IDENTIFYING THE FACTORS AFFECTING CONCEPTUAL COST ESTIMATION OF CONSTRUCTION PROJECTS IN VOLATILE ECONOMIES

by

Ronak Sadeghian

B.Sc. in Civil Engineering, K. N. Toosi University of Technology, 2020

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Civil Engineering
Boğaziçi University

2023

ACKNOWLEDGMENTS

First, I would like to express my profound gratitude and appreciation to my advisor, Assoc. Prof. Semra Çomu Yapıcı, for guiding and supporting me through my education journey. Getting through several courses with Dr. Çomu, I felt lucky to learn priceless things from her that I can carry with me during my engineering path. Her broad insight into engineering topics paves the road for me to find my interest. Her tips and comments pushed me to think deeper and taught me to look at a certain subject from a different point of view.

I would like to thank the committee members, Prof. Beliz Özorhon and Prof. Murat Günaydın, for their valuable time and their insightful comments. I am grateful for their advice and suggestions.

Additionally, I would like to thank whoever helped me with data collection and shared their valuable experience with me.

Finally, I would like to thank my family for their tremendous support in every stage of my life, especially when it gets hard. I must express my deepest gratitude to my mother, Roshanak Shabafrooz, for her unconditional love and support, and to my father, Majid Sadeghian, for his words of wisdom.

ABSTRACT

IDENTIFYING THE FACTORS AFFECTING CONCEPTUAL COST ESTIMATION OF CONSTRUCTION PROJECTS IN VOLATILE ECONOMIES

The construction industry plays a vital role in the country's economy. Increasing the success rate of construction projects has become more important, especially, in volatile countries. For that matter, conceptual cost estimation is used from the very first point of the project, when the feasibility study is done. After that, when more detailed information is available, more accurate cost estimation is done with the help of the project design.

In countries with volatile economies, predicting project costs beforehand is challenging due to economic fluctuations, resulting in more uncertainties when starting a construction project. Accurate conceptual cost estimation can help project owners overcome cost overruns and face economic instabilities. This study proposes a method for conceptual cost estimation for construction projects using historical data analysis as the foundation for creating a connection between the primary factors influencing the cost of building projects to produce estimations for new projects in Turkey.

The goal of construction conceptual cost estimation is to predict the cost of a construction project, including all temporary expenditures generated throughout the construction process, and to serve as a basis for budgeting, planning, and decision-making. This paper uses quantitative data analysis methods such as Neural Network, Decision Trees, and Regression to develop a cost model using the historical data of construction projects in Turkey. The models' accuracy was then evaluated by measuring the mean absolute percentage error. The effect of design and economic factors which were used as the input of the cost modes were then discussed in this study. This study further suggests using the economic factors alongside the design factors to develop more accurate cost models.

ÖZET

DALGALI EKONOMİLERDE İNŞAAT PROJELERİNİN KAVRAMSAL MALİYET TAHMİNİNİ ETKİLEYEN FAKTÖRLERİN BELİRLENMESİ

İnşaat sektörü ülke ekonomisinde hayati bir rol oynamaktadır. İnşaat projelerinin başarı oranının artırılması, özellikle istikrarsız ülkelerde daha önemli hale gelmiştir. Bu nedenle, kavramsal maliyet tahmini, fizibilite çalışması yapıldığında projenin ilk noktasından itibaren kullanılır. Daha sonra, daha ayrıntılı bilgiler mevcut olduğunda, proje tasarımı yardımıyla daha doğru maliyet tahmini yapılır.

Değişken ekonomilere sahip ülkelerde, ekonomik dalgalanmalar nedeniyle proje maliyetlerini önceden tahmin etmek zordur ve bu da bir inşaat projesine başlarken daha fazla belirsizliğe neden olur. Doğru kavramsal maliyet tahmini, proje sahiplerinin maliyet aşımlarının üstesinden gelmesine ve ekonomik istikrarsızlıklarla yüzleşmesine yardımcı olabilir. Bu çalışma, Türkiye'deki yeni projeler için tahminler üretmek amacıyla inşaat projelerinin maliyetini etkileyen birincil faktörler arasında bir bağlantı oluşturmak için temel olarak tarihsel veri analizini kullanarak inşaat projeleri için kavramsal maliyet tahmini için bir yöntem önermektedir.

İnşaat kavramsal maliyet tahmininin amacı, inşaat süreci boyunca ortaya çıkan tüm geçici harcamalar da dahil olmak üzere bir inşaat projesinin maliyetini tahmin etmek ve bütçeleme, planlama ve karar verme için bir temel oluşturmaktır. Bu makale, Türkiye'deki inşaat projelerinin geçmiş verilerini kullanarak bir maliyet modeli geliştirmek için Yapay Sinir Ağı, Karar Ağaçları ve Regresyon gibi nicel veri analizi yöntemlerini kullanmaktadır. Modellerin doğruluğu daha sonra ortalama mutlak yüzde hata ölçülerek değerlendirilmiştir. Bu çalışma ayrıca, daha doğru maliyet modelleri geliştirmek için tasarım faktörlerinin yanı sıra ekonomik faktörlerin de kullanılmasını önermektedir.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii

ABSTRACT iv

ÖZET…. v

LIST OF FIGURES ix

LIST OF TABLES xi

LIST OF ACRONYMS / ABBREVIATIONS xii

1. INTRODUCTION 1

1.1. Problem Definition and Statement 3

1.2. Aim and Objectives of the Study 4

1.3. Research Methodology 4

1.4. Scope and Limitations 5

1.5. Organization of the Study 5

2. BACKGROUND 7

2.1. Overview of Construction Cost Estimation 7

2.1.1. Definition 7

2.1.2. Accuracy 7

2.1.3. Limitations 8

2.2. Purpose of Cost Estimation 9

2.3. Types of Construction Cost Estimates 9

2.3.1. Cost Estimation Based on the Construction Phase 9

2.3.2. Construction Cost Estimation Based on Technique 11

2.3.2.1. Quantity Take-off Estimate 11

2.3.2.2. Unit Price Estimate 11

2.3.2.3. Parametric Estimate 12
2.3.2.4. Bottom-up Estimate 12
2.3.2.5. Comparative Estimate 12
2.3.2.6. Expert Judgment Estimate 12
2.4. Factors Affecting Conceptual Cost Estimates 12
2.4.1. Structural Design Variables Affecting Cost Estimation 12
2.4.2. Economic Variables and Indexes Affecting Cost Estimation 15
2.4.2.1. Inflation Rate 15
2.4.2.2. Exchange Rate 16
2.4.2.3. GDP 17
2.4.2.4. Consumer Price Index 18
2.4.2.5. Minimum Wage 18
2.5. Overview of Quantitative Cost Modeling in the Construction Industry 19
2.6. Types of Quantitative Algorithms and Their Efficiency in Cost Estimation 24
2.7. Limitations of Cost Estimation Algorithms 31
2.8. Identification of the Literature Gap 32
3. RESEARCH METHODOLOGY 34
3.1. Research Framework 34
3.2. Input Variables 37
3.3. Hypotheses Development 39
3.4. Data Collection 42
3.5. Quantitative Data Modeling Algorithms 46
3.5.1. Neural Network 46
3.5.2. Decision Tree 48
3.5.3. Regression 50
3.6. Evaluation of the Algorithms 52

4. RESULTS AND FINDINGS 55
4.1. Relative Performance of Algorithms 55
4.2. Result of Each Estimation Model 56
4.2.1. Neural Network 56
4.2.2. C&R Tree 60
4.2.3. Regression 63
4.2.4. CHAID 68
5. DISCUSSION 70
5.1. Interpretation of The Result 70
5.2. Limitations and Future Research 74
6. CONCLUSION 77
REFERENCES 80
APPENDIX A: DATA COLLECTION SHEET 94
APPENDIX B: COST MODELS RESULT 97

LIST OF FIGURES

Figure 3.1. Research methodology flowchart 36

Figure 3.2. Location of the projects 43

Figure 3.3. Foundation system of the projects 43

Figure 3.4. Floor system of the projects 44

Figure 3.5. Framing type of the projects 44

Figure 3.6. Contract type of the projects 45

Figure 3.7. Exterior finishing type of the projects 45

Figure 3.8. Usage of the buildings in the projects 46

Figure 3.9. Neural network model 48

Figure 3.10. Decision tree model 50

Figure 3.11. IBM SPSS 18.0 model setting 53

Figure 3.12. Algorithms selected in IBM SPSS 18.0 54

Figure 4.1. Inputs and target. 55

Figure 4.2. Top four cost models and their relative errors 57

Figure 4.3. Developing neural network model 58

Figure 4.4. Neural network model summary 58

Figure 4.5. Neural network 59

Figure 4.6. Cost drivers' impact on the NN model 60

Figure 4.7. C&R tree 62

Figure 4.8. Cost Drivers' impact on the C&R tree model 63

Figure 4.9. Cost drivers' impact on the regression model 65

Figure 4.10. Cost drivers' impact on the CHAID model 68

Figure 4.11. Schematic view of the CHAID model 69

Figure 5.1. Neural network model accuracy using the design variables 73

Figure 5.2. The effect of design variables on NN model 74

LIST OF TABLES

Table 3.1. Design factors considered as the cost driver 38

Table 3.2. Economic factors considered as the cost driver 39

Table 4.1. Label encoding of categorical variables 64

Table 4.2. Regression model summary 65

Table 4.3. Regression coefficients 66

Table 5.1. Importance of design variables 73

Table 5.2. Importance of economic variables 72

Table B.1. Neural network cost estimation result 97

Table B.2. C&R cost estimation result 98

Table B.3. Regression cost estimation result 99

Table B.4. CHAID cost estimation result 100

LIST OF ACRONYMS / ABBREVIATIONS

Abbreviation	Explanation
3D	Three Dimensional
4D	Four Dimensional
AHP	Analytic Hierarchy Process
AI	Artificial Intelligence
ANN	Artificial Neural Network
ANN	Artificial Neural Network
ARDL	Autoregressive Distributed Lag
BIM	Building Information Modeling
BoQs	Bills of Quantities
BPNN	Back Propagation Neural Network
C&R	Classification and Regression
CART	Classification and Regression Tree
CBR	Case-Based Reasoning
CHAID	Chi-squared Automatic Interaction Detection
CPI	Consumer Price Index
EFHNN	Evolutionary Fuzzy Hybrid Neural Network
EFNIM	Evolutionary Fuzzy Neural Inference Model
EIFS	Exterior Insulation and Finish System
EV&I	Economic Variables and Indexes
GA	Genetic Algorithm
GDP	Gross Domestic Production
HEPP	Hydro-Electrical Power Plant
IBM	International Business Machines
ID3	Iterative Dichotomiser 3
IT	Information Technology
KDD	Knowledge Discovery in Databases
KNN	K-Nearest Neighbor

LSVM	Lagrangian Support Vector Machine
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MLP	Multilayer Perceptron
MLR	Multiple Linear Regression
MRA	Multiple Regression Analysis
MSE	Mean Squared Error
NN	Neural Network
PMBOK	Project Management Body of Knowledge
PPI	Producer Price Index
RF	Random Forest
RFB	Radial-Basis Function
ROM	Rough Order of Magnitude
SOM	Self-Organizational Map
SPSS	Statistical Package for the Social Sciences
SVM	Support Vector Machine
TL	Turkish Lira
TurkStat	Turkish Statistical Institute
USD	United States Dollar
WPI	Wholesale Price Index
XGBoost	Extreme Gradient Boosting

1. INTRODUCTION

The construction industry directly and positively impacts the country's economic growth considering its connection to other sectors and its immense contribution to employment [1]. The construction industry is also responsible for a significant part of material production within a country which results in the development of all sectors; thus, the national income largely depends on the growth rates of the construction industry [2]. Being able to accurately predict the construction project's cost at the initial stages has become crucial to avoid cost overruns and increase the construction companies' profit. The goal of construction cost estimating is to accurately predict the cost of a construction project, including all temporary expenditures generated throughout the construction process, and to serve as a basis for budgeting, planning, and decision-making [3]. The major problems with conventional cost estimation strategies are the need for detailed project information, changes in the design parameters, uncertainties of project development, etc. Qualitative data analysis methods can reduce uncertainties in cost estimation and provide designers with rough estimates during the design phase. Furthermore, traditional cost estimation ignores macroeconomic variables such as the wholesale price index, liquidity, and building service index, which may result in cost variations, especially in an unstable economic environment [4]. Therefore, integrating the macroeconomic variables into the conceptual cost estimates while developing cost models can help the project owners predict their costs more realistically and lead to the project's success. Various strategies have been used to estimate construction costs accurately at an early stage when project information is scarce. Regression analysis, neural networks, support vector machines, and case-based reasoning are some these techniques that have been used recently.

The practical usage of conceptual cost estimate models in the construction industry is still limited. He et al. [5] assert that only 36% of construction companies and 21% of owners use advanced data analysis for project estimation. They emphasized the importance of developing early cost-estimating models and considering them as a data mining process in order to use data-driven methods accurately. There is a broad consensus among construction professionals that the traditional conceptual cost estimation is not because of the huge error

rates and subjective comparisons; yet, many construction companies and project owners find it difficult to utilize data-driven cost estimation models. There are some difficulties associated with using such data-driven construction, such as; the unavailability of enough historical data, weak data management practices, and lack of IT knowledge and systems [5]. Therefore, the success of conceptual cost estimation is primarily dependent on the historical data set of completed projects in the construction location and the knowledge of discovery from large data sets.

Cost estimation of a construction project typically starts with the rough order of magnitude (ROM), which refers to the initial state of cost estimation. It helps to achieve an approximate estimate at the conceptual stage to assess the project's feasibility. The accuracy of such estimates varies from -25% to 75% [6]; yet, they have a substantial impact on project management and impact the projects' success. Later in the projects, when more detailed project information is available, a more accurate estimation can be done by creating bills of quantities (BoQs) which have an accuracy of -5% to 10% [7].

Identifying the factors affecting the cost estimation accuracy and determining various methods to perform the cost estimation has become substantial over time, especially in an unstable economy. This study aims to propose an approach to increase the accuracy of conceptual cost estimates or rough order of magnitudes by integrating macroeconomic factors of the country of construction location. This study emphasizes that design variables are not solely sufficient to predict the construction projects' conceptual costs. This magnifies when the construction project is progressing in a developing country because of the volatile economies, which leads to economic uncertainties. Countries with economic volatility often experience macroeconomic shocks which directly impact the public finances and affect government assets and liabilities [8].

Macroeconomic factors are economic indicators of a country that affect the overall performance of an economy [9]. These factors focus on aggregate measures such as national income, inflation, employment rate, and overall economic growth, and they depend on government policies, global economic conditions, and social and cultural factors [10]. Some of the macroeconomic factors are; Gross Domestic Production (GDP), Unemployment,

Inflation Rate, Fiscal Policy, Monetary Policy, Exchange Rates, Interest Rates, Average Net Wages, and Economic Growth Rate [11].

1.1. Problem Definition and Statement

Estimating the project's cost is one of the most critical components of a construction project which provides the detailed cost of the individual resources in a tender. The complex nature of the construction projects and the number of uncertainties associated with that result in severe difficulty in accurately predicting the project's cost. Therefore, forecasting the budgets and completing the construction project within the preliminary cost estimate has been a challenge for project owners. To increase the accuracy of cost estimates, identifying the effective factors in a construction project is crucial. This accelerates when the construction project is taking place in a developing country with a volatile economy since the construction industry largely depends on the stability of the country [2]. When it comes to an unstable economy, construction companies, and project owners should make decisions with economic consideration and fluctuations in order to overcome cost overruns and increase profitability. The global financial crisis of 2008 is an excellent example of economic instability, which led to the bankruptcy of a significant number of construction companies [12].

Many previous studies proposed construction cost models to predict the total project cost by inputting project characteristics. These characteristics are mostly design factors that were used as the cost drivers of the models. There are many structural design variables affecting a construction project's cost. These variables include many items like the choice of the structural frame, whether or not to use high-quality material, whether or not no use elevator, the design of reinforced concrete materials, the use of composite frames, storey's height, floor system, and exterior finishing type. The initial construction cost of various structural systems is a significant factor in the selection of the structural system for residential buildings [13]. However, not many studies have integrated economic factors into the cost models. Semenova and Vitkova [2] asserted that external environmental factors such as exchange rates, imports, exports, tax rates, and income level play a crucial role in the financial decision-making of construction companies in the Czech Republic and Spain and give the companies and project owners the insight to create more effective financial

strategies for successful and stable development. However, there is a literature gap on conceptual cost estimation models using explicitly macroeconomic factors along with design variables.

1.2. Aim and Objectives of the Study

This study aims to develop a conceptual cost estimate using macroeconomic factors aside from the conventional structural design factors. The main question of this study is whether there is a considerable difference in the model's accuracy when integrating the economic factors into the cost model in a developing country with a volatile economy. The main objectives of the study are:

i. To discover the effective cost drivers of construction projects in Turkey.
ii. To develop a cost model that can be used as a conceptual cost estimate tool in future projects using both macroeconomic factors and design variables.
iii. To investigate how macroeconomic factors affect the cost model accuracy.

1.3. Research Methodology

This study includes an extensive literature review to identify the structural design variables used to develop cost models in the previous studies. Macroeconomic factors and their relation to construction projects have also been investigated in previous studies. Quantitative data analysis models such as Neural Network, Regression, and Decision Trees were used to develop conceptual cost models using the identified macroeconomic and design variable. In order to develop the cost models, data collection has been done by collecting construction cost data within Turkey. Sixty-four data points were gathered from the completed projects in Turkey with the total construction cost and technical specifications, and based on the construction year, the economic factors were added for each project. SPSS modeler 18.0 was used to develop quantitative models, analyzing and validating the models.

1.4. Scope and Limitations

The scope of this study is to investigate the factors affecting conceptual cost models and to identify macroeconomic factors that help increase the cost models' accuracy and reduce cost overruns.

Because of the unstable economy that developing countries face, it has become a struggle for construction firms to budget their project without overruns accurately. There is not enough research on these economic factors affecting the construction industry in Turkey; therefore, this study aims to investigate the influence of macroeconomic factors affecting the construction industry in Turkey.

Some limitations are associated with this study, such as the number of participants in the data collection process and the diversity of the construction projects. Furthermore, there is a need to investigate further the relationship between macroeconomic factors and how they affect each other since they are not independent variables. Therefore, the interdependency of the economic variables used as the model input should be further studied.

1.5. Organization of the Study

This study consists of six chapters. The first chapter (Introduction) briefly discusses the significance of the study along with the problem definition, research objectives, methodology, and research limitations. The second chapter (Background) focuses on the research background. It examines the existing literature on construction cost estimation in different project stages, focusing on initial cost estimates, factors affecting conceptual cost estimates, including structural design and economic variables, and quantitative data analysis approaches to develop cost estimate models. The third chapter (Methodology) introduces the research methodology by presenting the methodology flowchart, hypotheses, data collection process, and quantitative data analysis methods. The fourth chapter (Results and Findings) presents the findings and results of the cost models and the errors associated with each model. The fifth chapter (Discussion) dive into the research findings in more detail by presenting the interpretation of the results, research limitations, and recommendation for

future studies. Finally, the sixth chapter (Conclusions) summarizes the study's findings and includes the study's contributions.

2. BACKGROUND

2.1. Overview of Construction Cost Estimation

2.1.1. Definition

Cost estimation is the procedure of examining a project's scope of work and forecasting the likely cost of work completion [14]. Cost estimating is considered one of the most critical components of construction projects providing the detailed cost of the individual resources in a tender [15]. The complex nature of the construction projects and the inherent uncertainty of the project's financial performance and funding result in severe difficulty in forecasting the exact budget [16]. Completing the construction projects within the preliminary cost estimate has been a challenge for the construction industry; project cost overestimation results in higher tender prices and ultimately getting rejected by the client; on the other hand, underestimating the project cost results in profit loss. Identifying the factors affecting the cost estimation accuracy and determining various methods to perform the cost estimation has become substantial over time.

2.1.2. Accuracy

The accuracy of the cost predictions has always been a challenge. Most construction projects, especially the ones with larger scales, have experienced cost and schedule overruns in which the final project cost is greater than the preliminary cost estimation prepared during the project planning phase. Ssemwogerere [17] claims that most construction projects are completed with a cost usually 25-35 percent greater than the initial cost, thus leading to cost overruns. Based on the survey by Akintoye and Fitzgerald [18], contractors tend to undertake cost estimating for construction planning purposes such as tender preparation and project evaluation, regardless of their company size [18]. According to Vergara and Boyer [19], the accuracy of the cost estimates depends on both the method and the intended use of the estimate [19]. Many factors affect the accuracy of these estimates, such as project characteristics, project teams, estimating procedures, and cost-related information.

Throughout time, researchers have tried to assess the impact of different project characteristics on its cost estimation. Arif et al. [20] argued that among all the different drivers, the contractor's involvement in developing the estimation and labor rates was the most critical factor [20].

Akintoye and Fitzgerald [18] discovered the most significant reasons for inaccuracy in cost estimation: lack of practical construction knowledge by the estimators, poor tendering documentation, market conditions, insufficient time to prepare the estimation, and variability of subcontractors' prices [18].

Aibinu and Pasco [21] discussed that the accuracy of the pre-tender cost estimates varies according to the size of the project and the principal structural material. They further discussed that the cost estimates before the tendering process tend to be overestimated rather than underestimated. Also, the error percentage for overestimating was higher than for underestimating. They suggested utilizing probability estimation, simulation of past estimates, integrating market sentiments into the estimates, and proper documentation for the firms to increase the accuracy of the cost estimates.

2.1.3. Limitations

Predicting construction costs is subject to many constraints that can affect its accuracy. One of them is a lack of data quality and quantity, which can impair estimation accuracy [17]. Another drawback is the complexity of the construction process, which makes precise cost estimation difficult [22]. Furthermore, the lack of industry standardization might make it difficult to produce estimates that can be applied across multiple projects and circumstances. Using out-of-date cost data can also result in wrong estimations. Furthermore, the unpredictability involved with construction projects can make precise cost estimation difficult. Finally, estimators' engagement with specifications for construction cost estimation has a significant impact on the efficiency and accuracy of cost estimates.

2.2. Purpose of Cost Estimation

One of the most important aspects of construction project management is determining the cost of the project since it is important for all the parties involved in the project. A comprehensive project cost estimate is essential for the project's success. The goal of construction cost estimating is to accurately predict the cost of a construction project, including all temporary expenditures generated throughout the construction process, and to serve as a basis of budgeting, planning, and decision-making [3]. Accurate cost estimation is essential for completing a construction project within the schedule and within the predicted budget, as well as optimizing the project's economic benefits.

Because of the severe swings in the price of building materials and construction equipment, a realistic construction project requires the construction company to properly create the cost estimate and purchase the building materials and construction equipment. This estimation should be done in a way that maximizes savings; during construction, the temporary acquisition of equipment and materials is an efficient strategy to cut the construction enterprise's superfluous expenditures and save the project cost.

The estimated cost is heavily affected by the client's specifications and the project requirements. In some cases, the owner/client or the contractor may require information of the lowest tender price for the project in order to pick a contractor/subcontractor. As a result, the project cost estimate is considered an important aspect for clients before starting the construction project. It also analyzes the project's feasibility and provides an idea for budget control following tendering [16].

2.3. Types of Construction Cost Estimates

2.3.1. Cost Estimation Based on the Construction Phase

Construction estimates can be used for a feasibility study as well as planning the budget later on when the project starts. Therefore, the estimation can be done during the different phases of the project as planning, scoping, design, or construction. The accuracy of the estimates plays a significant role in projects' success [20]. Estimates developed later in

the delivery process have historically been more indicative of actual costs than those developed early in the process. However, early cost estimates help the owner, contractor, designer, etc., to determine the feasibility of a project, check the alternatives, and establish the initial budget. The accuracy of the construction cost estimate tends to increase as the project progresses toward the end [6].

Estimates which are done during a project's planning, scoping, or early design phases, when very little information is known about the project, have consistently been much lower than the actual final cost. This usually results in financial challenges for the owner of the project, whether it is private or public. Furthermore, the public has lost faith in the ability of the government to estimate and deliver a project for the estimated cost. This necessitates a more precise cost estimate at the planning stage. Furthermore, comparing the actual cost to the early cost predictions can be utilized to evaluate the performance and overall success of the project team [23].

Cost estimating typically begins with the rough order of magnitude (ROM), which refers to the initial state of cost estimation for a project. An approximate estimate is required at the conceptual stage to assess whether the project's feasibility should be studied [24]. The estimated accuracy of this form of estimation (ROM) is -25% to +75%, according to the PMBOK. Furthermore, conceptual cost estimates have a substantial impact on project management since they are used to study the feasibility and impact on the final project's success [25]. Later in the project, when more thorough information is available, definitive estimates with -5% to +10% accuracy are produced.

After passing the design stage, detailed cost estimates by contractors are done for creating bills of quantities (BoQs), which are done with standard measurement standards. Although there are standard procedures for contractors to estimate projects, when it comes to cost estimation in practice, they frequently develop their own ways [7].

Increasing profit margins while maintaining a high level of quality in construction projects has become difficult. Economic issues contribute to over 60% of project failures [26]. Designers employ cost estimation tools and rely on their experience and data from prior and current projects to reduce construction costs. Traditional cost estimating has relied on

specifications and blueprints; however, comparative cost estimation requires a linear link between the project's final cost and basic design characteristics, which is problematic and causes mistakes [27]. Cost management is based on conceptual cost estimation for planning, budgeting, and execution [5]. AI models are currently considered an alternative method for monitoring project costs and cash flows since they allow for the development of non-linear relationships and offer more realistic and accurate results [28].

Estimates made during the scoping phase of a project are used to establish the project's baseline cost. These estimates are critical since they are used to confirm feasibility and make preliminary budgetary decisions. At this stage, approximate quantities for items like outer walls, concrete slabs, excavation, and so on can be determined.

Several estimations are made during the design stage to track the cost dependent on the design modification. At this stage, documentation will aid the estimator in correctly tracking the project cost as the design changes. Furthermore, hazards should be identified and considered during the design development process, as well as when estimating costs.

2.3.2. Construction Cost Estimation Based on Technique

2.3.2.1. Quantity Take-off Estimate. The basis of quantity take-off estimates is to estimate the project's material and amount of each item and then price each item. This type of cost estimation is mostly done along with other types of estimates, such as preliminary or detailed estimates. Quantity take-off estimates are useful for establishing the detailed quantities of materials needed for a building project, but they can be time-consuming and require a high level of practical competence.

2.3.2.2. Unit Price Estimate. A unit price estimate is done based on the cost per unit of material or labor, and it is often used for repetitive tasks. According to research by Owoeye [29], unit price estimates are helpful for cost control and provide a quick and easy way to estimate the cost of a project. However, they may not be accurate if the unit prices used are outdated or if there are unforeseen changes in the project while proceeding with the construction.

2.3.2.3. Parametric Estimate. Parametric cost estimates use historical and statistical data to estimate a project's costs based on similar projects. This type of estimate is beneficial when there are not enough details available and there is limited information and specification available about the project. Parametric estimates can provide a quick and accurate estimate of the cost of a project, but they rely heavily on the quality of the data used.

2.3.2.4. Bottom-up Estimate. A bottom-up estimate involves breaking down the project into smaller sections and estimating the cost of each component separately before adding up all the costs for the entire project. Bottom-up estimates are useful for identifying potential cost savings opportunities, but they can be time-consuming and require a high level of technical knowledge.

2.3.2.5. Comparative Estimate. A comparative estimate is based on the costs of similar projects and is frequently used when no previous cost data for the project is available. Comparative estimates can provide a quick and accurate estimate of a project's cost, but they are strongly dependent on the similarity of the projects being compared.

2.3.2.6. Expert Judgment Estimate. To estimate costs, an expert judgment estimate relies on the skills and knowledge of the experts in the field. Expert judgment estimates, are effective for projects with a high level of uncertainty, but they can be subjective and rely primarily on the expertise of the individuals involved.

2.4. Factors Affecting Conceptual Cost Estimates

2.4.1. Structural Design Variables Affecting Cost Estimation

There are many structural design variables affecting a construction project's cost. These variables include many items like the choice of the structural frame, whether or not to use high-quality material, whether or not no use elevator, the design of reinforced concrete materials, the use of composite frames, storey's height, floor system, and exterior finishing type. The initial construction cost of various structural systems is a significant factor in the selection of the structural system for residential buildings [13]. The use of reinforced concrete walls can provide a significant increase in performance with a very small increase

in construction cost [30]. The implementation of green design strategies from the design, construction, and day-to-day running of the building can reduce its potential effect on the natural environment, the running costs of the structure, and the negative effects on the users. The use of cloud interaction technologies can save construction costs by up to 25%, and when building high-rise buildings and structures, up to 35% [31]. The cost of safety estimation is critical in construction, specifically for high-rise building constructions [32]. The choice of building materials and structural form chosen in the design phase not only affects the costs for construction but also decides the costs for maintenance [33].

Cheng et al. [28] identified significant factors affecting construction projects' cost and, consequently, developed the EFNIM (Evolutionary Fuzzy Neural Inference Model) cost model. These factors were; floor, area, duration, foundation system, structural system, and finishing.

Günaydın and Doğan [27] considered eight factors to develop a neural network model for cost estimations. These factors were; the total area of the building, the ratio of the typical floor area to the total area of the building, the ratio of ground floor area to the total area of the building, the number of floors, the console direction of the building, the foundation system of the building, the floor type of the building and the location of the core of the building. They further suggested that other factors can be considered in the model, such as the total height of the roof, the quality classification of the structural material, the ratio of the curtain wall to the total area, and the ratio of the secondary beams to the total number of beams.

Rafiei and Adeli [4] identified eight physical and financial (P&F) factors and nineteen economic variables and indexes (EV&Is). They further discussed that the EV&I variables could be estimated in a given time resolution, such as a week, a month, or a quarter depending on data availability. Typically, there is a temporal lag between changes in economic variables and their effects on construction costs. P&F factors employed in the model were; project locality defined in terms of zip codes, total floor area of the building, lot area, total preliminary estimated construction cost based on the prices at the beginning of the project, preliminary estimated construction cost based on the prices at the beginning of the project, equivalent preliminary estimated construction cost based on the prices at the

beginning of the project in a selected base year, duration of construction and price of the unit at the beginning of the project per square meter. EV&Is factors employed in the model were; the number of building permits issued, BSI for a preselected base year, WPI of building materials for the base year, total floor areas of building permits issued by the city or municipality, cumulative liquidity, private sector investment in new buildings, land price index for the base year, number of loans extended by banks in a time resolution, amount of loans extended by banks in a time resolution, interest rate for loan in a time resolution, average construction cost of buildings by private sector at the time of completion of construction, average of construction cost of buildings by private sector at the beginning of the construction, official exchange rate with respect to dollars, nonofficial (street market) exchange rate with respect to dollars, consumer price index (CPI) in the base year, CPI of housing, water, fuel, and power in the base year, stock market index, population of the city and gold price.

Cheng et al. [25] utilized six quantitative factors and four qualitative factors as the input to the EFHNN model. The quantitative factors were; floors underground, total floor area, floors aboveground, site area, number of households, and households in adjacent buildings. Qualitative factors were; soil condition, seismic zone, interior decoration, and electro-mechanical infrastructure.

Elmousalami [34] identified 17 input variables to develop a fuzzy genetic model to predict the cost of Field Canal Improvement Projects (FCIPs) which were; area served, average area served sections, total length of pipeline, equivalent diameter, duration, irrigation valves number, air and pressure relief valve number, sump, pump house, max discharge, electrical pump discharge, diesel pump discharge, orientation, construction year, rice existence, intake existence and ganabiaa canal.

Gunduz and Sahin [35] identified 12 input variables to estimate the cost of hydroelectric power plant projects in the early stages, which were, installed capacity, average discharge, project design discharge, project design head, length of the tunnel, length of the channel, length of energy transmission line, diameter of the penstock, length of the penstock, five-year occurrence flood discharge, hundred-year occurrence flood discharge and catchment area of the basin.

Location, construction year, building type, number of storeys, quality, and building technology were utilized as independent variables by Kouskoulas and Koehn [36] to explain variances in project cost. They contended that because the type of building has a significant impact on project costs, they should be assessed individually.

Juszczyk [37] identified 13 input variables for the developed artificial neural network. These variables were footprint, cubature, number of storeys, type of foundation, building structure, roof structure, number of structural segments of the building, number of elevator shafts, ground conditions, usable area of dwellings, usable area of commercial premises, usable area of underground garages and building finishing standard.

2.4.2. Economic Variables and Indexes Affecting Cost Estimation

2.4.2.1. Inflation Rate. The construction industry has a significant effect on the countries' GPD growth and overall economic growth. Countries' economic growth also depends on a number of other variables which inflation rate is one of them. Musarat et al. [38] assert that the inflation rate is getting neglected in most construction projects while budgeting, which causes cost overruns for the project as the building material price, cost of machinery, and labor wage change throughout the whole project timeline. The study also concludes the primary source of cost overrun is the change in material cost, which makes the final project cost fluctuate between 5% to 10% since the building materials impact 35% to 60% of overall construction cost.

Not only can the inflation rate affect the construction sector, but the construction industry's growth rate can also affect the economy of a country and the inflation rate [39]. Therefore, it can conclude that the effect is mutual with both the inflation rate and construction project costs.

Musarat et al. [38] also discuss the reasons for inflation. The leading causes are 1. Demand-Pull inflation, which happens because of excess aggregate demand, money, and credit boom, causes the economy closes to full capacity and a positive output gap. 2. Cost Push Inflation happens by growing wage expenses in the labor market, increasing raw

material costs from domestic and abroad suppliers, and rising import prices due to lower exchange rates. 3. Administered Prices that change as a result of changes in regulated prices, indirect taxes and subsidies, and environmental taxes.

Musarat et al. [38] discovered that inflation is one of the key factors determining the cost of construction projects. The study also believes that the inflation rate may construction material costs by up to 10% a year.

When estimating the cost of a project, construction project managers should take inflation into consideration in different phases of construction. When evaluating the cost of supplies, labor, and equipment, they must account for inflation as well. Failure to account for inflation can lead to underestimating the actual total cost of a project, resulting in budget overruns and delays of the project. As a result, calculating inflation before beginning a building project is critical to assuring its success and avoiding overruns.

Additionally, another effect of inflation is that it can bring higher interest rates, which also can increase the cost of financing for construction projects [40]. However, the relationship between inflation and construction costs is not always completely clear and might rely on other different factors, such as the type of construction project and the status of the economy [41]. As a result, while planning and budgeting construction projects, it is critical to examine the potential effect of inflation cost projections in different ways.

2.4.2.2. Exchange Rate. The currency exchange rates have huge effects on construction costs. For instance, an increase in the exchange rate can raise the cost of imported construction supplies, and raise the entire construction cost [42]. Second, fluctuations in exchange rates may impact the price of raw materials, which can affect the cost of construction projects [43]. Third, changes in exchange rates can affect labor costs, which in turn can affect building prices [44]. Finally, exchange rate changes can impact construction businesses' profitability, which can affect their ability to invest in new projects and, ultimately, construction costs [45].

2.4.2.3. GDP. GDP stands for Gross Domestic Product. It is a measure of the total value of goods and services which are produced within a country's borders in a specific time period, usually a year or a quarter [46]. GDP is considered one of the important indicators of a country's economic performance and is often used to compare the economic performance of different countries. It includes all final goods and services produced for consumption, government spending, investment, and net exports. The calculation of GDP takes into account the market value of all the goods and services that are produced, which means that both tangible goods and intangible services, such as education, healthcare, and software, are included [38]. GDP per capita represents inflation and living costs, and it is useful for comparing living standards between nations on the basis of GDP per capita at purchasing power parity (PPP). GDP can also be used to calculate counties' economic health [38].

The relationship between construction output and GDP has received great attention, and it is well recognized that the construction industry contributes a major portion to the gross domestic product (GDP). According to Yiu et al. [47], there is a positive correlation between GDP and various measures of construction output. However, the relationship between GDP and construction output is not always straightforward and can depend on various factors, such as the type of construction project and the state of the economy. The construction sector makes a significant contribution to the Gross Domestic Product (GDP) of many countries, including Nigeria and India [48] [49]. Many studies have concentrated on the simple relationship between the gross value (or value-added) of construction output and GDP (or GDP per capita). However, Yiu et al. [47] discuss the argument on the different proportions of construction output toward GDP in developing versus developed countries and how the share of construction output in GDP increased but ultimately declined in industrially advanced countries. The study suggests that the growth of construction output is a marginal diminishing function with the growth of real GDP in advanced industrialized cities and that financial cost is generally considered one of the factors affecting construction output. Therefore, the impact of GDP on construction costs is not as direct as the impact of inflation. Inflation can lead to an increase in the cost of construction materials and labor, which can increase construction costs [50]. Therefore, it is important to consider both GDP and inflation when estimating construction costs.

Windapo and Cattell [51] proposed a conceptual model of the factors affecting building material prices which result in the fluctuation in final project cost. According to the study, the factors that make the material cost deviation are; “Supply and Demand,” “Market Conditions,” “Transport Costs,” “Energy Costs,” “Raw Materials,” “Labor Costs,” “Inflation,” “Crude Oil Prices,” “Exchange Rates,” and “Import Duties.”

2.4.2.4. Consumer Price Index. The Consumer Price Index (CPI) can have an impact on construction costs. The CPI measures the average changes in the price paid by consumers for goods and services over time [52]. According to Shinghal and Nagar [53], diesel prices can be a price determinant for various commodities, including construction materials, and can impact the overall cost of a construction project. In India, the variation in project cost due to inflating diesel prices is calculated using the Wholesale Price Index (WPI) [53]. However, the CPI is a more accurate measure of inflation for consumers and can be used to estimate the impact of inflation on construction costs [52]. Additionally, Wong and Ng [54] found that the building cost index, which is closely related to construction costs, has explanatory value for predicting the Taiwan Producer Price Index (TPI) [54]. Therefore, it is essential to consider the CPI and other relevant price indices when estimating construction costs.

2.4.2.5. Minimum Wage. There is no direct evidence from the provided research results on how minimum wage affects construction costs. However, some of the results suggest that minimum wage increases can lead to an increase in prices [55]. Aaronson [55] found that restaurant prices generally rise with changes in the wage bill, and Lindner and Harasztosi [56] found that almost 75 percent of the minimum wage increment was paid by consumers. Additionally, König and Möller [57] found that minimum wage increases in the German construction sector led to an increase in wage growth for affected workers. However, the impact of minimum wage on construction costs is likely to be complex and context-dependent. Therefore, further research may be needed to fully understand the relationship between minimum wage and construction costs.

According to the provided research results, there is limited information on how minimum wage affects the construction industry in Turkey. However, some studies have analyzed the impact of minimum wage on the labor market in Turkey. For instance, Işık et

al. [58] found that the increment in the minimum wage had an economically considerable and statistically significant effect on wages, informality, and employment in Turkey. Additionally, Kemal and Kocaman [59] used the Autoregressive Distributed Lag (ARDL) model to study the impact of minimum wage on prices, growth, and unemployment, in Turkey. They found that minimum wage has a positive impact on prices and a negative impact on unemployment.

Moreover, Aydinli et al. [60] conducted a survey to determine the determinants of a fair wage and industry-wide wage policies in the Turkish construction industry. They found that unfair wage policies that fail to meet the needs of employees have been cited as one of the most important reasons for the high turnover rates of construction professionals.

In conclusion, while there is limited information on how minimum wage affects the construction industry in Turkey, studies have shown that minimum wage has a huge impact on the labor market in Turkey. Policymakers should consider the impact of minimum wage on the labor market and ensure that wage policies are fair and meet the needs of employees in the construction industry.

2.5. Overview of Quantitative Cost Modeling in the Construction Industry

In recent years, there has been an upsurge in the use of quantitative cos modeling in the construction industry. Quantitative algorithms can potentially increase productivity and efficiency in the construction industry in many ways. They can be used throughout the lifecycle of a construction project, from planning through execution and maintenance [61]. They can also be used to optimize the structural design based on sustainability and buildability requirements and automate tailored alterations to work surroundings and encourage healthy worker habits [62].

Construction schedules can be considered the most crucial step in managing a project. The scheduler’s background and experience can affect the process of project scheduling. Also, the lack of insufficient information can be problematic. To overcome these problems, automation in this process should be excessively focused on [63].

Kim and Soibelman [64] were the first ones in the construction industry that implement data mining processes in this industry. They proposed the KDD (Knowledge Discovery in Databases) principle that has the following steps: (1) identification of problems, (2) data preparation, (3) data mining, (4) data analysis, and (5) refinement process.

One of the main challenges to adopting quantitative algorithms in the construction industry is the lack of know-how and research. Another challenge is the lack of applied knowledge specific to construction activities, which makes risk analysis challenging [61]. One of the most undigitized industries in the world is the construction industry, which hinders its growth compared to other industries, such as manufacturing [65].

Despite the challenges, there are opportunities for the adoption of cost algorithms in the construction industry. They can improve construction processes, enhance competitiveness, and boost the industry's overall performance [62]. They can predict housing rates to improve building planning, construction, and sales [66] and can also improve safety on construction sites [62].

Of these quantitative methods, artificial intelligence and machine learning algorithms are of great importance. AI and ML algorithms have the potential to enhance the construction industry in many ways, such as increasing efficiency, productivity, and safety. However, the industry faces challenges in adopting AI due to the lack of know-how and research. The industry needs to invest in research and development to overcome these challenges and fully realize the potential of AI in construction.

Elmousalami [67] asserts that the construction cost estimating must be completed within a limited time. Therefore, the precise conceptual cost estimate is a challenging task for cost engineers, project managers, and decision-makers. Additionally, the research proposed that the main motivations to automate cost estimation are; 1. Quantity survey is a time and effort-consuming process 2. Cost estimation may be prone to human errors during assessment or personal judgment where biases and inaccuracies exist. 3. Highly accurate and reliable tool is required for Project managers and decision-makers to evaluate the conceptual cost of the proposed project.

Cheng et al. [28] claim that over 60% of the failures in construction projects are because of economic factors; therefore, a dynamic cash flow forecast is necessary to manage cash flow based on financial considerations. AI models are currently considered an alternative approach to controlling project costs and cash flows. They further discuss that, in order to develop AI models for project cost control, there are two essential tasks: to develop an appropriate approach and to collect and build a historical data pool with relevant parameters.

Günaydın and Doğan [27] discuss that AI models have invaluable benefits for design professionals at the early stages of the designing phase. To reduce construction costs, designers begin to use some cost estimation techniques and utilize their experience and data from their previous projects in their current projects. Therefore, utilizing all the previous data points and developing an AI model can be the solution to reduce construction costs. However, they discuss that the quantity and quality of the datasets are crucial to the precision and reliability of the conceptual cost models.

Traditional cost estimation has been done with the help of specifications and blueprints; however, comparative cost estimation, which is done by experts, assumes that the final cost and basic design variables of the projects have a linear relationship; however, this assumption itself is questionable. Artificial intelligence tools enable the development of the non-linear relationship and provide us the results that are more realistic and accurate. The major problems with conventional cost estimation strategies are the need for detailed project information, changes in the design parameters, uncertainties of project development, etc. AI can reduce uncertainties and provide designers with rough estimations during the design phase.

Selecting the materials and the design of a building structure can have a remarkable impact on its overall cost, roughly around 60% of the total construction cost. This cost consists of items such as structural system, mechanical system, finishings, etc. However, the relative weights of each of these items are different from one another according to the properties of each project, such as the type and usage of the building [27].

Rafiei and Adeli [4] point out that traditional cost estimation ignores some economic variables, such as the wholesale price index, liquidity, and building service index, which may result in variation in costs, especially in an unstable economic environment. By using machine learning models, these indexes can also be considered, which previously could not be included in the cost estimate. Consequently, the study encourages researchers around the world to collect such data in cities with a volatile and fluctuating economic environment. However, the model may not yield an accurate result in highly unstable economic environments since, for instance, government policies are not completely predictable. Although the model proposed in the study can be very beneficial for the final cost estimation of construction projects, it cannot be used for conceptual cost estimation since it requires preliminary cost estimation, and it did not take into account any structural information to predict the conceptual cost.

Elmousalami [34] argues that Building Information Modeling (BIM) provides data for cost estimation; however, predictive machine learning models can predict a project's cost on a macro level which is more beneficial. The study discusses that, at the conceptual stage of the construction projects, limited information is available; however, it is vital to perform the cost estimate at this stage since it can affect further decisions. It further argues that the most important step in cost model development is to identify the cost drivers. Many previous studies have identified key cost drivers by qualitative approaches such as questionnaire surveys, the Delphi method, or the analytical hierarchy method. So, it can be observed that identifying project features and key cost drivers need experts' opinion and question survey, which led to the model's uncertainty. Therefore, the study employs a fuzzy genetic model to decrease the uncertainties.

Gunduz and Sahin [35] conclude that cost estimation models for the HEPP project at the early stages assist the bidding process and should be done during the pre-construction phases since cost overruns and delays have been considered major uncertainties in construction projects. They further argue that the cost of HEPP (Hydro-Electrical Power Plant) projects is divided into three sections civil works cost, electro-mechanical cost, and other miscellaneous items and indirect cost. Indirect costs are considered 13% of the total cost of civil and electro-mechanical works. The cost escalation coefficient should be

considered according to the year of construction, and all costs should be changed based on this coefficient.

According to He et al. [5], the practical usage of conceptual cost estimation models in the industry is still limited in the construction industry. A global construction survey in 2016 found 36% of construction companies and 21% of owners use advanced data analysis for project estimation. There is a broad consensus among construction industry professionals that the traditional conceptual cost estimation is not satisfactory since it could be error-prone and involve manual efforts to compare a great number of previous samples. Some drawbacks of using data-driven cost estimation models are weak data management practices and the unavailability of historical data. Also, the cost of purchasing IT systems and the lack of data mining personnel are some other difficulties. Government data, online information, building contractors, project owners, and others might all provide data for conceptual cost estimation. The study emphasizes that it is critical to think of early cost estimating as a data mining process in order to use data-driven methods like multiple regression analysis (MRA) and neural networks (NNs). The study further argues discusses that the success of conceptual cost estimation mostly depends on the historical data sets of completed projects and that knowledge discovery from large data sets has become more important in the construction industry. Conceptual cost estimation has been utilized for planning, budgeting, and execution, and it is the basis of cost management. The reason that it's difficult to estimate costs accurately during the design phase is that there is not much information available as limited variables are available rather than aggregating cost items based on quantity take-off and unit prices.

Sonmez [68] believes that knowing the conceptual cost estimate in a project helps the owner, contractor, designer, etc., to determine the feasibility of a project, check the alternatives, and establish the initial budget. Range estimation was used in the study to quantify the uncertainties involved with the conceptual cost estimate. This method ultimately will improve the cost estimates in the early stages of projects.

As Jadid and Idrees [69] indicated, in construction projects, bidding documents play a vital role in winning the project. These documents consist of engineering drawings, specifications, and BoQs (Bills of Quantities), which aim to simplify the calculation process.

BoQ simplifies the calculation process in the construction industry by packaging the quantity of each project work required to complete the project. It is a detailed list of materials, labor, and equipment needed to complete a construction project. The BoQ can increase the accuracy of calculations without using assumptions that could increase the volume of work. The study also asserts that in the construction design phase, 70% of the life cycle cost of a product is determined. Pre-design cost estimates are done even before the completion of the project design and have some iterations along with the preliminary design. This estimation can be called a rough, preliminary, order of magnitude, or feasibility estimate. Errors in these estimates can range from -30% to 50%.

2.6. Types of Quantitative Algorithms and Their Efficiency in Cost Estimation

The use of machine learning algorithms in predicting and estimating construction costs in the early stages has been an interesting topic in recent years. Various strategies have been used to estimate construction costs accurately at an early stage when project information is scarce. Regression analysis, neural networks, support vector machines, and case-based reasoning are some of the machine-learning techniques that have been used.

Faghihi et al. [63] asserted that automation in budgeting and scheduling could be divided into some other sections: case-based reasoning and knowledge-based approaches, model-based approaches, GAs (Genetic Algorithm), expert systems, NN (Neural Networks), and other approaches. The study then provided a description of each approach and the method by which they can be helpful to the construction industry. CBR (Case-Based Reasoning) method can determine the specific data of formerly practiced situations. It investigates the database and finds a comparable situation to the present problem, and uses the data to find a solution for the current problem. Model-based approaches can be adopted for automation in scheduling based on the idea of adding models of construction methods. This method is based on extracting a project schedule from a 3D or 4D model of a building. Genetic algorithm models mimic the natural evolutionary process using a fitness function as a core metric and evolve into a series of optimized solutions. The expert system is a series of algorithms that imitates the human skills of decision-making. Artificial neural networks (ANN) are the models that imitate the function of the brains of animals which are able to do pattern recognition using all-or-none rules of the nerves. The study found the dominant

distribution of the genetic algorithm method as an optimization tool by researchers to solve scheduling problems; in other words, the method used more often to optimize project scheduling is the genetic algorithm. The contribution of these methods to the construction field can be divided into time, cost, and quality contribution.

Genetic algorithm is mostly optimizing the recourse and integrates the concept of the critical path method with the line of balance. Case-based reasoning and knowledge-based approaches mostly use similar historical data to solve the scheduling problem by developing a knowledge base. Model-based approaches try to establish some rules for automating the process of construction scheduling, such as spatial reasoning. Artificial neural networks attempt to perform cost optimization by using cost-schedule tradeoffs [63].

In a study by Park et al. [70], a two-level stacking heterogeneous ensemble algorithm combining RF (Random Forest), SVM (Support Vector Machine), and CatBoosting was suggested for the accurate cost estimation of a construction project in the early stage when the information necessary to predict construction cost is insufficient. The study showed that the proposed algorithm outperformed other machine learning algorithms in terms of accuracy.

Cho et al. [71] compared the multiple regression model with a neural network model in estimating construction costs for elementary schools. Kim et al. [72] also compared the multiple regression model with a neural network model, as well as a support vector machine model, in estimating school building construction costs. Hsiao et al. [73] proposed a neuro-fuzzy cost estimation model for semiconductor hookup construction. Juszczyk et al. [74] reported on the use of artificial neural networks, regression analysis, fuzzy logic, and case-based reasoning in cost estimation for construction projects. Barros et al. [75] developed a more accurate estimation technique for construction highway projects using ANNs. Tayefeh Hashemi et al. [76] conducted a systematic review of machine learning techniques for cost estimation in construction projects. Ali et al. [77] developed an integrative data intelligence model for construction cost estimation. Matel et al. [78] used ANNs for cost estimation of engineering services. Wu and Huang [3] studied construction cost estimation based on artificial intelligence technology.

Artificial neural networks have been utilized in various studies to estimate the cost of construction projects. ANNs have been found to be useful for predicting the cost of school building construction, highway projects [79], reinforced and prestressed concrete bridges [79], and sports fields [74]. ANNs have also been used to replicate past cost trends in highway construction and estimate future cost trends in Louisiana, USA [74].

Support vector machines (SVMs) have been used in some studies to estimate the quality of conceptual cost estimates [80]. SVMs have been found to be useful for assessing the quality of conceptual cost estimates of residential buildings [81] and for assessing the quality of conceptual cost estimates in general [80].

Regression analysis has also been used in several studies to estimate construction costs. Regression analysis has been found to be useful for predicting the cost of elementary school construction [82], and semiconductor hookup construction [76]. Regression analysis has also been used to predict the accuracy of cost estimates in the early stages of construction projects and to predict the cost of raising buildings in the UK [74].

Fuzzy logic algorithms have been used successfully in the construction industry to estimate construction costs. Monte Carlo simulation has been used extensively for addressing probabilistic uncertainty in range estimating for construction projects. However, subjective and linguistically expressed information results in added non-probabilistic uncertainty in construction management. Fuzzy logic has been utilized successfully for addressing such uncertainties in construction projects [83]. In practice, an approach that can handle both random and fuzzy uncertainties in a risk assessment model is necessary. Fuzzy logic plays an important role in many construction engineering and management applications, which are reviewed in the study conducted by Fayek [84]. Plebankiewicz et al. [85] presented three original models for planning the time, costs, and risk of a construction investment using fuzzy logic. The costing model uses fuzzy inference from CBR cases to increase the accuracy and correctness of the cost calculation performed for the investor. Therefore, fuzzy logic algorithms can be used to estimate construction costs by taking into account both random and fuzzy uncertainties.

Decision trees have also been used in some studies to estimate the cost of construction projects. Decision trees have been found to be useful for estimating the cost of adding a left-turn lane at signalized intersections [86]. Furthermore, decision tree algorithms could be used to estimate construction costs by analyzing various attributes of the project and making predictions based on those attributes. However, more research would be needed to determine the specific attributes that would be used and how they would be weighted in the decision tree algorithm.

The study conducted by Elmousalami [67] presents a comprehensive performance comparison for the AI model and evaluates different AI models' prediction accuracy. The result shows that out of 20 AI techniques, the most accurate and suitable method is XGBoost with 9.091% and 0.929 based on mean absolute percentage error and adjusted R^2, respectively.

Cheng et al. [28] used artificial intelligence approaches to predict cash flow in order to develop appropriate strategies for projects. For that matter, approaches such as K-means clustering, genetic algorithm fuzzy logic, and neural network were analyzed in their study. They used the k-means clustering approach to categorize similar projects and other approaches to develop an EFNIM. Genetic algorithm was used to optimize the membership functions of fuzzy logic and neural network parameters globally. It turned out that K-means clustering is the most straightforward approach for clustering data that starts with k centroids. After repeated iteration steps, every clustering feature can be determined. Although a perfect convergence cannot be guaranteed, cluster numbers can be set as a limitation for target convergence. K-means can be utilized in practical applications such as pattern recognition problems. The drawback of this approach is that it's required a specified setting to reach the optimal solution and its time consumption.

Chen and Tun [87] evaluated the usefulness of neural network technology for modeling non-mechanistic management. They pointed out that the process of the artificial neural network is similar to the human brain through imitation, which makes it a powerful tool. Furthermore, BPN (Back Propagation Network), which is also called a multi-layer feed-forward neural network, is a prevalent form of neural network model for a wide variety of tasks. The results indicate the usefulness of neural network technology for non-mechanistic

modeling in ecological research and management, and Artificial neural networks were recognized to be very helpful for non-linear ordination and visualization of ecological data in the proposed model.

Günaydın and Doğan [27] performed cost estimation in the early phases of building design processes for concrete structural systems of 4-8 story residential buildings with the help of the proposed ANN model. The average cost estimation accuracy of the developed tool by the study was reported to be 93%. Therefore, the developed cost estimation model has invaluable benefits for design professionals at the early stages of the designing phase.

Rafiei and Adeli [4] developed a model using machine learning approaches such as DBM-SoftMax, BPNN (Back Propagation Neural Network), and SVM. A sensitivity analysis was done to achieve the best result by considering EV&I (Economic variables and Indexes) factors in different time lags. They pointed out that some inputs of the model can be duration, delay status, labor, crew size, site operational requirements, and material estimations. One of the most important contributions of this paper is to pinpoint the importance of economic variables in accurate construction cost estimation.

Kim and Kim [88] developed a cost estimation model using GA and CBR. In the study, First, based on previous cases, a case base is developed, then the cases similar to the new case are retrieved by using similarity score S (multiplying the similarity of each attribute by its weight and summating all of them gives similarity score for each case). Then, the costs of retrieved cases are adapted to the new case by adjusting the costs by considering length, width, and construction cost index. By changing the important weights of attributes using a GA, the objective function, which is the error rate, is minimized. Then they used the model in an actual case and compared the error rates with other methods. Also, they performed a sensitivity analysis to identify the effects of the similarity evaluation criteria. The weight generation methods adopted by this paper will be available for a variety of facilities and can be beneficial in designing a more reliable cost estimation model at the early stage of any project. The study concluded that CBR is more practical compared to AHP (Analytic Hierarchy Process), gradient search methods, etc., for weight generation.

Cheng et al. [25] developed an Evolutionary Fuzzy Hybrid Neural Network (EFHNN) to improve the precision of conceptual cost estimation. First, they integrated neural networks and high-order neural networks into a hybrid neural network. Then, they used fuzzy logic in the high-order neural network to handle uncertainties which resulted in a fuzzy hybrid neural network. Lastly, they integrated a genetic algorithm with fuzzy logic and high order neural network to optimize the fuzzy hybrid neural network, which results in "EFHNN." According to the results, the proposed EFHNN can work as an accurate cost estimator during the early stages of construction projects. Overall, the construction cost estimator proposed by the paper was used in the preliminary stage before the detailed planning since estimating construction costs in the preliminary concept stage help engineers adjust the details and improve the chance of project success. The proposed EFHNN differs from the genetic algorithm and neural network approaches because of the unique hybrid layer connection types, different membership functions, and optimized parameters.

Elmousalami [34] developed a reliable parametric cost estimation model at the conceptual stage of the project. They conducted a literature review to identify the cost drivers of the model. Then they automatically determine the cost drivers among the ones they identified, based on quantitative approaches such as questionnaire survey, Delphi methods, and AHP, without just relying on experts' opinions. Then they collected historical cases to develop the model and integrated fuzzy logic with the model to maintain uncertainty in the predictions. Finally, they overcame the fuzzy logic problems by using hybrid fuzzy models. Based on the results, the fuzzy genetic model works better than the traditional fuzzy model in this methodology.

Gunduz and Sahin [35] developed two cost estimation models to estimate the cost of hydroelectric power plant projects in the early stages. They collected the data from fifty-four HEPP projects. Forty-nine projects were selected to develop the prediction model, and five projects were selected for validation purposes. The variables of a regression model are tested to select the ones that fit best with the available data. Then, correlation analysis was done to investigate the linear relationship between every two variables. A backward elimination method was used for developing the initial cost model, and regression analysis was done by a software package called Minitab. Then, an artificial neural network was utilized as an alternative methodology to develop the cost model by a software called Neural Power. The

neural network is composed of three layers which are the input layer, hidden layer, and output layer. These layers consist of neurons by which layers are connected to each other. Also, a weighted function exists, which is called a transfer function. The accuracy of the developed model was tested by five testing projects. This was done by using the Mean Absolute Percentage error (MAPE). The cost estimation model developed in this paper by regression model had a 9.94% prediction accuracy, and the neural network model had a 5.04% accuracy.

He et al. [5] investigated the frequency of methods used for developing estimate models, including regressions, ANNs, CBR, GA, simulation, expert system, and fuzzy theory. Among those methods, regression models have been used most frequently. Regression is a simple method for determining the impact of independent project characteristics on building costs. When linear regression models are insufficient to match the data set, higher-order regression models are used.

Sonmez [68] focused on the advantages and disadvantages of the current methods of conceptual cost estimation. Additionally, it investigated the usage of regression, neural networks, and range estimation techniques for conceptual cost estimation. Three regression models and two neural network models were developed in the study to examine the need for non-linear terms in the regression model. The level of uncertainty for the estimates was then investigated by prediction intervals. It was concluded that regression models use fewer parameters compared to neural networks; however, regression analysis requires making decisions about the class of relationships, e.g., linear, quadratic, etc. On the other hand, in neural network analysis, there is no need to put much effort into deciding the type of relations or the probability distribution of the variables. In this study, R^2 (coefficient of determination) and P (significance level) were used to measure the level of contribution of each variable. Furthermore, the study used two error measures; one was the MSE (Mean Squared Error), and the other one was the MAPE. The results show that the neural network model provided a better fit in terms of closeness to the data, but the regression model provided a more accurate prediction. Therefore, the regression model proposed in this study could be considered an adequate model.

2.7. Limitations of Cost Estimation Algorithms

Cost estimate models have limits in the construction industry. One of the barriers to adoption is a lack of data quality and quantity, which might impair algorithm accuracy [89]. Another restriction is the intricacy of the construction process, which makes accurate modeling and prediction difficult [90]. Furthermore, the absence of industry standardization can make building algorithms that can be used across multiple projects and settings challenging. The application of quantitative data analysis techniques necessitates considerable computer resources, which can be costly and time-consuming. Furthermore, interpreting the outcomes of these algorithms can be difficult, and the algorithms may not always provide apparent explanations. Therefore, it can be concluded that the most significant barriers adopt these quantitative techniques in the construction industry is a lack of knowledge and research. Another difficulty is a lack of applied knowledge unique to building activities, which makes risk assessments difficult [61]. The construction sector is one of the least digital in the world, which limits its growth compared to other industries, such as manufacturing [65].

Poor data maintenance practices and a lack of historical data are two major downsides of using data-driven cost-estimating methods. Other problems include the high cost of IT equipment and labor scarcity in data mining.

The accuracy of cost estimate models has long been debated, and a variety of factors can influence it. In the construction business, for example, the availability of materials data affects the accuracy of the models in predicting and estimating building costs in the early phases [91]. Construction cost databases are often smaller and more diversified than datasets from other disciplines, making it challenging to develop suitable predictive methods utilizing tiny materials datasets. As a result, while training quantitative models for forecasting and estimating building costs in the early phases, it is critical to consider data availability.

He et al. [5] have identified that the obstacles to implementing advanced conceptual cost estimation in the industry may lie in the steps of data preparation and model evaluation which again point out the difficulty of construction cost data collection and technical knowledge of quantitative model development.

2.8. Identification of the Literature Gap

As discussed in this chapter, conceptual cost estimation for construction projects is crucial for feasibility study purposes as well as projects' success. Therefore, many previous studies addressed the factors affecting these cost estimates aiming to present highly accurate models for cost prediction. However, the factors presented by most of the previous studies were mostly design factors that were used as the cost drivers of the models. There are many structural design variables affecting a construction project's cost. These variables include many items like the choice of the structural frame, whether or not to use high-quality material, whether or not no use elevator, the design of reinforced concrete materials, the use of composite frames, storey's height, floor system, and exterior finishing type. The initial construction cost of various structural systems is a significant factor in the selection of the structural system for residential buildings [13] and not many studies have investigated the effective structural factors in predicting the construction cost at early stages. In this respect, the first research question is as follows.

i. Which structural factors have the most effect on conceptual cost estimation models? Günaydın and Doğan [27] used design factors such as the total area of the building, the ratio of the typical floor area to total area, the ratio of the ground floor to the total area, the number of floors, console direction, foundation system, floor type and location of the core of the building. Considering the unstable economic situation of some developing countries, the current study aims to take into account also the economic factors along with the detected design characteristics.

Berk and Biçen [1] studied the relationship between GDP growth and the construction sector in emerging countries but did not consider other economic factors. Semenova and Vitkova [2] studied the relationship between macroeconomic factors of the economic health of the Czech Republic and Spain. The study emphasizes the importance of paying attention to all possible internal and external factors in the management of construction companies and the decision-making process in the Czech Republic and Spain. The factors investigated by the study were GDP, Inflation, and Unemployment Rate. However, the current study asserts that besides the factor proposed by Semenova and Vitkova [2] other economic factors can also affect cost estimation; such as the Average Rediscount Rate, Average Net Minimum

Wage, Construction Turnover, CPI (Consumer Price Index), and Exchange Rate as well. Consequently, the current study seeks to investigate the effect of these economic factors by developing an early cost estimate model by reviewing the economic factors in Turkey.

Rafiei and Adeli [4] pointed out the importance of integrating economic variables and indexes in addition to the design characteristics of projects. They integrated factors such as wholesale price index, exchange rates, CPI, land price index, and gold price to develop a cost model after detailed design characteristics of the project are determined. The main difference between the current study with the study done by Rafiei and Adeli [4] is that the current study aims to develop an early cost model before the detailed designs are available. This helps the project owners to do the project feasibility study and ease the decision-making process.

ii. What is the impact of economic factors on the cost estimation models' accuracy? This study focuses on the macroeconomic factors affecting conceptual cost estimation and aims to close the literature gap by comprehensively investigating the effectiveness of economic variables on the construction conceptual cost estimate along with the design characteristics of the project. The role of economic variables gets more important when it comes to predicting the construction cost in a volatile economy. According to the results of the study done by Bayır and Orak [92], throughout the period 2004-2020, overall prices in Turkey were heavily dependent on the exchange rates, and they are also dependent on the global variables but with a narrow margin, comparing to the domestic variables. The current study focuses on the macroeconomic variables of Turkey within the past two decades and investigates how they can be beneficial in predicting construction project costs.

3. RESEARCH METHODOLOGY

This chapter goes on to the methods that were employed in this study to propose a method for conceptual cost estimation for construction projects. The methodology utilized to complete this study used historical data analysis as the foundation for creating a connection between the primary factors influencing the cost of building projects to produce estimations for new projects.

This chapter discusses the study strategy and design, factors influencing the cost of building projects, and the data gathering, analysis, and evaluation method.

3.1. Research Framework

Figure 3.1. demonstrates the framework of the study. First and foremost, a comprehensive literature review identified the factors affecting the construction project's cost. These factors can be structural design factors, as most of the previous studies addressed these variables and their relationship with construction cost, and finally, the equation that can be used to estimate the total cost based on these design variables. However, some studies failed to consider also economic variables affecting construction costs.

Economic variables influence construction costs for several reasons, particularly in developing countries like Turkey. A country's economic development, stability, and performance fluctuate significantly in such an economically volatile country. Exchange and interest rate shocks impact government dept service and can lead to deterioration in different industries such as construction. Changes in global market circumstances, political instability, disasters, and government economic policies can all contribute to this matter. This uncertain environment can impact construction projects, such as increasing construction costs and delaying project completion. Because of the reasons stated previously, planning a building project in a developing country requires consideration country's economy. For this reason, a variety of economic measures are used to measure financial volatility, such as the Consumer Price Index, Gross Domestic Product, and Purchasing Managers Index.

In a developing country with a volatile economy, citizens and companies may experience serious issues. For instance, lower consumer confidence, a high unemployment rate, a high poverty level, and a lower number of foreign investors are some of the outcomes of economic volatility. Some countries in the Middle East and North Africa are exposed to fiscal risks and are at risk of economic volatility which makes them to be more vulnerable than other parts of the world [8].

Governments may use policies such as monetary and fiscal policies, trade policies, and incentives for investors to minimize this economic instability. These policies can help the economy's stability and mitigate the impacts of financial volatility on the population. Additionally, businesses and people can insulate themselves from economic instability by diversifying their investments, investing in foreign markets, and decreasing debt, resulting in fewer construction projects. Therefore, being able to predict and estimate a construction project cost is relatively more important in economically unstable countries.

As discussed to this point, economic variables can significantly impact a construction project's costs. Factors such as changes in the price of raw materials, labor, regulatory fees, energy cost, interest rates, market conditions, and demands based on economic conditions can affect the cost of construction projects. The size, structure type, usage, site condition, and building environment can also influence construction costs, which are considered design variables in this study. In addition, the availability and condition of construction machinery can affect the economic situation of construction companies. Other factors that can cause cost overruns in construction projects include inadequate equipment, financial complexities, labor shortage, alteration in material specification, lack of communication between parties, and delay in decision-making. Macroeconomic indicators have also been studied as potential predictors of construction material prices [93]. To improve the economics of construction projects, it is crucial to consider the life cycle cost theory and analyze the construction and operation costs comprehensively [94].

After selecting the design and economic variables as the input variables, data collection was done to collect the cost data of construction projects. Sixty-four data points were gathered by different contractors aiming to develop a conceptual cost estimation model using both design and economic variables. Consequently, using the data points, cost models

were developed using quantitative methods such as regression, neural network, and decision tree algorithms. IBM-SPSS Modeler 18.0 was used to analyze the cost models and select the best model based on the least mean absolute percentage error.

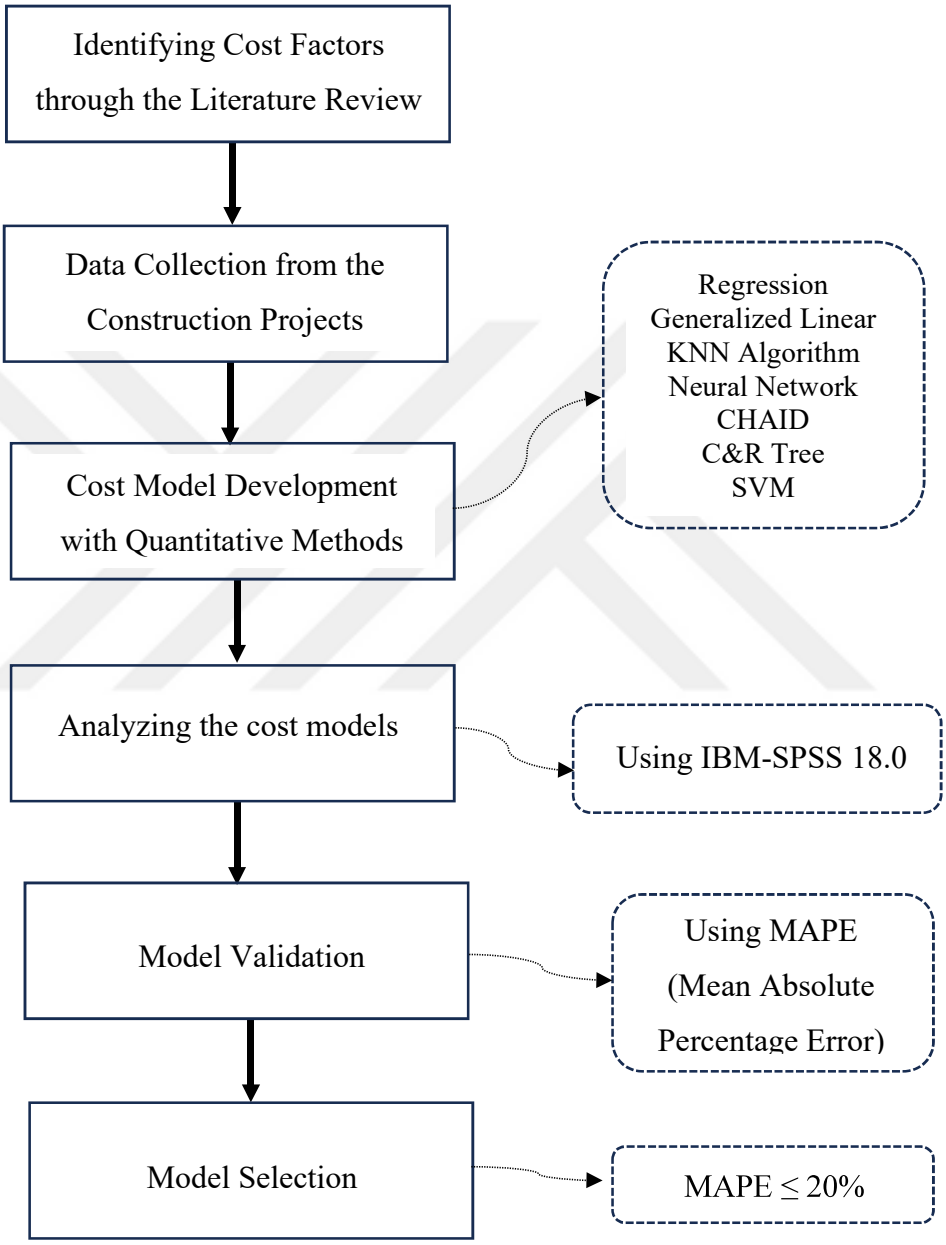

Figure 3.1. Research methodology flowchart.

3.2. Input Variables

Table 3.1. shows the design variables used in this study as the input of the construction cost estimate model. These variables were identified based on the literature review conducted in Chapter 2.

Table 3.2. shows the economic variables of the study, which were utilized to develop a more accurate construction cost estimate. This study emphasizes that design variables are not solely sufficient to predict the project cost. This magnifies when the construction project is progressing in a developing country because of the economic uncertainties that the stakeholders may experience.

These variables are then categorized into different measurement categories. Each variable could be either continuous, categorical, flag, nominal, ordinal or without a specific type. “Number of storeys” was identified as a continuous variable since it was a range of numbers, “Project location” was identified as a nominal variable since it was the name of the cities, “Construction year” was identified as an ordinal variable since it contains numbers with sequence, “Construction duration” detected as a continuous variable, “Area” was detected as a continuous variable, “Foundation system” was identified as a flag variable since it only had two categories, “Floor system” was detected as a nominal variable, “Framing type” was detected a flag variable because of having only two categories, “Contract type” considered as a nominal variable, “Exterior finishing type” was regarded as a nominal variable and “Usage of building” was considered as a nominal variable. Furthermore, all the economic factors, which include “Construction turnover,” “Average discount rate,” “Average net minimum wage,” “CPI,” “GDP per capita,” “Inflation rate,” and “Exchange rate,” were considered as continuous variables

Table 3.1. Design factors considered as the cost driver.

Cost driver	Definition	Value or unit
$X_{1,1}$	Number of storeys	Unit
$X_{1,2}$	Project location	İstanbul Antalya Ankara Diyarbakır İzmir Balıkesir Mardin
$X_{1,3}$	Total area	m^2
$X_{1,4}$	Foundation system	Isolated Mat
$X_{1,5}$	Floor system	Reinforced Concrete Pan Slab Steel Systems
$X_{1,6}$	Framing type	Reinforced Concrete structural steel
$X_{1,7}$	Contract type	Lump-sum Unit Price Cost Plus
$X_{1,8}$	Exterior finishing type	Curtain Wall & Stone Curtain Wall Curtain Wall & Plaster Plaster Brick Coating Aluminum Façade Precast Stone EIFS EIFS & Glass Glass & Stone EIFS & Stone EIFS & Composite EIFS & Plaster
$X_{1,9}$	Usage of the building	Residential Commercial Educational Hospital Residential & Commercial Dormitory

Table 3.2. Economic factors considered as the cost driver.

Cost driver	Definition	Value or unit
$X_{2,1}$	Construction Turnover	Billion TL
$X_{2,2}$	Average Rediscount Rate	%
$X_{2,3}$	Average Net Minimum Wage	TL
$X_{2,4}$	CPI	Index (2003=100)
$X_{2,5}$	GDP Per Capita	TL
$X_{2,6}$	Inflation Rate	%
$X_{2,7}$	USD Exchange Rate	TL

Table 3.2. indicates the economic factors which were used as the cost driver of the model. The information regarding each variable shows the economic condition of Turkey in different ways, and each was extracted from TurkStat. TurkStat, formerly known as the Turkish Statistical Institute, is Turkey's national statistical organization. Its primary objective is to gather, compile, evaluate, and distribute official statistics and information regarding the Turkish economy, society, and environment. As it is shown in Table 3.2., "Construction Turnover" is one of the economic factors used as the cost driver of the model. It is the total value of construction work in a country, which is Turkey, in this research, completed within a specific time period, usually a year. It includes the value of new construction projects, which in this study, it is based on Billion Turkish Liras. "Average Rediscount Rate" is another economic variable showing in percentage and yearly-based. "Average Net Minimum Wage" is the average yearly wage based on the Turkish Lira. "CPI" is the consumer price index compared to the year 2003. "GDP Per Capita" represents the monetary value of goods and services produced in a country, on average, per person based on the Turkish Lira. "Inflation Rate" was presented yearly-based and as a percentage. Finally, the "USD Exchange Rate" was introduced as the US dollar price in Turkish Lira.

3.3. Hypotheses Development

Identifying the conceptual cost model's input plays a vital role in the model's accuracy. Therefore, many studies have tried to integrate different cost drivers with the cost model. These input variables can be technical specifications, environmental conditions, or other factors. Selecting the materials, and specifying some design characteristics such as structural system, mechanical system, finishing, etc. can significantly impact the conceptual cost estimate [27].

This study identified effective structural design variables that can be observed in Table 3.1. including nine primary design specifications that are decided before the project starts. It is assumed that these structural design variables have a direct effect on the conceptual cost model.

- H1a: “Number of storeys” positively correlates with the conceptual cost estimate model.
- H1b: “Project location” has a significant relation with the conceptual cost estimate model.
- H1c: “Total area” positively correlates with the conceptual cost estimate model.
- H1d: “Foundation system” has a significant relation with the conceptual cost estimate model.
- H1e: “Floor system” has a significant relation with the conceptual cost estimate model.
- H1f: “Framing type” has a significant relation with the conceptual cost estimate model.
- H1g: “Contract type” has a significant relation with the conceptual cost estimate model.
- H1h: “Exterior finishing type” has a significant relation with the conceptual cost estimate model.
- H1i: “Usage of the building” has a significant relation with the conceptual cost estimate model.

The fluctuation of economic variables in a country, such as wholesale price index, liquidity, and building service index, can heavily affect the conceptual cost estimate and may result in cost variations [4]. These variations are hard to predict and often ignored in traditional cost estimation techniques. Companies should try to adopt appropriate medium-term fiscal frameworks to build up buffers. These frameworks should result in lessening the budgetary impact of risk and improve reliance on future shocks [8]. Considering macroeconomic variables while estimating and budgeting a construction shock can mitigate the fiscal risks and elevate the projects’ success rate.

This study emphasizes the importance of considering economic variables while estimating construction costs, especially in countries with volatile economies like Turkey. Economic variables and indexes are expected to have positive relationships with the

construction projects' cost; therefore, it is expected that these economic variables have a direct effect on the conceptual cost model. Annual construction turnover in a country can be an important index to show how active is the construction industry in that country, and can be effective for the project's cost prediction. Therefore, the next hypothesis will be as follows.

- H2a: "Construction Turnover" positively correlates with the conceptual cost estimate model.

Rafiei and Adeli [4] utilized economic factors such as interest rate, CPI, exchange rate, worker wage, and inflation rate to develop a cost model for construction projects after the design stage. Therefore, the next hypotheses are as follows.

- H2b: "Average Rediscount Rate" positively correlates with the conceptual cost estimate model.
- H2c: "Average Net Minimum Wage" positively correlates with the conceptual cost estimate model.
- H2d: "CPI" positively correlates with the conceptual cost estimate model.

The relationship between construction output and GDP has received great attention recently. Yildiz [95] and Semenova and Vitkova [2] assert that there is a positive relationship between GDP and the construction industry. Yiu et al. [47] discuss that the relationship between GDP and construction output is not always straightforward and can depend on the state of the economy. As this study focuses on Turkey's economy and the relationship between economic factors and construction outputs, the next hypothesis is as follows.

- H2e: "GDP Per Capita" positively correlates with the conceptual cost estimate model.

Musarat et al. [38] argue that the impact of the inflation rate is getting neglected in most construction projects while budgeting which results in cost overruns. The inflation rate can also affect the conceptual cost estimation at the early stages of the projects. Fluctuations in the inflation rate can be a result of the change in foreign currencies changes. Therefore,

these two factors, the inflation rate and exchange rate, can play important roles while estimating construction projects' costs. As a result of the previous discussion, the next hypotheses are as follows.

- H2f: "Inflation Rate" positively correlates with the conceptual cost estimate model.
- H2g: "USD Exchange Rate" positively correlates with the conceptual cost estimate model.

Finally, it is essential to emphasize that the design factors are not enough when it comes to predicting a construction project's cost, as is considered in many previous studies. In today's economy, where severe fluctuation can be observed, structural design variables cannot be sufficient to perform the project feasibility study and cost estimation. Therefore, the third and final hypothesis will be as follows:

- H3: The conceptual cost model's accuracy drastically increases when economic factors are integrated into the model.

3.4. Data Collection

Sixty-four data points were gathered from the construction projects that have been finished. Contractors and project managers were asked to fill out the data sheet consisting of 9 design factors shown in Table 3.1. along with the construction year and duration. The economic factors were further added to the dataset using the construction year.

Appendix A shows the 64 data points filled in by the participants. This data shows the construction cost of projects in 7 cities of Istanbul, Antalya, Ankara, Diyarbakir, Izmir, Baliksehir, and Mardin. The distribution of the location data is shown in Figure 3.2. The foundation system was either Isolated or Mat; however, most of the projects had the Mat foundation system as can be observed in Figure 3.3. The floor system was either Reinforced Concrete, Pan Slab, or Steel System. The majority of the data was regarding the Reinforced Concrete floor system as Figure 3.4 suggests. The framing Type of the projects was either Reinforced Concrete or Structural Steel, and the majority of the projects had Reinforced Concrete framing (Figure 3.5.). Contract Type was either Lum-sum, Unit price, or cost Plus;

the majority of the projects had Lum-sum contracts (Figure 3.6.). Exterior Finishing Type was the most diverse variable in the collected data; therefore, it is classified into 14 groups, Curtain Wall & Stone, Curtain Wall, Curtain Wall & Plaster, Plaster, Brick Coating, Aluminum Façade, Precast, Stone, EIFS, EIFS & Glass, Glass & Stone, EIFS & Stone, EIFS & Composite, and EIFS & Plaster (Figure 3.7.). Usage of the building was either Residential, Commercial, Educational, Hospital, Residential & Commercial, or Dormitory (Figure 3.8.).

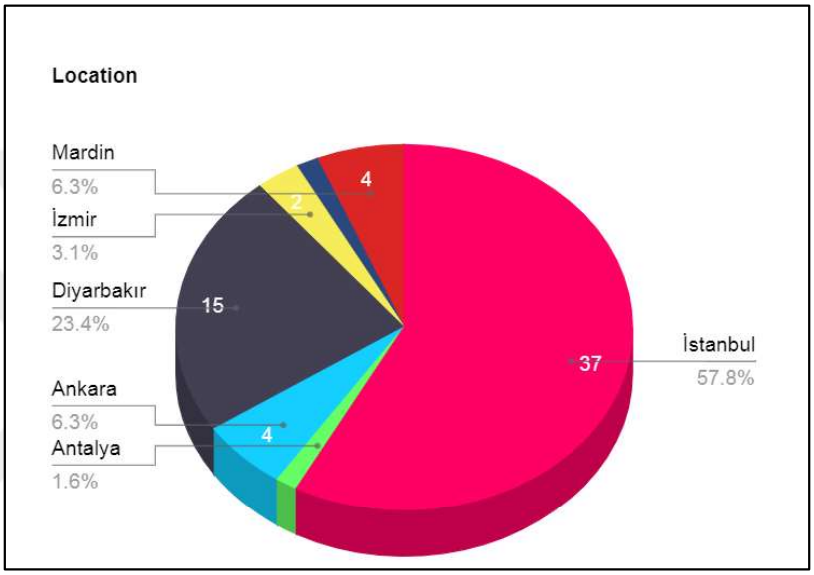

Figure 3.2. Location of the projects.

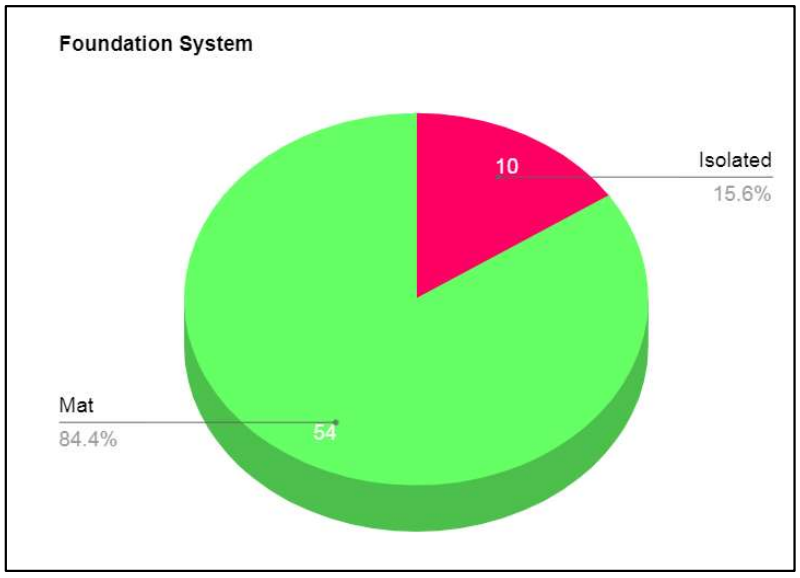

Figure 3.3. Foundation system of the projects.

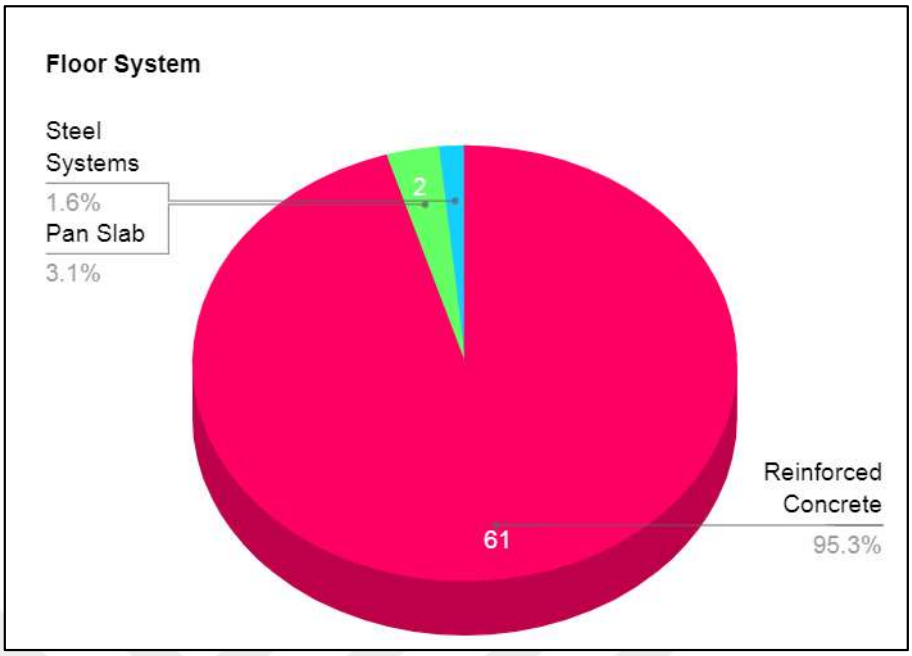

Figure 3.4. Floor system of the projects.

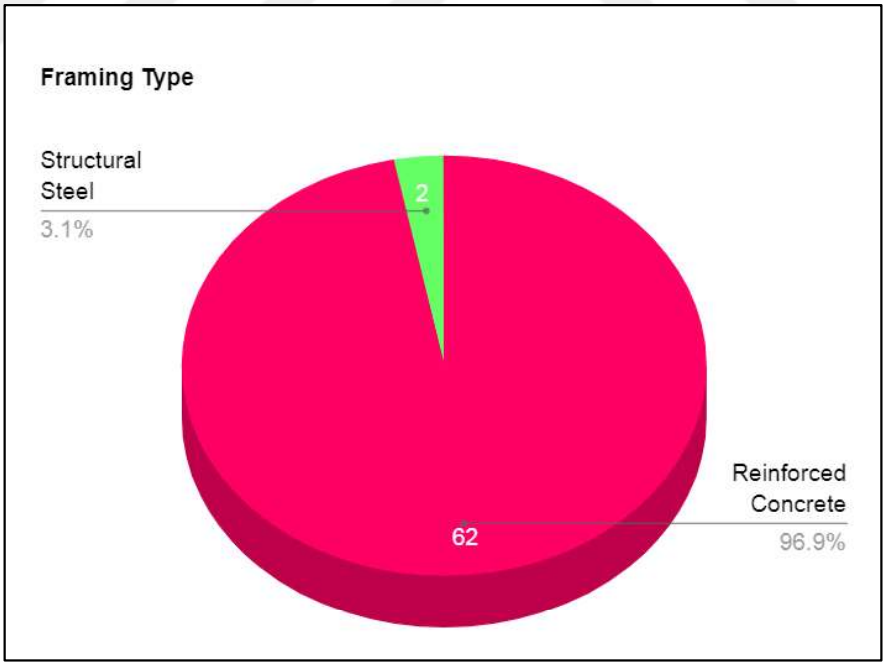

Figure 3.5. Framing type of the projects.

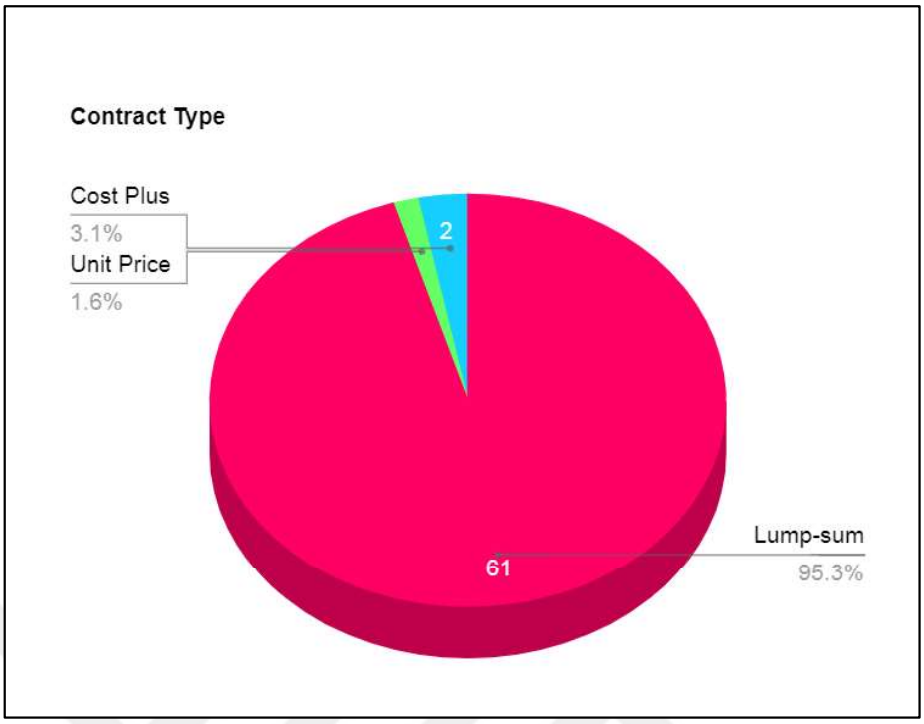

Figure 3.6. Contract type of the projects.

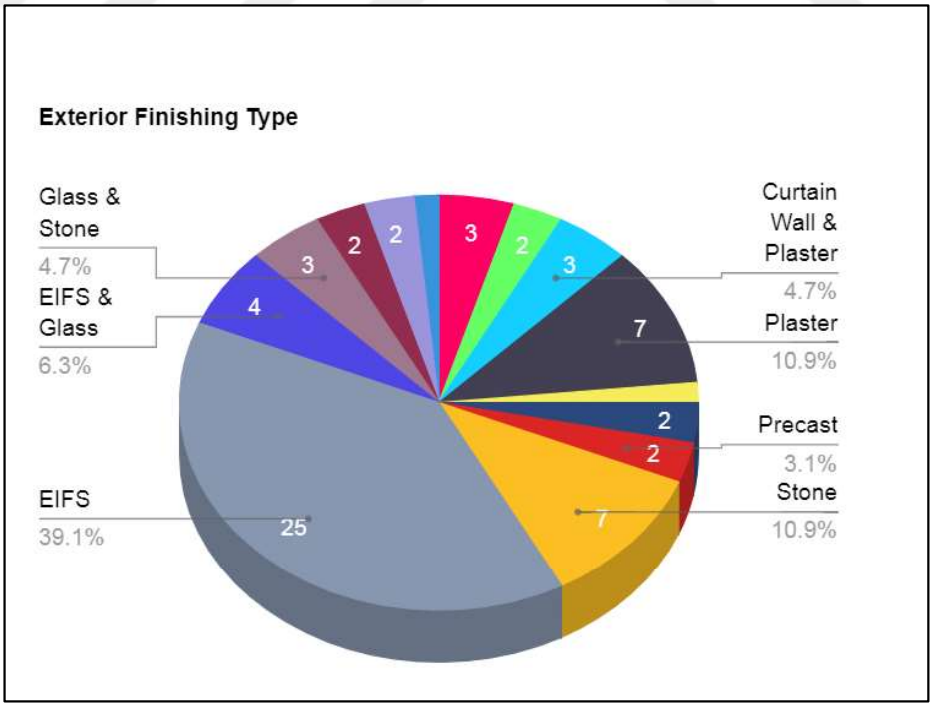

Figure 3.7. Exterior finishing type of the projects.

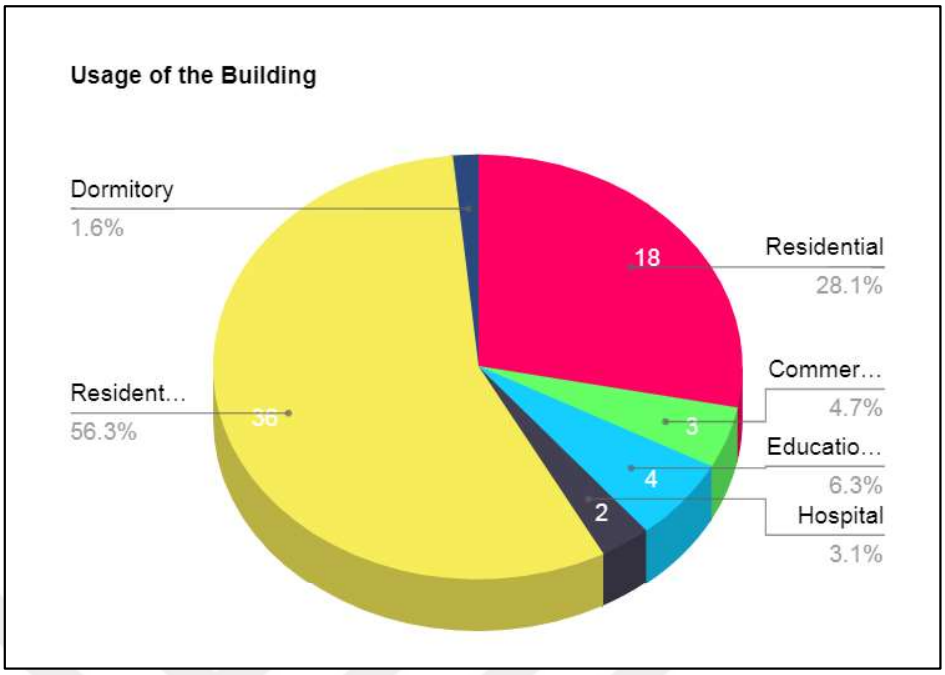

Figure 3.8. Usage of the buildings in the projects.

3.5. Quantitative Data Modeling Algorithms

3.5.1. Neural Network

A “Neural Network” (NN) which is occasionally referred to as an “Artificial Neural Network” (ANN), is a computational model inspired by the structure and functionality of biological neurons. They are used to model complex interactions between inputs and outputs as non-linear statistical data modeling techniques. NNs are reliable tools for developing parametric cost-estimating models due to their outstanding performance in modeling relationships between inputs and outputs. In this study, intelligent models were produced using NNs in addition to other models developed by other artificial intelligence algorithms.

In 1943, McCulloch and Pitts proposed a neuron model as a switch that could receive input from other neurons depending on the total weighted input. This is considered to be the first neural network [96].

Neural Networks can store trainable information and allow them to correctly identify and classify new and similar patterns. Therefore, NNs can be used in a variety of fields, such

as power systems, pattern recognition, signal processing, medicine, planning, military systems, and control and search [97].

Neural Network focuses on storing information in a structure composed of many neurons and imitates the human brain in terms of learning. Neural Network models learn from the previous data and generalize the solutions for future applications and are utilized for solving non-linear or ill-structured problems [28].

Basu et al. [97] classified the Neural Network as Single Layer Perceptron, Multilayer Perceptron (MLP), Radial-Basis Function Network (RBF), Hopfield Network, Boltzmann Machine, Self-Organization Map (SOM), Modular Network, Support Vector Machine, Bayesian Network, Probabilistic Graphical Models and Hidden Markov Models.

The neural network is usually made up of three layers which are the input layer, the hidden layer, and the output layer. These layers are made up of neurons that connect the layers together. There is also a weighted function known as a transfer function. As can be observed in Figure 3.9., The neurons in the Neural Network input are considered the dendrites of a biological neuron, and they are in charge of receiving information from the environment or from other neurons. The Neural Network input layer receives data from the outside environment. The neurons in the hidden layers connect the input and output layers, similar to a biological neuron's cell body, and are in charge of transferring processed information to other neurons. The hidden layers are in charge of transmitting information to neurons in the output layer. The neurons in the output layer function as the axons of biological neurons, delivering the processed information to other neurons or directly to the environment. Most of the time, in ANN models, the outputs of each neuron in the output layer directly go to the outside.

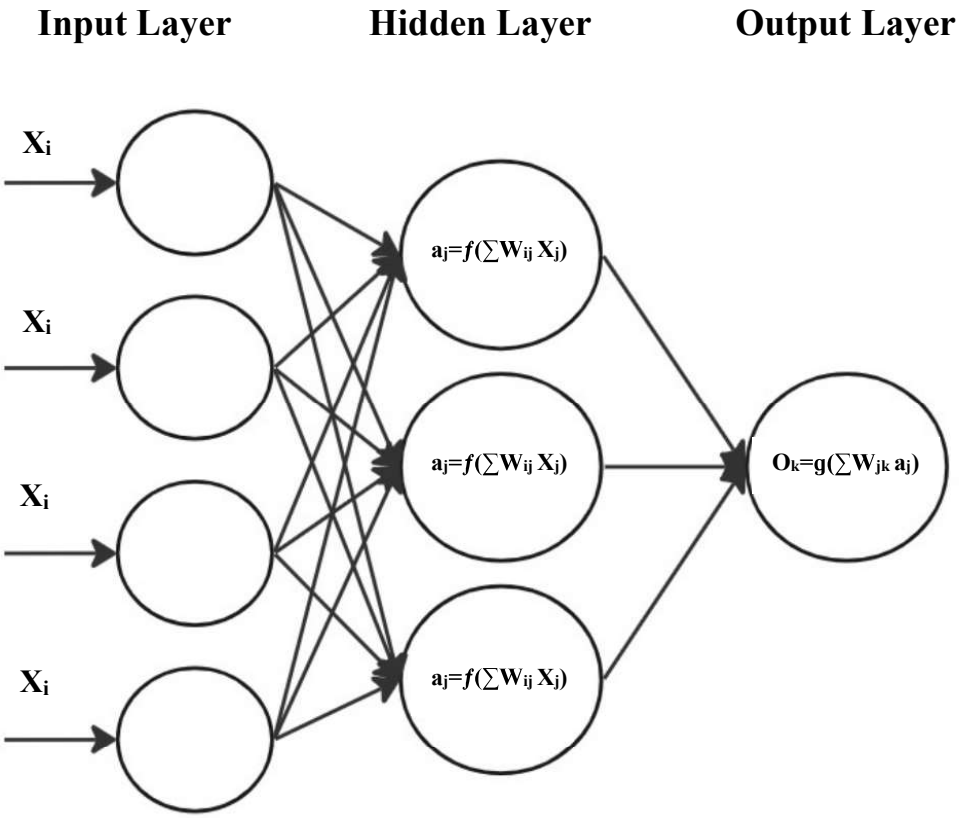

Figure 3.9. Neural network model.

The transfer function should be selected based on the target variable while developing a Neural Network Model. The transfer function can be considered in many forms, such as Hyperbolic Tangent Sigmoid, which maps the input values to a range between -1 and 1. A semi-linear function is another transfer function that acts as a linear function for positive values, and it is zero for negative values. A step transfer function is linear with saturation. The sigmoidal transfer function transfers the values to a range between 0 and 1 [98].

To develop the Neural Network model, IBM SPSS Modeler 18.0 was used. The model has got one hidden layer and eight neutrons. Another alternative method to establish the Neural Network is using a phyton code or any other software language.

3.5.2. Decision Tree

Decision Trees are models that are used to visualize and help decision-making in a sequential way. They consist of three types of nodes: decision nodes, chance nodes, and

terminal nodes. Figure 3.10. shows a schematic view of the decision tree algorithm. Decision trees are commonly used for data mining for classification systems based on multiple covariates and for developing prediction algorithms.

Decision Tree algorithms are considered supervised learning algorithms that use a white box model approach and classify the data in a hierarchical way. The algorithm is created in a way that can effectively reduce the decision tree level and enhance efficiency.

Decision Tree algorithms can be used in a variety of fields, such as data extraction, search engine optimization, and prediction. Choosing the right algorithm for the target task is necessary to get the most accuracy and efficiency out of all decision tree algorithms.

There are many types of decision tree algorithms each is used for a specific aim. “Regression Tree” is one of the types of decision trees used for predicting a continuous variable. “Classification Tree” is mostly used for predicting categorical values. “Ensemble Decision Tree” is another type of decision tree that is usually combined with other machine learning algorithms in order to elevate the performance. “CART,” also known as “C&R,” is the classification and regression tree that can do both classification and regression tasks.

“CHAID” (Chi-squared Automatic Interaction Detection” is mostly used to determine the best split for each node and is used when there is categorical data. “ID3” (Iterative Dichotomiser 3) is another decision tree algorithm that is vastly used for data classification. “C4.5” is another decision tree algorithm that extends the ID3 and can handle both categorical and continuous data. “CatBoost” is another decision tree algorithm that is designed for handling large categorical datasets [99].

This study has utilized C&R and CHAID decision tree algorithms in order to predict the effect of design and economic variables on the construction estimate and to present a construction cost estimation model based on the input data.

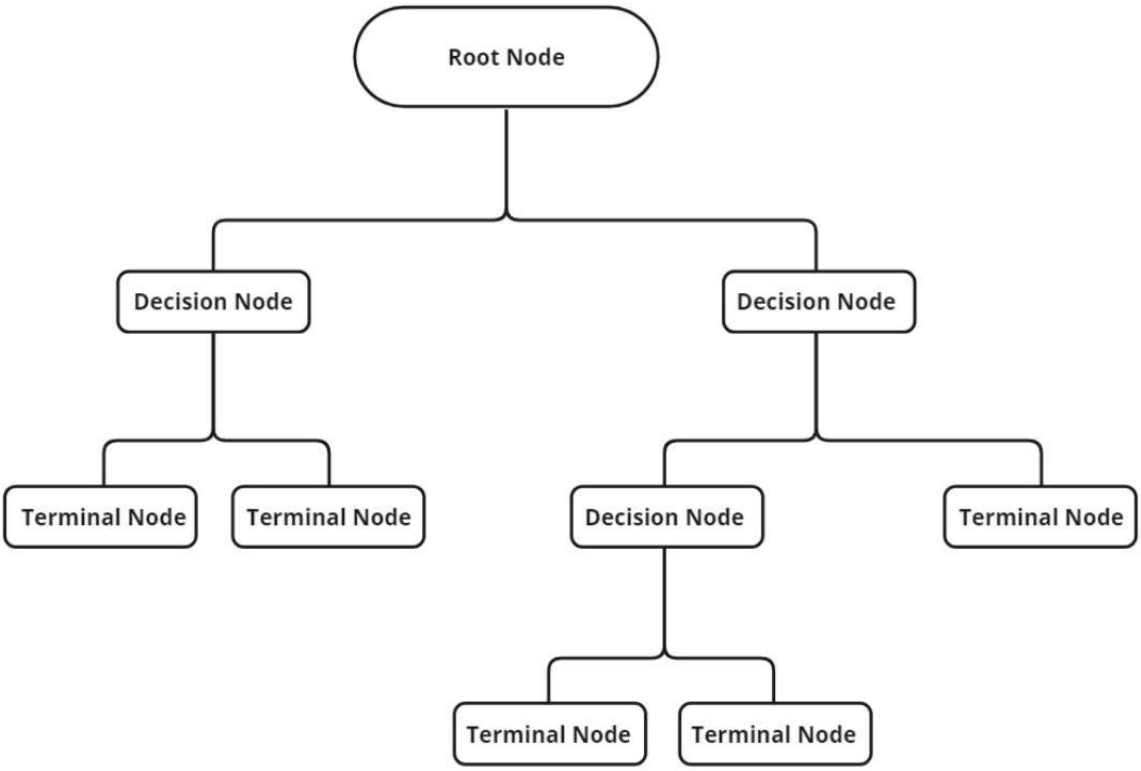

Figure 3.10. Decision tree model.

3.5.3. Regression

A regression model is another type of machine learning model used for predicting continuous values based on the input values. Regression is considered a supervised learning technique and involves training the model on an input dataset and using the trained model to predict the output for new input data. The goal of the regression model is to find the best-fit line that can indicate the output value in the most accurate way. The regression model is classified as simple linear regression, multiple linear regression, and polynomial regression [100].

Simple linear regression in which the model has a single independent variable is expressed as

$$y = \beta_0 + \beta_1 x + \varepsilon, \tag{3.1}$$

where y is the dependent variable, x is the independent variable, and β is the coefficient and Ɛ is the error. In other words, a simple linear regression model is used to develop the

relationship between two variables and assumes that there is a linear relationship between the two variables [101].

A multiple linear regression (MLR) which is used in this study, is expressed as

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m + \varepsilon, \tag{3.2}$$

where y is the dependent variable, β is the coefficient, x is the dependent variable and Ɛ is the error. Multiple linear regression is an extension of simple linear regression and is mostly used to develop a model between a dependent variable and two or more independent variables. The model assumes that there is a linear relationship between the independent variables and the output, which is a dependent variable. It estimates the best coefficients in order to get the best fit of the data [102].

The forward selection method is one of the methods of multiple linear regression that starts with no variables in the model and selects the variables by comparison criterion to improve model performance. The backward elimination method is another regression method that is the opposite of the forward selection method. In this method, the variables are selected at the beginning of the model, and then the unimportant independent variables are eliminated based on the t-test for each variable. The stepwise method is an extension of the forward selection method, and it is a combination of the forward and backward methods [34].

A polynomial regression model uses a polynomial function to analyze the relationship between a dependent variable and an independent variable and is expressed as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x^2 \cdots + \beta_h x^h + \varepsilon, \tag{3.3}$$

where y is the dependent variable, β is the coefficient, x is the dependent variable and Ɛ is the error In a polynomial regression model, the relationship between the variables is not linear; instead, it can be approximated by a polynomial function of degree n, that n is a positive integer. The coefficients are then estimated in a way that the model results in the best fit. The degree of polynomial regression is chosen based on the relationship complexity between the dependent variable and independent variables [103].

R and R^2 are the measures of the regression model's performance which show the strength of the relationship between the independent variables and the dependent variable. R is called the correlation coefficient, which ranges from -1 and 1 and shows the strength and direction of the linear relationship between the variables. When R is, -1 shows a perfect negative correlation. When it is 0, it indicates no correlation; when it is 1 it shows a perfect positive correlation. $R^{2,}$ on the other hand, is the coefficient of determination and measures the data fit. It shows how well the regression line fits the input data and represents the proportion of the dependent variable which is explained by the independent variables. R^2 ranges from 0 to 1; 0 shows that the regression line does not fit the data, and 1 indicates that the regression line fits the data well [104].

3.6. Evaluation of the Algorithms

To Evaluate the developed models, MAPE (Mean Absolute Percentage Error) is utilized to estimate each model's error and rank them based on that. MAPE is a metric that is commonly used to evaluate the accuracy of forecasting models. It measures the average percentage difference between the actual values and the predicted values. MAPE is calculated by taking the absolute difference between the actual and estimated values, dividing it by the actual values, and then taking the average by summing it all up and dividing by the number of values in the series. MAPE is presented as a percentage and can range from 0% to infinity [105].

MAPE has some limitations, such as a lack of ability to handle zero or negative values, which results in infinite or undefined MAPE values, and being sensitive to extreme values, which results in distortion of the accuracy. Therefore, it is often used with other metrics to evaluate the performance of the forecasting model [106]. MAPE is expressed as

$$M = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A-F}{A}\right|, \tag{3.4}$$

where M is the mean absolute percentage error, n is the number of times the summation iteration happens, A is the actual value and F is the forecasted value

With the required data to develop the cost model, IBM SPSS Modeler 18.0 was utilized to employ machine learning algorithms by dividing the data set into two training and testing

sets. Based on the size of the database, it was decided to split the dataset by 80% and 20% for training and testing sets, respectively. This results in having 52 data points for developing and training the machine learning model and 12 data points for testing and eventually evaluating the machine learning models.

IBM SPSS 18.0 can also detect and generate missing values, which was not the case in current research since, during the data collection, all the participants could successfully provide the project cost data.

Figure 3.11. illustrates the model setup on IBM SPSS 18.0. The setup was defined in a way that the output models are ranked based on their relative error. IBM SPSS then developed and examined 12 models based on 12 algorithms. As can be observed in Figure 3.12., these algorithms were "Regression," "Generalized Linear," "KNN Algorithm," "Linear-AS," "LSVM," "Random Trees," "SVM," "Tree-AS," "Linear," "CHAID," "C&R Tree," and "Neural Network." The top four algorithms were then chosen based on the lower relative error of the algorithms.

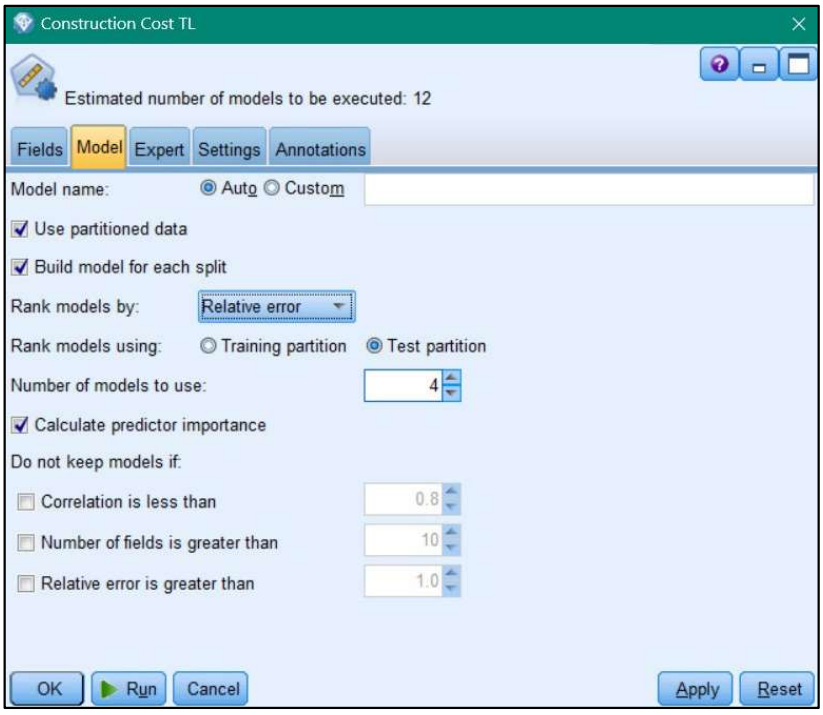

Figure 3.11. IBM SPSS 18.0 model setting.

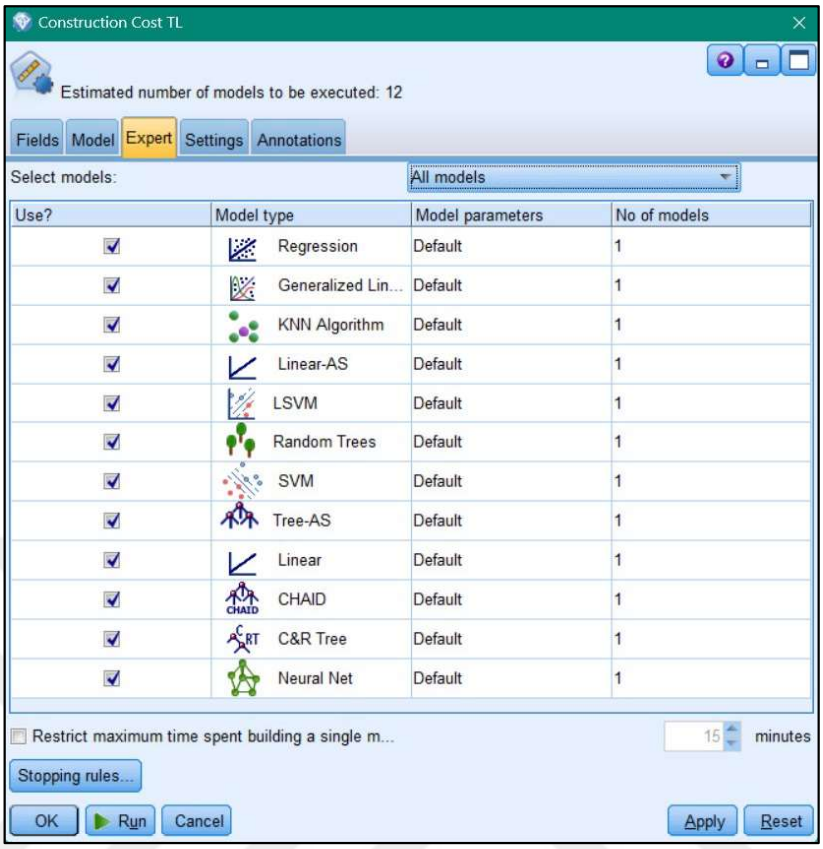

Figure 3.12. Algorithms selected in IBM SPSS 18.0.

4. RESULTS AND FINDINGS

In this Chapter, the result and findings of the quantitative analysis of the construction costs, using IBM SPSS Modeler 18.0, have been presented.

4.1. Relative Performance of Algorithms

Using the Auto Numeric Modeling tool in IBM SPSS Modeler 18.0, the best algorithms were identified through a number of Machine Learning Algorithms presented by the software. After inputting the 64 data points of construction costs with 18 cost drivers, measurement types and value ranges were defined for each cost driver. Each of the cost drivers can be continuous, nominal, ordinal, and flag values based on their definition. Eighteen cost drivers were assigned as the input to the model, and construction cost (TL) was given as the target of the model.

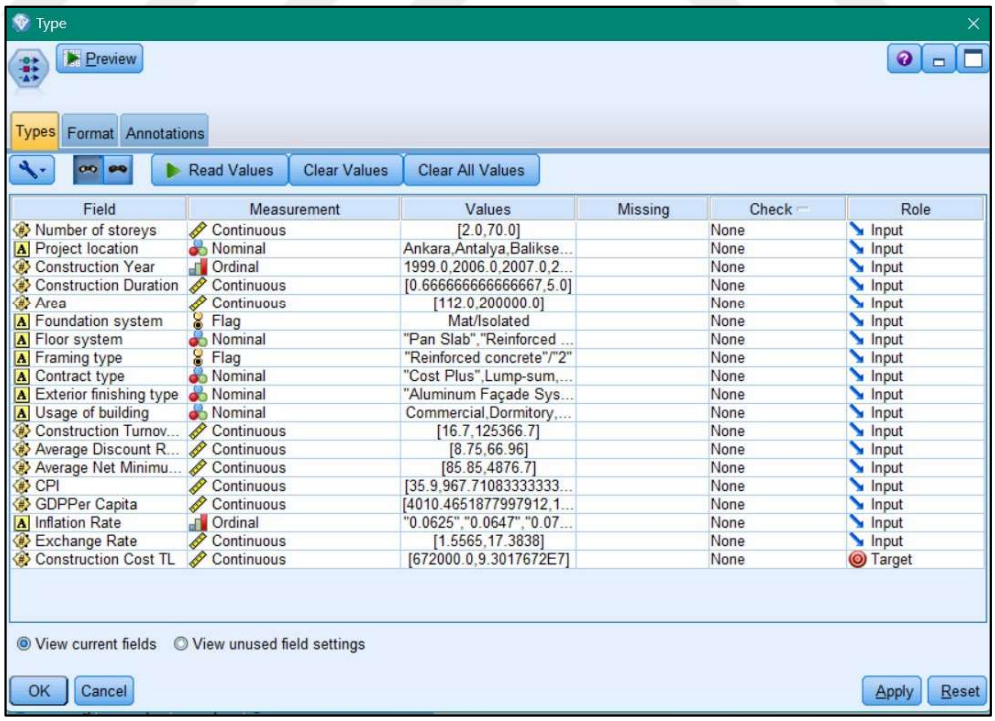

Figure 4.1. Inputs and target.

Twelve quantitative data models were selected in the software to analyze the data and develop a model with each of them. These ML models are as follows:

- Regression
- Generalized Linear
- KNN Algorithm
- Linear-AS
- LSVM
- Random Trees
- SVM
- Tree-AS
- Linear
- CHAID
- C&R Tree
- Neural Network

After applying the Auto Numeric Modeling to the data, the top four models were selected by IBM SPSS 18.0. Figure 4.1. shows these ML models: Neural Network, with a Relative Error of 3.8%; C&R Tree, with the Relative Error of 14.5%; Regression, with the Relative Error of 32.4% and CHAID, with the Relative Error of 68.6%.

By knowing the top four models with the best performance for the input data, the main analysis will be according to these models. Meaning that four different construction cost models will be developed, and the data will be analyzed based on them.

4.2. Result of Each Estimation Model

4.2.1. Neural Network

Knowing that Neural Network is one of the top four models to develop our cost estimate model (Figure 4.2.), detailed analysis has been performed to estimate the model

accuracy and error. By that means, the Neural Network model has been selected out of the modeling methods in IBM SPSS Modeler 18.0 (Figure 4.3.).

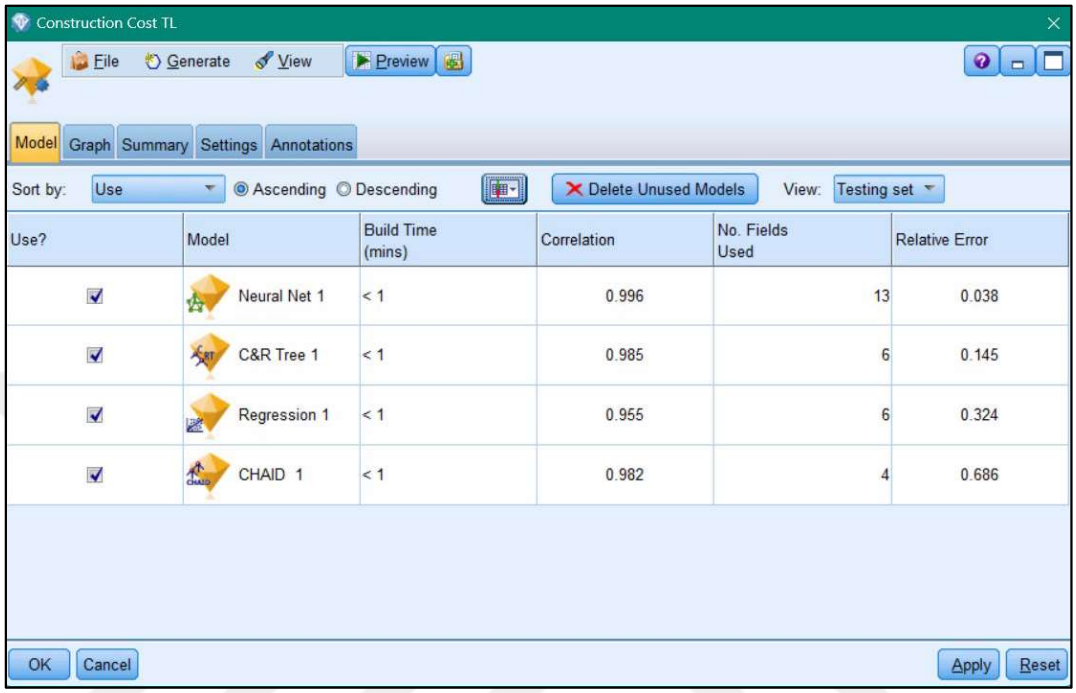

Figure 4.2. Top four cost models and their relative errors.

The Neural Network model was built based on Multi-layer Perception (MLP). The term “multi-layer” points out that the MLP has more than one layer of neurons, meaning hidden layers exist, as opposed to a single-layer perceptron with only one layer of neurons. The additional layers allow the MLP to learn more complex patterns and relationships in the data. Therefore, the neural network was built using one hidden layer with eight neurons (Figure 4.5). The stopping rule was assigned as “Error cannot be further decreased.” The result reveals an accuracy of 97.5% for the Neural Network model of cost estimation (Figure 4.4.).

Figure 4.6 shows the impact of each cost driver (Model Input) on the outcome (Construction cost). According to the result, “Number of storeys” has the most impact with the importance of 21%. Then it comes “Area” with the importance of 13%. After that, “Inflation Rate” has the most impact, with an importance of 7%; then “Average Discount Rate” with an importance of 6%; “Exterior finishing type” with an importance of 6%; “GDP Per Capita” with the importance of 6%; “CPI” with the importance of 6%; “Average net

minimum wage" with the importance of 5%; "Usage of the building" with importance of 5% and "Construction Year" with importance of 5%.

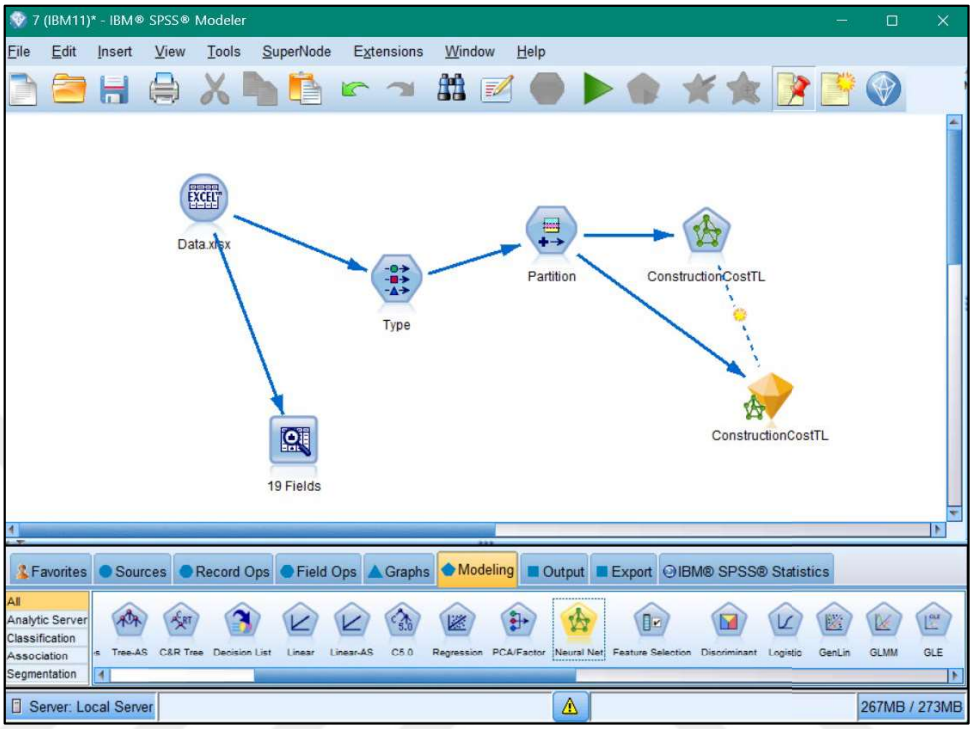

Figure 4.3. Developing neural network model.

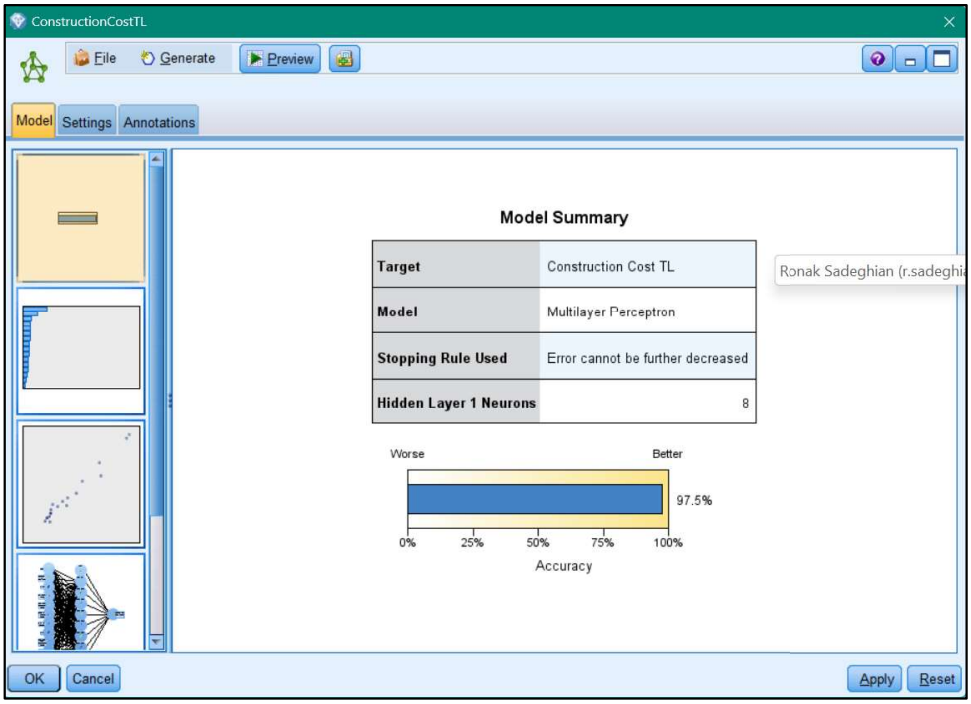

Figure 4.4. Neural network model summary.

By exporting the table, including all the values forecasted by the Neural Network model (Table A.1), .MAPE is calculated for the Neural Network model using the equation (3.4), which is $M = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A-F}{A}\right| = 16.7418\%$. In this equation M is the mean absolute percentage error, n is the number of times the summation iteration happens, A is the actual value and F is the forecasted value.

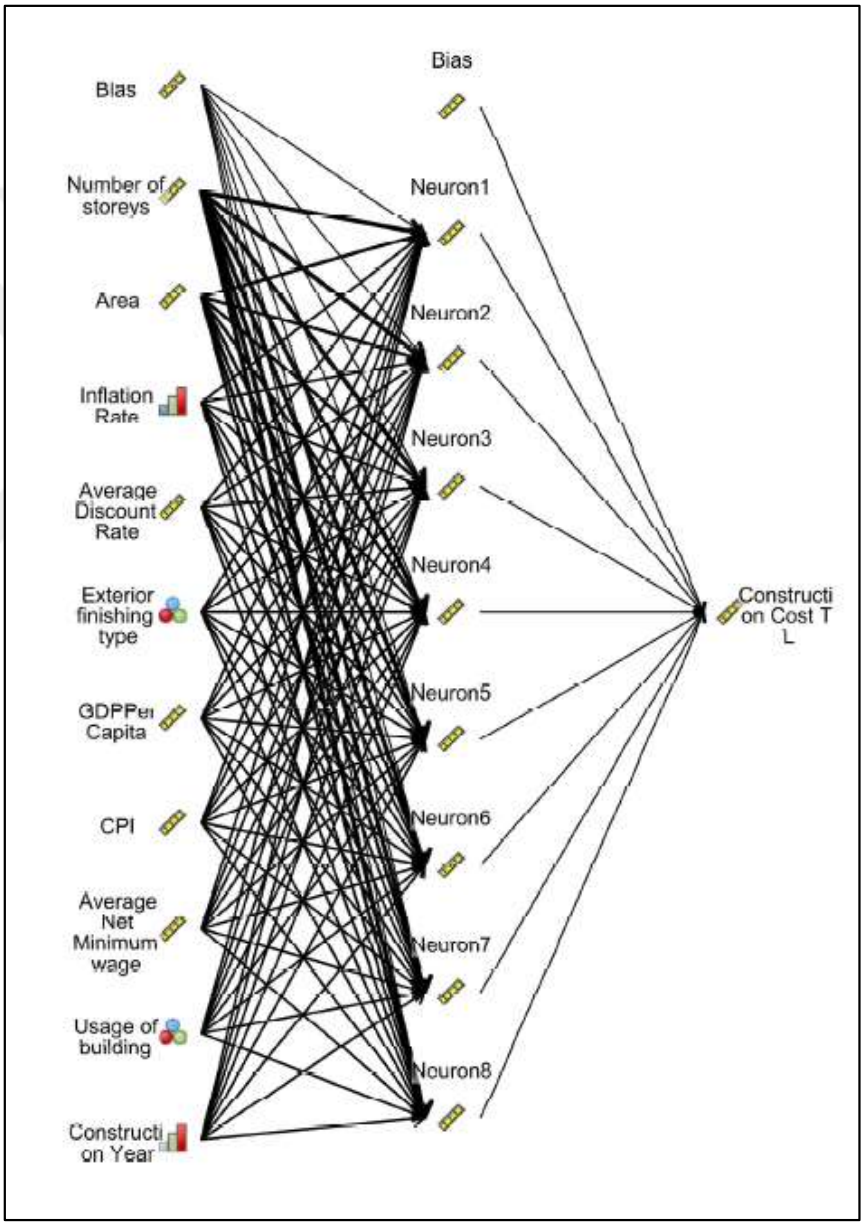

Figure 4.5. Neural network.

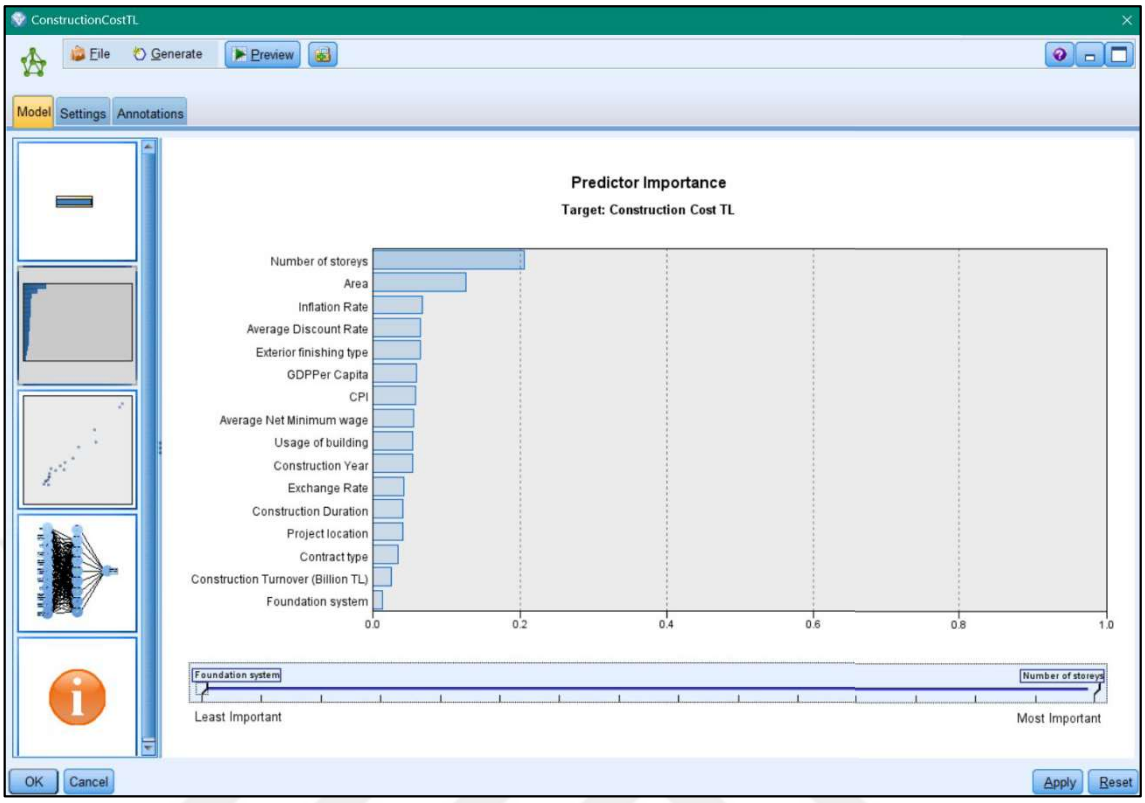

Figure 4.6. Cost drivers' impact on the NN model.

4.2.2. C&R Tree

The second-best model for forecasting construction cost using our data points was the C&R Tree (Classification & Regression Tree) model, which is a type of decision tree algorithm.

The C&R Tree is a machine learning algorithm that is used to build prediction models from data [107]. The models are created by iteratively splitting the data and fitting it with a basic prediction model within each partition. As a result, the partitioning can be graphically depicted as a decision tree [107]. The C&R Tree algorithm is a type of classification and regression tree algorithm that is used for dependent variables that have a finite number of values, with prediction error evaluated in terms of misclassification cost (Loh, 2011). C&R Tree, along with C4.5 and CART, is one of the most extensively used methods for generating decision trees; it is also used to solve classification and regression problems [108].

Both C&R Tree and Decision Tree are machine learning algorithms used to build prediction models from a dataset. They are not, however, identical, and there are minor differences between them. C&R Tree, as previously stated, is a classification and regression tree approach used for dependent variables with a finite number of unordered values, with prediction error evaluated in terms of misclassification cost. Decision Tree, on the other hand, is a data mining approach used for data classification and forecasting. Gulati et al. [109] describe the Decision Tree as a white box model that classifies data in a hierarchical framework. Both techniques use a tree structure to represent the model; however, C&R Tree works best for classification and regression problems with categorical dependent variables, whereas Decision Tree is a more generic technique for data mining and classification [109]. So, there are some differences in the critical aspects of the algorithm [110].

By conducting the C&R Tree algorithm to develop the construction cost estimate model (Figure 4.7.), not only could we get the cost predictions using our cost drivers, but also, we could determine which of the cost drivers should play the input role in the model.

According to Figure 4.8., the C&R Tree algorithm deploys 11 cost drivers and eliminates the other 7. The inputs that have been eliminated are; “Exterior finishing type,” “Usage of the building,” “Project location,” “Contract type,” “Foundation type,” “Floor System,” and “Framing type.”

The importance rates of those 11 cost drivers used in the C&R Tree model are as follows; “Area” has the most effect with the importance rate of 14%, “Average net minimum wage” has the importance rate of 9%, “Average discount rate” has the importance rate of 9%; “Construction turnover” has the importance rate of 9%, “Construction duration” has the importance rate of 9%, “GDP Per Capita” has the importance rate of 9%, “CPI” has the importance rate of 9%, “Construction year” has the importance rate of 9%, “Exchange rate” has the importance rate of 9%, “Inflation rate” has the importance rate of 9%, and “Number of storeys” has the importance rate of 9%,

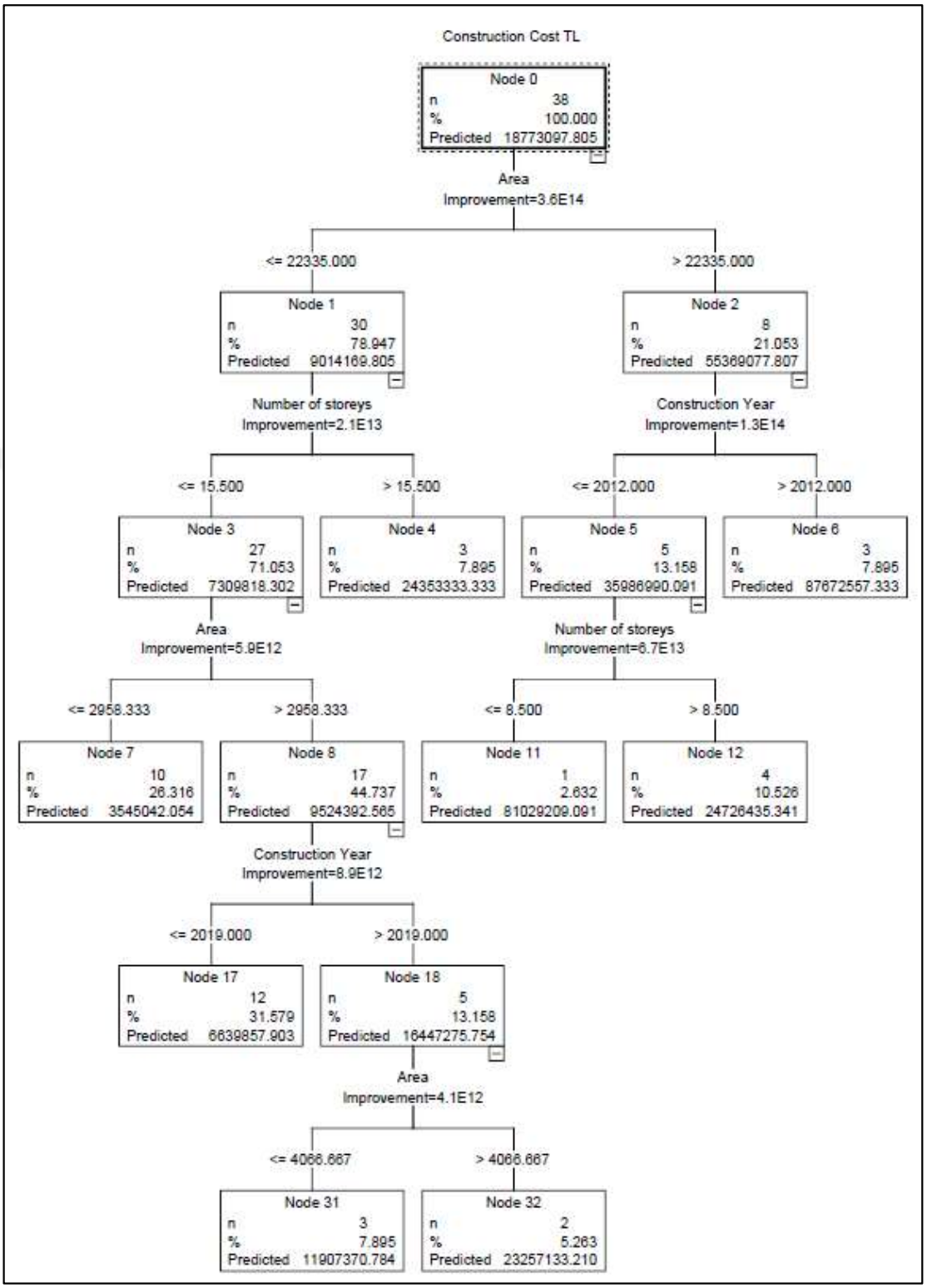

Figure 4.7. C&R tree.

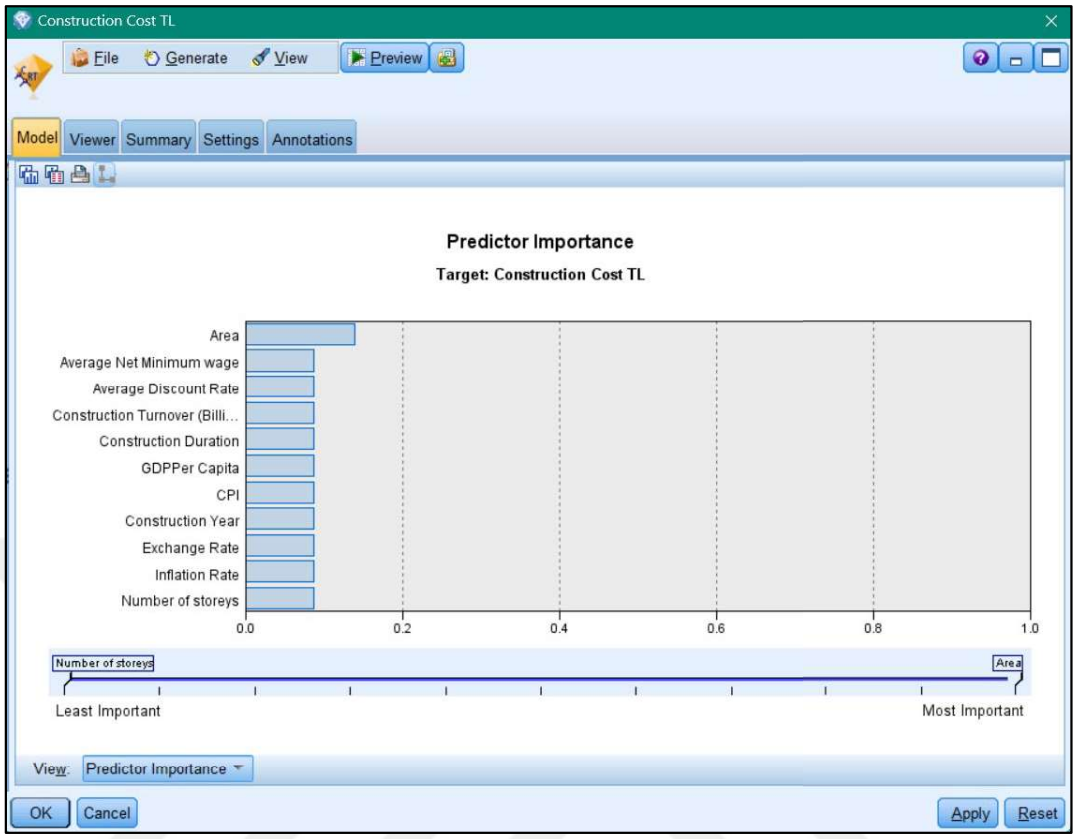

Figure 4.8. Cost Drivers' impact on the C&R tree model.

MAPE is calculated for the C&R tree model using the equation (3.4), which is $M = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A-F}{A}\right| = 24.2108\%$. In this equation M is the mean absolute percentage error, n is the number of times the summation iteration happens, A is the actual value and F is the forecasted value.

4.2.3. Regression

In order to make the regression model take into account all cost drivers, including the ones which are not numerical and instead are categorical, flag, nominal, or ordinal, we need to perform "Label Encoding."

Label encoding is a machine learning technique that converts categorical data into numerical data [111]. Label encoding is mostly used in regression algorithms to turn categorical data into numerical variables that may be employed as model predictors [112]. Label encoding gives each category of a categorical variable a distinct number value. Dummy variables are a type of label encoding in which a categorical variable is transformed

into a set of binary variables and are often used to represent categorical variables in regression models [111].

Table 4.1. demonstrates the Label Encoding performed for the categorical cost drivers of the data points. Each of the categorical cost drivers was given a range of numbers so that they can be used in the general regression formula, and regression coefficients can be assigned to them.

Table 4.1. Label encoding of categorical variables.

Cost Driver / Encoding	Location	Foundation	Floor system	Framing Type	Contract Type	Exterior Finishing Type	Building Usage
1	İstanbul	Isolated	Reinforced Concrete	Reinforced Concrete	Lump-sum	Curtain Wall & Stone	Residential
2	Antalya	Mat	Pan Slab	structural steel	Unit Price	Curtain Wall	Commercial
3	Ankara		Steel Systems		Cost Plus	Curtain Wall & Plaster	Educational
4	Diyarbakır					Plaster	Hospital
5	İzmir					Brick Coating	Residential&Commercial
6	Balıkesir					Aluminum Façade	Dormitory
7	Mardin					Precast	
8						Stone	
9						EIFS	
10						EIFS & Glass	
11						Glass & Stone	
12						EIFS & Stone	
13						EIFS & Composite	
14						EIFS & Plaster	

Figure 4.9. shows the impact of each cost driver on the output model of construction cost. According to the result, “Usage of building” and “Construction duration” have the most effect on the model, with an importance of 11%. After that, “CPI” with an importance of 10%, “Number of storeys” with an importance of 8%; “Average discount rate” with an importance of 7%; “Exterior finishing type” with the importance of 6%; “Exchange rate” with the importance of 6%; “Construction turnover” with the importance of 5%; “Framing type” with the importance of 5%; “Floor system” with the importance of 5%; “Foundation system” with the importance of 5%; “Average minimum wage” with the importance of 5%; “Construction year” with the importance of 5%; “Area” with the importance of 4%; “Project location” with the importance of 3%; “Contract type” with the importance of 2%, and “GDP per capita” with the importance of 1%.

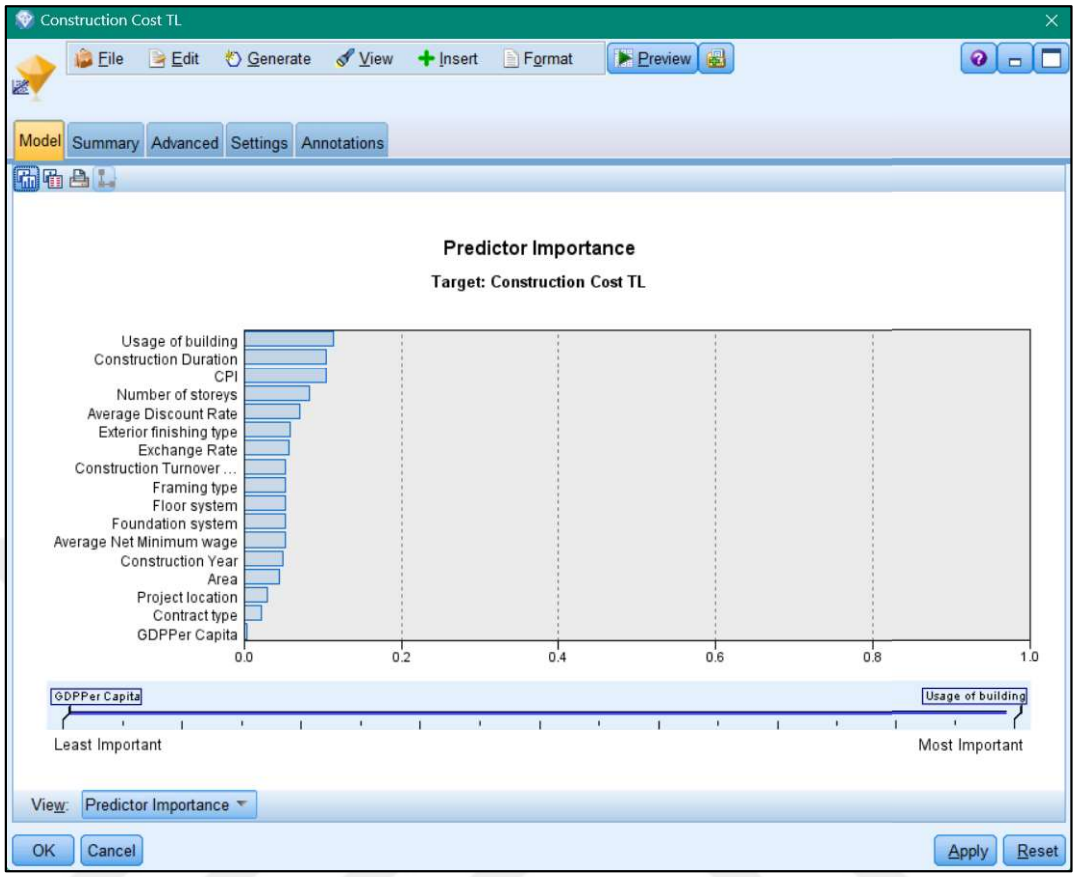

Figure 4.9. Cost drivers' impact on the regression model.

Table 4.2. Regression model summary

Model	R	R Square	Adjusted R Square
Regression	0.913	0.834	0.760

Table 4.2. demonstrates the R, R-squared, and Adjusted R-squared values for the developed regression model. In regression models, R, R-squared, and adjusted R-squared are all measures of the model's goodness of fit. R is the coefficient of correlation between the expected and actual values of the dependent variable [113]. R-squared is "the proportion of the variance in the dependent variable that can be explained by the independent variables in the regression model" [113]. R-squared has a value between 0 and 1, with a greater value suggesting that the model is better fitted to the data. An R-squared value of 0 indicates that the model explains no variability in the dependent variable, whereas an R-squared value of 1 indicates that the model perfectly explains all variability in the dependent variable. We can generate an Adjusted R-squared by changing the R-squared to account for the number of

independent variables in the model. When comparing models with various numbers of independent variables, adjusted R-squared penalizes the inclusion of unnecessary independent variables in the model and is a better measure to evaluate the model's goodness of fit [113].

Table 4.3. Regression coefficients.

Model	Unstandardized Coefficients	Standardized Coefficients	t	Sig.
	B	Beta		
(Constant)	-1262799371.000		-.459	.649
Number of storeys	851323.363	.458	5.426	.000
Project location	-1729404.504	-.139	-1.361	.182
Construction Year	606059.544	.125	.445	.659
Construction Duration	5058725.100	.168	1.979	.055
Area	363.616	.500	5.384	.000
Foundation system	-15194564.440	-.223	-2.110	.042
Floor system	-10061095.260	-.135	-.982	.332
Framing type	36436423.210	.283	1.975	.056
Contract type	2425767.547	.038	.498	.621
Exterior finishing type	129162.736	.017	.138	.891
Usage of building	1049308.587	.078	.993	.327
Construction Turnover (Billion TL)	-105.981	-.073	-1.010	.319
Average Discount Rate	107947.101	.042	.192	.849
Average Net Minimum wage	-1.026	-.090	-.762	.451
CPI	63410.129	.697	.718	.477
GDP Per Capita	1189.473	.060	.417	.679
Exchange Rate	-1667948.652	-.356	-.441	.662

The regression model coefficients are shown in Table 4.3. Unstandardized and standardized coefficients are two types of coefficients used in a regression model to interpret the relationship between the independent variables and the dependent variable. The change in the dependent variable for a one-unit change in the independent variable is represented by unstandardized coefficients, often known as beta coefficients. Unstandardized coefficients are given in the units of the dependent variable and can be used to compare the relative importance of the independent variables in the model [114]. Standardized coefficients, on

the other hand, describe the change in the dependent variable for a one-standard-deviation change in the independent variable. When the independent variables are measured on multiple scales, standardized coefficients are beneficial for assessing the relative relevance of different independent variables in the model [114]. The standardization of input variables is a straightforward method for improving the interpretability of regression coefficients. It balances the input variables, making the coefficients of untransformed binary predictors directly comparable [115].

Based on the regression coefficient values presented in Table 4.3. the mathematical relationship between the dependent variable (Construction cost) and independent variables (Cost Drivers) can be further obtained. The mathematical equation for construction cost is expressed as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n + \varepsilon, \quad (4.1)$$

where y is the dependent variable (Construction cost), x_1, x_2, ..., x_n are the independent variables (Cost Drivers), β_0, β_1, β_2, ..., β_n are the coefficients or regression parameters that represent the change in y (Construction cost), and ε is the error term or constant value, which represents the part of y that the independent variables cannot explain

Knowing the above regression equation, the construction cost equation will be Construction cost = [Number of storeys × 851323.4] + [Project location × -1729404.5] + [Construction Year × 606059.5] + [Construction Duration × 5058725.1] + [Area × 363.6] + [Foundation system × -15194564.4] + [Floor system × -10061095.3] + [Framing type × 36436423.2] + [Contract type × 2425767.5] + [Exterior finishing type × 129162.7] + [Usage-of building × 1049308.6] + [Construction Turnover × -106.0] + [Average Discount Rate × 107947.1] + [Average Net Minimum wage × -1.026] + [CPI × 63410.1] + [GDP Per Capita × 1189.5] + [Exchange Rate × -1667948.7] + (-1262799370.5)

MAPE is calculated for the Regression model using the equation (3.4), which is $M = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A-F}{A}\right|$= 24.7731%. In this equation M is the mean absolute percentage error, n is the number of times the summation iteration happens, A is the actual value and F is the forecasted value.

4.2.4. CHAID

The last model among the top four machine learning algorithms which could predict the construction cost using the data points was the CHAID model. The CHi-squared Automatic Interaction Detection (CHAID) model is a data mining ensemble approach based on decision trees that are frequently used for prediction analysis in a variety of applications. CHAID is a multivariate method for automatically classifying and analyzing vast numbers of variables. It is an exploratory tool that investigates the relationship between an outcome variable and a series of predictors and their interactions, progressively splitting a data set using a classification-tree structure by selecting from a set of specified variables, the one that maximizes a chi-square criterion in each split.

According to Figure 4.10., the most important independent variables in the CHAID model are; respectively; “Area” with the importance rate of 28%, “Usage of building” with the importance rate of 16%, “Average net minimum wage” with the importance rate of 11%, “Project location” with the importance rate of 11%, “Construction year” with the importance rate of 11%, “Construction duration” with the importance rate of 11% and “Number of storeys” with the importance rate of 11%.

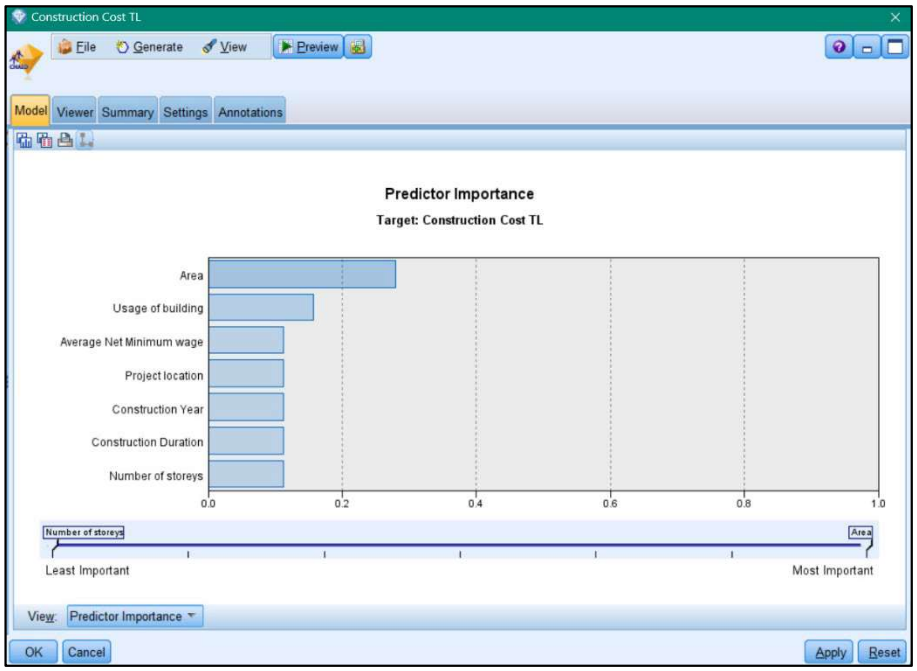

Figure 4.10. Cost drivers’ impact on the CHAID model.

Figure 4.11. illustrates a schematic figure of the CHAID model developed to predict the construction cost using the independent variables. It can be observed that the CHAID Tree has got a depth of 5 with a total of 34 nodes.

MAPE is calculated for the CHAID model using the equation (3.4), which is $M = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A-F}{A}\right|$= 37.5447%. In this equation M is the mean absolute percentage error, n is the number of times the summation iteration happens, A is the actual value and F is the forecasted value.

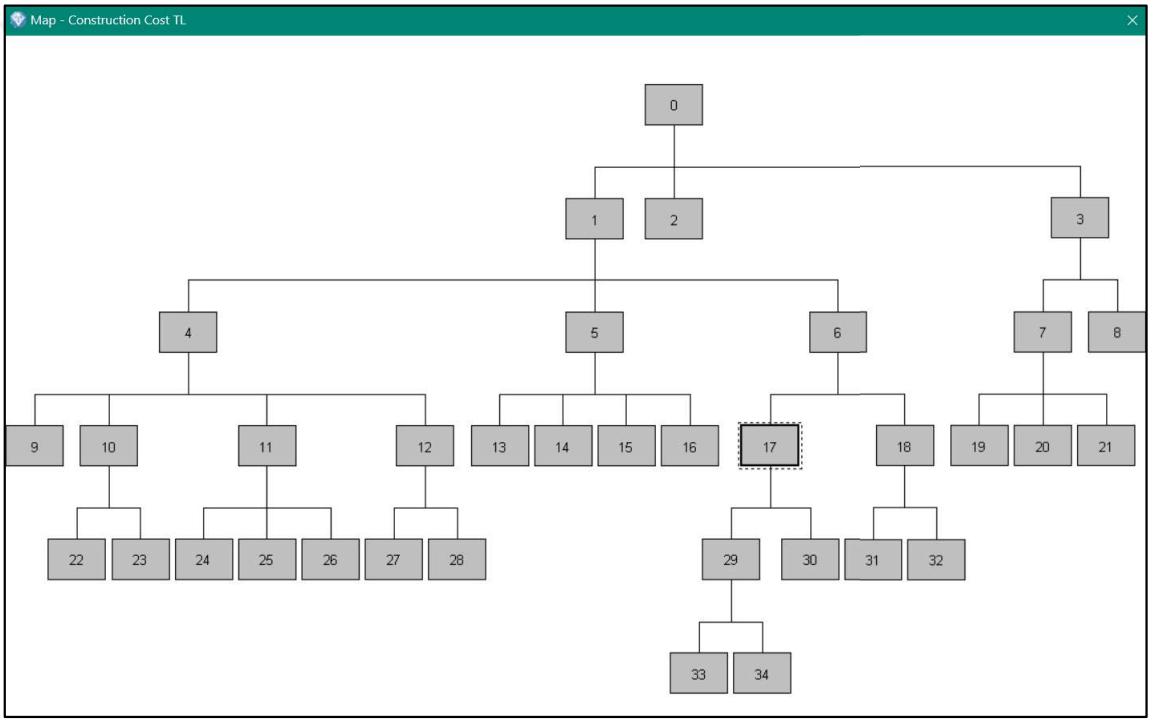

Figure 4.11. Schematic view of the CHAID model.

5. DISCUSSION

This Study investigated the factors affecting construction conceptual cost estimates and emphasized the economic factors along with the design factors. It also looked into the quantitative data analysis algorithms and compare the models' performance in cost estimation using sixty-four data points gathered within Turkey.

There are several benefits in automating cost estimation and developing statistical models in order to predict the conceptual cost estimates. Traditional cost estimation approaches are time and effort consuming processes. They are also more likely to be prone to human errors and biases because of personal judgments [67]. Precise conceptual cost estimation can be done in a limited time and ease the decision-making process in the primary stage of a construction project.

Over 60% of the construction project's failures are because of economic factors [28]. Fiscal risks caused by economic fluctuations should be taken seriously by the project owners to mitigate the financial risks. Therefore, macroeconomic factors analysis has gotten more importance in volatile economies. This study proposed an approach for conceptual cost estimation of the construction projects by considering the macroeconomic factors of the country along with the project specifications.

5.1. Interpretation of The Result

Many previous studies investigate the usefulness of statistical models in cost estimating. Chen and Tun [87] concluded that Neural Network works best for non-mechanistic modeling in ecological research and management and is helpful for non-linear ordination and visualization of data. The current study also concludes that Neural Network works the best in developing a construction cost estimate with a relative error of 3.8% compared to other models. This indicated that Neural Network can be used for performing conceptual cost estimates at the preliminary stage of the projects.

The predominant cost drivers can be identified at the early stage of the project. The impact of each design variable has been studied in the current research. Considering the Neural Network as the best fit for the database, the importance of the inputs then shows the effect of each variable on the cost model. NN model shows a positive impact of "Number of Storeys" (H1a), "Area" (H1c), "Exterior finishing type" (H1h) and "Usage of the building" (H1i). By that mean, it has been suggested that "Project Location," "Foundation System," "Floor System," "Framing type," and "Contract type" cannot be helpful for developing the cost model using the input database. However, based on the model developed by regression, there is a positive relationship between "Framing type" "Floor system" has, "Project location" and "Contract type" with the cost model.

The results support the finding of the study carried out by Günaydın and Doğan [27] and Cheng et al. [25]. Both studies suggest that the total area of the project and the number of storeys affect the conceptual cost estimation of construction projects. However, the results of the NN fail to support the relationship between the foundation system and floor system as opposed to the findings in the study carried out by Günaydın and Doğan [27] and Cheng et al. [25].

Neural network has found a relationship between exterior finishing type and conceptual cost estimation similar to [25]. Therefore, the current study supports the finding of the study carried out by Cheng et al. [25].

Table 5.1. Importance of design variables

Variable / Model	Number of Storeys	Project Location	Area	Foundation System	Floor System	Framing Type	Contract Type	Exterior Finishing Type	Usage of the Building
NN	21%	-	13%	-	-	-	-	6%	5%
C&R	9%	-	14%	-	-	-	-	-	-
Regression	8%	3%	4%	5%	5%	5%	2%	6%	11%
CHAID	11%	11%	28%	-	-	-	-	-	16%

Each model also presents a different impact of economic factors on the developed cost model (Table 5.2.). According to the Neural Network model, "Average rediscount rate," "Average net minimum wage," "CPI," "GDP per capita," and "Inflation rate" affect the

construction cost (H2b, H2c, H2d, H2e, H2f); however, the model did not consider "Construction turnover," and "Exchange rate." Alternatively, the C&R model considers all the economic factors in the cost model. Furthermore, the regression model considered all the economic factors except for the "Inflation rate." Finally, CHAID did not consider any of the economic factors except for the "Average net minimum wage" in the cost model.

Some previous studies pointed out the relationship between macroeconomic factors and project costs. Semenova and Vitkova [2] stated that there is a significant relationship between unemployment and liquidity of construction companies or in other words, the financial stability of the construction companies. The study shows the impact of GDP on the financial stability of construction companies in Spain. The result of the current study also supports the relationship between GDP and construction costs in Turkey.

Rafiei and Adeli [4] discussed the importance of integrating economic factors such as CPI, inflation and exchange rates, worker wage, and interest rate into the construction cost model and the result showed a positive relationship between economic factors and the cost estimate model. The current study supports that result and shows a significant correlation between CPI, inflation rate, average minimum wage, average rediscount rate, and cost model as can be observed in Table 5.2.

Yildiz [95] asserts that GDP is the granger cause of maritime exports, maritime imports, and the construction industry in Turkey. The result of this study supports their finding and asserts that there is a positive correlation between GDP and construction cost.

Table 5.2. Importance of economic variables

Variable / Model	Construction Turnover	Average Rediscount Rate	Average Net Minimum Wage	CPI	GDP Per Capita	Inflation Rate	Exchange Rate
NN	-	6%	5%	6%	6%	7%	-
C&R	9%	9%	9%	9%	9%	9%	9%
Regression	5%	7%	5%	10%	1%	-	6%
CHAID	-	-	11%	-	-	-	-

In order to investigate the effect of economic factors on the cost model's accuracy, the economic variables were then eliminated, and another cost model was developed to compare the model's accuracy using just the design variables.

Figure 5.1. shows the accuracy of Neural Network model when inputting just the design variables. As can be observed in the figure, the accuracy dropped to 73.9%; therefore, it can be concluded that economic variables play a vital role in the cost model's accuracy (H3).

By eliminating the economic variables, the importance factors of design variables change. As can be observed in Figure 5.2, "Number of storeys" now has the greatest impact of 36%, "Area" has the impact of 20%, "Exterior finishing type" has the impact of 14%, "Project location" has the impact of 11%, "Foundation system" has the impact of 7%, "Usage of the building" has the impact of 5%, "Construction duration" has the impact of 4%, and "Contract type" has the impact of 3%. The new model did not consider "Floor system" and "Framing type" among the effective design variables as the input of the model.

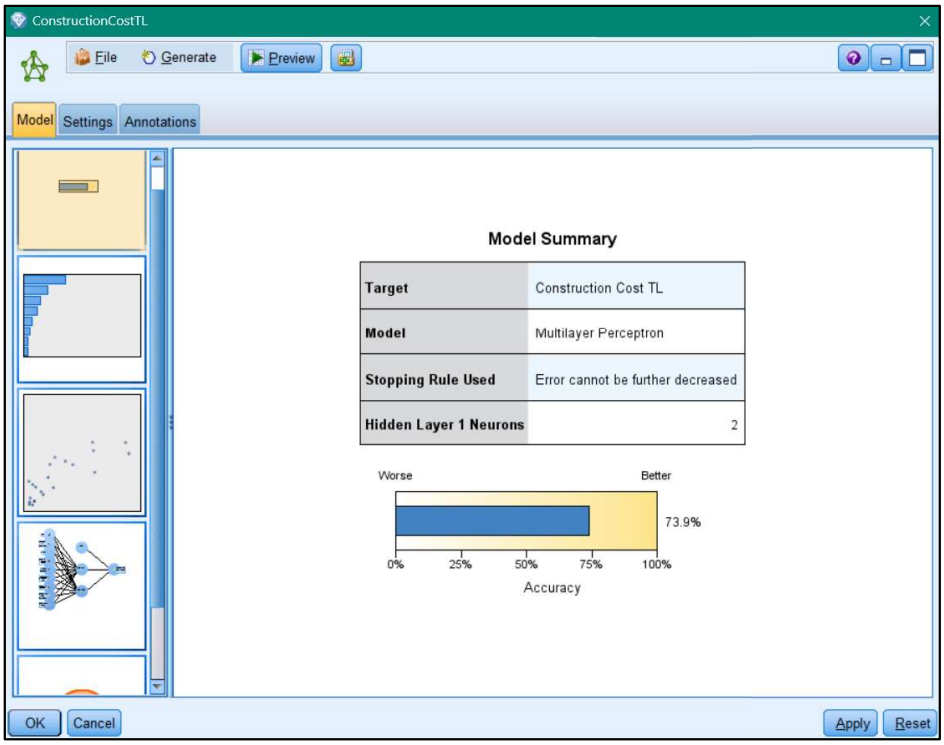

Figure 5.1. Neural network model accuracy using the design variables.

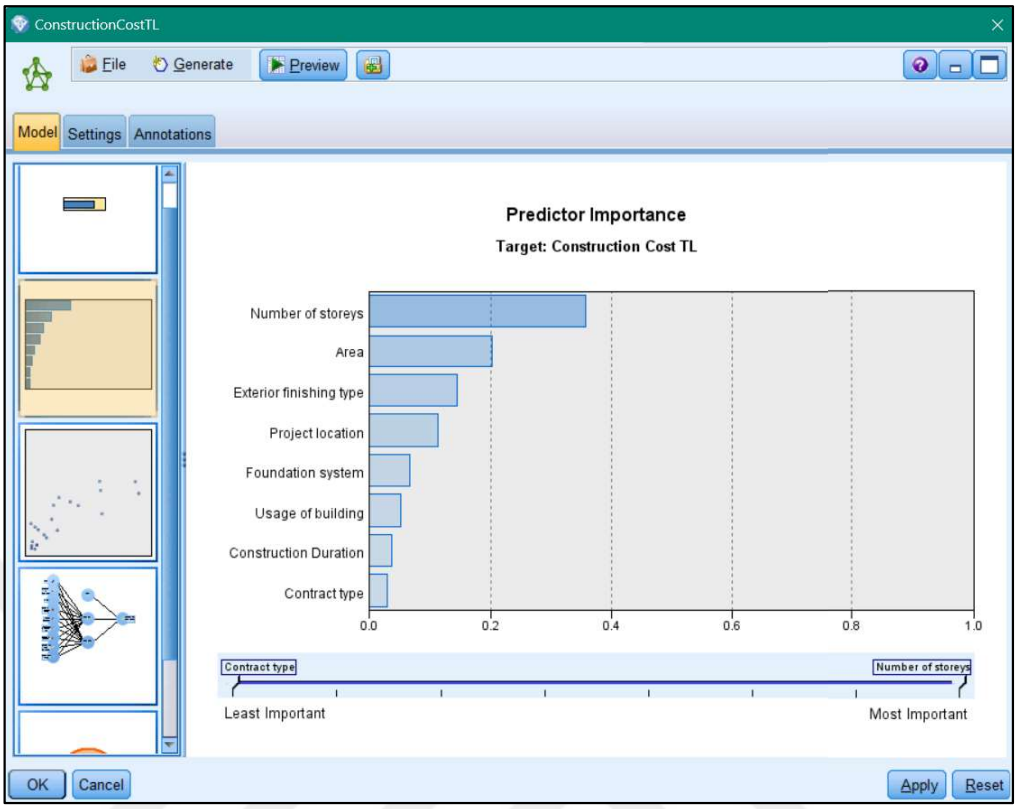

Figure 5.2. The effect of design variables on NN model.

5.2. Limitations and Future Research

This study has got limitations that affect the final research results. The first limitation is the number of participants. Totally, sixty-four data points were gathered that include projects' technical data and construction costs. The total project costs are considered confidential information that nobody is willing to share.

Turkey has witnessed the growth of the construction industry in many periods. As Gül [116] discussed, Turkey has observed construction growth mainly in two periods 1982 to 1988 and 2002 to 2008. The construction industry has always been one of the important pillars of economic growth and contributes approximately 8-9% to the country's GDP [117]. Therefore, it plays an essential role in terms of job opportunities. As a result, project managers and owners are not willing to share any technical or economic information.

Another limitation was the diversity of the projects in different aspects. For instance, out of sixty-four projects, sixty-two had reinforced concrete framing, and only two of them had structural steel framing. As a result, the accuracy of "Framing type" effect on the model is questionable. Similarly, the rate of the projects with the concrete floor system was too high. Out of sixty-four projects, sixty-one of them had reinforced concrete floor system, two of them had pan-slab floor system, and one had steel floor system. Therefore, the low impact of these variables on the model is justifiable.

Another shortcoming of the study is that it fails to consider the relationship between the variables themselves and their mutual effect on the cost model; this term is called multicollinearity. The cost drivers used as the model inputs in this study are not necessarily independent; as a matter of fact, some of them are interdependent, and the relationship between them should be studied. This can be seen mainly in economic factors; for instance, the exchange rate can affect the inflation rate, and consequently, the inflation rate can affect the average minimum wage. Multicollinearity is the linear relationship between two or more variables, which also implies a lack of orthogonality between them [118]. Therefore, a correlation analysis should be conducted to study the economic variables' relationship and measure their strength. This results in a more accurate cost model, which can be trusted.

Multicollinearity is a common problem in finance, particularly among macroeconomic variables, because the dependencies between variables can alter over time. Collecting more data is one of the simple ways to mitigate the multicollinearity effect; however, it is not always easy to access a large dataset. Therefore, other methods such as variables selections methods can be helpful in that matter. One of the common variable selection methods is stepwise regression which has two basic processes, forward selection, and backward elimination [119]. The forward selection method starts with developing the model by adding up the variables one at a time whereas the backward elimination method starts developing the model by considering all the variables and dropping them one by one. Another method to do variable selection is to perform correlation matrix analysis. In this method, the correlation matrix for all the variables in the dataset is computed. Consequently, highly correlated pairs of variables are identified and one of them can be eliminated, or alternatively, they can be combined into a single composite variable if appropriate [120].

Another way to address the multicollinearity while developing a model with dependent variables is to use modified estimator models. For instance, a convolutional neural network (CNN) can be utilized to detect the macroeconomic variables' correlation. CNN is very effective in terms of addressing the relationship between variables and reduces the number of network parameters. Unlike NN, CNN adopts multiple convolutional and pooling layers to detect the most effective variables before developing a fully connected network for prediction.

In conclusion, to address the problem of interdependency in future studies, two general methods can be employed. One is to apply variable selection methods such as forward selection or backward elimination methods, and pick the most effective inputs and eliminate the dependent ones. The other one is to modify the estimation model and employ modified algorithms such as CNNs and RNNs (Recurrent Neural Networks).

6. CONCLUSION

Cost estimation is a critical component of any construction project in every phase, from initial planning, when the rough order of magnitude takes place, to the very end point, project delivery. The project's cost estimate accuracy increases when the project moves toward the end. Before the detailed designing takes place, the estimates done at the planning and scoping help the project owners to decide on the project's feasibility and estimate the budget roughly. Therefore, these rough cost estimates significantly impact owners' decisions and projects' success.

In developing countries with volatile economies, however, predicting the project's cost beforehand is significantly more complex because of economic fluctuations. This results in more uncertainties when it comes to starting a construction project. Being able to perform an accurate conceptual cost estimate makes it easier for the project owners to overcome cost overruns and easily face economic instabilities.

Many previous studies focused on developing cost models in order to accurately predict a construction project's cost in the early phases. They mainly considered design variables and characteristics as the input of the cost model. This study, however, emphasized the fact that structural design variables are not solely enough to develop a cost estimation model for construction projects. Because of the construction projects' complexity and their direct dependence on the material, labor, and machinery costs, it is vital to have a clear sight of the country's future economy while starting a construction project. Therefore, integrating macroeconomic factors in the conceptual cost estimation model is crucial in economically unstable countries.

This study proposed cost estimate models for building structures in Turkey considering nine design factors which are; the number of storeys, project location, total construction area, foundation system, floor system, contract type, exterior finishing type, and usage of the building. Also, seven economic factors were taken into consideration while developing the models. These variables were; construction turnover, average rediscount rate, average net

minimum wage, CPI, GDP per capita, inflation rate, and USD exchange rate. These factors were integrated into the model considering the construction year of the building project and its duration.

The results indicate that among the different quantitative data analysis models, neural network, C&R tree, regression, and CHAID have the best performance in predicting the construction projects' cost. Therefore, the impact of the structural design variables and economic variables was further studied using these quantitative models.

The result of the neural network model demonstrates the significant impact of the economic variables on the conceptual cost model. It suggests that the average rediscount rate has the importance of 6% on the cost model. Furthermore, the average net minimum wage impacts the cost model by 5%, CPI impacts the model by 6%, GDP per capita impacts the model by 6%, and inflation rate impacts the cost model by 7%. However, when it comes to structural design variables, neural network fails to consider all the input variables proposed by the study. This is a result of limited data points that have been used as the model input. Neural network suggests that the number of storeys has an importance f 21% on the cost model, area has an importance of 13%, exterior finishing type has the importance of 6%, and usage of the building has an importance of 5% on the cost model. Regression model, on the other hand, considered all the structural design variables proposed by the study, as the model input, and proposed importance factors for all.

The accuracy of the proposed models can be enhanced by considering the interdependency of the economic variables and studying their relationship and effects on one another. Increasing the number of projects' data and their diversity also can be helpful in terms of improving the models' performance and can be investigated by future studies.

The main contribution of the current study to the construction industry is to provide insight into the influence of macroeconomic factors on the conceptual cost estimates in volatile economies. Economic factors proposed by the study can be further integrated by the construction companies into their cost estimations while predicting project costs and budgets. It is crucial to note that a historical economic data analysis is necessary when deciding to start a construction project in order to avoid cost overruns. This is frequently

neglected by project owners in developing countries since they assume that design variables are sufficient to predict the construction project's cost accurately.

The public sector can benefit from a conceptual cost estimation, which is developed from historical data, to ensure the effective use of public resources. It enables more informed decision making, better planning, and greater efficiency throughout the construction lifecycle. It improves accurate budgeting at the early stages of planning, and ensures that the projects are adequately funded. This way, the risk of cost overruns and unexpected expenses can be reduced by the public sector and can result in increasing transparency and accountability in the public sector. Furthermore, it is easier to justify the budget allocation to the stakeholders using such cost models at the early stages and reduce bidding errors. Conceptual cost estimation can also be used by subcontractors while bidding for construction projects. They can benefit from cost models and increase their chances at tendering stage by accurately predicting the tender price with the historical cost models.

REFERENCES

1. Berk, N. and S. Biçen, "Causality Between the Construction Sector and GDP Growth in Emerging Countries: The Case of Turkey", *Athens Journal of Mediterranean Studies*, Vol. 4, No. 1, pp. 19-36, 2018.

2. Semenova, T. and E. Vitkova, "Impact of Macroeconomic Indicators on the Financial Stability of Construction Companies in the Czech Republic and Spain", in *IOP Conference Series: Materials Science and Engineering*, Prague, Czech Republic, Vol. 471, no. 2, IOP Publishing, p. 022030, 2019.

3. Wu, W. and H. Huang, "Construction Engineering Cost Estimation Based on Artificial Intelligence Technology", *Journal of Intelligent & Fuzzy Systems*, Vol. 2, pp. 1-9, 2021.

4. Rafiei, M. H. and H. Adeli, "Novel Machine-Learning Model for Estimating Construction Costs Considering Economic Variables and Indexes", *Journal of Construction Engineering and Management*, Vol. 144, No. 12, p. 04018106, 2018.

5. He, X., R. Liu, and C. J. Anumba, "Data-Driven Insights on the Knowledge Gaps of Conceptual Cost Estimation Modeling", *Journal of Construction Engineering and Management*, Vol. 147, No. 2, p. 04020165, 2021.

6. Guide, A., "Project Management Body of Knowledge (pmbok® guide)", in *Project Management Institute*, Vol. 11, pp. 7-8, 2001.

7. Law, C., "Building Contractor Estimating: British Style", *Cost Engineering*, Vol. 36, No. 6, p. 23, 1994.

8. Boukezia, R., J. Charaoui, J. M. Frank, M. Harb, M. Queyranne, N. Reyes, P. F. Ryan, and A. F. Tieman, "Managing Fiscal Risks in the Middle East and North Africa", *Departmental Papers*, Vol. 2023, No. 005, 2023.

9. DORUK, Ö. T., "Macroeconomic Determinants of Firm Performance: Evidence From Turkey", *The Singapore Economic Review*, Vol. 68, No. 1, pp. 1-20, 2019.

10. Fischer, S., "The Role of Macroeconomic Factors in Growth", *Journal of Monetary Economics*, Vol. 32, No. 3, pp. 485-512, 1993.

11. Egbunike, C. F. and C. U. Okerekeoti, "Macroeconomic Factors, Firm Characteristics and Financial Performance: A study of Selected Quoted Manufacturing Firms in Nigeria", *Asian Journal of Accounting Research*, Vol. 3, No. 2, pp. 142-168, 2018.

12. Silva, A. C. and G. A. Chávez, "Microfinance, Country Governance, and the Global Financial Crisis", *Venture Capital*, Vol. 17, No. 1-2, pp. 191-213, 2015.

13. Kim, S. B., Y. H. Lee, and A. Scanlon, "Comparative Study of Structural Material Quantities of Multi-Storey Linear-Type Residential Buildings", *The Structural Design of Tall and Special Buildings*, Vol. 18, No. 6, pp. 673-686, 2009.

14. Choon, T. T. and K. N. Ali, "A Review of Potential Areas of Construction Cost Estimating and Identification of Research Gaps", *Jurnal Alm Bina*, Vol. 11, No. 2, pp. 61-72, 2008.

15. Akintoye, A., "Analysis of Factors Influencing Project Cost Estimating Practice", *Construction Management & Economics*, Vol. 18, No. 1, pp. 77-89, 2000.

16. Amade, B. and E. O. P. Akpan, "Project Cost Estimation: Issues and the Possible Solutions", *International Journal of Engineering and Technical Research (IJETR) ISSN*, Vol. 2, No. 5, pp. 181-188, 2014.

17. Ssemwogerere, K., "An Analysis of the Effectiveness of the Use of Contigency Sum in Project Cost Estimation & Management in the Ugandan Construction Industry", in *RICS Construction and Property Conference*, Greater Manchester, p. 1522, 2011.

18. Akintoye, A. and E. Fitzgerald, "A Survey of Current Cost Estimating Practices in the UK", *Construction Management & Economics*, Vol. 18, No. 2, pp. 161-172, 2000.

19. Vergara, A. J. and L. T. Boyer, "Probabilistic Approach to Estimating and Cost Control", *Journal of the Construction Division*, Vol. 100, No. 4, pp. 543-552, 1974.

20. Arif, F., S. H. Lodi, and N. Azhar, “Factors Influencing Accuracy of Construction Project Cost Estimates in Pakistan: Perception and Reality”, *International Journal of Construction Management*, Vol. 15, No. 1, pp. 59-70, 2015.

21. Aibinu, A. A. and T. Pasco, “The Accuracy of Pre-Tender Building Cost Estimates in Australia”, *Construction Management and Economics*, Vol. 26, No. 12, pp. 1257-1269, 2008.

22. Kim, G.-H., J.-M. Shin, S. Kim, and Y. Shin, “Comparison of School Building Construction Costs Estimation Methods Using Regression Analysis, Neural Network, and Support Vector Machine”, *Journal of Building Construction and Planning Research*,Vol. 1, No. 1, pp. 1-7, 2013.

23. Oberlender, G. D. and S. M. Trost, “Predicting Accuracy of Early Cost Estimates Based on Estimate Quality”, *Journal of Construction Engineering and Management*, Vol. 127, No. 3, pp. 173-182, 2001.

24. Liu, L. and K. Zhu, “Improving Cost Estimates of Construction Projects Using Phased Cost Factors”, *Journal of Construction Engineering and Management*, Vol. 133, No. 1, pp. 91-95, 2007.

25. Cheng, M.-Y., H.-C. Tsai, and E. Sudjono, “Conceptual Cost Estimates Using Evolutionary Fuzzy Hybrid Neural Network for Projects in Construction Industry”, *Expert Systems with Applications*, Vol. 37, No. 6, pp. 4224-4231, 2010.

26. Min-Yuan, C., T. Hsing-Chih, and H. Wen-Shan, “Web-Based Conceptual Cost Estimates for Construction Projects Using Evolutionary Fuzzy Neural Inference Model”, *Automation in construction*, Vol. 18, No. 2, pp. 164-172, 2009.

27. Günaydın, H. M. and S. Z. Doğan, “A Neural Network Approach for Early Cost Estimation of Structural Systems of Buildings”, *International Journal of Project Management*, Vol. 22, No. 7, pp. 595-602, 2004.

28. Cheng, M.-Y., H.-C. Tsai, and C.-L. Liu, “Artificial Intelligence Approaches to Achieve Strategic Control Over Project Cash Flows”, *Automation in Construction*, Vol. 18, No. 4, pp. 386-393, 2009.

29. Owoeye, J. O., Salako, O. A., & Daramola, A. O. , “Factors Influencing Unit Price Estimation in Construction Industry”, *Journal of Engineering, Project, and Production Management*, Vol. 7, No. 1, pp. 31-40, 2017.

30. Erkan, İ. H., T. P. Doğan, and M. H. Arslan, “An Investigation on Determining Optimum Wall Ratio–Cost Relationship of Shear Walled Reinforced Concrete Buildings”, *Challenge Journal of Structural Mechanics*, Vol. 6, No. 1, pp. 1-9, 2020.

31. Garyaev, N. and A. Rybakova, “Cloud Interaction Technologies in the Design and Construction”, in *MATEC Web of Conferences*, Vol. 170, EDP Sciences, p. 01076, 2018.

32. Hadwiansyah, R. and Y. Latief, “Structural Equation Model (SEM) Correlation Between Work Breakdown Structure (WBS), Work Method and Risk Towards Cost Of Safety On Low-Cost Apartments Project”, in *IOP Conference Series: Materials Science and Engineering*, Yogyakarta, Indonesia, Vol. 1232, no. 1, IOP Publishing, p. 012006, 2022.

33. Wei, C.-C., C.-C. Hsiang, and T.-C. Shan, “Decision Support Model with Life Cycle Assessment for Building in Design Phase”, in *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, Sydney, Australia, Vol. 31, IAARC Publications, p. 1, 2014.

34. Elmousalami, H. H., “Intelligent Methodology for Project Conceptual Cost Prediction”, *Heliyon*, Vol. 5, No. 5, p. e01625, 2019.

35. Gunduz, M. and H. B. Sahin, “An Early Cost Estimation Model for Hydroelectric Power Plant Projects Using Neural Networks and Multiple Regression Analysis”, *Journal of Civil Engineering and Management*, Vol. 21, No. 4, pp. 470-477, 2015.

36. Kouskoulas, V. and E. Koehn, “Predesign Cost-Estimation Function for Buildings”, *Journal of the Construction Division*, Vol. 100, No. 4, pp. 589-604, 1974.

37. Juszczyk, M., “The Use of Artificial Neural Networks for Residential Buildings Conceptual Cost Estimation”, in *AIP Conference Proceedings*, Melville, New York, Vol. 1558, no. 1, American Institute of Physics., pp. 1302-1306, 2013.

38. Musarat, M. A., W. S. Alaloul, and M. Liew, “Impact of Inflation Rate on Construction Projects Budget: A Review”, *Ain Shams Engineering Journal*, Vol. 12, No. 1, pp. 407-414, 2021.

39. Oikawa, K. and K. Ueda, “The Optimal Inflation Rate Under Schumpeterian Grow”, *Journal of Monetary Economics*, Vol. 100, pp. 114-125, 2018.

40. Chowdhury, I., M. Hoffmann, and A. Schabert, “Inflation Dynamics and the Cost Channel of Monetary Transmission”, *European Economic Review*, Vol. 50, No. 4, pp. 995-1016, 2006.

41. Lucas, J., Robert E, “Inflation and Welfare”, *Econometrica*, Vol. 68, No. 2, pp. 247-274, 2000.

42. Kandil, M., “Exchange Rate Fluctuations and Economic Activity in Developing Countries: Theory and Evidence”, *Journal of Economic Development*, Vol. 29, pp. 85-108, 2004.

43. Wiser, R. H., M. Bolinger, B. Hoen, D. Millstein, J. Rand, G. L. Barbose, N. R. Darghouth, W. Gorman, S. Jeong, and A. D. Mills, “Wind Technology Data and Trends: Land-Based Focus (Update 2020)[Slides]”, Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2020.

44. Goldberg, L. and J. Tracy, “Exchange Rates and Local Labor Markets”, in *The Impact of International Trade on Wages*, University of Chicago Press, pp. 269-307, 2000.

45. Alesina, A. F. and R. Perotti, “The Welfare State and Competitiveness”, ed: National Bureau of Economic Research Cambridge, Mass., USA, 1994.

46. Coyle, D., “GDP: A Brief But Affectionate History”, *The Quarterly Journal of Austrian Economics*, Vol. 20, No. 1, pp. 97-103, 2017.

47. Yiu, C., X. Lu, M. Leung, and W. Jin, “A Longitudinal Analysis on the Relationship Between Construction Output and GDP in Hong Kong”, *Construction Management and Economics*, Vol. 22, No. 4, pp. 339-345, 2004.

48. Saka, N., "An Assessment of the Impact of the Construction Sector on the Gross Domestic Product (GDP) Of Nigeria", *Journal of Surveying, Construction and Property*, Vol. 13, No. 1, pp. 42-65, 2022.

49. Maran, R., S. Rajendran, and S. Kalidindi, "Material Cost and Escalation Clauses in Indian Construction Contracts", *Proceedings of the Institution of Civil Engineers-Construction Materials*, Vol. 164, No. 2, pp. 95-108, 2011.

50. Akanni, P., A. Oke, and O. Omotilewa, "Implications of Rising Cost of Building Materials in Lagos State Nigeria", *SAGE Open*, Vol. 4, No. 4, p. 2158244014561213, 2014.

51. Windapo, A. and K. Cattell, "Examining the Trends in Building Material Prices: Built Environment Stakeholders' Perspectives", *Management of Construction: Research to Practice*, Vol. 1, pp. 187-201, 2012.

52. Diewert, W. E., "The Consumer Price Index and Index Number Purpose", *Journal of Economic and Social Measurement*, Vol. 27, No. 3-4, pp. 167-248, 2001.

53. Shinghal, A. K. and H. Nagar, "Testing Empirically the Impact of Inflating Diesel Prices in WPI and Cci", in *IOP Conference Series: Materials Science and Engineering*, Tamil Nadu, India, Vol. 872, no. 1, IOP Publishing, p. 012023, 2020.

54. Wong, J. M. and S. T. Ng, "Forecasting Construction Tender Price Index in Hong Kong Using Vector Error Correction Model", *Construction Management and Economics*, Vol. 28, No. 12, pp. 1255-1268, 2010.

55. Aaronson, D., "Price Pass-Through and the Minimum Wage", *Review of Economics and Statistics*, Vol. 83, No. 1, pp. 158-169, 2001.

56. Lindner, A. and P. Harasztosi, "Who Pays for the Minimum Wage?", *American Economic Review*, Vol. 102, No. 4, pp. i-vii, 2019.

57. König, M. and J. Möller, "Impacts of Minimum Wages: A Microdata Analysis for the German Construction Sector", *International Journal of Manpower*, Vol. 30, No. 7, pp. 716-741, 2009.

58. Işık, E., Ö. Orhangazi, and H. Tekgüç, "Heterogeneous Effects of Minimum Wage on Labor Market Outcomes: A Case Study From Turkey", *IZA Journal of Labor Policy*, Vol. 10, No. 1, pp. 10-16, 2020.

59. Kemal, B. M. and M. Kocaman, "The Impact of Minimum Wage on Unemployment, Prices, and Growth: A Multivariate Analysis for Turkey", *Economic Annals*, Vol. 64, No. 221, pp. 65-83, 2019.

60. Aydinli, S., O. Mustafa, and O. Emel, "Wage Determinants and Wage Inequalities–Case of Construction Engineers in Turkey", *Teknik Dergi*, Vol. 30, No. 2, pp. 8961-8985, 2019.

61. Regona, M., T. Yigitcanlar, B. Xia, and R. Y. M. Li, "Artificial Intelligent Technologies for the Construction Industry: How Are They Perceived and Utilized in Australia?", *Journal of Open Innovation: Technology, Market, and Complexity*, Vol. 8, No. 1, p. 16, 2022.

62. Mathern, A., K. Ek, and R. Rempling, "Sustainability-Driven Structural Design Using Artificial Intelligence", in *Proceedings of the IABSE Congress New York City—The Evolving Metropolis*, New York, NY, USA, pp. 4-6, 2019.

63. Faghihi, V., A. Nejat, K. F. Reinschmidt, and J. H. Kang, "Automation in Construction Scheduling: A Review of the Literature", *The International Journal of Advanced Manufacturing Technology*, Vol. 81, No. 9, pp. 1845-1856, 2015.

64. Kim, H. and L. Soibelman, "Knowledge Discovery and Supervised Machine Learning in a Construction Project Database", *WIT Transactions on Information and Communication Technologies*, Vol. 28, pp. 137-146, 2002.

65. Chen, X., "Research on the Application of AI Artificial Intelligence Technology in Intelligent Buildings", *Academic Journal of Architecture and Geotechnical Engineering*, Vol. 4, No. 3, pp. 37-41, 2022.

66. Abioye, S. O., L. O. Oyedele, L. Akanbi, A. Ajayi, J. M. D. Delgado, M. Bilal, O. O. Akinade, and A. Ahmed, "Artificial Intelligence in the Construction Industry: A

Review of Present Status, Opportunities and Future Challenges", *Journal of Building Engineering*, Vol. 44, p. 103299, 2021.

67. Elmousalami, H. H., "Comparison of Artificial Intelligence Techniques for Project Conceptual Cost Prediction: A Case Study and Comparative Analysis", *IEEE Transactions on Engineering Management*, Vol. 68, No. 1, pp. 183-196, 2020.

68. Sonmez, R., "Conceptual Cost Estimation of Building Projects With Regression Analysis and Neural Networks", *Canadian Journal of Civil Engineering*, Vol. 31, No. 4, pp. 677-683, 2004.

69. Jadid, M. N. and M. M. Idrees, "Cost Estimation of Structural Skeleton Using an Interactive Automation Algorithm: A Conceptual Approach", *Automation in Construction*, Vol. 16, No. 6, pp. 797-805, 2007.

70. Park, U., Y. Kang, H. Lee, and S. Yun, "A Stacking Heterogeneous Ensemble Learning Method for the Prediction of Building Construction Project Costs", *Applied Sciences*, Vol. 12, No. 19, p. 9729, 2022.

71. Cho, H.-G., K.-G. Kim, J.-Y. Kim, and G.-H. Kim, "A Comparison of Construction Cost Estimation Using Multiple Regression Analysis and Neural Network in Elementary School Project", *Journal of the Korea Institute of Building Construction*, Vol. 13, No. 1, pp. 66-74, 2013.

72. Kim, G.-H., J.-M. Shin, S. Kim, and Y. Shin, "Comparison of School Building Construction Costs Estimation Methods Using Regression Analysis, Neural Network, and Support Vector Machine", *Journal of Building Construction and Planning Research*, Vol. 1, No. 1, pp. 1-7, 2013.

73. Hsiao, F. Y., S. H. Wang, W. C. Wang, C. P. Wen, and W. D. Yu, "Neuro-Fuzzy Cost Estimation Model Enhanced by Fast Messy Genetic Algorithms for Semiconductor Hookup Construction", *Computer-Aided Civil and Infrastructure Engineering*, Vol. 27, No. 10, pp. 764-781, 2012.

74. Juszczyk, M., A. Leśniak, and K. Zima, "ANN Based Approach for Estimation of Construction Costs of Sports Fields", *Complexity*, Vol. 2018, pp. 1-11, 2018.

75. Barros, L. B., M. Marcy, and M. T. Carvalho, "Construction Cost Estimation of Brazilian Highways Using Artificial Neural Networks", *International Journal of Structural and Civil Engineering Research*, Vol. 7, No. 3, pp. 283-289, 2018.

76. Tayefeh Hashemi, S., O. M. Ebadati, and H. Kaur, "Cost Estimation and Prediction in Construction Projects: A Systematic Review on Machine Learning Techniques", *SN Applied Sciences*, Vol. 2, No. 10, pp. 1-27, 2020.

77. Ali, Z. H., A. M. Burhan, M. Kassim, and Z. Al-Khafaji, "Developing an Integrative Data Intelligence Model for Construction Cost Estimation", *Complexity*, Vol. 2022, pp. 1-18, 2022.

78. Matel, E., F. Vahdatikhaki, S. Hosseinyalamdary, T. Evers, and H. Voordijk, "An Artificial Neural Network Approach for Cost Estimation of Engineering Services", *International Journal of Construction Management*, Vol. 22, No. 7, pp. 1274-1287, 2022.

79. Kovačević, M., N. Ivanišević, P. Petronijević, and V. Despotović, "Application of Support Vector Machines in Assessing Conceptual Cost Estimates", *Građevinar*, Vol. 73, No. 1, pp. 1-13, 2021.

80. An, S.-H., U.-Y. Park, K.-I. Kang, M.-Y. Cho, and H.-H. Cho, "Application of Support Vector Machines in Assessing Conceptual Cost Estimates", *Journal of Computing in Civil Engineering*, Vol. 21, No. 4, pp. 259-264, 2007.

81. Juszczyk, M., "Residential Buildings Conceptual Cost Estimates With the Use of Support Vector Regression", in *MATEC Web of Conferences*, Rostov-on-Don, Russia, Vol. 196, EDP Sciences, p. 04090, 2018.

82. Khandani, A. E., A. J. Kim, and A. W. Lo, "Consumer Credit-Risk Models via Machine-Learning Algorithms", *Journal of Banking & Finance*, Vol. 34, No. 11, pp. 2767-2787, 2010.

83. Sadeghi, N., A. R. Fayek, and W. Pedrycz, "Fuzzy Monte Carlo Simulation and Risk Assessment in Construction", *Computer-Aided Civil and Infrastructure Engineering*, Vol. 25, No. 4, pp. 238-252, 2010.

84. Fayek, A. R., “Fuzzy Logic and Fuzzy Hybrid Techniques for Construction Engineering and Management”, *Journal of Construction Engineering and Management*, Vol. 146, No. 7, p. 04020064, 2020.

85. Plebankiewicz, E., K. Zima, and D. Wieczorek, “Modelling of Time, Cost and Risk of Construction With Using Fuzzy Logic”, *Journal of Civil Engineering and Management*, Vol. 27, No. 6, pp. 412-426, 2021.

86. Juszczyk, M., “Development of Cost Estimation Models Based on Ann Ensembles and the Svm Method”, *Civil and Environmental Engineering Reports*, Vol. 30, No. 3, pp. 48-67, 2020.

87. Chen, T.-P. and J.-X. Tun, “The Research of Artificial Neural Network as a Tool of Adjustment Prediction in Eco-City Construction”, in *2005 International Conference on Machine Learning and Cybernetics*, Guangzhou, China, Vol. 6, IEEE, pp. 3416-3420, 2005.

88. Kim, K. J. and K. Kim, “Preliminary Cost Estimation Model Using Case-Based Reasoning and Genetic Algorithms”, *Journal of Computing in Civil Engineering*, Vol. 24, No. 6, pp. 499-505, 2010.

89. Unke, O. T., S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K.-R. Müller, “Machine Learning Force Fields”, *Chemical Reviews*, Vol. 121, No. 16, pp. 10142-10186, 2021.

90. Gambatese, J. A., M. Behm, and J. W. Hinze, “Viability of Designing for Construction Worker Safety”, *Journal of Construction Engineering and Management*, Vol. 131, No. 9, pp. 1029-1036, 2005.

91. Zhang, Y. and C. Ling, “A Strategy to Apply Machine Learning to Small Datasets in Materials Science”, *Npj Computational Materials*, Vol. 4, No. 1, pp. 1-25, 2018.

92. Bayır, M. and Ç. Orak, “Globalization and Inflation”, *Hacettepe University Journal of Economics & Administrative Sciences/Hacettepe Üniversitesi Iktisadi ve Idari Bilimler Fakültesi Dergisi*, Vol. 41, No. 1, pp. 1-20, 2023.

93. Diya, D. K. and N. Jesna, “Macroeconomic Indicators as Potential Predictors of Construction Material Prices ”, *Sustainability, Agri, Food and Environmental Research*, Vol. 10, pp. 1-9, 2022.

94. Chen, H. G., C. Y. Li, and D. D. Xie, “Economic Study on Sewage Treatment Engineering Based on LCC”, in *Advanced Materials Research*, Vol. 838, Trans Tech Publ, pp. 2689-2693, 2014.

95. Yildiz, R., “A Case Study of Examining the Relationships Between Maritime Foreign Trade, GDP, and the Construction in Turkey”, *Sosyal Bilimler Arastirmalari Dergisi*, Vol. 12, No. 2, pp. 1213-1229, 2022.

96. Chakraverty, S., D. M. Sahoo, N. R. Mahato, S. Chakraverty, D. M. Sahoo, and N. R. Mahato, “McCulloch–Pitts Neural Network Model”, *Concepts of Soft Computing: Fuzzy and ANN with Programming*, Vol. 11, pp. 167-173, 2019.

97. Basu, A., J. Lowell, R. Kennedy, A. Hunter, and R. Ryder, “Development of an Artificial Neural Network (ANN) For Retinal Image Analysis”, *Diabetes*, Vol. 50, p. A430, 2001.

98. Espinoza-Sepulveda, N. F. and J. K. Sinha, “Robust Vibration-Based Faults Diagnosis Machine Learning Model for Rotating Machines to Enhance Plant Reliability”, *Maintenance, Reliability and Condition Monitoring*, Vol. 1, No. 1, pp. 2-9, 2021.

99. Myles, A. J., R. N. Feudale, Y. Liu, N. A. Woody, and S. D. Brown, “An Introduction to Decision Tree Modeling”, *Journal of Chemometrics: A Journal of the Chemometrics Society*, Vol. 18, No. 6, pp. 275-285, 2004.

100. Maulud, D. and A. M. Abdulazeez, “A Review on Linear Regression Comprehensive in Machine Learning”, *Journal of Applied Science and Technology Trends*, Vol. 1, No. 4, pp. 140-147, 2020.

101. Acharya, M. S., A. Armaan, and A. S. Antony, “A Comparison of Regression Models for Prediction of Graduate Admissions”, in *2019 International Conference on*

Computational Intelligence in Data Science (ICCIDS), Chennai, India, IEEE, pp. 1-5, 2019.

102. Zhang, Z., Y. Li, L. Li, Z. Li, and S. Liu, "Multiple Linear Regression for High Efficiency Video Intra Coding", in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, United Kingdom, IEEE, pp. 1832-1836, 2019.

103. Jie, H. and G. Zheng, "Calibration of Torque Error of Permanent Magnet Synchronous Motor Base on Polynomial Linear Regression Model", in *IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society*, Lisbon, Portugal, Vol. 1, IEEE, pp. 318-323, 2019.

104. Taylor, R., "Interpretation of the Correlation Coefficient: A Basic Review", *Journal of Diagnostic Medical Sonography*, Vol. 6, No. 1, pp. 35-39, 1990.

105. Khair, U., H. Fahmi, S. Al Hakim, and R. Rahim, "Forecasting Error Calculation With Mean Absolute Deviation and Mean Absolute Percentage Error", in *Journal of Physics: Conference Series*, Moscow, Russian Federation, Vol. 930, no. 1, IOP Publishing, p. 012002, 2017.

106. Coleman, C. D. and D. A. Swanson, "On Mape-R as a Measure of Cross-Sectional Estimation and Forecast Accuracy", *Journal of Economic and Social Measurement*, Vol. 32, No. 4, pp. 219-233, 2007.

107. Loh, W. Y., "Classification and Regression Trees", *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol. 1, No. 1, pp. 14-23, 2011.

108. ÇINAROĞLU, S., "Comparison of Performance of Decision Tree Algorithms and Random Forest An Application on OECD Countries Health Expenditures", *International Journal of Computer Applications*, Vol. 138, No. 1, pp. 37-41, 2016.

109. Gulati, P., A. Sharma, and M. Gupta, "Theoretical Study of Decision Tree Algorithms to Identify Pivotal Factors for Performance Improvement: A Review", *International Journal of Computer Applications*, Vol. 141, No. 14, pp. 19-25, 2016.

110. Quinlan, J. R., "Induction of Decision Trees", *Machine Learning*, Vol. 1, pp. 81-106, 1986.

111. Pham, T. T., I. Reid, Y. Latif, and S. Gould, "Hierarchical Higher-Order Regression Forest Fields: An Application to 3D Indoor Scene Labelling", in *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, pp. 2246-2254, 2015.

112. Ng, S. F., Y. M. Chew, P. E. Chng, and K. S. Ng, "An Insight of Linear Regression Analysis", *Scientific Research Journal*, Vol. 15, No. 2, pp. 1-16, 2018.

113. Karch, J., "Improving on Adjusted R-squared", *Collabra: Psychology*, Vol. 6, No. 1, 2020.

114. Schielzeth, H., "Simple Means to Improve the Interpretability of Regression Coefficients", *Methods in Ecology and Evolution*, Vol. 1, No. 2, pp. 103-113, 2010.

115. Gelman, A., "Scaling Regression Inputs by Dividing by Two Standard Deviations", *Statistics in Medicine*, Vol. 27, No. 15, pp. 2865-2873, 2008.

116. Gül, Z. B., "Construction Industry in Turkey: An Input-Output Analysis Using the World Input-Output Database (WIOD) For the 2002-2011 Periods", *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, Vol. 31, No. 1, pp. 157-174, 2017.

117. Bilgin, G., P. Bilgin, İ. Dikmen Toker, and M. T. Birgönül, "Innovation Vision of the Turkish Construction Industry: A Comparative Qualitative Content Analysis of Strategic Roadmaps", presented at the International Conference on Innovation, Technology, Enterprise and Entrepreneurship (ICITEE), Pattaya, Chonburi, Thailand, 2019, Paper No.

118. Alin, A., "Multicollinearity", *Wiley Interdisciplinary Reviews: Computational Statistics*, Vol. 2, No. 3, pp. 370-374, 2010.

119. Chan, J. Y.-L., S. M. H. Leow, K. T. Bea, W. K. Cheng, S. W. Phoong, Z.-W. Hong, and Y.-L. Chen, "Mitigating the Multicollinearity Problem and Its Machine Learning Approach: A Review", *Mathematics*, Vol. 10, No. 8, p. 1283, 2022.

120. Leech, N. L., J. A. Gliner, G. A. Morgan, and R. J. Harmon, “Use and Interpretation of Multiple Regression”, *Journal of the American Academy of Child & Adolescent Psychiatry*, Vol. 42, No. 6, pp. 738-740, 2003.

APPENDIX A: DATA COLLECTION SHEET

Number of storeys	Project location	Construction Year	Construction Duration	Area	Foundation system	Floor system	Framing type	Contract type	Exterior finishing type	Usage of building	Construction Cost TL
6	Istanbul	2012	1.5	200,000	Isolated	Reinforced concrete	Reinforced concrete	Lump-sum	Curtain Wall + Stone	Commercial	81,029,209
24	Istanbul	1999	1.5	50,000	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Curtain Wall	Commercial	1,784,280
8	Istanbul	2009	1.5	7,000	Isolated	Reinforced concrete	Reinforced concrete	Lump-sum	Curtain Wall	Commercial	7,896,150
3	Antalya	2009	2	17,000	Isolated	Reinforced concrete	Reinforced concrete	Lump-sum	Curtain Wall + Plaster	Residential	19,142,182
20	Istanbul	2009	2	34,600	Isolated	Reinforced concrete	Reinforced concrete	Lump-sum	Curtain Wall + Plaster	Residential	9,355,741
6	Istanbul	2006	2	7,900	Isolated	Reinforced concrete	Reinforced concrete	Lump-sum	Curtain Wall + Plaster	Residential	2,141,836
5	Istanbul	2011	2	13,000	Isolated	Reinforced concrete	Reinforced concrete	Unit Price	Plaster	Educational	5,157,971
4	Istanbul	2020	0.666666667	500	Mat	Reinforced concrete	Reinforced concrete	Cost Plus	Brick Coating	Residential	2,982,560
8	Istanbul	2020	1.333333333	4,000	Mat	Reinforced concrete	Reinforced concrete	Cost Plus	Aluminum Façade System	Hospital	8,947,680
15	Istanbul	2022	1.25	3,000	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Aluminum Façade System	Hospital	14,776,230
15	Ankara	2019	3	62,416	Isolated	Reinforced concrete	Reinforced concrete	Lump-sum	Plaster	Residential	60,722,435
10	Ankara	2021	1	21,670	Isolated	Reinforced concrete	Reinforced concrete	Lump-sum	Plaster	Residential and Commercial	20,839,687
15	Ankara	2022	2	41,658	Isolated	Reinforced concrete	Reinforced concrete	Lump-sum	Plaster	Residential	93,017,672
5	Ankara	2022	1	9,289	Isolated	Reinforced concrete	Reinforced concrete	Lump-sum	Plaster	Residential and Commercial	19,289,149
22	Istanbul	2010	5	45,000	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Precast Concrete	Residential and Commercial	40,000,000
11	Istanbul	2016	3	7,500	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Stone	Residential	7,000,000
44	Istanbul	2016	5	120,000	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Precast Concrete	Residential and Commercial	80,000,000
5	Istanbul	2022	2	4,800	Mat	Pan Slab	Reinforced concrete	Lump-sum	Stone	Educational	25,674,579
2	Istanbul	2020	1	1,200	Mat	Steel Systems	Structural Steel	Lump-sum	EIFS + Curtain Wall	Educational	5,108,754

2	Istanbul	2019	2	9,586	Mat	Pan Slab	Structural Steel	Lump-sum	Curtain Wall + Stone	Educational	43,343,386
5	Istanbul	2017	1.5	9,651	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Plaster	Dormitory	11,519,044
6	Istanbul	2020	1	2,500	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS + Curtain Wall	Residential and Commercial	5,200,000
15	Diyarbakir	2012	2	30,000	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	14,500,000
14	Diyarbakir	2012	2	44,500	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	24,200,000
15	Diyarbakir	2015	2	65,000	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	41,700,000
14	Diyarbakir	2010	2	13,322	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	11,200,000
7	Istanbul	2022	1	2,040	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Stone	Residential and Commercial	9,000,000
10	Diyarbakir	2007	1.5	4,070	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	1,125,000
11	Diyarbakir	2006	3	82,350	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	25,350,000
8	Istanbul	2022	1.5	3,333	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	14,705,882
5	Istanbul	2020	1.5	2,083	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential	4,333,333
7	Istanbul	2020	2	2,917	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS + Curtain Wall	Residential	6,066,667
7	Istanbul	2022	1	3,000	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential	6,240,000
10	Istanbul	2019	2	4,286	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	9,688,993
6	Diyarbakir	2018	1.5	5,709	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Stone	Residential	3,104,903
3	Izmir	2022	1	150	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential	750,000
3	Izmir	2022	1	112	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential	672,000
3	Baliksehir	2021	1	146	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential	803,000
15	Diyarbakir	2016	2	6,850	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS + Curtain Wall	Residential and Commercial	5,250,000
14	Diyarbakir	2016	2	6,580	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS + Curtain Wall	Residential and Commercial	4,335,000
15	Diyarbakir	2017	2	6,756	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	6,035,000
13	Diyarbakir	2019	2	5,980	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	7,400,000
9	Diyarbakir	2017	2	2,520	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	2,090,000

9	Diyarbakir	2017	2	2,615	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	2,400,000
9	Diyarbakir	2020	2	2,786	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	3,360,000
9	Diyarbakir	2019	2	2,160	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Curtain Wall + Stone	Residential and Commercial	2,100,000
17	Mardin	2018	2	11,475	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	5,500,000
16	Mardin	2019	2	9,600	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	5,560,000
9	Mardin	2018	2	4,500	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Curtain Wall + Stone	Residential and Commercial	2,600,000
7	Mardin	2019	2	2,730	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	3,150,000
70	Istanbul	2018	1.416666667	26,000	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Stone	Residential and Commercial	60,000,000
70	Istanbul	2021	1.416666667	23,000	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS + Stone	Residential and Commercial	90,000,000
24	Istanbul	2020	1.166666667	7,000	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Stone	Residential	35,000,000
20	Istanbul	2018	2	13,000	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Plaster	Residential and Commercial	25,000,000
16	Istanbul	2022	1	7,500	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Stone	Residential and Commercial	62,000,000
7	Istanbul	2018	1.5	10,000	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Curtain Wall + Stone	Residential and Commercial	10,500,000
6	Istanbul	2020	1.5	7,000	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	Curtain Wall + Stone	Residential and Commercial	21,700,000
11	Istanbul	2019	1.5	6,320	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential	6,076,704
11	Istanbul	2019	2	6,250	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	5,955,702
13	Istanbul	2019	2	7,409	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	8,190,619
15	Istanbul	2019	1.5	3,876	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential	3,227,715
15	Istanbul	2019	1.5	4,010	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential	3,615,337
13	Istanbul	2019	2	7,493	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	7,903,523
13	Istanbul	2019	2	7,388	Mat	Reinforced concrete	Reinforced concrete	Lump-sum	EIFS	Residential and Commercial	7,067,726

APPENDIX B: COST MODELS RESULT

Table B.1. Neural network cost estimation result

Number of Storeys	Project Location	Construction Year	Construction Duration	Area	Foundation System	Floor System	Framing Type	Contract Type	Exterior Finishing Type	Usage of Building	Consturction Turnover	Average Discount Rate	Average Net Minimum Wage	CPI	GDP Per Capita	Inflation Rate	Exchange Rate	Construction Cost	Partition	NN Estimation
6	1	2012	1.5	200000	1	1	1	1	1	2	129.4	16.38	720.45	206.835	11674.94	0.0889	2.13	81029209.1	2	90995801.809
24	1	1999	1.5	50000	2	1	1	1	2	2	16.7	66.96	85.85	35.9	4010.465	0.6487	1.556	1784279.63	1	1351555.690
8	1	2009	1.5	7000	1	1	1	1	2	2	94.6	19.97	679.7	164.323	9044.315	0.0625	2.34	7896150	1	8492728.080
3	2	2009	2	17000	1	1	1	1	3	1	136	19.97	679.7	164.323	9044.315	0.0625	2.34	19142181.8	2	22128362.182
20	1	2009	2	34600	1	1	1	1	3	1	136	19.97	679.7	164.323	9044.315	0.0625	2.34	9355741.36	2	7952380.159
6	1	2006	2	7900	1	1	1	1	3	1	83	23.13	531	128.757	7971.237	0.096	1.745	2141836.36	2	2645167.910
5	1	2011	2	13000	1	1	1	2	4	3	275	14.02	816.75	189.946	11289.13	0.0647	2.442	5157971.08	1	4254481.054
4	1	2020	0.667	500	2	1	1	3	5	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	2982560	1	3537231.102
8	1	2020	1.333	4000	2	1	1	3	6	4	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	8947680	2	7050771.840
15	1	2022	1.25	3000	2	1	1	1	6	4	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	14776230	1	15267266.814
15	3	2019	3	62416	1	1	1	1	4	1	523	18.08	2020	418.236	9194.697	0.1518	5.644	60722435	1	63479070.617
10	3	2021	1	21670	1	1	1	1	4	5	781.524	15.75	2825	561.614	9591.615	0.196	7.229	20839687	1	19567954.255
15	3	2022	2	41658	1	1	1	1	4	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	93017672	1	94099967.215
5	3	2022	1	9289	1	1	1	1	4	5	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	19289149	1	22230543.341
22	1	2010	5	45000	2	1	1	1	7	5	196	14.99	744.75	178.4	10629.47	0.0857	2.448	40000000	2	31160000.000
11	1	2016	3	7500	2	1	1	1	8	1	382	9	1300.99	280.846	10964.46	0.0778	3.034	7000000	1	9227822.653
44	1	2016	5	120000	2	1	1	1	7	5	382	9	1300.99	280.846	10964.46	0.0778	3.034	80000000	2	95840000.268
5	1	2022	2	4800	2	2	1	1	8	3	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	25674579.4	2	25168405.437
2	1	2020	1	1200	2	3	2	1	10	3	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	5108753.88	1	4073378.026
2	1	2019	2	9586	2	2	2	1	11	3	523	18.08	2020	418.236	9194.697	0.1518	5.644	43343385.8	2	36625160.995
5	1	2017	1.5	9651	2	1	1	1	4	6	442	8.75	1777.5	312.144	10696.34	0.1114	3.78	11519044.2	1	15596785.887
6	1	2020	1	2500	2	1	1	1	12	5	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	5200000	1	6011380.625
15	4	2012	2	30000	2	1	1	1	9	5	175	16.38	720.45	206.835	11674.94	0.0889	2.13	14500000	1	14985764.224
14	4	2012	2	44500	2	1	1	1	9	5	175	16.38	720.45	206.835	11674.94	0.0889	2.13	24200000	1	16227046.726
15	4	2015	2	65000	2	1	1	1	13	5	306	9	1000.54	260.585	11085.32	0.0767	2.973	41700000	1	43001649.906
14	4	2010	2	13322	2	1	1	1	13	5	196	14.99	744.75	178.4	10629.47	0.0857	2.448	11200000	1	12835200.287
7	1	2022	1	2040	2	1	1	1	8	5	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	9000000	1	17603033.469
10	4	2007	1.5	4070	2	1	1	1	9	5	91.3	16.98	573.75	140.031	9735.458	0.0876	1.785	1125000	1	60300.187
11	4	2006	3	82350	2	1	1	1	14	5	83	23.13	531	128.757	7971.237	0.096	1.745	25350000	1	24692670.496
8	1	2022	1.5	3333.333	2	1	1	1	9	5	125366.7	14.74	4876.7	967.711	10655.46	0.1561	17.384	14705882.4	2	20073529.412
5	1	2020	1.5	2083.333	2	1	1	1	9	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	4333333.33	1	3894759.623
7	1	2020	2	2916.667	2	1	1	1	10	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	6066666.67	1	7864224.343
7	1	2022	1	3000	2	1	1	1	9	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	6240000	1	7918276.728
10	1	2019	2	4285.714	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	9688992.68	1	5657266.738
6	4	2018	1.5	5709.429	2	1	1	1	8	1	448.5	13.72	1603.12	363.125	9793.028	0.1633	5.697	3104902.73	2	3319141.017
3	5	2022	1	150	2	1	1	1	9	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	750000	1	903614.140
3	5	2022	1	112	2	1	1	1	9	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	672000	1	923328.312
3	6	2021	1	146	2	1	1	1	9	1	781.524	15.75	2825	561.614	9591.615	0.196	7.229	803000	1	2299714.377
15	4	2016	2	6850	2	1	1	1	10	5	382	9	1300.99	280.846	10964.46	0.0778	3.034	5250000	2	4430557.009
14	4	2016	2	6580	2	1	1	1	10	5	382	9	1300.99	280.846	10964.46	0.0778	3.034	4335000	1	3137518.598
15	4	2017	2	6756	2	1	1	1	9	5	442	8.75	1777.5	312.144	10696.34	0.1114	3.78	6035000	1	8074572.640
13	4	2019	2	5980	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	7400000	1	8707811.953
9	4	2017	2	2520	2	1	1	1	9	5	442	8.75	1777.5	312.144	10696.34	0.1114	3.78	2090000	1	1873618.267
9	4	2017	2	2615	2	1	1	1	9	5	442	8.75	1777.5	312.144	10696.34	0.1114	3.78	2400000	1	1912123.094
9	4	2020	2	2786	2	1	1	1	9	5	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	3360000	1	3134718.973
9	4	2019	2	2160	2	1	1	1	11	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	2100000	1	2446500.000
17	7	2018	2	11475	2	1	1	1	9	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	5500000	1	11067774.914
16	7	2019	2	9600	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	5560000	1	6039591.948
9	7	2018	2	4500	2	1	1	1	11	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	2600000	2	2249000.000
7	7	2019	2	2730	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	3150000	1	558863.275
70	1	2018	1.417	26000	2	1	1	1	8	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	60000000	1	63373235.889
70	1	2021	1.417	23000	2	1	1	1	12	5	781.524	15.75	2825	561.614	9591.615	0.196	7.229	90000000	1	90087180.991
24	1	2020	1.167	7000	2	1	1	1	8	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	35000000	1	29800895.562
20	1	2018	2	13000	2	1	1	1	4	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	25000000	2	31447559.346
16	1	2022	1	7500	2	1	1	1	8	5	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	62000000	1	48994696.131
7	1	2018	1.5	10000	2	1	1	1	1	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	10500000	1	13072500.000
6	1	2020	1.5	7000	2	1	1	1	1	5	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	21700000	2	23761500.000
11	1	2019	1.5	6319.5	2	1	1	1	9	1	5.644	0.152	6076703.6	418.236	9194.697	0.1518	5.644	6076703.6	1	1538575.869
11	1	2019	2	6250.2	2	1	1	1	9	5	5.644	0.152	5955702.4	418.236	9194.697	0.1518	5.644	5955702.4	1	4457346.419
13	1	2019	2	7408.7	2	1	1	1	9	5	5.644	0.152	8190619.3	418.236	9194.697	0.1518	5.644	8190619.3	1	12408485.531
15	1	2019	1.5	3876	2	1	1	1	9	1	5.644	0.152	3227715.2	418.236	9194.697	0.1518	5.644	3227715.2	1	4908749.146
15	1	2019	1.5	4009.5	2	1	1	1	9	1	5.644	0.152	3615337	418.236	9194.697	0.1518	5.644	3615337	2	4088912.523
13	1	2019	2	7493.2	2	1	1	1	9	5	5.644	0.152	7903523.4	418.236	9194.697	0.1518	5.644	7903523.4	1	11807902.147
13	1	2019	2	7387.9	2	1	1	1	9	5	5.644	0.152	7067725.5	418.236	9194.697	0.1518	5.644	7067725.5	1	10033236.770

Table B.2. C&R cost estimation result

Number of Storeys	Project Location	Construction Year	Construction Duration	Area	Foundation System	Floor System	Framing Type	Contract Type	Exterior Finishing Type	Usage of Building	Consturction Turnover	Average Discount Rate	Average Net Minimum Wage	CPI	GDP Per Capita	Inflation Rate	Exchange Rate	Construction Cost	Partition	C&R Estimation
6	1	2012	1.5	200000	1	1	1	1	1	2	129.4	16.38	720.45	206.835	11674.94	0.0889	2.13	81029209.09	1	81029209.091
24	1	1999	1.5	50000	2	1	1	1	2	2	16.7	66.96	85.85	35.9	4010.465	0.6487	1.556	1784279.625	1	2419483.172
8	1	2009	1.5	7000	1	1	1	1	2	2	94.6	19.97	679.7	164.323	9044.315	0.0625	2.34	7896150	1	6639857.903
3	2	2009	2	17000	1	1	1	1	3	1	136	19.97	679.7	164.323	9044.315	0.0625	2.34	19142181.82	2	14222641.091
20	1	2009	2	34600	1	1	1	1	3	1	136	19.97	679.7	164.323	9044.315	0.0625	2.34	9355741.364	1	16653219.628
6	1	2006	2	7900	1	1	1	1	3	1	83	23.13	531	128.757	7971.237	0.096	1.745	2141836.364	1	3341264.728
5	1	2011	2	13000	1	1	1	2	4	3	275	14.02	816.75	189.946	11289.13	0.0647	2.442	5157971.08	2	6639857.903
4	1	2020	0.667	500	2	1	1	3	5	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	2982560	1	3545042.054
8	1	2020	1.333	4000	2	1	1	3	6	4	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	8947680	1	11907370.784
15	1	2022	1.25	3000	2	1	1	1	6	4	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	14776230	1	11907370.784
15	3	2019	3	62416	1	1	1	1	4	1	523	18.08	2020	418.236	9194.697	0.1518	5.644	60722435	1	87672557.333
10	3	2021	1	21670	1	1	1	1	4	5	781.524	15.75	2825	561.614	9591.615	0.196	7.229	20839687	1	23257133.210
15	3	2022	2	41658	1	1	1	1	4	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	93017672	1	87672557.333
5	3	2022	1	9289	1	1	1	1	4	5	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	19289149	1	23257133.210
22	1	2010	5	45000	2	1	1	1	7	5	196	14.99	744.75	178.4	10629.47	0.0857	2.448	40000000	1	24726435.341
11	1	2016	3	7500	2	1	1	1	8	1	382	9	1300.99	280.846	10964.46	0.0778	3.034	7000000	1	6639857.903
44	1	2016	5	120000	2	1	1	1	7	5	382	9	1300.99	280.846	10964.46	0.0778	3.034	80000000	1	87672557.333
5	1	2022	2	4800	2	2	1	1	8	3	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	25674579.42	1	23257133.210
2	1	2020	1	1200	2	3	2	1	10	3	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	5108753.877	1	3545042.054
2	1	2019	2	9586	2	2	2	1	11	3	523	18.08	2020	418.236	9194.697	0.1518	5.644	43343385.79	1	27002929.349
5	1	2017	1.5	9651	2	1	1	1	4	6	442	8.75	1777.5	312.144	10696.34	0.1114	3.78	11519044.23	1	6639857.903
6	1	2020	1	2500	2	1	1	1	12	5	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	5200000	1	3545042.054
15	4	2012	2	30000	2	1	1	1	9	5	175	16.38	720.45	206.835	11674.94	0.0889	2.13	14500000	2	20184000.819
14	4	2012	2	44500	2	1	1	1	9	5	175	16.38	720.45	206.835	11674.94	0.0889	2.13	24200000	1	24726435.341
15	4	2015	2	65000	2	1	1	1	13	5	306	9	1000.54	260.585	11085.32	0.0767	2.973	41700000	2	49950762.573
14	4	2010	2	13322	2	1	1	1	13	5	196	14.99	744.75	178.4	10629.47	0.0857	2.448	11200000	1	6639857.903
7	1	2022	1	2040	2	1	1	1	8	5	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	9000000	1	5337000.176
10	4	2007	1.5	4070	2	1	1	1	9	5	91.3	16.98	573.75	140.031	9735.458	0.0876	1.785	1125000	1	6639857.903
11	4	2006	3	82350	2	1	1	1	14	5	83	23.13	531	128.757	7971.237	0.096	1.745	25350000	1	24726435.341
8	1	2022	1.5	3333.333	2	1	1	1	9	5	125366.7	14.74	4876.7	967.711	10655.46	0.1561	17.384	14705882.35	1	11907370.784
5	1	2020	1.5	2083.333	2	1	1	1	9	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	4333333.333	1	3545042.054
7	1	2020	2	2916.667	2	1	1	1	10	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	6066666.667	1	3545042.054
7	1	2022	1	3000	2	1	1	1	9	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	6240000	1	11907370.784
10	1	2019	2	4285.714	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	9688992.681	2	6639857.903
6	4	2018	1.5	5709.429	2	1	1	1	8	1	448.5	13.72	1603.12	363.125	9793.028	0.1633	5.697	3104902.729	2	3461966.543
3	5	2022	1	150	2	1	1	1	9	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	750000	1	3545042.054
3	5	2022	1	112	2	1	1	1	9	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	672000	1	3545042.054
3	6	2021	1	146	2	1	1	1	9	1	781.524	15.75	2825	561.614	9591.615	0.196	7.229	803000	1	3545042.054
15	4	2016	2	6850	2	1	1	1	10	5	382	9	1300.99	280.846	10964.46	0.0778	3.034	5250000	1	6639857.903
14	4	2016	2	6580	2	1	1	1	10	5	382	9	1300.99	280.846	10964.46	0.0778	3.034	4335000	1	6639857.903
15	4	2017	2	6756	2	1	1	1	9	5	442	8.75	1777.5	312.144	10696.34	0.1114	3.78	6035000	1	6657964.509
13	4	2019	2	5980	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	7400000	1	6839877.951
9	4	2017	2	2520	2	1	1	1	9	5	442	8.75	1777.5	312.144	10696.34	0.1114	3.78	2090000	1	3545042.054
9	4	2017	2	2615	2	1	1	1	9	5	442	8.75	1777.5	312.144	10696.34	0.1114	3.78	2400000	1	3545042.054
9	4	2020	2	2786	2	1	1	1	9	5	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	3360000	1	3545042.054
9	4	2019	2	2160	2	1	1	1	11	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	2100000	2	2454900.874
17	7	2018	2	11475	2	1	1	1	9	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	5500000	1	24353333.333
16	7	2019	2	9600	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	5560000	1	24353333.333
9	7	2018	2	4500	2	1	1	1	11	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	2600000	1	6639857.903
7	7	2019	2	2730	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	3150000	1	3545042.054
70	1	2018	1.417	26000	2	1	1	1	8	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	60000000	1	87672557.333
70	1	2021	1.417	23000	2	1	1	1	12	5	781.524	15.75	2825	561.614	9591.615	0.196	7.229	90000000	1	87672557.333
24	1	2020	1.167	7000	2	1	1	1	8	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	35000000	2	27860000.000
20	1	2018	2	13000	2	1	1	1	4	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	25000000	1	24353333.333
16	1	2022	1	7500	2	1	1	1	8	5	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	62000000	1	24981653.359
7	1	2018	1.5	10000	2	1	1	1	1	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	10500000	1	6639857.903
6	1	2020	1.5	7000	2	1	1	1	1	5	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	21700000	1	23257133.210
11	1	2019	1.5	6319.5	2	1	1	1	9	1	5.644	0.152	6076703.6	418.236	9194.697	0.1518	5.644	6076703.6	1	6547362.903
11	1	2019	2	6250.2	2	1	1	1	9	5	5.644	0.152	5955702.4	418.236	9194.697	0.1518	5.644	5955702.4	1	6745796.275
13	1	2019	2	7408.7	2	1	1	1	9	5	5.644	0.152	8190619.3	418.236	9194.697	0.1518	5.644	8190619.3	1	7637542.875
15	1	2019	1.5	3876	2	1	1	1	9	1	5.644	0.152	3227715.2	418.236	9194.697	0.1518	5.644	3227715.2	1	3639857.903
15	1	2019	1.5	4009.5	2	1	1	1	9	1	5.644	0.152	3615337	418.236	9194.697	0.1518	5.644	3615337	1	2789145.573
13	1	2019	2	7493.2	2	1	1	1	9	5	5.644	0.152	7903523.4	418.236	9194.697	0.1518	5.644	7903523.4	1	7639873.276
13	1	2019	2	7387.9	2	1	1	1	9	5	5.644	0.152	7067725.5	418.236	9194.697	0.1518	5.644	7067725.5	1	6679657.574

Table B.3. Regression cost estimation result

Number of Storeys	Project Location	Construction Year	Construction Duration	Area	Foundation System	Floor System	Framing Type	Contract Type	Exterior Finishing Type	Usage of Building	Consturction Turnover	Average Discount Rate	Average Net Minimum Wage	CPI	GDP Per Capita	Inflation Rate	Exchange Rate	Construction Cost	Partition	Regression Estimation
6	1	2012	1.5	200000	1	1	1	1	1	2	129.4	16.38	720.45	206.835	11674.944	0.0889	2.13	81029209.09	1	81320345.304
24	1	1999	1.5	50000	2	1	1	1	2	2	16.7	66.96	85.85	35.9	4010.465	0.6487	1.556	1784279.625	1	2800962.155
8	1	2009	1.5	7000	1	1	1	1	2	2	94.6	19.97	679.7	164.323	9044.315	0.0625	2.34	7896150	1	5372692.365
3	2	2009	2	17000	1	1	1	1	3	1	136	19.97	679.7	164.323	9044.315	0.0625	2.34	19142181.82	2	12365179.478
20	1	2009	2	34600	1	1	1	1	3	1	136	19.97	679.7	164.323	9044.315	0.0625	2.34	9355741.364	1	13624988.815
6	1	2006	2	7900	1	1	1	1	3	1	83	23.13	531	128.757	7971.237	0.096	1.745	2141836.364	1	1590576.705
5	1	2011	2	13000	1	1	1	2	4	3	275	14.02	816.75	189.946	11289.129	0.0647	2.442	5157971.08	2	7045257.224
4	1	2020	0.667	500	2	1	1	3	5	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	2982560	1	2354581.992
8	1	2020	1.333	4000	2	1	1	3	6	4	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	8947680	1	11628880.547
15	1	2022	1.25	3000	2	1	1	1	6	4	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	14776230	1	21981060.643
15	3	2019	3	62416	1	1	1	1	4	1	523	18.08	2020	418.236	9194.697	0.1518	5.644	60722435	1	51398955.100
10	3	2021	1	21670	1	1	1	1	4	5	781.524	15.75	2825	561.614	9591.615	0.196	7.229	20839687	1	34257334.564
15	3	2022	2	41658	1	1	1	1	4	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	93017672	1	57133691.351
5	3	2022	1	9289	1	1	1	1	4	5	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	19289149	1	35989082.989
22	1	2010	5	45000	2	1	1	1	7	5	196	14.99	744.75	178.4	10629.467	0.0857	2.448	40000000	1	40068802.242
11	1	2016	3	7500	2	1	1	1	8	1	382	9	1300.99	280.846	10964.461	0.0778	3.034	7000000	1	11768414.993
44	1	2016	5	120000	2	1	1	1	7	5	382	9	1300.99	280.846	10964.461	0.0778	3.034	80000000	1	94954399.701
5	1	2022	2	4800	2	2	1	1	8	3	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	25674579.42	1	16036719.263
2	1	2020	1	1200	2	3	2	1	10	3	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	5108753.877	1	7912419.102
2	1	2019	2	9586	2	2	2	1	11	3	523	18.08	2020	418.236	9194.697	0.1518	5.644	43343385.79	1	33705525.637
5	1	2017	1.5	9651	2	1	1	1	4	6	442	8.75	1777.5	312.144	10696.342	0.1114	3.78	11519044.23	1	9159351.893
6	1	2020	1	2500	2	1	1	1	12	5	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	5200000	1	4667318.137
15	4	2012	2	30000	2	1	1	1	9	5	175	16.38	720.45	206.835	11674.944	0.0889	2.13	14500000	2	13490527.285
14	4	2012	2	44500	2	1	1	1	9	5	175	16.38	720.45	206.835	11674.944	0.0889	2.13	24200000	1	17911634.883
15	4	2015	2	65000	2	1	1	1	13	5	306	9	1000.54	260.585	11085.318	0.0767	2.973	41700000	2	32698083.390
14	4	2010	2	13322	2	1	1	1	13	5	196	14.99	744.75	178.4	10629.467	0.0857	2.448	11200000	1	7598299.520
7	1	2022	1	2040	2	1	1	1	8	5	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	9000000	1	15253392.816
10	4	2007	1.5	4070	2	1	1	1	9	5	91.3	16.98	573.75	140.031	9735.458	0.0876	1.785	1125000	1	548957.813
11	4	2006	3	82350	2	1	1	1	14	5	83	23.13	531	128.757	7971.237	0.096	1.745	25350000	1	23212370.963
8	1	2022	1.5	3333.333	2	1	1	1	9	5	125366.7	14.74	4876.7	967.711	10655.46	0.1561	17.384	14705882.35	1	14696795.445
5	1	2020	1.5	2083.333	2	1	1	1	9	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	4333333.333	1	2978959.477
7	1	2020	2	2916.667	2	1	1	1	10	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	6066666.667	1	6273313.418
7	1	2022	1	3000	2	1	1	1	9	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	6240000	1	11699725.440
10	1	2019	2	4285.714	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	9688992.681	2	14053807.975
6	4	2018	1.5	5709.429	2	1	1	1	8	1	448.5	13.72	1603.12	363.125	9793.028	0.1633	5.697	3104902.729	2	3594999.205
3	5	2022	1	150	2	1	1	1	9	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	750000	1	1123922.250
3	5	2022	1	112	2	1	1	1	9	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	672000	1	1115512.339
3	6	2021	1	146	2	1	1	1	9	1	781.524	15.75	2825	561.614	9591.615	0.196	7.229	803000	1	1187245.136
15	4	2016	2	6850	2	1	1	1	10	5	382	9	1300.99	280.846	10964.461	0.0778	3.034	5250000	1	9145979.296
14	4	2016	2	6580	2	1	1	1	10	5	382	9	1300.99	280.846	10964.461	0.0778	3.034	4335000	1	8196479.633
15	4	2017	2	6756	2	1	1	1	9	5	442	8.75	1777.5	312.144	10696.342	0.1114	3.78	6035000	1	9976448.888
13	4	2019	2	5980	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	7400000	1	12035633.824
9	4	2017	2	2520	2	1	1	1	9	5	442	8.75	1777.5	312.144	10696.342	0.1114	3.78	2090000	1	3328231.641
9	4	2017	2	2615	2	1	1	1	9	5	442	8.75	1777.5	312.144	10696.342	0.1114	3.78	2400000	1	3362775.154
9	4	2020	2	2786	2	1	1	1	9	5	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	3360000	1	4245688.608
9	4	2019	2	2160	2	1	1	1	11	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	2100000	2	2616177.133
17	7	2018	2	11475	2	1	1	1	9	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	5500000	1	8309894.214
16	7	2019	2	9600	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	5560000	1	10717680.061
9	7	2018	2	4500	2	1	1	1	11	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	2600000	1	1765257.354
7	7	2019	2	2730	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	3150000	1	1474141.859
70	1	2018	1.417	26000	2	1	1	1	8	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	60000000	1	66007895.100
70	1	2021	1.417	23000	2	1	1	1	12	5	781.524	15.75	2825	561.614	9591.615	0.196	7.229	90000000	1	77225694.077
24	1	2020	1.167	7000	2	1	1	1	8	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	35000000	2	47136561.850
20	1	2018	2	13000	2	1	1	1	4	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	25000000	1	21148991.936
16	1	2022	1	7500	2	1	1	1	8	5	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	62000000	1	33484026.589
7	1	2018	1.5	10000	2	1	1	1	1	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	10500000	1	6074089.681
6	1	2020	1.5	7000	2	1	1	1	1	5	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	21700000	1	17032688.968
11	1	2019	1.5	6319.5	2	1	1	1	9	1	5.644	0.152	6076703.6	418.236	9194.697	0.1518	5.644	6076703.6	1	4829864.027
11	1	2019	2	6250.2	2	1	1	1	9	5	5.644	0.152	5955702.4	418.236	9194.697	0.1518	5.644	5955702.4	1	7632807.990
13	1	2019	2	7408.7	2	1	1	1	9	5	5.644	0.152	8190619.3	418.236	9194.697	0.1518	5.644	8190619.3	1	7464543.904
15	1	2019	1.5	3876	2	1	1	1	9	1	5.644	0.152	3227715.2	418.236	9194.697	0.1518	5.644	3227715.2	1	6246066.789
15	1	2019	1.5	4009.5	2	1	1	1	9	1	5.644	0.152	3615337	418.236	9194.697	0.1518	5.644	3615337	1	5897059.552
13	1	2019	2	7493.2	2	1	1	1	9	5	5.644	0.152	7903523.4	418.236	9194.697	0.1518	5.644	7903523.4	1	7789718.742
13	1	2019	2	7387.9	2	1	1	1	9	5	5.644	0.152	7067725.5	418.236	9194.697	0.1518	5.644	7067725.5	1	8608635.191

Table B.4. CHAID cost estimation result

Number of Storeys	Project Location	Construction Year	Construction Duration	Area	Foundation System	Floor System	Framing Type	Contract Type	Exterior Finishing Type	Usage of Building	Consturction Turnover	Average Discount Rate	Average Net Minimum Wage	CPI	GDP Per Capita	Inflation Rate	Exchange Rate	Construction Cost	Partition	CHAID Estimation
6	1	2012	1.5	200000	1	1	1	1	1	2	129.4	16.38	720.45	206.835	11674.94	0.0889	2.13	81029209	1	66121164.511
24	1	1999	1.5	50000	2	1	1	1	2	2	16.7	66.96	85.85	35.9	4010.465	0.6487	1.556	1784280	1	1784279.625
8	1	2009	1.5	7000	1	1	1	1	2	2	94.6	19.97	679.7	164.323	9044.315	0.0625	2.34	7896150	1	16273595.329
3	2	2009	2	17000	1	1	1	1	3	1	136	19.97	679.7	164.323	9044.315	0.0625	2.34	19142182	2	16273595.329
20	1	2009	2	34600	1	1	1	1	3	1	136	19.97	679.7	164.323	9044.315	0.0625	2.34	9355741	1	9355741.364
6	1	2006	2	7900	1	1	1	1	3	1	83	23.13	531	128.757	7971.237	0.096	1.745	2141836	1	4064233.711
5	1	2011	2	13000	1	1	1	2	4	3	275	14.02	816.75	189.946	11289.13	0.0647	2.442	5157971	2	7683392.287
4	1	2020	0.667	500	2	1	1	3	5	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	2982560	1	2982560.000
8	1	2020	1.333	4000	2	1	1	3	6	4	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	8947680	1	8947680.000
15	1	2022	1.25	3000	2	1	1	1	6	4	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	14776230	1	14776230.000
15	3	2019	3	62416	1	1	1	1	4	1	523	18.08	2020	418.236	9194.697	0.1518	5.644	60722435	1	66121164.511
10	3	2021	1	21670	1	1	1	1	4	5	781.524	15.75	2825	561.614	9591.615	0.196	7.229	20839687	1	16273595.329
15	3	2022	2	41658	1	1	1	1	4	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	93017672	1	66121164.511
5	3	2022	1	9289	1	1	1	1	4	5	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	19289149	1	16273595.329
22	1	2010	5	45000	2	1	1	1	7	5	196	14.99	744.75	178.4	10629.47	0.0857	2.448	40000000	1	66121164.511
11	1	2016	3	7500	2	1	1	1	8	1	382	9	1300.99	280.846	10964.46	0.0778	3.034	7000000	1	16273595.329
44	1	2016	5	120000	2	1	1	1	7	5	382	9	1300.99	280.846	10964.46	0.0778	3.034	80000000	1	66121164.511
5	1	2022	2	4800	2	2	1	1	8	3	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	25674579	1	25674579.419
2	1	2020	1	1200	2	3	2	1	10	3	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	5108754	1	5108753.877
2	1	2019	2	9586	2	2	2	1	11	3	523	18.08	2020	418.236	9194.697	0.1518	5.644	43343386	1	16273595.329
5	1	2017	1.5	9651	2	1	1	1	4	6	442	8.75	1777.5	312.144	10696.34	0.1114	3.78	11519044	1	16273595.329
6	1	2020	1	2500	2	1	1	1	12	5	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	5200000	1	5200000.000
15	4	2012	2	30000	2	1	1	1	9	5	175	16.38	720.45	206.835	11674.94	0.0889	2.13	14500000	2	20117531.464
14	4	2012	2	44500	2	1	1	1	9	5	175	16.38	720.45	206.835	11674.94	0.0889	2.13	24200000	1	66121164.511
15	4	2015	2	65000	2	1	1	1	13	5	306	9	1000.54	260.585	11085.32	0.0767	2.973	41700000	2	66121164.511
14	4	2010	2	13322	2	1	1	1	13	5	196	14.99	744.75	178.4	10629.47	0.0857	2.448	11200000	1	16273595.329
7	1	2022	1	2040	2	1	1	1	8	5	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	9000000	1	9000000.000
10	4	2007	1.5	4070	2	1	1	1	9	5	91.3	16.98	573.75	140.031	9735.458	0.0876	1.785	1125000	1	2454166.667
11	4	2006	3	82350	2	1	1	1	14	5	83	23.13	531	128.757	7971.237	0.096	1.745	25350000	1	25350000.000
8	1	2022	1.5	3333.333	2	1	1	1	9	5	125366.7	14.74	4876.7	967.711	10655.46	0.1561	17.384	14705882	1	14705882.353
5	1	2020	1.5	2083.333	2	1	1	1	9	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	4333333	1	4333333.333
7	1	2020	2	2916.667	2	1	1	1	10	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	6066667	1	6066666.667
7	1	2022	1	3000	2	1	1	1	9	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	6240000	1	6240000.000
10	1	2019	2	4285.714	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	9688993	2	13543676.673
6	4	2018	1.5	5709.429	2	1	1	1	8	1	448.5	13.72	1603.12	363.125	9793.028	0.1633	5.697	3104903	2	4003824.231
3	5	2022	1	150	2	1	1	1	9	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	750000	1	741666.667
3	5	2022	1	112	2	1	1	1	9	1	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	672000	1	741666.667
3	6	2021	1	146	2	1	1	1	9	1	781.524	15.75	2825	561.614	9591.615	0.196	7.229	803000	1	741666.667
15	4	2016	2	6850	2	1	1	1	10	5	382	9	1300.99	280.846	10964.46	0.0778	3.034	5250000	1	5795140.480
14	4	2016	2	6580	2	1	1	1	10	5	382	9	1300.99	280.846	10964.46	0.0778	3.034	4335000	1	5795140.480
15	4	2017	2	6756	2	1	1	1	9	5	442	8.75	1777.5	312.144	10696.34	0.1114	3.78	6035000	1	5795140.480
13	4	2019	2	5980	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	7400000	1	5795140.480
9	4	2017	2	2520	2	1	1	1	9	5	442	8.75	1777.5	312.144	10696.34	0.1114	3.78	2090000	1	2454166.667
9	4	2017	2	2615	2	1	1	1	9	5	442	8.75	1777.5	312.144	10696.34	0.1114	3.78	2400000	1	2454166.667
9	4	2020	2	2786	2	1	1	1	9	5	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	3360000	1	2454166.667
9	4	2019	2	2160	2	1	1	1	11	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	2100000	2	2454166.667
17	7	2018	2	11475	2	1	1	1	9	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	5500000	1	16273595.329
16	7	2019	2	9600	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	5560000	1	16273595.329
9	7	2018	2	4500	2	1	1	1	11	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	2600000	1	2454166.667
7	7	2019	2	2730	2	1	1	1	9	5	523	18.08	2020	418.236	9194.697	0.1518	5.644	3150000	1	2454166.667
70	1	2018	1.417	26000	2	1	1	1	8	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	60000000	1	66121164.511
70	1	2021	1.417	23000	2	1	1	1	12	5	781.524	15.75	2825	561.614	9591.615	0.196	7.229	90000000	1	66121164.511
24	1	2020	1.167	7000	2	1	1	1	8	1	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	35000000	2	16273595.329
20	1	2018	2	13000	2	1	1	1	4	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	25000000	1	16273595.329
16	1	2022	1	7500	2	1	1	1	8	5	1569.474	14.74	4876.7	967.711	10655.46	0.1561	17.384	62000000	1	16273595.329
7	1	2018	1.5	10000	2	1	1	1	1	5	445.8	13.72	1603.12	363.125	9793.028	0.1633	5.697	10500000	1	16273595.329
6	1	2020	1.5	7000	2	1	1	1	1	5	697	10.92	2324.7	469.591	8599.767	0.1228	7.456	21700000	1	16273595.329
11	1	2019	1.5	6319.5	2	1	1	1	9	1	5.644	0.152	6076704	418.236	9194.697	0.1518	5.644	6076704	1	6076703.600
11	1	2019	2	6250.2	2	1	1	1	9	5	5.644	0.152	5955702	418.236	9194.697	0.1518	5.644	5955702	1	5795140.480
13	1	2019	2	7408.7	2	1	1	1	9	5	5.644	0.152	8190619	418.236	9194.697	0.1518	5.644	8190619	1	16273595.329
15	1	2019	1.5	3876	2	1	1	1	9	1	5.644	0.152	3227715	418.236	9194.697	0.1518	5.644	3227715	1	3227715.200
15	1	2019	1.5	4009.5	2	1	1	1	9	1	5.644	0.152	3615337	418.236	9194.697	0.1518	5.644	3615337	1	3615337.000
13	1	2019	2	7493.2	2	1	1	1	9	5	5.644	0.152	7903523	418.236	9194.697	0.1518	5.644	7903523	1	16273595.329
13	1	2019	2	7387.9	2	1	1	1	9	5	5.644	0.152	7067726	418.236	9194.697	0.1518	5.644	7067726	1	16273595.329