كلية الهندسة بشبرا
FACULTY OF ENGINEERING AT SHOUBRA
ERJ
PRINT ISSN 3009-6049
ONLINE ISSN 3009-6022

ENGINEERING RESEARCH JOURNAL (ERJ)
Volume (53) Issue (1)
January 2024, pp:158-169
https://erjsh.journals.ekb.eg

Developing a Detailed Prediction Model for Construction Site Overheads Using Artificial Neural Network in Egypt

Amr A.Gamal[1], Tarek M. Attia[2], Mai M.Ahmed[3]*

[1]Department of Civil Engineering, Faculty of Engineering at Shoubra Benha University, Egypt.
[2]Housing and Building National Research Center (HBRC), Giza, Egypt
[3]Civil Engineering Department, Faculty of Engineering, Fayoum University, Fayoum 63514, Egypt
* Corresponding Author
E-mail: maimaged.a@gmail.com, Dr.amrgamal@hotmail.com, tarek_attia@yahoo.com

Abstract: The estimation of costs is a crucial aspect of construction planning that should be carried out during the initial phases of a project to establish its budget. The term "project overheads" refers to the indirect expenses associated with a project, such as providing general services at the site or plant, including insurance, site accommodation, and other similar costs. The goal of this research is to investigate the variables that affect the precision of site overheads estimation. In this study, the effectiveness of artificial neural networks (ANNs) was examined in addressing the challenge of accurately estimating project overhead costs during the initial stages of building design. The research involved developing a comprehensive prediction model to estimate the percentage of site overhead costs. The primary objective of this paper is to examine the methods used for estimating project overheads in Egypt. This involves identifying factors that affect site overhead costs and developing a detailed prediction model for construction site overhead using artificial neural network in Egypt.For the purpose of achieving the goal of this research, A questionnaire was conducted among construction companies in Egypt. It was found that the most important factors (according to past studies and questionnaire results) were project's duration, Inflation and Interest rate in the Country and project's size. ANN model has been developed to estimate the overhead percentage depending on its characteristics. It was found that the use of the ANN model is an effective tool for minimizing the amount of work required to estimate the percentage of on-site overhead costs with greater accuracy.

Keywords: Site overheads, Estimating, Egyptian construction companies, ANN

1. 1. INTRODUCTION

Project overheads refer to the indirect expenses incurred in a project, such as providing general services at the site or plant, including insurance, site accommodation, and other similar costs [1]. In Egypt, project overheads are typically outlined as a separate section in the Bills of Quantities under the category of "Preliminaries." These overhead costs are usually estimated to represent a relatively small percentage of the total contract sum, ranging from around 15% to a maximum of 30% [2, 3], Despite its potential to boost bidding success rates and serve as a profit center for contractors, the estimation of project overheads receives less attention than the direct cost estimation, leading to an approximation rather than a precise estimation. This can be attributed to the misconception that these costs account for a relatively small percentage of the contract sum. Nonetheless, an accurate estimation of project overheads is critical in ensuring project profitability and success [1]. The research aims to identify a framework of significant project overhead factors within construction projects in Egypt. Hence, developing a detailed prediction model for construction site overhead using artificial neural network in Egypt. The study is limited to the construction sector in Egypt, yet the conclusions are supposed to generalize the results [4]. The current research focuses on the estimation of indirect costs (overheads) by contractor companies. These estimates are typically used for tendering purposes aimed at achieving project target profitability [5].

2. Literature review

2.1. Definition of overheads

Chan have mentioned that project overheads are commonly known as preliminaries [1]. Overheads are defined by the royal institution of chartered surveyors (RICS) as "the cost of managing a project such as the provision of plant, site staff, facilities for on-site services, and other miscellaneous items that are not covered by the rates for measured works." [6]. there are two types of overheads that apply to construction companies:

i. Office overhead costs

Office OH costs are generally the contractor's administrative costs for all projects that are ongoing to support the business and are more commonly referred to as general administrative costs [7]. The office OH costs are described as administrative costs to maintain the company's success [8]. Examples of costs included in the category of office overhead costs were given, including office rent, office salaries, office furnishings, taxation, permits, IT prices, and utilities. The cost of running a business and providing off-site services.

ii. Site overhead costs

(Also known as project overheads, general cost elements, or expenses). These site-specific costs associated with a project are not itemized for individual activities. These expenses may cover items such as site management, safety measures, work-related insurances, bonding charges, telephone, water, energy expenditures, and other similar costs [9]. The site OH involves costs that aren't directly connected to a particular activity but are still required to complete the project [10].

2.2. Factors that affect the accuracy of site overheads estimation

Many theorists have emphasized the value of project overhead estimation. Since direct costs were estimated similarly across bids, markups and indirect costs (project overheads) tended to be the main areas of conflict. Also, it was recommended that efforts be focused on enhancing the techniques for determining the indirect costs in order to increase the accuracy of the tender pricing. Recognizing costs is the first step towards reducing them [11]. Overhead costs, which account for a sizeable amount of project's costs, are one of the most critical cost centers that could enhance controlling as well as management of prices, raise the likelihood of winning a bid, and directly affect the company's financial status [12]. Additionally, it was shown that reducing overhead costs is a smart way of reducing overall costs. Although it is commonly known that overhead costs are not the main source of costs, they are a crucial factor in winning contract bids [13]. While it can be challenging to precisely track all overhead costs, taking the appropriate steps to estimate them accurately and documenting them separately or incorporating them into item rates can enhance the monitoring and management of cost structures, ultimately resulting in improved profitability [2, 5]. Due to the current competitive environment, underestimating project overheads is more common because an overestimate could result in an uncompetitive bid while an underestimate increases the likelihood that the bid will be successful [14, 15]. All the prior studies suggests that the decision of the individual estimator and the contractor's method are very important factors in the estimation of project overheads. The process of estimating indirect costs for a project is often considered time-consuming and prone to inaccuracy. As a result, many contractors tend to estimate indirect costs as a percentage of the direct costs to save time and simplify the estimation process. However, this method may not be accurate enough for most estimates [16]. The factors that impact overhead costs have been mentioned in previous studies are summarized below:

1. Project related factors

a. Project complexity

Project complexity can have a significant effect on the accuracy of overhead estimation. The more complex a project is, the more difficult it becomes to accurately estimate overhead costs. This is because complex projects typically involve a large number of variables, uncertainties, and interdependencies that can affect the allocation and distribution of overhead costs. For example, a complex construction project may require multiple subcontractors, specialized equipment, and extensive coordination between different teams, which can result in higher administrative expenses and lower productivity [17].

b. Project location

The location of the project can have an effect on various overhead costs, such as those related to travel, access, public property security, and the cost of setting up and maintaining temporary offices and other facilities [1, 18].

c. Project size

By expanding the project's scope, more personnel and resources are needed. As a result, it influences the size of overhead costs. In the study of Shash, the size of a project is the primary factor that affect the bids [19].

d. Type of project

The type of project can have a significant impact on overhead estimation, as different types of projects have unique characteristics and requirements that influence the allocation and distribution of indirect costs [17]. The nature of a project, such as whether it involves the construction of a road, dam, building, or other structure, can impact the number of jobs involved. Additionally, coordinating supervision and transportation can also affect overhead costs [1]. In a study conducted by El-Sawy, between 2002-2009, which analyzed 52 construction projects in Egypt, it was found that the type of project was the third most influential factor affecting overhead costs [20].

e. Duration of the project

Most research emphasizes that project duration is a critical factor which affects overhead cost, and that parts of the project directly related to duration account for more than 45% of the project's overall overhead costs. Additionally, the increased likelihood of project delays makes it even more crucial for this factor to have an impact on overhead costs [1]. El-Sawy showed that project duration is the most important effective factor in overhead costs [20].

f. Required quality level of the projects

While the necessary quality level for a project is occasionally influenced by the nature of the work, such as a fast transportation project requiring higher quality than a typical road project, this can significantly affect staffing and the required documentation for quality control, thereby impacting overhead costs [21]. To put it another way, the ability to achieve higher quality standards is contingent upon having sufficient management, tools, and resources, all of which can impact overhead costs.

2. Organizational factors

a. Past experience in similar projects

The project team can explain the project framework with greater accuracy since they have experience working on projects of a similar nature, which reduces the amount of trial and error. This consequently has an impact on reducing overheads. Due to the contractor's experience, the job could be completed successfully and to a high standard [22]. According to Nabil I. El-Sawalhi, the expertise involved in determining the percentage of overhead costs during the tender pricing process is a crucial factor in estimating overhead costs, given its sensitivity and pivotal role in securing a successful bid [23].

3. Client and government regulation

a. Type of contract

Standardizing the contract format or adapting it to the employer's conditions can impact the level of overhead costs associated with potential claims and the need to fulfill additional commitments, if any [24]. The type of contract used in a project can have a significant impact on overhead estimation, as different types of contracts have unique characteristics that can affect the allocation and distribution of indirect costs. For example, a cost-plus contract may have different overhead costs than a fixed-price contract, due to differences in risk, scope, and payment structure [17].

4. Environmental factors

a. Payment schedule

While contracts include provisions and stipulations regarding delays that could potentially impact the fulfillment of obligations, such as payments, these circumstances typically do not affect the performance of the project itself. Rather, they depend on the willingness of subcontractors and salespeople to adhere to the project schedule. Ultimately, this can lead to the imposition of extra costs and an increasing in overheads. In the research of Enshassi, the fifth real element driving up overhead costs is the delay in payments [5]. In order to complete the project more quickly, they claim that many contractors depend on these payments, and any delay in confirming these payments with the employers causes a pause in the contractors' work. According to research conducted by Memon, the primary factors that impact construction costs are the cash flow and financial challenges faced by contractors [22].

b. Regional economic condition

Companies have fewer projects during lean economic times, and cutting overhead costs is necessary to increase survival chances [1]. The region's economic situation has an impact on the wages paid to employees, the cost of services, the cost of machine leasing, etc., which in turn has an impact on the total amount of overheads.

c. Country of performing the project

The location of a project can have an impact on overhead costs due to factors such as cultural differences, governmental regulations, taxes, safety concerns related to war, and international recognition. For instance, in some countries, taxes are levied on construction projects carried out by different employers, resulting in increased overhead costs [25].

d. Number of competitors

According to a study by Assaf S. B.-S., the number of competitors ranked third among the significant factors that impact overhead costs [2]. However, in the research of Enshassi, these variable ranks seventh and has a less impact [5]. The number and quality of competitors can influence overhead costs, as increased competition may prompt contractors to lower their site overhead percentage.

e. Contractor's cash availability

In the research of Enshassi, After the project's location and complexity, the contractor's cash availability comes in third place in terms of the elements that have the biggest impact on overhead costs. However, this factor ranks fifth in terms of its impact on overhead costs [5]. Contractors with robust financial conditions tend to be less affected by it than those with weaker financial conditions. The contractor's cash availability, which has an impact on providing and supplying products and machinery, gives the opportunity to save money and complete the project more efficiently and has a significant impact on the total amount of overhead costs [2].

f. Companies' classification

The construction companies in Egypt are classified into

seven grades, ranging from 1 to 7, based on their financial and technical capabilities. Contractors at higher classification grades tend to have lower overhead percentages compared to lower-classification contractors. This is because higher-classification contractors have larger-scale
operations and are better equipped to handle larger projects, which can lead to economies of scale and lower overhead costs per project. Also, higher-classification contractors have more experience and better management systems, which can help them reduce their overhead costs.

g. Inflation and interest rate in the country

Inflation: A general rise in prices and a corresponding decline in the purchasing power of money steadily lowers purchasing power and has a serious impact on several economic sectors. Over the past decade, the Egyptian Construction Industry has experienced a notable rise in inflation, resulting in significant repercussions on ongoing projects. This can be attributed to various factors, including the high demand for certain construction materials, recent governmental regulations, fluctuating market conditions, and escalating fuel prices.
In the study of Enshassi, the inflation is the most important reason for increasing the site overhead costs.

3. Research objectives

This paper aims to examine the practices used in estimating project overheads in Egypt, with specific objectives that include:

- Identify factors that affect site overhead costs.
- Developing a detailed prediction model for construction site overhead using artificial neural network in Egypt.

4. Research methodology

For the purpose of achieving the goal of this research, a questionnaire was used to identify and assess factors that affect site overhead costs within construction companies in Egypt. Secondly, collect data from different construction projects in Egypt to be used in the model.
Artificial neural networks (ANNs) are constructed using input data and the desired output data required to train the model. The data set can be continuously updated by incorporating new examples to enhance the results and minimize errors. Therefore, an ANN may be employed to create a comprehensive prediction model for estimating the percentage of site overhead costs in Egypt [26].
Finally, a framework verification has made by applying the ANN prediction model using a case study in construction project in Egypt to verify the results and determine the OH percentage of accuracy.

5.Data collection and analysis

A structured questionnaire addressed to engineers, seniors, team leaders and managers from tender, estimation, controls and operation departments in construction companies and sent by mail. seventy respondents have been collected, each response is representing one project information, accordingly a list of seventy construction projects were collected. The questionnaire was conducted among construction companies in Egypt with different classes 1, 2, 3 and 4, as they are the most effective grades.
The questionnaire was distributed among the contractors to obtain their perspectives regarding the mentioned aspects of the overhead in construction projects in Egypt [23].

Sample size

Ensuring an adequate sample size is crucial for obtaining reliable conclusions based on research findings. This study encompasses contracting firms across the first, second, third, and fourth categories involved in construction projects. Random sampling was employed to select samples from each contractor category level, allowing for representative data collection. The formula which is shown in the below equation was used to determine the sample size of an unlimited population [38].

n = (Z^2 * p * (1 - p)) / (E^2) (1)

Equation 1. Sample size

Where:

n is the required sample size.

Z is the Z-score corresponding to the desired level of confidence (e.g., 1.645 for a 90% confidence level).

p percentage of picking a choice, expressed as a decimal (0.5 used for sample size needed).

E is the desired margin of error.

n = (1.645^2 * 0.5 * (1 – 0.5)) / (0.1^2) = 67.9~68

5.1. Results & findings

5.1.1. Factors identification: identify and assess factors that affect site overhead costs

The results of the questionnaire as illustrated in Table 1 shows that project’s duration, inflation and interest rate in the country are the most effective factors in site overhead cost, and type of contract and number of competitors have the lowest impact in site overhead cost.
In Table 1 the relative importance index for each factor are calculated to rank the factors from the most influence factor to the lowest one.

Table 1. Rank of factors that affect site overhead costs.

Factor	RII
project’s duration	0.9
Inflation and Interest rate in the Country	0.852
project’s size	0.796
Regional Economic Condition	0.792
project location	0.788
Contractor’s Cash Availability	0.768
Country of Performing the Project	0.764
Factor	**RII**
Past experience in similar projects	0.74

Payment Schedule	0.74
Project complexity	0.732
type of project	0.728
Companies' Classification	0.728
required quality level of projects	0.716
type of contract	0.704
Number of Competitors	0.688

5.1.2. Sampling of construction projects: collect data for different construction projects in Egypt to be used in the model.

Analyzing the collected data from the questionnaire of fifty different construction projects indicated the following:

- Twenty-one project were public sector, and twenty-nine were private sector.

- Thirty-seven project have a unit rate contract type, twelve were lump sum, and one project was cost plus.

-Twenty-four residential projects, twelve roads' projects, four industrial projects, three renovation projects, three educational projects, two hospital projects and two hotels.

-There were forty-two responses from companies grade 1, three from grade 2, two from grade 3 and three were lower than grade 3.

- There were twenty-seven projects with contract value of more than one billion, eight projects less than two hundred million, fourteen projects less than five hundred million and one project less than fifty million.

-There were twelve projects with a duration of less than 18 months, eleven project between 18 and 24 months, ten project between 24 and 36 months, eight project between 36 and 48 months and nine project with a duration of more than 48 months.

These statistics are illustrated in the below figure

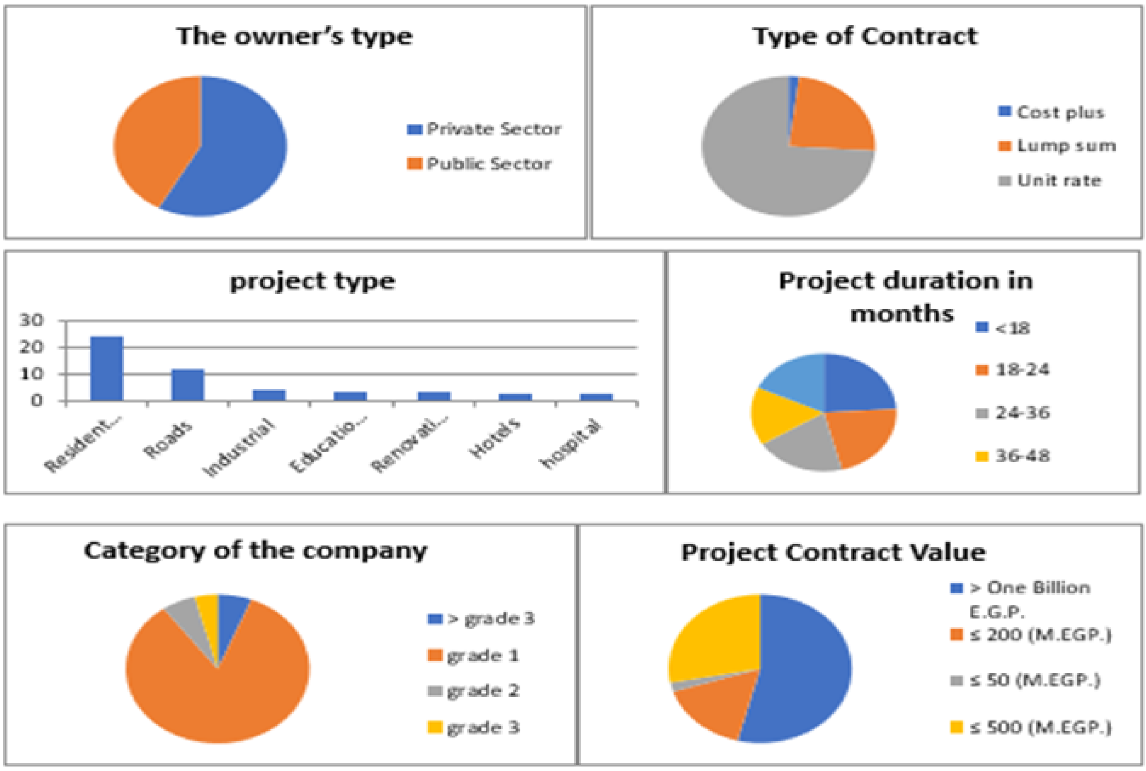


Fig 1. The collected data analysis

6.Artificial neural networks

6.1.ANN introduction

Artificial neural networks (ANNs) have been successfully applied in a variety of industries and are utilized as a means of solving complex and challenging problems [27]. The current research will concentrate on (ANNs), which are computer models inspired by the structure and function of biological neural networks. The architecture of the ANN is influenced by the flow of information through the network, and it learns by processing inputs and generating the desired output [28]. In a study conducted by Kulkarni the recent applications of artificial neural networks (ANNs) in various aspects of the construction industry, including productivity, cost, risk analysis, dispute resolution, duration, and unit rate, were analyzed and discussed. The findings of the study indicate that ANNs are effective tools for task classification, prediction, modeling, and optimization in the construction industry [26]. ANNs rely on input data and the desired output data required to train the model, and this data set can be updated continuously by incorporating new examples to enhance the results and minimize errors.

Therefore, ANNs can be utilized to develop a comprehensive prediction model for estimating the percentage of site overhead costs in Egypt.

This section presents the steps taken to develop the artificial neural network.

1. Collecting the needed project data for the learning process
2. Coding of each input
3. Design the network architecture and choosing the transfer function
4. Determine the learning algorithm
5. Training of the network
6. Validation of the network
7. Testing the network

6.1.1. Data set

The data set utilized in this research comprises the information required to construct the model. This data is organized in a tabular format, with rows and columns, and is typically stored in a data file, such as a CSV or Excel file. To make sure that the project data were being collected from a representative sample, a sample size calculation was made. The ability to import an excel data file is provided by "Neural Designer". It divides the data set to 60% for learning, 20% for validation, and 20% for testing and verification the network. Project data was coded before being entered into the data set file so that it could be correctly imported into the "Neural Designer" application. The six input variables in the model for estimating site overhead percentage are the rank of the company, project direct cost, project duration, project location, contract type, and ownership type of the company. After importing the data set file, the software categorises all data variables by default into inputs and outputs. And the output or target variable is the site overhead percentage.

6.1.2.Training strategy

Choosing the right training algorithm is the key component in developing a neural network model. Different learning techniques can be used to train the network and adjust the weights. In this study, all test models were trained using a supervised backpropagation algorithm. The algorithm identifies the difference between the desired output and the actual output, and automatically adjusts the network weight to minimize errors. As the training progresses, the network weights are continually adjusted until the observed performance reaches an acceptable level. The backpropagation process gradually reduces the error between the target output and the model output, resulting in the generation of output mapping data that reduces the root mean square error (RMSE). The algorithm decreases the RMSE, and the training concludes once the RMSE remains constant. The RMSE is used to measure model performance or accuracy [29].

$$\mathbf{RMSE} = \frac{\sqrt{\sum_{i}^{n}=1(\mathrm{Si}-\mathrm{Oi})^2}}{n} \qquad \mathbf{(2)}$$

Equation 2.Root Mean Square Error

The above equation represents the RMSE equation where:

(Si)= The ANN model's predicted output value.

(Oi)= The dataset's actual output value.

(n)= number of projects or samples.

The root mean square error serves as a reliable metric for determining the success of a training run. Also, relative percentage error would be calculated, it is a measure of the accuracy of an estimate or measurement. It is expressed as a percentage and represents the difference between the estimated or measured value and the true or expected value, as a proportion of the true or expected value [30].

RPE = (|estimated value - true value| / true value) x 100% **(3)**

Equation 3. Relative Percentage Error

6.2.Developing the model

6.2.1. Data set analysis using the software

6.2.1.1. Data set statistics

In the process of model design, basic statistics play a crucial role as they can help identify potential erroneous data points, and to make sure that they are reliable for use in the ANN modeling. They serve as an indicator to detect and address any incorrect or inaccurate inputs or information associated with each factor. It is imperative to thoroughly examine and verify the critical statistical measures of each variable to ensure their accuracy. Table 2 shows the basic statistics for each factor, which are minimum, maximum, mean, and standard deviations.

Table 2. Data statistics (Author)

	Minimum	Maximum	Mean	Deviation
Category	1	4	1.32	0.819
Budget	1	5	3.88	1.29
Duration	1	5	2.82	1.44
Type	1	7	2.74	1.93
Client (2)	0	1	0.42	0.499
contract type	1	3	2.5	0.863
OH%	0.15	0.35	0.206	0.0577

Company category

This factor was one of the most important factors in the questionnaire. According to the analysis, the mean was 1.32 and the standard deviation was 0.819.

Project budget

The total project budget, which indicates the total contract amount, shows a mean of

3.88, a standard deviation of 1.29.

Project duration

The project duration shows a mean of 2.82 and a standard deviation of 1.44. There are five categories of project duration, with a minimum of one and a maximum of five. The project will be assigned to one of these categories based on its duration.

Project type

A mean of 2.74 and a standard deviation of 1.93 are displayed for the project type factor, which might be either residential, commercial, or involve renovations of office buildings, hotels, roadways, or other structures. The dataset includes seven different types of projects as the minimum and maximum values are one and seven, respectively.

Nature of the client

The mean and standard deviation of this factor, which determines whether the client is private or public, are both 0.42 and 0.499, respectively. The minimum is one, and the maximum is two, which means there are only two categories of clients, public or private.

Type of contract

The contract type has a mean of 2.50 and a standard deviation of 0.863, indicating a lump sum, cost-plus, or unit rate contract.

6.2.1.2. Inputs correlation

Table 3 displays the absolute values of the input correlations among the various variables or inputs. These values, which range from 0 to 1, represent the strength and relationship between two variables.

- A correlation value near 1 indicates a strong relationship.
- A correlation value near 0 indicates weak or no relationship.

Table 3. Inputs correlation (Author)

	Category	Budget	Duration	Type	Client	contract type
Category	1	-0.404	-0.273	-0.0663	0.0637	-0.0603
Budget	-0.404	1	0.447	-0.153	-0.0462	-0.147
Duration	-0.273	0.447	1	-0.28	-0.399	-0.11
Type	-0.0663	-0.153	-0.28	1	0.522	0.297
Client	0.0637	-0.0462	-0.399	0.522	1	0.257
contract type	-0.0603	-0.147	-0.11	0.297	0.257	1

The project duration, project budget, and company category are the primary three factors that affect the site OH cost % in construction projects in Egypt, according to the software analysis of the inputs to quantify their influence on the output. The project type and client type, which are the other variables, have just a minor impact on the site OH costs %. The output is least affected by the client type and project type, and neither factor had an impact on the site OH percentage.

6.2.2.Model selection

This section employs a trial-and-error approach to train the network and identify the optimal model. During each trial, the perceptron layer is adjusted based on the following parameters:

- The number of layers
- The number of neurons in each layer
- The type of activation function used

The best model is selected based on the lowest testing RMSE value and relative percentage error. The trial-and-error process involves fifty trials, with variations made to the number of neurons in each trial.

-One hidden layer hyperbolic tangent activation function.
-One hidden layer logistic activation.
-Two hidden layers logistic activation function for each.
-Two hidden layers hyperbolic tangent activation function for each.

6.2.2.1. One hidden layer hyperbolic tangent activation function

The initial twelve models, each comprising one hidden layer with a hyperbolic tangent activation function, demonstrated a non-linear relationship between the number of neurons and the RMSE and RPE (**Fig 2**) among these models, the one with the lowest errors had an RMSE value of 0.4766 and a relative percentage error of 4.7358%, while the model with the highest errors had an RMSE value of 0.9438 and a relative percentage error of 7.7930%.

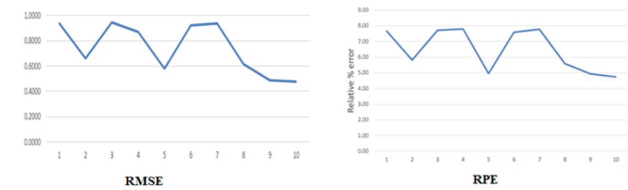


Fig 2. Model Trials with One hidden layer hyperbolic tangent activation function

6.2.2.2. One hidden layer logistic activation function

The subsequent models, ranging from 13 to 25, each with one hidden layer and a logistic activation function, demonstrated a non-linear relationship between the number of neurons and the RMSE and RPE (**Figure 3**) The model with the lowest errors in this range had an RMSE value of 0.5850 and a relative percentage error of 4.8746%, while the model with the highest errors had an RMSE value of 1.1539 and a relative percentage error of 9.7397%.

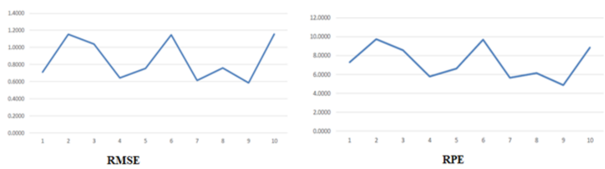


Figure 3. Model Trials with One hidden layer logistic activation function

6.2.2.3. Two hidden layers logistic activation function for each

Models 26 to 38, each consisting of two hidden layers with a logistic activation function, demonstrated a non-linear relationship between the number of neurons and the RMSE and RPE (**Fig 4**) among these models, the one with the lowest errors had an RMSE value of 0.2921 and a relative percentage error of 2.9298%, while the model with the highest errors had an RMSE value of 1.4636 and a relative percentage error of 9.8403%.

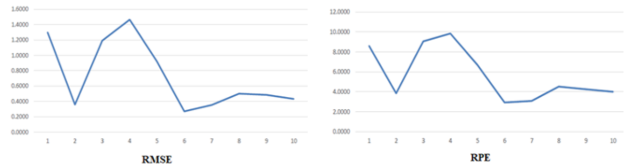


Fig 4. Model Trials with Two hidden layers Logistic activation function for each

6.2.2.4. Two hidden layers hyperbolic tangent activation function for each

Models 39 to 50, each comprising two hidden layers with a hyperbolic tangent activation function, demonstrated a non-linear relationship between the number of neurons and the RMSE and RPE (**Fig 5**), The model with the lowest errors in this range had an RMSE value of 0.4732 and a relative percentage error of 4.1252%, while the model with the highest errors had an RMSE value of 1.2974 and a relative percentage error of 10.720%.

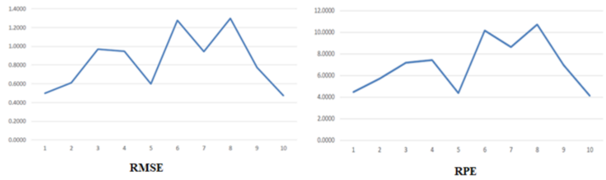


Fig 5. Model Trials with Two hidden layers hyperbolic tangent activation function for each

The optimal model, selected based on its lowest RMSE value, is model 26, which comprises two hidden layers with a logistic activation function. The first hidden layer has five neurons, and the second hidden layer has three neurons, as shown in (**Fig 6**), This model achieved the lowest RMSE for testing, with a value of 0.2920 and a relative percentage error of 2.9297%. The figure below depicts the structure of the chosen model, which includes a scaling layer represented by yellow circles, perceptron layers represented in blue circles, and an output layer represented in the red circle.

- Input variables: 6 variables
- First hidden layer: 6 neurons
- Second hidden layer: 5 neurons
- Output layer: 1 neuron
- Activation function: Logistic sigmoid activation function.
- RMSE: 0.2921
- Relative percentage error: 2.9298%

The linear regression analysis of the output, illustrated in **Figure 7**, represents the predicted percentage of site overhead costs for the chosen Model 26, using ten classified projects as testing instances. According to the chart depicted in **Figure 8** the linear regression analysis exhibited a robust correlation, with an R2 value of 0.888

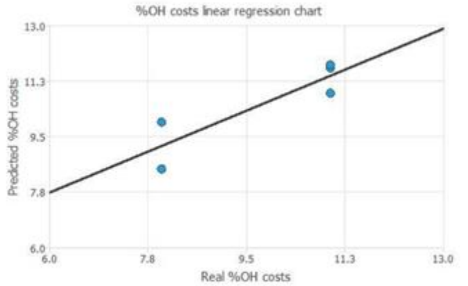


Fig 6.The selected model design

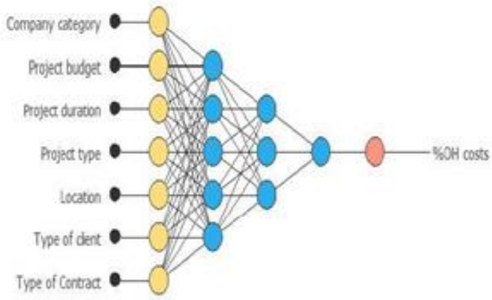


Fig 7.The selected model linear regression chart

6.2.3.Model deployment

Applying a model to forecast fresh data is referred to as deployment in machine learning. It is true that for the customer to use the knowledge obtained when developing a predictive model, it must be organized and presented. The knowledge acquired during the creation of a predictive model must be organized and presented so that users can benefit from it. The neural designer software allows the model to be exported in various formats to facilitate further research and comprehension. The resulting expression is presented below [29]:

- Python expression
- R expression
- Mathematical expression

In this study, the model will be exported to the python language, which is widely recognized and utilized in the field of machine learning.

6.2.3.1.Export to python

Currently, python is one of the most widely used programming languages, particularly in the realm of AI and neural network programming. It is an interpreted language, which eliminates the need to compile machine language instructions and allows developers to run programs directly. This feature enables the language to be easily managed in the native machine learning language,

which is comprehensible to hardware, making it a powerful tool for machine learning applications [31]. Python is a high-level programming language that can be applied to complex scenarios. It is considerably more versatile than other languages, as it can handle variables, objects, arrays, complex arithmetic, and other computer science concepts. Moreover, python is a general-purpose language that is compatible with various platforms and technologies. It also features an automated memory management system that supports diverse programming paradigms, including imperative, functional, and object-oriented. C-Python, an open-source variant of python, is extensively used across multiple operating systems [31].
Deep learning is a crucial aspect of machine learning, as it constitutes the primary feature of neural network development and serves as the foundation for various technologies in numerous industries. Python can be utilized to create artificial neural networks (ANN) for machine learning [32].
For this study, the chosen model was exported to the python language, and the python online compiler was used to examine the file. The python language model is presented in appendix A, and the result from entering the value of the six variables in the shell window to calculate the OH% is shown in the below figure:

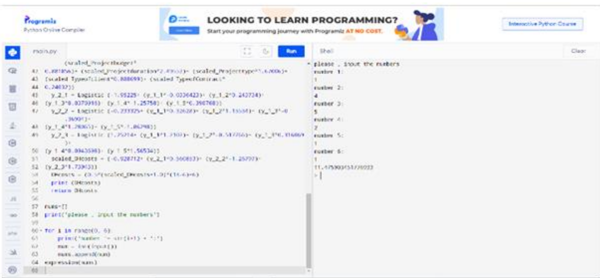


Fig 8. The calculated output from the selected model in Python language opened with Python Online Compiler (Author)

In this trial the researcher has chosen the following project characteristics:
1.Company category to be 1st grade which have code (1).
2.Project budget to be < 1 billion which have code (4).
3.Project duration to be > 48 months which have code (5).
4.Project type to be hotels which have code (2).
5.Client type to be private sector which have code (1).
6.Contract type to be lump sum which have code (1).
The OH percentage according to these inputs is 11.5%.
This code is eligible to be used by tender departments in the phase of bidding to help in estimating the overhead percentage for any project from the earlier defined categories.

6.3. Framework verifications

After the model was developed to predict the percentage of overhead costs for construction projects in Egypt, and its performance was demonstrated using data from a randomly selected project.
Framework verifications will be conducted to validate the data to verify the results and determine the percentage of accuracy, The verification would be conducted by two ways:

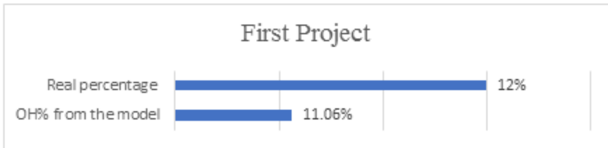
- •The first one is to try two different real construction projects in Egypt which are different in their characteristics, in order to compare the real OH percentage for these projects and the predicted one.
- •The second way is to compare the ranking of the factors which have resulted from the previous studies.

6.3.1. First verification method:

6.3.1.1. First project:

- Company category: 1st grade which have code (1).
- Project budget: < 1 billion which have code (4).
- Project duration: > 48 Months which have code (5).
- Project type: residential which have code (1).
- Client type: private sector which have code (1).
- Contract type: lump sum which have code (1).

According to the above data, the OH% resulted from the model is 11.06%, and the real percentage is 12%.

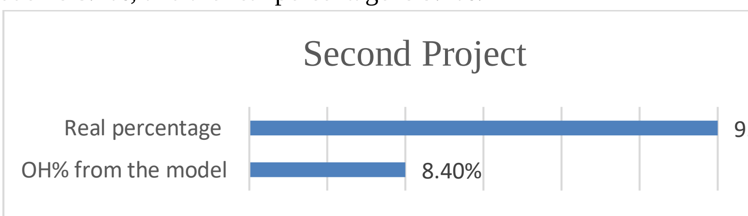


Accordingly, the percentage resulted from the model is relatively accurate.

6.3.1.2. Second project:

- Company category: 1st grade which have code (1).
- Project budget: < 200 million which have code (2).
- Project duration: < 18 Months which have code (1).
- Project type: residential which have code (1).
- Client type: private sector which have code (1).
- Contract type: Unit rate which have code (3).

According to the above data, the OH% resulted from the model is 8.4%, and the real percentage is 9.2%.



Accordingly, the percentage resulted from the model is accurate.

6.3.2. Second verification method:

From previous studies, I have compared the ranking of the factors which affected project overheads and found the following results:

•Project duration: The duration of the project was found to be the most influential factor affecting overhead costs, with a correlation coefficient of 0.529 indicating a direct proportional relationship between project duration and on-site overhead costs. Othman has ranked the same as project duration has a correlation of 0.87 in his results [33]. Bakr et al. gave weight to each factor, so project duration was the second one in the ranking as it took the same weight as the company category [34].

•Project budget: The project budget was identified as the second most significant factor affecting site overhead costs, with a correlation coefficient of 0.436 indicating a relationship between project budget (or actual contract value) and the percentage of on-site overhead costs. Othman has ranked the same as project budget has a correlation of 0.563 [33]. Bakr et al. said that project budget has the third importance [34].

•Company category: The company category was found to be the third most influential factor affecting on-site overhead costs, with a correlation coefficient of -0.2634 indicating a relationship between company category and the percentage of site overhead costs. For the purposes of this study, the company category was limited to first and second-grade companies only. Othman has ranked the same as company category has a correlation of -0.416 [33]. Bakr et al. said that company category was the second one in the ranking [34].

•Contract type: The contract type was found to have the least significant impact on on-site overhead costs, with a correlation coefficient of -0.081 indicating a weak relationship between contract type and the percentage of site overhead costs. Othman mentioned that it influences the site OH costs with a correlation of -0.217 [33]. Bakr et. al reported that contract type was the first one in the ranking [34].

•Client type: The type of client was found to have a slight impact on on-site overhead costs, with a correlation coefficient of -0.167 indicating a weak relationship between the nature of the client and the percentage of site overhead costs. Othman mentioned that it influences the site OH costs with a correlation of -0.174 which is also a weak relation [33]. Bakr et al. has ranked client type as the lowest one [34].

•Project type: The type of project was found to have a minor impact on site overhead costs, with a correlation coefficient of -0.11 indicating a weak relationship between project type and the percentage of site overhead costs. Othman mentioned that the project type was ranked the lowest factors among the factors that influence the site OH costs with a correlation of 0.13, which is almost near zero [33].

7. Conclusions

The goal of this research is to overview the variables that affect the precision of site overheads estimation. Accordingly, a questionnaire was conducted among the top four construction companies in the category in Egypt. The results of this questionnaire were analyzed to rank the factors and study their importance.

Also, the effectiveness of artificial neural networks (ANNs) was examined in addressing the challenge of accurately estimating project overhead costs during the initial stages of building design.

The results of the research can be summarized in the following points:

1. This study highlights the advantages of utilizing neural network technology for estimating overhead costs, which include its simplicity, fast calculation speed, and high accuracy in making predictions.
2. The proposed model was tested and found to be accurate, as demonstrated in the framework verification. This enhances contractors' ability to make precise cost predictions, thereby maximizing their profits.
3. Through data analysis, it was determined that there is no single ideal percentage for on-site overhead costs in Egypt.
4. The study found that the use of the ANN model is an effective tool for minimizing the amount of work required to estimate the percentage of on-site overhead costs with greater accuracy. To facilitate the implementation of this model, the research provides a general overview of ANN components and a guide for using the Neural Designer software. It should be noted that this model is based on the application of ANN and programming in Python.
5. It was found and verified that the most important factors were project’s duration, inflation and interest rate in the country and project’s size.
6. The results demonstrated a high level of reliability for the model, indicating its suitability for use in obtaining overhead percentages within acceptable limits.

References

[1]- Chan, C T W and Pasquire, C (2002) Estimation of project overheads: a contractor’s perspective. In: Greenwood, D (Ed.), 18th

Annual ARCOM Conference, 2-4 September 2002, University of Northumbria. Association of Researchers in Construction Management, Vol. 1, 53 62.

[2]- Assaf, SA, Bubshait, AA, Atiyah, S and Al-Shahri, M (1999) Project Overhead Costs in Saudi Arabia. Cost Engineering, 41(4), 33 –38.

[3]- Hegazy, T and Moselhi, O (1995) Elements of Cost Estimation: A Survey in Canada and the United States. Cost Engineering, 37(5), 27 – 33.

[4]- Dunna, B., & Burela, V. P. Success factors for effective implementation of project controls in contracting companies: A Qualitative study, 2008.

[5]- Enshassi, A., Rashid Abdul Aziz, A. and Karriri, A. (2008), "Investigating the overhead costs in construction projects in Palestine", Journal of Financial Management of Property and Construction, Vol. 13 No. 1, pp. 35-47, doi: 10.1108/13664380810882066.

[6]- RICS (2015), RICS New Rules of Measurement, NRM 2: Detailed Measurement for Building Works, Coventry, Published by Royal Institution of Chartered Surveyors (RICS), available at: https:// www.rics.org/globalassets/rics-website/media/upholding-professional-standards/sectorstandards/construction/nrm-2-detailed-measurement-for-building-works-1st edition-rics.pdf.

[7]- Zack, Jr. James G. (2002), "Calculation and recovery of home office overhead", Construction management Journal, International Cost Engineering Council, 3rd world Congress on cost Engineering, Project Management and Quantity Surveying, 6th Pacific Association of Quantity Surveyors congress, Melbourne, Australia, April 14-18.

[8]- Shelton, F. & Brugh, M., 2002. Indirect costs of contracts. Journal of Construction Accounting and Taxation, 12(4), pp. 3-9.

[9]- The Chartered Institute of Building (2018) Project Overheads. In: Code of Estimating Practice, 1sted. John Wiley & Sons Ltd.

[10]- Lowe, S., Bielek, R. & Burnham, R., 2003. Compensation for Contractors' Home Office Overhead, Washington, D.C: National Cooperative Highway Research Program.

[11]- Tah, JHM, Thorpe, A and McCaffer, R (1994) A Survey of Indirect Cost Estimating in Practice. Construction Management and Economics, 12, 31-36.

[12]- Plebankiewicz, E., and Leśniak, A. (2013). Overhead costs and profit calculation by Polish contractors. Technological and Economic Development of Economy, 19(1), 141-161.

[13]- Šiškina, A., Juodis, A., & Apanavičiene, R. (2009). Evaluation of the competitiveness of construction company overhead costs. Journal of Civil Engineering and Management, 15(2), 215-224.

[14]- Ioannou, P.G. and Awwad, R.E. (2010), "Below-average bidding method", Journal of Construction Engineering and Management, Vol. 136 No. 9, pp. 936-946, doi: 10.1061/(ASCE)CO.1943-7862. 0000202.

[15]- Shrestha, S.K. (2014), "Average bid method – an alternative to low bid method in public sector construction procurement in Nepal low bid method", Journal of the Institute of Engineering, Vol. 10 No. 1, pp. 125-129, doi: 10.3126/jie.v10i1.10888.

[16]- Peurifoy, R. a. (2002). Overhead. In: Estimating Construction Cost, 5th ed. 14.

[17]- Institute., Project Management Institute. (2017). A Guide to the Project Management Body of Knowledge (PMBOK® GUIDE) – Sixth Edition .

[18]- Cooke, B. (1981) Contract Planning and Contractual Procedures, Macmillan, Basingstoke.

[19]- Shash, A. A., and Abdul-Hadi, N. H. (1992). Factors affecting a contractor's mark-up size decision in Saudi Arabia. Construction Management and Economics, 10(5), 415-429.

[20]- ElSawy, I., Hosny, H. and Abdel Razek, M. (2011), "A Neural Network Model for Construction Projects Site Overhead Cost Estimating in Egypt". IJCSI International Journal of Computer Science Issues, Vol. 8 No. 1, pp. 273-283.

[21]- Chao, L. C. (2010). Estimating project overheads rate in bidding: DSS approach using neural networks. Construction Management and Economics, 28(3), 287-299.

[22]- Memon, A. H., Abdul Rahman, I., Abdullah, M. R., and Abdu Azis, A. A. (2011). Factors affecting construction cost in Mara large construction project: perspective of project management consultant. International Journal of Sustainable Construction Engineering and Technology, 1(2), 41-54.

[23]- Nabil I. El-Sawalhi, Ahmed El-Riyati. An Overhead Costs Assessment for Construction Projects at Gaza Strip. American Journal of Civil Engineering. Vol. 3, No. 4, 2015, pp. 95-101. doi: 10.11648/j.ajce.20150304.11

[24]- HESAMI, S.; LAVASANI, S. Amin. Identifying and classifying effective factors affecting overhead costs in constructing projects in Iran. International Journal of Construction Engineering and Management, 2014, 3.1: 24-41.

[25]- Shen, L. Y., Wu, G. W., and Ng, C. S. (2001). Risk assessment for construction joint ventures in China. Journal of Construction Engineering and Management, 127(1), 76-81.

[26]- Kulkarni, P. S., Londhe, S. N. & Deo, M. C., 2017. Artificial Neural Networks for Construction Management. Journal of Soft Computing in Civil Engineering, 1(2), pp. 70-88.

[27]- Lazarevska, M., Knezevic, M., Cvetkovska, M. & Gavriloska, A. T., 2014. Application of Aritifical Neural Networks In Civil Engineering. Tehnički vjesnik, 21(6), pp. 1353-1359.

[28]- Chauhan, N. S., 2019. Introduction to Artificial Neural Networks (ANN). [Online] Available at: https://towardsdatascience.com/introduction-to-artificial-neural-networks-ann-1aea15775ef9 [Accessed 05 Aug. 2020].

[29]- Neural Designer, 2023. What is Neural Designer? [Online] Available at: https://www.neuraldesigner.com/learning/tutorials/neural-networks-applications

[30]- Neil, S. P. & Hashemi, M. R.(2018)." Chapter 8 - Ocean Modelling for Resource Characterization."ed.,193-235.

[31]- Cuelogic Technologies Pvt. Ltd.(2016).Role of Python in Artificial Intelligence (AI). Retrieved August 26, 2020, from https://www.mybib.com/tools/harvard-referencing-generator

[32]- Magnimind. (2020, August 26). Why Learning Python Is Important For Machine Learning. Retrieved from https://becominghuman.ai/why-learning-python-is-important-for-machinelearning-

[33]- Othman, S. "Evaluation of Indirect Cost Estimation in the Egyptian Construction Industry." Master thesis, Joint Study Programme between HTW Berlin and Metropo-lia UAS, 2021.

[34]- Bakr, A.A., Ragab, A.A. and Yehia, N.A. ESTIMATING OF SITE OVERHEADS FOR RESIDENTIAL PROJECTS IN EGYPT USING NEURAL NETWORK. Proceedings of the 2nd International Conference of Sustainable Construction and Project Management, 2018.

[35]- Al-Tawal, D.R., Arafah, M. and Sweis, G.J., A model utilizing the artificial neural network in cost estimation of construction projects in Jordan. Engineering, Construction and Architectural Management, 2020.

[36]- Eksteen, B. and Rosenberg, D. (2002), "The management of overhead costs in construction companies" in Greenwood, D. (Ed.), 18th Annual ARCOM Conference, 2-4 September 2002, and University of Nothumbria, Vol. 1 Association of Researchers in Construction Management, pp. 13-22.

[37]- "A neural network approach for early cost estimation of structural systems of buildings" International Journal of Project Management 22 (2004) 595–602. H. Murat Gunaydın *, S. Zeynep Do_gan April

2004, Faculty of Architecture (Mimarlık Fakultesi), Izmir Institute of Technology (_Izmir Yuksek Teknoloji Enstitusu), Gulbahce Koyu,Urla, 35430 Izmir, Turkey.

[38]- Ibrahim, Ahmed H., and Lamiaa M. Elshwadfy. "Factors affecting the accuracy of construction project cost estimation in Egypt." Jordan Journal of Civil Engineering 15, no. 3 (2021).

Appendix A

```
#!/usr/bin/python
from math import exp
def Logistic(x) :
   return(1/(1+exp(-x)))
def expression(inputs) :
   if type(inputs) != list:
      print('Error: Argument must be a list')
      return None
   if len(inputs) != 6:
      print('Error: Incorrect number of inputs')
      return None
   Companycategory=inputs[0]
   Projectbudget=inputs[1]
   Projectduration=inputs[2]
   Projecttype=inputs[3]
   Typeofclient=inputs[4]
   TypeofContract=inputs[5]
 scaled_Companycategory = (Companycategory-1.175)/0.384808
 scaled_Projectbudget = (Projectbudget-4.025)/2.13022
 scaled_Projectduration = 2*(Projectduration-1)/(6-1)-1
 scaled_Projecttype = (Projecttype-2.9)/1.44648
 scaled_Typeofclient = (Typeofclient-1.075)/0.266747
 scaled_TypeofContract = 2*(TypeofContract-1)/(3-1)-1
y_1_1 = Logistic (0.648813+ (scaled_Companycategory*-0.098206)+ (scaled_Projectbudget*
0.475306) +  (scaled_Projectduration*0.56265)+ (scaled_Projecttype*-
1.42937)+ (scaled_Typeofclient*1.76293)+ (scaled_TypeofContract*
0.573361))
  y_1_2 = Logistic (0.421492+ (scaled_Companycategory*1.41238)+
    (scaled_Projectbudget*-
  0.796516)+ (scaled_Projectduration*-0.417334)+
    (scaled_Projecttype*-0.969011)+
  (scaled_Typeofclient*-0.221975)+ (scaled_TypeofContract*
  2.01511))
    y_1_3 = Logistic (-0.60516+ (scaled_Companycategory*-
    1.74375)+ (scaled_Projectbudget*
  0.625127)+ (scaled_Projectduration*-1.41318)+
    (scaled_Projecttype*-0.397128)+
  (scaled_Typeofclient*-1.67949)+ (scaled_TypeofContract*
  1.40095))
    y_1_4 = Logistic (-1.22115+ (scaled_Companycategory*-
    0.325518)+ (scaled_Projectbudget*-
  1.28677)+ (scaled_Projectduration*0.594703)+ (scaled_Projecttype*-
    0.246565)+
  (scaled_Typeofclient*1.13003)+ (scaled_TypeofContract*
  0.21472))
    y_1_5 = Logistic (1.98621+ (scaled_Companycategory*-
    0.975683)+ (scaled_Projectbudget*
  0.881856)+ (scaled_Projectduration*2.49552)+
    (scaled_Projecttype*1.67006)+
  (scaled_Typeofclient*0.888699)+ (scaled_TypeofContract*
  0.24032))
    y_2_1 = Logistic (-1.95225+ (y_1_1*-0.0336423)+
    (y_1_2*0.243734)+
   (y_1_3*0.0370916)+ (y_1_4*-1.25758)+ (y_1_5*0.390768))
     y_2_2 = Logistic (-0.233325+ (y_1_1*0.32628)+
(y_1_2*1.13534)+ (y_1_3*-0.36904)+
    (y_1_4*1.28065)+ (y_1_5*-1.86298))
       y_2_3 = Logistic (1.25214+ (y_1_1*1.2102)+ (y_1_2*-
0.517766)+ (y_1_3*0.316069)+
    (y_1_4*0.0943698)+ (y_1_5*1.56534))
       scaled_OHcosts = (-0.928712+ (y_2_1*0.560853)+ (y_2_2*-
1.26797)+
    (y_2_3*1.73943))
       OHcosts = (0.5*(scaled_OHcosts+1.0)*(13-6)+6)
       print (OHcosts)
       return OHcosts
    nums=[1, 1, 1, 1, 1, 1]
    expression(nums)
```

List of Abbreviations

AI	Artificial intelligence
ANN	Artificial neural network
MSE	Mean square error
OH	Overhead cost
RMSE	Root mean square error
RPE	Relative Percentage Error