

Article

An Innovative Approach Regarding Efficient and Expedited Early Building Renovation Cost Estimation Utilizing ANNs and the TOPSIS Methodology

Vasso E. Papadimitriou * and Georgios N. Aretoulis * 

Laboratory of Planning and Project Management, School of Civil Engineering, Faculty of Engineering, Aristotle University of Thessaloniki, Aristotle University of Thessaloniki Campus, 54124 Thessaloniki, Greece

* Correspondence: vpapad@civil.auth.gr (V.E.P.); garet@civil.auth.gr (G.N.A.)

Abstract

Early cost assessment is an essential part of building construction strategy; however, preliminary estimates are occasionally unreliable given incomplete data, which causes budgetary overruns. In general, traditional prediction techniques are imprecise and sluggish, particularly while the project scope is still unclear. By introducing a hybrid framework that utilizes ANNs for renovation cost estimation and features enhancements by the TOPSIS method to guarantee contextual relevance and input accuracy, the present research overcomes these drawbacks. Utilizing data from projects that are structurally and contextually comparable enhances the model's predicted reliability and robustness. The study builds, trains, and tests two ANN models using IBM SPSS Statistics software, which is based on a thorough literature review and actual renovation data from construction businesses. One model utilized 53 data points from prior building renovation projects, whereas the second model employed 11 data points from post-TOPSIS technique building renovation projects. The Radial Basis Function (RBF) procedure is the basis for models that include 14 input data such as total initial cost, estimated completion time, initial demolition drainage cost, initial cost of plumbing work, initial heating cost, initial cost of electrical work, initial cost of masonry coatings, initial cost of plasterboard construction, initial bathroom cost, initial flooring costs, initial frame cost, initial door cost, initial paint cost, and initial kitchen construction cost, and one output data, the total final cost. The models show excellent performance with near 0.5 relative error and up to 0.3 monetary units sum of squares error before applying the TOPSIS method and nearly 0.6 relative error and up to 0.8 monetary units sum of squares error after the TOPSIS implementation, proving the usefulness and demonstrating the speed of the ANN in estimating overall renovation costs in combination with the TOPSIS approach. By employing this hybridized approach, the entire contingent procedure is expedited and accomplished more rapidly.

Keywords: buildings renovation; cost estimation; building cost estimation models; TOPSIS method; artificial neural network; radial basis function algorithm



Academic Editor: Jaafar Alghazo and Ghazanfar Latif

Received: 22 September 2025

Revised: 15 October 2025

Accepted: 20 October 2025

Published: 3 November 2025

Citation: Papadimitriou, V.E.; Aretoulis, G.N. An Innovative Approach Regarding Efficient and Expedited Early Building Renovation Cost Estimation Utilizing ANNs and the TOPSIS Methodology. *Algorithms* **2025**, *18*, 696. <https://doi.org/10.3390/a18110696>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The building sector is one of the major contributors to the Gross National Product (GNP) and thus is essential to the financial progress and expansion of all countries. This industry includes a broad range of operations, such as infrastructure development and the maintenance and renovation of already existing structures, as well as building projects for

residential, commercial, and industrial use. The construction sector is a vital component of economic growth, generating jobs, stimulating adjacent businesses, building infrastructure, and directly contributing to the GNP. Its wide-ranging effects on many economic sectors highlight how important it is to promoting national prosperity and raising living standards [1].

In recent decades, ANNs have been widely used to address difficult non-linear issues in construction engineering and management. The use of ANNs in construction cost prediction, schedule estimation, efficiency forecasting, dispute occurrence, resolution outcome prediction, and contract performance demonstrates the potential and resilience of ANNs in solving problems that were challenging for conventional mathematical and statistical methods. Fundamentally, the quantity and accuracy of the data utilized for training algorithms is crucial for the results of forecasts, identification, and classification, as demonstrated by Waziri et al. [2], given the fact that ANN efficiency is data-sensitive.

In the fields of construction management and engineering, ANNs are an efficient method for parametric modeling, along with being frequently employed as a substitute to traditional modeling approaches, particularly for datasets with non-linear relationships. They also serve as the foundation for decision-support software that use algorithms that are learned under supervision to reach the optimal choices possible. Compared with additional parametric models like regression analysis, they perform exceptionally well in these kinds of applications. They can map learning uncertainty and are thought to be the best for short-term forecasting [2].

Despite its continuous growth, the construction industry still confronts long-standing issues including excessive risk, inadequate quality work, excess expenditure, and delays. The success of a project is determined by the triple limitations of time, money, and quality, according to all industry participants [3].

This study's innovation is the incorporation of a placement strategy from the discipline of Artificial Intelligence, specifically ANNs, and its application to the adequate construction projects sector. The implementation of ANNs as a methodological framework for the accurate prediction of the overall cost of buildings renovation projects is investigated in present research. The Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) is utilized as a decision-making tool to confirm that the most representative and relevant input data is selected. By employing verified project data as a foundation, this integrative approach seeks to improve the predictive model's robustness and dependability. Only a small number of scholars have conducted research into the feasibility of utilizing forecasting algorithms to estimate building renovation construction costs. Nonetheless, as far as the writers are aware, no thorough analysis of the usage of that kind of AI hybrid technology in the building sector has been performed [4,5].

Although budget overruns are a common occurrence in preliminary cost estimation for building renovation projects, the persistent variability and severity of the estimated inaccuracy present a fundamental issue. This large variation undermines excellent financial management since major cost variances impede optimal capital allocation, both positively (overruns jeopardize project viability) and negatively (excessive contingencies signify opportunity cost). The present research thus tackles the scientific requirement to reduce the sum of squares and relative error. Providing a solid and reliable single-point estimate is the goal, with the objective of reducing financial risk and improving the effectiveness of early project decision-making procedures.

Both the RBF-ANN and TOPSIS approaches are acknowledged as proven techniques. However, this research's methodological and scientific originality stems from the hybrid architecture and the specific information flow established to resolve a single, high-variance estimation problem. A unique scientific innovation has been made with the development

of this multi-stage forecast technique that reaches beyond simple component concatenation. TOPSIS is utilized specifically as a sophisticated, multi-criteria similarity filter rather than as a typical decision-making tool.

The empowered model is developed using this novel strategy, which offers a more dependable, unbiased, and contextually relevant method for choosing the RBF training inputs. The financial issue of uncertain expenditure volatility is effectively addressed by the proposed model, which yields a robust, highly accurate forecasting tool for the volatile field of renovation cost estimation. It offers a quantifiable reduction in sum of squares and relative error when compared to benchmark models.

This article's sections are arranged as follows: Section 2 provides a thorough analysis of the literature on ANNs as a tool for building projects and the construction sector in general, and specifically in building renovation projects. The formulation of the problem is demonstrated in Section 3. Section 4 describes the methodology applied to the data collection and analysis procedures for creating RBF ANN models to forecast and compare the final costs of building renovations. To reduce the failure rate when using ANNs, the authors recommend using the TOPSIS method to assist in gathering the most pertinent input data. Section 5 presents the proposed solution and results and provides clarification. Section 6 contains discussion, and, lastly, Section 7 describes conclusions, limitations, and suggestions for further research.

2. Literature Review

An outline of recent research on the use of ANNs and associated AI methods for cost prediction in building projects is given in this section.

Artificial Intelligence (AI) and Machine Learning (ML) techniques have become effective tools for solving enduring challenges such as cost estimation in building construction projects. They offer robust ways to handle non-linear difficulties, map intricate input/output inter actions, and produce more precise forecasts through training on previous data [6,7].

Due to its intrinsic complexity, which includes unique projects, many variable components, and substantial uncertainties, the construction sector usually has difficulties in accurately estimating costs. In the early phases of a project, when there is a lack of specific information, typical estimation techniques are frequently laborious, subjective, and prone to errors [8].

Specific usage and artificial revolution applications of AI/ML are being incorporated into more general online environments and are also applicable to specific building scenarios.

The use of ANNs in a variety of construction management domains is thoroughly investigated by Kulkarni et al. [9] It summarizes research from more than 100 published sources, emphasizing the value of ANNs in fields including cost forecasting among others.

Through an examination of more than 100 resources, Waziri et al. [2] analyses the patterns and emerging prospects of ANN usage for building and leadership. It discusses how ANNs have been successfully utilized in fields including risk evaluation, organizing, cost forecasting, and optimization, frequently pointing out better outcomes when combined with additional soft computing techniques.

Also, in their 2019 study, Hasware et al. [10] concentrate on the application of ANNs and other methods of AI to anticipate the cost of building materials. The authors emphasize the substantial influence that variations in material costs have on the performance and cost of projects. The article suggests that AI technologies, especially ANNs, are ideal for analyzing historical pricing and the factors that influence them. The result allows for more precise projections of potential expenses for materials and is an essential aid for project planning decision-making processes.

A thorough comparative analysis of six popular machine learning algorithms—Logistic Regression (LR), Decision Tree (DT), Support Vector Machine (SVM), Naïve Bayes (NB), Artificial Neural Network (ANN), and Random Forest (RF)—was carried out by Amornsamankul et al. [11]. Their study focused on the utilization of algorithms, advantages, disadvantages, and efficacy in a variety of industries, including machine learning techniques to create a model for predicting property prices and increase forecast accuracy more effectively. The tabular synthesis of algorithm performance across over 40 empirical investigations, which emphasizes software utilized, training efficacy, and precision, is a significant addition. To highlight the value of accessible systems like Weka and R in easing algorithm completion, the authors stress the significance of choosing machine learning algorithms according to information dimensions, complexity, and intended use.

A survey of many unconventional techniques used to forecast cost overruns in building projects is presented by Gotlur et al. [12]. The authors address the significance of figuring out what causes cost overruns and how various approaches—like fuzzy logic, regression analysis, and ANNs—are utilized to measure and examine pertinent data. The evaluation compares the efficacy of each approach at creating forecasting models for price increases in building projects. It is a useful tool for learning how sophisticated analytical methods are applied to project economic protection.

An overview of the use of ANNs for building material cost prediction is presented by Dilip and Jesna [13]. Considering the regular and substantial swings in material prices, the researchers stress how crucial precise material price prediction is to the profitability of construction companies. Their research highlights the value of ANN models in assisting contractors in tracking historical pricing and predicting future trends by synthesizing prior research that has effectively used them for such forecasts.

Abioye et al. [14] analyzes whether AI, like visual recognition and ANNs, might address important building concerns like safety, overinflated prices, and even delays. It emphasizes how AI is becoming increasingly important in the management of resources, reducing waste, predicting risks and mitigation, surveillance, and mechanization. Authors stress that data fragmentation, lack of AI knowledge, and worker opposition are the main obstacles to adoption.

Additionally, the Comprehensive AI-Driven Cost Dynamics Model (AICD-CDM) combines several modern methods for machine learning, such as Linear Regression (LR), ANNs, Random Forest (RF), XGBoost, Light Gradient Boosting (LGBost), and Natural Gradient Boosting (NGBoost), to provide and demonstrate probabilistic predictions for estimating the cost of sustainable green buildings, while successfully managing inherent unpredictability [15].

A thorough review of the developments in AI technologies for cost estimates in projects is also given by Islam Shamim et al. [16]. The authors examine and evaluate the contributions of many AI approaches, such as machine learning, deep learning, regression, and hybrid models, to project performance. The study summarizes the body of work to pinpoint the main themes, advantages, and difficulties of incorporating AI in a cost-estimating process, providing insightful information such as how ANNs could improve forecast precision and flexibility in intricate, changing project settings.

Furthermore, ANNs can anticipate the most affordable tender price for certain types of buildings, such as primary and secondary school buildings, with high accuracy. Research has demonstrated that models with additional variables frequently perform more effectively, with average prediction percentages of 79.3% and 82.2% in tests [17].

An accurate ANN model for estimating building project expenses at an early stage was created by Arafa et al. [18]. The model had been trained on past pricing information from 71 construction endeavors in the Gaza Strip using a multi-layer neural network that feeds

forward using the back-propagation technique, proving its ability to accurately anticipate both the fundamental foundation and actual expenses.

Another approach is a Fuzzy Neural Network (FNN) method for calculating the determination of the completion expenses of building projects, which is presented by Feylizadeh et al.'s research [19]. Compared to conventional earned value estimation approaches, this method aims to deliver a more thorough and accurate projection of project completion costs by integrating both qualitative and quantitative factors.

The Multilayer Perceptron Neural Network (MLPNN) algorithm for estimating the length of construction projects is presented by Petruseva et al. [20] to show that the use of ANN models considerably increases the accuracy of predictions when compared to linear regression techniques, utilizing data from 75 construction projects in Bosnia and Herzegovina.

In addition, using multiple regression and neural network technology, Gulcicek et al. [21] examined the effects of several parameters on the load-carrying system cost of strengthened concrete residential estate structures, including the building significance factor, seismic zone, soil type, floor size, and number of stories. It indicated that the two most important cost factors were the seismic zone and the building significance factor.

A new model for the life-cycle cost evaluation (LCCA) of building tasks was created by Alqahtani and Whyte [6]. It incorporates ANNs with the idea of Cost-Significant Items (CSIs). Their model demonstrated that ANNs may handle complicated, non-linear issues and uncertainties, owing to their outstanding precision (1–2% inaccuracy) in calculating overall running expenses.

Also, Lhee et al. [22] provide a new two-step neural-network-based approach to estimating the ideal cost contingency for transport building projects. The model learns from historical project data from the state's government department of infrastructure to increase the reliability of contingency forecasting, which will enhance financial performance and resource utilization.

The efficacy of many ANN topologies, such as the Generalized Regression Network (GRNN), Probabilistic Neural Network (PNN), and Back Propagation Neural Network (BPNN), is compared to conventional regression analysis for order-of-magnitude cost estimation by El Sawah [23]. Real information collected from 35 low-rise structurally steel structures was utilized to create models.

An ANN model for predicting early-design construction costs of building projects is presented by El-Sawalhi et al. [24]. The model seeks to offer a more accurate and effective cost-estimating tool by using 11 important input characteristics and 169 case studies gathered from the Gaza Strip's building industry. The study shows how ANNs might help with improved financial planning and decision-making in the construction industry by enhancing initial cost estimates for building projects.

With the goal of estimating the initial costs of public sector office buildings in Sri Lanka, Dissanayake et al. [25] also created an ANN model, which they found to be beneficial. Creating an initial construction-expense-estimating model especially for Sri Lankan government building projects improved the precision of the initial stage's cost projections in this context by using an ANN model that was trained using data gathered from 50 finished tasks.

Likewise, even with sparse beginning data, ANNs were successfully utilized to forecast the ultimate cost of building projects. An ANN model for building projects in Iraq had a correlation value of 100% and an accuracy of 94.19% [26].

Additionally, Juszczuk [27] created and evaluated ANN ensembles and SVMs for basement framework cost forecasting, discovering comparable prediction accuracy.

An effective substitute for conventional techniques was provided by an ANN model for government buildings in Thailand that showed excellent precision ($R^2 = 0.914$) in cost predictions [8].

Also, for early cost assessment aimed at tall structures, machine learning approaches, such as ANNs, have been investigated. A multi-classifier approach that uses k-Nearest Neighbors (KNN) has demonstrated excellent accuracy ($R^2 = 0.81$) [3].

To anticipate important Earned Value Management (EVM) measures such as the Schedule Performance Index (SPI), Cost Performance Index (CPI), and Complete Cost Performance Indicator (CCPI), ANNs are efficiently employed. In residential developments in Iraq, ANN models revealed excellent correlation coefficients (e.g., CPI at 93.00%) and high average accuracies (e.g., CPI at 90.83%), offering proactive project control [28].

By suggesting an ANN method to forecast the most appropriate post-contract price control methods (PCCTs), Omotayo et al. [29] tackle the enduring problem of overcharges in building projects. The authors point out that cost overruns continue to occur despite a variety of strategies, and they attribute the situation to construction project managers' (CPMs') and quantity surveyors' (Qs's') incorrect use of PCCTs. To enhance cost management and lower the frequency of cost overruns in construction projects, the research surveys 135 CPM and QS specialists to create an ANN model which will effectively detect appropriate PCCTs.

Regarding long-term project management in building projects across the Indian sub-continent, Thakuria et al. [30] concentrate on the critical component of conceptual cost forecasting. The constraints of traditional approaches that primarily rely on the information that is accessible as well as the competence of estimators are emphasized by the authors. To address these issues, the study suggests and creates a thorough ANN model that is intended to precisely estimate and predict conceptual costs. By offering a more systematic and effective way to anticipate costs at an early stage, this strategy aims to enhance project planning and success in the area.

Significantly, an ANN model for building endeavors in India, utilizing a limited dataset, obtained 87% accuracy in initial cost estimate [31].

ANNs can generate accurate cost projections for abstract engineering services, which are distinct from material-based building expenses in the context of engineering services' cost estimation. Even with fragile databases, a model created using data from 132 projects demonstrated a 14.5% increase in accuracy (MAPE of 13.65%) when contrasted with earlier models [7].

While individual ANNs seem effective, the creation of hybrid and composite AI/ML models represents a distinct and growing phenomenon. To enhance accuracy, robustness, and the capability to manage intricate, non-linear interactions more skillfully, these methods incorporate many methodologies or integrate them alongside sophisticated feature selection methods [27,32].

During their 2021 study, Al-Tawal et al. [33] created and evaluated ANN simulation algorithms for estimating building construction expenditures at the conceptual, diagram, and comprehensive design phases. Utilizing a study of the literature and Delphi methodologies, the scientists chose 53 design parameters from actual data across 104 finished buildings in Jordan. These were then reduced to 41 and 27 to be applied in the basic and conceptual design phases, respectively.

To optimize time and cost estimations in prefabrication methods management for building, Challa and Rao [34] introduce a model based on ANNs. This kind of approach was selected because of their capacity to handle data of all kinds, allowing for precise forecasts even in the case of incomplete or ambiguous data. The information from 30 prefabrication building projects was utilized to teach the model, which included variables like projected

versus actual time and cost. The results showed that, in terms of reducing inaccuracy during estimate, ANNs in conjunction with Particle Swarm Optimization (PSO) performed better than Genetic Algorithms (GA) and Ant Colony Optimization (ACO). In the initial planning stages of projects, the suggested ANN model demonstrated its promise as a trustworthy decisions tool by achieving forecast accuracy levels above 90%.

A two-phase ANN model was created by Aguinaldo and Silva [35] to predict the cost and length of projects involving public school buildings in the Philippines. This model, which was built on Work Breakdown Structure (WBS) sub-models, showed that ANNs may produce significant results regardless of little data.

To solve the recurring problem of overinflated costs in building projects, Deepa et al. [36] provide a hybrid machine learning approach for the earlier cost estimates of pile foundations. The authors increase the reliability and model clarity by combining data mining methods with ANNs. To determine seven important characteristics, the technique starts with expert surveys (e.g., site location, water table, grade of concrete, foundation depth and diameter). The model uses a Bayesian Regularization-trained ANN using Levenberg–Marquardt optimization and is trained and validated utilizing data from 176 real-world examples among five projects for infrastructure in India, utilizing MATLAB R2021a and R. The resulting model outperforms traditional estimating techniques with a cost-forecasting precision of 97.42% (MAPE = 2.58%). In complicated geotechnical environments, the study shows that the conceptual estimation of costs may be improved by integrating machine learning with domain knowledge.

According to Trijeti et al. [37], ANNs may be utilized to estimate a building project's duration and cost. Acknowledging that construction is inherently unpredictable, making it extremely difficult to complete projects within predetermined budgets and timelines, the authors suggest ANNs as a useful tool. The goal of their research is to indicate how ANNs may enhance the forecasting of these crucial project parameters, which will increase construction management's effectiveness and caliber.

For the cost estimate of sustainable building projects, Al-Somaydai et al. [38] suggest a hybrid methodology that utilizes ANNs. This approach is intended to address issues in areas like Iraq that have little data and antiquated estimating techniques. The study uses ANNs to simulate the estimating process to improve forecasting accuracy as well as efficiency. To deliver state-of-the-art, accurate, and flexible technology for sustainable building cost prediction, it explores several factors in the creation of ANNs, such as network architecture and the influence of internal features on model efficacy.

By utilizing the connection between the intended building area and site location, Alabdali [39] created an ANN model to enable the interactive tracking of building costs and progress. The model demonstrated good forecasting accuracy utilizing real-world data and geometric parameters gathered from Revit, specifically when the building-to-site surface ratio reached 0.5. The model's capacity to replicate real budget situations and offer early warnings of cost deviations was validated utilizing a case study including concrete and wall analysis. The study confirms the efficacy of ANNs in early-stage cost estimation for building projects and emphasizes the importance of integrating geographic information to improve forecasting accuracy.

Specialized hybrid versions have emerged because of the growing focus on environmentally friendly buildings. The overall cost projection precision for typical green tall buildings was greatly enhanced by a model that combined regression analysis and ANNs. This model achieved a Mean Absolute Percentage Error (MAPE) of 15.09%, which was within suitable company norms and exceeded independent regression and ANN models alike [40].

The preparation, design, building, and handling phases of a building endeavor are all covered by the AI technologies mapped out by Egwim et al. [41]. Among the main conclusions are that AI is most frequently applied in the planning and management phases of buildings. Because of the fragmented systems of information, there is little integration throughout the value system. Training and compatible tool investments are essential for greater usage.

The utilization of ANNs for analyzing and forecasting overinflated costs in residential building projects is the main topic of Yadav and Pipaliya [42]. The authors intend to draw attention to ANNs' ability as a strong forecasting tool to assist with the common problem of expense overruns, which has a considerable influence on project participants and the construction sector. Their study lays the foundation for more precise pricing and risk reduction by synthesizing the body of research on the application of ANNs in this field.

In the effort to overcome the drawbacks of conventional estimates, such as linear constraints and expert biases, Padala and Goyal [43] created an ANN-based early-stages cost-prediction system for Indian building construction projects. In MATLAB, they trained a multi-layer feed-forward network utilizing data from 377 finished buildings and 17 important cost factors, such as the built-up area, foundation category, material amounts, and labor productivity. As observed by the low MSE and R-values ≈ 0.91 throughout the training, validation, and test groups, the model obtains great accuracy following the meticulous maintenance of information, choice of characteristics, and hyperparameter tweaking. This ANN framework provides owners and contractors with a strong, data-driven tool for accurate cost projections throughout conceptual design by incorporating intricate non-linear interactions.

The rising difficulty of cost management in contemporary high-rise and multifunctional buildings is addressed by Chen et al. [44], who examined the use of algorithm-optimized backpropagation neural networks (BPNN) and support vector machines (SVM) for office building cost estimation. The study creates six prediction models: BPNN, GA-BPNN, PSO-BPNN, GA-SVM, PSO-SVM, and GSA-SVM, utilizing grey relational analysis (GRA) for modeling and principal component analysis (PCA) for indicator reduction. The best model for predicting the cost of office buildings utilizes the PCA-GSA-SVM model, which shows consistent, quick, and accurate results.

Considering around 80% of the studies were conducted utilizing ANN-based models, the examined literature highlights how prevalent ANNs have become in construction cost estimation research. Backpropagation Neural Networks (BPNN), Radial Basis Function (RBF) networks, Multilayer Perceptrons (MLP), and hybrid configurations are among the most frequently employed architectures. In a variety of project types and geographical settings, these models continuously exceed conventional techniques in controlling uncertainty, regulating non-linear connections, and enhancing prediction accuracy. The results validate the methodological stability and flexibility of ANNs in assisting with early decision-making and maximizing the use of resources in buildings construction management.

ANNs in Building Renovation Projects

Renovation projects pose special difficulties for cost prediction because of financial limitations and the requirement for in-depth familiarity with existing buildings [45].

The use of Artificial Intelligence (AI) methods and applications in renovation projects in the Architecture, Engineering, and Construction (AEC) industry is examined by Boceneala et al. [46]. To find frequently used AI methods, data sources, and procedures, the research does a thematic analysis of the body of current literature. It addresses the potential benefits and inherent obstacles of using AI in renovation projects, providing

important insights into optimizing value and tackling the industry's sustainability and efficiency issues.

In their study, Khodabakhshian et al. [47] investigate the use of Machine Learning (ML) approaches to calculate the cost of renovating school buildings. The authors examine how machine learning may be utilized to forecast the amount of funds needed to upgrade the current school architecture. To promote sustainable growth and infrastructure development activities, the research will analyze a variety of building attributes and revenue information to show whether machine learning (ML) may deliver more precise and effective cost projections for renovation projects.

Prior studies by authors evaluate several ANN designs, including Radial Basis Function (RBF) and Multilayer Perceptron (MLP), for certain uses, such as estimating the cost of building renovation projects. When compared with MLP models for remodeling expenses, RBF models frequently show better accuracy (up to 6% sum of squares error and almost 0% relative error after training). To the best of the authors' knowledge, these were the first ANN models reported for predicting the ultimate cost of building renovation projects [4,5].

In summary, research continuously shows how AI and ML may revolutionize early approaches to specialized renovation initiatives by improving reliability, effectiveness, and decision-making in a variety of building-cost-estimating domains. The building sector is expected to have a more accurate, flexible, and comprehensive control of costs in the years ahead due to the growing richness of hybrid models as well as integrated online tools.

To the best of the authors' knowledge, there is a considerable gap in the literature regarding the limited application of predictive techniques in the building renovation industry, as shown in the previous paragraphs. The use of ANNs as a methodological framework for the precise assessment of overall renovation costs is examined in the present study with the goal of closing this gap. As a decision-making process with several factors, the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) is integrated to ensure the contextual relevance and accuracy of the input data. The model's empirical basis is reinforced by this integrative method, which also improves the model's robustness and prediction reliability by using information from projects that are highly comparable to the goal example, in terms of structure and context.

3. Problem Formulation

3.1. Artificial Neural Network Architectures in Building Construction Sector

A machine learning technique that is frequently stated as being intended to detect dynamic or non-linear performance in construction is the ANN [48]. Numerous undefined non-linear issues related to the building sector have been successfully solved by ANNs. Based on the vast amount of research, it appears that ANNs are among the most flexible scientific computational methods presently. Their architecture, including hidden layers and non-linear activation functionality, appears sufficient in flexibility to handle challenging cognitive issues [49].

ANNs learn from numerous combinations of inputs and their corresponding output structures, which gives them analogy-based decision-making skills. This process is somewhat motivated by organic brain systems. In the field of civil engineering, ANNs have been used since the late 1980s to solve a variety of issues, such as process optimization, building simulation, cost estimation, and challenges with pattern categorization and selection [22].

In building construction, ANNs are frequently utilized for estimation and forecasting tasks. The goals of forecasting and estimation algorithms for statistics are distinct. While a forecasting model aims to anticipate the result of a random factor, an estimating model aims to determine the properties of the real situation in nature. Consequently, in estimate modeling, the data utilized to develop the model is also utilized to test it, but in forecast

modeling a random quantity of data is substituted into the algorithm in circumstances where the data value had not been utilized to build the model [48].

An ANN's structure is built of multiple nodes, or neurons, organized in interrelated layers and combinations of layers. There are several different kinds of layers, including input, hidden, and output layers [42].

While there are a variety of network designs, the subsequent guidelines should be taken into consideration when choosing a suitable neural network. The quantity of nodes that are input depends on the total amount of variables that are independent in this kind of model. To estimate costs that are too complex to be resolved by conventional methods, ANNs may be employed to extract significance from complex data. ANNs are considered to have numerous advantages in buildings' total cost estimation. Compared to other conventional non-linear statistical techniques, ANNs may build forecasting models with greater ease since they may autonomously learn whether to perform tasks, for instance, according to the data collected for training. Although ANNs cannot directly describe the links between input variables and their outcomes, estimating and further precisely explaining correlations amongst variables is a primary goal when developing a decision approach. The quantity of neurons should be enough for the network to reach a consensus, but not so many that they become memorized [50].

An increase in the training sample could produce an improved outcome for the ANN. A network that can effectively generalize new data and display accurate forecast values is produced with the aid of extra data. The machine learning system could learn to memorize the training examples, yet it lacks the ability to generalize to novel scenarios, thus as novel information is introduced, the mistake margin grows [31].

Whereas technologies based on conventional computer logic involve comprehensive software to accomplish a specific task, ANNs are intelligent systems that are based on simplified computational models of the biological structure of the nervous system of humans. Whenever full models required to perform traditional computing approaches either happen to be excessively large or overly complicated to adequately symbolize, or merely absent, ANNs' capacity for autonomous learning remains very helpful. ANNs offer a great degree of generalization capability and noise protection due to their widely dispersed, linked architecture.

In situations in which the process to be modeled is so complicated that its details are not possible to explicitly state in a mathematical manner or when an explicit formulation results in an impairment of sensitivity owing to simplifying issues, ANN systems provide several benefits above standard modeling techniques [28], specifically Radial Basis Function Networks that are described in the next section.

Radial Basis Function (RBF) Networks

The unique and increasingly popular relatives of feed-forward neural network topologies are Radial Basis Function (RBF) networks, which are particularly beneficial for problems requiring function estimation, identifying patterns, and time-series analysis prediction. The distinctive feature of RBF networks is the way they utilize the spatial proximity of input directions to a group of prototype centers to decide when to activate hidden units.

Three layers constitute an RBF network's architecture: an input layer, one hidden layer, and an output layer. A center vector in the input field space is linked to every single radial basis neuron that populates the hidden layer. As localized fields of reception, these neurons react significantly to impulses that come into proximity to them. Higher activation values are obtained from nearer inputs. The activation for every hidden unit is calculated as a function of the distance between the input vector and the unit's center. Localized features

and non-linear interactions among the data might be captured by the network thanks to this process.

The network may provide ongoing or discrete outputs according to the application due to the output layer's linear interaction with the hidden layer activations. The ideal values for the resultant weights, radial function widths (or spreads), and centers are usually established during the training phase. RBF networks' fast convergence characteristics and very simple structure make them ideal for issues wherein the input space displays region-specific or clustering behavior.

Furthermore, RBF networks' interpretability, which results from their transparent decision limits and localized activation, makes them an appealing option for fields that need comprehensible AI models. By using hybrid approaches, like integrating RBFs with different learning frameworks or optimization methods, their performance might be further optimized.

The RBF, which connects the units in one layer to the values of units in the layer that follows, is the activation function for the hidden layer. The output units represent the weighted summation of the hidden units since this identity function operates as the activation mechanism for the output layer. Radial Basis Function normalization utilizes the SoftMax activation function to normalize all concealed unit activations so that they add up to one [51].

A multiplier applied to the Radial Basis Functions' width is the overlapping factor. When d is the quantity of input units (the total of the number of categories across all factors and the number of variables), the automatically determined value of the overlapping factor is $1 + 0.1d$, [51].

In RBF ANNs, the number and location of hidden neurons must be carefully chosen to reduce the chance of overfitting. Due to their reliance on Radial Basis Functions centered at data stages, RBF systems are more sensitive to the density and dispersion of input samples than, for instance, Multilayer Perceptron (MLP) models. Generally, overfitting can be prevented by limiting the number of basic functions, identifying ideal centers using clustering approaches like k-means, and evaluating the model performance on different datasets. Cross-validation in conjunction with regularization during the output layer training phase improves generalization and keeps the model from fitting noise in the input information [52].

3.2. TOPSIS Method

For the purpose of identifying the optimal option according to the principles of the consensus solution, Hwang and Yoon in 1981 devised TOPSIS, the Technique for Order of Preference by Similarity to Ideal Solution. Selecting the answer with the minimum Euclidean distance from the perfect solution and the furthest Euclidean distance from the ideal solution that is negative may be thought of as the balanced approach [52].

With this approach, choices are mostly based on which alternative is nearest to the ideal spot and the farthest from an anti-ideal option. Because this procedure is straightforward and customizable, TOPSIS is a frequently utilized and user-friendly technique [53].

The goal of the TOPSIS approach is to obtain a sequence of preference that is comparable to the optimal solution, meaning a hypothetical answer has the highest amount of advantages and the lowest number of costs of qualities or options. The perfect negative answer, on the other hand, is a hypothetical solution which would minimize the advantages of the characteristics or requirements while raising the associated expenses [53].

As a result, the option that is closest to the ideal positive answer and furthest from the opposing viewpoint is the best one, which also serves as the answer to the issue. The optimal answer and Euclidean distance define the analogy or variation, while the dataset's

highest and lowest scores will be utilized to analyze the opposite of what is desirable. Alternatives among criteria are conceivable using this approach's more accurate modeling structure since it permits ignoring an inadequate result while examining one criterion in favor of excellent outcomes in other [53].

The fundamental tenet of TOPSIS is that the alternative which is selected should be least related to the negative ideal solution (the non-optimal solution) and most closely related to the positive ideal solution (the optimal answer). Because of its straightforward mathematical model and quick computing method, TOPSIS is utilized extensively by researchers [54].

In the following section the ANN RBF models will develop, initially utilizing data from 52 plus 1 new building renovation projects to estimate the total cost, and subsequently, after applying the TOPSIS method, utilizing the 10 plus 1 new projects selected by the above method to estimate the total cost of the new project.

4. Methodology

4.1. Development of Artificial Neural Network Models for Cost Estimation in Building Renovation Projects

As mentioned above, as a critical method inspired by the biological brain's capacity to generate solutions and gain knowledge from training, ANNs have emerged as a prominent and extensively used AI method in building expenditure assessment [7,28,50].

Because of their ability to manage noisy or insufficient data and simulate intricate, non-linear connections, they are particularly suitable for the multifaceted character of building endeavors [6,24,26].

Utilizing both historical and current data, ANNs may serve as a useful tool for building renovation cost overrun research. They can improve the overall management of these projects in building and other industries, offer more precise cost estimates, evaluate risks, and allocate resources optimally [42].

Numerous ANN models were created for the purpose of this study. The research examined 53 renovation enhancements claimed by two different buildings renovation construction corporations. Both companies have been engaged in the industry for a lengthy time. Data processing was simple given that they approached every project in a similar procedure.

Every project was an individual endeavor. An analysis of cost was determined after every building was thoroughly surveyed and measured. Both the proprietor and the construction team agreed with this first estimate. The additional expenses allowed by Greek legislation for private ventures were also included in the overall cost of every project.

For every building's renovation project, fourteen different parameters were methodically found and documented. The total initial cost, estimated completion time, initial demolition drainage cost, initial cost of plumbing work, initial heating cost, initial cost of electrical work, initial cost of masonry coatings, initial cost of plasterboard construction, initial bathroom cost, initial flooring costs, initial frame cost, initial door cost, initial paint cost, initial kitchen construction cost, and one output data, the total final cost, were among them. A systematic database was carefully created for each of these variables to make statistical analysis straightforward subsequently.

IBM SPSS Statistics 30 software was used to create the database, which included the previously indicated technological standards [51], guaranteeing consistency, dependability, and analytical rigor. The collection consisted primarily of residential buildings located in cities, specifically in the Greek municipalities of Katerini and Thessaloniki. Mostly apartment complexes, these buildings face special architectural and infrastructure issues because many of them are several decades old, especially those in the historic city centers.

These remodeling initiatives were primarily driven by the requirement to enhance energy efficiency in compliance with modern sustainability regulations. Yet, there were also other goals, including updating the amenities, redesigning the buildings' appearances, and, in particular instances, changing the buildings' intended function. During the six-year period from 2018 to 2024, these projects were scheduled to be implemented.

4.2. Data Collection and Validation

ANN models' efficacy and predicted precision are strongly impacted by the caliber and volume of data used for training. Massive datasets are normally advantageous, but recent research shows that even less significant datasets of exceptional quality may produce accurate findings, highlighting the significance of data uniformity and value over mere bulk. The quality, the amount, and proper preparation associated with input data are critical factors that impact the efficiency and dependability of AI/ML models in the building industry [7,31].

Choosing from the most important input elements is essential for maximizing model accuracy, lowering processing overhead, and enhancing generalizability [7,24,32].

Non-obvious links that empirical intuition can overlook might be revealed by data-driven feature selection techniques [8]. The ground-level area, whole floor area, and price inflation, for example, have all been found to constitute important factors in estimating building expenses [33]. The number of floors and typical surface area have a significant impact on preliminary construction costs [24].

As construction companies consider this knowledge exclusive and therefore are reluctant to share it with other competing firms, gathering data regarding real projects built in Greece during the last six years was a challenging task. Furthermore, since businesses compete in the economy, this type of knowledge helps them stand out. To create an appropriate ANN model, however, sincere attempts have been made to gather sufficient data, in addition to the quantity of projects and the quality of the data provided. Personal interactions with construction companies, organizations, and government agencies around Greece served as the foundation for the data collection process.

To address the shortcomings in the data gathered, fundamental hypotheses and requirements were established, such as the requirement that projects be fully completed and constructed, that they take place between 2018 and 2024, that all unfinished, missing, or false information be removed, that replicate data—similar initiatives with the same values—be eliminated, as well as that every set of data maintains consistency, with minimal variation in project cost, as Al-Tawal et al. also stated in their research [33].

An ANN model allows for the removal of nonpredictive or wasteful variables by removing them from the model, in contrast to regression methods that need the covariance between the variables to be verified. Conversely, the variable selections are supported by less correlated values. This significantly accelerates the training process and justifies the inclusion in the model because only the most important variables are included.

Until the model is trained, it is challenging to establish what quantity of data is required for a model to be precise. Determining whether additional data is needed can be aided by the network's efficiency evaluation [7].

In the context of the current research, the correlation analysis was completed initially. Results are shown in Table 1.

Table 1. Correlation Analysis.

Total Final Cost	Initial Total Cost	Estimated Completion Time in Days	Initial Demolition–Drainage Costs	Initial Cost of Plumbing Work	Initial Heating Costs	Initial Cost of Electrical Work	Initial Cost of Masonry Coatings	Initial Cost of Plasterboard Construction	Initial Bathroom Cost	Initial Flooring Cost	Initial Frame Cost	Initial Door Cost	Initial Paint Cost	Initial Kitchen Construction Cost
Pearson Correlation (ρ)	0.989 **	0.927 **	0.167	0.346 *	0.792 **	0.985 **	0.927 **	0.167	0.346 *	0.792 **	0.908 **	0.904 **	0.922 **	0.284 *
Sig. (2-tailed) (p)	<0.001	<0.001	0.233	0.011	<0.001	<0.001	<0.001	0.233	0.011	<0.001	<0.001	<0.001	<0.001	0.039
N	53	53	53	53	53	53	53	53	53	53	53	53	53	53

Notes: ** Correlation is significant at the 0.01 level (one-tailed), * Correlation is significant at the 0.05 level (one-tailed).

The dependent variable, final cost, has a significant relationship with the independent variables of the project, including the total initial cost, estimated completion time, initial demolition drainage cost, initial cost of plumbing work, initial heating cost, initial cost of electrical work, initial cost of masonry coatings, initial cost of plasterboard construction, initial bathroom cost, initial flooring costs, initial frame cost, initial door cost, initial paint cost, and initial kitchen construction cost, according to Table 1.

Fifty-three building projects' final total costs are examined in this analysis, along with the parametric dependence of various initial project cost components. The study measures the direction and intensity of these interactions using Pearson correlation coefficients, with a one-tailed threshold of statistical significance established. The results shed important light on which starting cost criteria are best at predicting the ultimate cost of a project.

A broad variety of correlation strengths is revealed by the results, which are compiled in Table 1. The ultimate total cost shows a very significant positive link with several important variables. The initial total cost ($\rho = 0.989$, $p < 0.001$), the initial cost of electrical work ($\rho = 0.985$, $p < 0.001$), the initial cost of masonry coatings ($\rho = 0.927$, $p < 0.001$), the initial cost of paint ($\rho = 0.922$, $p < 0.001$), the initial cost of the frame ($\rho = 0.908$, $p < 0.001$), and the initial cost of the door ($\rho = 0.904$, $p < 0.001$) all exhibit exceptionally high Pearson correlation coefficients, all of which are significant at the $p < 0.01$ level. According to this, these variables are accurate indicators of the project's ultimate cost. It is assumed that there will be a significant link between the original and final total costs because the first estimate is the main baseline for the project budget.

Substantial and noteworthy correlations also occur with additional factors, although to a significantly lesser extent. These consist of the initial flooring cost ($\rho = 0.792$, $p < 0.001$) and the projected completion time in days ($\rho = 0.927$, $p < 0.001$). Given that unforeseen postponements or scope changes frequently result in longer project durations and greater labor, overhead, and material costs, the positive correlation between time and cost follows logically.

Furthermore, there is a lower correlation between some cost components and the project's ultimate cost. The initial plasterboard construction cost ($\rho = 0.167$, $p = 0.233$) exhibits a statistically negligible and extremely weak association. Likewise, there is a lower but still statistically significant positive association between the initial bathroom cost ($\rho = 0.346$, $p = 0.011$) and the initial kitchen construction cost ($\rho = 0.284$, $p = 0.039$). These results imply that variations in those project elements could not have as significant an

influence on the total cost as other variables, possibly because they are more standardized or flexible parts of a building project.

The most crucial factors for forecasting the ultimate total cost of a building project seem to be the variables with the strongest correlation, such as the initial total cost and the electrical, masonry, and paint costs. When creating RBF ANN prediction models, these indicators should be given special consideration. On the other hand, factors with weak or negligible correlations, like the cost of plasterboard construction, are not likely to be extremely helpful in these models and might even be eliminated to simplify the model architecture without significantly lowering accuracy.

An ANN model could estimate a project's total beginning cost utilizing all the previously mentioned parameters as input data. These variables have varied degrees of association with the final project cost, according to the correlation study that was supplied. Even though certain correlations are powerless, it remains advantageous for the model's performance to include all of them.

A measure of the linear relationship between each independent variable and the ultimate total cost was obtained from the correlation analysis. With Pearson correlation coefficients near 1, variables like the beginning total cost, initial electrical work, and initial masonry coverings demonstrated an immensely robust, statistically significant positive relationship with the final cost. This implies that the ANN ought to provide these features a high weight because they are reliable forecasters.

Conversely, there existed a minimal and statistically insignificant association between variables such as the initial cost of plasterboard construction. Nevertheless, employing them as inputs is still an excellent choice for a variety of reasons, such as the following:

Usually, a linear relationship is measured by the Pearson correlation coefficient. However, a simple correlation analysis is unable to identify complicated non-linear correlations amongst variables; an ANN might. A variable might still be able to provide pertinent data to the model's overall prediction in a manner that is not linear even though it exhibits a low linear correlation.

All the inputs together contribute to the model's effectiveness. On an individual basis, a single weakly correlated variable might not have significant predictive ability; however, paired with additional variables, it may assist in creating a more complex and precise forecast. With the ability to learn the relative relevance of each input, an ANN may effectively give strongly linked variables greater importance while still utilizing any small information that merely correlated variables may supply.

There may be an elimination of potentially helpful information when variables are excluded based only on a weak linear association, which might reduce the accuracy of the model overall. The model can comprehend the situation by retaining all the accessible data points.

4.3. Implementation of TOPSIS Method

With the goal of determining the most comparable endeavors from a collection of historical data, the given program is a condensed implementation of a similarity-based matching method. The approach uses Euclidean distance as the metric for evaluating similarity and is founded on the ideas of the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS). In project management and cost-estimating, where past data can be utilized to guide decisions regarding new endeavors, this method is very pertinent [55–57].

Following the standard six-step process, the TOPSIS approach was applied: (1) decision matrix formulation; (2) normalization; (3) weighted normalization; (4) Positive Ideal Solution (PIS) and Negative Ideal Solution (NIS) identification; (5) calculation of separation measures; and (6) determination of relative closeness. This methodological order is

consistent with the standard implementation developed by Lai et al. [56] and Tzeng and Huang [58,59].

The TOPSIS approach was utilized as the classic definition states, using vector normalization to guarantee that the criteria are comparable and unitless. The $N = 52$ renovation projects and $M = 14$ criteria were subjected to the following procedure:

First, define the decision matrix where is the performance value of the i -th alternative (project) with respect to the j -th criterion (e.g., Initial Cost, Completion Time, Cost Plumbing, etc.).

Second, normalize the decision matrix which can be obtained by applying vector normalization:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{k=1}^N x_{kj}^2}} \quad (1)$$

where r_{ij} is a normalized value, more specifically the value of x_{ij} after vector normalization. This value is now unitless and scaled identically across all criteria and x_{ij} is the decision matrix element, the initial performance value of the i -th alternative (project) with respect to the j -th criterion.

In the next third step, the normalized values should be multiplied by weight vector w , which is predetermined based on the significance hierarchy and characteristics of the project:

$$v_{ij} = w_j r_{ij} \quad (2)$$

where v_{ij} is the weighted normalized value, the final value of the i -th project on the j -th criterion after applying the weight, and w_j the weight vector specified to the j -th criterion, representing its relative significance.

Fourth, the Negative Ideal Solution (NIS, A^-) and Positive Ideal Solution (PIS, A^+) should be determined. From the v matrix, which must be getting the best (v_j^+) and poorest (v_j^-) values for each criterion.

Afterwards, in the fifth step the S_i^+ and S_i^- separation measures should be established (where S_i^+ is the Euclidean Distance of the i -th project from the PIS and S_i^- is the Euclidean Distance of the i -th project from the NIS).

And last, every option's Euclidean distance utilizing the PIS and NIS, respectively, should also be determined.

The CC_i , as the relative closeness which is reported as the final ranking score (closeness coefficient) for the i -th project, establishes the project's proximity to the ideal solution in comparison to the anti-ideal solution:

$$CC_i = \frac{S_i^-}{S_i^+ + S_i^-} \quad (3)$$

The choices are then listed in decreasing order of CC_i .

The decision matrix, built from a historical database of 52 finished renovation projects ($N = 52$), serves as the framework for the present TOPSIS analysis. The total initial cost, estimated completion time, initial demolition drainage cost, initial cost of plumbing work, initial heating cost, initial cost of electrical work, initial cost of masonry coatings, initial cost of plasterboard construction, initial bathroom cost, initial flooring cost, initial frame cost, initial door cost, initial paint cost, and initial kitchen construction cost of each project is represented by 14 different quantitative evaluation criteria ($M = 14$). The complete, unedited decision matrix is archived and made accessible to the corresponding author upon request with the goal of ensuring complete transparency and reproducibility.

To illustrate the significant variation in project scope and expense that guides our distance-based similarity computation, Table 2 displays the descriptive statistics (mean, standard deviation, minimum, and maximum) for each criterion.

Table 2. Descriptive statistics.

Criterion	Mean (μ)	Standard Deviation (σ)	Minimum	Maximum
Total Initial Cost (€)	41,995.47	55,475.07	11,500.00	359,135.00
Estimated Completion Time (Days)	66.42	30.37	26.00	208.00
Initial Demolition–Drainage Cost (€)	1483.65	791.42	0.00	4000.00
Initial Cost of Plumbing Work (€)	2277.00	2443.50	0.00	13,304.00
Initial Heating Cost (€)	3119.62	3765.73	0.00	25,850.00
Initial Cost of Electrical Work (€)	3835.38	6423.21	400.00	45,000.00
Initial Cost of Masonry Coatings (€)	2235.10	4173.63	0.00	29,000.00
Initial Cost of Plasterboard Construction (€)	838.88	2190.83	0.00	15,241.82
Initial Bathroom Cost (€)	3414.85	3916.64	0.00	27,190.00
Initial Flooring Cost (€)	3523.11	3257.64	200.00	16,253.00
Initial Frames Cost (€)	3599.08	6882.04	0.00	48,000.00
Initial Doors Cost (€)	1939.53	2393.51	0.00	14,450.00
Initial Painting Cost (€)	3440.75	5639.34	0.00	38,000.00
Initial Kitchen Construction Cost (€)	2488.45	2893.64	0.00	13,600.00

The exact descriptive statistics associated with the 14 assessments were calculated utilizing a dataset of 52 renovation projects. Prior to beginning the TOPSIS normalization procedure, it is essential to comprehend the size and variability of the data.

Significant background information for TOPSIS' normalization stage is provided by the fact that several cost categories (such as heating, plumbing, and demolition) presented zero ratings for the minimum, indicating that a few projects did not require work for those specific categories.

The program utilizes two datasets, one with the attributes of a brand-new construction project and another with the attributes of fifty-two complete projects. Each finished project is given a matching score in relation to the new project, which forms the basis of the technique.

For data ingestion, two Excel files, one carrying the attributes of the 52 finished projects and the other featuring the attributes of the new project, are first loaded into pandas Data Frames by the program.

The next step remains the computation of Euclidean Distance. The Euclidean distance between the attribute vector of each finished project and the attribute vector of the new project is computed. Each finished project's attribute vector and the new project's attribute vector are measured for Euclidean distance d . With the attributes $\{a_1, a_2, \dots, a_n\}$ and $\{b_1, b_2, \dots, b_n\}$ for each project, respectively, the Euclidean distance (d) between P_1 and P_2 is determined by the following Equation (4):

$$d = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad (4)$$

A lower Euclidean distance suggests that the two projects are more comparable, Where the similarity score is presented in Equation (5):

$$\text{Similarity score} = \frac{1}{1 + d} \quad (5)$$

This alteration implies that an improved result can be achieved with less distance (greater resemblance). In accordance with this evaluation, the completed endeavors are subsequently organized in descending sequence to demonstrate an apparent order of resemblance to the new project.

Using Google Colaboratory [60], a cloud-based development environment that supports Jupyter notebooks and GPU (Graphics Processing Unit) acceleration, the methodological framework for project matching was written in Python 3.11. Euclidean distance measures were utilized in the present study to determine the similarity scores between a new project and a collection of finished projects. The Excel-formatted input data was submitted interactively through Colab's file interface. Data processing was performed using the pandas library, while numerical calculations were supported by the math and NumPy modules. To find the most pertinent finished projects, matching scores were calculated and rated. Reproducibility, flexibility, and accessibility to computational resources were provided by the usage of Colab.

4.4. Application of RBF-ANN in Cost Estimation

4.4.1. Implementation of an RBF-Based ANN Model Utilizing 53 Input Variables Before Applying TOPSIS Method

Starting with 53 input data with a simple setup of just one neuron in the hidden layer, the Radial Basis Function (RBF) neural network model was gradually developed. The addition of one hidden neuron every analytical iteration was carried out in subsequent investigations, enabling a systematic comparison of model complexity and performance.

In the hidden layer, the normalized radial basis function is the activation function. The input data is assigned to the hidden layer via this method, wherein every data vector's proximity to the core of a hidden unit affects the moment of activation of the unit. All hidden units undergo this normalization by applying a SoftMax activation function, which ensures that the total amount of their activations approaches one.

The identity function is utilized by the output layer as its mechanism for activation. The results are therefore a direct weighted summation of the hidden unit activation events enabling the simulation to linearly convert the non-linear feature space which the hidden layer has produced.

The dispersion of every hidden neuron, which depends on the amount of distance away it is from the center, determines its effect. The overlapping factor represents the factor that is applied to this dispersion. In the present research, the model's training algorithm effectively identified such overlapping factors. The scale of the model is indicated by the number of input data points, d , which is expressed as $1 + 0.1d$.

The 32-neuron hidden layer ANN became the final result of the search. Forty models were essentially produced by the study. The study revealed that the ANN with 32 neurons in the hidden layer generated the MOST accurate results when calculating the total of the relative errors and sum of squares errors with near 0.5 relative error and up to 0.3 monetary units sum of squares error in the testing sample.

With a limited quantity of construction data readily accessible and sans the requirement for an additional intricate architecture, the trained model's data demonstrated that the ANN-based approach was productive in estimating the structures' cost. To optimize the model's learning capacity from the meager input, researchers meticulously chose its properties. Selecting the most relevant independent variables was necessary to achieve this goal. They also used high-quality, as exact as feasible data for verification to remove flaws or variations. The first objective was to make sure that data source was accurate and consistent. High-quality data from two building renovation construction companies with decades of experience can greatly enhance model performance, particularly when the

amount of data seems low to avoid overfitting. In addition, regularization techniques such as early halting were used.

4.4.2. Implementation of a Radial Basis Function Artificial Neural Network Model After TOPSIS Method

Through identifying a homogeneous group of renovation projects that closely resemble the new project, TOPSIS was applied as a strategic data filtering tool. According to the fundamentals of primarily deductive thinking, this procedure prioritizes the most pertinent previous instances to provide a new forecast.

A considerable increase in computational efficiency is the main advantage of this strategy. The RBF-ANN model's training phase computational load is significantly reduced by cutting the training dataset from 53 projects to 11 (the new project plus the 10 most comparable projects). Since each data point helps determine the hidden neuron centers and the ensuing distance metrics, the RBF-ANN's training duration and the quantity of data points are directly correlated. Therefore, a smaller number of data points leads to a quicker convergence and a speedier cost estimate creation.

From the original collection of 53 projects (52 plus the one new building renovation project), the TOPSIS approach is a highly efficient and reliable method for pre-selecting the 10 most comparable initiatives with the one new project.

The nine-neuron hidden layer ANN formed the search's ultimate outcome. The research project effectively yielded 40 models. By analyzing the sum of squares and the relative errors, (nearly 0.6 relative error and up to 0.8 monetary units sum of squares error in the testing sample), the study revealed that the ANN with nine neurons in the hidden layer produced the most accurate findings.

Compared with utilizing the entire dataset, this method has several advantages when training an ANN model with an RBF kernel:

It strengthens the RBF-ANN framework. The significant decrease in the size of the dataset is the main factor causing the RBF-ANN training process to accelerate. 11 projects—10 comparable ones plus 1 new project—rather than 53 were utilized to train the model. This decrease has a direct effect on the amount of computation.

The RBF-ANN is built on hidden layer centers, each of which represents a training data point. Each additional data point's Euclidean distance from these centers is determined throughout the training procedure. The necessary computational operations are reduced whenever there are fewer points of data since there are fewer centers. Improved model completion and a notable reduction in training time result from this approach.

Avoiding overfitting and improving accuracy by selecting the ten most similar projects, the model's accuracy is increased in addition to performance. The TOPSIS analysis did not include the other 43 projects because they might contain noise or unrelated data that could cause the model to overfit. This occurs whenever the model discovers the unique features of the unconnected projects and is incapable of appropriately generalizing to a novel, untested project.

Alternatively, the RBF-ANN can discover the most accurate and representative associations among project parameters and the final cost by training the model on a homogeneous dataset (the 10 similar projects). By focusing on data that is directly comparable to the project being estimated, the model produces estimates that prove more accurate and dependable.

The following Figure 1 displays the flowchart for the suggested analytical strategy:

The model's current focus on the 14 detailed cost and time attributes is driven by the necessity for quantifiable, empirical data to ensure methodological rigor within the TOPSIS framework. Cost minimization is not the inherent objective; rather, the model functions as a feasibility and budgeting tool to match a new project's scope to a historically

executed, proven cost structure. Quality and safety are regarded as essential legislative limitations; before analysis, it must be assumed that all 52 historical projects and each proposed project comply with the minimal regional building and safety requirements. The itemized cost criteria could be utilized as a high-fidelity stand-in for technical complexity and required expenditure by neutralizing technical variation (such as seismic or climatic demand) under this *ceteris paribus* condition. The model's usage is restricted to projects within the same precise regulatory and climatic jurisdiction, ensuring that it remains data-driven and repeatable.

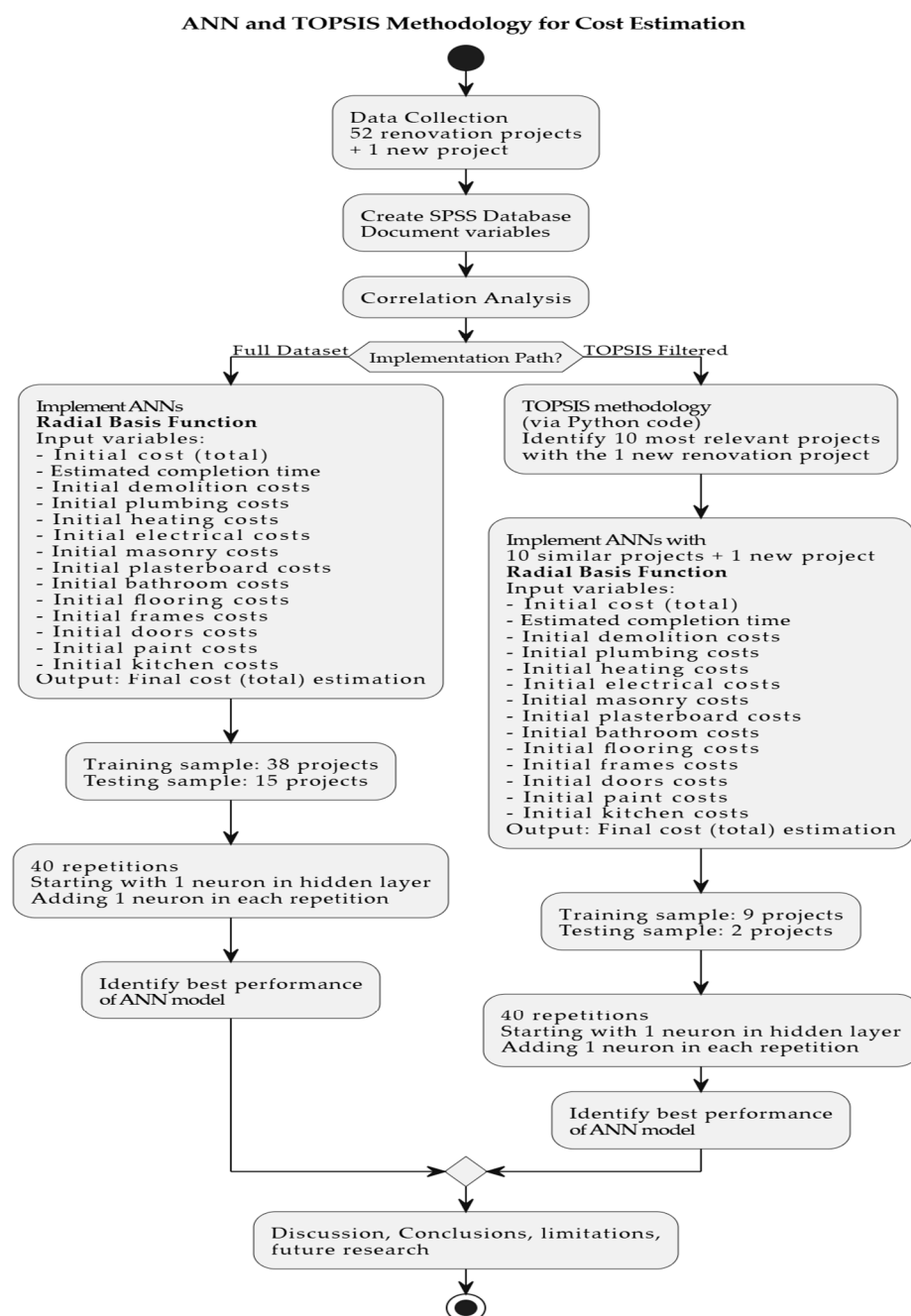


Figure 1. The suggested analytical process.

5. Proposed Solution and Resulting Findings

Developing two ANN models, one with 53 input data and another after the implementation of TOPSIS methodology with 11 input data, to forecast building expenses at

the planning and design phases beginning of a project's life cycle was the goal of this research. The comparison of the two robust ANN models and their results is remarkable. Whenever project data is scarce or not accessible, the models are meant to assist shareholders, contractors, and other stakeholders in building projects in estimating the cost of the endeavor.

Every single one of the models that has been provided has excellent forecasting performance when measured by general performance measures. The models meet expectations for cost-prediction accuracy in the early stages of the building project since the majority of cost-forecast percentage mistakes lie within the range of -20% ; $+20\%$ [61,62].

The current study utilized relative errors and sum of squares errors to evaluate the precision of the ANN model. A measurement's precision in relation to its size is indicated by its relative error. This Equation (6) is applied to determine the relative error:

$$\text{Relative error} = \left(\frac{\text{Absolute Error}}{\text{Actual Value}} \right) \quad (6)$$

where Actual Value represents the true value (actual final cost) and Absolute Error is the variance among the measured or projected value and the actual value (actual and estimated cost) [51].

The sum of squares error, or SSE, is a metric used to measure the variation among the values estimated by a model and the actual data values. The sum of squares demonstrating the differences among the expected and actual outcomes is utilized in the analysis. It can be determined by employing Equation (7) that follows:

$$\text{SSE} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (7)$$

where y_i refers to actual observed values (actual final cost), $\hat{y}_i$ to expected outcomes (estimated final cost) derived from the model, and n is the amount of observations [53].

Figures 2 and 3 indicate how the training sample's sum of squares error and relative error are displayed when RBF is implemented with 53 input points of data (prior to the implementation of the TOPSIS method).

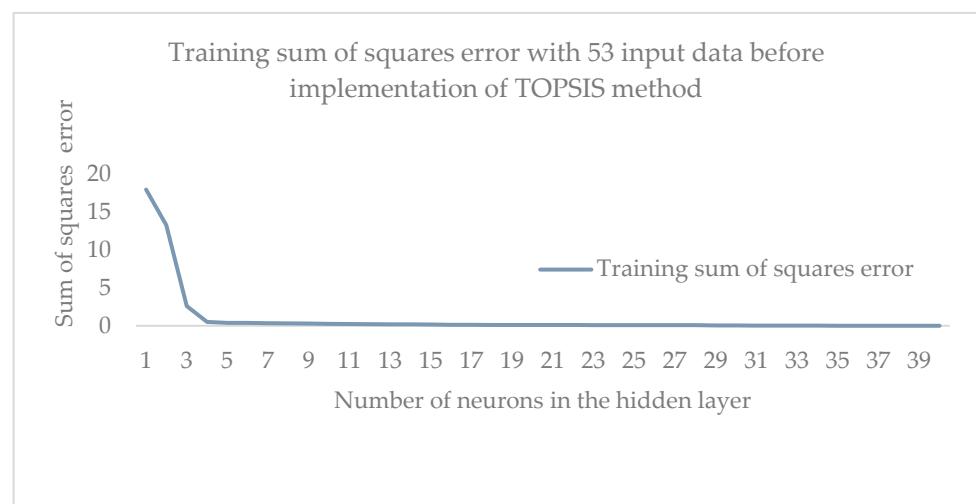


Figure 2. Training sample sum of squares error with 53 input data before implementation of TOPSIS method.

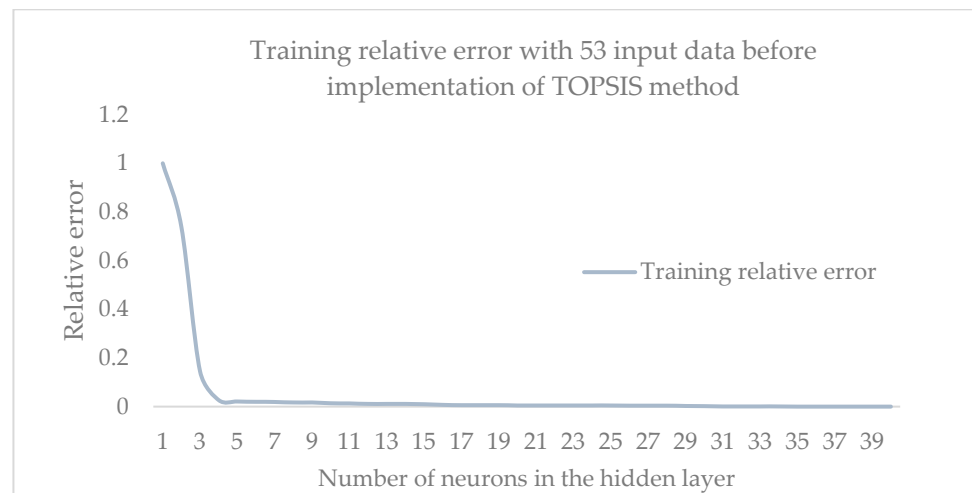


Figure 3. Training sample relative error with 53 input data before implementation of TOPSIS method.

With concern about overfitting after 40 neurons in the hidden layer after RBF implementation, the testing sample's sum of squares error and relative error were kept low, as shown in Figures 4 and 5.

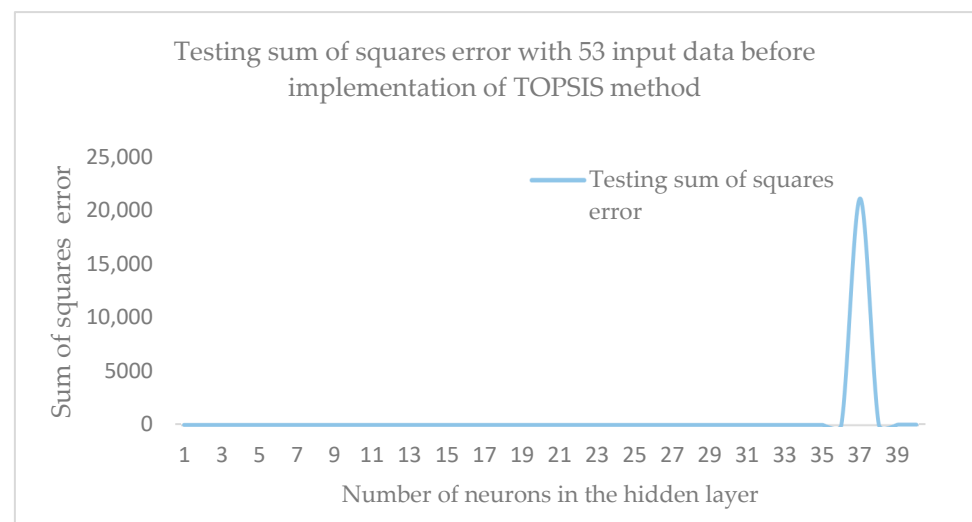


Figure 4. Testing sample sum of squares error with 53 input data before implementation of TOPSIS method.

The following charts relate to the application of the ANN RBF model after the implementation of the TOPSIS methodology (via python program):

In this context, the improved performance of the RBF approach in collaboration with the TOPSIS method demonstrates its robustness and reliability in cost-estimating, especially in the face of significant project cost variability.

As demonstrated in Figures 2–5, when the entire dataset of 53 renovation projects was employed (52 projects and the new one), the ANN-RBF model performed adequately. The above figures indicate that the testing sample's relative error and sum of squares error (SSE) were kept low. This implies that despite having sophisticated architecture with up to 32 neurons in the hidden layer, the model was successful at extrapolating from the training data and did not experience overfitting.

A considerably smaller dataset of 11 data points (10 relevant projects and the new one) was obtained by applying the TOPSIS methodology to reduce the dataset down to the

10 most pertinent projects for the new project. This condensed dataset was then subjected to another application of the ANN-RBF model. According to the results given, the model's reapplication to the improved dataset also produced outcomes that are regarded as accurate. A similar degree of performance to the original model, but with a more targeted dataset, is implied by Figures 6–9, which display the SSE and relative error for both the training (9 projects) and testing (2 projects) samples.

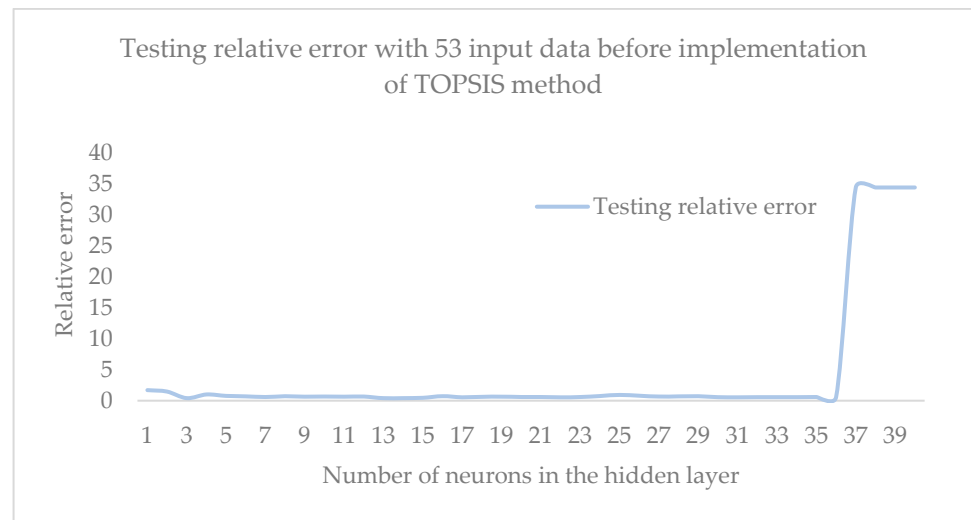


Figure 5. Testing sample relative error with 53 input data before implementation of TOPSIS method.

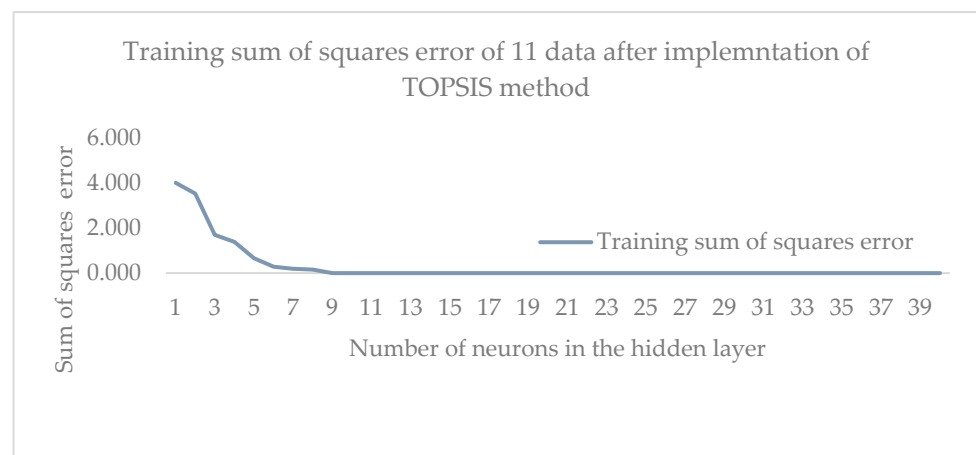


Figure 6. Training sample of sum of squares error with 11 input data after implementation of TOPSIS method.

According to the primary finding, the ANN-RBF model's performance was not negatively impacted by pre-processing the data using the TOPSIS methodology and choosing a more manageable and pertinent subset of projects. It appears, as could be seen in Figures 10 and 11, to have utilized a significantly smaller training and testing set and still managed to maintain a comparable degree of accuracy.

Subsequently, this implies that the effectiveness and applicability of prediction models can be raised by combining neural networks with a concentrated, multi-criteria decision-making technique such as TOPSIS. This makes it feasible to train the model on a more representative and focused dataset, which could result in more precise predictions for a particular new project. A better, more useful strategy for cost estimation in an actual building setting is shown by this hybrid methodology.

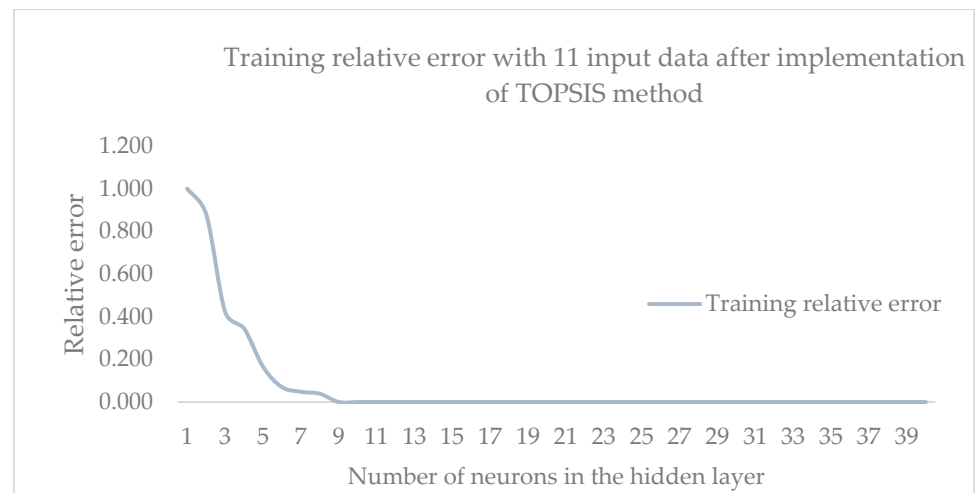


Figure 7. Training sample of relative error with 11 input data after implementation of TOPSIS method.

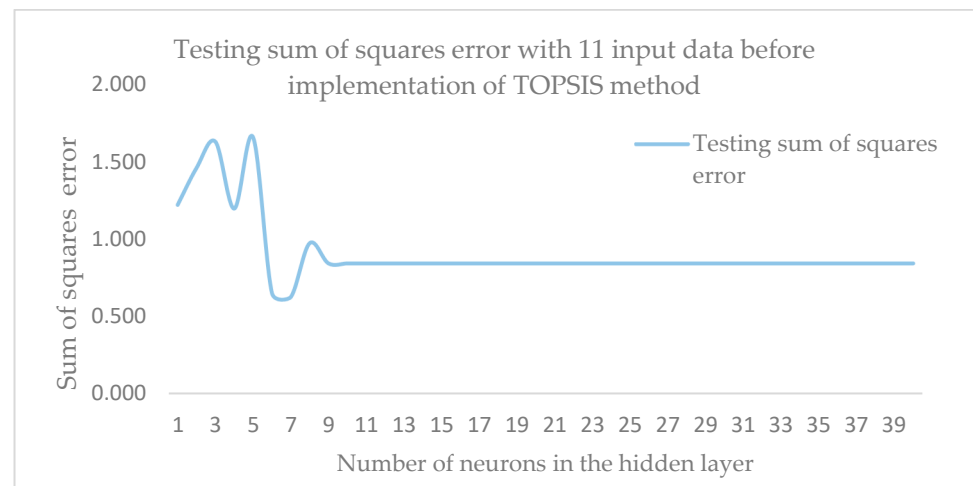


Figure 8. Testing sample of sum of squares error with 11 input data after implementation of TOPSIS method.

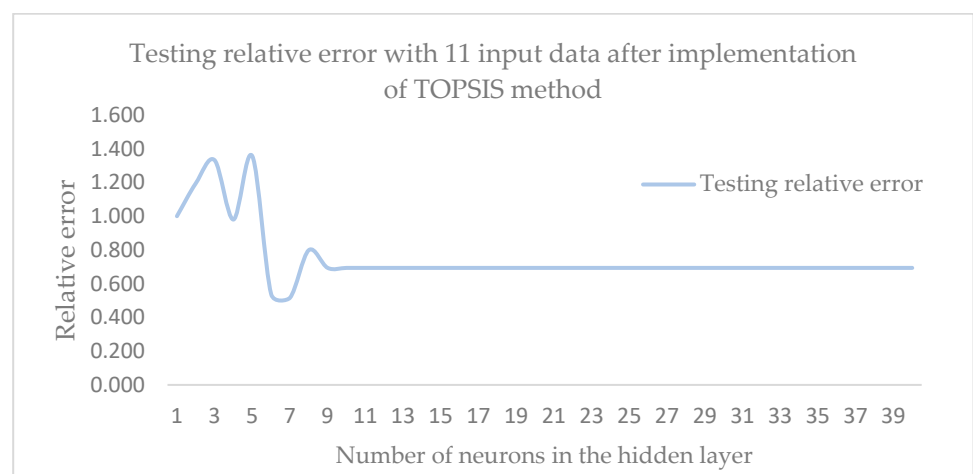


Figure 9. Testing sample of relative errors with 11 input data after implementation of TOPSIS method.

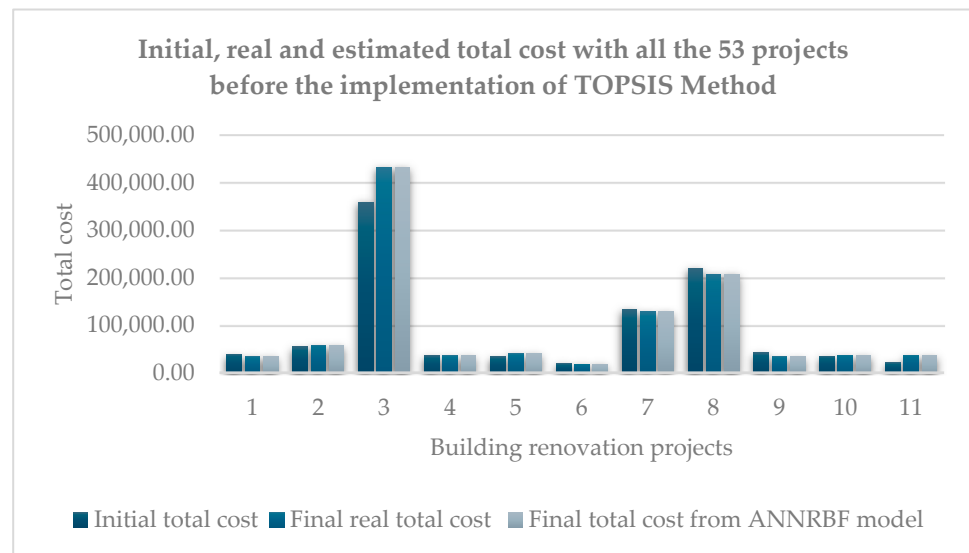


Figure 10. Initial, real and estimated total cost with all the 52 + 1 projects before the implementation of TOPSIS Method.

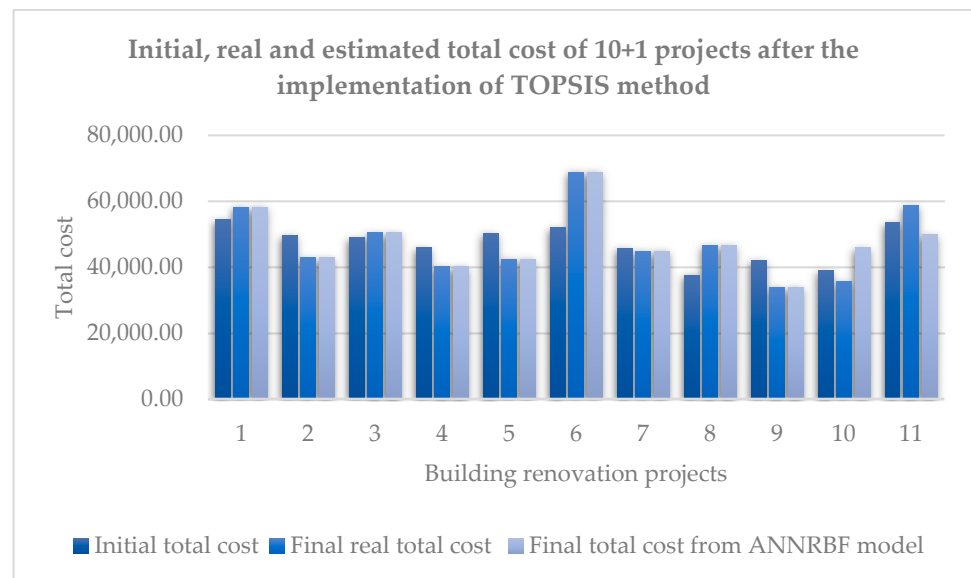


Figure 11. Initial, real and estimated total cost of 10 + 1 projects after the implementation of TOPSIS method.

6. Discussion

Utilizing an increasingly durable, multi-stage decision-support structure, the research effectively illustrates a hybrid approach for building cost estimation, surpassing single-technique models. According to the primary findings, an RBF-ANN model offers a precise and effective way to predict project costs and, when combined with the TOPSIS multi-criteria decision-making method, it can perform noticeably faster. The inherent complexity and data restrictions of early-stage project management are addressed by this method, which is a significant advancement.

Evaluation of the RBF-ANN model's performance, especially as recorded in the findings prior to TOPSIS application, demonstrates how the model can provide dependable predictions even with a small dataset of 53 projects. That the model's accuracy is sufficient but also exceptional is in line with an increased amount of the academic literature, as shown in Section 2, that acknowledges ANNs' improved performance over traditional techniques

like regression analysis and case-based reasoning. Cost forecasting is a notably valuable application for ANNs due to their capacity to learn by experience and generalize solutions for upcoming projects while making presumptions regarding their data structure. This is an important benefit since standard models frequently miss the complex, non-linear interactions between elements that affect costs.

The model's sufficient but also exceptional accuracy is supported by the results that are supplied in Section 5, which show how well the model performs according to quantitative criteria. The RBF-ANN model, which was trained on 53 projects, displayed a variety of performance measures prior to the TOPSIS method's implementation. The testing sum of square errors, for instance, varied; the most successful models had values between 0.25 and 0.45, whereas several of the models with more than 10 hidden neurons had corresponding testing relative errors between 0.411 and 0.735. A smaller selection of 11 projects was utilized to create a different model when the TOPSIS method was implemented into practice.

The performance was undoubtedly enhanced by this hybrid process. A relative testing error of 0.694 was frequently attained by the most robust models, while values as low as 0.517 were consistently found for the optimized models. The combined effect of the two approaches is demonstrated by this quantitative data, where the TOPSIS method serves as a crucial refining stage that raises the RBF-ANN model's predicted accuracy.

The two-step, hybrid approach employed in this study is the research's most convincing contribution. The main predictive engine, the RBF-ANN model, produces a cost forecast. The TOPSIS approach is then implemented as a ranking layer and post-prediction validation. TOPSIS is a tool for multi-criteria decision-making that evaluates options according to how far they are geometrically from the ideal answer. In this instance, it converts the ANN's raw numerical output into a more intricate, useful recommendation for construction managers.

A verifiable, data-driven method replaces subjective, experience-based evaluation by offering a reasonable basis to determine the most trustworthy estimate from a range of possibilities. This is not simply a combination of methods; rather, it is a major shift from prediction to evaluating. Stakeholder confidence in the estimated costs is raised, and the model's general applicability is improved by this combination.

From a pragmatic standpoint, the consequences for those involved in the building industry are substantial. Cost overruns and disputes are two major industrial hazards that the strategy directly mitigates. It makes it possible to make better financial plans and make more informed decisions by providing extremely accurate cost estimates early in a project. When complicated forecasting duties are automated, project managers can focus their attention on higher-value, important responsibilities. This approach has the potential to develop into an internal strategic asset, as seen by the research proposal to establish a well-built database of past initiatives from two companies.

Although the outcomes indicate assurance, the necessity of extending the data subset utilized for training and testing is an obvious restriction and a key subject for further research. It is frequently acknowledged that the amount and quality of data are crucial constraints for ANN usage. To address data inadequacy, future studies should investigate data enhancement strategies, especially for project types that are still in their infancy, such as environmentally friendly structures, where historical data is limited.

A reasonable and essential improvement would be to broaden the model's application to new fields, such as power plant and highway projects, proving that it is not just for building renovations.

In summary, this study effectively demonstrates the effectiveness of a hybrid ANN-TOPSIS framework for building cost estimation, bridging the gap between theoretical modeling and real-world applications. The effectiveness of ANNs as a better substitute for

traditional techniques is confirmed, and an integrated method that converts unprocessed forecasts into multi-criteria, actionable assessments is also established. The construction sector is encouraged to undertake major expenditures on these data-driven approaches and develop proprietary databases to produce dynamic, self-improved models that significantly enhance project predictability and promote operational efficiency over time.

7. Conclusions, Limitations, and Future Research

A novel, data-driven method of estimating building costs is persuasively supported by the research that has been provided. This project's primary contributions are threefold. First, it demonstrates that ANNs—specifically the RBF model—are a highly effective tool for cost forecasting, proving that they are more accurate than traditional techniques. Second, by fusing the evaluative capabilities of the TOPSIS technique with a forecasting ANN model, it develops a complex hybrid approach. A basic numerical forecast is transformed into a logical, multi-criteria decision-making aid that fulfills a crucial requirement for stakeholders according to this synergistic approach, which marks a substantial methodological development. The research concludes with a realistic strategy for development that highlights the establishment of an ongoing, self-enhancing mechanism built around a unique database of past project data.

When an ANN and TOPSIS are combined, the goal of the model is fundamentally changed. By producing a forecast for project expenses, the ANN serves as a predictive tool. As a tool for decision-making, TOPSIS offers a logical foundation for choosing or prioritizing the best option. This powerful combination transforms the algorithm from a straightforward, quantitative output to a qualitative, decision-maker-actionable advice.

With the goal of evaluating how to determine the most suitable method using the data, it is crucial to employ a variety of algorithms in software to assess correctness.

Implementing such data-dependent approaches is, without a doubt, the direction of the future for managing construction projects. This model is a vital tool for companies due to its capacity to produce precise estimates, reduce project risks, and simplify repetitive processes. To establish an economic advantage and establish the foundation for future efficiency improvements, stakeholders should make an effort to develop distinctive, customized models utilizing carefully chosen databases. The present study provides a solid framework for future research that attempts to fill the knowledge gap between academic theory and real-world, everyday application, thereby enhancing the accuracy and performance of building projects.

Additionally, this hybrid approach enables the researchers to navigate around the subjective drawbacks of conventional, experience-based assessment. Employing a methodical approach such as TOPSIS assures consistency and reduces the impact of individual prejudices, offering a more dependable and verifiable structure for making decisions.

The present study aims to stimulate more contemplation among building owners by providing them with a tangible (renovation alternatives, constructive details) and quantifiable viewpoint. In this way, they could encourage homeowners to consider renovating more and help put governmental initiatives to lower utility costs into practice.

Furthermore, it concluded that one of the fastest options for technological advancement is enhancing data. An enormous quantity of data is needed to create training patterns for machine learning models so they can produce forecasts that are trustworthy and precise. This inquiry might have gathered partially insufficient instances because of human capability limits.

Additionally, there also exists a dearth of research evaluating the study's applicability in additional regions or areas, considering that all the data utilized by this research originated in Central Macedonia, Greece. Through utilizing data from various regions and

worldwide environments, scientists need to improve the amount and diversity of training samples in further research. The precision, stability, and generalizability of the model will all be enhanced as a result.

Another goal of the present study is to provide governmental organizations with data collection guidelines as additional information to help them with the documentation procedure.

Further studies can significantly increase the robustness of the overall infrastructure. The enhancement of the current remodeling project estimation method is one of the study's primary advantages. Future outcomes for built-environment sustainability will undoubtedly be enhanced, provided that this kind of innovative method could quicken making choices for renovating building projects.

Instances of this include focusing on projects with similar cost requirements, within a single organization's database, or a particular building utilization type. Lastly, the models used in this study did not include variables that had to do with the features of the project, like the market, the economy, inflation, changes in labor productivity and expenditure, variations in the cost of materials, and changes in schedule. Consequently, it would be quite beneficial to create models that incorporate all these components.

Data augmentation approaches are a possible answer for new project types, like green buildings, when previous data are limited. Researchers can effectively mitigate data insufficiency and construct strong models for these emerging domains by combining data from small historical bases. This strategy guarantees that the model can be successfully applied to novel problems where conventional data collection techniques are not feasible.

The application of this RBF-ANN cost-estimating method to various project kinds and areas should also be investigated by researchers. For many cost-estimation situations where traditional approaches would not be sufficient, the methodology—in specific, the integration of TOPSIS for the selection of data and the RBF-ANN for prediction—could be an effective solution.

In hybrid architecture, the main obstacle is the influence of the multi-criteria preselection procedure on the ANN component's evaluation result. By filtering and ranking the most relevant previous projects employing the TOPSIS approach, the current design process effectively achieves both regional relevance and prognostic precision. This method guarantees that a restricted, highly similar portion of the range of features is inevitably the focus of the framework's reported dependability measures. Even though this targeted selection improves accuracy right away, further research is necessary to maintain the outer validity of the ANN's effectiveness when assessed globally throughout the whole database. Therefore, the smooth incorporation of the similarity score into the ANN's loss function ought to be the top priority for future investigations. These techniques will enable the system to successfully strike a balance between the two goals of complete global forecast accuracy and local relevance filtering.

In conclusion, a broader implementation of this technique in the building industry, along with increased comprehension among interested parties, may result in more precise cost estimates and fewer project delays.

Author Contributions: Conceptualization, V.E.P.; methodology, V.E.P.; software, V.E.P., G.N.A.; validation, V.E.P.; investigation, V.E.P.; data curation, V.E.P.; writing—original draft preparation, V.E.P.; writing—review and editing, V.E.P.; visualization, V.E.P.; supervision, G.N.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: The authors thank the *Algorithms* journal for its support and the reviewers for valuable feedback.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Abam, U.; Oduma, O.; Mbadike, M. Analysis of Cost Control Challenges in the Construction Industry, Using Artificial Neural Network. *Poljopr. Teh.* **2024**, *49*, 39–53. [\[CrossRef\]](#)
2. Waziri, B.S.; Bala, K.; Bustani, S.A. Artificial Neural Networks in Construction Engineering and Management. *Int. J. Archit. Eng. Constr.* **2017**, *6*, 50–60. [\[CrossRef\]](#)
3. Sanni-Anibire, M.O.; Mohamad Zin, R.; Olatunji, S.O. Developing a Preliminary Cost Estimation Model for Tall Buildings Based on Machine Learning. *Int. J. Manag. Sci. Eng. Manag.* **2021**, *16*, 134–142. [\[CrossRef\]](#)
4. Papadimitriou, V.E.; Aretoulis, G.N. A Final Cost Estimating Model for Building Renovation Projects. *Buildings* **2024**, *14*, 1072. [\[CrossRef\]](#)
5. Papadimitriou, V.E.; Aretoulis, G.N.; Papathanasiou, J. Radial Basis Function (RBF) and Multilayer Perceptron (MLP) Comparative Analysis on Building Renovation Cost Estimation: The Case of Greece. *Algorithms* **2024**, *17*, 390. [\[CrossRef\]](#)
6. Alqahtani, A.; Whyte, A. Artificial Neural Networks Incorporating Cost Significant Items towards Enhancing Estimation for (Life-Cycle) Costing of Construction Projects. *Australas. Constr. Econ. Build.* **2013**, *13*, 51–64. [\[CrossRef\]](#)
7. Matel, E.; Vahdatikhaki, F.; Hosseinyalamdary, S.; Evers, T.; Voordijk, H. An Artificial Neural Network Approach for Cost Estimation of Engineering Services. *Int. J. Constr. Manag.* **2022**, *22*, 1274–1287. [\[CrossRef\]](#)
8. Sitthikankun, S.; Rinchumphu, D.; Buachart, C.; Pacharawongsakda, E. Construction Cost Estimation for Government Building Using Artificial Neural Network Technique. *Int. Trans. J. Eng.* **2021**, *12*, 1–12. [\[CrossRef\]](#)
9. Kulkarni, P.S.; Londhe, S.N.; Deo, M.C. Artificial Neural Networks for Construction Management: A Review. *J. Soft Comput. Civ. Eng.* **2017**, *1*, 70–88. [\[CrossRef\]](#)
10. Hasware, M.; Baviskar, S.; Narwade, R. Prediction of Construction Materials Cost Using Artificial Intelligence Tools. *Int. J. Eng. Res. Technol. (IJERT)* **2019**, *8*, 276–279.
11. Amornsamankul, S.; Pimpunchat, B.; Triampo, W.; Charoenpong, J.; Nuttavut, N. A Comparison of Machine Learning Algorithms and Their Applications. *Int. J. Simul. Syst. Sci. Technol.* **2019**, *20*, 8.1–8.17. [\[CrossRef\]](#)
12. Gotlur, K.; Pujar, G.; Nayak, S.; Bh, J.R. Review On Prediction of Cost Overrun in Construction Projects Using Non-Traditional Methods. *Int. Res. J. Eng. Technol.* **2020**, *7*, 5728–5733.
13. Dilip, D.K.; Jesna, N.M. A Review on Application of Artificial Neural Networks for Predicting the Prices of Construction Materials. *Int. Res. J. Eng. Technol.* **2020**, *07*, 832–834.
14. Abioye, S.O.; Oyedele, L.O.; Akanbi, L.; Ajayi, A.; Davila Delgado, J.M.; Bilal, M.; Akinade, O.O.; Ahmed, A. Artificial Intelligence in the Construction Industry: A Review of Present Status, Opportunities and Future Challenges. *J. Build. Eng.* **2021**, *44*, 103299. [\[CrossRef\]](#)
15. Zhang, Y. Research on the Application of Artificial Intelligence-Based Cost Estimation and Cost Control Methods in Green Buildings. *Scalable Comput. Pract. Exp.* **2024**, *25*, 4772–4778. [\[CrossRef\]](#)
16. Shamim, M.M.I.; Hamid, A.B.b.A.; Nyamasvisva, T.E.; Rafi, N.S.B. Advancement of Artificial Intelligence in Cost Estimation for Project Management Success: A Systematic Review of Machine Learning, Deep Learning, Regression, and Hybrid Models. *Modelling* **2025**, *6*, 35. [\[CrossRef\]](#)
17. Elhag, T.M.S.; Boussabaine, A.H. An Artificial Neural System for Cost Estimation of Construction Projects. In Proceedings of the 14th Annual ARCOM Conference, Reading, UK, 9–11 September 1998; Association of Researchers in Construction Management. University of Reading: Reading, UK, 1998; Volume 1, pp. 219–245.
18. Arafa, M.; Alqedra, M. Early Stage Cost Estimation of Buildings Construction Projects Using Artificial Neural Networks. *J. Artif. Intell.* **2010**, *4*, 63–75. [\[CrossRef\]](#)
19. Feylizadeh, M.R.; Hendalianpour, A.; Bagherpour, M. A Fuzzy Neural Network to Estimate at Completion Costs of Construction Projects. *Int. J. Ind. Eng. Comput.* **2012**, *3*, 477–484. [\[CrossRef\]](#)
20. Petrusheva, S.; Zujo, V.; Zileska-Pancovska, V. Neural Network Prediction Model for Construction Project Duration. *Int. J. Eng. Res. Technol. (IJERT)* **2012**, *1*, 1646–1654.
21. Gulcicek, U.; Ozkan, O.; Gunduz, M.; Demir, I.H. Cost Assessment of Construction Projects through Neural Networks. *Can. J. Civ. Eng.* **2013**, *40*, 574–579. [\[CrossRef\]](#)
22. Lhee, S.C.; Flood, I.; Research, U.F.; Professor, H. Development of a Two Step Neural Network-Based Model to Predict Construction Cost Contingency. *J. Inf. Technol. Constr. (ITcon)* **2014**, *19*, 399–411.
23. El-Sawah, H.; Moselhi, O. Comparative Study in the Use of Neural Networks for Order of Magnitude Cost Estimating in Construction. *J. Inf. Technol. Constr. (ITcon)* **2014**, *19*, 462–473.

24. El-Sawalhi, N.I.; Shehatto, O. A Neural Network Model for Building Construction Projects Cost Estimating. *J. Constr. Eng. Proj. Manag.* **2014**, *4*, 9–16. [\[CrossRef\]](#)
25. Dissanayake, D.M.S.M.; Fernando, N.G.; Jayasinghe, S.J.A.R.S.; Rathnaweera, P.H.S.B. Preliminary Project Cost Estimation Model Using Artificial Neural Networks for Public Sector Office Buildings in Sri Lanka. In Proceedings of the Faculty of Architecture Research Unit (FARU) Research Conference 2015, Moratuwa, Sri Lanka, 11–12 December 2015; pp. 1–12.
26. Abd, A.M.; Naseef, F.S. Predicting the Final Cost of Iraqi Construction Project Using Artificial Neural Network (ANN). *Indian J. Sci. Technol.* **2019**, *12*, 1–7. [\[CrossRef\]](#)
27. Juszczak, M. Development of Cost Estimation Models Based on ANN Ensembles and the SVM Method. *Civ. Environ. Eng. Rep.* **2020**, *30*, 48–67. [\[CrossRef\]](#)
28. Mohammed, S.J.; Abdel-Khalek, H.A.; Hafez, S.M. Predicting Performance Measurement of Residential Buildings Using an Artificial Neural Network. *Civ. Eng. J.* **2021**, *7*, 461–476. [\[CrossRef\]](#)
29. Omotayo, T.; Bankole, A.; Olanipekun, A.O. An Artificial Neural Network Approach to Predicting Most Applicable Post-Contract Cost Controlling Techniques in Construction Projects. *Appl. Sci.* **2020**, *10*, 5171. [\[CrossRef\]](#)
30. Thakuria, R.; Parida, A.P. Study of Conceptual Cost in Construction Project in Indian Sub-Continent Field Using ANN. *Int. J. Res. Appl. Sci. Eng. Technol.* **2022**, *10*, 1650–1654. [\[CrossRef\]](#)
31. Gupta, A.; Debnath, P. Artificial Neural Network Analysis for Cost Estimation of Building Projects in India. In Proceedings of the International Conference of Contemporary Affairs in Architecture and Urbanism-ICCAUA, Alanya, Turkey, 11–13 May 2022; Volume 5, pp. 193–205. [\[CrossRef\]](#)
32. Ali, Z.H.; Burhan, A.M.; Kassim, M.; Al-Khafaji, Z. Developing an Integrative Data Intelligence Model for Construction Cost Estimation. *Complexity* **2022**, *2022*, 4285328. [\[CrossRef\]](#)
33. Al-Tawal, D.R.; Arafah, M.; Sweis, G.J. A Model Utilizing the Artificial Neural Network in Cost Estimation of Construction Projects in Jordan. *Eng. Constr. Archit. Manag.* **2021**, *28*, 2466–2488. [\[CrossRef\]](#)
34. Challa, R.K.; Rao, K.S. An Effective Optimization of Time and Cost Estimation for Prefabrication Construction Management Using Artificial Neural Networks. *Rev. D'intelligence Artif.* **2022**, *36*, 115–123. [\[CrossRef\]](#)
35. Aguinaldo, E.S.; Silva, D.L. Digitization and Digitalization of Cost and Duration Estimation in Philippine Public School Building Projects Using Artificial Neural Network. In Proceedings of the 2023 6th International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, 26–29 May 2023; Institute of Electrical and Electronics Engineers Inc.: New York, NY, USA, 2023; pp. 506–511. [\[CrossRef\]](#)
36. Deepa, G.; Niranjana, A.J.; Balu, A.S. A Hybrid Machine Learning Approach for Early Cost Estimation of Pile Foundations. *J. Eng. Des. Technol.* **2025**, *23*, 306–322. [\[CrossRef\]](#)
37. Trijeti; Irwanto, R.; Rahayu, T.; Panudju, A.T. Artificial Neural Networks for Construction Project Cost and Duration Estimation. *Rev. D'intelligence Artif.* **2023**, *37*, 1449–1460. [\[CrossRef\]](#)
38. Al-Somaydai, J.A.; Albadri, A.T.; Al-Zwainy, F.M.S. Hybrid Approach for Cost Estimation of Sustainable Building Projects Using Artificial Neural Networks. *Open Eng.* **2024**, *14*, 20220485. [\[CrossRef\]](#)
39. Alabdali, N. Area Base of Cost Estimation for Building Construction Projects Using an Artificial Neural Network. *AURUM J. Eng. Syst. Archit.* **2024**, *8*, 113–125. [\[CrossRef\]](#)
40. Kamilah, F.; Isvara, W. Conceptual Cost Estimation In Green Buildings By Using Regression Analysis and Artificial Neural Network Methods To Improve Accuracy. *J. Res. Soc. Sci. Econ. Manag. (JRSEM)* **2024**, *04*, 1687–1698.
41. Egwim, C.N.; Alaka, H.; Demir, E.; Balogun, H.; Olu-Ajayi, R.; Sulaimon, I.; Wusu, G.; Yusuf, W.; Muideen, A.A. Artificial Intelligence in the Construction Industry: A Systematic Review of the Entire Construction Value Chain Lifecycle. *Energies* **2024**, *17*, 182. [\[CrossRef\]](#)
42. Yadav, A.; Pipaliya, J. Review on Cost Overrun Analysis Using Artificial Neural Network in Residential Construction Projects. *Rabindra Bharati Univ. J. Econ.* **2024**, *XVI*, 69–73.
43. Padala, S.P.S.; Goyal, A. Early Stage Cost Prediction Model for Indian Building Construction Projects Using Artificial Neural Networks. *J. Financ. Manag. Prop. Constr.* **2025**, ahead-of-print. [\[CrossRef\]](#)
44. Chen, G.; Zheng, S.; He, X.; Liang, X.; Liao, X. Machine Learning-Based Cost Estimation Models for Office Buildings. *Buildings* **2025**, *15*, 1802. [\[CrossRef\]](#)
45. Cha, H.S.; Lee, D.G. Determining Value at Risk for Estimating Renovation Building Projects by Application of Probability-Based Fuzzy Set Theory. *J. Asian Archit. Build. Eng.* **2018**, *17*, 63–70. [\[CrossRef\]](#)
46. Bocaneala, N.; Mayouf, M.; Vakaj, E.; Shelbourn, M. Artificial Intelligence Based Methods for Retrofit Projects: A Review of Applications and Impacts. *Arch. Comput. Methods Eng.* **2024**, *32*, 899–926. [\[CrossRef\]](#)
47. Khodabakhshian, A.; Rampini, L.; Vasapollo, C.; Panarelli, G.; Cecconi, F.R. Application of Machine Learning to Estimate Retrofitting Cost of School Buildings. In *Resilient and Responsible Smart Cities*; Advances in Science, Technology and Innovation; Springer: Cham, Switzerland, 2023; pp. 215–228. [\[CrossRef\]](#)

48. Sacks, R.; Girolami, M.; Brilakis, I. Building Information Modelling, Artificial Intelligence and Construction Tech. *Dev. Built Environ.* **2020**, *4*, 100011. [[CrossRef](#)]
49. Flood, I.; Kartam, N.; Members, A. Neural Networks in Civil Engineering. II Systems and Application. *J. Comput. Civ. Eng.* **1994**, *8*, 149–162. [[CrossRef](#)]
50. Luu, V.T.; Kim, S.-Y. Neural Network Model for Construction Cost Prediction of Apartment Projects in Vietnam. *J. Korean Soc. Constr. Manag.* **2009**, *10*, 139–147.
51. IBM. *IBM SPSS Statistics 30 Core System User's Guide, Release 0, Modification 0 of IBM® SPSS®*, 30th ed.; IBM Corporation, Ed.; International Business Machines Corporation: New York, NY, USA, 2021.
52. Nong, J. The Design of RBF Neural Networks and experimentation for solving overfitting problem. In Proceedings of the ICEOE 2011: 2011 International Conference on Electronics and Optoelectronics, Dalian, China, 29–31 July 2011; IEEE: New York, NY, USA, 2011; pp. V1-75–V1-78.
53. Madanchian, M.; Taherdoost, H. A Comprehensive Guide to the TOPSIS Method for Multi-Criteria Decision Making. *Sustain. Soc. Dev.* **2023**, *1*, 1–6. [[CrossRef](#)]
54. Rohmatulloh, R.; Winarni, S. TOPSIS Method for Determining The Priority of Strategic Training Program. *Int. J. Adv. Sci. Eng. Inf. Technol.* **2014**, *4*, 67. [[CrossRef](#)]
55. Yoon, K. A Reconciliation among Discrete Compromise Solutions. *J. Oper. Res. Soc.* **1987**, *38*, 277–286. [[CrossRef](#)]
56. Lai, Y.-J.; Liu, T.-Y.; Hwang, C.-L. Theory and Methodology TOPSIS for MODM. *Eur. J. Oper. Res.* **1994**, *76*, 486–500. [[CrossRef](#)]
57. Jahanshahloo, G.R.; Lotfi, F.H.; Izadikhah, M. Extension of the TOPSIS Method for Decision-Making Problems with Fuzzy Data. *Appl. Math. Comput.* **2006**, *181*, 1544–1551. [[CrossRef](#)]
58. Tzeng, G.-H.; Huang, J.-J. *Multiple Attribute Decision Making: Methods and Applications*, 2011th ed.; CRC Press: Boca Raton, FL, USA, 2011; Volume 186.
59. Shih, H.S.; Shyur, H.J.; Lee, E.S. An Extension of TOPSIS for Group Decision Making. *Math. Comput. Model.* **2007**, *45*, 801–813. [[CrossRef](#)]
60. Google. Google Colaboratory. Available online: <https://colab.research.google.com> (accessed on 6 September 2025).
61. Zhang, G.; Zhang, Q. Machine Learning-Based Construction Cost Prediction Model and Optimization Design. *Int. J. High Speed Electron. Syst.* **2025**, 2540774. [[CrossRef](#)]
62. Sadikoglu, E.; Demirkesen, S. Review of Machine Learning and Artificial Intelligence Use for Cost Estimation in Construction Projects. In Proceedings of the 2025 European Conference on Computing in Construction-CIB W78 Conference on IT in Construction, Porto, Portugal, 14–17 July 2025.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.