modelling

MDPI

Systematic Review

Advancement of Artificial Intelligence in Cost Estimation for Project Management Success: A Systematic Review of Machine Learning, Deep Learning, Regression, and Hybrid Models

Md. Mahfuzul Islam Shamim [1,*], Abu Bakar bin Abdul Hamid [1], Tadiwa Elisha Nyamasvisva [2] and Najmus Saqib Bin Rafi [3]

[1] Faculty of Business, Information and Human Sciences (FBIHS), Infrastructure University Kuala Lumpur (IUKL), Kajang 43000, Selangor, Malaysia; abubakarhamid@iukl.edu.my
[2] Center for Postgraduate Studies, Infrastructure University Kuala Lumpur (IUKL), Kajang 43000, Selangor, Malaysia; tadiwa.elisha@iukl.edu.my
[3] Public Administration, Jahangirnagar University, Dhaka 1342, Bangladesh; nsaqib@allacademicresearch.com
* Correspondence: m.shamim@research.wust.edu

Abstract: This systematic review investigates the integration of artificial intelligence (AI) in cost estimation within project management, focusing on its impact on accuracy and efficiency compared to traditional methods. This study synthesizes findings from 39 high-quality articles published between 2016 and 2024, evaluating various machine learning (ML), deep learning (DL), regression, and hybrid models in sectors such as construction, healthcare, manufacturing, and real estate. The results show that AI-powered approaches, particularly artificial neural networks (ANNs)—which constitute 26.33% of the studies—, enhance predictive accuracy and adaptability to complex, dynamic project environments. Key AI techniques, including support vector machines (SVMs) (7.90% of studies), decision trees, and gradient-boosting models, offer substantial improvements in cost prediction and resource optimization. ML models, including ANNs and deep learning models, represent approximately 70% of the reviewed studies, demonstrating a clear trend toward the adoption of advanced AI techniques. On average, deep learning models perform with 85–90% accuracy in cost estimation, making them highly effective for handling complex, nonlinear relationships and large datasets. Machine learning models achieve an average accuracy of 75–80%, providing strong performance, particularly in industries like road construction and healthcare. Regression models typically deliver 70–80% accuracy, being more suitable for simpler cost estimations where the relationships between variables are linear. Hybrid models combine the strengths of different algorithms, achieving 80–90% accuracy on average, and are particularly effective in complex, multi-faceted projects. Overall, deep learning and hybrid models offer the highest accuracy in cost estimation, while machine learning and regression models still provide reliable results for specific applications.

Keywords: cost prediction models; artificial intelligence (AI); machine learning (ML); project management; project efficiency

check for updates

Academic Editor: Wei Gao

Received: 23 February 2025
Revised: 23 March 2025
Accepted: 21 April 2025
Published: 24 April 2025

Citation: 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. https://doi.org/10.3390/modelling6020035

1. Introduction

Cost estimation is a critical function in project management, influencing decisions related to budgeting, resource allocation, and financial performance throughout the lifecycle of a project. The success of any project, regardless of industry, hinges on the accuracy and reliability of cost predictions. Poor cost estimation can lead to budget overruns, delays, and compromised project outcomes [1,2]. Cost estimation is vital for the success of every project because it directly affects budgeting, allocation of resources, and ultimately performance. There were techniques that evolved from a very primitive stage until the introduction of more sophisticated statistical and machine learning methods to enhance the accuracy and efficiency of cost estimation [1]. Today, the discipline of cost estimation is being revolutionized with the introduction of artificial intelligence (AI), which has proven to enhance the accuracy of probable costing scenarios and allow for real-time adjustments based on real-time complex parameters. This paper presents a systematic review of applications of artificial intelligence in cost estimation for project management success [3,4].

Over time, various techniques have been developed to improve cost estimation, ranging from simple rule-of-thumb methods to more complex statistical models. Traditionally, these techniques relied on expert judgment, historical data, and manual calculations, but as projects' complexity and scale increased, these methods became insufficient for accurately predicting costs, particularly in dynamic and data-driven environments [4,5]. Consequently, the need for more sophisticated and reliable approaches to cost estimation has become more pressing [3].

The advent of artificial intelligence (AI) in recent years has introduced a transformative shift in cost estimation practices across various industries. AI, particularly machine learning (ML) and deep learning (DL), has emerged as a powerful tool to enhance the accuracy, efficiency, and scalability of cost estimation models [6]. Unlike traditional methods, AI-driven techniques have the ability to process large volumes of data, identify complex patterns, and continuously improve predictions based on real-time information [7]. These capabilities are crucial in industries like construction, healthcare, manufacturing, and real estate, where the project parameters are often complex and constantly changing, making accurate cost prediction a daunting challenge. One of the key motivations behind the integration of AI into cost estimation is the ability to address the limitations of traditional methods. While conventional models rely on expert input and historical data, they often fail to capture the intricate and nonlinear relationships between various cost-driving factors [5]. AI models, on the other hand, leverage advanced algorithms and vast datasets to make predictions based on both historical and real-time project data. These models not only improve accuracy but also allow for the dynamic adjustment of predictions as project conditions evolve [8].

AI helps predict material costs by finding patterns in market trends. By implementing machine learning and deep learning algorithms, it predicts changes in material prices caused by supply and demand, economic conditions, and past price trends, thus assisting in providing accurate and timely estimates of prices [9,10]. Artificial intelligence models can estimate the cost of labor on the basis of variables such as skill levels, time requirements, and labor demand fluctuations. AI algorithms based on contextual variables from the current labor market can help achieve higher accuracy in predicting labor costs by considering issues such as project complexity, labor availability, and regional wage variation [11].

AI enhances decision-making in cost allocation by analyzing various parameters, such as project scope, resource availability, and potential risks. Through advanced algorithms, AI can optimize budget distribution by identifying cost-effective solutions and allocating resources where they are most needed. This leads to more efficient project management and reduces the likelihood of cost overruns.

This paper presents a systematic review of the application of AI-driven cost estimation models in project management, focusing on ML, DL, regression models, and hybrid AI approaches. This review synthesizes findings from studies published between 2016 and 2024, examining the performance of these AI models in diverse project environments and across various sectors. The objective of this review is to evaluate the performance of AI technologies that have enhanced cost estimation practices, identify the challenges associated with their adoption, and provide insights into the future potential of these models in transforming project management.

A significant aspect of this review is its focus on the practical applications of AI in real-world projects. By analyzing studies across different industries, this review explores how AI models are being utilized to predict costs more accurately, optimize resource allocation, reduce cost overruns, and improve decision-making throughout the project lifecycle. The ability to leverage AI for real-time adjustments and scenario-based predictions has the potential to revolutionize cost estimation, particularly in large and complex projects, where traditional methods often fall short. Additionally, AI models can improve the responsiveness of project managers to unexpected changes, such as fluctuations in material costs, labor availability, or regulatory changes, by providing more accurate forecasts and actionable insights.

2. Literature Review

2.1. Historical Evolution of Cost Estimation Techniques

Cost estimation has always been one of the most important aspects of project management and resource allocation. Throughout the years, it has also reflected the development of cost estimation techniques for advanced technology and increased computational power, as well as the heightening difficulty of projects undertaken by human beings. In the early civilizations, the method of cost estimation was largely rudimentary and empirical. Builders or artisans estimated costs according to personal experience, rules of thumb, and the historical accumulated experience of everyone. Building projects such as the Pyramids of Giza and the Great Wall of China illustrate this [2]. In the Middle Ages, guilds and master craftsmen devised standardized measurement units and prices for labor, resulting in slightly more formalized approaches for estimating some cathedrals and castles [1]. This kind of project grew from the 17th to the 19th century, increasing in complexity and size. Cost estimation methods really took off during this period for project works, through elementary arithmetic calculations and material costing [3]. Seymour and Hussein (2014) indicated that the modern era of project management started in the 20th century, where Gantt charts and other technological tools were integrated with project management and cost management [4]. Systematic documentation and formal procedures of cost management have made cost management more and more formalized. Mathematical advances in statistical and probabilistic models entered the arena of cost estimation in the 20th century. Techniques such as parametric estimation, which uses historical data, developed mathematical relationships between project parameters and costs. Cost estimation has undergone changes with the advancement of computers [12]. The use of machine learning algorithms, such as regression models, decision trees, and support vector machines, is growing for predicting project costs from historical data. These methods enhance performance by capturing complex and nonlinear relations among variables [13]. Figure 1 represents the evolution of the cost estimation technique.

2.2. Techniques and Tools for Cost Estimation

In recent decades, the cost estimation discipline has evolved from its traditional method of advanced AI-assisted approaches [8]. Traditional estimates rely on manual calculations, expert judgment, and previous experiences and data [6]. Techniques such as analogous estimation, parametric estimation, bottom–up estimation, etc., are widely used, from construction to manufacturing [14]. Analogous estimation takes data from previous projects for comparison with projects that are common in nature. Parametric estimation is based on statistical data and relationships of varying parameters (e.g., project size, project scope), which mostly give a somewhat better estimate if a reasonable model is chosen. More detailed estimates are fed into bottom–up estimation based on the lowest levels of work breakdown structure (WBS) segments [5]. Such techniques consume time and present more chances for errors, as they rely directly on human judgment in managing the data collected from project plans. These methods are often labor-intensive, with a high probability of errors, and heavily dependent on the individual's judgment. Conversely, AI-based cost estimation methods are greatly powered by machine learning (ML), deep learning (DL), and predictive analytics, which require less time and effort for cost estimation. The advanced setups automate the analysis of large, complex datasets, far exceeding traditional methodologies in producing accurate, real-time cost predictions [7]. Essentially, AI is being used by project managers to reveal hidden patterns and insights that might be readily neglected during manual analysis. AI algorithms analyze historic project data, market trends, inflation rates, labor costs, or material price fluctuations to generate very accurate cost forecasts. They may also adjust to changing project patterns and adjust their predictions over time. AI-based cost estimation models could provide enhanced reliability and applicability in complex and large domain projects, with the learning of new patterns and trends in various fields [15]. The combination of a multitude of data across various sources, analyzed in real time, has transformed AI-based cost estimation processes, which are quicker, more efficient, and can handle modern-day project management issues with high accuracy and confidence in the decision-making path [7,16]. Figure 2 shows the approaches of cost estimation.

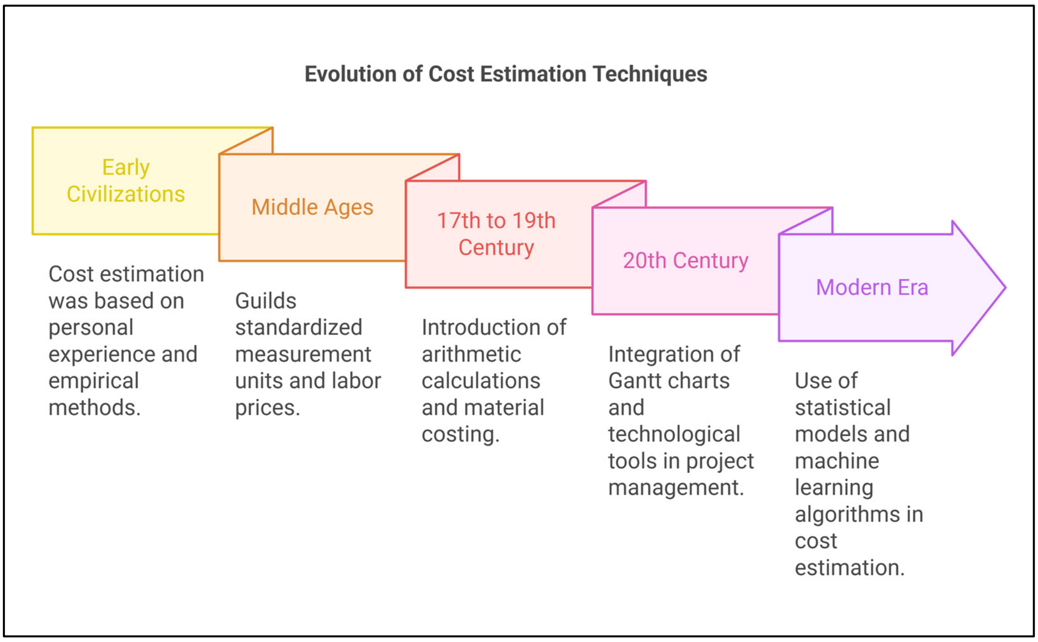

Figure 1. Evolution of cost estimation techniques.

2.3. AI-Enabled Decision Support Systems in Cost Prediction

The economy of the present world is changing rapidly with the integration of technology-driven resources. Project management's success depends on the perfect use of resources that are relevant to the project implementation method. Project cost is one of the catalysts that influence success from the start to the end of the project lifecycle. It is necessary to predict the cost of the project in the very early stages of the project's initiation [17]. Artificial intelligence has brought significant changes in every field, including the various project management fields, as it has permitted organizations to operate in dynamic and uncertain environments with utmost precision and flexibility. AI-powered decision support systems integrate machine learning (ML), deep learning, data mining, and predictive analytics, incorporating large amounts of historical and real-time data [18]. Again, these tools extract actionable insights by identifying hidden patterns, trends, and correlations that often remain hidden under the radar of more conventional techniques. Popular models in data-based decision-making, which are used to predict costs for the labor market, material prices, or fluctuations in the market, include neural networks and regression models [16]. AI systems keep learning and adapting over time to increase their predictability, unlike conventional statistical methods. One major area of cost prediction where AI has proven useful is in dynamic environments characterized by swiftly changing variables [19], because disparate material price fluctuations, shortages of labor, and alterations in legislation can create a burden in construction and manufacturing projects' implementation [20]. Artificial intelligence can draw upon its access to real-time data from an array of sources—such as the IoT and market-database-oriented deployments—to provide timely and accurate cost forecasts [21]. When organizations can learn based on the processing of dynamic inputs, they will always be in the position to make the best decisions that provide solutions for optimizing the utilization of resources and ensuring that risks are mitigated [22]. AI-enabled DSSs can make decision-making a lot more meaningful, diminishing uncertainty while providing scenario-wise analysis, thus empowering the decision-makers with the option to choose the most cost-effective plan by exploring scenarios through the simulation of various outcome options based on different input variables [23].

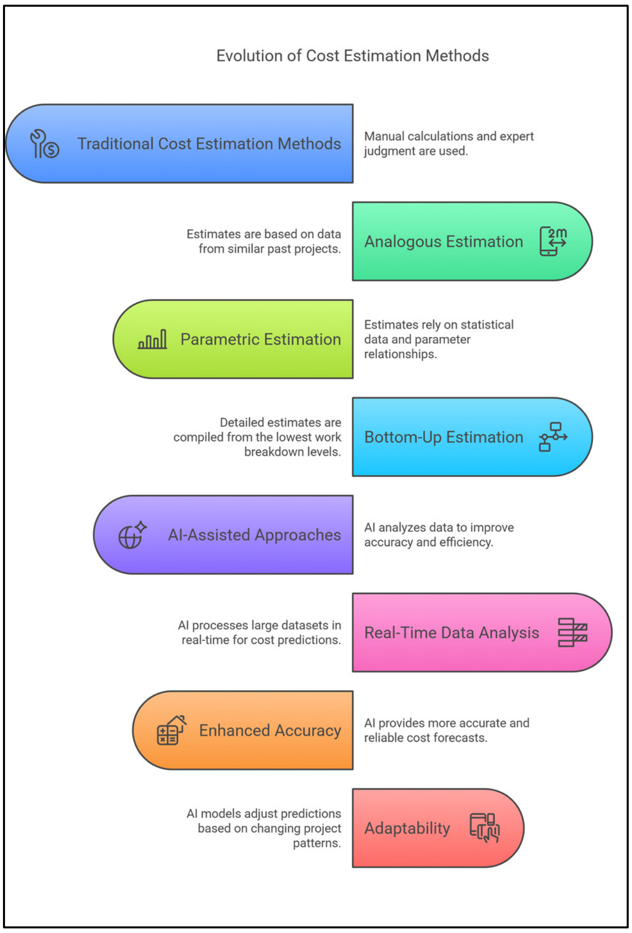

Figure 2. Evolution of cost estimation models.

2.4. AI Techniques Used in Cost Estimation

AI in cost estimation mostly depends on the collaboration of sophisticated algorithms and the perfect use of technologies. AI models have increased accuracy in cost estimation due to integrating machine learning (ML) and deep learning (DL) with other computational techniques that are relevant to project management processes. Various AI methods can improve reliability and efficiency in cost estimation after reducing the human error factor and fine-tuning the budgeting approach. Artificial neural networks (ANNs) are considered to be among the effective models that could replicate and make decisions similar to the function of the human brain. ANNs are very effective on complex datasets with nonlinear relationships, applying to sectors such as construction, manufacturing, and healthcare, where project dynamics might prove unpredictable. Decision trees, support vector machines (SVMs), and artificial neural networks (ANNs) are some of the other useful supervised learning algorithms that are employed for cost prediction [24]. These models are trained on historical cost data to discern certain patterns and correlations, hence increasing the accuracy of estimation. Regression models, such as linear regression and polynomial regression, are also helpful for predicting costs against several project parameters [25].

For more complex data processing projects, deep learning models can utilize CNNs and RNNs to predict and identify trends and actual costs [26,27]. RNNs and Long Short-Term Memory (LSTM) networks are effective in time-series forecasting, making them suitable for predicting cost trends over time [28,29]. CNNs, combined with NLP, can assist with textual data such as project specifications to extract cost-related knowledge [26,30]. Expert systems leverage rule-based AI to simulate human expertise in cost estimation, integrating predefined rules for different project scenarios. Fuzzy logic systems can handle uncertainty in cost estimation by assigning degrees of truth to various input variables, making them useful for projects with incomplete or vague data [31].

2.5. Enhancing Cost Estimation Processes with AI: Key Areas of Application

AI has shown enormous potential in improving project-management-related cost estimation. However, some important aspects of the cost estimation process for which the literature is scarce include pre-measurements, price level determination, and forecasting of valorization. To address this gap, this section will briefly describe the fundamental ways in which AI can assist in such stages of cost estimation.

The pre-measurement stage is very important for establishing a baseline for cost estimation. Traditionally, this operation has depended heavily on expert judgment and historical data [5]. Here, AI-powered processes like machine learning and deep learning algorithms could automate the extraction and analysis of data from historical projects, minimizing human error, and allowing for a smooth and efficient process. Again, these algorithm techniques can develop image recognition with any unstructured data source, e.g., architectural drawings or blueprints that could yield accurate measurement values, and help assure consistency in pre-measurements and accuracy from project to project [1].

Dynamic pricing for materials, labor, and other project components undergoes constant changes based on prevailing market conditions. AI facilitates the exact pricing for real-time data of varied sources, ranging from commodity markets to labor rates, as well as the trends followed in regional costs [9,11]. Advanced machine learning models, such as regression analysis or support vector machines (SVMs), can use patterns to establish future predictions of air prices, enabling project managers to determine the probability of price changes for their cost estimates to be more accurate, open, and adaptable for project budgets [31,32].

Valorization refers to the process of determining the value or worth of a project or its components over time [33]. Most artificial intelligence models, especially specific deep learning techniques such as artificial neural networks and recurrent neural networks, make these engines particularly relevant for forecasting valorization, due to their capability of handling vast amounts of time-series data and finding recurrence patterns [19]. A trained model can include historical values and enable the comparison of how these changes might affect other external variables, for example, trends in the market, inflation, or project-based scope changes [10]. Aside from these stated applications, AI could automate the routine, provide better predictions, and enable close-to-real-time adjustments depending on project conditions. With AI tools, the cost estimation process can be made more efficient, accurate, and adaptable—aspects that are not to be overlooked in light of the great complexity and size of present-day projects.

The use of AI integration in these specific phases demarcates the scope of this work, focusing on the application of AI technology in the cost estimation process specific to the project management context [23]. By incorporating AI tools, the cost estimation process can be made more efficient, accurate, and adaptable, which is critical for managing the increasing complexity and scale of modern projects [34].

3. Methodology

The systematic review in this manuscript adheres to the guidelines outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). The review methodology was conducted following the PRISMA framework, which includes stages of identification, screening, eligibility assessment, and inclusion. All relevant studies meeting the inclusion criteria were included, and methods for data extraction and thematic analysis were rigorously followed (Figure 3). The review is in compliance with the PRISMA guidelines, and the study's registration details, if applicable, will be provided in the Supplementary Materials.

3.1. Identification Phase

We conducted a systematic screening process to determine which studies would be included in the synthesis, in line with the PRISMA guidelines. Only those studies that explicitly focused on AI-driven cost estimation techniques were selected. To identify the relevant literature on AI-focused cost estimation in project management, a search strategy was developed, which we carried out across many databases—Scopus, IEEE Xplore, Web of Science, ScienceDirect, and Google Scholar. The search strategy was restricted to peer-reviewed journal articles, conference papers, and technical reports published from 2016 to 2024. The primary search terms included "artificial intelligence" OR "AI" AND "cost estimation" OR "cost prediction", along with additional terms such as "project management" AND ("cost management" OR "cost estimation") and "machine learning" OR "ML" AND "cost prediction". The goal was to capture studies discussing AI in cost estimation for project management success. Initially, 917 articles were retrieved. To ensure a comprehensive dataset, manual searches were also performed by reviewing the reference lists from key publications.

3.2. Screening Phase

Duplicate articles were removed using Zotero reference management tools, reducing the dataset to 726 unique records. The remaining articles underwent a title and abstract screening, where studies not related to AI-driven cost estimation in project management were excluded. Two independent reviewers assessed each article based on predefined inclusion and exclusion criteria. In the screening and selection process of this systematic review, two independent reviewers were involved to ensure the objectivity and consistency of article assessment. Each article was independently evaluated by both reviewers using a predefined set of inclusion and exclusion criteria, which were developed to ensure the relevance and quality of the studies included in the review. The inclusion criteria focused on articles that explicitly discussed AI-driven cost estimation techniques, such as machine learning, deep learning, and hybrid models, in the context of project management. The exclusion of 510 articles in this study was based on several factors. Many articles were excluded for lacking a clear focus on AI-driven cost estimation in project management, instead addressing other aspects of project management, such as scheduling or resource allocation. Some studies did not provide sufficient application details of AI techniques, or their datasets were too small or lacked statistical power, reducing their reliability. Articles that presented niche, non-generalizable case studies or were redundant with other included research were also excluded. Additionally, studies that focused solely on traditional cost estimation methods without incorporating AI, those without measurable results or performance metrics, and those published before the study's 2016 cutoff were excluded to ensure that the review remained focused on relevant, up-to-date, and scientifically rigorous sources. Based on these criteria, 510 articles were excluded from the total 726 articles. After this phase, 216 articles were selected for full-text review.

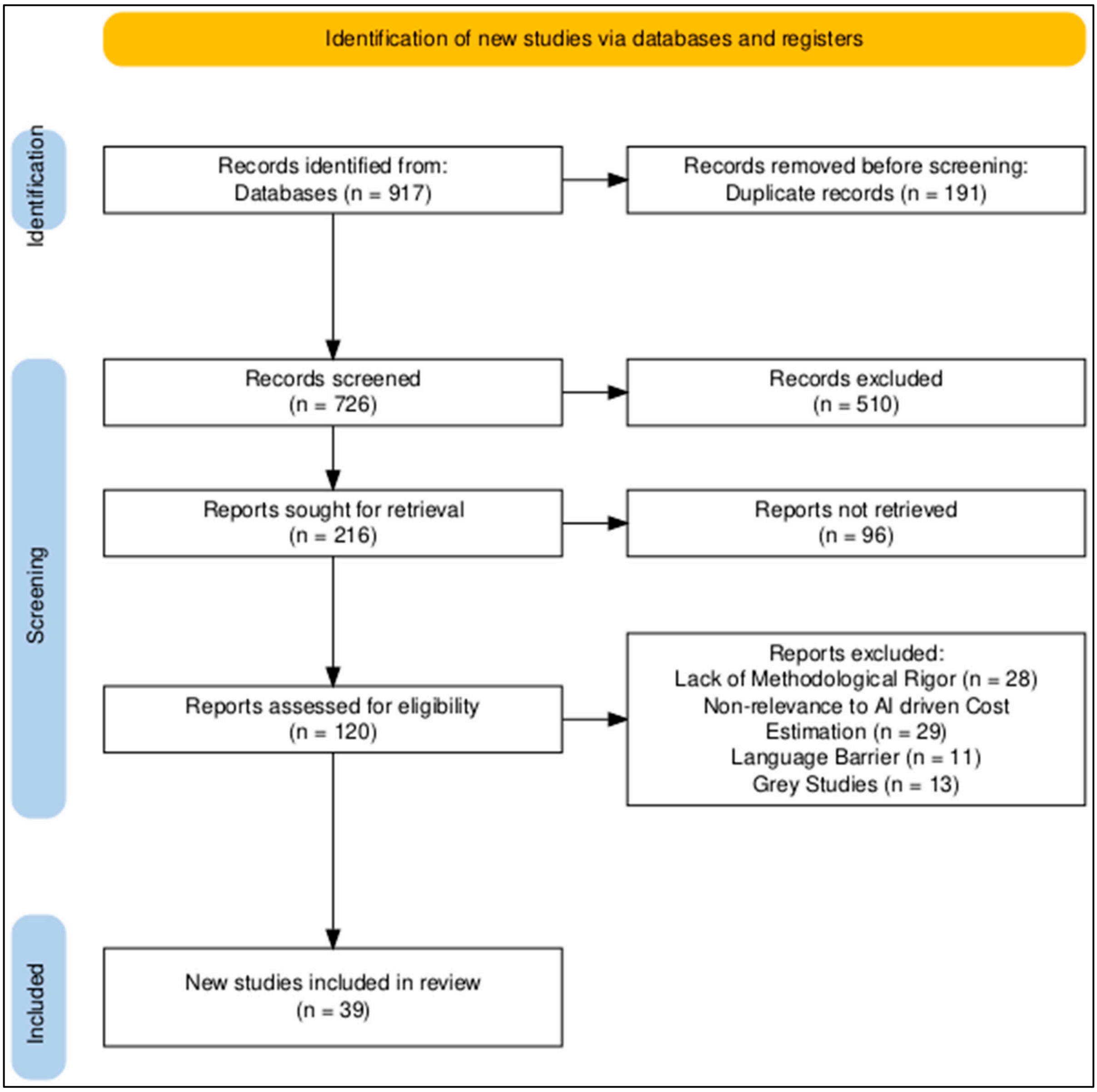

Figure 3. PRISMA flow diagram of the study.

3.3. Eligibility Assessment Phase

A full-text review was conducted on the 216 articles to assess their methodological quality, relevance, and contribution to AI-driven cost estimation in project management. To ensure consistency, the Critical Appraisal Skills Programme (CASP) checklist and the Cochrane Risk-of-Bias Tool were used. Articles were evaluated based on their methodological rigor, data reliability, relevance to engineering and cybersecurity, and the clarity of their risk assessment framework applications. After this phase, 39 high-quality articles were selected for detailed analysis.

3.4. Data Extraction and Thematic Analysis

A data extraction sheet was developed to document key details from the selected articles, including their title, author(s), publication year, research methodology, cybersecurity framework discussed, and key findings. The extracted data were analyzed through comparative analysis and narrative synthesis to identify common themes and the effectiveness of various cybersecurity risk frameworks. The studies were categorized based on their approach to AI-driven cost estimation in project management. For synthesis, we focused on key trends, such as the superiority of hybrid models and AI-driven techniques like artificial neural networks (ANNs) and XGBoost in cost estimation. To identify the presence of statistical heterogeneity, a subgroup analysis was conducted based on the industry type and model complexity. This allowed for an assessment of how various factors, such as the project type or dataset size, influenced the models' performance. To ensure the robustness and reliability of the synthesized results, sensitivity analyses were performed to examine the potential impact of varying the inclusion criteria—such as study designs, data types, and analysis methods—on the overall findings. The results of the sensitivity analyses indicated that the main findings remained robust, even when the inclusion criteria were adjusted, suggesting that the synthesized results are not overly sensitive to the variations in the study parameters.

4. Findings

4.1. General Information of the Studies

This segment (Table 1) presents general information about the studies included in the review. The table compiles all shortlisted papers selected through the PRISMA methodology. Each entry provides key details such as author, year, study design, and main findings.

Table 1. General study information of selected papers.

Study	First Author	Year of Publication	Citation	Study Design	Data Location Collection	Data Location Collection
[35]	Fangwei Ning	2020	56	Experimental study	China	121,980 cases
[36]	Debaditya Chakraborty	2020	184	Experimental study	Data from RSMeans Assemblies Books	Medium- and high-rise buildings consisting of 4477 data points
[37]	Abolfazl Jaafari	2021	27	Experimental study	Iran	4811 data samples collected from 300 road segments
[8]	Erik Matel	2022	133	Experimental study	Data collected from engineering consultancy firms	132 engineering projects

Table 1. *Cont.*

Study	First Author	Year of Publication	Citation	Study Design	Data Location Collection	Data Location Collection
[38]	Jean-Loup Loyer	2016	161	Case study approach	-	254 jet engines' data
[39]	Odey Alshboul	2022	95	Experimental study	Various locations across North America	283 building data
[40]	Haytham H. Elmousalami	2020	126	Experimental study	Egypt	144 FCIPs
[41]	Agnieszka Leśniak	2018	123	Case study approach	Poland	143 construction projects' data
[42]	Michał Juszczyk	2018	83	Experimental study	Poland	115 construction projects' data
[43]	Hongquan Guo	2021	87	Experimental study	N/A	74 open-pit mining projects
[44]	Zainab Hasan Ali	2022	24	Experimental study	Iraq	90 building projects' data
[45]	Ran Wang	2022	52	Experimental study	Hong Kong	98 public school construction projects
[46]	Mohammad Hossein Rafiei	2018	212	Experimental study	Iran	372 low- and midrise building
[47]	Marcin Relich	2018	113	Case study approach	Poland	61 new product development projects
[48]	Zaher Mundher Yaseen	2020	159	Experimental study	Iraq	40 completed construction projects
[49]	Alexandre Vimont	2022	38	Experimental study	France	510,182 subjects
[50]	Madhu Mazumdar	2020	43	Simulation study	United States	4205 subjects
[51]	Ch. Anwar ul Hassan	2021	43	Experimental study	N/A	1338 subjects
[52]	Mohammad Amin Morid	2020	47	Simulation study	United States	91,000 individuals
[53]	Laís B. Barros	2018	24	Experimental study	Brazil	14 highway projects
[54]	Winky K.O. Ho	2021	254	Experimental study	Hong Kong	40,000 housing transactions
[55]	Haishuai Wang	2018	245	Predictive modeling	United States	41,503 patient visits
[56]	Seokheon Yun	2022	23	Case study approach	South Korea	908 construction cases
[57]	Amol Tatiya	2018	101	Case study approach	United States	530 deconstruction projects
[34]	Jiacheng Dong	2020	23	Experimental study	China	143 months of construction cost index data

Table 1. *Cont.*

Study	First Author	Year of Publication	Citation	Study Design	Data Location Collection	Data Location Collection
[58]	Viren Chandanshive	2019	58	Experimental study	India	78 building projects
[59]	T.Q.D. Pham	2023	43	Experimental study	N/A	10,000 parametric building configurations
[60]	Hosam Elhegazy	2022	57	Decision-making study	United States	More than 900 data points
[61]	Ashraf Abdulmunim Abdulmajeed	2021	18	Case study approach	United States	60 completed software projects
[62]	Othman Subhi Alshamrani	2017	68	Experimental study	North America	320 construction projects
[63]	Xiaonan Chen	2020	19	Experimental study	N/A	22 samples of general aviation aircraft
[64]	Haytham H. ElMousalami	2018	126	Experimental study	Egypt	144 field canal improvement projects
[65]	Miljan Kovačević	2021	42	Experimental study	Serbia	181 bridge construction projects
[66]	Mohammad Amin Morid	2019	40	Experimental study	United States	3.8 million medical claims and 780,000 pharmacy claims from 24,000 patients
[67]	Bo Rang Park	2018	41	Experimental study	South Korea	11 monitoring areas from a test building in winter
[68]	Gagan Kumar Patra	2024	37	Experimental study	N/A	1388 health insurance records
[32]	Igor Peško	2017	80	Experimental study	N/A	166 road construction projects
[69]	Ahmed I. Taloba	2022	47	Experimental study	Japan	24,353 patient records
[70]	Jesam Ujong	2022	59	Experimental study	Nigeria	78 responses from different construction professionals

4.2. Cost Estimation Trends

Figure 4 presents the year-wise distribution of studies focused on various cost estimation models, including deep learning, hybrid models, machine learning, and regression models. The chart highlights a clear upward trend in the number of studies from 2016 to 2024, particularly in the years 2021 and 2022, marking a significant shift toward more advanced AI models in project management. Initially, the years 2016 to 2018 saw a low number of studies, but starting in 2019, research began to ramp up, with deep learning and machine learning models gaining prominence. Deep learning stood out in 2022 as the dominant model, with hybrid models contributing to this increase. Notably, machine learning experienced notable growth in 2021, reflecting its growing adoption for cost estimation in project management. In contrast, regression models showed a steady but much smaller presence, without the significant spikes seen in the other models. The steady rise in AI-driven methodologies, particularly deep learning and hybrid models, suggests a transition toward more complex and accurate predictive techniques, aiming to address the challenges of cost estimation in fields like construction, healthcare, and manufacturing. The tapering off of studies toward 2023–2024 indicates a possible plateau in the research interest or a shift toward refining these models for broader implementation. This trend underscores the increasing reliance on data-driven and automated approaches for cost estimation, signaling a move away from traditional methods.

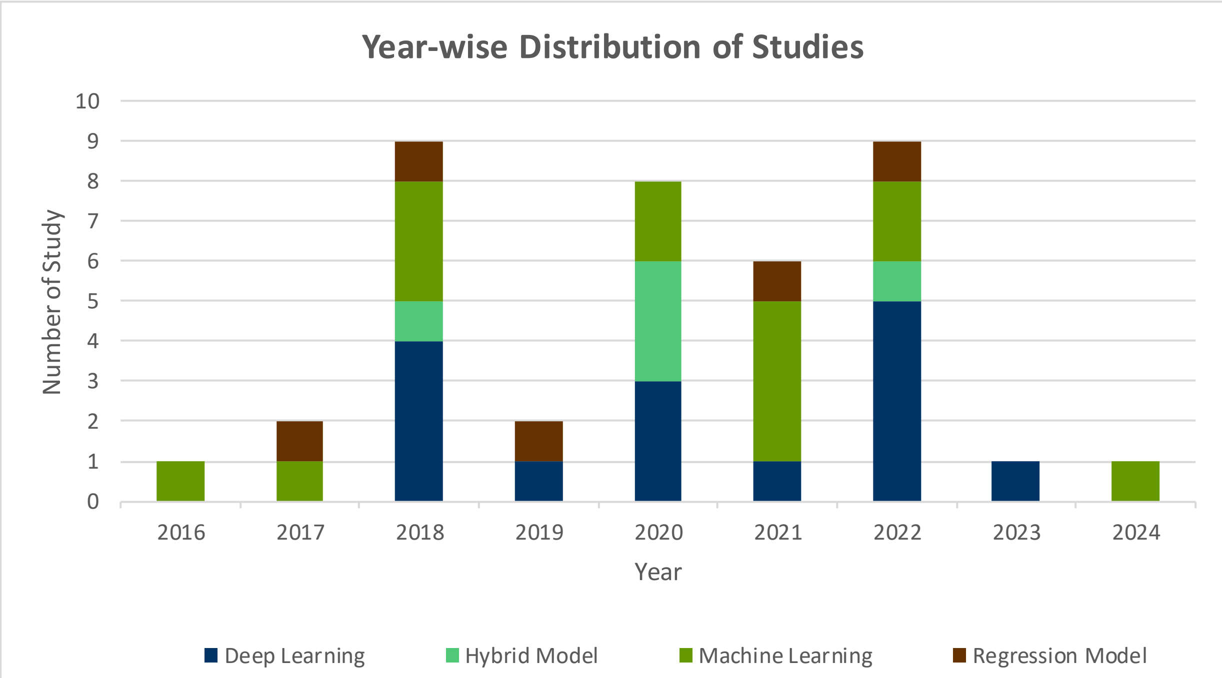

Figure 4. Trends in cost estimation models: year-wise distribution of studies.

The trend suggests that there is a developing inclination to use more sophisticated models, especially deep learning, which has received considerable attention in the last few years. This migration indicates that the research community is venturing into and adopting intricate and accurate means of predicting costs in, inter alia, the fields of construction and project management. This chart evidences the explosion of advanced machine learning and hybrid models in the past few years, accentuating a transition toward increasingly automated, data-driven paths to cost estimation.

4.3. Relationship Between Study Count and Citation Impact in Cost Management

The chart provides an analysis of the relationship between the study count and citation impact in the field of cost estimation and management. It is evident that some topics attract a larger number of studies, while others generate a higher citation impact. For example, "Improving cost management" stands out with the highest number of studies (seven), and it also shows a high citation count (close to 900). "Reducing cost estimation errors" and "Preventing cost overruns" show a similar trend, with a relatively higher citation sum, although the study count is lower (around 3–4 studies).

On the other hand, "Optimizing cost estimation" and "Stabilizing cost prediction" have fewer studies (3–5 and show more moderate citation counts (around 600–700). Interestingly, "Enhancing risk adjustment" has a minimal number of studies (only two), yet it maintains a decent citation count (about 300), indicating that fewer studies in niche areas can still make a significant academic impact.

The comparison between study count and citation impact highlights that a greater number of studies does not necessarily correlate with a higher citation impact (Figure 5). "Reducing bias in cost estimation" and "Supporting strategic financial planning" are examples where the study count is moderate yet the citation impact is significantly higher, suggesting that these areas have attracted more attention and relevance in the academic community. This emphasizes that the citation impact can often reflect the importance and influence of specific topics within cost estimation and management.

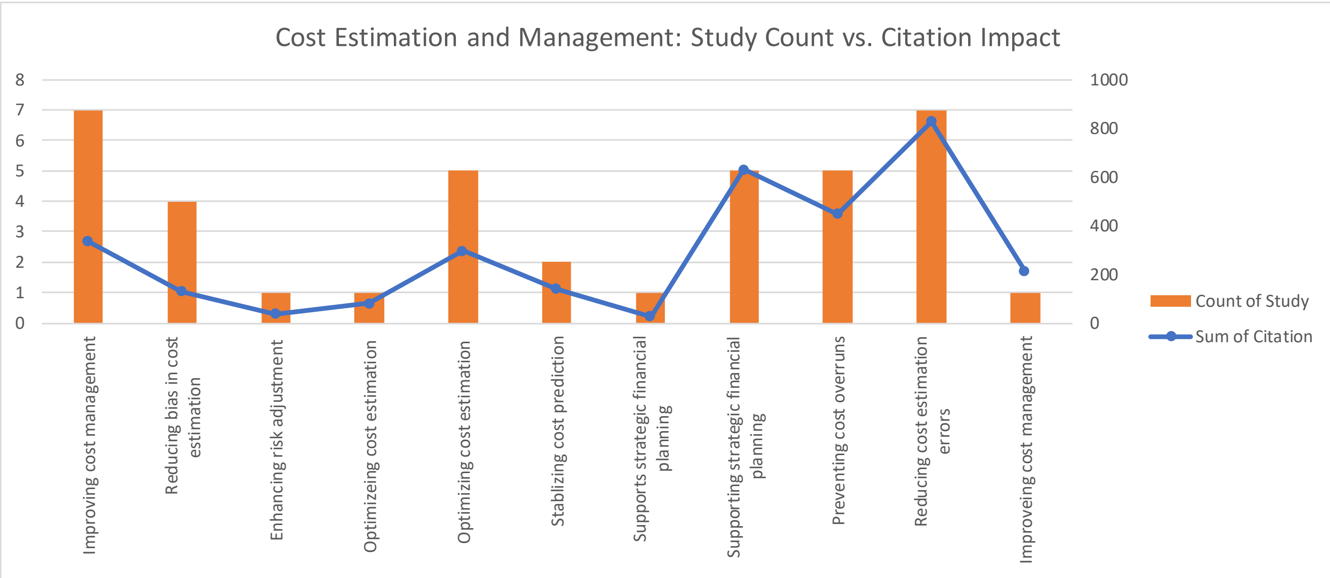

Figure 5. Distribution of studies and citations in cost estimation and management research.

Overall, the findings demonstrate that the most influential topics are not always the ones with the highest number of studies. Instead, topics like "Improving cost management" lead in both quantity and citation impact, while other focused areas like "Enhancing risk adjustment" might have fewer studies but still hold valuable academic weight. This underlines the importance of both quantity and quality in the research output for shaping future directions in cost estimation.

4.4. Using of AI Cost Estimation Models in Different Projects

Deep learning, according to the data, is the most widely used technique among various industries, with 15 studies, followed by machine learning with 14, while hybrid models and regression models are covered within 5 studies each (Figure 6). Among the most common applications in the building construction sector are deep learning (7), hybrid models (3), and machine learning (2). Similar approaches are observed in the healthcare sector, with decent employments of machine learning (4) and deep learning (2). Road construction, field construction, and general construction have moderate usages, mostly towards machine learning and deep learning. Meanwhile, other industries, like aviation, bridge construction, and mining, have relatively less use of these advanced techniques in their studies, suggesting these areas may have room for improvement by trying new methods. The data show that machine learning is gradually extending its application into the infrastructure studies field; for instance, it is being used widely in canal and highway construction. This also suggests a possible wider use for hybrid models and regression models in less-involved industries. In summary, however, although machine learning is still the kingpin, there is still potential for research and wide adoption across various industries.

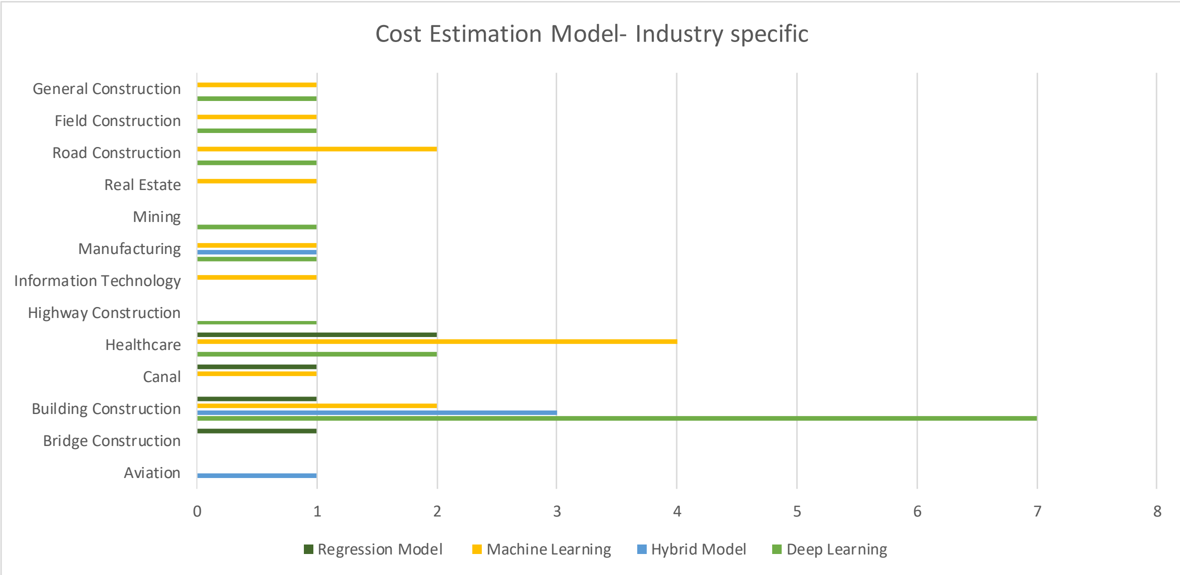

Figure 6. Distribution of cost estimation models used in different projects.

4.5. Percentage of AI Models in Industries

The pie chart (Figure 7) depicts the distribution of some machine learning models used in studies on cost estimation. The most widely applied model is the artificial neural network (ANN), with a contribution of 26.33% to the total studies. Some other models, such as convolutional neural networks (CNNs), XGBoost, and support vector machines (SVMs), are represented by 7.90% each. The remainder comprises K-Nearest Neighbors (KNN), principal component analysis with ANN (PCA-ANN), and regression models, with each making a 2.63% contribution.

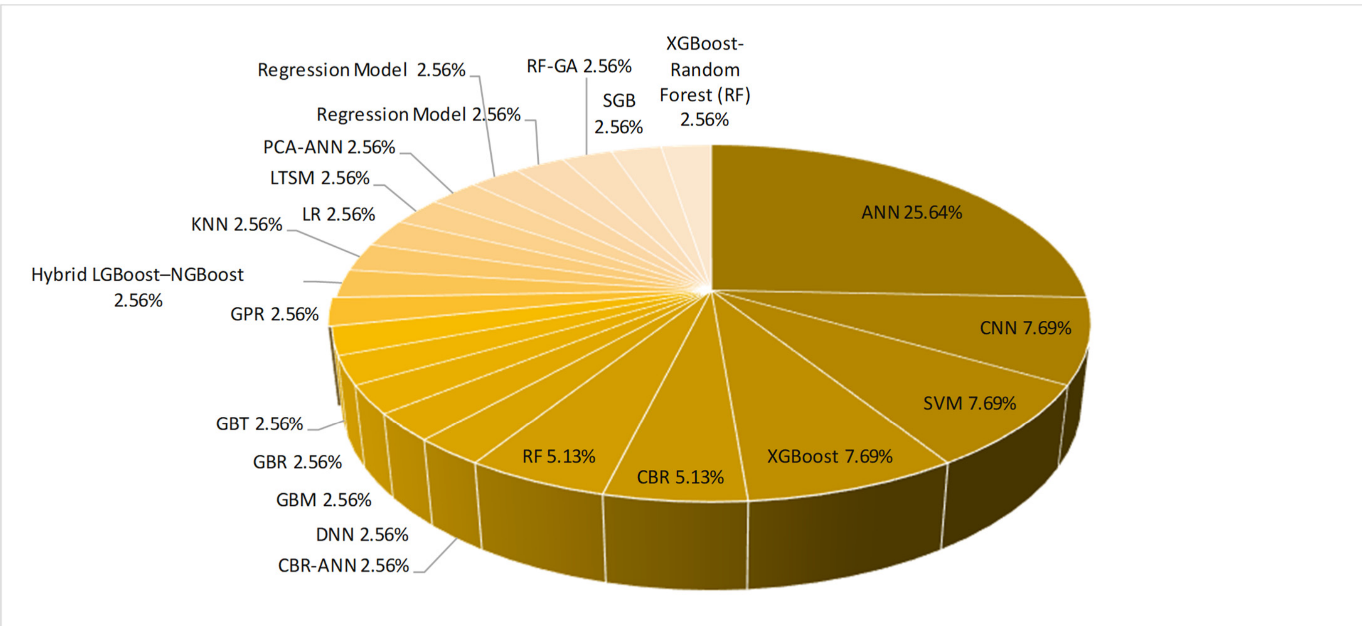

Figure 7. Distribution of AI models used in cost estimation studies.

Hybrid and specialized models such as Hybrid LGBoost–NGBoost and XGBoost–Random Forest (RF) also constitute 2.63%, owing to their specific applications. RF and CBR models account for 5.27% in the aggregate, highlighting moderate applications in the field of cost estimation. These data show a strong predominance toward ANN-based models and a few other lesser-used ones, like SVM and CNN. One can conclude that ANN remains the most prominent model for cost estimation research, albeit with several competencies led by other models.

4.6. Assessment of Confidence in the Findings

The certainty in the body of evidence for the outcomes assessed in this systematic review is moderate–high, with varying degrees based on the study design, risk of bias, consistency, precision, and directness. While most of the studies are experimental and of high quality, some are limited by factors such as specific datasets, lack of real-world validation, and high computational requirements, which reduce the generalizability of the findings. AI models, particularly ANN and XGBoost, consistently outperform traditional methods in sectors like construction and manufacturing, showing high precision and reliability. However, the applicability of these models to other industries, such as healthcare, remains uncertain due to differences in data and complexity. Overall, while the evidence supports the effectiveness of AI in cost estimation, further validation and refinement are needed to increase confidence in the broader applicability of these models across various project types and industries. Figure 8 represents the percentage of models which are used throughout the industries.

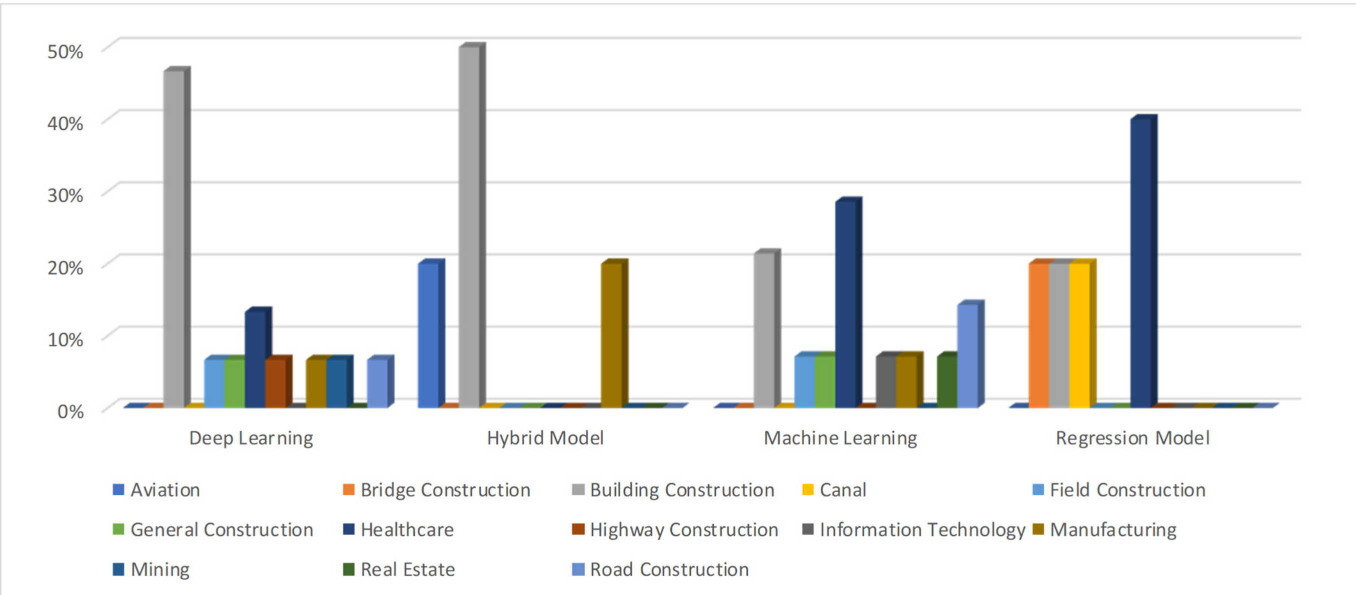

Figure 8. Distribution of AI model applications across various industries.

5. Discussion

5.1. Performance of the Models Industry-Wise

5.1.1. Manufacturing

Ning et al. [35] conducted an experimental study on 121,980 cases from the manufacturing industry, where two models of CNN were compared. They figured out that, with extensive datasets and computational resources, CNNs perform better in cost estimation in manufacturing projects. Case-based reasoning with artificial neural networks [47] has been applied in manufacturing new products to figure out the performance of the model. Using data from 61 industrial luminaire design products in Poland, this hybrid model effectively retrieved past cases to predict costs and reduce the design risk. This model showed great accuracy, at 94.05% on average. Jean-Loup Loyer's study [38] compared various machine learning methods to estimate the manufacturing costs of jet engine components during the early stages of design. Five statistical models—multiple linear regression (MLR), generalized additive models (GAMs), artificial neural networks (ANNs), support vector regression (SVR), and Gradient-Boosted Trees (GBT)—were evaluated using data from Rolls-Royce, with GBT being the most accurate and efficient. The paper notes that the more modern machine learning approach (GBT) would prove all the more effective from a standpoint of accuracy and fit. These models not only help with predicting costs but also provide important engineering insight into the key cost drivers. The final recommendations suggest using a combination of these models for comprehensive cost estimation and point out the need for further integration of these methods within design optimization tools in the aerospace industry.

5.1.2. Road Construction

Road construction costs are relatively high due to the complexity of the materials used, based on the area, along with the fluctuation of prices. A study in Iran [37] found that support vector machine (SVM) performed well. This study picked 300 forest road construction projects, comparing five machine learning models—linear regression (LR), K-Star, Multilayer Perceptron (MLP), support vector machine (SVM), and instance-based learning (IBL); SVM achieved 97.56% accuracy. ANN models have been identified [8] to predict costs accurately, as they have the ability to handle nonlinear relationships better than the traditional method, which is a prime need in cost prediction. This model was tested on 132 road construction projects, where it achieved 86.35% accuracy.

5.1.3. Building Construction Projects

Chakraborty [36] has demonstrated that the hybrid LGBoost–NGBoost model is useful for evaluating the early estimation of costs during building construction projects, where they found 95% accuracy. AI-powered cost estimation has become more transparent and reliable by integrating SHapley Additive exPlanations (SHAP) analysis to mitigate the black-box challenges of the ML-driven cost models. XGBoost–Random Forest is a hybrid cost estimation model for projects with complex input dependencies in the construction sector [44]. Although the accuracy rate is not up to the mark, at only 75%, in comparison to the other models that have been discussed, it could show enhanced performance with large datasets and with more clearance when selecting parameters. Ran wang [45] applied DNNs for a dataset of 98 public school building projects and achieved 87.09% accuracy. He also said that, like ANNs, DNNs are capable of handling economic data and complex linear relationships in the building construction industry. By incorporating economic fluctuations, SVM has enhanced the reliability of cost estimation in residential building construction projects [46]. This study emphasized the importance of integrating economic indicators while implementing AI-driven decision-making in cost management. SVM effectively handles high-dimensional data, with 88–93% accuracy in the cost prediction of residential projects. An experimental study [48] was conducted in Iraq with 40 building construction projects, using another hybrid model, Random Forest with Genetic Algorithm (RF-GA), which achieved 91.67% accuracy in predicting costs and outperformed other RF models in the experiments. It was also added that, by minimizing cost overruns and improving cost scheduling, this model showed significant results. Machine learning algorithms such as Gradient-Boosting Machine were used [54] to derive predictions of property prices using a dataset of approximately 40,000 transactions from 1996 to 2014. This study demonstrated the possibilities for machine learning applications with regard to property valuation and its effects on their computation–real performance trade-off. The ANN [56] is a multioutput regression method, which was used to predict construction costs based on data from Korea's Public Production Service (PPS). Scaling techniques like min–max and Z-score, along with "L1" and "L2" regularization, improve the accuracy and stability of cost estimation. ANNs proves their worth very well in predicting both overall as well as sub-construction costs, thus providing a rich source for budget formulation in the construction industry. The authors of [58] essentially constructed a multilayer feedforward neural network with backpropagation for predicting early-stage building construction costs based on data collected from 78 projects in Mumbai, India. The study incorporated dummies of influential costing parameters and further utilized techniques such as Bayesian regularization, which could enhance prediction accuracy. Thus, the findings talk about the potential of ANNs as a possible tool for cost estimation for better financial planning and decision-making in the construction industry.

The application of ANNs [60] in preliminary accurate cost estimation of composite flooring systems in multi-story houses has been examined by Hosam Elhegazy. An ANN was developed, considering the RSMeans Assemblies Books data for the years 1997–2019 (USA), for predicting flooring system costs with respect to structural variables. Traditional costing procedures are now quite inferior to the ANN model in terms of accuracy. In [67], an artificial neural network (ANN) model was used to foresee the energy costs for Variable Refrigerant Flow (VRF) heating systems based on the operational data from a test building in Seoul, South Korea. The model achieved high accuracy, indicating its ability to optimize energy consumption and reduce heating costs. This model shows great potential for application in smart building management, and future research will deal with real applications and the integration of the system with real-time energy pricing. In a study by Jesam Ujong [70], an ANN approach was tailored to predict construction costs and times using data on building projects within Calabar Municipal, Nigeria. The ANN model showed far better performance when compared with multiple linear regression. Hence, the results showed that the ANN improved the accuracy of the cost and schedule estimation, thus helping the contractors and clients in project planning and decision-making. Future work will be geared towards increasing the size of the dataset and integrating deep learning for further advancement. A study by Alshamrani [62] presented a multiple regression model for predicting the initial construction costs of both conventional and sustainable college buildings in North America. The model drew from RSMeans data and considered factors like the building area, floor height, number of floors, and type of structure (steel or concrete) in estimating the costs. A total of 320 scenarios were tested, and the model was validated with real data to an accuracy of 94.3%.

This makes the model developed very useful for universities to compare the economic advantages of the sustainable buildings against the regular ones. Alshboul [39] investigated machine learning methods for the prediction of the construction costs of green buildings. Three models were compared: XGBOOST, DNN, and RF, and the minimum error metrics were achieved with XGBOOST, along with the maximum accuracy value. The results highlighted the importance of the different features, people, technical aspects, and technological types that influence green building costs. The predictive accuracy of XGBOOST (96%) proves it to be a very strong statistical tool in the hands of green builders, compared to DNN and RF models. This shows that, for certain things, machine learning models can be good at predicting green building costs and lowering uncertainties in green construction.

5.1.4. General Construction

A case-based reasoning (CBR) model [57] was used as a way of predicting building deconstruction costs, with the limitations incurred in traditional demolition cost estimation including both subjectivism and non-individualization of removal scenarios. The model was tested on some real-world cases of deconstruction of some buildings in the state of Michigan, scoring up to 98.8% in accuracy. The results indicate cases where deconstruction can be cheaper than demolition when material recovery is considered. An application of

LSTM neural networks [34] to predict construction cost indices from 143 datasets collected from various Chinese data sources spanning the years 2007 to 2019. As far as long-term dependencies are concerned, LSTM has almost proven to be effective, and it also improves the accuracy to above 99.34% by removing limitations from the existing traditional forecasting methods. This method has great potential in engineering cost estimation and decision-making in project planning and budgeting.

5.1.5. Highway Construction

In highway construction projects, ANNs have performed well in terms of accurate cost estimation [53]. Here, the authors used different ANN architectures with 10, 15, and 20 neurons to evaluate the potential performance of the model. It was found that ANNs have the ability to identify key cost-driving factors that are relevant to the project, along with adaptation to nonlinear patterns, making this the preferred model in highway construction projects. Fourteen highway projects from Brazil were selected to determine the most suitable model.

5.1.6. Canals

Elmousalami [21] compared 20 AI techniques for 140 FCIPs in building construction. XGBoost, due to its high accuracy, scalability, and ability to handle missing data and complex patterns without overfitting, was best suited for this study. This study compared XGBoost with artificial neural networks (ANNs), decision trees, support vector machines (SVMs), fuzzy logic, genetic–fuzzy models, case-based reasoning (CBR), and XGBoost successfully estimated the costs with 90% accuracy. A quadratic regression model [64] was developed for the cost estimation of field canal improvement projects (FCIPs) using 144 Egyptian records (2010–2015). The model surpassed standard regression techniques, boasting an R^2 of 0.86 and a mean absolute percentage error of 7.82%, thereby being considered a reliable tool for predicting costs at the irrigation projects' initiation stage. It outperformed linear, semi-log, and artificial neural network models in reflecting nonlinear cost relationships, thus enhancing the accuracy and clarity of decision-making. The authors plan to extend this work by working on a larger database and further hybridizing with artificial neural networks under the regression model.

5.1.7. Field Construction

Case-based reasoning (CBR) [41] was explored in field construction projects in 143 fields in Poland, integrating sustainability factors. CBR achieved an accuracy of 86% in predicting costs based on historical data and adapting new cases. On the other hand, Juszczyk [42] performed a study on a similar dataset but with 115 field projects, and he identified ANNs' ability to handle complex cost relationships and improve early-stage cost prediction along with cost overruns. This did not require predefined equations or a developed correlation between cost predictors and final construction costs, with 85% accuracy in estimation.

5.1.8. Mining

Estimating capital costs in open-pit mining projects with 74 pits, ANNs [43] outperformed traditional models. ANNs can leverage their multilayer structure to adapt to varying mining conditions to predict costs, along with maintaining complex, nonlinear cost dependencies in mining.

5.1.9. Healthcare

Mazumdar [50] found that RF models performed well against generalized linear models and Partially Linear Additive Quantile Regression (PLAQR) models in predicting healthcare costs. RF demonstrated superior performance by detecting nonlinearity and interactions without pre-specifying the model parameters. Stochastic Gradient Boosting (SGB) is another machine learning model that has been used in cost prediction [51]. By evaluating multiple machine learning projects with a Kaggle dataset, SGM performed better than linear regression, ridge regression, XGBoost, and Random Forest models. SGB has potential in insurance calculations, fraud detection, and risk assessment when calculating budgets in the healthcare industry. There is potential risk and bias in cost modeling while designing and implementing any project in the real world. CNNs [52] remove this sort of uncertainty by balancing better estimation reliability with its handling of different cost parameters in the healthcare industry. CNNs perform well in handling nonlinearity and interactions in healthcare cost modeling. For hospital readmissions, using Electronic Medical Records (EMRs) from Barnes-Jewish Hospital with a DNN model worked better to estimate costs. This deep learning technique is able to handle imbalanced class distribution and improve predictive power. DNNs [45] perform far better than their conventional counterparts by minimizing the costly errors of misclassification and enabling early action within the hospital, and this method has already been deployed for real-time clinical decision support. In his research, Mohammad Amin Morid [66] developed the Gradient-Boosting Model (GBM) for predicting healthcare costs using 24,000 patient records, 3.8 million medical claims, and 780,000 pharmacy claims from the University of Utah Health Plans (2013–2016). The model itself yielded a high degree of accuracy, which indicated that it could assist health insurers and providers in optimizing resource allocation and models based on value-based payments, while future work would involve exploring several deep learning techniques for further enhancement. This study found XGBoost regression [68] the be the most suitable, as it achieved greater accuracy with respect to ridge, lasso, and KNN, while also accommodating nonlinear relationships quite well in health insurance cost estimations. One noteworthy feature of XGBoost regression is its strong performance-boosting technique, which contributes to a high level of confidence in predictive analytics for insurance pricing and risk estimation. The study of Peško [32] looked into employing AI-driven predictive modeling for the total project cost and duration estimation within the construction industry, using data from 166 projects. This model surpassed traditional regression-based methods, thus providing better cost estimation and decision-making. The results indicate the role of AI in enhancing planning in construction; future studies will be directed toward addressing the validation of the model for different types of projects and the integration of more sophisticated machine learning techniques. In [69], based on the patient records of Tsuyama Chuo Hospital for 24,353 patients, a linear regression model was developed solely to predict the healthcare costs attributable to obesity. The model was found to yield the best results in estimating medical expenses, as indicated by an accuracy of 97.89%. The findings state that BMI and other correlates of obesity play an important role in predicting healthcare costs, thus providing insight into financial planning and risk assessment for the insurance industry and policymakers.

5.1.10. Software

Abdulmajeed [61], using data for software projects, described an artificial intelligence technique to predict the cost of software projects. They used an extensive dataset developed by NASA, which included 60 software projects. Cost estimations using three machine learning models—KNN, CNN, and ENN—were checked and compared for their accuracy. The KNN model gave the highest prediction accuracy, at 90.24%. The findings indicate that machine learning techniques are a plausible alternative in software cost estimation, which will significantly aid in budgeting and resource allocation.

5.1.11. Aviation

In this work [63], Xiaonan Chen set out to develop a PCA-ANN (principal component analysis + artificial neural network) model for the cost estimation of general aviation aircraft from 22 literature examples. The model outperformed the multiple linear regression (MLR), stepwise regression (SR), and ANN-only models, with considerably high accuracy values (R = 0.9994, MAPE = 0.009–0.015, RMSE = 1.667–3.416). The results demonstrated how PCA-ANN can be effectively used to reduce dimensions and improve cost prediction, thus aiding aviation manufacturers in pricing and financial planning. Future work could investigate the expansion of the dataset and hybrid regression–ANN models.

5.1.12. Bridge Construction

Kovačević's [65] study focused on Gaussian Process Regression (GPR) with the ARD–exponential covariance function for estimating the construction costs of 181 reinforced and prestressed concrete bridges located throughout Serbia's Pan-European Transport Network from 2009 to 2014. The model proved superior to ANN, regression trees, and SVR in terms of accuracy (R = 0.89, MAPE = 11.60%), while also providing uncertainty quantification and reliability. Therefore, the findings support objective cost forecasting, minimizing human bias in budgeting, while future investigations will presumably be geared more toward the expansion of the datasets in terms of volume, and there is an inclination to approach cost estimation as a classification problem. Table 2 explains the real-life implications of the best-suited model.

Table 2. Comparison of preferred and best-suited models, along with their real-life implications.

Model	Preferred Technique	Best-Suited Model	Real-Life Implications
Deep Learning Techniques	Convolutional neural network (CNN) [27,52,55]	CNNs learn complex relationships and features from large datasets, along with handling imbalanced class distribution while predicting costs	CNNs have automated price quotations and reduce dependency on expert-driven cost estimation
	Artificial neural network (ANN) [8,20,42,53,56,58–60,67,70]	ANNs were chosen for their ability to handle nonlinear relationships and build complex relationships between cost predictors and final costs, without requiring predefined equations	Enhances capital planning accuracy, reduces cost overruns, and improves investment decision-making by leveraging historical data for systematic cost estimation and reducing dependence on expert judgment
	Deep neural network (DNN) [45]	DNNs are recommended for projects where economic factors and project characteristics interact in a complex manner	Used for predicting early-stage construction costs of public school projects, assisting stakeholders in financial decision-making
	Long Short-Term Memory (LSTM) [34]	It effectively captures long-term dependencies in time-series data, overcoming issues	Improved accuracy in project cost estimation and planning

Table 2. *Cont.*

Model	Preferred Technique	Best-Suited Model	Real-Life Implications
Machine Learning Technique	Support vector machine (SVM) [32,37,46]	SVM handles high-dimensional data and complex terrain-dependent costs between construction costs and economic variables	Enables early cost prediction for road construction projects in forestry; supports bid selection processes and alternative road alignment choices
	Random Forest (RF) [49,50]	RF showed superior performance by detecting nonlinearity and interactions without pre-specifying the model parameters	Enhances predictive accuracy in healthcare cost estimation, reducing bias in provider reimbursement under value-based care models
	Gradient-Boosted Trees (GBT) [38]	GBT was chosen for its superior predictive accuracy and ability to handle complex data structures	Used for early-stage cost estimation for perfect manufacturing decisions
	Extreme Gradient Boosting (XGBoost) [21,39,68]	XGBoost is best suited due to its scalability and ability to handle missing data, along with the best trade-off between accuracy and computational efficiency	XGBoost can predict conceptual costs at early project stages, aiding financial planning and decision-making
	Hybrid LGBoost–NGBoost [36]	Hybrid LGBoost–NGBoost model has the ability to estimate uncertainty by automating cost forecasting at a large scale	Enables construction professionals to perform value engineering (VE) efficiently by evaluating different design alternatives based on accurate cost predictions
	Stochastic Gradient Boosting (SGB) [51]	Enables construction professionals to perform value engineering efficiently by evaluating different design alternatives based on accurate cost predictions	Enhances pricing accuracy and improves cost management for selecting better plans
	Categorical Boosting (CBR) [57]	CBR can learn from past cases, requiring less training, while offering better clarity and interpretability in cost estimation	Supports decision-makers in shifting from demolition to sustainable deconstruction by providing accurate cost estimates and potential material recovery values
	K-Nearest Neighbors (KNN) [61]	KNN effectively learns cost patterns from historical project data, making it a reliable model for estimating costs without requiring complex training processes	Provides accurate cost estimation, reducing financial risks and optimizing project planning
	Gradient-Boosting Machine (GBM) [54]	GBM can minimize prediction errors, making it the most effective model for property price estimation	Provides accurate property price predictions, aiding real estate investors and policymakers

Table 2. *Cont.*

Model	Preferred Technique	Best-Suited Model	Real-Life Implications
Regression-Based Models	Regression model [62,64]	This model can handle large, nonlinear projects with complex dependencies	Automates and optimizes cost estimation processes; improves projects with early-stage financial planning
	Logistic regression (LR) [69]	Linear regression is best for its simplicity and high interpretability	Policymakers are benefitted by this model through cost forecasting
	Generalized Boosting Regression (GBR) [66]	Gradient boosting can capture complex temporal dependencies in cost data	Managing resources efficiently and optimizing value-based payment models
	Gaussian Process Regression (GPR) [65]	Handling the complex, nonlinear relationships in cost estimation	Helps allocate care management resources efficiently and optimize value-based payment models
Hybrid and Optimization-Based Models	XGBoost–Random Forest (RF) [44]	This model performs by balancing complex input dependencies while estimating project costs	Enhances cost estimation accuracy, reduces errors, and improves decision-making in early project planning
	CBR-ANN [47]	CBR-ANN effectively selects relevant past cases, improving cost prediction accuracy	Improves cost estimation accuracy, reduces design uncertainty, and aids decision-making in new product development projects
	Random Forest–Genetic Algorithm (RF-GA) [48]	RF-GA optimizes Random Forest hyperparameters using Genetic Algorithms, enhancing learning efficiency with complex, nonlinear delay dependencies	Improves delay prediction accuracy, enables better risk management, and enhances project planning
	Principal component analysis–ANN (PCA-ANN) [63]	PCA-ANN reduces dimensionality while retaining key information, improving ANNs' performance in cost prediction	Provides more accurate and practical cost estimations, enabling better pricing and financial planning

5.2. Performance of the Models in Different Industries

5.2.1. Deep Learning Models

This particular study looks into the efficiency of different neural network models for cost estimation in various industries and presents the advantages of each model for different types of projects. CNNs ([35,52]) perform exceptionally in the manufacturing and healthcare industries because of their reliance on big datasets and the recognition of complex nonlinear systems. In other words, CNNs have been performing very well in estimating manufacturing costs, mainly because of the large datasets and computer resources used. ANNs [8,42] have been widely implemented in road construction, general construction, and mining, where their ability to capture nonlinear relationships and improve some early cost predictions is regarded as quite valuable. For instance, the ANN gave an accuracy rate of 86% in predicting costs in road construction, while in general construction it helped predict cost overruns and early-stage costs, with an accuracy of 85%. For projects involving complex data relationships, such as public school buildings and hospital readmissions, DNNs [45] give better predictive power than conventional models, and with fewer errors. Long Short-Term Memory [34] provides for long-term dependency and, thus, proves to be a great model for predicting the construction cost index with high accuracy for longer periods of time compared to other conventional methods. ANNs have been successfully applied in various fields: predicting composite flooring systems' costs [60], the energy costs of heating systems in smart buildings [67], and time and costs for construction jobs in Nigeria [70], with almost perfect accuracy (99.9995%). Thus, it can be concluded that ANNs are the most flexible model, providing sound and accurate predictions for costing in different industries, while CNNs, DNNs, and LSTM are specialized to cater to certain advantages in specific areas, such as healthcare, public infrastructure, and long-term forecasting.

5.2.2. Machine Learning Models

Across various industries, different types of machine learning model have thus flourished in their application in cost prediction and decision-making. In the area of building construction, XGBoost has been said to have provided utmost accuracy in cost prediction for projects, with an accuracy of about 90%, mainly due to its capabilities of handling missing data and complex patterns [40]. Case-based reasoning (CBR) has been effectively applied in cost estimation for construction projects in the field, obtaining an accuracy of 86% while factoring in sustainability [41]. Support vector machines (SVMs) were used on residential building projects, and an accuracy of 88–93% was obtained by these models, accounting for economic fluctuations [46]. In the health sector, Random Forest (RF) was found to be better in cost prediction, estimating nonlinearity without the need for pre-specifying the model parameters [50]; moreover, Stochastic Gradient Boosting (SGB) appeared to be successfully used in fraud detection and risk assessment in insurance [51]. Support vector machine, Random Forest, and Gradient-Boosting Machine models were used in the price prediction of real estate properties, balancing the trade-off between computational efficiency and real-world performance [54]. In the deconstruction industry, CBR achieved 98.8% predictive accuracy in the cost estimation of building deconstruction, thus providing less expensive alternatives to traditional opposition [57]. K-Nearest Neighbors (KNN) provided an accuracy of 90.24% in the cost estimation of software development projects, thus providing a workable alternative to conventional methods [61]. XGBoost regression was found to be the most suitable model for estimating health insurance costs, with an R2 of 86.81 and RMSE of 4450.4, efficiently capturing the nonlinear association [68]. Artificial intelligence-powered predictive modeling, with XGBoost, was noted to outperform traditional methods for estimating total project costs and durations, enhancing their respective capabilities in construction project estimation [32]. Finally, in the construction of green buildings, XGBoost was found to be the most accurate in prediction (96%) in comparison with deep neural networks (DNNs) and Random Forest (RF) models, after multiple considerations of factors like people, technology, and technical aspects [39]. Such studies confirm the versatility and accuracy of machine learning models across various industrial sectors, with XGBoost standing tall as the best performer in construction, healthcare, and green building cost prediction.

5.2.3. Regression Models

The different applications of machine learning-generated models for cost estimation across different industries have been highly efficacious. For instance, in Egypt, irrigation projects were highly predicted using quadratic regression, with an R^2 value of 0.86 and an MAPE value equal to 7.82%, which would be higher than those of any traditional regression method [64]. GPR with the ARD–exponential covariance function had an R^2 of 0.89 and MAPE of 11.60%, surpassing other models in estimating the costs of concrete bridges for the transport network in Serbia, while also accounting for uncertainty quantification [65]. The Gradient-Boosting Model (GBM) predicted healthcare costs in the U.S. quite impressively, achieving an MAPE of 2.04%, substantially better than other models such as linear regression and Random Forest [66]. In the case of predicting healthcare costs for obesity, linear regression achieved a fantastic 97.89% accuracy, which also stressed BMI's importance in prediction [69]. Reliable multiple regression testing showed the capacity to estimate initial construction costs for sustainable buildings in North America, with an accuracy of 94.3%, offering an opportunity for universities to evaluate the economic feasibility of green projects [62]. Last but not least, Gradient-Boosted Trees (GBT) proved the best in estimating the manufacturing costs of components of jet engines in the aerospace industry, overcoming traditional models like MLR and ANNs [38]. The above discussion articulates the fact that such models can demonstrate the tremendous and compelling role that machine learning plays in enhancing the accuracy of cost predictions across all sectors.

5.2.4. Hybrid Models

Hybrid machine learning models are used for cost predictions across various industries due to their complex data handling abilities. The hybrid LGBoost–NGBoost model showed 95% accuracy in performing early cost estimation for building construction projects [36], demonstrating the possibility of accurate and reasonable cost predictions. XGBoost–Random Forest (RF), on the other hand, showed an observed accuracy of 75%, while still being heavily applied in projects with complex input dependence within the construction sub-sector [44]. Case-based reasoning (CBR) with artificial neural networks (ANNs) have shown amazing success in manufacturing for cost prediction, with accuracy of 94.05% using data from industrial luminaire designs in Poland [47]. The Random Forest–Genetic Algorithm (RF-GA) hybrid model revealed a high predictive capability, with an accuracy of 91.67% for cost predictions and schedule optimization on construction projects in Iraq, signifying its ability to curtail costs [48]. In aviation, the PCA-ANN model outperformed regression models, with high accuracy for the cost estimation of general aviation planes (R = 0.9994) [63]. The models showed strong performances in the industries of construction, manufacturing, and aviation, thus marking a significant advancement in cost prediction and management. Table 3 summarizes the results of the preferred models which are used in the different studies.

The risk-of-bias (RoB) assessment of the 39 studies revealed a diverse range of challenges that affect the reliability of their findings. Studies with moderate risk commonly face challenges like high computational requirements, data imbalance, and issues related to external factors (e.g., labor costs, market fluctuations, political instability). Additionally, many studies are limited by a lack of external validation, data variability, or uncertainty handling in the models, which could reduce their generalizability to real-world scenarios. Several studies also experience difficulties with model interpretability and the high complexity of data, which may hinder the effectiveness of these models in diverse conditions.

Table 3. Results of the cost prediction models.

Study	Preferred Technique	Model	Evaluation Metrics	Reported Accuracy Performance	Proposed Future Directions/Recommendations
[35]	CNN	Deep learning techniques	the MAPE value was estimated at 10.02%; with 400,000 cases it reached 6.34% while predicting the cost	89.98%	Optimize CNN architectures for better computational efficiency
[8]	ANN		Engineering costs were accurately estimated through the ANN, with an MAPE of 14.5%.	86.35%	Conduct real-world validation and pilot testing in active engineering consultancy projects
[42]	ANN		MAPE errors were expected to be smaller than 15%	85%	Aim to expand dataset size, refine input variable selection, and explore deep learning approaches
[43]	ANN		The ANN model achieved RMSE = 138.103, R^2 = 0.990, MAE = 114.589, and APE = 7.770%	92.23%.	Integrate external economic factors, expand dataset size, and refine ANN architectures
[45]	DNN		DNN model achieved R^2 = 0.95 and MAPE = 12.91%	87.09%	Refine input variables, integrate real-time economic data, and test alternative deep learning models
[52]	CNN		The proposed CNN model achieved the best predictive performance, with an MAPE of 1.67%	94.53%	Focus on optimizing CNNs for cost modeling and exploring hybrid models
[53]	ANN		Mean absolute percentage error (MAPE) of 5.84%	99%	Expand datasets and refine ANN architectures for better generalization
[55]	CNN		The area under the curve (AUC) for the proposed model was 0.70	92%,	Improving interpretability of deep learning-based predictions and enhancing cost-sensitive learning for different clinical applications
[56]	ANN		R^2 score for sum cost: 0.80; R^2 score for total cost: 0.65	81% to 89%	Improving machine learning model stability and accuracy through better data preprocessing techniques
[58]	ANN		The best ANN model (11-3-1 architecture with Bayesian regularization) achieved R^2 = 0.9922 and RMSE = 0.02469	99.22%.	Development of ANN-based cost estimation tools for real-world application
[59]	ANN		R^2 is likely between 0.98 and 1.00, and MAE $\approx$ 0.1 to 1 (depending on the scale)	99%	Enhancing the framework for broader applications
[60]	ANN		The mean squared error (MSE) for optimal ANN models ranged between 0.0026 and 0.09, confirming the model's reliability	99.9%	Using real project cost data instead of relying solely on RSMeans
[70]	ANN		This model showed mean absolute error (MAE): 0.2952 and root-mean-square error (RMSE): 0.5638	99.9995%	Further exploration of nonlinear and complex factors affecting cost estimation

Table 3. *Cont.*

Study	Preferred Technique	Model	Evaluation Metrics	Reported Accuracy Performance	Proposed Future Directions/Recommendations
[67]	ANN	Deep learning techniques	ANN model structured with one hidden layer and 15 neurons achieved an R^2 of 0.8417	84.17%	Implementing ANN models into real-world VRF heating control systems
[34]	LTSM		The LSTM model predicted engineering cost indices, achieving MAE: 0.96, MSE: 1.03, and MAPE: 0.71	of 99.29%.	Further optimization of LSTM models for broader applications
[37]	SVM	Machine learning techniques	Support vector machine (SVM) cost prediction, with an accuracy of R = 0.993 and RMSE = 2.44%.	97.56%.	Apply more machine learning methods and metaheuristic optimization to improve cost estimation accuracy
[46]	SVM		Significant error reduction compared to BPNN and SVM	88% to 93%.	Future research should explore larger datasets, integrate more economic indicators, and test alternative machine learning architectures
[32]	SVM		SVM achieved an MAPE of 7.06% for cost estimation	92.94%	Further validation on diverse project types; integration of additional machine learning techniques
[49]	RF		The RF model achieved the best predictive performance, with an adjusted R^2 of 47.5%, mean absolute error (MAE) of EUR 1338, and hit ratio (HiR) of 67%	67%	Use real item data to increase the accuracy of the model
[50]	RF		In terms of MAPE, RF ranged from 12% to	88.48%	Future research should explore improving risk adjustment models and assessing RF's computational efficiency for real-world applications
[68]	XGBoost		XGBoost achieved R^2 = 86.81 and RMSE = 4450.4	86.81%	Incorporating deep learning and metaheuristic approaches
[39]	XGBoost		XGBOOST achieved the highest accuracy, with R^2 = 0.96 and MAPE = 19.9%	80.1%	Future work should incorporate a broader dataset, more attributes, various building types, and an expanded lifecycle cost analysis
[40]	XGBoost		XGBOOST achieved the best accuracy: MAPE = 9.091%,	90.909%	Recommendation to develop hybrid models incorporating fuzzy logic to handle uncertainties, improve rule generation for fuzzy systems, and explore deep learning for cost prediction
[41]	CBR		It generated a mean absolute estimate error of 14%.	86%	Improve case similarity calculations and integrate sustainability metrics into other estimation models

Table 3. *Cont.*

Study	Preferred Technique	Model	Evaluation Metrics	Reported Accuracy Performance	Proposed Future Directions/Recommendations
[51]	SGB	Machine learning techniques	The SGB model achieved the best performance, with an RMSE of 0.340 and cross-validation score = 0.858	86%	Use nature-inspired algorithms and deep learning models for cost prediction
[54]	GBM		This model had the best R^2 score of 0.90365 for GBM	90.37%	Incorporating additional property transaction data from a larger geographical region
[57]	CBR		Paired t-tests showed a high correlation (0.999, p = 0.03), confirming the model's statistical significance at a 95% confidence level	98.8%	Investigating the deconstruction supply chain and market development
[61]	KNN		The KNN model had the lowest MMRE (0.101), RMSE (0.547), and BRE (0.205), confirming its superior cost prediction accuracy	90.24%	Expanding the dataset with real-world software project data
[62]	Regression Model	Regression models	MAPE was evaluated at 5.7%	94.3%	Incorporating real-time cost data beyond RSMeans
[64]	Regression Model		Quadratic regression model: R^2 = 0.86, MAPE = 9.12% (training), and MAPE = 7.82%	86%	Exploring hybrid ANN–regression models
[65]	GPR		Best model (GPR with ARD–exponential): R = 0.89 and MAPE = 11.60%	88.4%	Exploring cost estimation as a classification problem instead of regression
[66]	GBR		Improves healthcare cost prediction accuracy, with Gradient Boosting achieving a MAPE of 2.04%	97.96%	Evaluating deep learning models for cost prediction
[69]	LR		The MAPE value was 2.11%, and R^2 =0.9789	97.89%	Integration of deep learning techniques for enhanced prediction
[63]	PCA-ANN	Hybrid and optimization-based models	During the test, the MAPE values were 0.009 (training) and 0.015 (testing)	99%	Expanding the dataset for improved accuracy
[47]	CBR-ANN		The CBR-ANN model (ANN-PA), with an MAPE value of 5.95%, achieved the best accuracy, demonstrating its effectiveness in improving cost estimation	94.05%	Refine attribute selection, expand the case base, and test deep learning models
[48]	RF-GA		RF-GA model showed a classification error = 8.33%	91.67%,	Refine delay factor classification, integrate real-time data, and enhance optimization techniques
[44]	XGBoost–Random Forest (RF)		The XGBoost-RF model achieved R^2 = 0.87 and MAPE = 0.25	75%	Refine input selection, integrate deep learning, and explore economic impacts on cost estimation

On the other hand, studies with high risk typically involve subjectivity in the assessment process, variability in training data, and a lack of real-world validation. These challenges can significantly compromise the study's external validity and generalizability. Issues like multi-collinearity, optimization constraints, and training data limitations also contribute to the higher risk in some studies. While a few studies manage to offer solutions in their future directions (such as expanding the datasets or optimizing the algorithms), the overall risk suggests that further validation and model adjustments are needed to enhance the robustness and applicability of the findings in broader contexts.

5.3. Limitations of This Study

One limitation of this manuscript is its reliance on existing studies from a limited number of industries, which may not fully represent the potential of AI-driven cost estimation across all sectors. The data used in many of the studies are specific to particular regions or project types, potentially limiting the generalizability of the findings. Furthermore, while this review covers a range of machine learning and AI techniques, the models included often face challenges related to data quality and size. Many studies rely on small or highly specific datasets that may not capture the full variability found in larger or more diverse projects. Additionally, there are concerns regarding the external validity of the models, as several of the studies have not been externally validated in real-world settings, raising questions about the robustness of these AI-driven approaches when applied outside controlled environments. Lastly, many of the AI models reviewed here lack interpretability, with complex deep learning techniques being difficult to explain to non-expert stakeholders, which could hinder their acceptance in practical applications.

5.4. Future Directions

Future research could focus on expanding the datasets used in AI-driven cost estimation, particularly by incorporating data from a wider range of industries, regions, and project types to improve the models' generalizability. Incorporating external factors such as economic trends, labor costs, and material price fluctuations into the models could enhance their adaptability and accuracy in dynamic environments. Additionally, validating the AI models in real-world projects across various industries would help assess their practicality and ensure that their findings are robust and applicable in diverse contexts. Another area for future development lies in improving the interpretability of AI models. Techniques like SHAP analysis or other explainable AI methods could be incorporated to make the models more transparent and acceptable to stakeholders. Finally, the potential for hybrid models that combine machine learning techniques with traditional cost estimation methods should be explored further to address the limitations of both approaches and improve the overall accuracy and reliability of cost predictions.

Table 4 shows the risk- bias assessments of the study.

Table 4. Risk-of-bias assessment of the studies, based on study limitations.

Model Domain	Study	Bias Level	Reason for Assessment
Deep Learning Techniques	[35]	Moderate	High computational requirements for 3D CNN training; lack of processing requirement data (e.g., surface roughness, machining precision)
	[8]	Moderate	The model has not been externally validated on new projects beyond the training dataset
	[42]	Moderate	High variance in project specifications, and sensitivity to input variable selection
	[43]	High	Variability in mining conditions, and lack of supply chain influence
	[45]	Moderate	Difficulty in incorporating real-time economic trends, and project complexity
	[52]	Low	CNNs require higher computational time; however, they improve cost prediction accuracy
	[53]	Moderate	Potential overfitting with increasing neuron counts
	[55]	Low	Imbalanced class distribution of readmission cases, need for higher sensitivity in predicting rare but costly misclassifications, and complexity of heterogeneous data integration
	[56]	Low	High data deviation led to unstable learning performance in some cases
	[34]	Moderate	Noise, anomalies, missing information, and data frequency differences
	[58]	Moderate	Neural networks require high computational resources
	[59]	High	Optimization constraints and data limitations
	[60]	Low	Sensitivity to variations in cost estimation due to local unit cost differences
	[67]	Moderate	Potential variability in real-world heating conditions; requires extensive data for model training
	[70]	Moderate	Exclusion of bidding, tender negotiation, supply chain, and safety constraints
Machine Learning Models	[37]	Moderate	The model is limited to low-volume forest roads and may not be generalizable to highway or urban road construction
	[39]	Moderate	Lack of external support considerations; exclusion of social and environmental factors
	[40]	Moderate	Handling uncertainty, high computational complexity, and model interpretability issues
	[41]	High	Subjectivity in similarity assessment; lack of consideration for external economic factors
	[46]	Moderate	Data availability constraints; economic fluctuations affecting construction costs
	[49]	High	Potential difficulty in generalizing results to larger populations
	[50]	Low	RF is computationally expensive; all models underpredicted high-cost cases, leading to bias against high-cost hospitals
	[51]	High	Model training time, and the impact of missing data
	[54]	Moderate	GBM and RF models showed high bias for extreme property values
	[57]	High	Lack of a well-established supply chain for salvaged materials, and the need for greater adoption of deconstruction methods
	[61]	High	KNN requires optimal K-value selection for optimal performance
	[68]	Moderate	Potential difficulty in generalizing results to larger populations
	[32]	Moderate	Normalization process required to ensure uniformity in data; limited dataset may impact model generalization

Table 4. *Cont.*

Model Domain	Study	Bias Level	Reason for Assessment
Regression Model	[38]	High	Data imbalance affected model performance across different component categories
	[62]	Moderate	RSMeans cost data may not fully reflect local variations
	[64]	Moderate	Multi-collinearity in cost data affecting regression models
	[65]	Moderate	Complexity of infrastructure cost estimation due to multiple influencing factors
	[66]	Moderate	Medical predictors did not significantly improve accuracy; fine-grained features increase dimensionality
	[69]	Moderate	Medical predictors had a minimal impact on forecast accuracy
Hybrid and Optimization-Based Models	[44]	Moderate	Data variability; political and economic instability affecting construction projects
	[47]	High	Limited case base for training, sensitivity to attribute weighting, and need for expert validation
	[48]	Moderate	External factors affecting delays, and sensitivity to parameter selection
	[63]	High	Multi-collinearity in cost data impacts traditional regression models

6. Conclusions

This systematic review underscores the transformative impact of artificial intelligence (AI) in the field of cost estimation for project management. By synthesizing insights from the studies, it is evident that AI-driven models, particularly those utilizing machine learning (ML), deep learning (DL), and hybrid approaches, have performed significantly well for cost estimation methods. This study highlights the pivotal role of advanced algorithms, such as artificial neural networks (ANNs), support vector machines (SVMs), and XGBoost, in improving predictive accuracy and adaptability to dynamic project environments. These AI models offer a far more robust solution for managing the complex variables in modern projects compared to conventional methods, which were often limited by human error and lack of scalability.

The findings indicated in this systematic review underscore the transformative role of AI-driven cost estimation models, particularly focusing on machine learning (ML), deep learning (DL), regression models, and hybrid models, all of which significantly outperform traditional methods in terms of accuracy and efficiency. Deep learning models like convolutional neural networks (CNNs), artificial neural networks (ANNs), deep neural networks (DNNs), and Long Short-Term Memory (LSTM) show exceptional performance, with ANNs achieving up to 85–99% accuracy across various industries. In machine learning, techniques such as support vector machines (SVMs), Random Forest (RF), Gradient-Boosted Trees (GBT), and Extreme Gradient Boosting (XGBoost)—especially XGBoost—consistently deliver high accuracy (up to 90%) in cost predictions for sectors like construction and healthcare. Regression models such as linear regression (LR), Generalized Boosting Regression (GBR), and Gaussian Process Regression (GPR) offer reliable results, with GPR achieving an R^2 of 0.89 in construction cost estimation. Furthermore, hybrid models like XGBoost–Random Forest (RF) and ANN–case-based reasoning (CBR) significantly enhance predictive performance, with CBR-ANN reaching 94.05% accuracy in manufacturing. Overall, these AI models provide remarkable improvements in cost estimation by reducing errors and increasing the reliability of predictions, with future research focusing on refining these models, expanding the datasets, and integrating economic variables to improve their real-world applicability across industries. The results underline a clear trend towards the adoption of AI methods, with ANNs remaining the most widely used technique. This review also identified challenges in model performance, particularly with data variability, the need for high-quality datasets, and high computational costs. To maximize the potential of AI in cost estimation, further research should focus on refining model accuracy, expanding real-world validation, and enhancing the integration of external factors such as market fluctuations.

The limitations of the current body of research are primarily related to the generalizability of the findings across diverse industries and real-world project types. Many studies are constrained by region-specific data, small sample sizes, and a lack of external validation, which reduces the applicability of their results in broader, more varied contexts. Furthermore, the reliance on datasets that may not fully capture the complexity and variability of real-world projects calls for more extensive and diverse data in future research. There is also a significant gap in addressing the interpretability of AI models, as project stakeholders often require clear, understandable insights from these systems to foster trust and make informed decisions.

Looking forward, there is a critical need for future research to expand the datasets used in AI-based cost estimation, incorporating more diverse industries, regions, and project types. Such an approach would increase the robustness and generalizability of these models. Additionally, integrating external economic factors—such as market fluctuations, labor costs, and material price variations—into AI models would improve their adaptability to real-world scenarios. Real-world validation of these models is essential to ensure their practicality, and this should be a key focus in future studies. Another promising direction is the development of more interpretable AI models. Techniques such as SHAP (SHapley Additive exPlanations) and other explainable AI methods can help bridge the gap between technical accuracy and stakeholder trust, making AI models more accessible and useful in decision-making processes.

AI has demonstrated considerable promise in revolutionizing cost estimation, and overcoming the challenges highlighted in this review will be key to its wider adoption. By addressing these technical limitations and enhancing model transparency, AI can play an increasingly central role in transforming project management practices, making cost estimation more accurate, efficient, and adaptable in the face of ever-evolving project dynamics.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/modelling6020035/s1.

Author Contributions: M.M.I.S. (first author, corresponding author): Conceived the research idea, led the methodology design, conducted the systematic review, and wrote the majority of the paper. He is the corresponding author for all communications regarding this paper. A.B.b.A.H.: Contributed to the theoretical framework, provided critical insights into the methodology, and reviewed and edited the final manuscript. T.E.N.: Assisted with the methodology development and analysis of findings, contributed to the interpretation of results, and provided feedback on the manuscript. N.S.B.R.: Provided assistance in the literature review and analysis, contributed to the writing and editing of certain sections of the manuscript, and assisted with formatting and referencing. All authors have read and agreed to the published version of the manuscript.

Funding: There are no external funding sources supporting the research and development of this paper.

Data Availability Statement: Dataset available on request from the authors.

Conflicts of Interest: N.S.B.R. was employed by the company All Academic Research Journal. The remaining authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial neural network
CBR	Case-based reasoning
CNN	Convolutional neural network
DNN	Depp neural network
GBM	Gradient-Boosting Machine
GPR	Gaussian Process Regression
GBT	Gradient-Boosted Trees
IBL	Instance-based learning
KNN	K-Nearest Neighbors
LGBoost	Light Gradient Boosting,
NGBoost	Natural Gradient Boosting
LTSM	Long Short-Term Memory
LR	Linear regression
PCA	Principal component analysis
RF	Random Forest
RME	Relative mean error
RMSE	Root-mean-square error
RNN	Recurrent neural network
SGB	Stochastic Gradient Boosting
SVM	Support vector machine
XGBoost	Extreme Gradient Boosting

References

1. Sargentis, G.-F.; Defteraios, P.; Lagaros, N.D.; Mamassis, N. Values and Costs in History: A Case Study on Estimating the Cost of Hadrianic Aqueduct's Construction. *World* **2022**, *3*, 260–286. [CrossRef]
2. Walker, D.; Dart, C.J. Frontinus—A Project Manager from the Roman Empire Era. *Proj. Manag. J.* **2011**, *42*, 4–16. [CrossRef]
3. Weave, P. The Origins of Modern Project Management. In Proceedings of the Fourth Annual PMI College of Scheduling Conference, Vancouver, BC, Canada, 15–18 April 2007; pp. 15–18.
4. Seymour, T.; Hussein, S. The history of project management. *Int. J. Manag. Inf. Syst.* **2014**, *18*, 233–240. [CrossRef]
5. Archer, S.; Lesczynski, M. Project Estimation—Go Parametric to Reduce 'Hectic'. Available online: https://www.pmi.org/learning/library/project-estimation-reduce-hectic-deloitte-6061 (accessed on 28 January 2025).
6. Hueber, C.; Horejsi, K.; Schledjewski, R. Review of cost estimation: Methods and models for aerospace composite manufacturing. *Adv. Manuf. Polym. Compos. Sci.* **2016**, *2*, 1–13. [CrossRef]
7. Tayefeh Hashemi, S.; Ebadati, O.M.; Kaur, H. Cost estimation and prediction in construction projects: A systematic review on machine learning techniques. *SN Appl. Sci.* **2020**, *2*, 1703. [CrossRef]
8. 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]
9. Hosny, S.; Elsaid, E.; Hosny, H. Prediction of construction material prices using ARIMA and multiple regression models. *Asian J. Civ. Eng.* **2023**, *24*, 1697–1710. [CrossRef]
10. Ou, T.-Y.; Cheng, C.-Y.; Chen, P.-J.; Perng, C. Dynamic cost forecasting model based on extreme learning machine—A case study in steel plant. *Comput. Ind. Eng.* **2016**, *101*, 544–553. [CrossRef]
11. Momade, M.H.; Durdyev, S.; Dixit, S.; Shahid, S.; Alkali, A.K. Modeling labor costs using artificial intelligence tools. *Int. J. Build. Pathol. Adapt.* **2024**, *42*, 1263–1281. [CrossRef]
12. Chou, J.-S. Cost simulation in an item-based project involving construction engineering and management. *Int. J. Proj. Manag.* **2011**, *29*, 706–717. [CrossRef]
13. Lowe, D.J.; Emsley, M.W.; Harding, A. Predicting Construction Cost Using Multiple Regression Techniques. *J. Constr. Eng. Manag.* **2006**, *132*, 750–758. [CrossRef]
14. Green, C. Estimating as an Art—What It Takes to Make Good Art. Presented at the PMI® Global Congress 2006—EMEA, Madrid, Spain: Project Management Institute, October 2006. Available online: https://www.pmi.org/learning/library/project-estimating-accurate-labor-costs-8207 (accessed on 22 January 2025).
15. Niederman, F. Project Openings management: For disruption from AI and advanced analytics. *Inf. Technol. People* **2021**, *34*, 1570–1599. [CrossRef]

16. Shamim, M.I. Exploring the success factors of project management. *Am. J. Econ. Bus. Manag.* **2022**, *5*, 64–72.
17. Sanchez, O.P.; Terlizzi, M.A.; De Moraes, H.R.D.O.C. Cost and time project management success factors for information systems development projects. *Int. J. Proj. Manag.* **2017**, *35*, 1608–1626. [CrossRef]
18. Yun, Y.; Ma, D.; Yang, M. Human–computer interaction-based Decision Support System with Applications in Data Mining. *Future Gener. Comput. Syst.* **2021**, *114*, 285–289. [CrossRef]
19. Tijanić, K.; Car-Pušić, D.; Šperac, M. Cost estimation in road construction using artificial neural network. *Neural Comput. Appl.* **2020**, *32*, 9343–9355. [CrossRef]
20. Guo, K.; Yang, Z.; Yu, C.-H.; Buehler, M.J. Artificial intelligence and machine learning in design of mechanical materials. *Mater. Horiz.* **2021**, *8*, 1153–1172. [CrossRef]
21. Bento, S.; Pereira, L.; Gonçalves, R.; Dias, Á.; Costa, R.L.D. Artificial intelligence in project management: Systematic literature review. *Int. J. Technol. Intell. Plan.* **2022**, *13*, 143–163. [CrossRef]
22. Aziz, S.; Dowling, M. Machine Learning and AI for Risk Management. In *Disrupting Finance*; Lynn, T., Mooney, J.G., Rosati, P., Cummins, M., Eds.; Palgrave Studies in Digital Business & Enabling Technologies; Springer International Publishing: Cham, Switzerland, 2019; pp. 33–50. [CrossRef]
23. Gupta, S.; Modgil, S.; Bhattacharyya, S.; Bose, I. Artificial intelligence for decision support systems in the field of operations research: Review and future scope of research. *Ann. Oper. Res.* **2022**, *308*, 215–274. [CrossRef]
24. Tayyab, M.R.; Usman, M.; Ahmad, W. A Machine Learning Based Model for Software Cost Estimation. In *Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016*; Bi, Y., Kapoor, S., Bhatia, R., Eds.; Lecture Notes in Networks and Systems; Springer International Publishing: Cham, Switzerland, 2018; Volume 16, pp. 402–414. [CrossRef]
25. Hammann, D. Big data and machine learning in cost estimation: An automotive case study. *Int. J. Prod. Econ.* **2024**, *269*, 109137. [CrossRef]
26. Krichen, M. Convolutional Neural Networks: A Survey. *Computers* **2023**, *12*, 151. [CrossRef]
27. Ning, F.; Shi, Y.; Cai, M.; Xu, W.; Zhang, X. Manufacturing cost estimation based on a deep-learning method. *J. Manuf. Syst.* **2020**, *54*, 186–195. [CrossRef]
28. Fischer, T.; Krauss, C. Deep learning with long short-term memory networks for financial market predictions. *Eur. J. Oper. Res.* **2018**, *270*, 654–669. [CrossRef]
29. Van Houdt, G.; Mosquera, C.; Nápoles, G. A review on the long short-term memory model. *Artif. Intell. Rev.* **2020**, *53*, 5929–5955. [CrossRef]
30. Xue, X.; Jia, Y.; Tang, Y. Expressway Project Cost Estimation with a Convolutional Neural Network Model. *IEEE Access* **2020**, *8*, 217848–217866. [CrossRef]
31. Upreti, K.; Singh, U.K.; Jain, R.; Kaur, K.; Sharma, A.K. Fuzzy Logic Based Support Vector Regression (SVR) Model for Software Cost Estimation Using Machine Learning. In *ICT Systems and Sustainability*; Tuba, M., Akashe, S., Joshi, A., Eds.; Lecture Notes in Networks and Systems; Springer Nature: Singapore, 2022; Volume 321, pp. 917–927. [CrossRef]
32. Peško, I.; Mučenski, V.; Šešlija, M.; Radović, N.; Vujkov, A.; Bibić, D.; Krklješ, M. Estimation of Costs and Durations of Construction of Urban Roads Using ANN and SVM. *Complexity* **2017**, *2017*, 2450370. [CrossRef]
33. Tajani, F.; Morano, P. A Systematic Analysis of Benefits and Costs of Projects for the Valorization of Cultural Heritage. In *Cultural Territorial Systems*; Rotondo, F., Selicato, F., Marin, V., Galdeano, J.L., Eds.; Springer Geography; Springer International Publishing: Cham, Switzerland, 2016; pp. 107–118. [CrossRef]
34. Dong, J.; Chen, Y.; Guan, G. Cost Index Predictions for Construction Engineering Based on LSTM Neural Networks. *Adv. Civ. Eng.* **2020**, *2020*, 6518147. [CrossRef]
35. Ning, F.; Shi, Y.; Cai, M.; Xu, W.; Zhang, X. Manufacturing cost estimation based on the machining process and deep-learning method. *J. Manuf. Syst.* **2020**, *56*, 11–22. [CrossRef]
36. Chakraborty, D.; Elhegazy, H.; Elzarka, H.; Gutierrez, L. A novel construction cost prediction model using hybrid natural and light gradient boosting. *Adv. Eng. Inform.* **2020**, *46*, 101201. [CrossRef]
37. Jaafari, A.; Pazhouhan, I.; Bettinger, P. Machine Learning Modeling of Forest Road Construction Costs. *Forests* **2021**, *12*, 1169. [CrossRef]
38. Loyer, J.-L.; Henriques, E.; Fontul, M.; Wiseall, S. Comparison of Machine Learning methods applied to the estimation of manufacturing cost of jet engine components. *Int. J. Prod. Econ.* **2016**, *178*, 109–119. [CrossRef]
39. Alshboul, O.; Shehadeh, A.; Almasabha, G.; Almuflih, A.S. Extreme Gradient Boosting-Based Machine Learning Approach for Green Building Cost Prediction. *Sustainability* **2022**, *14*, 6651. [CrossRef]
40. Elmousalami, H.H. Comparison of Artificial Intelligence Techniques for Project Conceptual Cost Prediction: A Case Study and Comparative Analysis. *IEEE Trans. Eng. Manag.* **2020**, *68*, 183–196. [CrossRef]
41. Leśniak, A.; Zima, K. Cost Calculation of Construction Projects Including Sustainability Factors Using the Case Based Reasoning (CBR) Method. *Sustainability* **2018**, *10*, 1608. [CrossRef]

42. Juszczyk, M.; Leśniak, A.; Zima, K. ANN Based Approach for Estimation of Construction Costs of Sports Fields. *Complexity* **2018**, *2018*, 7952434. [CrossRef]
43. Guo, H.; Nguyen, H.; Vu, D.-A.; Bui, X.-N. Forecasting mining capital cost for open-pit mining projects based on artificial neural network approach. *Resour. Policy* **2021**, *74*, 101474. [CrossRef]
44. 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]
45. Wang, R.; Asghari, V.; Cheung, C.M.; Hsu, S.-C.; Lee, C.-J. Assessing effects of economic factors on construction cost estimation using deep neural networks. *Autom. Constr.* **2022**, *134*, 104080. [CrossRef]
46. Rafiei, M.H.; Adeli, H. Novel Machine-Learning Model for Estimating Construction Costs Considering Economic Variables and Indexes. *J. Constr. Eng. Manag.* **2018**, *144*, 04018106. [CrossRef]
47. Relich, M.; Pawlewski, P. A case-based reasoning approach to cost estimation of new product development. *Neurocomputing* **2018**, *272*, 40–45. [CrossRef]
48. Yaseen, Z.M.; Ali, Z.H.; Salih, S.Q.; Al-Ansari, N. Prediction of Risk Delay in Construction Projects Using a Hybrid Artificial Intelligence Model. *Sustainability* **2020**, *12*, 1514. [CrossRef]
49. Vimont, A.; Leleu, H.; Durand-Zaleski, I. Machine learning versus regression modelling in predicting individual healthcare costs from a representative sample of the nationwide claims database in France. *Eur. J. Health Econ.* **2022**, *23*, 211–223. [CrossRef]
50. Mazumdar, M.; Lin, J.-Y.J.; Zhang, W.; Li, L.; Liu, M.; Dharmarajan, K.; Sanderson, M.; Isola, L.; Hu, L. Comparison of statistical and machine learning models for healthcare cost data: A simulation study motivated by Oncology Care Model (OCM) data. *BMC Health Serv. Res.* **2020**, *20*, 350. [CrossRef] [PubMed]
51. Ul Hassan, C.A.; Iqbal, J.; Hussain, S.; AlSalman, H.; Mosleh, M.A.A.; Sajid Ullah, S. A Computational Intelligence Approach for Predicting Medical Insurance Cost. *Math. Probl. Eng.* **2021**, *2021*, 1162553. [CrossRef]
52. Morid, M.A.; Sheng, O.R.L.; Kawamoto, K.; Abdelrahman, S. Learning hidden patterns from patient multivariate time series data using convolutional neural networks: A case study of healthcare cost prediction. *J. Biomed. Inform.* **2020**, *111*, 103565. [CrossRef]
53. Barros, L.B.; Marcy, M.; Carvalho, M.T.M. Construction Cost Estimation of Brazilian Highways Using Artificial Neural Networks. *Int. J. Struct. Civ. Eng. Res.* **2018**, *7*, 283–289. [CrossRef]
54. Ho, W.K.O.; Tang, B.-S.; Wong, S.W. Predicting property prices with machine learning algorithms. *J. Prop. Res.* **2021**, *38*, 48–70. [CrossRef]
55. Wang, H.; Cui, Z.; Chen, Y.; Avidan, M.; Abdallah, A.B.; Kronzer, A. Predicting Hospital Readmission via Cost-Sensitive Deep Learning. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2018**, *15*, 1968–1978. [CrossRef]
56. Yun, S. Performance Analysis of Construction Cost Prediction Using Neural Network for Multioutput Regression. *Appl. Sci.* **2022**, *12*, 9592. [CrossRef]
57. Tatiya, A.; Zhao, D.; Syal, M.; Berghorn, G.H.; LaMore, R. Cost prediction model for building deconstruction in urban areas. *J. Clean. Prod.* **2018**, *195*, 1572–1580. [CrossRef]
58. Chandanshive, V.; Kambekar, A. Estimation of Building Construction Cost Using Artificial Neural Networks. *J. Soft Comput. Civ. Eng.* **2019**, *3*, 91–107. [CrossRef]
59. Pham, T.Q.D.; Le-Hong, T.; Tran, X.V. Efficient estimation and optimization of building costs using machine learning. *Int. J. Constr. Manag.* **2023**, *23*, 909–921. [CrossRef]
60. Elhegazy, H.; Chakraborty, D.; Elzarka, H.; Ebid, A.M.; Mahdi, I.M.; Aboul Haggag, S.Y.; Abdel Rashid, I. Artificial Intelligence for Developing Accurate Preliminary Cost Estimates for Composite Flooring Systems of Multi-Storey Buildings. *J. Asian Archit. Build. Eng.* **2022**, *21*, 120–132. [CrossRef]
61. Abdulmajeed, A.A.; Al-Jawaherry, M.A.; Tawfeeq, T.M. Predict the required cost to develop Software Engineering projects by Using Machine Learning. *J. Phys. Conf. Ser.* **2021**, *1897*, 012029. [CrossRef]
62. Alshamrani, O.S. Construction cost prediction model for conventional and sustainable college buildings in North America. *J. Taibah Univ. Sci.* **2017**, *11*, 315–323. [CrossRef]
63. Chen, X.; Yi, M.; Huang, J. Application of a PCA-ANN Based Cost Prediction Model for General Aviation Aircraft. *IEEE Access* **2020**, *8*, 130124–130135. [CrossRef]
64. Elmousalami, H.H.; Elyamany, A.H.; Ibrahim, A.H. Predicting Conceptual Cost for Field Canal Improvement Projects. *J. Constr. Eng. Manag.* **2018**, *144*, 04018102. [CrossRef]
65. Kovačević, M.; Ivanišević, N.; Petronijević, P.; Despotović, V. Construction cost estimation of reinforced and prestressed concrete bridges using machine learning. *J. Croat. Assoc. Civ. Eng.* **2021**, *73*, 1–13. [CrossRef]
66. Morid, M.A.; Sheng, O.R.L.; Kawamoto, K.; Ault, T.; Dorius, J.; Abdelrahman, S. Healthcare cost prediction: Leveraging fine-grain temporal patterns. *J. Biomed. Inform.* **2019**, *91*, 103113. [CrossRef]
67. Park, B.R.; Choi, E.J.; Hong, J.; Lee, J.H.; Moon, J.W. Development of an energy cost prediction model for a VRF heating system. *Appl. Therm. Eng.* **2018**, *140*, 476–486. [CrossRef]

68. Patra, G.K.; Kuraku, C.; Konkimalla, S.; Boddapati, V.N.; Sarisa, M.; Reddy, M.S. An Analysis and Prediction of Health Insurance Costs Using Machine Learning-Based Regressor Techniques. *J. Data Anal. Inf. Process.* **2024**, *12*, 581–596. [CrossRef]
69. Taloba, A.I.; El-Aziz, R.M.A.; Alshanbari, H.M.; El-Bagoury, A.-A.H. Estimation and Prediction of Hospitalization and Medical Care Costs Using Regression in Machine Learning. *J. Healthc. Eng.* **2022**, *2022*, 7969220. [CrossRef] [PubMed]
70. Ujong, J.A.; Mbadike, E.M.; Alaneme, G.U. Prediction of cost and duration of building construction using artificial neural network. *Asian J. Civ. Eng.* **2022**, *23*, 1117–1139. [CrossRef]