# Machine Learning Based Automated Construction Planning System for Sri Lanka

Project ID: 24-25-201

Project Proposal Report

**Ahamed. R.A – IT21158018**

B.Sc. (Hons) Degree in Information Technology Specializing in Information Technology

Department of Information Technology
Sri Lanka Institute of Information Technology
Sri Lanka

August 2024

# Machine Learning Based Automated Construction Planning System for Sri Lanka

Project ID: 2024-201

Project Proposal Report

**Ahamed. R.A – IT21158018**

Supervised by Mr. N.H.P. Ravi Supunya Swarnakantha

Co-supervised by Dr. Dharshana Kasthurirathna

B.Sc. (Hons) Degree in Information Technology Specializing in Information Technology

Department of Information Technology

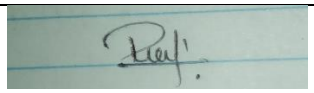
Sri Lanka Institute of Information Technology

Sri Lanka

August 2024

## DECLARATION OF THE CANDIDATE AND SUPERVISOR

I certify that this submission is my own work, and, to the best of my knowledge, it contains no material previously published or written by another person (except where explicitly defined in due acknowledgment statements), nor materials which substantially have been accepted for the award of any other degree or diploma of an educational institution. I also hereby given to the Sri Lanka Institute of Information Technology exclusive rights including but not limited to: /a) To reproduce, publish and distribute my dissertation in whole or part upon any nationalization methods now known/releasing media forms. I only have the right to use this work in whole or edited format in upcoming works (articles, books or talks).

| NAME | STUDENT ID | SIGNATURE |
|---|---|---|
| R.A. AHAMED | IT21158018 |  |

Name of supervisor: Mr. Supunya Swarnakantha

Name of co-supervisor: Dr. Dharshana Kasthurirathna

The above candidate has carried out research for the bachelor's degree dissertation under my supervision.

Signature: Date: 8/22/2024
Supervisor: Mr. Supunya Swarnakantha

Signature: Date: 8/22/2024
Co-Supervisor: Dr. Dharshana Kasthurirathna

## ABSTRACT

Construction well-known to face significant challenges in cost estimating a key factor towards project overruns and delays! Conventional approaches to estimation use static data and are not designed for the dynamic nature of ongoing projects, often leading to large financial discrepancies; these can affect everyone involved. This solves a critical problem that many people want to solve, the accuracy of materials used in construction projects by providing accurate and real-time estimates. This project introduces a new machine learning system for calculating costs endorsing this changeouts patch there from cost. Utilizing a blend of historical data and real-time feeds, the process leverages advanced machine learning algorithms to predict cost estimates on-the-fly as project conditions change. It also ensures estimating is based on both up-to-date market conditions and variables that are unique to the project, which in turn greatly improves forecast accuracy.

The machine learning system continuously learns and adjusts, rather than relying on an outdated static model; it continually uses new input to make better predictions. This capability enables it to be made in real time, giving project managers and stakeholders a constantly updated overview of potential costs which is very important for managing your budget and ensuring optimal execution of the construction. The introduction of the system has shown a great improvement in estimation accuracy when compared to standard methods, confirming its usefulness for the construction process. Delivering accurate, timely financial data that helps in more efficient project planning and resource allocation, it materially mitigates the risk of exceeding budgeted amounts while also fostering better practice around overall project management. Its flexibility and scalability can address a wide range of construction project types, paving the way for consistent pricing process across the industry as well as promoting best practice solution adoption in sustainable modalities.

Key Words – Automated Construction Planning, Machine Learning, Cost Estimation, Real-Time Data Integration, Construction Project Management, Predictive Analytics, Budget Management, Real-Time Updates, Historical Data, Machine Learning Models, Risk Mitigation, Project Scheduling, Data Quality, Real-Time Cost Predictions, Data Sources, Construction Industry Challenges,

- **Construction Cost Estimation**
- **Machine Learning**
- **Real-Time Data**
- **Material Cost Analysis**
- **Predictive Modeling**
- **Artificial Neural Networks (ANN)**
- **IoT in Construction**
- **Project Budgeting**
- **Construction Management**
- **Dynamic Cost Estimation**
- **Data-Driven Decision Making**
- **Building Information Modeling (BIM)**
- **Cost Overruns**
- **Project Planning**
- **Regression Analysis**
- **Smart Construction**
- **Digital Construction Tools**
- **Cost Efficiency**
- **Construction Automation**
- **Scalable Estimation Models**

## ACKNOWLEDGMENT

Without the constant support of our supervisors, Mr. Supunya Swarnakanth and Dr. Dharshana Kasthurirathan, Instructor, Department of computing, Sri Lanka Institute of Information Technology, I would not have

been able to complete this research. By offering essential comments and helping us through our blunders, they have helped me make my project a success and generate more efficient outcomes. I'd also like to thank our panel members for giving us invaluable advice in helping me comprehend the research's flaws and proposing ways to improve the research's functionality.

I'd also like to express our gratitude to coordinators, the course's comprehensive design and analysis lecturer, for giving us the required assistance throughout this project module enabling me to address the flaws we encountered effectively. Finally, I'd like to express my appreciation to everyone who helped me complete this study, whether directly or indirectly. Your invaluable help and suggestions inspired me to achieve this task

## TABLE OF CONTENTS

# 1. INTRODUCTION

The construction industry struggles to predict costs effectively because of fluctuating material prices, labor rates and job specifications. Such traditional methods have static data, so they are unable to keep an accurate value causing cost and inefficiency overruns. In this project, we have established a system that can help Architecture, Engineer and construction firms to prepare accurately budget or even getting approximate idea of material uses & estimating the cost on time by performing Machine Learning based predictions which is capable enough for fetching in real-time data suitable output. The suggested system will estimate reliable costs to construction professionals by using historical data and current market trend which were not available in olden times It leads to better planning of project.

## 1.1 Background

The construction sector occupies a significant place in the world economy, but many problems are associated with it, first of all, with cost estimation and, consequently, budgeting. Cost control is significant in construction projects due to prophetic and vital implications to budget and actual resource provisions and management. Previously, to assess the cost of construction, there were techniques such as manual methods, and expert judgments with simulation based on historical records and this techniques are not dynamic. These estimates are made formidable by other factors like varying costs of the raw materials, costs of employing labour and other related unknown additional project requirements.

The identification of different moving parts that cost estimation can go wrong can be seen below: Such issues are not only financially burdensome , but also have a negative impact organizational image and hinder the successful and efficient delivery of construction projects by a firm . With the increasing industry size of construction and the more complex projects, there is a need to develop advance methods that will enable estimators to come up with accurate, real-time estimates.

Continuous advancements in technology especially the ‘machine learning’ offers a chance to enhance cost estimation in construction. Machine learning opens the pathways for increased computing capability, capacity to make predictions and making of decisions from large datasets. Due to the use of machine learning algorithms it will be possible to create an adaptive model which can be improved over time as new data is received and thus afford construction managers more accurate cost estimates.

The idea behind this project is to extend the possibilities of current linear and linear-programming based estimations of construction costs as well as selection of appropriate

materials with the application of machine learning. It is the hope for the proposed system to incorporate real-time market information, as well as the history to give dynamic and flexible cost estimates which will be useful at every stage in construction professions.

### 1.2 Literature Review

Literature review on the incorporation of machine learning in construction cost estimation shows studies showing that incorporating machine learning improves the precision and speed of estimates. The two main conventional methods of cost estimating that have been the conventional methods in the construction industry include Parametric and Analogy estimates. But these methods have a significant drawback because they presuppose the use of fixed data and formulas that do not consider fluctuations in construction projects.

Artificial neural networks, support vector machines and decision trees are some of the models which have shown higher accuracy in construction cost estimation. There are these models as they can calculate more numerous and different sorts of coefficients such as material costs, labor costs of specific project phases, timeframes and other, environmental conditions are also considered to reach higher accuracy in cost estimates. As an example, in the study by Kim et al. (2019), it was noted that cost estimation by using ANN model was more accurate, and less fluctuating than conventional approaches in the context of high-rise building projects.[1]

The utilisation of real-time data has also been found to be a significant aspect influencing the enhancement of cost estimation reliability. Sources have indicated that integrating actual time data from IoT gadgets for instance; sensors that track way material is being used or the rate at which workers are producing output, is highly helpful in improving the performance of machine learning. Zhang et al. (2020) also showed the possibility of using IoT data in connection with machine learning to forecast the construction costs and improve construction management adaptability.[2]

However, there are still some challenges that can be encountered hence they include the following ones. The source of the data fed into a machine learning model also has to be the best as wrong or incomplete data will lead to wrong predictions. Another disadvantage of some models of machine learning is a high level of sophistication that may limit understanding of the model and its direct application in construction processes. Further research is also required as to the efficiency and effectiveness of such models especially in large construction projects that generate a huge amount of data.

Altogether, machine learning is a promising field in the construction context regarding cost estimate enhancement; however, it is not without challenges regarding data quality, model explanation, and data size. These challenges can be solved by creating a more complex system of machine learning that is combined with real-time data feed, which should give a more precise and real-time costs estimations for the construction industry.

### 1.3 Research Gap

Notwithstanding the encouraging new developments in the application of machine learning to construction cost estimation, we identify several significant research and practical gaps. Firstly, we identified a significant gap between recent findings from research and application to practical use in the construction industry. As highlighted in the literature, there is an urgent need for consideration of real-time data sources in machine learning modelling. However, while much of the literature has focused on demonstrating the potentially improved prediction accuracy that can be obtained if IoT devices and real-time market data are integrated with machine learning models, we are yet to see comprehensive systems that actually achieve this integration of data streams from these differing data sources into a unified environment. Such a system would enable construction managers to receive updated estimates of costs in construction as they work with data that reflects real-time construction markets and on-site realities.

A second significant gap concerns the lack of attention paid to various types of projects – from residential to commercial buildings – and a paucity of research exploring other project types (eg, infrastructure, renovation) that have distinctive characteristics, and hence different models and approaches.

Furthermore, while academic research suggests many opportunities to use machine learning models in the construction industry, in reality their field deployment has been hampered by model complexity and the lack of transparency in some popular approaches aspects that can potentially make them difficult to adopt by practitioners who value simpler, more interpretable techniques. These gaps highlight how research on machine learning can also benefit critically from efforts to address both the technical capabilities of the developed models, and their applicability in practice, as well as the acceptance of their results by the end-users.

Finally, some of us must turn to more technical and philosophical questions about how these technologies can be scaled and tailored to different kinds of construction projects, and how – once the lessons have been learned – the benefits of ML can be applied throughout the industry.[2]

## 1.4 Research Problem

The problem is that this research will focus on the real-time cost estimation for construction projects. Conventional techniques of cost estimation often fail to cope with the active nature of construction projects resulting in overrunning of budget, delays in project and misallocation or underutilization resource. Historically, these methods use static data — historical costs and standard pricing lists that do not accurately reflect changes in material prices or labor cost as your construction projects evolve. This can lead to large disparities between quoted and real costs, presenting a big drain on the bottom line for construction managers.

Integrating machine learning with real-time data might be a good answer to this. Machine learning models include complex algorithms with up to the minute information, making cost estimations more exact and timelier. Well, in theory a system like that sounds great but the practical implementation presents problems. This includes verifying quality, ensures that model accuracy remains relevant for the duration of time as well obtaining user acceptance from construction professionals who may or may not be familiar with advanced technologies and how they can help.

The goal of this research will be to overcome the above challenges with a real-time interactive construction cost and material estimation system, which uses machine learning for making quick predictions on costs. By taking into account historical projects data, real-time market trends and on-site sensor data amongst other sources available, the system will deliver trusted cost estimations across the project cycle. Furthermore, the proposal will investigate how the system can be adopted in a deployable manner within real-life construction contexts such that it is user-friendly and sustainable across various classes of construction activities.

Through challenging and avoiding these cases, the research attempts to be an addendum in promoting employment of cost estimations methods within construction at large for better project control without budget overruns like too many instances ending dire.

# 2. OBJECTIVE

## 2.1 Main Objective

The primary goal of this research is to make a flexible and strong construction cost estimation designed by using machine learning for real-time prediction with higher accuracy. Using historical data, real-time information and cutting-edge predictive models directly from the cloud, Cost Port + is a modern platform that seeks to expand traditional cost estimation options. It will also feature a system that can revise cost forecasts when additional data is released, equipping construction experts to make informed choices all through the project.

At the core of this goal is creating a simple entryway for users to get their project data into it, check live cost predictions and tweak things as needed. We are designing this interface for the end-user, and by doing so will make sure that a person with no technical ability whatsoever can operate in an easy and efficient manner. The system will also be designed to work with a wide range of other construction management tools (scheduling software, procurement platforms) in order for this solution to offer an end-to-end project planning and management suite.

These models are tested using rigorous testing and real-life research to verify the accuracy and reliability of these working systems. The research will investigate its ability to mitigate the increases in project costs relative to actual budget overruns and overall project outcomes. This system is designed to be able to absorb new data as it comes in and then adjust its predictions, a change that should lead to better cost estimates which when translated through will help make construction projects more proficient.

To streamline the estimation, allocation and management practice in construction arena to generate a proposed recommended tool aiming at increasing efficiency; mounting accuracy while ensuring adaptability contributing significantly towards project success. In addition to helping construction professionals budget and allocate resources more effectively, this tool will also assist the entire construction industry in establishing a new benchmark for cost estimating methods.

## 2.2 Specific Objectives

The research is formulated over a number of specific objectives, which collectively support the major goal of designing machine learning construction cost estimation system. All of these goals correspond to a crucial phase in the life cycle development, deployment and verification of system.

1. Create a Machine Learning Model for Building-Cost Estimation: Here the ultimate goal is to generate high quality machine learning predictive model that can analyze both past historical data and instantaneous in-time order to produce precise estimate of cost associated with several projects on constructions. This includes identifying the relevant machine learning algorithm like linear regression, decision trees to neural networks and fine-tuning them for construction cost estimation specificities. This model will be tested and refined for accuracy using more types of projects in George County.

2. Build Real-Time Data Sources within the Cost Estimation System: The second goal is to add real-time data sources in place so that when estimating costs using the system, price predictions are continually revised as time advances i.e. taking current market situation & project progress into account here). This will involve incorporating the data from IoT-enabled sensors monitoring material usage, labor productivity and live alerts on suppliers/ market trends. With this, the system can provide more accurate and current cost estimates using real-time data making it a stronger tool for dynamic construction environments. ·

3. Enable the Cost Estimation System with a User-Friendly Interface: As another major aim, you should aspire to implement an intuitive and friendly user interface that would make it simple for construction professionals on job sites to enter their project data easily as well as view real-time costs — all add-ons considered (and keeping in line with market values) which they can also track changes along the course of development. The team, within a month will design the interface in order to enable users to interact with it thus creating actionable insights through visuals of data, dashboards and reports. The system needs to provide a new user interface that is easy for someone on site, who may not have high technical knowledge to understand.

4. Obtain Real-World Case Studies and Demonstrate the System Works: The system is evaluated based on how well it produces accurate costs for real construction projects as an ultimate goal of validation. This will check the control cost process and comparing it with actual project costs, also evaluate how accurately,reliably & scalability that system perform The study will also analyse how the system can be implemented in different forms of construction works thus making it a sustainable solution for wider use across the industry. The team will continue to use what they learn from this case study to improve the system, making it more fault-tolerant and general-purpose enough for everyday applications.

All these objectives together ensure that system developed is not only technically correct but also practical to use and of value for construction professionals.

# 3. METHODOLOGY

## 3.1 Research Area

This research is specifically aimed at developing an ML based real-time construction cost estimation system that could address the needs from dynamic and complex industry like CORE. The research area selected covers the interfacing between construction management and digital technology, focusing on how cutting-edge data analytics and machine learning can be used in managing challenges of project planning cost control, resource utilisation.

This provided the necessary background information as to why this research area is of high importance today, given that it still encounters problems related to cost overruns and time slippages in addition to resource mismanagement within construction projects. The high level of variability in modern construction projects is a fact which traditional cost estimation techniques are not always equipped to handle. Usually, the used estimation methods are based on historical data and static pricing algorithms that do not take into account changes in material prices, labor fees or project scope.

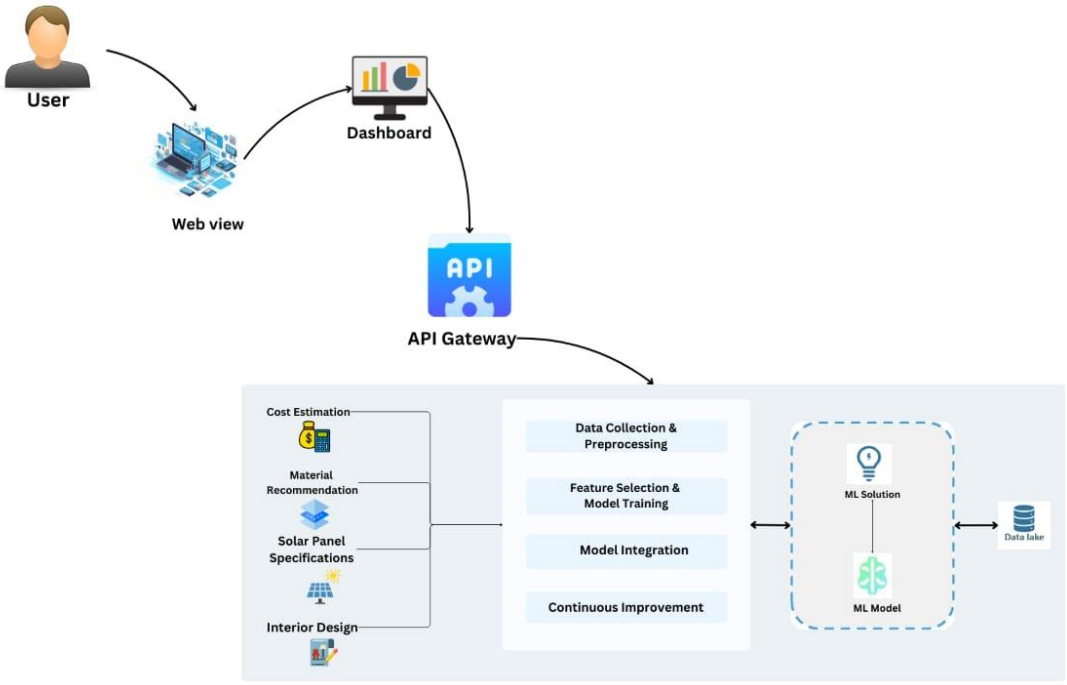
Specifically, we aim to use machine learning for more accurate and timely cost forecasting with this research aiming at improving the processes of decision-making in construction projects so that potential project inefficiencies are minimized. There are tons and tons of data available and these machine learning models can use that amount of 10*5 N Data with each Giga Bytes, identify the patterns in this huge volume of a dataset which were not possible to diagnose for Traditional methods. And in the area of construction, where accurate cost estimation plays a crucial role for a project's success, this is one such capabilities.

This research area also involve the use of data from different sources, i.e., historical project data (e.g., cost or productivity records), real time market development and sensor information on-site that may be used to ensure the system can adapt to other scenarios. Building a model from these disparate data sources and using them together is the step change in cost estimating for construction.

The study thus fills a niche in the extant literature as it is devoted to an emerging research topic that adopts real-time data and predictive analytics, instead of using classical cost estimation techniques. The study seeks to improve project efficiency and reliability in the construction industry through a better balanced application of these modern technologies into planning.

### 3.2 Architecture

The architecture of our proposed system is purposefully built to accommodate the seamless integration of machine learning algorithms with real-time data arrangements that ultimately helps in accurate and up-to-date construction cost estimations. Main components The system architecture included as follows: Data Collect, Clean and Pre-Process Machine Learning Models Training evaluation User Interface




1. Data Collection: This part of the engine is responsible for collecting information from multiple sources (past project data, live market trends and IoT devices/sensors on site). These sources facilitate the estimation of costs by providing necessary data upfront. With more data available from across the spectrum of your organization, this system can provide a holistic perspective into what influences construction costs. As a real-life example of how it works, IoT devices on construction sites can send data in real time about material usage and labor productivity; the system also gets market trends (price) that allow to understand current availability and prices for particular materials.

2. Data Preprocessing: The raw data is collected and it needs to be processed (clean) in order for the model to run properly. Train data prep — Getting the data ready for training involves multiple tasks, i.e. Normalizing Data, Feature extraction and handling missing or partial fill dat While data normalization guarantees that the input is consistent and comparable across different sources, feature extraction helps to find out which variables are useful for predicting construction costs. The formatted data is then saved in a manner which the model can read.

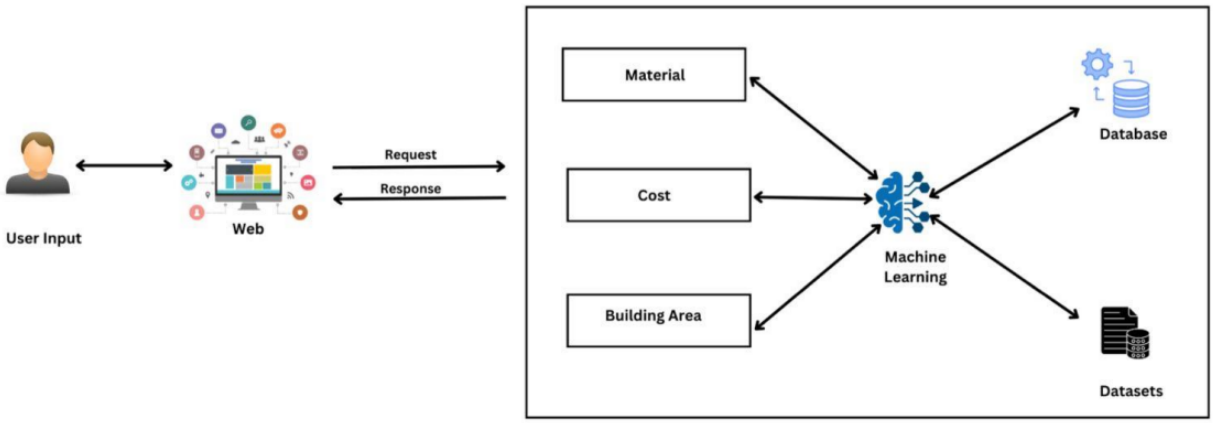
3. The machine learning model Machine Learning Model Training the central part of our system architecture. Following are the components this model consists of, It takes input data first analyze that one and recognize patterns among them based on pattern detect the construction cost. The model is created to be adaptive, trained from new data continuously in order to let the accuracy increase over time. We will see different machine learning algorithm like regression models, decision trees etc which we have to prepare. A historical validation of the model would be trained and deployed on an old set of data to check if is reliable enough.

4. User Interface – The user interface component offers an intuitive and easy-to-understand system for users to work with. The interface lets users enter relevant project data, see live cost estimates and adjust as necessary. The front-end will also have data visualization tools in shape of graphs to analyze and understand the predictions made by system. In order for the system to be picked up by those in construction, user-friendliness and ease of access are key.

This architecture is designed to modular and scalable, as new features can be added per sub module while working on a progress of how the data gets enriched when appended more recent tools or resources contributing informations in use cases evolve over time. The envisioned system is capable of accurate and reliable cost estimation using modern machine learning methods on real-time data, aiding construction professionals in process planning as well management.

### 3.3 Software Architecture

In this software architecture of the proposed construction cost estimation system used for processing and analysis a large volume of data to make our model process timely predictions that will be accurate. It provides a sustainable modular and scalable architecture for new functionality, allowing the system to grow with respect to both offered functions and data set size as it matures.




1. The data layer is the first part of how to build SEO software and it deals with collecting, storing, retrieving information from different sources. On this lower layer resides a resilient database management system (DBMS) made for operations supporting structured and unstructured data, perpetuation of records and sensor data; your fundamentally enterprise stuff. The DBMS is designed to quickly retrieve data used in real-time predictions, making it optimized for fast queries. You might use technologies like PostgreSQL or MongoDB here depending on what kind of structured/unstructured data processing is necessary.

2. Processing layer: Above the data is a processing layer, which cleans and transforms it to shapewhat the machine learning models consume. Lowest layer data processing pipelines: These are series of different derivations/processing happening over the raw features, such as normalization, information extraction techniques and imputation of missing values. The pipelines are implemented to be massively parallelizable, providing the capacity of processing large datasets efficiently and at low latency. Processing this data in the processing layer by preparing it into a usable form for performing machine learning can give more accurate and reliable predictions.

3. ML Layer: The machine learning layer is a set of the actual predictive models five in number. This layer sits on top of any machine learning framework (e.g Tensorflow, Pytorch) and decouple the underlying algorithms from schedulingFormsModule traditional regression models to neural networks & ensemble methods Its machine-learning layer is

modular  meaning you can swap in different models as needed. This flexibility allows it to be tailored for a broad range of construction projects and data sources. These models will be trained on historical data and validated with current inputs to ensure that they can predict the cost of construction in a reliable manner.

5. Deployment layer: Lastly, the deployment layer manages how an application is deployed and scaled across three different environments (on-premises, cloud). This also ensure that the system is strong enough to handle real time data processing and prediction. It has performance monitoring tools built in to monitor the health of your system and ensure nothing goes wrong.

In general the software architecture was designed to be scalable, flexible and resilient.It means that in order to provide value for construction professionals demanding fast accurate real time cost estimates. The architecture makes use of advanced technologies and state-of-the-art CD/CI practices to build the platform, it provides a workbench for implementation which helps in successful realization.

### 3.4 Requirement Gathering and Analysis

For this proposed construction cost estimation system, the requirement gathering and analysis are considered as a key part of development so that final product fulfill all needs of users and give true real time output in terms on accurate cost details. The first step is the identification of stakeholders, some individuals and groups involved in construction projects such as project managers, cost estimators or site engineers. They intimately understand the exact requirements and obstacles they confront in their duties; inputting directly into shaping these system features.

1. Following are some of the steps for requirement gathering Step 1: Stakeholder Identification The first step in requirements elicitation is identifying who uses or will be affected by a system. Construction project managers, cost estimators, site engineers and administrative personnel like procurement officers or financial analysts are among these stakeholders. Involving these stakeholders early in the process will help you understand their needs, expectations and challenges they face with current cost estimation methods.

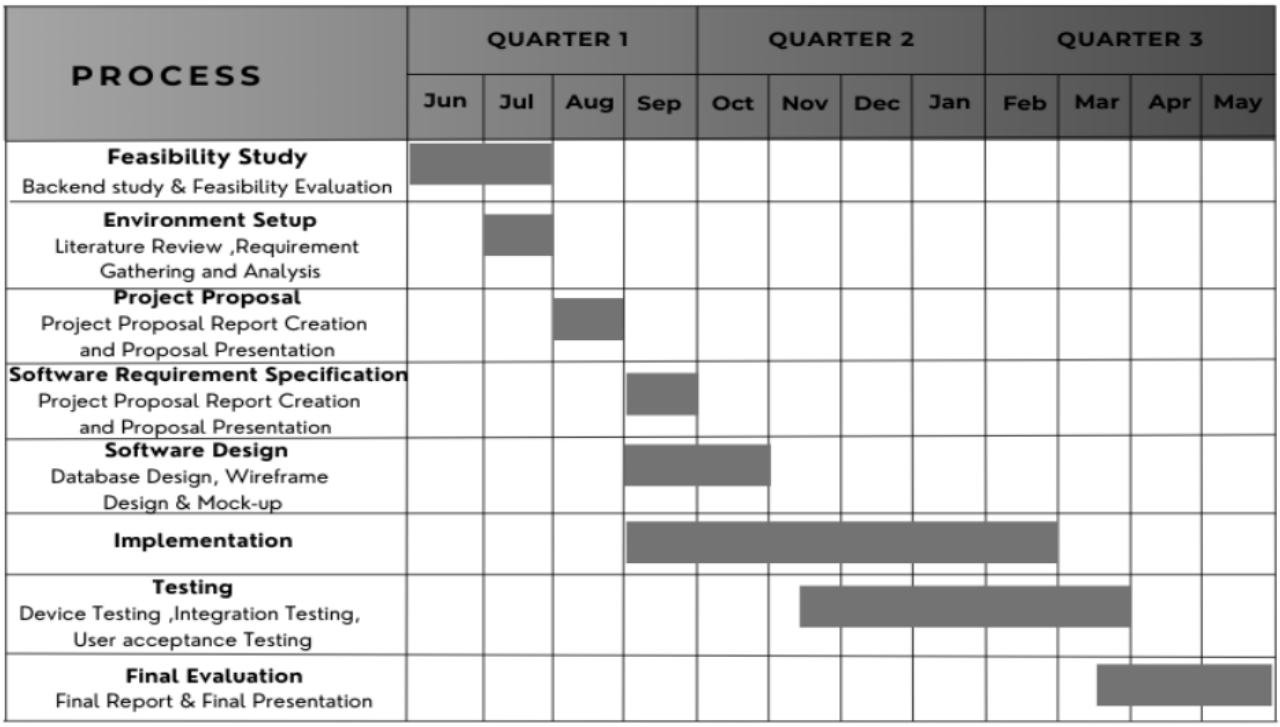
2. Information Gathered: The process of gathering requirements consists interviews, surveys and some other ways that includes stakeholders to understand their needs in detail label X Width x HEIGHT nuts edit resemblance Embed font Name These interactions help in establishing the main functionalities the system should have, such as real-time data integration, ease of use in terms of a graphical interface and capacity for processing high volume off data. Also, users can indicate what all data sources they want to integrate in the system like project historical data, real time sensor streams from machinery and market trend.

3. Functional and Non-Functional Requirements: After collecting the requirements, we analyze if there is a conflict or redundancy gaps found in such recruitment. Functional requirements describe details about what the system is supposed to be able to do: allowing users to input project data, generate real-time cost estimates and visualizing data on interactive dashboards. Non-functional requirements deal with system performance, load, and security. These can consist of the scalability that perform with huge data, responsiveness toward user input functionality and snaps contain privacy security own fields.

4. Analysis also generates issues that use cases and user stories can answer with specifics on how the system will function in real-world scenarios. By illustrating these cases, you allow developers to gain insight into the requirements of diverse user communities that then guide them toward a natural design and easy system navigation. For instance, a use case might define how the system operates for the project manager: after entering new details of his or her project within few minutes he obtains another cost estimate that reflects current market prices.

5. Such documents are most often referred to in a system development project as the “Requirements Document” which is developed during analysis phase and it supposes, though not always considered or understood this way by many who only see these requirements document abstractly. This document specifies every feature, performance and aspect that the system must satisfy in terms of processing details which is easy to understand for a development team. This document makes certain that all the deliverable needs are attended to by maintaining a balance between the business objectives and purpose of software development.

### 3.5 Gantt Chart



# 4. PROJECT REQUIREMENTS

## 4.1 Functional Requirements and Non-Functional Requirements

### 4.1.1 Functional Requirements

1. Data Collection and Input:

The interface should be able to accept information such as project-specific data and past cost, material prices, labor rates or in practice real-time site information from IoT sensors.

The system must be able to support bulk data uploads via CSV files or API integrations with existing project management systems, which allows the seamless integration of and updating content.

2. Real-Time Cost Estimation:

From input data, the system should be able to create real-time cost estimates for projects that are routinely under construction. Both material and labor costs will have a forecast based on market trends along with the conditions of building.

Estimates should be a feedback process where users can change material types, labor rates and get another cost estimate quickly.

3. User Interface and Visualization:

There should be an easy-to-use user interface where the system is able to navigate between features easily, and provide cost estimates. The dashboards in the interface that contain visual data (graphs, charts & tables) is to enhance cost breakdowns and trends for users.

Additionally, the interface should have a way for users to create and export detailed reports that aggregate all the cost estimates, broken down by project phase but also separate based on material type or labor category.

4. Integration with Machine Learning Models

This requires the system to have a suite of machine learning models that can be trained on historical and real-time data in order to predict future costs with high accuracy. This requires models that are constantly learning, adapting based on new data to make statistically stronger predictions as well.

There should be an implemented mechanism that allows for updating and training machine learning models over time as more data is made available coming from different sources.

### 4.1.2 Non-Functional Requirements

1. Performance:

Estimations should be generated in real time, within seconds of inserting new data the system responsible for them needs to be incredibly fast. Solution should be able to deal large volumes of data and there shall not delay or performance degradation.

2. Scalability:

It also needs to be scalable as the amount of data and users continues to grow. Its capacity to be able in expand without major changes the underlying architecture.

3. Security:

The system needs to have high-level robust security so that sensitive data, project details etc and even proprietary business data should not be seen by cheating person. This encompasses data en-route and at-rest, user authentication as well as role-based access control, all complemented by per security audits.

4. Reliability:

In terms of the system has a very high availability, 99.9% up time This includes a failover with redundancy to avoid data loss and maintain continuous operation.

5. Usability:

The system should have high usability, using an interface that is easy to understand and requires little training for construction professionals. Users should have access to help documentation, tooltips and a user-support portal which will further enable users in case of any difficulties.

## 4.2 Technology and Tool Selection

The selection of appropriate technologies and tools is crucial for the successful implementation of the proposed construction cost estimation system. The following considerations were made in selecting the most suitable options

1. Programming Language:

The primary reason for Python is it has a wide range of libraries available in machine learning and data analysis as well also that can be used to develop web projects. Python has libraries like TensorFlow, scikit-learn, and Pandas that are very helpful for machine learning implementation model as well as data processing purposes. The language is highly flexible, easy to learn and offers clean code readability which suits the requirements of building a sophisticated system storing data that requires frequent updates.

2. Machine Learning Framework:

The machine learning framework chosen for our project was TensorFlow due to the robustness, flexibility and support of deep learning models. It utilized apaches park to offload spark for large data processing, and TensorFlow gives the power required in practicing machine learning models that can be fit into a dataset or accommodate continuous prediction making real-time. The prevalence of its use within the industry also means there are plenty of resources and documentation for developers.

3. Database Management System:

PostgreSQL was chosen as the database management system (DBMS) for its reliability, scalability, and support for complex queries. PostgreSQL is well-suited for managing the diverse types of data required for the system, including structured and unstructured data. Its strong support for ACID (Atomicity, Consistency, Isolation, Durability) properties ensures data integrity and reliability, which is crucial for handling sensitive construction cost data.

4. Cloud Services:

Selection of Amazon Web Services (AWS) for Cloud Hosting: AWS has been selected because it provides large amounts services and a wide range support to store data, utilize machine learning algorithms, process real time. Big data and the cloud: The AWS infrastructure ensures that we build our system to scale appropriately with demand. Amazon S3 for data storage, Amazon EC2 to run our code in a scalable manner and Amazon Sage Maker (to deploy machine learning models), are all components of the system.

5. User Interface Development:

React. Front-end The front-end part was built using js because it is a library with component-based architecture that allows for faster construction of dynamic and responsive user interfaces. React. It is very user-friendly as well responsive in all devices with a better UI performance as compared to it alternatives. Thanks to how well it fits into modern web development practices, Elm can be assimilated with other tools and libraries signed some deal?

6. Data Visualization:

D3.js excels in data visualization, allowing complex and interactive charts using familiar HTML, CSS, and JavaScript. It ensures data is clear and easy to understand, aiding accurate decision-making. These technologies make the system fault-tolerant, scalable, and capable of providing real-time construction cost estimates. Python, TensorFlow, and React are used for backend, machine learning, and frontend, respectively, to create a powerful and user-friendly system.

## 5. BUDGET

| **Description** | **Cost (LRK)** |
|---|---|
| Data gathering | $40 |
| Internet | $5 |
| Documentation | $35 |
| System designing | $200 |
| Implementation and Testing | $35 |
| **Total** | **$315** |

## 6. COMMERCIALIZATION ASPECTS

A number of key factors need to be taken into account in order to successfully commercialize the machine learning-powered construction cost estimation system. This is your first business plan and start by other aspects such like; market analysis, pricing strategies, target customers.. Etc.

1. Market Analysis:

The construction is one of the biggest and fastest growing sectors in every countries. The requirement for cost estimation tools are always on a rise due to constant demand from anyone involved with buildings or public works. The traditional methods are more manual as opposed to automated, time-consuming and error-prone. These needs are all met with a complex system that leverages machine learning to provide live cost and material estimations, which specifically construction companies find extremely valuable. With digitalization as an already booming trend in construction, it promises a huge market potential.

2. Target Customers:

This system is primarily designed for construction companies from small contractors to large multinational firms. These companies need accurate cost estimation tools to effectively manage their projects and remain competitive in the market. Among other potential clients are project management companies, government infrastructure development departments and educational institutions with construction management programs. However, it will be necessary to target a marketing strategy for the customers. This may include sales efforts involving those within the industry or through an association, and online marketing campaigns to create awareness of its benefits.

3. Pricing Strategy:

Pricing is another critical success factor in the commercialisation of the system and there is therefore need to come up with a well-developed pricing strategy. The pricing model could be a per-ticket fee which involves the users having to pay a periodical fee for the operating of the system, monthly or annually. This model guarantees constant revenues for the system and opens it to a great number of users on the other side. On the other hand, what is known as the ‘degree-of-service’ model with definite price structures could be adopted. For instance, a low-risk package may contain fundamental functionalities like cost estimating, reporting, and data analysis within the tool Or, a high-risk package may contain additional features regarding machine learning customization, integration of data, and timely support.

4. Competitive Advantage:

The potential competitive advantage of the system is the capacity to generate precise estimates of cost in real time by wide use of machine learning technologies. As opposed to other systems used for estimation, which incorporate fixed data to produce estimates, this system re-calculates the estimates based on the latest data which the users input and therefore the estimates supplied always correspond to the most recent cost estimates. Also, because of the application's interface and the integration with other construction management applications, the system can be easily implemented and utilised. Among the strengths, the application of sophisticated data visualization methods, enabling users of the system to make adequate use of their data and thus come up with valid conclusions quickly, should be considered.

5. Revenue Streams:

Apart from subscription fees other sources of income that can be realized in the system are; One is to sell professional services, which pay consultants to work with the companies to configure it to their needs, or train them or help them thereafter. Another source of income may be provided by the provision of data analysis services. Using the data that would be obtained through the system, the company could offer other information regarding trends within the industry, the cost, and efficiency standards. Last but not the least; other affiliations with other software firms or construction industry players can create other streams of revenue.

The marketing and distribution strategy of the construction cost estimation system concerns the general approach and plans of achieving maximum market coverage and ensuring the viability of the system in future. In this respect, the system enhances an element that the construction industry gravely needs, especially for real time cost estimates during project delivery. Thus, the system has all the major prospects of being successful in the market if the top priorities in the process of setting the pricing policy are to be aimed at defining the target market segments and creating effective marketing messages in the target market segments, as well as at identifying the key competitive advantages of the company´s offering.

## REFERENCES


[1] Kim, J., Lee, H., & Park, S. (2019). "Application of Artificial Neural Networks in Construction Cost Estimation: A Case Study of High-Rise Building Projects. " *Journal of Construction Engineering and Management, 145*(3), 04019010.

[2] Zhang, X., Li, Q., & Wang, J. (2020). "IoT-Driven Construction Cost Estimation Using Machine Learning: A Real-Time Approach. " Automation in Construction, 112, 103094.

[3] McMahon, S., & Smith, R. (2018). "Enhancing Construction Cost Estimation Accuracy Through Machine Learning and Big Data Integration. " *Journal of Construction Research, 26*(4), 567-583.

[4] Kim, G. H., An, S. H., & Kang, K. I. (2004). "Comparison of Construction Cost Estimating Models Based on Regression Analysis, Neural Networks, and Case-Based Reasoning. " *Building and Environment, 39*(10), 1235-1242.

[5] Son, H., & Kim, C. (2010). "Hybrid Principal Component Analysis and Support Vector Machine Model for Predicting the Cost Performance of Commercial Building Projects Using Pre-Project Planning Variables." *Automation in Construction, 19*(7), 844-850.

[6] Cheng, M. Y., & Roy, A. F. (2011). "Evolutionary Fuzzy Decision Model for Construction Material Procurement. " *Automation in Construction, 20*(5), 636-644.

[7] El-Sayegh, S. M., & Mansour, M. H. (2015). "Risk Assessment and Allocation in Highway Construction Projects in the UAE. " *Journal of Management in Engineering, 31*(6), 04015002.

[8] Yu, J., & Skibniewski, M. J. (1999). "A Neurofuzzy Computing Approach to Construction Project Cost Estimation. " *International Journal of Project Management, 17*(1), 9-16.

[9] Chou, J. S., & Lin, C. W. (2013). "Predicting Disputes in Public Construction Projects: A Machine Learning Approach. " *Journal of Civil Engineering and Management, 19*(1), 21-30.

[10] Akintoye, A., & Fitzgerald, E. (2000). "A Survey of Current Cost Estimating Practices in the UK. " *Construction Management and Economics, 18*(2), 161-172.

[11] Wibowo, M. A., & Alfen, H. W. (2015). "A Systematic Review of Cost Overrun Research in Construction. " *Procedia Engineering, 125*, 89-94.

[12] Memon, A. H., Rahman, I. A., & Azis, A. A. A. (2012). "Time and Cost Performance in Construction Projects in Southern and Central Regions of Peninsular Malaysia." *International Journal of Advances in Applied Sciences, 1*(1), 45-52.