Al-Azhar University Civil Engineering Research Magazine (CERM)
Vol. (42) No. (3) October . 2020

Developing an ANN model to predict the Contractor's actual Cost Overruns for Lump Sum Contract Projects in Egypt

A. M. Ragheb, H.H. Massoud, M.T. Hamisa

*Construction and Engineering Building Department, College of Engineering and Technology. Arab Academy for Science, Technology and Maritime Transport, Alexandria, Egypt.

**Projects Department" Egyptian Natural Gas Company-Gasco", Alexandria, Egypt.

***MSc student, Construction and Engineering Building Department, College of Engineering and Technology. Arab Academy for Science, Technology and Maritime Transport, Alexandria, Egypt.

ملخص عربي

من أهم المشاكل التى تواجه شركات المقاولات عند قيامها بدراسة أى عطاء لمشروع جديد , مشكلة تقدير قيمة الزيادة المتوقعه لتكلفة المشروع الفعلية عن التكلفة المقدرة له خلال مرحلة الدراسة و تقديم العطاء للمالك , و تعد هذه اشكالية كبيرة بخاصة فى نوعية المشروعات القائمة على نظام المقطوعية Lump Sum حيث انه غالبا ما لا يكون مذكورا بالتعاقد اى استحقاق لتعويض المقاول عن اى زيادات فى اسعار الطاقة او خلافه.

فى البحث الحالى تم إستخدام الشبكات العصبية الأصطناعية كأداة يمكن من خلالها تقدير الزيادة المتوقعة فى تكلفة المشروع عن التكلفة المقدرة له خلال مرحلة دراسة وتقديم العرض للمالك , مع ملاحظة عدم توافر المعلومات الكافية عن عناصر التكلفة المختلفة و القرارات المتعاقبة الغير متوقعة الأكثر تأثيرا على إقتصاديات المشروع.

Abstract

Since September 2016, the Egyptian construction market is facing events that were not usually taken into consideration at projects early stages. These events causes insufferable cost overrun that, in some cases, makes the contractor is subjected to loss; his profit. Estimating the construction cost overrun with the traditional methods is impracticable under the new market's circumstances. For this, AI technique is applied to investigate its ability to predict the cost overrun more accurate than the traditional techniques. This research aimed to develop an artificial neural network (ANN) model for predicting the possible cost overrun. The top factors affecting on the construction cost are identified through questionnaire that was distributed to selected professionals in the construction field to be used as inputs for the targeted model. Forty-Nine real life projects were collected to be used in training, validating and testing the neural network. The best network architecture

is determined using N-connection 2.0 software. The results show that ANN is a useful tool that facilitates the prediction of the cost.

Key Words: Construction, Cost Estimation, Cost overrun, Artificial neural network.

1. INTRODUCTION

Attention to the construction project's cost overrun is essential for the construction company/contractor. Normally, no project progress could be achieved during its different phases without adequate cash, which is needed to cover the day by day project expenditures. A cost control system is mandatory in order to help in observing and eliminate the excess of the project actual cost over its estimated one. Most of the construction companies fail due to lack of liquidity which supports their daily activities than because of in-adequate management of any other resource. The project cost elements are likely to be changed during its construction as: the raw material prices, sub-contractors contract, equipment cost, labor wages, transportation cost and preliminaries and general cost. These changed circumstances caused disputes and confusions to all project parties, where every project's stakeholder has a different perspective in determine their responsibility for these changes and as a consequence who can they be compensated [3].

This research aims to develop a tool that can allow the construction company/contractor to predict their lump sum price contract cost overrun at the project tendering stage for the Egyptian private sector projects. This model decreases sources of disputes in future and also, increases the chances of on time project completion. A competent financial management necessitates accurate identification for the project's cost potential allocation and this can be done through a reliable cost prediction [20].

Through the presented research, a careful review for the previously published researches related to the project cost overrun problem is conducted. A number of factors have been concluded as to affect the construction project cost overrun. Then, these concluded factors are being specified more after consulting a number of experts related to different project parties: owner, consultant and contractors who are working into the Egyptian construction market. Taking their opinion into consideration, a structured explementary questionnaire is designed in order to rank the identified factors which affect the construction project cost overrun, according to priorities. The questionnaire was distributed among an addressed number of respondents. Respondents were being selected carefully, as to cover different project fields into the Egyptian construction market as from: construction companies' general managers, main/sub-contractors, consultant and owners' representatives and also, they have different roles in the entities they work at in Egypt.

The questionnaire results have been analyzed, factors are being ranked and then, the first top ten ranked factors are then, being used to develop an ANN model that allow the construction company/contractor to predict the construction project cost overrun at its tendering stage for the Egyptian private sector construction projects whom have lump sum contract price. These factors are the: time and cost estimates, un-stable of the local currency in relation to dollar value, poor information availability during bidding, fuel cost increase, fluctuation of materials prices, changes in material specifications and type, changes in the project scope, in-adequate project preparation, planning and implementation, in-accurate material estimating and delays in decision making.

Data regarding the designated top ten-ranked factors is being extracted from a collected number of forty nine previously constructed projects. These forty nine projects are being constructed in Egypt during the time frame (2007-2019). The average estimated contract value for these collected forty nine projects ranged from under 15 million to over a billion Egyptian pounds.

Artificial Neural Network has ability to mapping and generalizing the non-linear relationship between the project's conditions and variables like tendering method, contractors need for the project, location, site access and project type in developing cost models to predict the final project's cost. Artificial neural network has gained considered a valuable application in construction time and cost management in recent time. There are many reviewed and screened researches that can use the ANN applications in cost prediction, optimization and scheduling, risk assessment, claims and dispute resolution outcomes and decision making. Accordingly, it is expected that an artificial neural network model can be considered as a useful tool that facilities the prediction of the cost overrun for the Egyptian private sector lump-sum contract at their tendering stages where little information is available about the project financing plan.

The N connection professional software version 2.0 which is based on the feed forward back propagation learning algorithm had been used to develop the proposed ANN model. A number of trials made utilizing different network parameters for each trial, in order to arrive at the best network architecture; characterized by the minimum calculated percentage of error.

2. LITERATURE REVIEW

The project construction cost is an essential part to make an offer for any new construction project. Efforts are always being exerted in order to keep the project overall cost during construction within its estimated value at the project tendering stage. Many researchers recorded a cost overrun dilemma in their countries.

Hienze (1992) conducted a research to identify the factors that may cause the project cost to overrun. A number of 466 projects cost data were being collected. Hienze explored that 25% of these projects suffered from the cost overrunning problem with a rate of 10% above budget, while another 10% of these projects suffer from cost overrun by a rate of 18% over their budget. His conclusion was that, the: project type, project location, bidding information and project size are the most significant factors that may cause the construction project cost to overrun [16].

Kaming (1997) investigated the cost information for a number of construction sites located in Jakarta and Yogyakarta in Indonesia. A questionnaire was developed focusing on the high-rise construction projects. The result was that factors are being ranked as: the material cost inflation, inaccurate material prices estimation and the project degree of complexity, the material cost increase, estimation lack of accuracy and lack of experience, to be the most influential factors that cause the project cost to overrun. He recommended to sponsoring more researches with aim of exploring more productive methods to help the construction firms to avoid the impact of the un-foreseen conditions that affect the construction projects performance regarding cost and time [18].

Adnan Enshassi (2010) developed a study for projects constructed in Gaza. 152 factors were being identified as the most affecting on the construction project cost overrun. A questionnaire was designed and distributed it to a number of 80 contractors in order to factors be ranked for priorities. He used the importance index for this mission. His conclusion was that factors as the: increment of materials prices due to continuous border closures, delay in construction, supply of raw materials and equipment by contractors, and the fluctuations in the cost of building materials, project materials monopoly by some suppliers, un-settlement of the local currency in relation to dollar value, low commitment from owner to compensate any bad result that may come from, and the owner policy in biding tender to the lowest price one and design changes, are the most highly ranked factors that affect the construction project cost overrun[11].

Memon (2011) identified 87 factors as the most affecting on the construction project cost overrun. He developed a comprehensive questioner and used the average index method to rank the factors. With a total number of 15 respondents, he concluded that: poor design, design delay, un-realistic contract duration & scope, lack of experience, and late delivery of materials and equipment, in-convenient relationship between management and labor, delay in preparation & drawings approval , in-adequate planning & scheduling, and poor site management & supervision, and mistakes during construction, these are all constitute the most significant factors that cause the project cost to overrun for the Malaysian construction projects[22].

Sweis (2013) conduct a research to identify the factors that constitute a major role in the construction project cost overrun. He concluded that: the design changes, lack of experience of the project type and location, governmental delay, severe weather conditions, these are the main factors that affect the construction project cost overrun. Data related to 57 construction projects was collected from the Ministry of Housing and Public Works, register. This collected data was analyzed, and the conclusion was that, 65% of these projects had not been completed within their estimated budget. The governmental delay, the weather conditions and the design changes constitute 73% from the reasons that cause the project cost to overrun [29].

Ismail (2013) studied the factors that affect the project cost and time along its different phases: planning, design, construction and completion phases. A number of 35 factors were being identified from literature reviews which are the: in-competent subcontractors, schedule delay, mistakes during construction, delay preparation and approval of drawings, poor financial control on site, delay in progress payment by owner, delay payment to supplier/subcontractor, lack of coordination/communication between parties, poor labor productivity, weakness of personal technicality, equipment availability and failure, delays in decisions making, in-adequate monitoring and control, mistakes and error design, in-complete design at time of tender, contractual claims, high labor cost, labor absenteeism & availability, fluctuation of material prices, in-accurate quantity take-off, in-adequate planning and scheduling, lack of experience, in-accurate time and cost estimates, frequent design changes, cash flow and financial difficulties related to contractors, financial difficulties related to owner, slow information flow between parties, shortage of site workers, material shortages, late delivery of materials and equipment, poor project management and project scope changes. Ismail concluded that: project scope changes, late delivery of materials and equipment, and lack of experience are the most important affecting factors on the project cost overrun [17].

EAD Muianga (2014) performed a systematic literature review in the Mozambicans construction projects market. A number of 92 studies previously published had been reviewed, and as a consequence a number of 92 factors were derived. Then, these 92 factors were categorized into eleven groups. These groups are the: governmental relations, contractual issues, organization, management, financing, design and documentation, schedule and control scope changes, environmental and economy, materials and labor/equipment. A questioner was developed and distributed between a numbers of expert project managers in major construction firms on Mozambique. As a result, the following factors: payments delayed, material delivery delay into site, payments procedure, financial problems, cash flow problems, lack of time control and cost inputs, scope/ specifications changes, re-work and project scope increase, are being identified as the most affecting factors on the construction project cost overrun[23].

Shehu (2014) gathered data for a 392 recently completed construction projects in the Malaysian construction market. He found that the cost ratio (CR), which indicates to the relation between the project final cost and the overall contract value, exceeds 1.04% for the chosen projects sample. He developed a questionnaire in order to rank the factors as the most affecting on the construction project cost overrun. Then, data regarding the factors: the project start and completion dates, project location, number of stories, gross floor area, pre-contract budget, contract sum and final account, were being extracted from the collected projects. The collected data are categorized according into groups as: the project sector, the project type, nature of project, procurement method, and tendering. He concluded that, 55 % of the construction projects in Malaysia are exposed to cost overrun [27].

Laila Khodir (2014) conducted a study for the purpose of measuring the probabilities of risk occurs due to political and economic events happened in Egypt after January 2011 which have led to a dramatic change in the risk management map into the Egyptian construction market. She developed a questionnaire and a number of 7 risk keys were being identified based on the risk-index score which are: currency price changes, new tax rates, lack of fuel, and un-secured roads, official changes, workers' strikes and fire risk [19].

M.El-Kholy (2015) used two different methodologies to predict the percentage of construction project cost overrun increase. Through the first one he utilized the regression model approach while through the second one, he utilized the case-based reasoning CBR. A number of 44 factors were being considered as the main causes for the construction project cost overrun. A questionnaire was developed to rank these 44 factors through using the relative importance weight method. Accordingly, a number of 11 factors are being selected as the most influencing factors, which are the: owner financial situation, contractor cash flow, procurement method: open/selective tender, material cost increase, inflation, type of tender competition: comprehensive/aggressive, currency floatation, project size, delay in design & approval, and risk retained by client for contract quantity variations, level of drawings detail, and in-accurate material estimating [9].

Ahmed Kholif (2013) reviewed a number of previous researches to identify the main factors causing time and cost overrun in educational buildings projects in Egypt. Then, a questionnaire survey was conducted based on a number of 53 factors to detect the most significant factors. According to their importance index, a number of 40 factors are being selected which are the: political in-security & in-stability, difficulties in getting work permit from government, practice of assigning contract to lowest bidder, high cost of skilled labor, financial difficulties of contractor, and slow payment for completed works, high insurance & high interest rates, bureaucracy in bidding/tendering procedures , and financial difficulties of owner, in-accurate bill of quantities, high transportation costs, in-accurate cost estimation and mistakes related to soil investigation issues[20].

Neural computing is a known dilated area from the artificial intelligence research different methods. Artificial intelligence aims to imitate the human brain through traditional computing methods one instruction is to be carried out after another while when using neural computing, it can be done far more rapidly than what is possible for the human brain. ANN is able to learn from previous examples, and then gain the experience to infer solutions for future problems which makes it a suitable tool in solving the complicated problems [5].

ANNs offer a powerful means of solving poorly defined problems that eluded solution by classic digital computing techniques. What sets ANN apart from the other applications is the quality of the training data besides the type and structure of the neural network adopted, method of training and the way in which both input and output data are structured and interpreted [5].

Generally, ANNs are used for classification and prediction. In classification, the input to the network is a description of an object to be recognized and the network output will be classification or recognition of that object while in prediction, the neural networks is used as a function approximation tool to replace lengthy procedural algorithms. A process that involves mapping from input to output variables to minimize the error between the network response and the target output until satisfactory results are achieved. After training; the network will be capable of producing approximate solutions for future identical problems [5].

Many NN techniques and applications have been applied for different areas in the construction management filed, such as resource allocation, scheduling problems, time-cost overrun problem, cost forecasting, portfolio selection, insolvency prediction and progress evaluation [14].

For example, Chester (2005) used the MATLAB software on developing an ANN model to forecast the changes in highway construction cost. After determinate the highway cost indices, Chester developed a multi-layer feed-forward neural network structure with back-propagation learning algorithm and a tangent sigmoid transfer function able to forecast the highways construction cost overrun with reasonable accuracy [7].

In order to estimate the percentage of contractor's markup, Shen and Love (1999) used thirty projects to train and test an ANN network that used market conditions, number of competitors, working cash requirements, overhead rates, current work load, labor availability, project type, project size, project location and project complexity as parameters to forecast the suitable markup percentage[12].

Elhag and Boussabaine (1998) proved that ANN model could be also used to determine the lowest tender price of school buildings. They concluded that the ANN suits the cost forecasting problems [8].

For the financial issues, Hazem Masoud (2009) developed a neural network model that predicts construction projects financing cost based on data collected from three types of construction projects which are: pipe line, industrial and building projects. The developed model's architecture was using an input layer that contains fourteen node, one hidden layer that contains five neurons and an output layer that contains one node expresses the predict project's financing cost with the use of tanged sigmoid transfer function [14].

Odeyinka et al. (2002) proposed model to forecast the variation between planned and actual cash out flow uncertainty and unpredicted risks. Neural Network model was developed using Back Propagation technique to estimate the percentage of forecast and actual cash flow. Using trial and error methodology it was found that 14 hidden nodes is the optimum form for the model. Data from 20 construction building used to validate and test the constructed model and the resulted shoed that ANN could be useful in mapping the complex relationships between the construction risks and the variation between actual and predict cash out flow [25].

Neural Network was proved as a successful tool in predicting the projects' cost in the early stage only by determine the inputs' parameter that the most effective parameters influencing the building's cost estimation by Arafa and Alqedra (2011). Taking in consideration the diversion of the types of building, a large numbers buildings and data collect from 71 projects were kept to train the ANN model, Arafa and Alqedra (2011) developed Multi-layered feed forward neural network architecture consisted of input layer, one hidden layer consist of seven neurons and output layer representing the building's cost in the early stage [3].

In order to enhance the estimation of life cycle cost, Alqahtani and Whyte (2013) used the ANN technique. In their study, they used both back-propagation method and spread sheet optimization using Excel solver. Data was collected from 20 construction project to develop the required framework. Using cost significant items, the various factors that affecting on construction cost were identified to be used as an input in the generated framework while the trial and error traditional method to define the number of hidden nodes. The best network used nineteen hidden nodes with tangent sigmoid used as a transfer function [1].

Accordingly, it is expected that an artificial neural network model can be considered as a useful tool that facilities the prediction of the cost overrun percentage for the Egyptian private sector construction projects from whom have lump-sum contract price at the project tendering stages where little information is available about the project financing plan.

3. AIM & RESEARCH OBJECTIVE

The presented research aims to develop an ANN model to predict the private sector construction projects cost overrun at their tendering stage. The proposed model helps to improve the bid accuracy of the new construction projects related to the Egyptian private sector from who has lump sum price contracts. The model benefits to: decrease project duration, save the project parties expanded effort, enhance project quality, and demolish time spent on construction disputes. The model guarantees an optimum bid price which is important for both project owner and contractor. The owner is concerned about the project budget while the contractor is concerned about the construction cost.

The research objectives can be outlined as follow:

- Assign and highlight the most effective cost divers that lead project cost to overrun focusing on the Egyptian private construction projects sector with whom has lump sum contract price.
- Utilize the neural network technique to predict the effect of the increase/ decrease of these identified cost divers on the project overall cost.

This is done through extracting data from a collected 49 previously constructed projects. These projects are related to the Egyptian construction private sector from the type whom has lump sum contract price. 35 collected projects data are used as learning patterns for the proposed ANN model. While, 10 collected projects data are used to test the ANN model and finally 4 projects data are used to validate the proposed ANN model and check its level of accuracy.

3. FACTORS AFEECT THE CONSTRUCTION COST OVERRUN

A comprehensive literature review was conducted in order to identify the most important factors that affect that construction project cost overrun. As a conclusion, a number of 80 factors were being identified. Then, with the advice and consultation a number of the Egyptian private sector construction market experts, these 80 factors are being eliminated and a short list of only 49 factors was derived

Then, a questionnaire was comprehensively structured, distributed among a number of 83 engineers. Those engineers are working into a known Egyptian private sector construction company. They are having different rolls into their companies as, they are: project managers, technical office and planner engineers. They also, have different experience ages in the construction field.

The questionnaire objective was to rank the short listed 49 factors according to their importance to affect the construction cost overrun. The respondents answer was collected and analyzed. Then, the top ten ranked factors are decided to be the only factors used to develop the proposed ANN model. This decision is taken in regard to the availability of the data needed to develop the ANN model after consulting again a number of Egyptian construction experts.
The top ten ranked factors are the: time and cost estimates, un-stable of the local currency in relation to dollar value, poor information availability during bidding, fuel cost increase, fluctuation of materials prices, changes in material specification and type, change in the project scope, in-adequate project preparation, planning and implementation, in-accurate material estimating and delays in the decision making, explained as follow:

3.1 IN-ACCURATE TIME AND COST ESTIMATES

The contractors' engineers especially, ranked this factor as the most significant influence factor on the construction project cost overrun. And as, the presented research is focusing on the Egyptian private sector construction projects that have lump sum contracts, through this contract type the contractor is the only party that burdens the project cost estimation risk.

3.2 UN-STABLE OF LOCAL CURRENCY IN RELATION TO DOLLAR VALUE

Many financial decisions have been taken in the last few years by the Egyptian government in order to reform the country's economy. One of these sudden and harmless decisions was the local currency floatation that occurred in November 2016 which, it was the cause of losing too much value from the purchasing power of the Egyptian pound against the US dollar, illustrated in Figure (1).

As a consequence, this made the future evaluation of the local currency against other currencies rate, very hard to be predictable. This factor has a major effect in the construction cost as it is directly affecting the material prices, foreign employees' salaries and consulting offices fees.

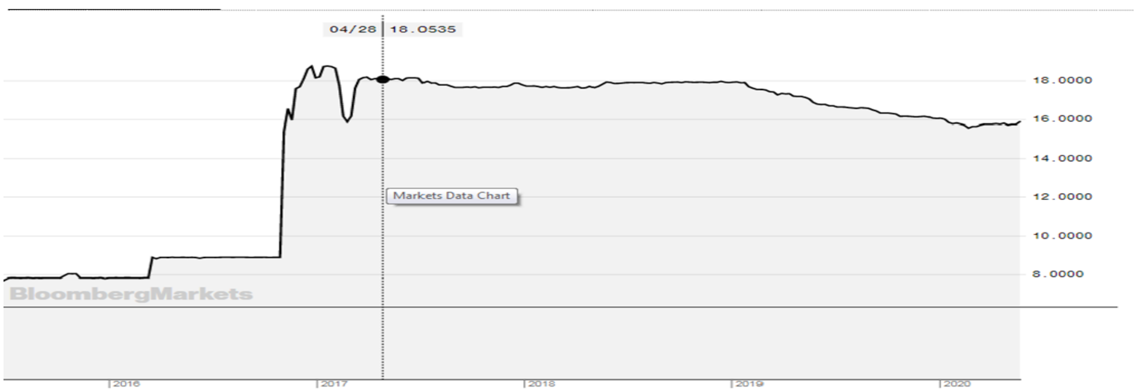

Figure (1): US dollar against the local currency form 2015 until 2019 [30]

3.3 POOR INFORMATION AVAILABILITY DURING BIDDING

Both contractors' and consultants' engineers considered this factor as one of the most significant influencing factors on the construction projects cost overrun. Proper lump sum contract projects pricing requires a high level of detailed drawings and specifications, un-fortunately, this is not the way things done in the Egyptian construction market. Then, any missed valuable data as the: contract general or special conditions, license documents, governmental provision/constrain, soil recommendation reports or any other relative document leads to un-realistic project pricing causes the construction project cost to overrun.

3.4 FUEL COST INCREASE

The Egyptian governmental announced prices for fuel and energy's which is charged on the different customers are extremely increased since 2011 till know. Fuel and energy prices are directly/indirectly affecting the construction projects cost to overrun. For instance: the transportation cost and the accommodation cost, administration and site facilities, others……. etc. Moreover, the construction projects which extremely depend on mechanical equipment during construction are affected more than others.

3.5 FLUCTUATION ON THE MATERIAL PRICES

A producer price index PPI is generated through the CAPMAS on January 2019 valued one hundred ninety-five against a value of only one hundred on January 2016 which indicates the massive inflationary trend occurred in the Egyptian country. Figure (2) shows the fluctuations in the PPI during the year 2018.

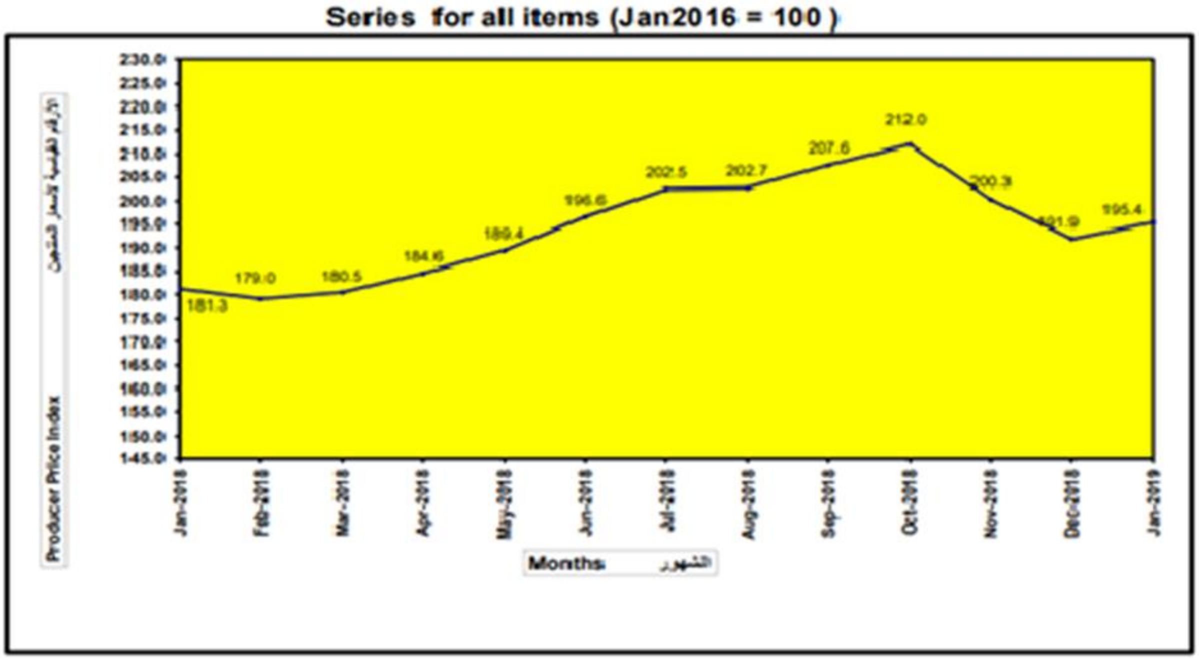

Figure (2): Fluctuation in the PPI during year 2018 [6]

3.6 CHANGES IN MATERIAL SPECIFICATIONS AND TYPE

Material specifications/type change in lump sum contract price projects especially when convergent with no detailed information and in-complete drawings constitutes a disaster for the contractor cost estimation.

In such, circumstances the project cost overrun must occur and that is the reason why especially contractors' engineers ranked this factor as one of the most significant influence factors that affect the construction projects cost overrun.

3.7 CHANGES IN THE SCOPE OF THE PROJECT

Change in a construction project scope may have two different forms. In the first one, an up-normal increase in the performed quantities occurs during construction. The estimated quantities were uncertain and the actual increased a lot than what was mentioned in the BOQ. In the second one, the activity material/nature may change during construction by client or his representative as for instance: aluminum doors to be erected instead of wooden ones. Either of the two forms led to some kind of: re-working, delay the activity duration, decrease labor productivity rate. All these have a significant influence on the project cost overrun.

3.8 INADUCATE PROJECT PREPARATIONS, PALNNING AND IMPLEMENTATION

Proper scheduling and elegance cost control leads any construction to success. These two items can only be guaranteed if, the construction company/contractor prepares a project construction plan with a time cost loaded schedule sticking with the contracted specifications and project related standards. Poor planning and misused quality control process leads to problems like: re-working, material high waste percentages, project time overrun, waste of project resources and liquidity damages. All, these significantly influence the construction project cost overrun.

3.9 IN-ACCURATE MATERIAL ESTIMATING

For a lump sum contract project, the level of accuracy on both time and cost must be as high as possible. Material estimating accuracy shall be taking into consideration at both post/pre-construction stages. Through the pre-construction stage, in-accurate material estimating leads to false pricing while, through the post construction stage, in-accurate material estimating leads to: delay on the project time, re-working and material damaging. All, these significantly influence the construction project cost overrun.

3.10 DELAYS IN DECISIONS MAKING

A lot of what happens in the construction process is the decision making. The construction companies' mangers/contractors have to think about how many choices are made each day through their superintendents, suppliers/subcontractors and employees at their offices. The decision making is made up of two things:

Experience and data. The more construction companies' managers/contractors have from both, the better their decisions. In order to make the most informed decision, the construction companies must understand the goal of their business. In fact, it's a good practice for the company to step back at least once a year to identify the areas which need more focus in its business. Delay in making the right decision on the right time, is one of the ten cost divers chosen from all contractors', clients' and consultants' engineers that significantly influence the construction project cost overrun.

4. RESEARCH METHODOLGY

4.1 QUESTIONNAIRE DESIGN PROCESS

A number of 80 factors were being identified from literature reviews which are believed to be most affect the construction projects cost overrun with taking into consideration, the factors which conformist the period subsequent to the huge sudden economic changes occurred in Egypt at November 2016. These 80 factors are being classified into twelve main categories: contractual issues, design and documentation, environmental and economy, financing, governmental relationships, labor and equipment, management, materials, organization, project nature requirements, schedule and control, scope changes. The purpose was to structure a comprehensive questionnaire that allows the respondents to specify the factors in a simple wording to be clear for getting respondents answers.

The 80 factors list seems to be a very long list based on opinion of a number of engineering experts' consultation and then as a consequence it has been eliminated to only 49 short list factors. This short was used to structure the questionnaire then, it has been distributed between a number of 83 engineers that have different roll in their Egyptian construction companies: project managers, technical office and planners' engineers and also, related to different construction parties: owner, contractor and consultant representatives.

It has to be noted that, the sample size which concerns to the number of the questionnaires answers/number of respondents that have to be collected in order to get a satisfying and trusted results from the questionnaire must be determined then, the infinite population method is being used, illustrated in the following formula [4].

$$N = \frac{(Z_{1-\alpha})^2 \times \sigma^2}{e^2}$$

Equation no. 01

Where N refers to the sample size, Z 1-α is pointing to the desired level of confidence 1-α that determines the critical Z value, σ is the standard deviation and e is the acceptable sampling error.

Considering 95% degree of confidence level for the presented research, then α will equal 0.05, Figure (3) shows that, each shaded zone has the half of α's area equal 0.025; then, the region's area equals 0.5-0.025= 0.475. Using slandered normal distribution's table, Z value is 1.96σ value that was calculated from random samples of replies = 0.7922. Accordingly, through assuming e value is 0.20, the N value will be calculated as following:

$$N = \frac{1.96^2 \ \times .7922^2}{.2^2} \ = \ 44.81$$

Therefore, the minimum sample's size required is 45 samples to attain 95% confidence level.

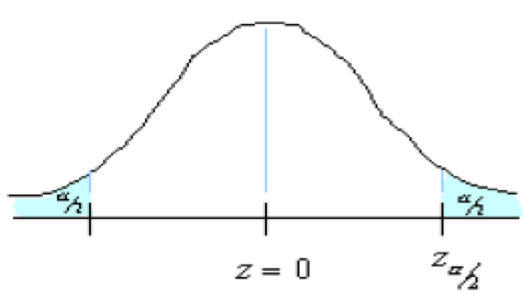

Figure (3): Standard normal distribution [4]

The number of respondents was as being expected: 35 respondents related to the contractors' party which constitute 70% response rate from the total number of the contractor's engineers. And, only 13 respondents related to the consultant's party which constitute 26% percent response rate from the total number of the consultant's engineers. Finally, only 2 respondents related to the client's engineers out of a total number of 11 clients' engineers, illustrated in Figure (4).

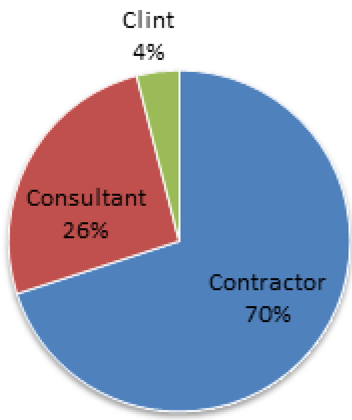

Figure (4): Type of respondents according to organization activity

The questionnaire results are measured using the Likert scaled as presented in Table (1). Each respondent is being asked to determine the impact of each factor occurrence as follow:

- V.L: the factor has very low impact in the construction cost.
- L: the factor has low impact in the construction cost.
- M: the factor has medium impact in the construction cost.
- H: the factor has high impact in the construction cost.
- V.H: the factor has very high impact in the construction cost.

The Relative Importance Index (RII) is used to rank the respondent's answers. The respondent's answers were numerically coded and the data was analyzed using Excel Microsoft office software. The ordinal scale was used as numerical labels to be used to rank the factors as following:

Table (1): Likert scale

Impact	Very low	Low	Medium	High	Very High
Scale	1	2	3	4	5

To determine the relative importance index, the respondents' answers scales were used to infer each item RII based in the following formula:

Relative importance index (RII) = $\frac{\sum W}{AN} = \frac{5\,n1+4\,n2+3n3+2n4+n5}{5N}$ Equation no.02

Where W is the weight given to each factor by the respondent, ranging from 1 to 5, (n1 number of respondents for very high impact, n2 = number of respondents for high impact، n3 = number of respondents for medium impact, n4 = number of respondents for low impact, n5 = number of respondents for Very low impact. A is the highest weight (i.e. 5 in the study) and N is the total number of respondents. The RII equals ranges from 0 to 1.
The result after analyzing the respondents answer is ranking the 49 factors that affect the construction companies as illustrated in Table (2).

Table (2): Factors affecting the construction project cost overrun rank

No.	Factors	Group	Overall	Contractors	Consultants	Clients	Overall ranking
1	Inaccurate Time and Cost estimates	Management	0.82	0.85	0.75	0.90	1
2	Unstable of the local currency in relation to dollar value.	Environment and economy	0.81	0.78	0.86	1.00	2
3	Poor information availability during bidding	Contractual issues	0.78	0.79	0.78	0.70	3
4	Fuel cost increase	Environment and economy	0.78	0.79	0.75	0.80	4
5	Fluctuation of prices of materials	Environment and economy	0.78	0.77	0.77	1.00	5
6	Changes in Material Specification and type	Scope changes	0.78	0.78	0.74	0.90	5
7	Change in the scope of the project	Scope changes	0.77	0.78	0.74	0.90	7
8	Inadequate project preparation, planning and implementation	Management	0.76	0.76	0.74	1.00	8
9	Inaccurate material estimating	Materials	0.76	0.79	0.69	0.70	8
10	Delays in decisions making	Management	0.76	0.74	0.78	0.90	10
11	Change in taxation/new tax rates	Environment	0.75	0.76	0.72	0.80	11
12	Late delivery of materials and equipment	Management	0.75	0.73	0.80	0.90	11

No.	Factors	Group	Overall	Contractors	Consultants	Clients	Overall ranking
13	Poor site management and supervision	Management	0.75	0.76	0.69	1.00	11
14	Lack of cost planning/monitoring during pre-and post-contract stages	Schedule and control	0.75	0.74	0.80	0.70	11
15	Schedule Delay	Schedule and control	0.75	0.75	0.74	0.80	15
16	Lack of experience in project type	Inflexibility (rigidity) of consultant	0.74	0.75	0.71	0.80	16
17	Mistakes and Errors in design	Design and documentation	0.74	0.72	0.80	0.70	17
18	Financial difficulties of contractor	Financing	0.74	0.76	0.66	0.90	17
19	Delay in Material procurement	Materials	0.74	0.74	0.72	0.80	17
20	Delay in project's handing over	Management	0.74	0.73	0.75	0.80	20
21	Poor Contract management	Management	0.74	0.72	0.75	0.90	20
22	Inaccurate quantity takeoff/ Bill of quantities	Design and documentation	0.73	0.71	0.77	0.80	22
23	Lack of communication/coordination between parties	Management	0.73	0.73	0.69	1.00	22
24	Financial difficulties of the owner	Financing	0.72	0.75	0.65	0.80	24
25	Lack of experience in project location	Rigidity of consultant	0.72	0.71	0.72	0.70	25
26	Inadequate planning and scheduling	Schedule and control	0.72	0.71	0.71	0.80	25
27	Shortages of materials	Materials	0.71	0.70	0.72	0.90	27
28	Poor labor productivity	Labor and equipment	0.71	0.70	0.71	0.80	28
29	Shortage of skilled workers	Labor and equipment	0.71	0.69	0.72	0.90	28
30	Effect of contract type on construction cost overrun	Contractual issues	0.70	0.70	0.71	0.70	30
31	Delays in costing variations and additional works	Schedule and control	0.70	0.71	0.68	0.60	31
32	Delay Preparation and approval of drawings	Design and documentation	0.69	0.70	0.69	0.60	32
33	Lack of cost reports during construction stage	Management	0.69	0.69	0.66	0.90	32
34	Contractual claims, such as, extension of time with cost claims	Contractual issues	0.69	0.70	0.66	0.60	34
35	Inflexibility (rigidity) of consultant	Rigidity of consultant	0.68	0.69	0.66	0.70	35
36	Additional work at owner's	Scope changes	0.68	0.69	0.63	0.90	35

	request						
37	Lack of experience of technical consultants	Rigidity) of consultant	0.68	0.70	0.60	0.80	37
38	Accurate prediction of equipment	Labor and equipment	0.67	0.66	0.68	0.80	38
39	Method of procurement	Contractual issues	0.66	0.69	0.60	0.60	39
40	Poor relationship between management and labor	Management	0.66	0.68	0.63	0.60	39
41	Quality standards and specifications	project nature and requirements	0.66	0.69	0.60	0.70	39
42	Difficulties in getting work permit from government	Governmental relation ships	0.66	0.65	0.65	0.80	42
No.	**Factors**	**Group**	**Overall**	**Contractors**	**Consultants**	**Clients**	**Overall ranking**
43	Equipment availability and failure	Labor and equipment	0.65	0.63	0.66	0.80	43
44	project size	project nature and requirements	0.64	0.66	0.54	0.80	44
45	Unavailability of utilities in site (such as, water, electricity, telephone, etc.)	Environment and economy	0.62	0.63	0.58	0.80	45
46	Improper advanced payment amount	Financing	0.60	0.62	0.54	0.60	46
47	Delay in inspection and approval of completed works	Design and documentation	0.60	0.60	0.58	0.60	47
48	High insurance and high interest rates	Environment and economy	0.59	0.58	0.62	0.70	48
49	Re-measurement of provisional works	project nature and requirements	0.58	0.58	0.55	0.60	49

The top ten ranked factors will then be used to develop the proposed ANN model.

4.2 DATA COLLECTION

The number of projects needed to: train, test and validate any neural system is a very important aspect in the development process of the ANN model. It reflects how well the developed model is being trained and determines its prediction accuracy level extent and also, the collected number of previously constructed projects has a direct relation with both the neural model topology and the complexity of the problem that is being modeled.

A number of 49 previously constructed projects have been collected from which being executed by the Egyptian private sector construction companies. All the collected projects are from the type whom have lump sum contract price. These projects are constructed during the time frame (2007-2019). Their contract values ranged from under 15 million to over a billion Egyptian pounds and their duration ranged between (11-133) months. The largest construction firms in the Egyptian private sector which are qualified and registered in the Egyptian Federation for Construction and Building Contractors as first-class companies executed these projects: (Orascom, Redcon, Rowad, Hassan Allam Sons and Kharfi Company).

Data regarding the top ten-ranked factors are being extracted from the 49 collected projects in order to be used in developing the ANN model. These data constitute the proposed ANN model inputs while its targeted output is the construction project cost overrun percentage calculated as follow:

Project cost overrun percentage = [Project overall actual cost - Project tendering estimated cost]/ Project tendering estimated cost Equation no.03

The top ten-ranked factors are being classified into: seven internal factors and three external factors. Data regarding the internal factors are extracted directly from the collected 49 projects. For the external factors needed data, it has been extracted through using the reports that are always issued from the Central Bank of Egypt [26] and also, by the Central Agency for Public Mobilization and Statistics in Egypt CAPMAS [6].

4.3 CREATE AN EXCEL SHEET

The presented research uses the N connection professional software version 2.0 [28] which is based on the feed forward back propagation learning algorithm to develop the proposed ANN model and the data needed to develop the model, has to be introduced to the software on an excel sheet format, Micro Soft Excel 5.0/95 and saved in partition (C) on an IBM computer.

The created excel is illustrated in Figure (5) and follow the next main rules:

- All the data extracted has to be encoded in a format compatible with the neural model. This step means to convert any symbolic/categorical fields found into a numeric format before being introduced to the proposed network architecture.
- Each row represents data regarding to only one project. Then, the designed excel sheet has to contain 49 rows where each row represents one project data. On other hand, each column represents only one piece of data related to only one factor from the top ten-ranked designated factors. Then, the designed excel sheet must has ten columns where each column represent a data related to one of the top ten ranked factors. In other meaning these ten excel sheet columns represents the input data of the proposed ANN model.
- At end of the ten excel sheet columns come column number eleven, which represents the construction project cost overrun calculated percentage. This column contains the proposed ANN model targeted output.

Data regarding 35 projects was used to train the model the proposed ANN model while 10 project data was used to test the model and finally 4 projects are used as the developed model validation data set.

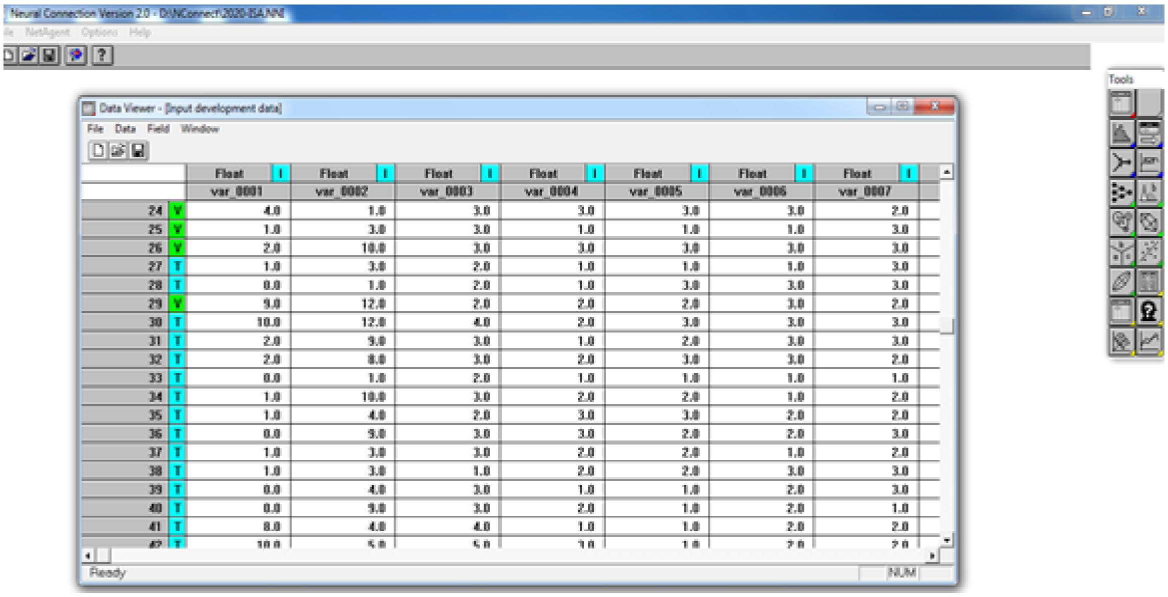

Figure (5): The ANN model input/output data created excel sheet

4.4 ARCHITECTING THE NEURAL NETWORK

Architecting the neural network means to identify its parameters. The parameters refer to the different main elements that constitute the network as: its hidden layers number, the number of neurons within each hidden layer and also, the minimum number of projects/facts that is needed to train the network. These parameters are application dependent, but they also have rules which can be followed in order to nearly identify its reasonable values.

According to the Neural connection 2.0 software user guide [28], there is a subtle relationship between the number of projects needed to train the NN model and its hidden neurons number. The number of projects needed to train the ANN model ranges from (2-10) times the total number of neurons constitute the network structure. Furthermore, almost all the NN previous construction management applications agreed upon using only one hidden layer and also, agreed upon that the number of hidden neurons is to be equal to 0.5 of the network input/output neurons [10].

The proposed model has an input layer that contains 10 neurons each one express one of the top ten-ranked factors identified to affect the construction project cost overrun. It has also, an output layer that contains only one neuron represents the construction project cost overrun calculated percentage. Then, based on the Neural connection guide line, the hidden neurons number and the minimum number of projects data needed to develop the ANN model, can be calculated as follow:

Number of hidden neurons = (Number of input layer neurons + Number of output layer neurons) / 2 = (10+1)/2 = 6 neurons

Minimum number of projects = 2*(Input layer neurons number + Hidden layer neurons number + Output layer neurons number) = 2*(10+6+1) = 34 Projects

Based on that, the collected 49 projects are found to be confidently sufficient to be used to develop the proposed ANN model.

4.5 TRAINING THE PROPOSED ANN MODEL

The top ten ranked factors data extracted from 35 collected projects were being used to train the model. A number of 55 trials had been attempted using different number of hidden layers/hidden neurons in order to arrive at the best network architecture, characterized by minimum percentage of error calculated. These 55 trials had been utilized in groups as follow:

- Group (1): Utilizing one hidden layer & a sigmoid transfer function, illustrated in Table (3).
- Group (2): Utilizing one hidden layer & a Tanah transfer function, illustrated in Table (4).
- Group (3): Utilizing two hidden layers & a sigmoid transfer function in each layer, illustrated in Table (5).
- Group (4): Utilizing two hidden layers & sigmoid in one layer and Tanah transfer function in the other layer, illustrated in Table (6).

Table (3): Utilizing one hidden layer & sigmoid transfer function

Sr	Number of hidden layers	number of hidden layers' neurons		Transfer function	Results		
		First layer	Second layer		RMS	Mean Absolute	mean absolut diff.
1	1	2	0	Sigmoid	0.09737	0.063616	26.96%
2	1	3	0	Sigmoid	0.057342	0.038927	16.49%
3	1	4	0	Sigmoid	0.039869	0.03297	13.97%
4	**1**	**5**	**0**	**Sigmoid**	**0.031125**	**0.026209**	**11.11%**
5	1	6	0	Sigmoid	0.080599	0.069914	29.62%
6	1	7	0	Sigmoid	0.094996	0.074808	31.70%
7	1	8	0	Sigmoid	0.088366	0.069553	29.47%
8	1	9	0	Sigmoid	0.076382	0.063951	27.10%
9	1	10	0	Sigmoid	0.069517	0.064764	27.44%
10	1	11	0	Sigmoid	0.066288	0.056106	23.77%
11	1	12	0	Sigmoid	0.092723	0.074513	31.57%
12	1	13	0	Sigmoid	0.096435	0.090029	38.15%
13	1	14	0	Sigmoid	0.066628	0.056556	23.96%
14	1	15	0	Sigmoid	0.064629	0.058914	24.96%

Table (3), presents a total number of fourteen trials made for group (1) through utilizing one hidden layer & a sigmoid transfer function. It shows that, the RMS and the absolute difference have changed for each trail as a consequence of changing the number of the hidden neurons. It is obvious that, the RMS and the absolute difference changes are non-linear. This means that, the trials must not stop if their results get worst but it must continue. For example, for trial number thirteenth in Table (3), it has a better result than for trial number twelve as the RMS values equals 0.066628 and 0.096435 respectively.

Table (3) concludes that for group (1), the best network architecture consisted of: an input layer contains ten neurons, only one hidden layer contains five neurons and an output layer contains one neuron with a sigmoid transfer function. This architecture results to an RMS value of 0.031125 and an absolute difference value of 11.11%. Table (3), shows also that, the worst trial result is for the first trial which, results to an RMS value of 0.09737 and an absolute difference value of 26.96%.

Table (4): Utilizing one hidden layer & Tanah transfer function

Sr	Number of hidden layers	number of hidden layers' neurons		Transfer function	Results		
		First layer	Second layer		RMS	Mean Absolute	mean absolute diff.
1	1	3	0	Tanh	0.067005	0.054033	22.89%
2	1	4	0	Tanh	0.104072	0.0846	35.85%
3	1	5	0	Tanh	0.184528	0.127183	53.89%
4	1	6	0	Tanh	0.057027	0.050318	21.32%
5	1	7	0	Tanh	0.098491	0.085634	36.29%
6	1	8	0	Tanh	0.099629	0.082163	34.81%
7	1	9	0	Tanh	0.121254	0.105885	44.87%
8	**1**	**10**	**0**	**Tanh**	**0.056204**	**0.048599**	**20.59%**
9	1	11	0	Tanh	0.101367	0.08818	37.36%
10	1	12	0	Tanh	0.105805	0.082964	35.15%
11	1	13	0	Tanh	0.111361	0.090875	38.51%
12	1	14	0	Tanh	0.154528	0.137323	58.19%
13	1	15	0	Tanh	0.112242	0.084105	35.64%

Table (4), presents a total number of thirteen trials made for group (2), utilizing one hidden layer & a Tanah transfer function. It shows the best network architecture consisted of: an input layer contains ten neurons, one hidden layer contains ten neurons and an output layer contains one neuron with a Tanah transfer function. This architecture results to an RMS value of 0.056204 and an absolute difference value of 20.59%. Table (4) shows that, the worst trial made is the last one on the table which, results to an RMS value of 0.184528 and an absolute difference of 53.89%.

By comparing group (1) and group (2) best network architecture considering the RMS and the absolute difference as a base obviously conclude that group (1) best network architecture is still the best.

Table (5): Utilizing two hidden layers & sigmoid transfer function in each layer

Sr	Number of hidden layers	number of hidden layers' neurons		Transfer function	Results		
		First layer	Second layer		RMS	Mean Absolute	mean absolute diff.
1	2	2	2	sigmoid/sigmoid	0.050139	0.039247	16.63%
2	2	3	1	sigmoid/sigmoid	0.218069	0.173688	73.60%
3	2	3	2	sigmoid/sigmoid	0.094639	0.077207	32.71%
4	2	3	3	sigmoid/sigmoid	0.08055	0.076055	32.23%
5	2	4	2	sigmoid/sigmoid	0.086563	0.078327	33.19%
6	2	4	3	sigmoid/sigmoid	0.078796	0.069165	29.31%
7	2	4	4	sigmoid/sigmoid	0.141671	0.097298	41.23%
8	2	5	3	sigmoid/sigmoid	0.063215	0.059747	25.32%
9	2	5	4	sigmoid/sigmoid	0.046527	0.039715	16.83%
10	2	5	5	sigmoid/sigmoid	0.097524	0.08489	35.97%
11	2	6	3	sigmoid/sigmoid	0.076552	0.066223	28.06%
12	**2**	**6**	**4**	**sigmoid/sigmoid**	**0.032086**	**0.030479**	**12.91%**
13	2	6	5	sigmoid/sigmoid	0.049947	0.04601	19.50%
14	2	6	6	sigmoid/sigmoid	0.096576	0.075435	31.96%

Table (5), presents a total number of fourteen network architecture trials made for group (3), utilizing two hidden layers & a sigmoid transfer function for each layer. In general, the results shown in Table (5) look more promised than the previous two groups' results. The table shows that, the best network architecture consisted of: an input layer contains ten neurons, two hidden layers contain six and four neurons respectively and an output layer contains one neuron with a sigmoid transfer function for each layer. This architecture results to an RMS value of 0.032086 and an absolute difference value of 12.91%. The worst trial appears in the table is the second one which, results to an RMS value of 0.0218069 and an absolute difference of 73.60%.

By comparing group (1), group (2) and group (3) best network architecture considering the RMS and the absolute difference as a base, it can be concluded that that group (1) best network architecture is still the best.

Table (6): Utilizing two hidden layers with sigmoid and Tanah transfer functions

Sr	Number of hidden layers	number of hidden layers' neurons		Transfer function	Results		
		First layer	Second layer		RMS	Mean Absolute	mean absolute diff.
1	2	2	2	sigmoid/Tanh	0.053247	0.043372	18.38%
2	2	3	1	sigmoid/Tanh	0.188472	0.134439	56.97%
3	2	3	2	sigmoid/Tanh	0.08483	0.077097	32.67%
4	2	3	3	sigmoid/Tanh	0.103029	0.091286	38.68%
5	2	4	2	sigmoid/Tanh	0.117214	0.097022	41.11%
6	2	4	3	sigmoid/Tanh	0.075501	0.058434	24.76%
7	2	4	4	sigmoid/Tanh	0.056026	0.044455	18.84%
8	2	5	3	sigmoid/Tanh	0.055634	0.052044	22.05%
9	2	5	4	sigmoid/Tanh	0.117194	0.101949	43.20%
10	2	5	5	sigmoid/Tanh	0.074049	0.065843	27.90%
11	2	6	3	sigmoid/Tanh	0.055841	0.049189	20.84%
12	2	6	4	sigmoid/Tanh	0.056395	0.052999	22.46%
13	2	6	5	sigmoid/Tanh	0.058844	0.050367	21.34%
14	2	6	6	sigmoid/Tanh	0.152274	0.115898	49.11%

Table (6), presents total number of fourteen network architecture trial made for group (4), utilizing two hidden layers with a sigmoid transfer function for the first layer and Tanah transfer function for the second layer. The table shows, that the best network architecture consisted of: an input layer contains ten neurons, two hidden layers with each one of them contains two neurons and finally, an output layer contains one neuron with a sigmoid transfer function for the first layer and Tanah transfer function for the other. This architecture results to an RMS value of 0.053247 and an absolute difference value of 18.38%. The worst trial appears in table is the second one, results to an RMS value of 0.188472 and an absolute difference of 56.97%.

Again , by comparing group (1), group (2), group (3) and group (4) best network architecture considering the RMS and the absolute difference as a base, it can be concluded that, the best architecture is corresponding to group (1) and results to an RMS value of 0.031125 and an absolute difference value of 11.11%. Then, the developed ANN model architecture consists of: an input layer contains ten neurons, one hidden layer contains five neurons and an output layer contains one neuron with a sigmoid transfer function.

4.6 TESTING THE DEVELOPED ANN MODEL

The supervised learning method is used to train the neural network model. Through the training process, all the data that regards to the top ten-ranked factors extracted from the collected forty nine projects including the calculated cost overrun percentage, are being introduced to the network input/output neurons through making use of the previously created excel sheet.

After, running the neural software, the model output results are to be compared with their corresponding actual values. The differences are to be calculated and then, used to adjust the network's connection weights. This adjustment happened automatically by the used software without any interference from the applicator. Connections weight continually adjusted until arriving to an acceptable limit. The back-propagation algorithm develops the inputs to the targeted output map by minimizing the root mean square error (RMS), represented by the following equation:

$$\text{RMS}=\sqrt{\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}(\boldsymbol{O_i}-\boldsymbol{P_i})\ /\text{n}}$$

Where: (n) refers to the number of the training samples

(Oi) refers the actual output related to the sample.

(Pi) refers the predicted output.

This value is automatically calculated and training stopped when this value remains un-changed for two chronic iterations.

Then the network is to be tested. Testing the network is the same as training it, except for that during the test a new set of data which the network has not seen before is to be introduced to it. Another difference is that during test, no corrections are to be made when the network is wrong. Hence, it is safe to say that, it is very important to evaluate the performance of the network just after training it and before performing the test. If the test results are in the range of the calculated absolute difference percentage between the model output and its actual corresponding values, then the network is ready to be used. If not, then the network needs more or better data representation through its training phase or it may need to be re-designed. The proposed model used ten projects data to be tested, for each project outcome; the absolute difference is to be calculated as follow:

$$\text{Absolute difference\%} = \frac{(\text{Targeted output} - \text{Calculated output})}{\text{Targeted output}} \times \mathbf{100}$$

This percentage constitutes one of the developed ANN model main characteristics.

4.7 VALIDATING THE DEVELOPED ANN MODEL

The main purpose from validating the developed ANN model is to determine its prediction accuracy level. This step comes after training and testing the developed ANN model. A number of 4 projects were being set aside from the collected 49 projects in order to be used to perform the validation task. The data extracted from these 4 projects is presented in Table (7).

Table (7): The developed model validation projects data

Project Name		Hub building- North Coast	Group of Villas- North Cost	Hotel- Cairo	Residential Buildings- North coast
INPUT DATA					
1	Inaccurate Time and Cost estimates	Medium	Medium	High	Medium
2	Percentage of Local Currency floatation	81.34%	83.86%	37.17%	-1.91%
3	Poor information availability during bidding	Low	High	High	High
4	Percentage of Fuel's Price Increase	79.49%	58.88%	81.49%	33.43%
Project Name		**Hub building- North Coast**	**Group of Villas- North Cost**	**Hotel- Cairo**	**Residential Buildings- North coast**
INPUT DATA					
5	Percentage of Materials prices' Fluctuation	1.99%	2.13%	0.84%	1.36%
6	Changes in Material Specification and type	Low	High	High	Medium
7	Change in the scope of the	High	High	High	Low

	project				
8	Inadequate project preparation, planning and implementation	Medium	Medium	High	Medium
9	Inaccurate material estimating	High	High	High	Medium
10	Delays in decisions making	High	High	High	High
	TARGETED DATA				
1	Construction Over Cost Percentage	30.00%	33.00%	36.00%	15.00%

Through validation, only the input data to be fed to the developed network then, the model allowed running and the targeted output for the four projects resulted. This targeted output is being compared with their corresponding actual values.

This comparison concluded that, for the four projects under investigation, three of them their predicted values are achieved with absolute difference percentage less than 10% while only one project's difference percentage is little more than 10% (11.07%). Hence, it is safe to say that there is a probability of 100% that the results of the proposed model will be located within their acceptable limits. For the all testing sample, the average absolute difference is 7.39%.

Table (8): Actual versus predicted outputs for the validation projects data

Project no.	Actual cost overrun %	Network output	Absolute difference	Comments
1	30.00%	27.81%	7.31%	Correct
2	33.00%	32.61%	1.20%	Correct
4	36.00%	39.99%	11.07%	Correct
5	15.00%	16.50%	9.97%	Correct

Finally, with 92.6% prediction accuracy level, the developed ANN model can be considered as an applicable tool to be used by the construction companies/contractors to predict their lump sum contract price construction projects cost overrun at the tendering stage.

5. COMCLUSIONS & RECOMMENDATION

The work presented shows that using the neural network technique could be effective in determining the percentage of the construction project cost overrun at tendering stage especially for projects that has lump sum contract price with an average error less than 7.4

%. On other hand, using one network to predict the cost overrun percentage for a very wide range project sizes, decreased the prediction accuracy for the developed ANN model. It seems that the prediction accuracy can be enhanced if enough data is available to develop several ANN models for grouped project sizes in a narrow range. Also, it must be noted that, increasing the number of hidden layers to two or more is not justifiable and it leads to slow the speed of the network training and operation and as a result, the accuracy of the network might be changed, however there still requires further study in this field as:

1. In future the study could be extended to include other cost divers from which had been identified through the presented study in the short listed forty-nine factor affect the construction project cost overrun or new cost drivers that are not covered with the literature review made in this research.

2. In future the study could be extended to collect more projects in a narrow project size range and also in a narrow construction duration range. Then, in such case multiple neural networks could be developed for each group of projects, as consequence the accuracy level of the developed models is to be increased.

3. In future this same idea for the NN developing could be applied on other types of construction projects like: infrastructure, steel structure, factories and airports and also, for a different types of project contracts base.

4. In future work it will be benefit to develop sub ANN models to predict the input data of the presented ANN model. As for instance, developing a model that could predict the average increase in the material prices during construction, the fuel cost and other several networks that can be integrated to consolidate the final targeted output of the developed ANN model presented through this research.

References

1- A. Alqahtani, A. Whyte, Artificial neural networks incorporating cost significant Items towards enhancing estimation for life-cycle costing of construction projects, Australasian Journal of Construction Economics and Building, 13 (3) (2013) 51-64.

2- Ahiaga-Dagbui, D.D. and Smith, S.D. (2012) neural networks for modelling the final target cost of water projects. In Smith, S. D. (ed.) Procs 28th Annual ARCOM Conference. Edinburgh, UK 3-5th September 2012 Association of Researchers in Construction Management, pp.307-316.

3- Arafa, M., and Alqedra, M. (2011). “Early stage cost estimation of building construction projects using artificial neural networks.” J. Artif. Intell. 4(1), 63–75.

4- Asal, E. M., 2014. Factors affecting building construction projects cost estimating. A Thesis Submitted To AASTMT, 1-95.

5- Beale R.and Jackson T. (2005) "Neural Computing: An Introduction", IOP Publishing Ltd.,Bristol,England.

6- CAPMAS. 2012. Central Agency for Public Mobilization and Statistics, Egypt. http://www.capmas.gov.eg/

7- Chester G. Wilmot and Bing M. (2005) Neural Network Modeling of High Way Construction Costs Journal of Construction Engineering and Mangment.July 765-771.

8- Elhag, T. and Boussabaine, A. An artificial neural system for cost estimation of construction projects. Proceedings of the 14th Annual ARCOM Conference. Pp.219–226.

9- El-Kholy, A. M., 2015. Predicting Cost Overrun in Construction Projects. International Journal of Construction Engineering and Management, 95-105.

10- El-Sawy, I., Hosny, H., and Abdel Razek, M. (2011). “A neural network model for construction projects site overhead cost estimating in Egypt.” IJCSI, 3(1), 273–283.

11- Enshassi, A., Kumaraswamy, M., & Al-Najjar, J. (2010). Significant factors causing time and cost overruns in construction projects in the Gaza Strip: Contractors’ perspective. International Journal of Construction Management, 10(1), 35–60.

12- H.Li, L.Y.Shen, and P. E.D. Love (1999) "ANN-Based Mark – Up Estimation System with Self- Explanatory Capacities." Journal of Construction Engineering and Management. May/June.185-189.

13- Hatem A. A. (2009), "Developing a Neural Networks Model for Supporting Contractors In Bidding Decision In Egypt", A thesis submitted to Zagazig University in partial fulfillment to the requirement for the Master of Science Degree.

14- Hazem Masoud (2009) " The neural network model for predicting the financing cost for construction projects.” International Journal Project Organization and Management, Vol. 1 No.3, 321-334.

15- Hegazy T. and Moselhi O. (1995). “Elements of Cost Estimation: A Survey in Canada and the United States.” Cost Engineering. 37(5), 27-31.

16- Hinze, J., Selstead, G., and Mahoney, J. P. (1992). “Cost overruns on State of Washington construction contracts.” Transportation Research Record 1351, Transportation Research Board, Washington, DC, 87–93.

17- Ismail, I., Rahman, I.A., Memon, A.H., 2013. Study of factors causing time and cost overrun throughout life cycle of construction project. Proceedings of Malaysian Technical Universities Conference on Engineering & Technology (MUCET), 3- 4 December, Kuantan, Pahang, Malaysia, 3-4.

18- Kaming P, Olomolaiye P, Holt G, Harris F., 1997. Factors influencing construction time and cost overruns on high-rise projects in Indonesia. Construct Manage Econom; 15:83-94.

19- Khodeir, L. M., & Mohamed, A. H. M. (2014). Identifying the latest risk probabilities affecting construction projects in Egypt according to political and economic variables. From January 2011 to January 2013. HBRC Journal.

20- Kholif, W., Hosny, H., & Sanad, A. (2013). Analysis of Time and Cost Overruns in Educational Building Projects in Egypt. International Journal of Engineering and Technical Research, 1(10).

21- Magdum SK, Adamuthe AC (2017) Construction cost prediction using neural networks. ICTACT J Soft Compute 8(1):1549–1556.

22- Memon, A.H., Ismail, A.R., Azis, A.A., 2011. Preliminary study on causative factors leading to construction cost overrun. Int. J. Sustain. Constr. Eng. Technol. 2 (1), 57–71.

23- Muianga, D.A.E., Granja, A.D. & Ruiz, J.A. 2014. Influence factors on cost and time overruns in Mozambican construction projects: Preliminary findings. In: Tamosaitiene, J., Panuwatwanich, K., Mishima, N. & Ko, C.H. Proceedings of the 5th Conference on Engineering, Project and Production Management, 24-28 November 2014, Port Elizabeth, South Africa. Port Elizabeth, South Africa: Nelson Mandela Metropolitan University, Department of Construction Management, pp.10-21.

24- N-Connection V2.0 Professional Software User Guide and Reference Manual (1997), California Scientific Software.

25- Odeyinka, H., Lowe, J., and Kaka, A. (2002). "A construction cost flow risk assessment model." ARCOM 18th Annual Conference, Northumbria, United Kingdom.

26- Official Exchange Rates Historical .Economic Research, Centeral Bank of Egypt, https://www.cbe.org.eg/en/EconomicResearch/Statistics/Pages/OfficialRateshistorical.aspx

27- Shehu, Z., Endut, I.R., Akintoye, A., Holt, G.D., 2014. Cost overrun in the Malaysian construction industry projects: a deeper insight. Int. J. Proj. Manag. 32 (8), 1471–1480.

28- SPSS Inc. and Recognition System Inc. "Neural Connection 2.0", SPSS Inc. and Recognition Systems Inc., 1997.

29- Sweis, G. H., Sweis, R., Rumman, M. A., Hussein, R. A., & Dahiyat, S. E., 2013. Cost overruns in public construction projects-The case of Jordan. Journal of American Science, 9(7), 134–141.

30- USD-EGPX-Rate. Bloomberg Markets, Bloomberg, https://www.bloomberg.com/quote/USDEGP:CUR.