IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 6, June 2015.
www.ijiset.com
ISSN 2348 – 7968

Software Cost Estimation Using Artificial Neural Network

Devbrat Sahu
M.Tech Research Scholar, CSE,Chattrapati Shivaji Institute of Technology
Durg, India

Mrs. Deepty Dubey
Associate Professor, CSE,Chattrapati Shivaji Institute of Technology
Durg, India

Abstract— **The software project cost and effort estimation is to scientifically estimate the required workload and its corresponding costs in the life cycle of software system. Software cost estimation is a complex activity that requires knowledge of a number of key attributes that affect the outcomes of software projects, both individually and in concert. The most critical problem is the lot of data is needed, which is often impossible to get in needed quantities. Artificial Neural Network model is one of building design that is utilized all the more as a part of software cost estimation .In this paper Artificial Neural Network (ANN) has been used using dataset for prediction of software effort. The performance indices such as Sum-Square-Error (SSE), Mean-Square-Error (MSE), Root-Mean-Square-Error (RMSE) and Mean-Absolute-Error (MAE) are calculated from ANN technique and the results obtained from these parameters are discussed for software cost estimation.**

Keywords—Software effort estimation, Artificial Neural Network, Source lines of code.

I. INTRODUCTION

Software has played an increasingly important role in systems acquisition, engineering and development, particularly for large, complex systems. For such systems, accurate estimates of the software costs are a critical part of effective program management [4]. The bulk of the cost of software development is due to the human effort and most cost estimation methods focus on cost and effort aspect and give estimates in terms of person-months. Accurate cost estimates are critical to both developers and customers. They can be used for request for proposal, contract negotiations, scheduling, monitoring and control. Underestimating the costs may result in management approving proposed systems that then exceed their budgets, with underdeveloped functions and poor quality, and failure to complete on time. Overestimating may result in too many resources committed to the projects or during contract bidding result is not winning the contract, which can lead to loss of jobs. Software project failures have been an important subject in the last decade. Software projects usually don't fail during the implementation and most project fails are related to the planning and estimation steps. Hence accurate cost estimation is become a challenge for IT industries.

Cost estimation is usually measured in terms of effort [2]. The most common metric used is person months or years (or man months or years). The effort is the amount of time for one person to work for a certain period of time.

An artificial neuron network (ANN) is a computational model based on the structure and functions of biological neural networks.

II. RESEARCH METHODOLOGY

A. ARTIFICIAL NEURAL NETWORK

A neural network is a simplified model of the biological neuron system. It is a massively parallel distributed processing system made up of highly interconnected neural computing elements that have ability to learn and thereby acquire knowledge and make it available for use. McCulloch and Pitts (1943) proposed the first computational model of a neuron, namely the binary threshold unit, whose output was either 0 or 1 depending on whether [3] its net input exceeded a given threshold. The most common algorithm for training or learning is known as error back-propagation algorithm. The error back propagation learning consists of two phases: a forward pass and a backward pass, an input is presented to the neural network, and its effect is propagated through the network layer by layer. This is also called Testing Phase. During the forward pass the weights of the network are all fixed. During the backward pass or "Training Phase", the weights are all updated and adjusted according to the error computed. An error is composed from the difference between the desired response and the system output. This error information is feedback to the system and adjusts the system parameters in a learning rule. The process is repeated until the performance is acceptable.

B. EVALUATING THE PERFORMANCE OF THE MODELS

The main measures used for evaluating the performance of machine learning techniques for predicting the software effort are as follows[3]:-

ISSN 2348 – 7968

1. **Sum Squared Error (SSE):-** The sum squared error is defined as.

$$E = \left(\sum_{i=1}^{n} (Pi - Ai)^2 \right)$$

Where Pi = Estimated value for data point i;
Ai =Actual value for the data point i;
n = Total number of data points.

2. **Mean Squared Error (MSE):-** The mean squared error is defined as.

$$E = \left(\frac{1}{n} \sum_{i=1}^{n} (Pi - Ai)^2 \right)$$

Where Pi = Estimated value for data point i;
Ai =Actual value for the data point i;
n = Total number of data points.

3. **Root Mean Squared Error (RMSE):-** The root mean squared error is defined as.

$$E = \left(\sqrt{\frac{1}{n} \sum_{i=1}^{n} (Pi - Ai)^2} \right)$$

Where Pi = Estimated value for data point i;
Ai =Actual value for the data point i;
n = Total number of data points.

4. **Mean Absolute Error (MAE):-** The mean absolute error measures of how far the estimates [3] are from actual values. It could be applied to any two pairs of numbers, where one set is "actual" and the other is an estimate prediction.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left| Pi - Ai \right|$$

Where Pi = Estimated value for data point i;
Ai =Actual value for the data point i;
n = Total number of data points.

Table 1 Data Set of Software Cost Estimation

KLOC	ME	a	b	AE	EE
902000	30	3.3602	0.8116	115.8000	115.4923
462000	20	3.3820	0.7920	96	95.6978
465000	19	3.4520	0.7890	79	78.3312
545000	20	3.7400	0.8240	90.8000	82.5803
311000	35	3.3240	0.8120	36.6000	18.3678
675000	29	3.2480	0.8116	98.4000	88.9900
128000	26	3.1580	0.8470	18.9000	10.1283
105000	34	3.3240	0.8321	10.3000	7.1042
215000	31	3.3167	0.8321	28.5000	20.4246
31000	26	3.2980	0.8120	7	5.8316

III. RESULT ANALYSIS

The Artificial Neural Network (ANN) has been used for predicting the software efforts using dataset. The performance indices and the results obtained from these models. These indices are Sum-Square-Error (SSE), Mean-Square-Error (MSE), Root-Mean-Square-Error (RMSE), Mean-Absolute-Error (MAE).
MATLAB programs were developed for training and testing of various models and also for computation of performance indices. The results are tabulated in Table 2.

Table 2 Performance indices with Artificial Neural Network

S. No.	Performance Measures	Artificial Neural Network (ANN) Ten Hidden Layer
1.	**Sum Squared Error (SSE)**	3404.9442
2.	**Mean Squared Error (MSE)**	340.4944
3.	**Root Mean Squared Error (RMSE)**	18.4524
4.	**Mean Absolute Error (MAE)**	3404.9

IV. CONCLUSION

The Artificial Neural Network (ANN) has been used to analyze the results using dataset for software development effort. A similar study can be carried out to predict software effort using prediction models based on other machine learning algorithms such as Genetic Algorithms. Cost benefit

ISSN 2348 – 7968

analysis of models may be carried out to determine whether a given effort prediction model would be economically viable.

References

[1] J. S. Pahariya, V. Ravi, M. Carr, M. Vasu, Computational Intelligence Hybrids Applied to Software Cost Estimation, International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM),Vol. 2, pp. 104-112, 2010

[2] Zeynab Abbasi khalifelu, Farhad Soleimanian Gharehchopogh, Comparison and evaluation of data mining techniques with algorithmic models in software cost estimation, 2012

[3] Prabhakar, Maitreyee Dutta, Prediction of software effort using artificial neural network and support vector machine, IJARCSSE, MARCH 2013.

[4] Jyoti G .Borade, Vikas R. Khalkar, software project effort and cost estimation techniques , IJARCSSE, august 2013.

[5] http://www.techopedia.com/definition/5967/artificial-neural-network.

[6] www.sciencedirect.com

[7] www.wikipedia.com

[8] Vahid Khatibi , Dayang A.N.Jawawi, software cost estimation methods,JETCIS,2010-11.

[9] P. Shubitsha , J. Kowski Rajan , Artificial Neural Network Models for software effort estimation , IJTEEE,2014.

[10] Xiaotie Qina , Miao Fangb , Summarization of Software Cost Estimation , Elsevier Ltd,2011.