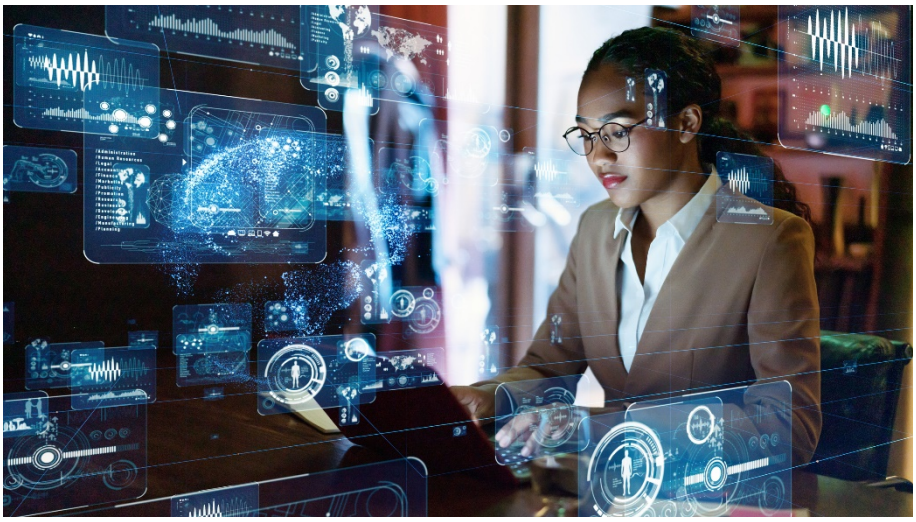
Machine Learning in Construction Cost Estimating: Progressing from the Mean

Written by: Mostafa Khard, Project Manager at Hill International



Pre-design cost estimates are critical to the success of any major construction project. An accurate initial cost estimate supports project managers and their clients in making informed decisions during the early phases of a project, impacting work across the entire project lifecycle.

During pre-design, designers and project managers use several tools and techniques to arrive at reasonable cost estimates. These methods include their own intuitive judgement, their experience, and libraries of data from projects with similar materials, locales, and characteristics. Combined with such standard cost estimating methods as detailed breakdowns, simplified breakdowns, activity costs, unit costs, and cost-per-square-foot estimates, pre-design estimates can provide realistic and viable estimates.

However, during the early planning and design stages, often only the most basic and functional project decisions have been made, and it may not be clear what specific data is needed to develop an estimate. For example, integrating client preferences, designer and contractor characteristics, procurement methods, and stakeholder involvement are difficult, if not impossible, to define at the conceptual stage. Similarly, for particularly ambitious or novel projects, the data needed to generate estimates may be incomplete or simply unavailable. This lack of early cost clarity can impede projects once work is underway and the true costs of a project become apparent. The result, unfortunately, is often overruns and delays.

Machine learning offers a potential solution for teams looking to improve their early stage estimates. Machine learning, simply put, is a set of algorithms that groups data, learns from this data, and then applies what it has learned to make intelligent decisions. Outside of the construction industry, machine learning is used by Netflix to suggest programs you may want to watch and by your email to identify and segregate spam messages. In the field of predictive modeling, which is most applicable to cost estimating, machine learning focuses on making the most accurate predictions possible.

While machine learning can deliver these predictions with impressive precision, it does so at the expense of *explainability*, which is the extent the internal mechanisms of the algorithm set can be explained. In short, machine learning can deliver a reliable cost estimate, but we don't always understand how. In a relatively traditional industry like construction, this lack of explainability may be

why machine learning is ubiquitous in everything from air travel to online advertising but is still a new tool for many construction cost estimators.

Nevertheless, machine learning offers several models for construction estimators to consider. One benefit is linear regression. Linear regression is a statistical tool estimators use to arrive a linear total cost function for a set of costs, based on past cost data. The estimate can then employ the function to predict total costs for an activity, such as per unit or per hour. Calculating someone's salary based on their total years of experience is one common example of a linear regression. This model is simple to build, but to date has had only limited use in construction estimating as it represents only a linear relationship between variables. That said, linear regression can still be helpful in predicting future rates such the costs of excavation, concrete, and steel, which are published annually. For such indices, a linear regression model can be an effective tool.

An artificial neural network (ANN) is another type of machine learning algorithm often used in prediction models. Each ANN is essentially a computing system of interconnected nodes that mimics human brain function. The network clusters large, raw data sets and, by identifying patterns, then solves very complex problems, classifies inputs, and makes difficult predictions. Researchers have implemented a neural network approach in various engineering fields with some success, and this same technique is worth studying in the construction industry as well, especially on complex mega-projects with many "known unknowns."

ANN could potentially be the most appropriate application of machine learning to early stage cost estimating for several reasons. First, unlike linear regression, ANNs can model interdependencies between input data, which will inevitably occur given the significant variables in construction, such as the number of stories, floor area, and number of elevators, to name only a few. Neural networks can deal more readily with non-linear relationships between these and other variables. In addition, neural networks can more effectively handle incomplete data sets, which are likely during the conceptual phase of the project. These benefits give ANN a significant advantage over other predictive models in creating meaningful estimates early phases of a project.

Beyond ANNs, research has also suggested other algorithms to predict costs, such as multiple regression analysis, case-based reasoning, fuzzy logic, and rough sets theory. In fact, researchers are working to combine these methods. One study combined ANNs, multi-factor evaluation, and fuzzy logic for construction cost estimating. The combination of these three logics delivered a more reliable cost figure than each one of them individually.

Confidence in cost prediction is one of highest determinants of overall project success, and improving accuracy and reliability in construction estimates, especially during the critical early planning phases of a project, is invaluable as owners look for greater precision and accountability from their teams. Machine learning offers a way to focus and prioritize efforts to control costs when decisions can have the greatest impact. Construction managers and owners should experiment with these new techniques to identify the best ways to predict the costs for their own unique projects.

###

About the Author:

Mostafa Khard, LEED AP, a CMAA member and project manager at Hill International, Inc., has more than five years of experience in the construction industry. Specializing in construction program management, Mostafa brings a blend of skills in document controls, quality assurance inspections, architecture, sustainable design, building information modeling, and cost estimating. Mostafa's project experience includes new construction, additions and renovations for corporate office, education, government, transportation, and medical clients in the U.S. and abroad.

Hill International, Inc., is a leading PM/CM firm headquartered in Philadelphia, Penn.