JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Construction Progress Prediction and Cost and Time Analysis

Harshwardhan Atkare
Department of Computer Science
Pune Institute of Computer Technology
Pune, India

Atharv Bhadange
Department of Computer Science
Pune Institute of Computer Technology
Pune, India

Raj Shende
Department of Computer Science
Pune Institute of Computer Technology
Pune, India

Prajwal Toundakar
Department of Computer Science
Pune Institute of Computer Technology
Pune, India

Abstract – This research employed artificial neural networks (ANNs) to analyze the intricate relationships between various input factors and the cost and duration of construction projects. The dataset for the study was generated based on six input factors: Project Type, Project Size, Location, Complexity, Number of Floors, Start Date, End_Date, Construction Cost, Material Cost, Labor Cost, Equipment Cost, Permit Cost, Contingency Cost, Overhead Cost, Number of Workers, Weather Conditions, Previous Project Completion Time, Previous Project Cost, and two output parameters (duration and cost). These factors were identified through a thorough review of relevant literature, expert opinions, and an extensive field survey. The research revealed that insufficient technological advancements and a lack of skilled personnel for effective project coordination contribute to cost overruns and extended durations, particularly in projects awarded to sole and mini-contractors compared to medium and multi-company firms. Clients with greater financial capacity, such as corporate organizations and government entities, tended to engage and negotiate with multi- and medium-sized businesses.

Keyword – Construction management · Project cost and duration · Artificial neural network (ANN) · MATLAB

Ⅰ. INTRODUCTION

A nation's level of development is often gauged by the state of its infrastructure assets, which play a crucial role in the progress and prosperity of society. Infrastructure development significantly influences various aspects of societal advancement (Elmousalami, 2020)[1]. The challenges of time and cost instability are persistent hurdles in construction operations, stemming from the unique nature of each building project. Unforeseen uncertainties make it challenging to adhere to project timelines and budgets. The inherent variability in construction projects arises from diverse factors such as location, clients, regulations, labour, technology, equipment, subcontractors, experience, stakeholders, and project teams (Chudley & Greeno, 2016)[2]. Accurate cost estimation is paramount for securing a project's financial viability, and the use of effective models is crucial for precise cost calculations and construction durability forecasts in the industry.

Despite advancements in project management, completing projects on time and within budget remains a complex task. Project complexity often leads to cost overruns, attributed to various factors like labour-related delays, delays from clients and contractors, and consultant-related delays.

Ⅱ. LITERATURE SURVEY

Early in the project, it becomes evident that obtaining an accurate estimation of construction costs and duration is imperative. The significance of precision in these estimates lies in the fact that any deviation, whether it be an underestimation or overestimation, can result in cost overruns and negatively impact project performance. This is particularly manifested in the project's failure to meet its quality standards and delivery schedule objectives.

Soft computing techniques are considered apt for modelling time-cost constraints in construction projects. This suitability stems from the inherent non-specific patterns in building projects, arising from environmental and logistics factors, as well as the non-linear and discrete dependencies involved (Wang et al., 2016)[3][4]. Analytical approaches are chosen due to the perceived lack of alignment between costs and activities, given the scarcity of representational data for investigating the relationship between activity duration and costs.

In situations where there is either an insufficient quantity or quality of data, the use of computer-induced systems becomes crucial. Construction simulation, facilitated by these systems, entails designing construction activities to better discern their behavioural patterns. This, in turn, contributes to the development of intelligent decision-making models. These models are focused on understanding the tasks associated with a project and identifying the requisite materials for their completion.

Artificial neural network

In contrast to traditional rule-based artificial intelligence methods, neural networks autonomously extract expert details from data without the need for predefined rules. Essentially, the system learns to become an "expert" in a given field through a trial-and-error approach. Artificial Neural Networks (ANN) serve as an information processing framework inspired by how the brain and nervous system of living organisms process information. Comprising a vast number of interconnected processing elements (neurons), these networks collaborate systematically to find solutions to specific problems.

Biological neurons, also known as nerve cells, are fundamental components of the brain and nervous system. They receive sensory input through dendrites, process it within the cell body (Soma), and transmit output through axons. Dendrites, branching out in a tree-like structure around the cell body, receive incoming signals, while axons function as long, thin, tubular structures similar to transmission lines. Neurons possess hair-like tube-shaped extensions known as (Alaneme et al., 2022a, 2022b;)[5].

Neurons form complex spatial arrangements, and as the axon reaches its destination, it transforms into a nerve fibre that carries impulses away from the cell body. Synapses, intricate structures at the tip of the axon, facilitate the connection between neurons. Interactions between neurons occur at these synapses, where dendrites receive input from other neurons. The soma progressively processes incoming signals and transforms the processed value into an output sent to other neurons through axons and synapses, as illustrated in Fig. 1 (Ujong, Jesam & Mbadike, Elvis & Alaneme, George,2022)[6].

The generated output signal becomes the input for the subsequent layer within the stack. The soma systematically processes incoming signals, transforming the processed value into an output dispatched to other neurons through the axon and synapses, as depicted in Fig. 1(Ujong, Jesam & Mbadike, Elvis & Alaneme, George, 2022)[6]. The foundational perceptron, introduced in 1943 by McCulloch and Pitts, emulates the functioning pattern of a living neuron.

A Perceptron, illustrated in Fig. 2, refers to a single-layer neural network with a lone output. It aggregates bias and inputs, each multiplied by their respective weights, making decisions based on the accumulated results. For a single observation, x0, x1, x2, x3…x(n) represent various input parameters to the network, and their weights are denoted as w0, w1, w2, w3….w(n) (Alaneme et al., 2021b;)[5]. The bias value, represented as 'b,' enables the adjustment of the activation function's position, preventing plots from originating solely at the origin.

In the simplest scenario, the products are summed, and fed into a transfer function (activation function), generating an output forwarded as a result. Mathematically, the summation/transfer function is presented in Eq. 1 (MT Alqershy, R Kishore 2023 - Taylor & Francis)[7]. The activation function is crucial for an Artificial Neural Network (ANN) to learn and comprehend complex information. It decides whether a neuron should activate by summing the weighted sum and adding bias. This introduces non-linearity into the neuron's output, preventing it from being a simple linear function (a first-degree polynomial). Without an activation function, the model cannot effectively learn and model intricate data (Alaneme et al., 2020a, 2020b)[5].

Fig. 2 Perceptron

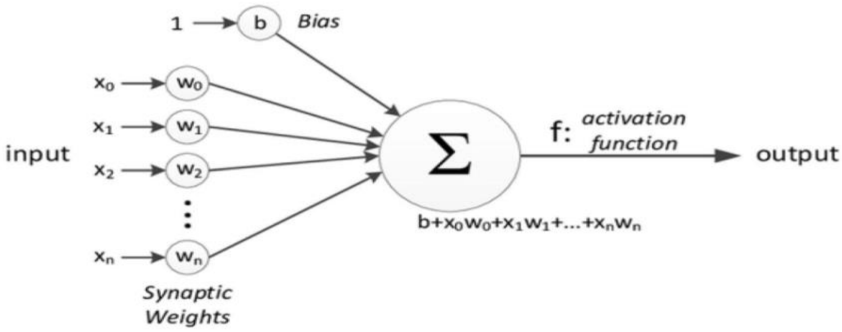
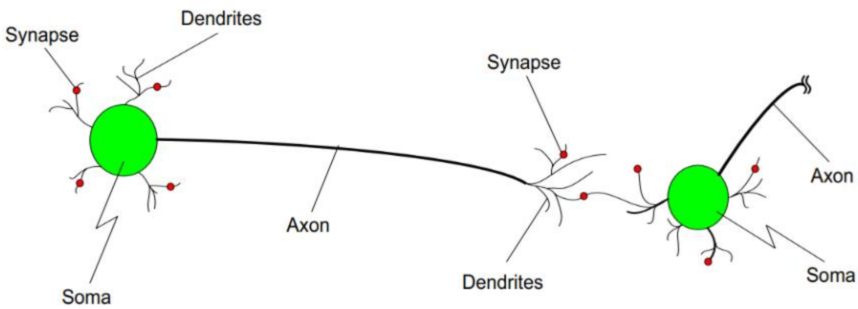


Fig. 1 Biological neuron



III. REAL-WORLD APPLICATION OF MACHINE LEARNING IN CONSTRUCTION COST PREDICTION.

A. Property Valuation

*I*n revolutionizing decision-making processes within construction cost prediction, the amalgamation of data analysis and machine learning brings about a profound transformation. Machine learning algorithms meticulously construct predictive models, taking into account variables like project dimensions, materials, labour, and prevailing market trends, thereby enhancing the precision of budgeting efforts. The utilization of data analysis techniques, such as cost sensitivity analysis, optimizes key cost elements with remarkable efficiency. The implementation of real-time expense monitoring, driven by machine learning capabilities, swiftly identifies instances of cost overruns, enabling prompt and effective interventions. Furthermore, the automation of change order analysis, facilitated by machine learning-powered natural language processing, significantly reduces manual workloads. Predictive models play a pivotal role in recognizing and mitigating risks and uncertainties, providing invaluable support for proactive risk management. The machine learning-driven optimization of resource allocation serves to minimize wastage, marking a pivotal advancement in the field.

B. Investment Decisions:

In the context of making investment choices within the construction industry, the integration of machine learning and data analysis emerges as a pivotal factor, furnishing valuable insights to guide well-informed decisions. Through the scrutiny of historical project data and the application of predictive models, algorithms driven by data delve into the trends of construction project progress and costs. This analytical prowess equips investors with the tools to formulate strategic decisions. Machine learning algorithms, encompassing regression and time series analysis, go a step further by forecasting future construction progress and estimating costs based on a multitude of parameters, ensuring a level of precision in predictions. The incorporation of real-time data sourced from construction sites refines these predictions, taking into account factors like material consumption, labour efficiency, and equipment performance. By leveraging these cutting-edge technologies, investors acquire a holistic comprehension of project risks and potential returns, thereby optimizing resource allocation, mitigating financial risks, and facilitating judicious investment decisions within the construction sector.

C. Market Analysis

In the domain of market analysis for construction projects, the fusion of machine learning and data analysis techniques transforms decision-making processes. Machine learning algorithms efficiently process extensive datasets, encompassing historical project costs, regional economic indicators, and market trends, facilitating predictive modelling for construction costs and market demands. Data analysis, complemented by machine learning, provides profound insights into customer preferences, vendor behaviours, and market fluctuations. Through the application of predictive analytics, construction companies can anticipate *future market* demands, optimize pricing strategies, and make real-time data-driven decisions. This synergy between machine learning and data analysis not only enhances the accuracy of cost estimations but also provides a comprehensive understanding of market dynamics, enabling construction businesses to remain competitive and adaptive in a swiftly changing market landscape.

IV. CHALLENGES AND FUTURE DIRECTIONS IN CONSTRUCTION COST PREDICTION.

A. Challenges in construction cost prediction

In the realm of estimating construction project costs and predicting progress, a major hurdle arises from the varied and varying nature of the data at our disposal. Construction endeavours produce copious data from disparate origins like sensors, project documents, and historical archives. The veracity and uniformity of this data become paramount in establishing dependable forecasts. Moreover, the amalgamation of heterogeneous data types—ranging from financial statistics and project timelines to real-time progress updates—poses a formidable challenge. These datasets arrive in diverse formats and structures, hindering the creation of a cohesive dataset for thorough analysis. Overcoming these obstacles necessitates the deployment of robust data validation methodologies, adept data cleansing techniques, and the statement of standardized data formats to streamline integration and analytical processes.

The complexity inherent in construction projects adds another layer of difficulty to the prediction process. Construction endeavours entail a multitude of interwoven elements, including project specifications, environmental nuances, labour intricacies, and material availability. Navigating these intricate connections poses a formidable challenge in crafting precise predictive models. Furthermore, unforeseen occurrences like weather disruptions, alterations in regulations, or disturbances in the supply chain have the potential to significantly sway project timelines and expenses. Effectively incorporating and accommodating these intricacies into predictive models demands the utilization of sophisticated machine-learning algorithms capable of handling nuanced patterns and diverse variables. Leveraging advanced methodologies, such as deep learning, becomes imperative for capturing nonlinear relationships within the data, ultimately elevating the precision of predictions.

The exploration of construction project data is not without its ethical and privacy intricacies. Sensitive details concerning contracts, financial transactions, and employee information often intertwine with these projects. It becomes imperative to navigate these complexities to both foster trust among stakeholders and adhere to stringent data protection regulations. Achieving an equilibrium between leveraging data for predictive insights and upholding the privacy rights of individuals involves the application of encryption, secure data transmission protocols, and techniques for anonymization. Furthermore, an essential component in this ethical journey involves imparting knowledge on ethical data usage practices to all stakeholders engaged in data analysis, thereby upholding ethical standards within the industry.

B. Future Directions

In the foreseeable future, the progression of construction project cost estimation and progress prediction is poised to hinge on the amalgamation of cutting-edge technologies and innovative methodologies. The continuous evolution of artificial intelligence and machine learning algorithms is anticipated to facilitate more precise predictions, adept at processing extensive and varied datasets. Predictive analytics is slated to transcend conventional forecasting, delving into prescriptive analytics territory, thereby providing actionable insights for informed decision-making. The cultivation of collaborative platforms and the establishment of industry-wide norms for data formats are expected to nurture a culture of knowledge exchange, fostering a collective intelligence approach to the realm of construction project management.

Table Ⅰ. A SUMMARY OF RESEARCH REVIEWED IN SECTION Ⅱ

Title	Algorithm(s) used	Key Observation(s)	Application(s)	Author(s)
Predicting the Cost and Duration of Building Construction Projects Using Artificial Neural Networks: A Case Study of Cross River State, Nigeria	Artificial Neural Networks (ANNs), Multiple Linear Regression (MLR)	- Building information details were observed to influence the project construction duration and cost - ANNs outperformed MLR in predicting construction project cost and duration - The study enhances productive decision-making during project planning	1. Projected estimations of construction project expenses and timelines utilizing artificial neural networks. 2. Assessment of the mechanical properties of expansive soil treated with hydrated-lime-activated rice husk ash employing gene expression programming. 3. Uncovering the origins and consequences of project construction delivery time delays. 4. Streamlining project expenses and timelines while enhancing holistic performance and efficiency. 5. Augmenting effective decision-making in the course of project planning processes.	Wang et al. [4]
Home Construction Cost Estimation Using ML	Random Forest Gradient Boosting Artificial neural networks Linear regression	- ML models generate more accurate predictions compared to traditional methods - Random Forest achieved the highest prediction accuracy among the tested methods - Effective data pre-processing and feature engineering techniques can improve the accuracy of the models - Further research is needed to address challenges related to data quality and model interpretability - Analysis of the five algorithms, conducted with varying training and testing sizes	- Home cost estimation - Real estate industry	-B. Balakumar -P. Raviraj [9]
Using Machine Learning to Predict Cost Overruns in	- Random Forest (RF) -Support Vector Regression (SVR) -Artificial Neural Network (ANN)	- The RF algorithm outperformed the other two algorithms in predicting cost overruns. - The study found that the most significant factors contributing to cost	- Accurately predicting cost overruns in construction projects can help project managers make informed decisions	-Abolfazl Jafari -Iman Pourhaji -Pete

Construction Projects: A Case Study of Large-Scale Projects in Malaysia		overruns were changes in project scope, design changes, and delays in project completion.	and take corrective actions to prevent or mitigate cost overruns. - The machine learning models developed in this study can be used as a tool for cost estimation and risk management in construction projects.	Burnap [10]
A Systematic Review of Machine Learning Algorithms for Construction Cost Estimation	-Artificial neural networks (ANNs) -Support vector machines (SVMs) - Decision trees - Random forests - Bayesian networks -Deep learning algorithms (e.g. convolutional neural networks, recurrent neural networks)	- The paper notes that machine learning algorithms have shown promise in improving the accuracy of construction cost estimation, particularly when compared to traditional methods. - However, the authors also caution that there are still challenges to be addressed, such as the need for high-quality data and the potential for bias in algorithmic decision-making. - The paper also highlights the importance of considering factors beyond just cost, such as project safety and environmental impact.	- The paper suggests that machine learning algorithms could be used in a variety of construction-related applications, including: - Cost estimation for individual projects - Predictive maintenance for construction equipment - Quality control and defect detection - Risk assessment and management	-Yasamin Ghadbhan Abed -Taha Mohammed Hasan -Raquim Nihad Zehawi

VII. CONCLUSION

The survey investigates the extensive body of research on integrating data analysis and machine learning in construction project cost estimation and progress prediction, marking a significant advancement for the industry. Overcoming challenges related to data quality, project complexity, and ethical considerations is key to achieving more precise predictions and informed decision-making. The future of construction projects relies on combining predictive and prescriptive analytics through advanced AI algorithms, enabling proactive responses to project dynamics. Utilizing real-time data from IoT devices alongside collaborative platforms is ushering in a new era of agile project management. Embracing these advancements will lead to more efficient, transparent, and adaptable construction projects, driving a shift towards data-driven excellence in construction management.

REFERENCES

[1] Elmousalami, H. H. (2020). Artificial intelligence and parametric construction cost estimate modelling: State-of-the-art review. Journal of Construction Engineering and Management, 146(1), 03119008.

[2] Chudley, R., & Greeno, R. (2016). Building Construction Handbook. Abingdon: Routledge.

[3] W. T. Lim, L. Wang, Y. Wang, and Q. Chang, "Housing price prediction using neural networks," in 2016 12th International Conference on natural computation, fuzzy systems and knowledge discovery (ICNC-FSKD), 2016, pp. 518-522.

[4] Wang, W.-C., Bilozerov, T., Dzeng, R.-J., Hsiao, F.-Y., & Wang, K.-C. (2017). Conceptual cost estimations using neuro-fuzzy and multifactor evaluation methods for building projects. Journal of Civil Engineering and Management, 23(1), 1–14.

[5] Ujong, Jesam & Mbadike, George. (2022). Prediction of cost and duration of building construction using artificial neural network. Asian Journal of Civil Engineering.

[6] Elvis & Alaneme, George. (2022). Prediction of cost and duration of building construction using artificial neural network. Asian Journal of Civil Engineering.

[7] MT Alqershy, R Kishore - International Journal of Construction …, 2023 - Taylor & Francis

[8] R. E. Febrita, A. N. Alfiyatin, H. Taufiq, and W. F. Mahmudy, "Data-driven fuzzy rule extraction for housing price prediction in Malang, East Java," in 2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2017.

[9] B. Balakumar, P.Raviraj, S.Essakkiammal, "Predicting Housing Prices using Machine Learning Techniques".

[10] Breiman L. Random Forests. SpringerLink. https://doi.org/10.1023/A:1010933404324 (accessed September 11, 2019)

[11] Tayefeh Hashemi, S., Ebadati, O.M. & Kaur, H. Cost estimation and prediction in construction projects: a systematic review on machine learning techniques. SN Appl. Sci. 2, 1703 (2020).

[12] Yu, W., & Skibniewski, M. J. (2010). Integrating a neurofuzzy system with conceptual cost estimation to discover cost-related knowledge from residential construction projects. Journal of Computing in Civil Engineering, 24(1), 35–44.

[13] Lowe, D., Emsley, M., & Harding, A. (2006). Predicting construction costs using multiple regression techniques. ASCE Journal of Construction Engineering and Management, 132(7), 750–758

[14] Remon, F. A., Sherif, M. H., & Yasser, R. A. (2014). Smart optimization for mega construction projects using artificial intelligence. Alexandria Engineering Journal, 53(3), 591–606.

[15] Feng, G. L., & Li, L. (2013). Application of genetic algorithm and neural network in construction cost estimate. Advanced Materials Research, 756–759, 3194–3198.

[16] Khosrowshahi, F., & Kaka, A. P. (1996). Estimation of project total cost and duration for housing projects in the UK. Building and Environment, 31(4), 375–383.

[17] Samuel, E. I., & Snapp, O. J. (2015). Cost model for unit rate pricing of concrete in construction projects. International Journal of Construction Engineering and Management, 4, 149–158.

[18] Thomas, N., Mak, M. Y., Skitmore, M., Lam, K. C., & Varnam, M. (2001). The predictive ability of Bromilow's time–cost model. Construction Management and Economics, 19(2), 165173.

[19] AbouRizk, S., & Halpin, D. (1992). Statistical properties of construction data. Journal of Construction Engineering and Management, 118, 525–544.