Analysis and Comparison of Neural Network Models for Software Development Effort Estimation

Kamlesh Dutta, Department of CSE, National Institute of Technology, Hamirpur, India
Varun Gupta, Amity School of Engineering and Technology, Amity University, Noida, India
Vachik S. Dave, Department of CSE, National Institute of Technology, Hamirpur, India

ABSTRACT

Prediction of software development is the key task for the effective management of any software industry. The accuracy and reliability of the prediction mechanisms used for the estimation of software development effort is also important. A series of experiments are conducted to gradually progress towards the improved accurate estimation of the software development effort. However, while conducting these experiments, it was found that the size of the training set was not sufficient to train a large and complex artificial neural network (ANN). To overcome the problem of the size of the available training data set, a novel multilayered architecture based on a neural network model is proposed. The accuracy of the proposed multi-layered model is assessed using different criteria, which proves the pre-eminence of the proposed model.

KEYWORD
Feed-Forward Neural Network, Neural Networks, Radial Basis Function Neural Network, Regression Test, Software Development Effort Estimation

1. INTRODUCTION

An accurate estimation is the key objective of any prediction model. Software development effort estimation is one of the important research domains for software organizations. However, there are several problems associated with making an accurate estimation. Stutzke (1996) a) attributed ambiguity about the meaning of estimation; b) Confliction of project goal; c) Lack of requirement information; d) Inclusion of Reusable code; e) New development procedures and tools; and f) Relation of inputs to output. Software development effort prediction depends on several factors such as software size, people involved in software development, use of technology, software complexity, and many more. Further, these factors are correlated to each other and influence the software development process directly or indirectly. The effort estimate can be derived from the size of software products, if it is estimated reliably. Defining software development lifecycle and development process that is followed to specify, design, develop, and test the software is necessary for the estimation of the software development effort from software size. Developing new software is not just a task of coding the software functionalities. Rather, coding the software is just a single part of the whole software development process and its effort. Designing the software or deliverables, implementation of the prototypes, writing & reviewing documentation, and reviewing & testing the code take up the larger portion of overall project effort.

DOI: 10.4018/JCIT.2019040106

Copyright © 2019, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
The project effort estimate requires identifying and estimate, and then sum up all the activities, one must perform to build a product of the estimated size.

The main objective of this paper is to develop neural network based software effort estimation model. A series of experiments are conducted to explore the suitable model and reported in the subsequent sections. The paper is organized as follows. In Section 2 we have enlighten the portion related to the neural networks based estimations of software development effort as reported in the literature. Section 3 gives an overview of various criteria’s used for assessing the suitable prediction model for software development estimation. Section 4 presents a series of experiments conducted to analyze the performance of various models. Most of the researchers emphasized the benefit or dominance of the ANN model over other traditional models. Here, the experimental studies are conducted to find out the best suitable ANN model for software effort estimation. Based on the analysis, a multilayered ANN model for the software effort estimation task is proposed in section 5. The model makes use of three categories for input parameters. The categories are based on the impact factor of the inputs. The suitability of the proposed model is further tested and verified through a series of simulation. Section 6 summarizes the results and concludes with the future scope of work in improving the accuracy of software development effort estimation.

2. RELATED WORK

Various approaches have been adopted in literature to make an accurate estimation of software effort. In recent years, application of machine learning approaches has been attempted. This has been possible due to the availability of data sets of a large number of completed projects. Among various machine learning techniques, neural networks based models are newly emerging models. Though several researchers (Venkatachalam, 1993; Finnie et al., 1997; Samson et al, 1997; Lee et al.,1998; Heiat, 2002; Ideri et al, 2002; Idri et al, 2004; Tadayon, 2005; Idri et al, 2006; Kanmani et al, 2007; Tronto et al., 2007; Park and Beak, 2008; Tronto et al, 2008; Iwata et al, 2009; Reddy and Raju, 2009; Ajitha et. al, 2010; Kaur et al, 2010; Bhatnagar et al, 2010; Balich and Martin, 2010; Reddy et al, 2010; Pendharkar, 2010; Attarzadeh and Ow, 2010; Dave and Dutta, 2011a, 2011b, 2011c; Shepperd and MacDonell, 2012; López-Martín, 2014) have worked on Neural Networks based models. However, there is still a need for more research work and attempts to find out the most suitable model for software effort prediction in term of accuracy, configurability and reasoning associated with the suitability of model. The detailed review of 21 articles related to the effort/cost estimation of software using neural networks is presented in Dave and Dutta (2014). The paper also suggests to research on neural networks based models for more accuracy.

Malhotra et. al. (2011) have used dataset of 499 projects for comparing the regression model with Machine learning methods and concluded that machine learning methods are giving satisfactory results. Mahajan et. al. (2011) have derived a model from neural networks and shown its dominance over COCOMO II model. Kaur et al. (2010) have shown that a neural network is better than Halstead model, Walston-Felix model, Bailey-Basili model and Doty model, same as Vahib Khatibi et al. (2011) have investigated 15 research articles and compared it with neural networks and concluded that neural networks outperforms most algorithmic methods.

Bhatnagar et al. (2010), and Barcelos Tronto et al. (2008) have also proved that neural network models are better than Regression analysis model. Kaur et. al.(2012) conducted a study for the comparison of neural network with other traditional methods and concluded the same that neuron based model is better than other methods.

3. EVALUATION CRITERIA FOR ACCURACY ANALYSIS

As a software development effort estimation is an important task in the software development project, the accurate effort estimation is still a great challenge for the software management team as well as