Using Artificial Neural Networks to Model Cost Overruns in Real Estate Projects

Kareem Mohammad Mahmoud Mostafa, Dr. Ayman Hamdy Nassar

Civil Engineering Program, Faculty of Engineering and Material Sciences, The German University In Cairo, Cairo, Egypt

Abstract: - **Cost overrun is one of the most important problems and risks that encounter Real Estate projects success, since it reduces the contractor's profit and sometimes lead to enormous losses, and leaving the project in great troubles. Construction cost is one of the peak criteria of success for a real estate project throughout its lifecycle and is of high concern to those who are involved in the real estate industry. All real estate projects, regardless of their size, complexity are saddled by targets and uncertainties. Mostly in developing countries real estate projects are characterized by overruns in cost. Cost overruns occur in every real estate project while the magnitude varies significantly from project to project. This leads to severe need of addressing the acute issue of cost overrun.**

Throughout this research we tried to gather and analyze the main factors which cause cost overruns in real estate projects. According to past researches 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, fluctuation in the currency that payments will be made by, 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, design changes, location of the project with respect to vendors, time needed for decisions to be made, what was known about the project at tender stage, the client's characteristics, unknown geological conditions, ignorance and lack of knowledge of the parties, suitability of the project schedule to the project, conflict among participants in the project, design complexity, scope change by the owner, 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, the inspection and testing procedures , 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 experience of the contractor in similar projects, site access ease by the contractor, the effectiveness of the planning and scheduling , the availability of equipment, delay in arrival of material, 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 category of the contractor and finally the method of procurement.

Two Questionnaires were created to determine the probability and impact of each factor and hence rank the factors according to Relative importance index besides determining the percentage overrun each factor could cause according to practitioners in the field of real estate. Not to mention, data of real projects which encountered overruns were included. Using Artificial Neural networks which is a branch of Artificial Intelligence, a model was developed including the 43 factors which can predict the percentage overrun for real estate project. The model was later tested by one of the projects and the variance between the actual overrun and the predicted overrun was calculated.

I. 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.

II. LITERATURE REVIEW

Construction industry is considered as one of the most important industries for the economic development of any country. However, this industry has been facing major issues regarding failure to complete projects within the estimated Budget.

Cost overrun can be defined as when the project objectives are not achieved within estimated budget **Avots(1983)**. The causes of cost overrun in construction projects are many,

some of them are not only hard to predict but also difficult to manage **Morris(1991).**

It is not something odd to see construction projects failing to achieve their target of creating facilities within the specified cost. Very few projects get completed within the estimated budget since construction projects are exposed to uncertain environments. The trend of overrun is more severe in developing countries where the cost overruns sometimes exceed 100% of the estimated cost of the project. In the past years, many construction projects were well known for their cost overruns **(Pickrell 1990)**;. Many researchers have conducted several attempts to realize, study and interpret the main reasons causing cost overruns in residential & different construction projects.

A study conducted on 8000 projects in 1997 , **(Frame, 1997)** found that only 16% of the projects were completed within the specified Budget, on time, and meeting the quality requirements). A research conducted in Nigeria, **(Omoregie & Radford, 2006)** reported that the minimum average percentage of cost escalation was 14%.While according to another research carried out in Portugal, construction projects faces a minimum of 12% of cost overrun **(Moura, Teixeira, & Pires, 2007).** Other researches reported that there are more building projects that had cost overruns than those which were completed within budget **(Chan and Kumaraswamy 1996).**

As a start, in turkey a study revelaed that the main causes of cost overruns were the increase of materials prices, fast growth of inflation, and delays caused by changes in design specification**(Arditi, et al.,1985)** Secondly, Okpala and Aniekwu searched for the main causes of cost overruns in Nigeria, and found out that the main causes were shortage of materials, problems regarding finance and payment for completed works, poor contract management, price fluctuations and fraudulent practices **(Okpala and Aniekwu,1988)**

Moreover, the Researcher Nassar et al. conducted a survey to estimate cost overrun and the analysis results showed that the average of cost overrun was 4% above the bid price**.(Nassar et al.,2005)**Furthermore, Olawale and Sun indicated that although the causes of overrun in projects may be close somehow but the results may vary dramatically from one county to another**. (Olawale and Sun,2010)**

Besides, it was reported that according to **(Art Chaovalitwongse et al.,2011),** the main causes of cost overrun were the increase in material prices due to inflation, inaccurate material estimating, and increased project complexity, these factors were mentioned and studied based on their observed importance and frequencies of occurrence.

On the other hand the two researchers **Jahren and Ashe (1990)** claimed that cost overrun are affected by factors such as : project size, difference between the selected bid and the government estimate, type of construction, level of competition. Surprisingly they also found that factors such as the quality of the contract document, the nature of interpersonal relations on the project and the policies of the contractor, could also have a great impact on cost overrun. Not to mention that they also found that cost overrun rate was more likely to occur on larger projects than smaller ones.

Moreover, **Kaming, Olomolaiye, Holt and Harris (1997)** conducted a questionnaire survey among project managers working on high-rise construction projects in two Indonesian cities: Jakarta and Yogyakarta, to determine the factors that impact construction cost and lead to overrun. Their results showed that the most important factors are material cost increases due to inflation, inaccurate material estimating and degree of project complexity. Similarly, a study was done in Kuwait by researchers Koushki, Al-Rashid and Kartam (2005) to investigate the main causes of cost overrun in construction project, they performed a person-interview survey among 450 private residential projects' owners and developers. They stated that the occurrence of cost overrun usually increased with the increase in the total cost of a project. Moreover, their survey results had shown that owners who spent more time on the pre-planning phase had experienced less cost overrun during the execution phase of their projects. Also, they found that the main causes of cost overrun in construction projects are: 1- Contractor-Related Problems, 2- Material-Related Problems, 3- Owners' Financial Constraints.

Additionally, **Le-Hoai, Lee and Lee (2008)** had also performed an investigation among previous studies to identify the main causes of cost overrun, they identified 21 of cost overrun causes and categorized them into 6 different groups: 1- Owner Related Factors, 2- Contractor Related Factors, 3- Consultant Related Factors, 4- Project Related Factors, 5- Material and Labor Related Factors, 6- External Factors.

Not to mention that**, Park & Papadopoulou (2012),** checked many research papers to determine the causes of cost overrun. As per their review they found 27 significant causes of cost overrun, then they summarized and analyzed the most common ten causes which are: 1- Inaccurate estimates, 2- Shortage of material, 3- Price fluctuations, 4- Inappropriate procurement route, 5- Delayed payment of completed works, 6- Unforeseen site conditions, 7- Poor site management by contractor and poor planning, 8- Change orders by client, 9- Lack of communication between parties, 10- Inadequate duration of contract period.

Morris (1990) claimed that the factors that impact cost overrun in construction projects are inefficient project planning, delay in construction progress, change in project's scope, planning and implementation, supply of raw materials and equipment by contractors, foreign trade not ready, delay in decision making by government and failure of specific coordinating bodies.

Bubshait and Al-Juwairah (2002) reviewed 42 factors that influence construction cost in Saudi Arabia. They used the

severity index to rank the factors. Response from owners, consultants, contractors and combination of were collected to identifying their degree of importance the study concluded that poor financial control on site, inexperience in managing contracts incorrect planning, materials cost, are the main factors that can lead to cost overrun.

Le-Hoai et al. (2008) reached a conclusion where he stated that the 5 main significant factors influencing cost overruns in Vietnam were financial difficulties of owner, poor site management and supervision, poor project management assistance, financial difficulties of contractors, design changes.

Enshassi et al. (2009) reviewed many researches and stated that the major factors causing cost overruns includes fluctuations in the cost of building materials, increase of materials prices, delay in construction, delay in supply of raw materials and equipment by contractors, unsettlement of the local currency in relation to dollar value, resources constraint, lack of cost planning/monitoring during pre-and post- contract stages, funds not ready, design changes, improvements to standard drawing during construction stages and inaccurate quality take-off.

Serdar Durdyev (2012) identified 40 factors in Turkey that causes cost overrun in construction of residential projects. The major factors are improper planning, inaccurate project cost estimation, high cost of needed resources, lack of skilled workforce, price of construction materials and high land prices. Finally, Chan **(2002)** identified other reasons causing cost overruns and classified them into 3 aspects such as owner's control, consultant's control, and beyond the control.

III. THEORETICAL BACKGROUND

1- 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:

1- Resource Planning
2- Cost Estimating
3- Cost Budgeting
4- Cost Control

1.1- 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.

1.2- 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.

Table 3-1 shows the resource planning process

Inputs	Tools and Techniques	Outputs
1- Historical information 2- Scope statement 3- Work breakdown structure 4- Organizational policies 5- Resource pool description.	1- Expert judgment 2- Alternatives identification	1- Resource requirement

Table 3-2 shows the cost estimation process

Inputs	Tools and Techniques	Outputs
1- Work breakdown structure 2- Resource requirements. 3- Resource rates 4- Activity duration estimates 5- Historical information 6- Chart of accounts.	1- Expert judgment 2- Alternatives identification	1- Resource requirement

1.3- 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.

Table 3-3 shows the cost budgeting process

Inputs	Tools and Techniques	Outputs
1- Cost estimates 2- Work breakdown structure 3- Project schedule	1- Cost estimating tools and techniques	1- Cost baseline

1.4- 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.

Table 3-4 shows the cost control process

Inputs	Tools and Techniques	Outputs
1- Cost baseline 2- Performance reports 3- Change requests 4- Cost management plan	1- Cost change control system 2- Performance measurement 3- Additional planning 4- Computerized tools	1- Revised cost estimates 2- Budget updates 3- Corrective action 4- Estimate at completion 5- Lessons learned

2- Artificial intelligence

Artificial intelligence (AI), is a computer-controlled system or digital computer that has the ability to perform tasks associated with intelligent beings perform processes and mimic the characteristics and behavior of humans, such as the ability to reason, discover meaning, generalize, or learn from past experience.

Artificial Intelligence consists of the following branches:

1- Natural Language processing

2- Expert Systems

3- Vision Systems

4- Machine learning(ANN)

5- Robotics

2.1- 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.

2.2- Types of Artificial Neural Networks

There are two Artificial Neural Network topologies used:

1- Feed-Forward ANN

2- Feed-back ANN

2.2.1- 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.

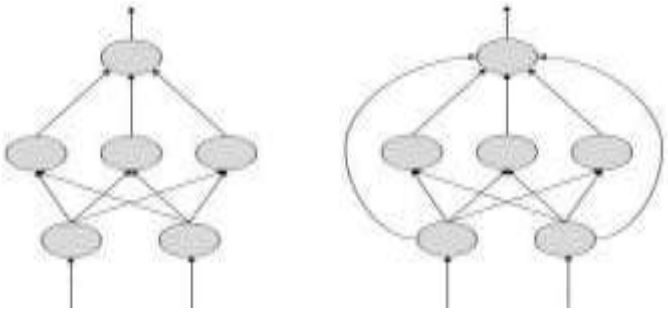

Figure 3-4 shows feed forward ANN

2.3- 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.

$$\mathbf{Y = W_0 + W_1X_1 + W_2X_2 + W_3X_3}$$

$$\mathbf{Y = [\ W \quad b\ ][\ x \ \ 1]^T}$$

2.4- 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.

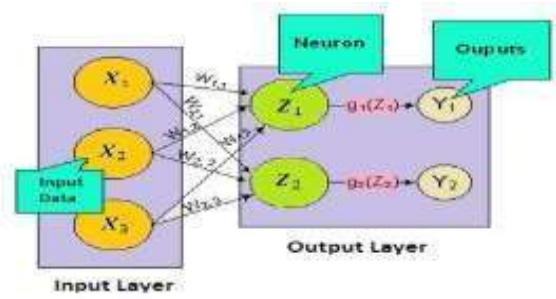

Figure 3-8 shows ANN working

In this single hidden layer feed forward neural network, the network's inputs are directly connected to the output layer nodes, Z1and Z2. The output nodes use activation functions g1 and g2 to yield the outputs Y1 and Y2.

$Z_1 = \sum_{i=1}^{3} W_{1,i}X_i + \mu_1 \quad Z_2 = \sum_{i=1}^{3} W_{2,i}X_i + \mu_2$

$Y_1 = g_1(Z_1) = g_1(\sum_{i=1}^{3} W_{1,i}X_i + \mu_1)$

$Y_2 = g_2(Z_2) = g_2(\sum_{i=1}^{3} W_{2,i}X_i + \mu_2)$

2.5- Activation functions

Activation function: The result of the summation function, that is the weighted sum, is transformed to a desired output by employing a non linear function (Fnl), 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.

$P' = \int \mathrm{NL}\ (y)$

$Y = \sum (\text{weights} * \text{input}) + \text{bias}$

The output value of Y can be anything ranging from –infinity to +infinity. 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).

2.6- 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.

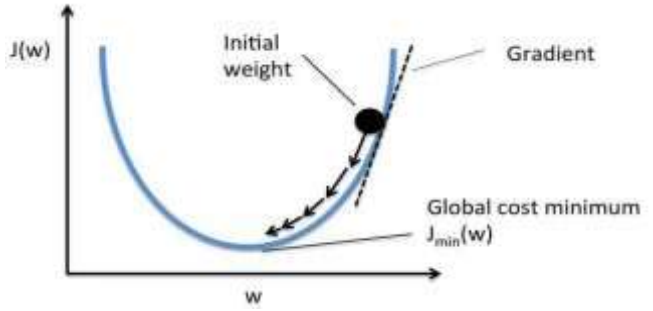

Figure 3-14 shows optimization

2.7- 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.

$$J(w) = P \hat{?}$$

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.

2.8- Artificial Neural networks limitations

All in all, one of the major challenges facing ANN is the amount of time taken to train networks, which can require a considerable amount of compute power for more complex tasks. However, the top challenge is that neural networks are black boxes, in which the user feeds in data and receives answers. They can fine-tune the answers, but they don't have access to the exact decision making process which makes it impossible to track the decision or determine its basis.

IV. METHODOLOGY

The methodology followed in this study includes the identification of factors affecting the cost overrun in Real Estate projects from past researches and literature beside conducting 2 Questionnaire surveys and finally creating a program that estimates the overruns in real estate projects based on the database created.

1- Identification of Various Factors Affecting Cost Overrun

As the initial step previous research papers were reviewed to investigate various factors causing cost overrun for different real estate projects. . Based on previous researches 43 factors were identified to have an impact on cost overrun for projects. The 43 factors were categorized in 6 different groups which are : Financial Management related factors, resource related factors, project management and contract related factors, project related factors, design and consultancy related factors and finally contractor related factors. The factors causing overrun include but are not limited to the financial condition of the owner , the cash flow of the contractor, material cost increase, competition at tender stage, fluctuation in the currency that payments will be made by, the project size, delays in design approval, quantity variations etc.

2- Preparation of Questionnaire Surveys

Two Questionnaires were prepared having the mentioned above 43 factors. The first questionnaire included the 43 factors, the occurrence frequency of each factor ranked from 1 to 5(lowest=1 , highest=5) , the impact of each factor ranked from 1 to 5(lowest=1 , highest=5) and the percentage overrun that could be caused by each factor according to the participants' experience. The following scales were used:

Table 4-1 shows probability scale

Scale of Probability	
1	Rare
2	Unlikely
3	Moderate
4	Likely
5	Very Likely

Table 4-2 shows impact scale

Scale of Impact	
1	Trivial
2	Minor
3	Moderate
4	Major
5	Extreme

Another Questionnaire was prepared containing the 43 factors but assigned for projects, where the participants were asked to specifically mention a project that experienced cost overrun and answer the questions according to that project. The participants were asked to state by yes/no whether the project experienced a problem in each of the 43 factors and the total % overrun experienced by the project was mentioned at the end of the survey.

3-Data Collection

- *Personnel Questionnaire*

A total of 150 questionnaires were distributed to participants of different ages and experience. The questionnaire survey was carried out among three groups of construction practitioners which are architects, engineers and contractors related to real estate industry. Questionnaires were distributed in different firms and 116 were returned as response..94 out of the 116 questionnaires were answered and collected successfully giving 63% response.

- *Projects Questionnaire*

A total of 50 questionnaires were distributed and 34 were successfully collected.

4- Data Analysis

4.1- Ranking of the factors

The response of both questionnaires were considered. The first questionnaire was used to rank the factors from highest to lowest according to the relative importance index method (RII).

Relative Importance index = Probability Index (PI) * Severity Index(SI)

PI = (Number of participants * Probability Scale 1 + Number of Participants * Probability Scale 2 + Number of participants * Probability Scale 3 + Number of participants * Probability Scale 4 + Number of participants * Probability Scale 5)

/ 94

SI = (Number of participants * Severity Scale 1+ Number of participants * Severity Scale 2 + Number of participants * Severity Scale 3 + Number of participants * Severity Scale 4 + Number of participants * Severity Scale 5) / 94

The factors were ranked from highest to lowest according to the relative importance index.

4.2- Creation of Database for Percentage Cost Overrun

The Percentage overrun data were collected from both questionnaires. The first questionnaire represents the opinion of the participant regarding the overrun each factor could cause while the second questionnaire represents the actual percentage overrun occurred as a result of interaction between a multiple of factors. The standard deviation for percentage overrun results collected from the first questionnaire was calculated from each factor and noisy data were rejected. The mean percentage overrun for each factor was calculated.

To merge percentage overrun data from both questionnaires, a weighting factor was assumed where 1 project input was assumed to equal 3 inputs from that of participants' opinion. The mean percentage of each factor was considered as a separate project having an overrun of that mean average and the number of projects was multiplied by 3.

Number of assumed projects = 43 projects

Number of actual projects = 34 projects

Adjusted number of projects = 34 * 3 + 43 = 145

% of data from the first questionnaire = 43/145 *100 = 30%

% of data from the second questionnaire= 34*3 / 145 = 70%

Using the above method the percentage overrun data was successfully merged to create the database used in estimating the percentage overrun for real estate Projects.

5- Developing the model using Artificial Neural Networks

A model was developed using Python language on Pycharm to predict the percentage cost overrun based on the data collected. A network having 43 neurons in the input layer , only one hidden layer was developed having 43 neurons and finally one neuron in the output layer.

6- Testing of the model

The model was tested using one of the projects given in the database to check the variance between the predicted percentage overrun of the model and actual percentage overrun.

V. DATA COLLECTION AND ANALYSIS

5.1- Relative Importance of Factors causing Cost Overrun

This table shows the average probability, impact, relative importance index and rank for each factor

Table 5-1 shows the relative importance index of the factors

No.	Factor Description	Average Probability	Average Impact	Relative Importance Index	Rank
F1	Financial Condition of the Owner	2.85	2.57	7.34	33
F2	Cash flow of Contractor	2.80	3.02	8.45	22
F3	Material Cost Increase	3.27	3.46	11.29	1
F4	Competition at Tender Stage	2.29	1.82	4.16	43
F5	Fluctuation In Currency that payments will be made by	2.85	3.71	10.59	4
F6	Project Size	3.01	3.41	10.28	6
F7	Delay in Design Approval	3.26	3.24	10.56	5
F8	Quantity Variations	3.31	1.95	6.44	40
F9	Detailed Degree of the Drawings	2.65	3.29	8.71	19

F10	Material Estimate Accuracy	2.71	2.51	6.81	37
F11	Quality requirements being hard to obtain	2.87	1.98	5.68	42
F12	Design Changes	2.38	2.55	6.08	41
F13	Location of the Project with respect to Vendors	2.61	2.53	6.60	38
F14	Time needed for Decisions to be taken	2.28	3.37	7.68	30
F15	What is know about the project in Tender Stage	2.95	3.43	10.09	8
F16	Client's Characteristics	2.96	3.11	9.19	18
F17	Unknown Geological Conditions	2.45	3.94	9.63	13
F18	Ignorance and Lack of Knowledge	2.56	3.71	9.52	15
F19	Suitability of the Schedule to the Project	3.32	2.99	9.92	10
F20	Conflict Among Participants	2.37	3.11	7.37	32
F21	Design Complexity	2.78	2.97	8.24	24
F22	Scope Change of the Owner	3.12	3.43	10.68	3
F23	Low/incompatible Advanced Payment	2.84	2.88	8.19	25
F24	Prequalification of the Contractor	2.68	2.82	7.56	31
F25	Work load in the Project	2.88	2.46	7.08	35
F26	Contract Type	3.11	3.61	11.20	2
F27	Dispute Settlement Procedure	2.46	3.36	8.26	23
F28	Inspection and Testing Procedures	2.91	2.47	7.19	34
F29	Site Management	2.69	3.55	9.56	14
F30	Adequacy of the Equipment for the Project	2.37	3.29	7.80	27
F31	Adequacy of the Safety Procedures Followed	2.87	2.76	7.91	26
F32	Similar Projects Experience	2.78	3.59	9.95	9
F33	Site Access Ease	2.81	2.73	7.68	29
F34	Effectiveness of Planning and Scheduling	2.94	3.18	9.34	17
F35	Availability of Equipment	2.26	3.80	8.57	21
F36	Delay in Arrival of Materials	3.03	3.35	10.16	7
F37	Shortage of Labors	2.72	3.44	9.36	16
F38	Labors are unskilled	3.00	3.24	9.73	11
F39	Inappropriate Method of Construction	2.72	3.15	8.58	20
F40	Decreased Productivity of Labor and Equipment	2.99	3.23	9.67	12
F41	Contractor Not Having Cost Control Engineers	3.04	2.55	7.77	28
F42	Low Category of the Contractor	2.74	2.35	6.45	39
F43	Method of Procurement(Open-Selective-Negotiated)	2.47	2.80	6.91	36

5.2- Statistics of the factors causing overrun

The results of the participants questionnaire for the cost overrun each factor could cause was filtered using standard deviation as shown below. The values of overrun having high deviation were eliminated as they were considered noisy data. Finally, an average of the overrun each factor could cause was determined. The 43 factors were categorized into six different groups as shown below were data of each factor was analyzed.

1. Financial Management Factors Statistics

Table 5-3 shows financial management factors

Factor ID	Name of the Factor
F1	Financial Condition of the Owner
F2	Cash flow of Contractor
F5	Fluctuation In Currency that payments will be made by

Table 5-4 shows financial management statistics

	F1	F2	F5
Mean	8.41	11.41	14.63
Mode	10.00	10.00	20.00
Standard Dev.	3.08	5.37	6.98
Lowest %	5.00	4.00	5.00
Highest %	16.00	20.00	30.00
Highest - Mean	7.59	8.59	15.37
Mean - Lowest	3.41	7.41	9.63

2- Resource Related Factors Statistics

Table 5-5 shows resource related factors

Factor ID	Name of the Factor
F3	Material Cost Increase
F30	Adequacy of the Equipment for the Project
F36	Delay in Arrival of Materials
F37	Shortage of Labors
F38	Labors are unskilled
F40	Decreased Productivity of Labor and Equipment

Table 5-6 shows resource related statistics

	F3	F30	F36	F37	F38	F40
Mean	16.08	8.60	9.57	5.24	5.75	4.66
Mode	10.00	10.00	10.00	5.00	5.00	10.00
Standard Dev.	5.95	3.83	4.39	2.48	1.86	3.31
Lowest %	8.00	3.00	2.00	2.00	2.50	1.00
Higest %	30.00	15.00	20.00	10.00	10.00	10.00
Highest - Mean	13.92	6.40	10.43	4.76	4.25	5.34
Mean - Lowest	8.08	5.60	7.57	3.24	3.25	3.66

3- Project Management and Contract Administration Related Factors

Table 5-7 shows project management and contract administration factors

Factor ID	Name of the Factor
F4	Competition at Tender Stage
F10	Material Estimate Accuracy
F19	Suitability of the Schedule to the Project
F20	Conflict Among Participants
F26	Contract Type
F27	Dispute Settlement Procedure
F34	Effectiveness of Planning and Scheduling
F41	Contractor Not Having Cost Control Engineers
F43	Method of Procurement(Open-Selective-Negotiated)

Table 5-8 shows project management and contract administration statistics

	F4	F10	F19	F20	F26	F27	F34	F41	F43
Mean	4.25	3.61	7.72	8.50	8.83	14.12	4.52	5.43	3.29
Mode	2.00	3.00	5.00	10.00	5.00	10.00	5.00	1.00	2.00
Standard Dev.	2.97	2.74	4.50	5.66	4.88	3.75	3.01	3.66	1.39
Lowest %	0.50	0.00	1.50	1.00	1.00	8.00	0.00	1.00	2.00
Higest %	10.00	10.00	20.00	20.00	20.00	20.00	12.00	12.00	6.00
Highest - Mean	5.75	6.39	12.28	11.50	11.17	5.88	7.48	6.57	2.71
Mean - Lowest	3.75	3.61	6.22	7.50	7.83	6.12	4.52	4.43	1.29

4- Project Related Factors Statistics

Table 5-9 shows project related factors

Factor ID	Name of the Factor
F6	Project Size
F8	Quantity Variations
F11	Quality requirements being hard to obtain
F13	Location of the Project with respect to Vendors
F15	What is know about the project in Tender Stage
F16	Client's Characteristics
F17	Unknown Geological Conditions
F22	Scope Change of the Owner
F25	Workload in the Project
F33	Site Access Ease

Table 5-10 shows project related factors statistics

	F6	F8	F11	F13	F15	F16	F17	F22	F25	F33
Mean	9.94	6.84	1.69	3.83	11.22	9.26	11.13	8.43	3.38	4.44
Mode	10.00	10.00	5.00	1.00	15.00	10.00	10.00	15.00	1.00	5.00
Standard Dev.	6.24	5.08	1.77	2.85	5.94	5.21	4.94	6.24	4.07	2.57
Lowest %	0.00	0.00	0.00	0.00	2.50	3.00	2.50	1.00	0.00	2.00
Higest %	20.00	20.00	5.00	10.00	20.00	20.00	20.00	20.00	15.00	11.00
Highest - Mean	10.06	13.16	3.31	6.18	8.78	10.74	8.88	11.57	11.63	6.56
Mean - Lowest	9.94	6.84	1.69	3.83	8.72	6.26	8.63	7.43	3.38	2.44

5- Design and Consultancy Factors Statistics

able 5-11 shows group D factors

Factor ID	Name of the Factor
F7	Delay in Design Approval
F9	Detailed Degree of the Drawings
F12	Design Changes
F21	Design Complexity
F28	Inspection and Testing Procedures

Table 5-12 shows group D statistics

	F7	F9	F12	F21	F28
Mean	9.28	3.24	7.54	5.24	2.27
Mode	5.00	5.00	3.00	10.00	5.00
Standard Dev.	4.96	2.72	5.70	5.46	1.83
Lowest %	2.00	0.00	1.00	0.00	0.00

Higest %	20.00	10.00	20.00	20.00	5.00
Highest - Mean	10.72	6.76	12.46	14.76	2.73
Mean - Lowest	7.28	3.24	6.54	5.24	2.27

6- Contractor Related Factors Statistics

Table 5-13 shows Contractor Related factors

Factor ID	Name of the Factor
F14	Time needed for Decisions to be taken
F18	Ignorance and Lack of Knowledge
F23	Low/incompatible Advanced Payment
F24	Prequalification of the Contractor
F29	Site Management
F31	Adequacy of the Safety Procedures Followed
F32	Similar Projects Experience
F35	Availability of Equipment
F39	Inappropriate Method of Construction
F42	Low Category of the Contractor

Table 5-14 shows Contractor Related statistics

	F14	F18	F23	F24	F29	F31	F32	F35	F39	F42
Mean	7.94	15.85	3.24	4.46	8.53	3.29	5.83	10.79	4.04	3.57
Mode	5.00	15.00	5.00	5.00	5.00	1.00	5.00	10.00	3.00	5.00
Standard Dev.	3.89	3.79	1.86	2.58	3.14	3.17	5.22	3.98	1.18	1.52
Lowest %	0.50	10.00	0.50	0.50	5.00	0.50	1.00	5.00	2.50	1.00
Higest %	15.00	25.00	6.00	9.00	15.00	10.00	20.00	18.00	6.00	6.00
Highest - Mean	7.06	9.15	2.76	4.54	6.47	6.71	14.17	7.21	1.96	2.43
Mean - Lowest	7.44	5.85	2.74	3.96	3.53	2.79	4.83	5.79	1.54	2.57

5.3- The projects Questionnaire

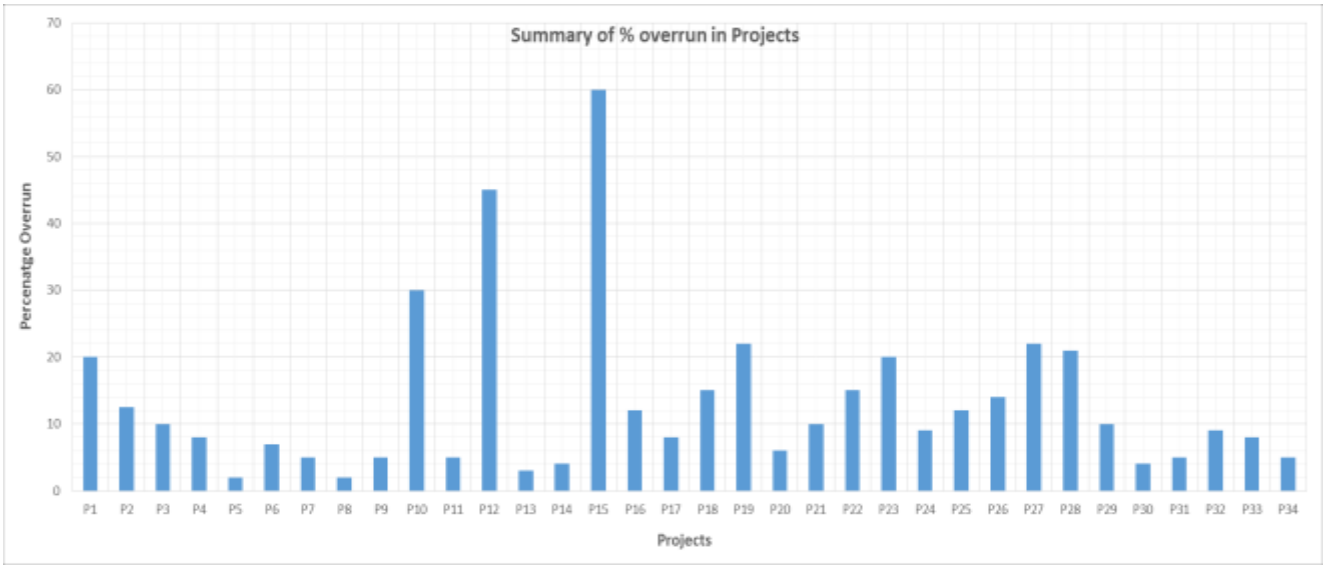

Figure 5.2 shows the results of the projects questionnaire

5.4- The personnel Questionnaire

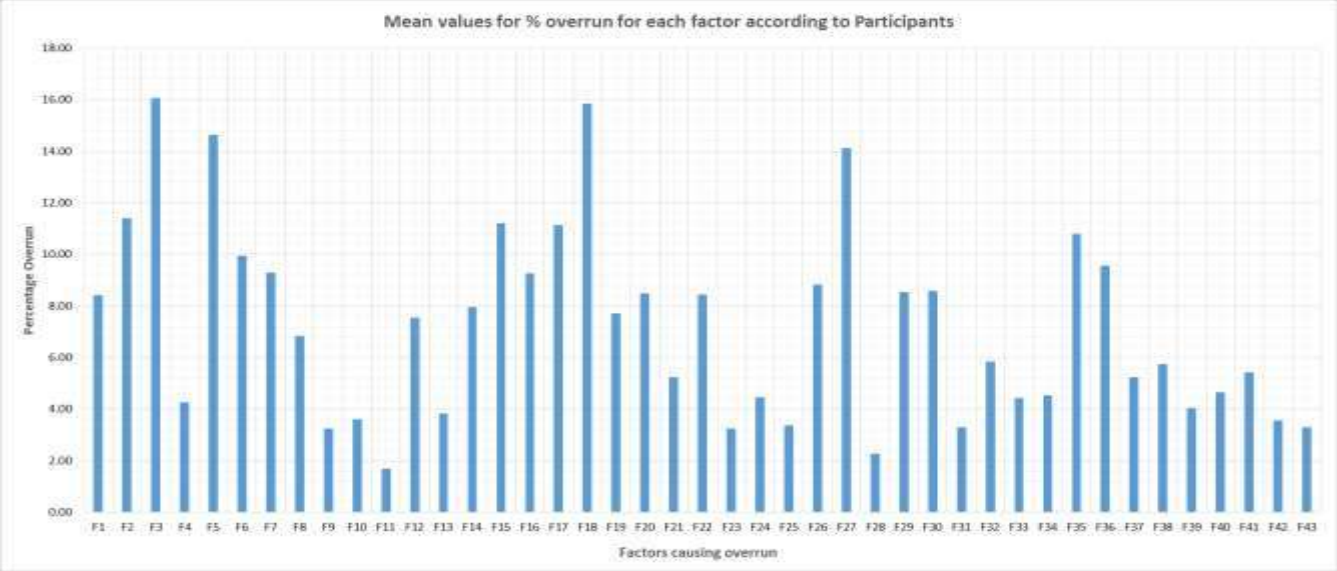

Figure 5.3 shows the average overrun for each factor according to participants

VI. THE SOFTWARE

6.1- Outline of the software

A software was developed to model the cost overruns for real estate projects in Egypt. The model was developed using Artificial Neural Networks, where the input layers included 43 neurons, one hidden layer was used having 43 neurons and one output which is the percentage overrun.

Each input neuron(from I1 to I43) , representing factors from F1 to F43, are connected to all hidden layer neurons(from H1 to H43) and all hidden layer neuron are connected to the output neuron. The network below shows a simplified setting for the network used in developing the software.

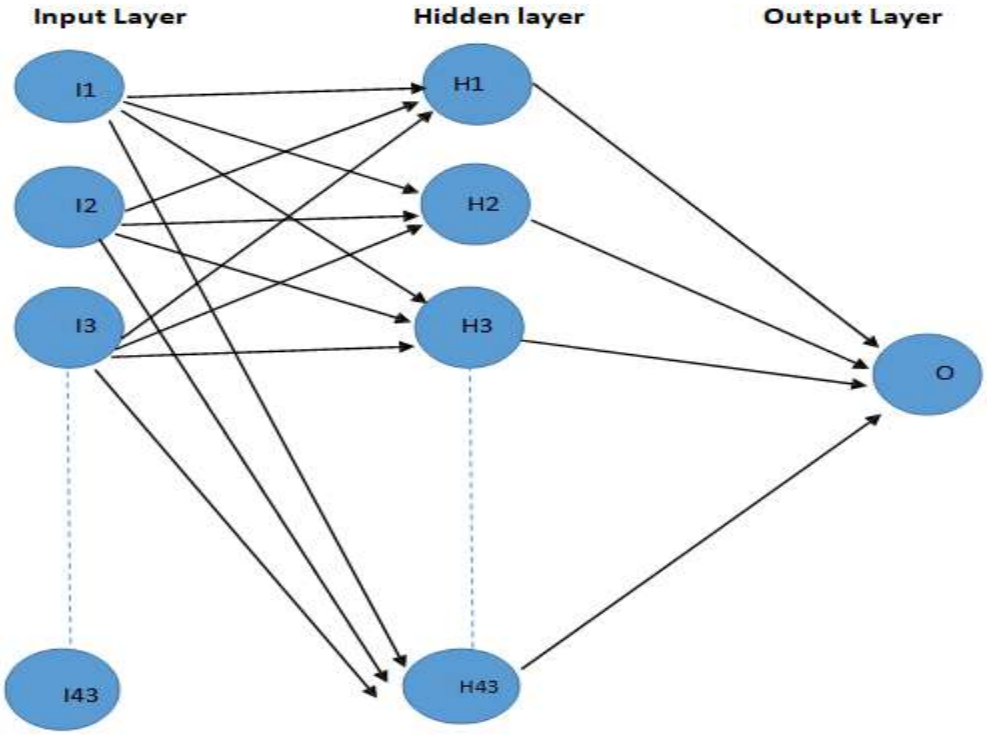

Figure 6-1 shows structure of ANN used

6.2- Program Screenshots

6.2.1- Interface of the program

This appears once the user launches the program, the user then specifies whether he wants to add a new project to the database and in this case he would click on train or clicks on predict to predict the overrun of a project.

Figure 6-3 shows description of interface of program

6.2.2- Sample of the Questions

The Screenshots below shows a sample of the questions to be answered by the user.

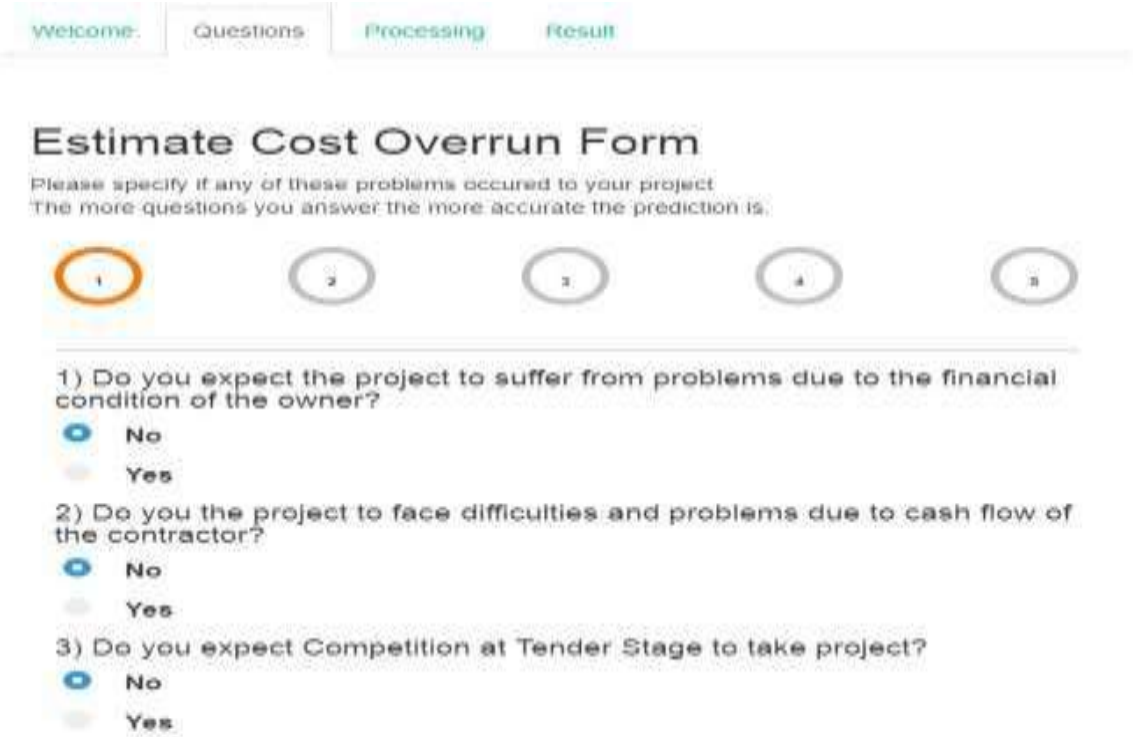

Figure 6-5 shows a the iterations running

Figure 6-4 shows a sample of the program questions

6.2.3- Processing and Iterations

The software runs the network and starts the iterations after the user clicks on **SUBMIT**

6.2.4- The Software Result

The software gives the user the result as shown below and a brief summary of the project including the number of factors and project ID generated.

VII. CONCLUSION AND FUTURE RECOMMENDATION

7.1 Conclusion

This research investigates, analyzes and ranks the main causes of cost overruns in real estate projects gathered from previous researches and predicts the percentage overrun for real estate projects according to a model developed using Artificial Neural networks. The developed model was based on the data gathered from practitioners' opinions (94 Personnel Questionnaires) and actual cases of overrun for Real Estate projects (34 Projects' Questionnaire). The participants specify the frequency of occurrence and impact of each of the 43 factors gathered on the project and specifies the percentage cost overrun each factor could produce. While for the projects' questionnaire, the participant answers the questions for a specific project and states the percentage overrun faced by the project. Based on the response of the personnel questionnaire, the relative importance index for each factor is calculated and the factors are ranked from highest to lowest, the percentage overrun for factors is filtered using standard deviation and the mean cost overrun for each factor is calculated.

According to the relative importance index calculated, 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, unavailability of required equipment, cash flow of the contractor, dispute settlement procedure, design complexity, incompatible advanced payment, adequacy of safety procedures, adequacy of equipment used in the project, absence of cost control engineers, site access ease, time needed for decisions to be made, prequalification of the contractor, conflict among participants, financial condition of the owner, complexity of inspection and testing procedures, workload in the project, method of procurement, material estimate accuracy, location of the project respect to vendors, class of the contractor, quantity variations, design changes, tough quality requirements and competition at tender stage.

Percentage cost overrun data from both questionnaires are used to form a database for the model. An assumption was made in order to merge percentage overrun data from both questionnaires, where the weight of the data from actual projects is assumed to be equal 3 times that of participants' opinions hence the output cost overrun accounts for 70% from the actual projects' overrun and 30% from the practitioners response for each factor.

The data collected was used to develop an Artificial Neural Network model that predicts the cost overrun in real estate projects. The Networks was developed using python language. The network's structure consists of 43 input neurons, 43 neurons in one hidden layer and an output(percentage cost overrun). The model was tested using one of the projects where the variance between the actual percentage overrun and the predicted by the model was found to be 1%.

This research provides an approach for industry practitioners to predict cost overrun percentage for real estate projects. On the other hand, it provides researchers with a methodology to build Artificial Neural Network models for cost overrun percentage prediction.

7.2- Future Recommendation

1. The model could be further upgraded by importing a scaling for each problem instead of yes/no answers. The user can specify the degree of the factor which would help building a model that relies on continuous data instead of discrete data which would enhance the performance achieving higher accuracy.
2. The train option in the model could be used to add new projects to the database of the model hence attaining better accuracy.
3. A formula could be obtained from statistics to merge data of projects with the response from practitioners, this will overcome the error due to my assumption (the weighting used) hence obtaining better results.
4. The input from the user for each factor could be upgraded to include text data, but in this case more hidden layers will be needed.

REFERENCES

[1]. **Oberlender, G.D., and Trost**, S.M. "Predicting Accuracy of Early Cost Estimates Based on Estimate Quality." J. of Constr. Eng. and Manage., 2001.

[2]. **Sharma S and Goyal** P K 2014 Cost overrun factors and project cost risk assessment in construction industry—A state of the art review. Int. Acad. Sci. Eng. Technol.

[3]. Azhar N, Farooqui R U and Ahmed S M 2008 Cost Overrun Factors In Construction Industry of Pakistan. Advancing and Integrating Construction Eduction Research & Practice.

[4]. Okpala D C and Aniekwu A N 1988 Causes of high costs of construction in Nigeria. J. Constr. Eng. Manag.

[5]. P.F. Kaming, P.O. Olomolaiye, G.D. Holt and F.C. Harris.1997 ―Factors influencing construction time and cost overruns on high-rise projects in Indonesia‖, Journal of Construction management and Economics. Vol. 15, No.1, pp.: 83-94.

[6]. Long Le-Hoai, Young Dai Lee, and Jun Yong Lee.2008: ―Delay and cost overruns in Vietnam large construction projects: a comparison with other selected countries‖. KSCE J. civil engineering 12(6): 367-377.

[7]. P.A. Koushki, AL-Rashid, Khalid and Kartam, Nabil.2005, ―Delays and cost increases in the construction of private residential projects in Kuwait‖, Journal of Construction Management and Economics. Vol. 23, No.3, pp. 285-294.

[8]. Ismail Abdul Rahman, Aftab Hameed Memon and Ahmad Tarmizi Abd.Karim. 2013, ―Significant Factors causing cost overruns in Large construction projects in Malaysia‖, Journal of Applied Sciences, Vol 13(2), pp. 286-293.

[9]. Mills, A. (1997) Analysis of the accuracy of building price forecasts. *The Australian Institute of Quantity Surveyors*, 1 (2), 31-38

[10]. Morrison, N. (1984) The accuracy of quantity surveyors' cost estimating.*Construction Management and Economics*, 2, 57 - 75.

[11]. Trost, S. M. and Oberlender, G. D. (2003). "Predicting Accuracy of Early Cost Estimates Using Factor Analysis and Multivariate Analysis." *Journal of Construction Engineering and Management* 129(2): 198 - 204.

[12]. Sørensen, O. (1994) Neural Networks in Control Applications. *PhD Thesis, Department of Control Engineering, Institute of Electronic Systems, Aalborg University, Denmark*

[13]. ZHU Dongli,(2005)“ Optimization model calculation of the project bidding based on capital-time value”

[14]. significant factors affecting time and cost overrun in Indian construction projects. *International Journal of Emerging Technology and Advanced Engineering*, 3(10), pp. 734-740.

[15]. Subramani, T., Sruthi, P.S. & Kavitha, M. 2014. Causes of cost overrun in construction. *IOSR Journal of Engineering (IOSRJEN)*, 4(6), pp. 2278- 8719.

[16]. Ubani, E.C., Okorocha, K.A. & Emeribe, S.C. 2011. Analysis of factors influencing time and cost overrun on construction projects in South Eastern Nigeria. *International Journal of Management Sciences and Business Research*, 2(2), pp. 73-84.

[17]. Baloyi, L. & Bekker, M. 2011. Causes of construction cost and time overruns: The 2010 FIFA World Cup stadia in South Africa. *Acta Structilia*, 18(1), pp. 51-67.

[18]. Cantarelli, C.C., VanWee, B., Molin, E.J. & Flyvbjerg, B. 2012. Different cost performance: Different determinants? The case of cost overruns in Dutch transport infrastructure projects. *Journal of Transport Policy*, 22, pp. 88-95.

[19]. Endut, I.R., Akintoye, A.A. & Kolley, J. 2009. Cost and time overrun of projects in Malaysi.

[20]. Kaliba, C., Muya, M. & Mumba, K. 2009. Cost escalation and schedule delay in road construction projects in Zambia. *International*

[21]. Memon, A.H., Abdul-Rahman, I. & Abdul-Aziz, A. 2011. Preliminary study on causative factors leading to construction cost overrun. *International Journal of Sustainable Construction Engineering and Technology*, 2(1), pp. 57-71.

[22]. Olawale, Y.A. & Sun, M. 2010. Cost and time control of construction projects: Inhibiting factors and mitigating measures in practice. *Construction Management and Economics*

[23]. Subramani, T., Sruthi, P.S. & Kavitha, M. 2014. Causes of cost overrun in construction.