers connected to the corresponding neurons called as hidden neurons. This helps to extract higher-order statistics and these hidden neurons are used particularly where the input layer is large. The source node gives the respective elements of the activation function. The output of the second layer is used as the input for the third layer. This process continuous and it depends upon the number of hidden layers. The most common algorithm used in training is the error backpropagation in MLP. In training the weights and bias are adjusted.

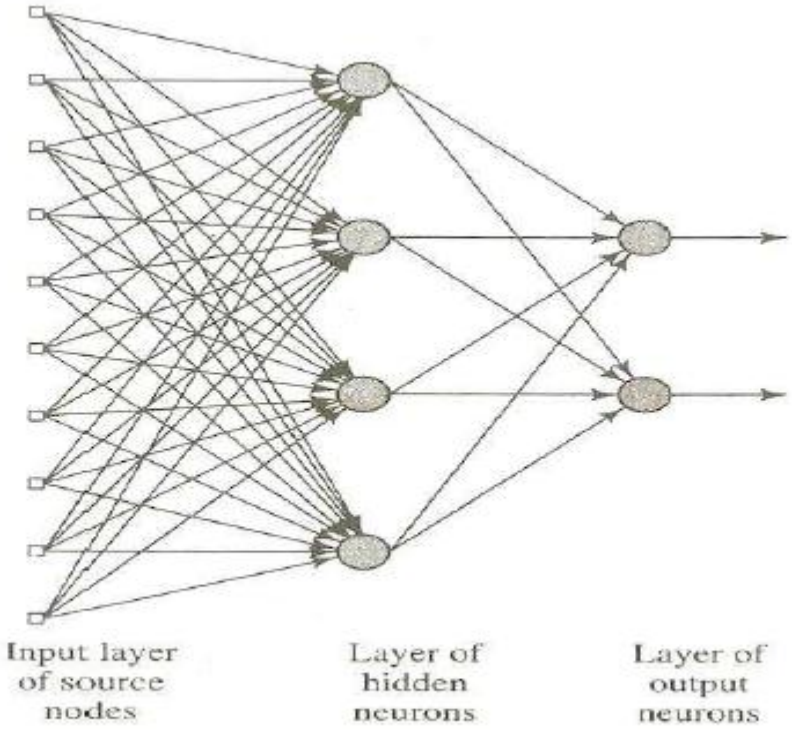

Figure 2 Multilayer FeedForward Network

3. Recurrent Neural Network: Each neuron will feed the output neuron and again the output neuron will feed the input neuron. It is a looping process.

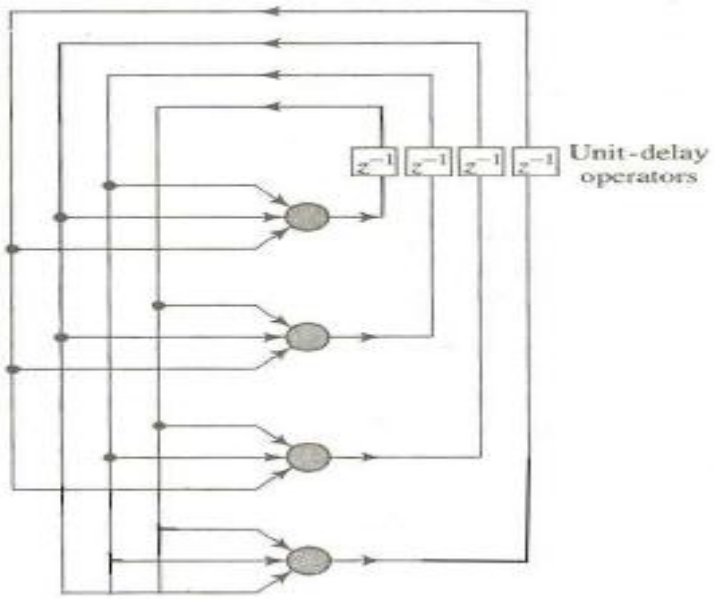

Figure 3 Recurrent Neural Networks

Prediction and Classification Using ANN: Neural Networks has been a successful tool for classification because of their ability to provide optimal solution for any random occurring problems. MLP with error backpropagation has been found to be widely used for classification problems. The reason is MLP has a simple algorithm and the non-linearity between data can be learnt by training them.

Few methods used for training MLP in backpropagation are Gradient Descent, Gradient Descent with Momentum, Gradient Descent with Adaptive Learning Rate and Momentum.

CHAPTER-2

LITERATURE REVIEW:

Arfa (2012), Data were collected with certain constrains to ensure the collected data are of similar nature. They are project implementation period should be less than one year, project has to be completely finished and should be in use, Construction year should not be later than 2002, project should be implemented under full time supervision of engineer under a construction firm. So a total of 92 data where statically analyzed to check the reliability. Frequency distribution was done based on the criteria of input variables. The frequency distribution based on type of building was plotted and it was found that high distribution for schools and low for kindergarten. 12projects like mosque and hospitals were excluded because their frequencies were low and their structural system is different from others. Remaining 71 building project data was collected from construction industry of Gaza strip. It was observed that 70% had column from 30 to 60, 7% had 61-90 column, 14.1% had 91 to 120 column and 8.5% had more than 120 column. It was observed that 77.5% of building do not have elevators and 22.5% had 1 or 2 elevators. . 7 input layers (ground floor, typical floor area, number of storey, number of column, type of footing, and number of rooms). ANN model was developed with one hidden layer and seven input neurons and one output layer. The one output layer which represents the early cost estimate of the building. They used backproagation technique to achieve the output. The weights were adjusted using number of training inputs and corresponding target values. Network error is the difference between the calculated and expected target. MATLAB version 2009 b was used. tan-sigmoid transformation function was used in hidden layers. Linear transformation function was used for output layer. The return value of tan-sigmoid function varies from -1 to +1. Trial and error method was followed to find the optimum hidden layer. The data were divided into three sets. 35 for training, 36 for validation and remaining were for test set. After successful training of ANN it was tested with a new set of data and it was observed mean was 0.960, SD was 0.420, Coefficient of determination was 97% between actual and the output data. Further sensitivity analysis was done by varying one parameter and observing the changes

in the output. Thus it was found Ground Floor area, Number of Elevators, Number of Storey and Type of Foundation had significance influence on the output.

Hany 2014, Considered BPNN (BackPropagation Neural Network) , PNN (Probabilistic Neural Network) , GRNN(Generalized Regression Neural Network) and Regression Analysis. The models were developed for low-rise structural steel building and short span timber bridges. The data used were obtained from previous research conducted by (Creese and Li , 1995) for timber bridges. Three input variables (volume of web, volume of bridge deck and weight of steel) were taken and one output variable (actual cost). Three NN models were developed (PBNN,GRNN,PNN) for different input variables to determine how practically NN is for the cost of the construction. Same inputs were used for Regression analysis. Coefficient of Determination R^2 for all the tree models having input variables as 1,2 and 3 where found using Regression, BPNN, GRNN and PNN was found. R^2 for was found always greater for NN as compared with Regression method. NN output was formed based on least square analysis. Estimating error decreased as the input increased. MAPE(Mean Absolute Percentage Error) of estimated cost was compared with the actual cost. The result obtained from MAPE was that PNN showed best performance among all the models.35 low-rise structural steel building with sophisticated models and analysis were chosen and the data were collected from large manufacturer firms. . The collected data was from 1993 to 1999. Mainly 4 types of documents were collected (contract including project specification and change in order, Blueprints, Detailed Estimate and Final cost estimate). The final cost estimate with and without markups. It was assumed only direct cost of the building was used for development of NN inorder to increase the consistency in cost estimating. A data entry sheet was designed for standard data collection, Organization, indexing, recording and analyzing. The data was analyzed for consistency. Individual calculation of total structural cost was calculated per square feet for every building. Project with different characteristic was rejected (4 projects). 35 projects were used for development of NN model. Building Area, Perimeter, Joist span and height were identified as main parameters which were correlated with the fabrication cost of steel structural building. In order

to improve the performance of NN the input data were scaled and transferred within (1, 10) intervals (Moselhi and Seqeira, 1998).

2 layer architecture with 4 nodes as input layer and 1 as output were developed for BPNN design. Levenberg-Marquardt [LM] training algorithm was used as it is fast in learning and converge the data. 14 different architectures were developed by varying the hidden nodes (3 to 16) to get the best performing model. The data were randomly divided by MATLAB. MSE and R^2 were calculated. It was observed with NN having 12 nodes as hidden layer gave least MSE. The training time was 37 seconds on IBM compatible computers with RAM of 2GB and the best validation check was 0.11395. Interruption of network occurred after 1055 learning epochs (certain time period). Least MSE with training set was 0.00497.NN was less sensitive than Regression analysis when inputs were changed. PNN was most stable with MAPE difference of 2.46%.

Irem Dikmen (2004), Used ANN as a strategic decision supporting tool that can classify international project based on their Attractiveness of the project and the Competitiveness of the company. This increases the organizational intelligence using the experience of Turkish contractors in international market. Sixteen criteria were taken based on the questionnaires survey conducted on the factors affecting the competitiveness of the project. Data from 600projects completed abroad by Turkish contractors from 1995 to 2000 was collected. It is assumed that the company first time enters the international market. Backpropagation algorithm was used with twenty five input nodes, 16 input criteria and two output nodes. The output parameters were over defined by Likert Scale of 1-5. The database was divided into two sets (training and testing). 75% was used for training and 25% was used for testing. The number of hidden layers was decided based on trial and error method. The RMS(Root Mean Square) error less than 1 was taken as acceptable performance of the network. Software used to develop the network was "Neural Networks Professional II Plus (2000)". This software was used because it decreases the risk of over training. To test the sensitivity of the network, each input value was increased 5% and the change in the output was monitored. This result

will give the most important factors affecting attractiveness of the project (market volume, availability of fund, country risk and economic prosperity). The revised network has 10 input nodes, 5 hidden node and 2 output nodes.

The researcher also tried using tanh instead of sigmoid function and the difference in the learning rule like delta-bar-delta rule instead of general delta rule were used. But none of them led to significant improvement. Turkish contractors considered availability of fund was one of the major variables of the project attractiveness. They also selected the international market with high growth rate potential. They compared this result with British contractor research results. Turkish contractors preferred low-risk countries whereas British contractors were on high-risk countries. A NN model was developed for entry decision for international Turkish contractors. This research will guide the contractors to know the type of data to be collected to enter into the international market. To improve the model performance, fuzzy function could be incorporated into the model, which may better represent the opinion of the experts.

CHAPTER-3

Development of Hypothesis:

From the literature review conducted, it has been decided to prepare a model using Artificial Neural Network (ANN) for calculating the total construction cost of the project using a programming language. Software used in past research papers was MATLAB,C++,Neural Works Professional II Plus(2000) It is decided to collect structural building data for low-rise steel building and concrete construction building from the construction firms located at the selected area. The number of input variables to be used is taken as 7 for concrete buildings (Ground Floor Area, Typical Floor Area, Number of Storey, Number of Column, Type of Footing, Number of Elevators and Type of Building) and 4 (Building Area, Joist Span, Perimeter and Height) inputs for steel buildings. The number of input variable may increase or decrease after analyzing the data collected from the decided location as the input variable depends on various aspects like Market Volume, Availability of Fund, Country Risk Rating, Economic Prosperity, Attitude of Host Government, Experience of the Company, e.t.c)

The architecture decided to be used is Multilayer FeedForward Network. The activation functions will squashes then variables between 0 to 1 or -1 to +1. It depends upon the activation function used. Here the activation function used will be tanh activation function or sigmoid activation function. The number of hidden layers will be decided based on trial and error method. Bias percentage will also be decided based on the data collected. The number of output layer will be one (total cost of construction).

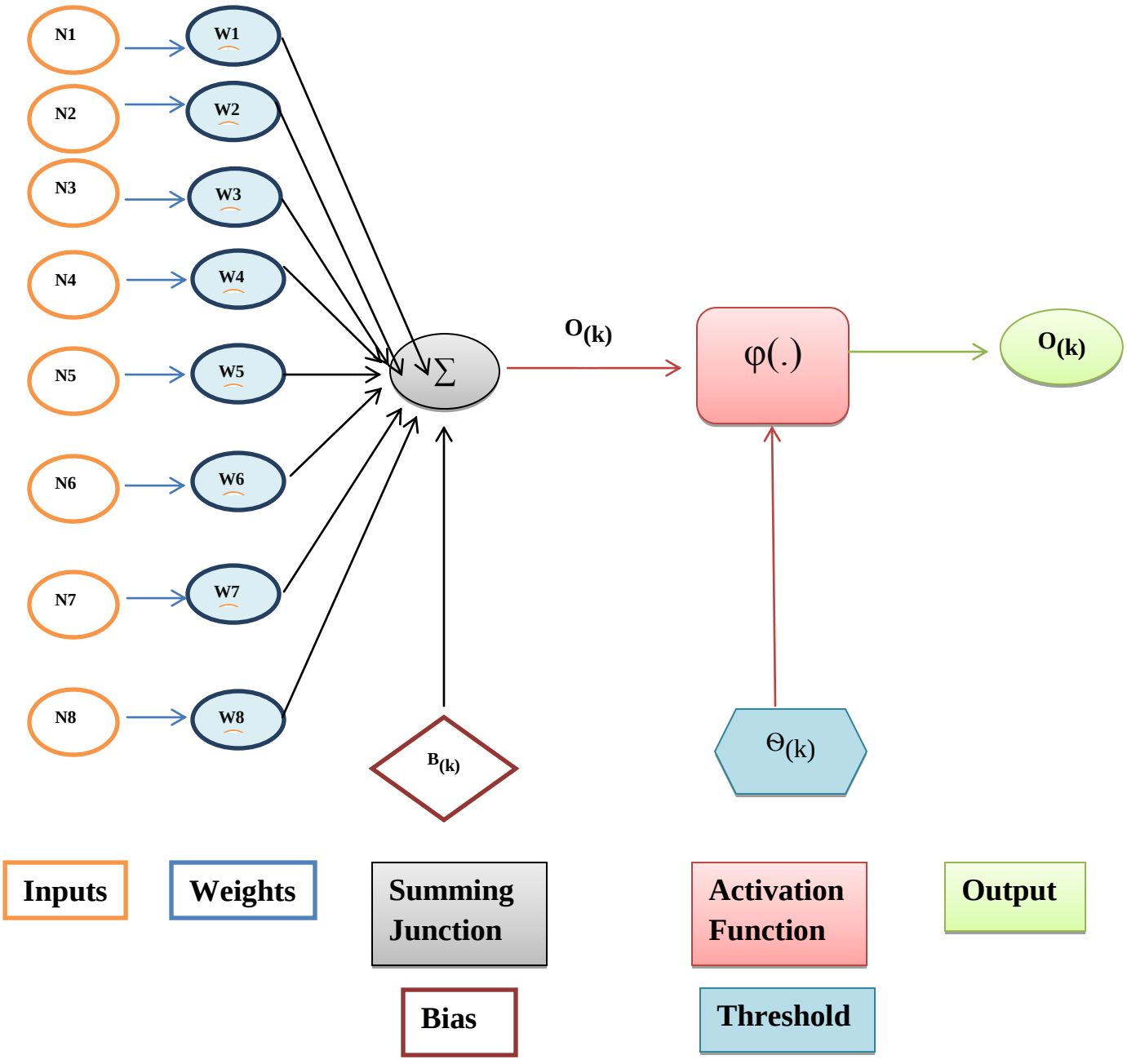

Figure 4 Structure of Artificial Neural Network

CHAPTER-4

RESEARCH DESIGN

4.1 Preparation of Research Design:

It is decided to collect maximum number of construction building data form the construction firms from 2010 to 2016 in order to have a consistency in the data collected. The main documents to be collected are

a) Contract including project specification and change in orders
b) Blueprints
c) Detailed Estimate
d) Final Cost Report

These data collected will be manually crosschecked for eliminating false data and odd characteristic building data.

4.2 Data Collection:

The data is planned to be collected from the construction firms in Chennai. If the data collected are insufficient, the network will start memorizing instead for learning the importance of the variable.

4.3 Data Processing and Analysis:

The collected data are fed into the computer using a data entry sheet so as the data collected can be organized, analyzed, indexed and recorded. It is decided to use ANN in a programming language. The data will be divided into three sets (Training, Validation and Test Set)

- Training: The set of data is used for modifying the weights and bias in order to minimize the network error.
- Validation: Error will be monitored and if the error rises the limits then the training is stopped and a reasonable level of performance error function will be adapted.
- Test Sets: Further checking and Generalization of Network is done.

It is planned to use 50% of the data for training, 30% for validation and 20% for testing.

4.4 Data Interpretation:

Analysis of data will be done after training, validation and testing is finished. A new set of data will be given to the network. The given data should not be a data which the network has already faced. Correlation Coefficient, Mean Square Error and Mean Absolute Percentage Error are calculated using mathematical formulas respectively. Thus based on these values the performance of the developed Network will be judged.

Further the sensitivity of the network will be checked based on the varying the inputs and the changes caused in the output. This will give us the importance percentage of each input variable. This can also lead to reduce the input variable that in turn helps in reducing the complexity of the network developed.

CHAPTER-5

CONCLUSION:

From the study of the past research paper it is clear that pre-bidding process is subjected to bias due to many assumptions that we are forced to make because of the limitation in data and time constraints. So predictions have to be made with limited data on hand. Many methods are used for prediction in many situations out of which regression analysis and artificial neural network are said to be the most effective methods. The developing and current method used in this field is Artificial Neural Network method. Occasionally it is coupled with Fuzzy logic and Genetic algorithms also. In ANN method the computer is trained to solve problems based on the past data available. Simulation software like MATLAB and other software are being used in application of ANN in various fields. Artificial Neural networks offer a number of advantages, including requiring less formal statistical training, ability to implicitly detect complex nonlinear relationships between dependent and independent variables, ability to detect all possible interactions between predictor variables, and the availability of multiple training algorithms.

The main advantage of using ANN is to study the influence of various parameters on the final cost of the construction project by dealing effectively with limited input data within limited time frame and by examining and assigning weightage suitably for the parameters which affect the cost of the construction project adversely.

Thus ANN model can be developed and used in big construction industries so as to improve the standard of estimation which in turn benefits the organization by achieving success in the bidding process among the competitors.

REFERENCES:

Adnan Enshassi, Sherif Mohamed and Munther Abdel-Hadi(2013) Factors affecting the accuracy of pretender cost estimate in Gaza strip. Journal of Construction in Developing Countries 18(1), 73-94, 2013

Afsath K, Linjiesh Sebastin (2015) Fuzzy MATLAB approach to construction contractor selection. International Journal of Scientific Engineering and Research (IJSER) input factor ISSN:2347-3878(2015):3.791

Ajibade A Yodeji Aibinu and Dharma Dassanayake(2011) Use of artificial intelligence to predict the accuracy of pretender building cost estimate. Management and Innovation for a Sustainable Built Enviroinment 20-23 June 2011, Amsterdam, The Netherland ISBN:9789052693958

Al-Bulushi, P.R King, M.J.Blunt, M.Kraaijveld(2012) Artificial Neural Networks Workflow and its applications in petroleum industry. Neural Comput & Applic (2012) 21:409-421

Alexander Laufer, Aviad Shapira, Dora Cohenca-Zall and Gregory A.Howell (1993)Prebid and Preconstruction planning process .ASCE

Chen, Kw Tang, R Fellows(2005) Construction project tendering in china: An exploration of theory and practice relating to article 33 of china's tendering law. Hong Kong Surveyor Vol 16(1), 11-22 June 2005 ISSN 1812-3953

Cliff J.Schexnayder, SandraL.Weber, Christine Fiori (2003) Project Cost Estimating a synthesis of highway practice June 2003

Datta, R.Roy (2009) Cost modeling technique for availability type service support contract: a literature review and empirical study. Proceeding of the 1st CIRP Industrial Product Service System (IPS2) Conference, Cranfield University, 1-2 April 2009, PP216

David Baccarini (2014) The maturing concept of estimating project cost contingency – A review 18 December 2014

Hojjat Adeli (2000) Neural Network in Civil Engineering: 1989-2000. Computer-Aided Civil and Infrastructure Engineering 16(2001)126-142

Jason Ongpeng (2015) An artificial neural network approach to structural cost estimation of building in Philippines. Presented at the DSLU Research Congress

Martin R, Skitmore (1998) Fundamental research in Bidding and Estimating in proceeding British/Israeli seminar on building economics. The International Council for Construction Research Studies and Documentation CIB W-55

Martin Skitmore, David Picken(2000) The accuracy of pretender building price forecast: An analysis of USA data. The International Journal Council for Building Research Studies and Documentation (CIB)W-92 procurement system symposium 1, pages pp 595-560

Michal Juszczyk, Agnieszka Lesniak, Krzyszt Zima (2014) Designing and Build system in polish public sector – Case study. International Scientific Conference People, Building and Environment (PBE2014) 15-17 October, kromeriz Czech Republic

Peter Raisbeck and Ajibade Ayodeii Aibinu(2010) Early stage cost estimation and relationship of architects to quantity surveyors. Egbu, C.(Ed)procs26th Annual ARCOM Conference 6-8Septembe 2010,Leeds,U.K, Association of Researchers in Construction Management , 53-61

Sung-Hoon An, U-Yeol park, Kyung –In Kang, Moon-Young Cho and Hun-Hee Cho(2007) Application of support vector machines in assessing conceptual cost estimate DOI:10.1061(ASCE)0087-3801(2007) 21.4(259)

SC Ugochukwa and K.C Okolie (2013) Investigation into factors influencing tendering duration stipulations by public building clients Nigeria. IOSR Journal of Engineering (IOSRJEN) October

Tarek Hegazy and Amar Ayed (1998) Neural Network Model for parametric cost estimation of highway project. J.Constr.Eng.Manage, 124(3): 210-218

Ajibade A Yodeji Aibinu and Dharma Dassanayake(2011) Use of artificial intelligence to predict the accuracy of pretender building cost estimate. Management and Innovation for a Sustainable Built Enviroinment 20-23 June 2011, Amsterdam, The Netherland ISBN:9789052693958

Hany El-Sawah Comparative study in the use of Neural Networks for order of magnitude cost estimating in construction. www.itcon.org – Journal of Information Technology in Construction- ISSN 1874-4753

Irem Dikmen and M.Talt Blrgonul, Neural Network Model to Support International Market Entry Decision. Journal of Construction Engineering and Management, Vol. 130, No. 1, February 1, 2001. @ASCE, ISSN 0733-9364/2001/1-59-66

Mohammed Arafa and Mamoun Alqedra , Journal of Artificial Intelligence, 2011. ISSN 1994-5450/DOI: 10.3923/jai.2011