Analysis of Overturns in Real Estate Project Using Artificial Network

Kavya S Raj

1Sivaji College of Engineering

ABSTRACT
Cost overrun is one of the most important problem that encounter Real Estate projects success, since it reduces the contractor’s profit and sometimes lead to enormous losses. Cost overruns occur in every real estate project while the magnitude varies significantly from project to project. Throughout this research we tried to gather and analyze the main factors which cause cost overruns in real estate projects such as the factors determining percentage overrun include the financial condition of the owner ,the cash flow of the contractor, material cost increase, competition at tender stage, the project size, delays in design approval, quantity variations, the detailed degree of the drawings used for estimating the budget, , material estimate accuracy, quality requirements being hard to reach, location of the project with respect to vendors, time needed for decisions to be made, the client’s characteristics, unknown geological conditions, ignorance and lack of knowledge of the parties, incompatible advanced payment, the prequalification of the contractor, workload in the project, the contract type , whether the parties agreed on dispute settlement procedure or not, whether the site was properly managed or not , the adequacy of the equipment used in the project, the adequacy of the safety procedures followed, the effectiveness of the planning and scheduling, the availability of equipment, shortage of labors, whether the labors used were skilled or not, the adequacy of the method of construction, the decreased productivity of labors and equipment, the availability of cost control engineers assigned for the project. The objectives were achieved through valid questionnaire survey conducted in contracted companies, consultants, owners.

1 Introduction
The construction industry has a tremendous impact on the economy of most of countries which is due to the fact that it provides the mandatory elements for the development of an economy. The construction industry in many countries accounts for 6-9 % of the Gross Domestic Product (GDP) reaching up to 10 % of the GDP of a few countries. For that, the construction industry is a vital element of the economy and has a significant effect on the efficiency and productivity of other industry sectors. As of Egypt the real estate industry is one of the most stable industries in the few past years as it accounts for the savings of the majority of the Egyptians .Not to mention, the increasing demand for new real estate projects due to the rapid increase in population. In developing countries the trend is more severe where the cost overruns sometimes exceed 100% of the estimated cost of the project. It is essential to have control on cost performance of projects to ensure that the completion of the project is within the estimated
budget. Therefore, project cost management is needed to keep the project within its estimated budget. The processes of project cost management include project resource planning, cost estimating, cost control and cost budgeting. The major thrust of this study was to identify, analyze and model the factors that contribute to building construction cost overruns for Real Estate projects in Egypt. For that, a model is developed using Artificial Neural Networks which predicts the cost overrun percentage based on a database of projects supplied and the opinions of practitioners in the market.

An Artificial neural network (ANN) is a mathematical model which predicts the system performance (i.e., system output). It is inspired by the structure and function of human biological neural networks. The ANN is developed and derived to have a function similar to the human brain by memorizing and learning various tasks and behaving accordingly. It is trained in such a way that it can predict specific behavior and also remembers that behavior in the future like the human brain does. As far as functionality and inter-neuron connection is concerned, ANN’s architecture is similar to human neuron layers in the brain.

2 Work Explanation

2.1 Theoretical Background

Cost Management

Cost Management in a project includes the processes required to make sure that the project is completed within the assigned and approved budget. Cost management includes the following 4 major processes:

a. Resource Planning
b. Cost Estimating
c. Cost Budgeting
d. Cost Control

Resource Planning

Resource planning process involves determining the resources including labor, equipment, materials) and the quantities of each that should be used to complete project activities.

Cost Estimating

Cost estimating is the process where an approximate calculation for the costs of all the resources needed to complete project activities is estimated.

Cost Budgeting

Cost budgeting is a process that involves allocating the overall cost estimates to individual work items in order to establish a cost baseline which is later used for measuring project performance using different techniques.
Cost Control

Cost control is the process of dealing with the influencing factors which create changes to the cost baseline to try to ensure that the changes made are beneficial, determining that the cost baseline has changed, and managing the actual changes when and as they occur.

2.2 Artificial Neural Networks

Artificial neural networks are computer programs inspired from biology that are designed to simulate the way in which the human brain processes information. Artificial neural networks develop knowledge by determining the patterns and relationships in data. They get trained through experience, not from programming rules. Artificial Neural Networks has been broadly used for estimating costs as it presents itself as an approach of calculation and decision making. Artificial Neural Network (ANN) can be used here in cost forecasting because it is an analogy based process that works well with limited data- in the same way as the human brain. Our brain has been developed after thousands of years of evolution, and holds several advantages over the von Neumann or modern parallel computers. These include massive tolerance, learning and generalization ability, adaptively and low energy consumption (Jain, 1996). Neural networks use a process which is analogous to the brain, wherein the existing data undergoes “training”. By using an adaptive learning rate for training, a trained neural network is developed for a better test Accuracy. (Takas’, 2018) The trained network becomes an “expert” in the category of information that it has been given to analyze. This “expert” can then be used to provide projections for new scenarios. Artificial neural networks are purely data driven models which transit from a random state to a final model through iterative training. Artificial neural networks are non-parametric statistical estimators which have proven to be universal approximations, and thus have potential for use in cost estimation modelling. There are two Artificial Neural Network topologies used:

- Feed-Forward ANN
- Feed-back ANN
- Feed-Forward ANN

The information flow is unidirectional. A unit sends information to other unit from which it does not receive any information. There are no feedback loops. They are used in pattern generation/recognition/classification. They have fixed inputs and outputs.
2.2.1 Structure of a neuron

The Input connections have an input value that is received either from the outside in the case of the input layer or from the previous neuron. The Bias is not linked to the other available neurons in the network and is assumed to have an input value of 1 for the summation function. Weights are real numbers representing the importance or strength of the input connection to each neuron. Weighing factors: Each of the inputs in the feature vector is assigned its own relative weight, which determines the impact of that input in the summation function.

\[ Y = W_0 + W_1X_1 + W_2X_2 + W_3X_3 \]

\[ Y = [W \ b] [x \ 1]^T \]

2.2.2 Working of ANNs

For the network topology diagram shown below, each arrow indicates a connection between two neurons and represents the pathway for the flow of information. Each connection has a weight, an integer number that controls the signal between the two neurons. If the network generates a desired output, there is no need to adjust the weights. However, if the network generates an undesired output or an error, then the system alters the weights in order to improve subsequent results.

**Figure 1:** shows ANN Working
In this single hidden layer feed forward neural network, the network's inputs are directly connected to the output layer nodes, Z1 and Z2. The output nodes use activation functions $g_1$ and $g_2$ to yield the outputs $Y_1$ and $Y_2$.

$$Z_1 = W_1, X_i + \mu_1$$
$$Z_2 = W_2, X_i + \mu_2$$
$$3i=1$$

$$Y_1= g_1 (Z_1) = g_1 (1, iX_i + \mu_1 3i=1)$$
$$Y_2= g_2 (Z_2) = g_2 (2, iX_i + \mu_2 23i=1)$$

2.2.3 Activation functions

Activation function: The result of the summation function, that is the weighted sum, is transformed to a desired output by employing a nonlinear function, also known as activation function. An artificial neuron simply calculates a weighted sum of its input, adds a bias and then decides whether it should be fired or not.

$$P^\hat{} = \int \text{NL} (y)$$
$$Y = \Sigma (\text{weights} \ast \text{input}) + \text{bias}$$

The output value of $Y$ can be anything ranging from $-\infty$ to $+\infty$. To limit the neuron value there had to be something else that could be used. The activation functions are introduced to check the $Y$ value produced by a neuron and decide whether outside connections should consider this neuron as fired or not (activated or not).

2.2.4 Optimization

The assigned optimization algorithm keeps on repeating a two phase cycle, propagation and weight update. When an input vector is entered to the network, $t$ is propagated forward through the network, one layer by layer, until it reaches the output layer. The output of the network is then compared to the desired output, using a loss function. The variance is used to decrease loss hence improve accuracy.

**Figure 1:** shows optimization

2.2.5 Loss function

For back propagation, the loss function calculates the difference between the network output and its expected output, after a case propagates through the network.
The function that is used to compute this error is known as Loss Function $J (w)$. The mean square error method is adopted in this research.

### 2.3 Artificial Neural Networks in Real Estate Appraisal

Researches exploited the computational capacity and robustness of the artificial neural networks and used them in real estate valuation or appraisal. This review has successfully classified the application into three different areas. They are:

A) Estimation of the value of real estate.
B) Price prediction or forecasting of real estate.
C) Other applications.

The research articles used for the classification were sourced from different peer-reviewed academic databases.

**Estimation of Value of Real Estate**

This area is further classified into three, namely; evaluation models, mass appraisal, and risk assessment and evaluation. All are different dimensions of real estate valuation and appraisal.

- **Evaluation Models**

Evaluation models are dedicated to estimating the price of real estate properties. Many quantitative and qualitative variables affect the value of an estate. Estate values or practitioners using evaluation models routinely do the estimation of the value of an estate. The models capture different quantitative and qualitative variables that affect the value of the given property. This means that multivariate models are preferred in estate appraisal. Regression analysis, time series and quantitative comparative approach are some of the available tools used in this aspect. A combination of regression analysis and ANN performed better than the individual models. It has also been shown that some data mining models and evolutionary computational methods have shown to be effective when combined with regression analysis or used in solitude for real estate valuation. These include Support Vector Machine K nearest neighbor, decision tree, particle swarm optimization algorithm and genetic algorithm. The precision of the data mining models could be improved by bagging or a combination of bagging and other algorithms. The use of regression analysis or other similar models is plagued with nonlinearity of relationship among the variables, the presence of unnecessary information and the absence of key information that greatly affects the evaluation. Autocorrelation of some variables also presents some challenges during real estate valuation. Hence, ANN tackles the problem, for instance, when it is combined with geographic information system (GIS). Failure to capture the necessary information by the given model impedes the accurate determination of the financial viability of real estate, increases the risk of borrowing and erodes the credibility of real estate valuation. Traditional valuation methods are handicapped in estimation of property value arising from huge (big) data that captures the location, socioeconomic, physical,
environmental and demographic characteristics of real estate. ANN has shown to be very keen on tackling the problem. The advantage of ANN over the traditional evaluation tools is that the learning function and nonlinear processing ability of ANN can improve the randomness and uncertainty of existing evaluation methods thereby minimizing the likelihood of information loss. Hence, ANN is one of the major evaluation tools that guarantees accurate property value estimation, even though in the presence of large datasets. The application of ANN has been extended to the evaluation of real estate companies in order to determine their competitive advantage, viability and strength. The application has been extended to the appraisal and monitoring of real estate market. Government agencies depend on accurate property valuation estimation for computation of property taxes.

- **Mass Appraisal**

Conventional evaluation methods cannot be applied in the estimation of prices of large amount of properties. The complexity of the nature of the characteristics underneath the mass houses means that advanced computational methods are needed to create a reasonable evaluation model. Artificial neural networks have been used in this context to estimate values between properties. The big data usually encountered in mass appraisal can be handled by the application of ANN and a combination of other data mining tools such as decision trees, random forest, boosted trees, and others. The application of ANN in the mass appraisal of properties has improved the credibility and efficiency of the evaluation system. Unarguably, this area remains a fertile ground for more research activities as revealed.

- **Risk Assessment and Evaluation**

Apart from the usual valuation of real estate, values also determine the risk associated with properties. The risk of the individual variables (components) is assessed to determine the net risk of the property. Risk evaluation is very crucial since real estate carries a lot of risks and adequate knowledge of the risk is essential for investment and business decisions. The level of risk determines the extent of investment. Traditionally, qualitative methods are used in the risk assessment in this context, which is highly subjective and often inaccurate. ANN can be used in this aspect by leveraging on its self-organization and learning capabilities in order to accurately quantify the risk by excluding the subjective factors in real estate and credit risk of real estate in banks. The adoption of ANN in risk Assessment ensures credibility and is necessary to avoid waste of scarce resources and foreclosures. In the case of credit risk, ANN helps to determine the state of credit rating and to handle complex relationships that are inbuilt in credit management in banks. ANN and its combination with fuzzy comprehensive evaluation have been used to determine the weight of every risk factor and relate them to the overall risk of the real estate. Risk assessment of contracting in real estate has also been reported. This was done by the use of ANN and SVM.
The real estate risk assessment remains an integral part of the general risk assessment of which the ANN has proven to be an efficient tool for estimating the former.

- **Price Prediction And Forecasting Of Real Estate**

Estate valuation is not just limited to valuing or appraising a property, it also predicts the price using historical data on various variables affecting the price or value of the property. The act of prediction involves the extraction of useful information from a given raw data in other to forecast or predict the unknown. Comparisons are subsequently made with the current selling prices. Predictions are often cumbersome using the traditional evaluation methods but the advent of computers and algorithms such as ANN has made it easy to explore the correlations between the variables that affect the price of real estate and create patterns within multi-dimensional variables. Data mining tools are widely used in this aspect. Often, the evaluation models used in real estate appraisal are used in price prediction. As stated earlier, the randomness of the variables reduces the predictive accuracy of evaluation models which are mainly modeled as deterministic. The prediction has to be accurate and a pointer to the success or failure of real estate. High positive correlation coefficient is desirable to indicate that the prediction is accurate. Different residential real property variables can be used. Examples are location, prior prices, building conditions, housing area and number of rooms, consumer spending, value of money, demand and supply, government policies, security, disposable income and rediscount rate. ANN has been combined with fuzzy method, ordinary least squares regression, logistic regression, Markov chain, support vector regression (SVR), principal components analysis, particle swarm optimization algorithm, decision tree, gradient boosting machines and genetic algorithm in prediction of price of properties. The unique strength of ANN is that it captures the uncertainties of the variables that constitute the predicting model at a reduced computational time and increased accuracy. Unique cycles or seasonal variations in the price of real estate can be precipitated using the ANN. ANN model help to identify the unique factors that are most likely to affect the price of properties. This helps to facilitate quick decisions by investors, buyers, government and financial institutions. Comparison of ANN over other methods showed that despite the predictive capability of ANN, some other models performed better than it. SVM performed better than ANN for small training data. A combination of SVR and rough sets yielded better prediction than ANN. Nero-Fuzzy Inference System, performed better than ANN. On the other, it has been reported that ANN performed better than linear regression and random forest algorithm in real estate price prediction. An emerging trend in this area is the use of virtual characteristics in the prediction of price of properties using ANN. Datasets used here are the photos of the interior and exterior of the property. This approach has also shown to be better in real estate evaluation than other notable offline and online methods. The use of satellite images, street views and the accessibility in real estate price predictions have been reported.
2.4 Application of ANN in Construction Management

2.4.1 Construction Cost

Estimation of the cost of a construction project is an important task in the management of construction projects. The quality of construction management depends on accurate estimation of the construction cost. Construction costs are very noisy and the noise is the result of many unpredictable factors. A regularization neural network is formulated and neural network architecture is presented for estimation of the cost of construction projects. The model is applied to estimate the cost of reinforced-concrete unit of structure as an example. The new computational model is based on a solid mathematical foundation making the cost estimation consistently more reliable and predictable. Further, the result of estimation from the regularization neural network depends only on the training examples. It does not depend on the architecture of the neural network, the learning parameters, and the number of iterations required for training the system.

2.4.2 Safe Work Behavior

The safety of construction workers and employees is a major social responsibility and it is a challenging task to ensure zero incidents at construction sites all over the world. Annually, nearly 60,000 fatal accidents happen at construction sites worldwide. Hence model has been developed employing an ANNs to predict the safe work behavior of employees using 10 safety climate constructs determined through literature review. The model utilizes safety climate constructs (determinants) as inputs and safe work behavior as an output. Two hundred twenty-two responses from several construction projects across India were collected through a questionnaire survey. A three-layer feed-forward back-propagation neural network was appropriate in building this model which has been trained, validated, and tested with sufficient data sets. The model predicts the safe work behavior of employees reasonably well. In addition, a sensitivity analysis was carried out to study the impact of each construct on the safe work behavior of employees. As a result, safety climate constructs like supervisory environment, work pressure, employees’ involvement, personal appreciation of risk, and supportive. Environment was significantly associated with the safe work behavior of employees. This model has great potential in aiding contractors and clients in promoting safe work behavior and the efficient management of the safety of employees in construction projects.

2.4.3 Safety

It is necessary to adopt safe working methods on site to reduce the no of accidents. To ensure safe work environment a smart ANNs-based slip-trip classification method, which integrates a smart sensor and an ANN. It was trained to identify the slip and trip events that occur while a worker walks in a workplace. It encourages preventive and collective actions to reduce
construction accidents by identifying the type of near miss, i.e., slip or trip, and the exact time that it occurs. The variation in the energy released by a worker is measured using a triaxial accelerometer embedded in a smart phone. This study is of value to researchers because the method measures a near miss quantitatively using acceleration. It is also of relevance to practitioners because it provides a computerized tool that records each and every moment of a near-miss event. A test was performed by collecting the three-axis acceleration streams generated by workers wearing a smart phone running the classifier as they walked around a simulated construction jobsite. It identified the type of near miss and the exact time of its occurrence. The test case verified the usability and validity of the computational methods.

2.5 Application of ANN in Construction

This article provides an overview of the application of ANN in the construction industry, including energy efficiency and energy consumption, structural analysis, construction materials, smart city and BIM technologies, structural design and optimization, application forecasting, construction engineering and soil mechanics.

2.5.1 Energy efficiency and energy consumption

Neural network technologies are widely used in tasks related to energy efficiency and energy consumption in buildings. First of all, ANN are used in the tasks of forecasting the loads on heating and cooling systems, electricity consumption and analysis of energy consumption. Neural networks are also used to study the thermal insulation properties of materials, the thermal and heat insulation property of building wall. Neural networks are used in tasks heating, ventilation and air conditioning for example for analysis and optimization of control systems of HVAC and for rapidly predict non-uniform indoor pollutant concentration.

2.5.2 Construction Materials

The use of neural networks in building materials to predict the characteristics of building materials, for example, for studying the effect of two types of materials including micro-silica and also calcium in silicate minerals on the compressive strength of mortars for estimating the volumetric water content in different times and positions during the water imbibition inside the porous building materials for predicting performance of lightweight concrete with granulated expanded glass and ash aggregate for designing the composition of cement stabilized rammed earth for studying on adiabatic temperature rise reflecting hydration degree of concrete for predicting the compressive strength of cement-based materials exposed to sulfate attack for prediction of chloride diffusion in cement mortar. Neural networks are used to predict corrosion. Neural networks are also used to study the fire resistance of building materials.

2.5.3 Safety in construction

In addition to studying the fire resistance of building materials, neural networks are also used in other areas of construction safety, such as safety assessment of megaprojects evacuation...
tasks forecasting the safety of building structures prevention various injuries at construction sites and others.

2.5.4 Construction cost estimate

One of the important aspects of construction is prediction of the behavior of the investment project model in construction at each stage of the life cycle. Neural networks have found application in the tasks of performance prediction of construction projects at different stages of projects and cost estimation in construction, predicting the quantity of materials for constructions and production rate.

2.5.5 Smart city

In smart house systems, neural networks are used for security systems, equipment control systems (heating, water supply, and ventilation, light), reminder systems, and monitoring and evaluation systems for objects, identification systems. Neural networks are also used as part of urban planning and for the analysis of aerial photographs in the interests of urban economy and construction in waste management systems.

2.5.6 BIM

The creation of information models spans all the time from the ideological concept of the project to the completion of operation and demolition of the building. Neural networks are also used in tasks related to BIM, such as creating model for spatial planning of site and building multi LOD modeling approach for the early phases of building design macro BIM cost analyzes researching on individual thermal comfort and others.

2.6 Artificial Neural Networks Limitations

The major challenges facing ANN are:

- The amount of time taken to train networks
- It can require a considerable amount of compute power for more complex tasks.
- The top challenge is that neural networks are black boxes, in which the user feeds in data and receives answers.
- It needed high processing time for big neural networks

3 Conclusion

The factors causing cost overrun in projects from highest to lowest are material cost increase, contract type, scope changes by the owner, fluctuation in the currency, delay in design approval, project size, delay in arrival of materials, lack of enough information about the project at tender stage, experience of the contractor in similar projects, suitability of the schedule to the project, unskilled labors, decreased productivity of equipment and labors, unknown geological conditions, poor site management, ignorance and lack of knowledge, shortage of labors, ineffective planning and scheduling, bad client’s characteristics, undetailed drawings, inappropriate method of construction etc. The project on the topic overrun in real
estate project is finished up to 40% as per methodology. Various aspects of factors regarding overrun in real estate project were studied and listed, the questionnaires to conduct surveys were prepared. The main factors causing overruns in real estate project are listed as contractor’s factors, client factors and owner related factors.

References