

A NEURAL NETWORK APPROACH TO DESIGN REALITY ORIENTED COST ESTIMATE MODEL FOR INFRASTRUCTURE PROJECTS

Nirajkumar Suneja, Prof. Jagruti P. Shah, Dr. Zankhana H. Shah, Dr.
Mehfuza S. Holia

•
B.V.M. Engineering College, Vallabh Vidyanagar, Gujarat, India
nirajsuneja1179@gmail.com

Abstract

Great infrastructure projects are flagrant for costing more than in the beginning intended. Road project cost overrun (under estimation) and cost revision is common major problems reported by MoSPI (Ministry of Statistics and Programme) 20% cost overrun in India. It is difficult to maintain infrastructure progress within cost, time and quality because of their unpredictability, enormous scope, long length, high speculation and longitudinal site condition. Framework Estimate engineers need to create cost estimation before time on in development to determine its practicability safe and sound financial support, meet up superiority standards and worth for money. this research focus on cost of road infrastructure projects in early stage before construction. Various literatures government report, private companies report Cost estimate guides is studied and it's study gives the idea of current cost estimation methods and causes of cost overrun in projects. Be that as it may the prior on in project's life cycle, the less is known around the particular points of interest of what really has to be built and a site's interesting characteristics and how they will influence, Research gives effort on the identifying proper attributes or factors that are readily available at an early stage for fast, easy and accurate detailed cost estimate from the past related research work. Different organizations have plenty of historical project information and lots of don't have the tool the other way around, very fewer resources to estimate compare to traditional methods. Machine learning (ML) cost estimation model advance uses algorithms to discover relationship between diverse variables of project and their cost and developed model.

Keywords—Road infrastructure, Cost estimate, Cost overrun, Cost Estimate Model, Machine learning, Artificial intelligence

I. Introduction

The purchase and retail and manufacturing growth of whatever nation is mostly dependent relative on the advance of its infrastructure and India is no invulnerability. India has the one of the largest road system across the ball, across over a sum of 5.5 million km. This road system transport 64.5% of all products in the country and 90% of India's absolute traveler traffic utilizes the street framework to move.[4]

The likely expense of doing for 1,424 Projects is fixed at Rs. 21.34 lakh crore marks. These undertakings were before all else expected to be done at an expense of Rs. 18.17 lakh crore, of which Rs. 8.07 lakh crore or 37.8% of unsurprising expense have been now brought about. And the percentage of cost overrun in infrastructure project is shown in Chart 1.1 (below) from year review of 2014 to 2018.[7]

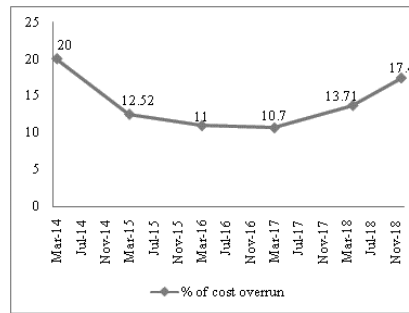


Chart 1 Percentage of cost overruns

The detailed estimate is the mainly precise form of estimates, other than it is moreover the majority lengthy; require the whole data to dispatch. Still to some extent small cost over-runs can reason to disturbance when a project is part of the pack of a wider program of outflow.

Following are cost drivers of infrastructure projects:-

1. Design & Development cost,
2. Material & Labour cost,
3. Land cost,
4. Insurance,
5. Professional services.
6. Legal& regulatory requirement cost (EIA: Environment impact Assessment).

The major factor responsible for cost overrun reportedly been (MoSPI Flash Report: 2019) [6]:-

1. Scope Changes
2. Design change/iterations
3. Inadequate DPR
4. Weak procurement planning
5. Unidentified risk
6. Balanced finance
7. Changes in rates of foreign exchange
8. Statutory duties/Taxes
9. General Price rise / inflation

Over-estimating the scope of projects and Under-estimating project costs are the two main reasons for cost overrun cost estimate / Underestimate.

Inside the worth estimation of road infrastructure comes that are there are three potential things where computable price figures either match or don't match with actual cost figures as shown in Table IX.

Situation	Remarks
Est. = Act.	Ideal
Est. < Act.	Under-estimation of funds leading to fund shortage
Est. > Act.	Over-estimation of funds leading to fund surplus

A. Aim and Objectives

The Aim of this study is “Designing Reality Oriented Cost Estimate Model for Transport Infrastructure Projects” develop model for Gujarat state roads. The objectives of this study are to identify influence factors and used them into in model to which is develop using python 3.7 language and Jupyter notebook with Machine learning (Deep Neural Network) Algorithm.

B. Study Area Profile

To design reality oriented cost estimate model for Transport infrastructure projects (Road) of Gujarat Region with Machine learning (Deep Neural network) algorithm with python 3.7 language Jupyter Notebook IDE which are easy to use and open source.

II. Related Works

The literature review provides knowledge of value estimation ways, Models and varied techniques that are utilized in road building value estimation by engineers. The worth estimate might be a most important task done by price engineer Major program structure comes associated programs suffer from an inclination to cost heaps of or carry longer than beginning assessments plot. The reasons for this are advanced nature of significant assignments and projects infer that, their assessment helpful and plan typically conveys confined precision.[3]

At intervals the Project Management word, it meant that there are massive variations at intervals the project attained value metrics, which may or will not be monitored closely. In fact, the effectiveness of the danger management can replicate at intervals the attained value metrics similar to price variance and schedule variance.[10]

Value estimates are elementary to any or all or any project-related branch of data also, enormously impact arranging, plan, request, esteem the executives/planning and building the board. Trademark table to the obliged helpfulness of data directly through the principle levels of an undertaking, development administrators much of the time impact their mindfulness, understanding and constant PC to gauge venture costs. This paper presents thorough depictions of the extended cycle Fuzzy Hybrid Neural Network (EFHNN) and its application in theoretical worth assessing for building comes.[1]

In preparing the model, the creators beginning affirm the constituents that decide the hugeness of innovation administrations, consequently utilize partner worthy ANN for a worth appraisal model. The model predicts the significance of innovation administrations as a territory of development cost upheld venture type, designing administrations class, venture area, and task degree.[2]

Construction worth prediction is extraordinarily important for construction corporations to contend and grow at intervals the business. Correct construction worth prediction at intervals the primary stage of the project is extraordinarily important for project utility studies and booming completion. Several factors have an effect on the price prediction.[5]

Estimation of the worth of development comes may be an imperative assignment at interims the administration of these comes. The quality of development administration depends intensely on their adjust esteem estimation. Esteem Evaluating is one in each of the first essential angles for legitimate execution of any development company.[12]

Extended a way that appraises the difference in size of course development costs after some time. an engineered neural organization model was built up that relates generally course development costs, diagrammatical as far as a course development indicant, to the value of development material, work, and instrumentation, the attributes of the agreement and moreover the getting environmental factors winning at the time the agreement was let. Results show that the model is set up to copy past course development esteem patterns in American state with sensible precision.[11]

III. Neural Network

Neural Network (Machine Learning: ML) could also be a spread of knowledge analysis tools that pulls out info whereas not being expressly programmed to behave thus. NN is closely involving engineering (AI). Through NN, computer systems are developed to do to odd jobs love categorization, clustering, predictions, pattern recognition, etc.

1. Supervised Learning

2. Un-Supervised Learning
3. Reinforcement Learning

The neural organization is another technique for data preparing. Numerous ample and basic preparing units (neurons) structure the unpredictable organization framework. The neural organization is one sort of enormous scope equal association system with versatile displaying capacity, which reenacts the structure of human mind. In numerous sorts of neural organization models, the back-proliferation neural organization model is the most mainstream network on account of its better elements of self-study and self-affiliation. The any standard organization is made out of three sorts of neurons layer. The main layer is known as the information layer. The center one is named as the concealed layer (can be multi-layer). Furthermore, the main one is known as the yield layer. Each layer of neurons frames completely association, and the neurons in each layer have no association. The information data is moved and prepared through info layer and shrouded layer. The condition of each neural unit layer just influences the condition of next layer. In the event that the normal data can't be got in the yield layer, the course will transform into the enactment capacity and standardization of sources of info and return the blunder signal along the previous association way. Through adjusting the estimation of weight between every neuron, the mistake signal is communicated organized into the information layer, and afterward be sent into the following layer. The rehashed use of these two courses makes the mistake more and a lot more modest, until it meets the prerequisites.

IV. Methodology of model development

So as to accomplish the Aim of the investigation, a technique was created in three stages. In the primary stage, other cost assessment examines utilizing NN were investigated, and afterward the plan of the model for application in roadway development was characterized. In the last Phase, the model was tried to research the best organization arrangement so as to get more prominent exactness. The stages will be introduced in the accompanying areas.

A. Information available during conceptual project development level

The information obtainable is one amongst the distinctive options of the various forms of price estimates. The estimate level among the various forms of estimates obtainable does not provide clear and definite boundary. Because of this reason, some literatures use the terms abstract and preliminary estimate interchangeably. Throughout the abstract section, general info regarding the project are going to be obtainable comparable to the road length, road carriageway, terrain project, location of the project and therefore the project scope, right of way etc.[7] The identification of the influential factors or parameters affecting total project construction cost is necessary to develop cost-estimating model. The input parameters identified through literature review of road construction cost estimation.

B. Description and Analysis on preliminary data collection

The Government official site database was utilized for this research study. This database contains Road project cost data Gujarat state. Sources are representing in Table X from the source collected 124 road projects data.

Table X Data collection source

Government Organization
Gujarat State Road Development Corporation Limited
World bank
Infrastructureindia.gov.in : Data base of Infrastructure projects in India

Road and Building Department, government of Gujarat
Gujarat state highway project – 2
Asian development bank
PPPindia.gov.in
<i>Total 123 Road Projects data collected</i>

The data analyses with removing null values and Interquartile range (IQR) method in Excel. The final data collection resulted in production of set of data to be used as historical feed for develop cost estimation model. Totally 124 data were collected and out of these 113 data were used for conceptual estimation model development. Rejection of data is based on the project based on rigid pavement, Curves having null values. The rejected data could not be used for model development due to incompleteness of data, inconsistency of data, the percentage distributions of the data set in terms of factors are shown in charts and tables.

Table XI. Data proportion in collection

Type of Project	No of Projects	Percentage
NH	12	11%
SH	94	83%
MDR	7	6%
Total	113	100%
Scope of Project	No of Projects	Percentage
Widening and Improvement	113	100%
Flexible	113	100%
ROW	No of Projects	Percentage
0-30	74	65%
30-60	39	35%
Total	113	100%
Thickness of Pavement	No of Projects	Percentage
0 – 650	20	18%
650 -1300	93	82%
Total	113	100%
Overlaying	No of Projects	Percentage
0-450	106	94%
450-900	7	6%
Total	113	100%

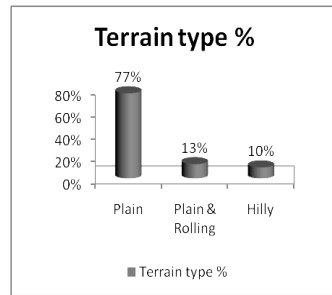


Fig. 1 Terrain Type

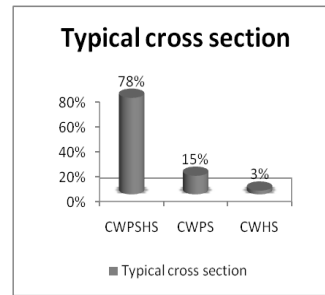


Fig. 2 Typical Cross section

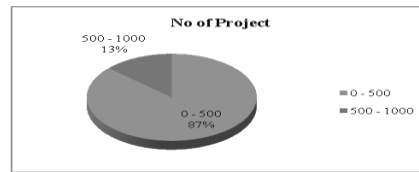


Fig. 3 Cost Distribution

C. The design and Modeling phase

A system for building up the model has been utilized to take care of the current issue. This philosophy consolidates five fundamental stages:

1. Assortment of Tool
2. Design Structure
3. Model implementation
4. Training and Testing
5. Discussion of results

This paper gives a model utilizing Regression Keras, which is based Deep learning rule. The computational instrument utilized was the Jupyter Notebook (IDE) that is embedded in the Python 3.7 Language. Firstly insert Required Libraries and modules.

Assigned input and output variables, the data should be in such the simplest way that the organization will decide the numerical connections between the information sources and their few yields. Then normalize data between zero to one.

Table XII Hot encoding

No.	Chosen Factors	Code
1	Type of road	NH = 1 SH = 2 MDR = 3
2	Scope of Work	New = 1 Widening & improvement = 2
3	Number of Lane	1L = 1 2L = 2 4L = 4 6L = 6
4	Length of Road	In Kilometres
5	Right of Way	In Metre
6	Type of pavement	Flexible = 1 Rigid = 2
7	CBR (%)	In Percentage
8	Terrain Type	Plain = 1 Plain & Rolling = 2 Hilly = 3
9	Thickness of Pavement	In Millimetres
10	Over Laying Thickness	In Millimetres

11	Width of paved Shoulders	In metre
12	Width of Hard Shoulders	In metre
13	Cross Drainage Structure	Yes = 1 No = 0
14	Protection work	Yes = 1 No = 0
15	Duration of Work	In years
16	Cost of Project	In crore

After creating array split the Training and Test Datasets in the ration of 0.60% and 0.40% respectively. The state of the preparation set (67 perceptions of 15 factors) and test set (46 perceptions of 15 factors).

After the split the data creating the Architecture of model built model with sequential creator 15 input dimensions with (ReLu) Activation function for passing one layer to another layer 3 hidden layer and one Output layer with one Out put dimension all the neurons are inter connected with each other which is done by using Dense module. The subsequent stage is to characterize an enhancer and the misfortune measure for preparing. The mean squared blunder is our misfortune measure and the "Adam" enhancer is our minimization calculation. The primary preferred position of the "Adam" analyzer is that we don't have to indicate the learning rate, similar to the case with angle plunge to Fits the model on the preparation dataset likewise give the contention, ages, which speaks to the quantity of preparing cycles which taken 2000 ages.

Table XIII Architecture of model

Model Type	Multilayer Perceptron (DNN)
Transfer Function	ReLu
Update Methods	Batch
Optimizer	Adam
No. of hidden layer	3
No. of PEs in the input layer	15
No. of PEs in the output layer	1

D. Experimental result and Analysis

Anticipate on the Test Data and Compute Evaluation Metrics of the Model Performance dependent on the Number of concealed layers, enactment capacity, and number of ages and learning rate has sway on the exhibition of neural organization. Different artificial neural network and multilayer preceptor models are develop by varying boundaries: No of concealed layers, No of shrouded hubs and actuation work.

This part presents the planned model, outcome and discussion. The projected model for prediction of cost estimation difficulty trouble utilizing NN is executed in Python adaptation 3.7 utilizing the Tensor-Flow system with Keras library. Keras is the library utilized to create profound learning show with Tensor-Flow as backend. Perfect way to assess the figure show is compare its genuine and anticipated comes about with the taking after Conditions.

$$MSE = \frac{\sum_{j=0}^p \sum_{i=0}^N (d_{ij} - y_{ij})^2}{NP} \quad (1)$$

$$MAE = \frac{\sum_{j=0}^p \sum_{i=0}^N |dy_{ij} - dd_{ij}|}{NP} \quad (2)$$

$$MAPE = \frac{100}{N} \sum_{j=0}^P \sum_{i=0}^N \frac{|dy_{ij} - dd_{ij}|}{dd_{ij}} \quad (3)$$

$$\begin{aligned} \text{Total MAPE} \\ = \frac{(MAPE_{Tr} \times N_{Tr} + MAPE_{C.V} \times N_{C.V}) / (N_{Tr} + N_{C.V}) + MAPE_{Test}}{2} \end{aligned} \quad (4)$$

$$TAP = 100 - \text{Total MAPE} \quad (5)$$

Table XIV Model Result

	Data set	RMSE	MAPE	TAP
DNN	Training	3.73	2.3%	63.70%
	Testing	274.40	70.3%	

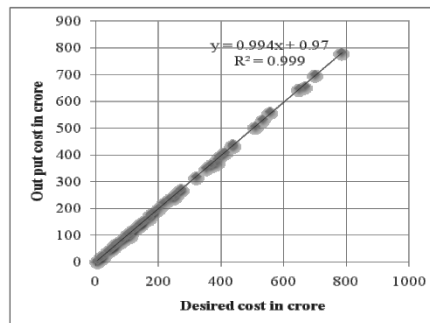


Chart 2 Desired output and actual network output for training set

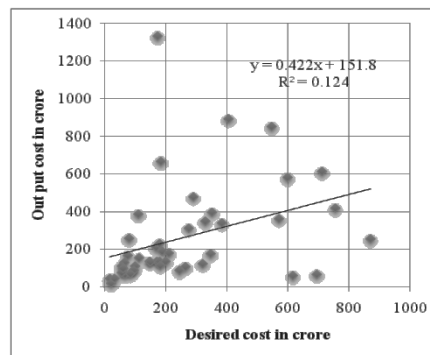


Chart 3 Desired output and actual network output for test set

We have assembled Regression models utilizing the profound learning structure, Keras. The guide utilized the Gujarat Road development cost information and fabricated a profound learning relapse model to anticipate the Cost in development in Crores.

Our model is accomplishing a steady exhibition with very little fluctuation in the train and test set RMSE. The ideal outcome would be a RMSE estimation of zero, yet that is practically unimaginable in genuine Cost datasets. Additionally, since the unit of the objective variable is in Crores, that likewise influences the RMSE esteem. There are different emphases, for example, changing the quantity of neurons, adding more concealed layers, or expanding the quantity of ages, which can be given a shot to see the effect on model execution.

Proposed the error of price estimation at the initial stage is more or less between ± 25 and ± 50 . During this study was projected error categorization supported on MMRE as shown in Table VII. Agreeing to the table one can select expectation precision level and diminish time for preparing and selecting network parameters. High prediction accuracy needs longer to coach the network and look

for advanced ANN models.[9]

Table XV Error Categorization in Percentage

MMRE	<i>Good</i>	Less than 25%
	<i>Fair</i>	25% – 50%
	<i>Poor</i>	More than 50%

V. Conclusion

Cost estimation done in the hypothetical stage of the extend grouping as a rule calculated generally which leads to huge incorrectness. The thought process for that has to have rugged thought of a extend in money related conditions. On the other side, venture subtle elements are not settled, well characterized due to need of data. In this paper, an attempt was made to prove that cost estimation inaccuracy at the initial planning phase. In conclusion, DNN might be a fitting instrument to assist unravels issues, which originate from various weaknesses, for example, taken cost assessment at the calculated stage with the precision of 63.70% to desire. Future work will be centered on creating a visual fundamental interface was created to encourage information passage for clients, where no past information of ANN is required to function DNN demonstrate. The interface given the client with numerous choices agreeing to the fifteen input parameters that depict the extend.

References

- [1] Cheng, M. Y., Tsai, H. C., & Sudjono, E. (2009). Evolutionary Fuzzy Hybrid Neural Network For Conceptual Cost Estimates In Construction Projects. 2009 26Th International Symposium On Automation And Robotics In Construction, Isarc 2009, 37(6), 512–519. <https://doi.org/10.1016/j.eswa.2009.11.080>
- [2] Hyari, K. H., Al-Daraiseh, A., & El-Mashaleh, M. (2016). Conceptual Cost Estimation Model For Engineering Services In Public Construction Projects. *Journal of Management In Engineering*, 32(1), 1–9. [https://doi.org/10.1061/\(asce\)me.1943-5479.0000381](https://doi.org/10.1061/(asce)me.1943-5479.0000381)
- [3] ICE (2019). Reducing The Gap Between Cost Estimates And Outturns For Major Infrastructure Projects And Programmes. *Institute of Civil Engineering*, 1–16
- [4] India Brand Equity Organisation (Ibef) <https://www.ibef.org/industry/infrastructure-sector-india.aspx>
- [5] Magdum, S. K., & Adamuthe, A. C. (2017). Construction Cost Prediction Using Neural Networks. *Ictact Journal On Soft Computing*, 8(1), 1549–1556. <https://doi.org/10.21917/ijsc.2017.0216>
- [6] Ministry of Statistics and Programme Implementation <http://www.mospi.gov.in>.
- [7] Niraj Suneja, Jagruti Shah, Zankhana Shah (2020), A Review of Cost Estimation: Methods and Models for Transport Infrastructure. *Tathapi*, 19(8), 681-690. <https://www.tathapi.com/index.php/2320-0693>
- [8] Sapp, C. E. (2017). Preparing And Architecting For Machine Learning. *Gartner Technical Professional Advice*, (January), 1–37. <https://doi.org/G00317328>
- [9] Schexnayder, C. J. and Mayo, Richard E., (2003) *Construction Management Fundamentals*, McGraw-Hill Higher Education, Boston, MA.
- [10] Venkata Suresh Addepali, M. Y. (2017). Integrating Risk Management And Cost Management to Arrive At A Realistic “Estimate At Completion”. *Project Managenet National Conference, India*, 1–13. Retrieved From www.pmi.org.in/conference
- [11] Waikar, A. S., & Kulkarn, T. V. (2019). Implementation of Artificial Neural Networks to

Forecast The Cost of Infrastructure Projects : A Review. *International Journal of Research In Advent Technology, Special Issue, Icimce 2019*, 138–142.

[12] Wilmot, C. G., & Mei, B. (2005). Neural Network Modeling of Highway Construction Costs. *Journal of Construction Engineering And Management*, 131(7), 765–771. [https://doi.org/10.1061/\(ASCE\)0733-9364\(2005\)131:7\(765\)](https://doi.org/10.1061/(ASCE)0733-9364(2005)131:7(765)).