XXI ENTAC – Encontro Nacional de Tecnologia do Ambiente Construído
14 a 16 de outubro de 2026 | Goiânia-Brasil
DOI: https://doi.org/10.46421/entac.v21i1.8669

Inteligência Artificial na estimativa prévia de custos de obras civis: uma revisão sistemática de métodos, variáveis e lacunas

Artificial Intelligence in the preliminary estimation of civil works costs: a systematic review of methods, variables and gaps.

Paola de Melo Silva
Universidade de Brasília (UnB) | Brasília | Brasil | paaolams79@gmail.com
Michele Tereza Marques Carvalho
Universidade de Brasília (UnB) | Brasília | Brasil | micheletereza@unb.br

Resumo

O artigo tem como objetivo identificar as variáveis e os métodos de Inteligência Artificial (IA) aplicados à análise prévia de custos de obras, bem como levantar as principais limitações associadas a esse uso. A metodologia baseou-se em uma revisão sistemática da literatura nas bases *Scopus* e *Web of Science*, utilizando o protocolo PRISMA para seleção de estudos publicados a partir de 2015. Os resultados indicam a predominância de Redes Neurais Artificiais, *Support Vector Machines*, *Deep Learning* e modelos *ensemble*, aplicados principalmente a variáveis físicas do empreendimento, indicadores econômicos, séries históricas e métricas de desempenho físico-financeiro. Observou-se que a escolha do modelo de IA está relacionada à disponibilidade, à natureza e à qualidade dos dados. Como contribuição científica, o estudo sistematiza o estado da arte ao propor uma análise comparativa entre métodos preditivos e tipos de variáveis utilizados na previsão de custos, além de apontar limitações e lacunas de pesquisa relacionadas à necessidade de bases de dados mais estruturadas, validação dos modelos em diferentes contextos e desenvolvimento de abordagens explicáveis para a estimativa prévia de custos.

Palavras-chave: Inteligência Artificial. Orçamento de Obras. Gestão de Custos. Previsão de Custos. Modelos Preditivos.

Abstract

This paper aims to identify the variables and Artificial Intelligence (AI) methods applied to the preliminary cost analysis of construction works, as well as to identify the main limitations associated with their use. The methodology was based on a systematic literature review conducted in the Scopus and Web of Science databases, using the PRISMA protocol to select studies published from 2015 onward. The results indicate the predominance of Artificial Neural Networks, Support Vector Machines, Deep Learning, and ensemble models, mainly applied to physical project variables, economic indicators, time series, and physical-financial performance metrics. It was observed that the choice of the AI model is related to the availability, nature, and quality of the data. As a scientific contribution, the study systematizes the literature by proposing a comparative analysis between predictive methods and types of variables used in cost prediction, while also highlighting limitations and research gaps related to the need for more structured databases, model validation in different contexts, and the development of explainable approaches for preliminary cost estimation.

Keywords: Artificial Intelligence. Construction Budgeting. Cost Management. Cost Forecasting. Predictive Models.

SILVA, P.M., CARVALHO, M.T.M. Inteligência Artificial na estimativa prévia de custos de obras civis: uma revisão sistemática de métodos, variáveis e lacunas. ENTAC 2026. In: ENCONTRO NACIONAL DE TECNOLOGIA DO AMBIENTE CONSTRUÍDO, 21, 2026, Goiânia. **Anais**... Goiânia: ANTAC, 2026.

INTRODUÇÃO E FUNDAMENTAÇÃO

A análise prévia de orçamentos em obras civis constitui etapa estratégica para apoiar decisões como viabilidade, definição de escopo e antecipação de riscos [1].

Inicialmente, utilizavam-se métodos simplificados baseados em atributos físicos, mas a crescente complexidade dos projetos evidenciou limitações dessas abordagens, especialmente em cenários com pouca informação inicial, alta variabilidade de insumos e baixa padronização [2][3].

Sob essa perspectiva, a Inteligência Artificial (IA) surge como uma alternativa mais robusta, permitindo modelar relações complexas e não lineares entre características do projeto, mercado e desempenho executivo [4]. Técnicas como redes neurais, SVM (*Support Vector Machine*), *ensembles* e *deep learning* ampliam a precisão e a capacidade de generalização das estimativas [5][6]. Entretanto, ainda persistem desafios como a dependência de bases históricas consistentes e a baixa interpretabilidade dos modelos preditivos [4][7].

Assim, o artigo busca responder quais são as variáveis e os métodos de Inteligência Artificial aplicados na análise prévia de custos de obras civis, além das principais limitações associadas ao uso dessas técnicas nesse contexto. Para isso, realiza-se uma revisão bibliográfica focada nos métodos de previsão de custos, nas variáveis utilizadas e nas técnicas de IA mais empregadas, apresentado um quadro-síntese das variáveis mais recorrentes em cada modelo de inteligência artificial.

METODOLOGIA

A revisão sistemática da literatura foi realizada nas bases *Scopus* e *Web of Science*, devido à sua relevância e abrangência na área de Engenharia Civil. A busca ocorreu em 20 de março de 2026, considerando publicações a partir de 2015. Utilizou-se a seguinte *string*:

(TITLE-ABS-KEY ("Construction Project" OR "Civil Engineering" OR "building projects") AND TITLE-ABS-KEY ("Machine Learning" OR "Artificial Intelligence" OR "AI" OR "Deep Learning") AND TITLE-ABS-KEY (forecasting OR prognosis) AND TITLE-ABS-KEY (Cost OR "Cost Management" OR "cost estimate")) AND PUBYEAR > 2014 AND PUBYEAR < 2027.

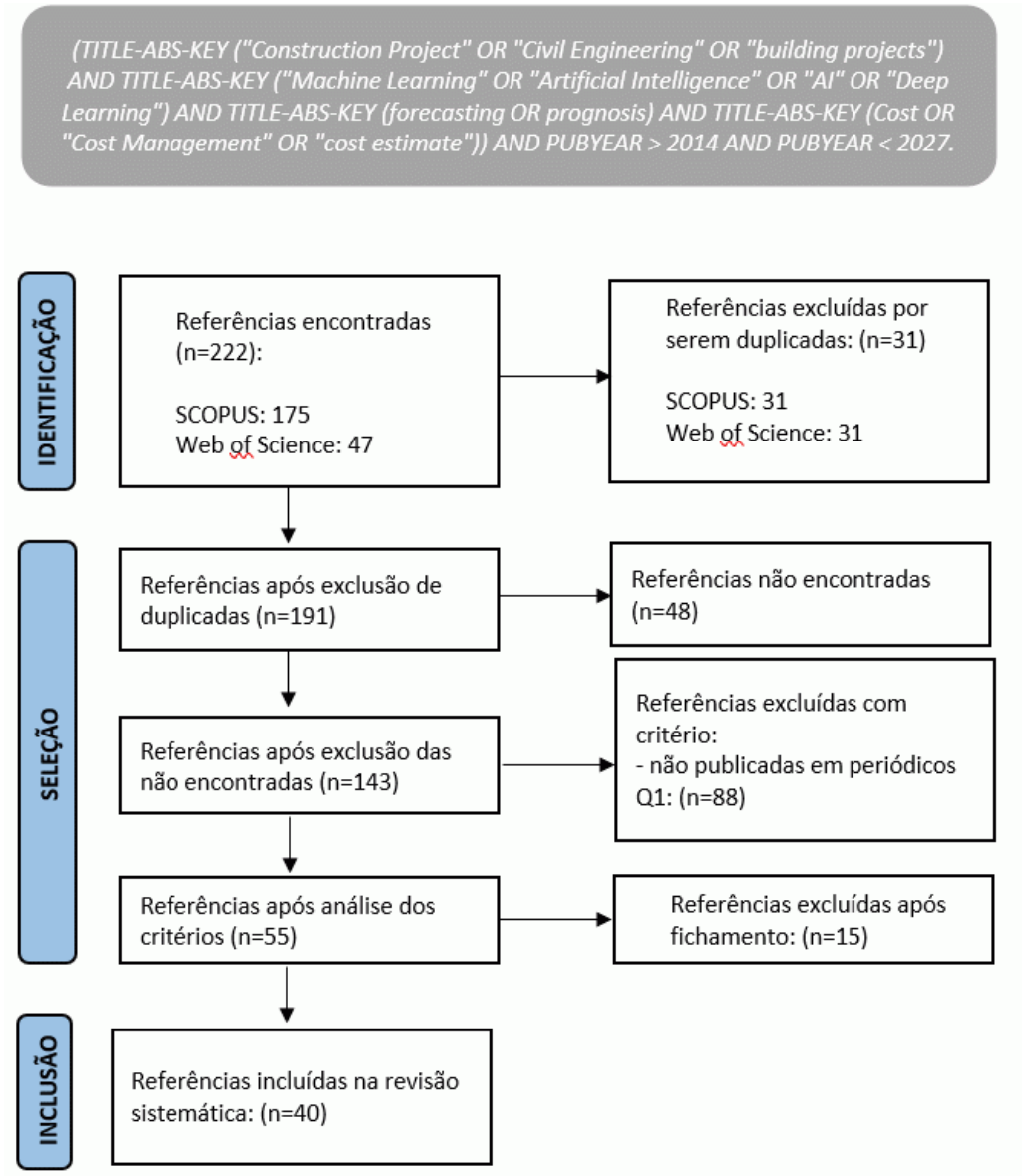
A revisão sistemática identificou 222 referências nas bases *Scopus* e *Web of Science*. Após a exclusão de duplicatas e trabalhos sem acesso restaram 143 estudos.

Como critério preliminar de qualidade científica, adotou-se a inclusão exclusiva de artigos publicados em periódicos classificados como Q1 no *SCImago Journal & Country Rank* (SJR), por apresentarem maior rigor editorial e maior confiabilidade metodológica. Assim, 88 artigos foram excluídos, resultando em 55 estudos.

Posteriormente, realizou-se um fichamento sistemático conforme o protocolo PRISMA, avaliando a aderência ao tema, a qualidade metodológica e mitigando o risco de viés com base em critérios como base de dados, técnica de IA empregada, validação dos modelos, métricas de desempenho e limitações metodológicas. Nessa etapa, 15

artigos foram excluídos por apresentarem fragilidade metodológica, baixo vínculo com previsão de custos ou alto risco de viés, resultando em uma amostra final de 40 artigos incluídos na revisão sistemática, conforme apresentado na Figura 1.

Figura 1: Fluxograma PRISMA



Fonte: o autor.

REVISÃO BIBLIOGRÁFICA

ANÁLISE PRÉVIA DE CUSTOS

A análise prévia de custos constitui, no âmbito da literatura estudada, a base conceitual para comparar e validar métodos de Inteligência Artificial aplicados à previsão orçamentária, especialmente diante dos diferentes níveis de maturidade informacional antes da execução [2][8].

Do ponto de vista metodológico, a literatura indica que a estimativa prévia de custos evolui conforme o grau de desenvolvimento das informações do projeto [2]. Nas fases iniciais, como concepção e estudos de viabilidade, essa análise subsidia decisões de continuidade, redirecionamento ou abandono do investimento, ainda que baseada em informações preliminares e com menor nível de detalhamento, como localização, tipologia, características físicas gerais e requisitos de desempenho [8].

Nesses estágios, predominam métodos tradicionais baseados em atributos físicos simplificados, como custo por área, volume e analogia, como apresentados por [8], por exemplo. Essas abordagens apresentam rápida aplicação, mas perdem robustez em projetos de elevada complexidade e baixa padronização, como os edifícios verdes estudados por [3].

Com o avanço do escopo, a lógica da estimativa prévia passa a incorporar variáveis quantitativas mais detalhadas, séries históricas, indicadores econômicos e parâmetros técnicos do empreendimento, aumentando a precisão e a capacidade explicativa da análise [3][4].

Essa condição é mais relevante em projetos de alta especificidade, como edifícios verdes, uso de materiais inovadores ou contextos com grande variação de preços [3][9]. Nesses casos, a incerteza decorre da combinação entre ausência de quantitativos historicamente consolidados, indefinição parcial de soluções construtivas e baixa disponibilidade de preços de referência [4].

Em níveis mais avançados de informação, tornam-se frequentes modelos estatísticos determinísticos, como regressão linear e múltipla. Na revisão sistemática apresentada por [2], a regressão linear aparece como um dos métodos mais recorrentes para previsão de custos e prazos em edifícios e infraestrutura.

Quando fatores econômicos e temporais influenciam o custo, os modelos passam a incorporar indicadores macroeconômicos, séries históricas e relações de dependência entre variáveis. Nos estudos de preços de materiais apresentados por [4], modelos causais utilizam CPI (*Consumer Price Index*) e PPI (*Producer Price Index*) como variáveis explicativas para prever tendências futuras.

Nesse sentido, as técnicas multivariadas têm sido empregadas para tratar a elevada quantidade de fatores que influenciam os custos de empreendimentos, como em [3], em que a PCA (*Principal Component Analysis*) é utilizada para redução de dimensionalidade dos dados.

Ainda em [3], a PCA transforma um conjunto de variáveis originais, frequentemente correlacionadas entre si, em menos componentes independentes que preservam a maior parte da variabilidade das informações. Segundo [3], essa etapa permite eliminar redundâncias, reduzir o problema da alta dimensionalidade e fornecer vetores de características mais representativos para a modelagem preditiva, aumentando a eficiência e a precisão do modelo de previsão de custos.

Complementarmente, métodos probabilísticos são empregados para explicitar a incerteza inerente à análise prévia, substituindo a lógica de valor único por distribuições de resultados [4]. Essa abordagem permite incorporar a variabilidade

associada aos fatores que influenciam os custos e explicitar o grau de incerteza das previsões.

Diante disso, a simulação de Monte Carlo é apontada por [4] como uma técnica utilizada na literatura para previsão de preços e quantificação de incertezas, permitindo a geração de múltiplos cenários a partir de distribuições de probabilidade e produzindo uma distribuição de resultados futuros em vez de uma estimativa pontual. Assim, métodos probabilísticos subsidiam a análise de risco ao estimar não apenas o valor esperado, mas também a faixa provável de variação dos custos e a ocorrência de cenários mais onerosos.

Além disso, há estudos, como o de [9], que utilizam abordagens como PERT (*Program Evaluation and Review Technique)* e simulação de risco para reforçar a relação entre incertezas de prazo e impacto no custo final. Nesses estudos, o custo final é continuamente ajustado conforme o avanço físico-financeiro do projeto.

Em síntese, observa-se que já existe uma base metodológica consolidada para a análise prévia de custos, composta por estimativas paramétricas, regressões, modelos causais, técnicas multivariadas e tratamento probabilístico da incerteza. Dessa forma, a Inteligência Artificial pode ser utilizada para expandir, automatizar e aprimorar a capacidade de generalização dos modelos, especialmente em cenários de informação limitada e nas fases iniciais do projeto.

RESULTADOS E DISCUSSÕES

MODELOS PARA ANÁLISE PRÉVIA

A evolução do prognóstico de custos na construção civil tem sido marcada pela migração de métodos determinísticos e regressivos tradicionais para abordagens baseadas em inteligência artificial capazes de representar relações complexas, não lineares e dinâmicas entre variáveis de projeto, mercado e desempenho executivo [1].

Nos estudos analisados nesta revisão bibliográfica, observa-se que a IA deixou de atuar apenas como ferramenta de estimativa pontual e passou a sustentar previsões preliminares, controle de custos, previsão de índices, otimização e apoio à tomada de decisão em aquisição de insumos [6][5].

As Redes Neurais Artificiais (RNAs) constituem a família mais tradicional e recorrente no prognóstico de custos. Nos artigos desta revisão, aparecem principalmente arquiteturas MLP (*Multi-Layer Perceptron*), RBF (*Radial Basis Function*) e DNN (D*eep Neural Networks)*.

O uso de MLP é frequentemente aplicado à estimativa preliminar de edifícios, aprendendo com obras passadas a relação entre características do projeto, como área construída, número de pavimentos, padrão construtivo e prazo contratual, e custo final [1][5]. Enquanto isso, as RBF são usadas como modelos comparativos para avaliar desempenho preditivo em custo e prazo, especialmente em relações não lineares [10][5]. Já as DNN aparecem em estudos mais recentes e são voltadas à análise de

dados reais e à identificação dos principais componentes que influenciam o desempenho dos custos durante a execução [6].

Os métodos baseados em kernel aparecem também nos artigos revisados, sobretudo SVR (*Support Vector Regression*), LSSVM (*Least-Squares Support Vector Machines*), PCA-SVM (*Principal Component Analysis - Support Vector Machines*) e outros modelos híbridos.

O SVR é utilizado em estimativas preliminares de custos de edifícios, infraestrutura urbana e previsão de índices, com destaque para modelos híbridos como o PCA-SVM, que reduzem redundâncias entre variáveis econômicas antes da regressão [1].

Por sua vez, o LSSVM é descrito em [4] como método causal recorrente na previsão de preços de materiais, utilizando indicadores econômicos e séries históricas, além de servir como modelo de comparação em estudos com métodos mais avançados de previsão temporal de preço de insumos, especialmente aço e vergalhão, como em [11].

A literatura estudada nesta revisão mostra também um crescimento de métodos de árvore e *ensemble*, como: *Random Forest, XGBoost* e *LightGBM.*

Um exemplo representativo do uso de *Random Forest* aparece no estudo de [8], sintetizado por [5], no qual o algoritmo é empregado na forma de *Random Forest Regressor* para previsão de custos de segurança.

Nos estudos comparativos de [5], métodos como *XGBoost, LightGBM* são aplicados a bases heterogêneas compostas por dados históricos do projeto, atributos categóricos e variáveis derivadas por engenharia de atributos para prever variáveis associadas a custo, prazo e contingência. O *LightGBM*, em especial, é combinado com SHAP *(SHapley Additive exPlanations)*, permitindo interpretar quais atributos históricos mais influenciam as previsões de desempenho

Pode-se destacar que os estudos mostram o uso de métodos *ensemble* e híbridos em bases heterogêneas de custo, especialmente quando integram múltiplas fontes de dados, como históricos de projeto, variáveis técnicas, dados de controle em tempo real e fatores externos [5].

Em [5], o EVM-ANN integra métricas de desempenho como SPI e CPI a redes neurais para prever simultaneamente custo e prazo, enquanto o híbrido *CatBoost* + SVM é aplicado à previsão de custos de mudanças de projeto, combinando atributos categóricos e relações não lineares.

Em resumo, a escolha do modelo depende menos da sofisticação da arquitetura e mais da compatibilidade entre algoritmo, volume de dados e natureza das variáveis analisadas [1]. Embora as redes neurais artificiais sejam as mais frequentes na literatura, sua superioridade não é absoluta.

Em [1], modelos baseados em kernel, como SVR e LSSVM, frequentemente superam MLPs em bases pequenas e médias, pois apresentam melhor capacidade de generalização e menor sensibilidade ao sobreajuste (*overfitting*). Enquanto o MLP se

destaca em relações fortemente não lineares e grandes volumes de dados, o SVR tende a ser mais estável quando há poucas obras comparáveis [1].

De forma semelhante, observa-se desempenho superior de ensembles baseados em árvores, especialmente em bases heterogêneas com dados históricos, variáveis categóricas e fatores externos [5].

Em [5], destaca-se que métodos ensemble e híbridos apresentaram maior acurácia, com o *LightGBM*, que alcançou baixo erro em comparação aos demais modelos, além de maior interpretabilidade com uso de SHAP, reforçando a migração de modelos black-box, como DNN profundas, para abordagens mais explicáveis.

Quanto a métrica de avaliação e comparação dos modelos predominam MAE (Erro Absoluto Médio), RMSE (Raiz do Erro Quadrático Médio), MSE (Erro Quadrático Médio) e MAPE (Erro Percentual Absoluto Médio), sendo o R^2 (Coeficiente de Determinação) utilizado em parte dos estudos comparativos, sobretudo em *benchmarks* entre diferentes famílias algorítmicas [1][5].

VARIÁVEIS MAIS UTILIZADAS NOS MODELOS DE IAs

Em termos gerais, os tipos de variáveis adotados dependem da técnica de IA aplicada, o nível de maturidade das informações do projeto e o objetivo preditivo do modelo. Destacam-se variáveis físicas do empreendimento, valores históricos de preços e produtividade, indicadores econômicos, métricas gerenciais de desempenho e atributos derivados por engenharia de dados, como fatores de risco, contingências atributos categóricos e componentes extraídos por PCA, mais presentes em *ensembles* e modelos híbridos.

O Quadro 1 agrupa artigos por variáveis recorrentes e relaciona cada grupo aos modelos de IA apresentando as referências correspondentes.

Quadro 1: Quadro-Síntese de variáveis por técnica de modelo de inteligência artificial

Tipos de variáveis predominantes	Exemplos de variáveis mais recorrentes	Técnica de IA	Frequência	Papel no prognóstico	Referências
Variáveis físicas, paramétricas, geométricas e atributos do empreendimento	Área construída; número de pavimentos; tipologia; padrão construtivo; prazo contratual	Redes Neurais Artificiais	16	Estimativa preliminar de custo total e generalização não linear	[A1 - A16]
Variáveis físicas, econômicas e séries de preços	Área acima do solo; custo total; índices econômicos; preços históricos de aço/vergalhão	Support Vector Machines	7	Regressão robusta para custo preliminar e índices de materiais	[B1 - B7]
Séries históricas, produtividade, sequências temporais, custo acumulado	Janela histórica de produtividade; custo acumulado; séries de preços; atraso; consumo de recursos	Deep Learning / Redes Recorrentes	7	*Forecast* dinâmico de custo, produtividade e atraso	[C1 - C7]
Variáveis heterogêneas, atributos categóricos, risco, contingência e histórico executivo	Fatores de risco; contingência; atributos categóricos; histórico de desvios; produtividade passada	Ensembles Baseados em Árvores	6	Robustez preditiva e interpretabilidade por importância das variáveis	[D1 - D6]
Séries temporais + indicadores econômicos + atributos derivados	Séries de custo; CPI/PPI; embeddings; componentes PCA; variáveis derivadas	Híbrido Deep Learning + SVM	3	Captura simultânea de dependência temporal e não linearidade	[E1 - E3]
CPI, SPI, SV, CV, progresso físico-financeiro	CPI; SPI; custo real; custo planejado; progresso físico acumulado	Neuro-Fuzzy / Híbridos gerenciais	1	Controle preditivo de custo e prazo	[F1]

A1: CHOU et al. (2015) doi: 10.1016/j.autcon.2015.02.006; A2: HAMDAN et al. (2025) doi: 10.1108/ECAM-09-2024-1209; A3: JEZZINI et al. (2025) doi: 10.1061/JCEMD4.COENG-16055; A4: JUSZCZYK et al. (2019) doi: 10.3846/jcem.2019.10534; A5: KHALEF; EL-ADAWAY (2021) doi: 10.1061/(ASCE)ME.1943-5479.0000959; A6: MAHDAVIAN et al. (2021) doi: 10.1061/(ASCE)CO.1943-7862.0001991; A7: MEHARIE et al. (2021) doi: 10.1108/ECAM-02-2020-0128; A8: MIR et al. (2021) doi: 10.1016/j.jobe.2021.102288; A9: NINDARTIN et al. (2025) doi: 10.3846/jcem.2025.24913; A10: PAZOUKI et al. (2025) doi: 10.1016/j.istruc.2025.109531; A11: NAGARAJ et al. (2022) doi: 10.1016/j.conbuildmat.2022.127933; A12: WANG et al. (2023) doi: 10.1016/j.jobe.2023.107150; A13: WANG et al. (2025) doi: 10.1108/ecam-02-2025-0259; A14: OCAK; ŞEKER; ROSTAMI (2018) doi: 10.1016/j.tust.2018.07.030; A15: ZHENG et al. (2024) doi: 10.1016/j.eswa.2024.124691; A16: DARKO et al. (2023) doi: 10.1061/JCEMD4.COENG-13101.
B1: ZHANG et al. (2020) doi: 10.1016/j.conbuildmat.2020.118803; B2: SABILLON et al. (2020) doi: 10.1061/(ASCE)CP.1943-5487.0000863; B3: MAHMOODZADEH et al. (2020) doi: 10.1016/j.tust.2020.103514; B4: CHENG; KHITAM; CHANG (2026) doi: 10.1108/ECAM-06-2025-0972; B5: CHENG; CAO; JAYA MENDROFA (2021) doi: 10.1016/j.jobe.2020.101973; B6: NESVOLD; BRATVOLD (2023) doi: 10.2118/212834-PA; B7: WU et al. (2025) doi: 10.1108/ECAM-09-2024-1247.
C1: CHENG; KHASANI (2024) doi: 10.36680/j.itcon.2024.023; C2: CHENG et al. (2024) doi: 10.1108/ECAM-03-2024-0324; C3: ALZAHRANI; SENOUCI; EL-RAYES (2025) doi: 10.1061/JCEMD4.COENG-16650; C4: KIM; MAKHMALBAF; SHAHANDASHTI (2025) doi: 10.1108/ECAM-06-2023-0544; C5: LI et al. (2025) doi: 10.3390/su172411097; C6: ALZARA et al. (2025) doi: 10.1371/journal.pone.0333189; C7: ALTALHONI et al. (2025) doi: 10.1061/JCEMD4.COENG-16622.
D1: PEIMAN et al. (2023) doi: 10.1108/ECAM-12-2022-1170; D2: HUANG; HSIEH (2020) doi: 10.1016/j.autcon.2020.103280; D3: SADICK; GURMU; GUNARATHNA (2024) doi: 10.1108/ECAM-11-2023-1108; D4: SONG; CHOI; CHO (2024) doi: 10.1061/JCEMD4.COENG-14567.
E1: REGONA et al. (2022) doi: 10.3390/joitmc8010045; E2: OOMMEN; GODAWATTE; NEILSEN (2026) doi: 10.1016/j.rineng.2026.109862; E3: SHARMA et al. (2021) doi: 10.3390/s21020463.
F1: MHADY; GÜRGÜN (2025) doi: 10.1061/JCEMD4.COENG-16487

Fonte: o autor.

DISCUSSÃO DOS RESULTADOS

O Quadro 1 mostra que a escolha do modelo de IA depende diretamente do tipo de variável utilizada no prognóstico de custos. As RNAs, mais frequentes na revisão (19 artigos), concentram-se em variáveis físicas e paramétricas, como área construída, número de pavimentos e padrão construtivo, sendo mais adequadas para estimativas preliminares de custo, onde predominam relações não lineares e menor detalhamento do projeto [8].

Modelos de *Deep Learning* e redes recorrentes, como LSTM, são mais aplicados a séries históricas, produtividade e preços acumulados, demonstrando maior aderência a previsões dinâmicas, como *forecast* de insumos e controle de atrasos [12]. Já os ensembles baseados em árvores, como *Random Forest* e *Gradient Boosting*, trabalham melhor com variáveis heterogêneas, fatores de risco e contingências, além de oferecerem maior interpretabilidade, aspecto relevante para apoio à decisão e auditorias [10].

Modelos híbridos, SVM e sistemas *Neuro-Fuzzy* aparecem em aplicações mais específicas, como séries econômicas, índices de preços e controle executivo por meio de indicadores como CPI e SPI [13].

A principal contribuição analítica do Quadro 1 é mostrar que não existe uma IA universalmente superior, mas sim uma adequação entre o tipo de variável e o modelo utilizado. Assim, a qualidade e padronização dos dados são mais determinantes que a sofisticação algorítmica.

LIMITAÇÕES DA IA PARA ANÁLISE PRÉVIA DE CUSTOS

Os estudos convergem em três limitações principais no uso de IA para prognóstico de custos:

- a elevada complexidade estrutural dos modelos, especialmente em ANN, SVR, *ensembles* e *deep learning*, apesar do bom desempenho em relações não lineares [1][5];
- a baixa interpretabilidade do processo decisório, sobretudo em arquiteturas profundas, o que tem impulsionado o uso de técnicas explicativas como SHAP em modelos como *LightGBM, XGBoost* e *Random Forest* [5][7] e;
- a forte dependência de bases históricas consistentes e representativas, uma vez que bases pequenas ou pouco robustas aumentam o risco de sobre ajuste, reduzem a generalização e comprometem a confiabilidade das previsões em novos cenários [1][4][5].

LACUNAS E OPORTUNIDADES DE PESQUISAS FUTURAS

A síntese das limitações recorrentes mostra que pesquisas futuras devem priorizar modelos mais robustos para cenários de baixa maturidade informacional, especialmente nas fases iniciais dos projetos, marcadas por variáveis qualitativas e alta incerteza [8][2].

No Brasil, essa lacuna é ampliada pela ausência de bases nacionais abertas e estruturadas, com dados consolidados de obras públicas para treinamento de algoritmos de IA, além da baixa padronização dos históricos dessas obras, o que limita a validação externa e a aplicação prática dos modelos.

Soma-se a isso o fato de que a maioria das referências é internacional, dificultando a generalização direta para o Brasil devido às diferenças de produtividade, contratação pública e composição de insumos, reforçando a necessidade de estudos aplicados ao contexto nacional.

Como oportunidade, destaca-se a construção de bases nacionais abertas, padronizadas e auditáveis, integrando medições, aditivos contratuais, cronogramas, achados de auditoria e composições oficiais de custos, aproximando os modelos preditivos da realidade das obras públicas brasileiras.

Também se observa a ausência de modelos híbridos que combinem *transfer learning*, para reduzir os efeitos da escassez de dados históricos, com métodos de interpretabilidade como SHAP, capazes de aliar maior desempenho preditivo e transparência analítica, ainda pouco explorados nos artigos revisados.

CONCLUSÃO

A revisão bibliográfica consolidou o estado da arte sobre o uso da Inteligência Artificial na análise prévia de custos de obras civis, evidenciando a evolução dos métodos tradicionais para modelos preditivos capazes de representar relações complexas entre variáveis técnicas, econômicas e executivas.

Os estudos analisados destacam o uso predominante de redes neurais artificiais, redes recorrentes, máquinas de vetor de suporte e modelos *ensemble*, aplicados principalmente a variáveis físicas do empreendimento, séries históricas, indicadores econômicos, métricas de desempenho físico-financeiro e atributos derivados por engenharia de dados.

Essa sistematização constitui importante contribuição científica, pois permite compreender não apenas quais modelos apresentam melhor desempenho, mas também quais fatores exercem maior influência sobre a precisão preditiva.

Apesar dos avanços, ainda persistem desafios relacionados à baixa interpretabilidade de modelos complexos, à dependência de bases históricas consistentes, à dificuldade de generalização e à escassez de dados estruturados.

Assim, pesquisas futuras poderiam priorizar o desenvolvimento de bases interoperáveis de dados nacionais, modelos explicáveis e a incorporação de dados dinâmicos de custo, prazo e desempenho ao longo do ciclo de vida do empreendimento.

Conclui-se, portanto, que a IA representa uma evolução metodológica relevante para tornar a estimativa prévia de custos mais robusta em cenários complexos.

INCLUSÃO DE DECLARAÇÃO DE USO DE IA

Durante a preparação deste artigo, o primeiro autor utilizou o ChatGPT, da OpenAI, em versões da família GPT-5 para a tradução, revisão gramatical do texto e no apoio à interpretação de dados. Após a utilização da ferramenta, os autores revisaram e editaram o conteúdo conforme necessário e assumem total responsabilidade pelo conteúdo do artigo publicado.

REFERÊNCIAS

[1] HASHEMI, S. T.; EBADATI, O. M.; KAUR, H. Cost estimation and prediction in construction projects: a systematic review on machine learning techniques. **SN Applied Sciences**, v. 2, n. 1703, 2020. DOI: https://doi.org/10.1007/s42452-020-03497-1

[2] SADIKOGLU, E.; DEMIRKESEN, S. Mapping the state-of-the-art: bibliometric and systematic analysis of machine learning and AI-based cost and time prediction in construction. **Journal of Asian Architecture and Building Engineering**, p. 1-28, 2026. DOI: https://doi.org/10.1080/13467581.2026.2641274

[3] WU, Z.; LIU, M.; MA, G.; JIANG, S. A hybrid forecasting model to improve cost prediction accuracy in green building projects with machine learning. **Engineering, Construction and Architectural Management**, v. 33, n. 4, p. 3375, 2026. DOI: https://doi.org/10.1108/ECAM-09-2024-1247

[4] LIU, Q.; HE, P.; PENG, S.; WANG, T.; MA, J. A survey of data-driven construction materials price forecasting. **Buildings**, v. 14, n. 10, p. 3156, 2024. DOI: https://doi.org/10.3390/buildings14103156

[5] SADEGHI, N.; NISU, C.; MARJANI, T.; LOTFI, R. Machine learning-enabled construction project management: systematic review and comparative performance. **Built**

Environment Project and Asset Management, p. 1-23 2026. https://doi.org/10.1108/BEPAM-11-2024-0264

[6] GAO, Y.; ANTWI-AFARI, M. F.; HUANG, Y.; CHEN, Z. Artificial intelligence in construction project management: a systematic literature review of cost, time, and safety. **Buildings**, v. 16, n. 5, p. 1061, 2026. DOI: https://doi.org/10.3390/buildings16051061

[7] REGONA, M.; YIGITCANLAR, T.; XIA, B.; LI, R. Y. M. Opportunities and adoption challenges of AI in the construction industry: a PRISMA review. **Journal of Open Innovation: Technology, Market, and Complexity**, v. 8, n. 1, p. 45, 2022. DOI: https://doi.org/10.3390/joitmc8010045

[8] SADICK, A.-M.; GURMU, A.; GUNARATHNA, C. Artificial intelligence-based pre-conception stage construction budget decision-making model. **Engineering, Construction and Architectural Management**, v. 32, n. 7, p. 4552, 2025. DOI: https://doi.org/10.1108/ECAM-11-2023-1108

[9] OTTAVIANI, F. M.; DE MARCO, A.; NARBAEV, T.; BALLESTEROS-PÉREZ, P. Work rate-based indicators for improving project performance regression models**. International Journal of Construction Management**, p. 99-112, 2025. DOI: https://doi.org/10.1080/15623599.2025.2508293

[10] XIA, X.; XIANG, P.; KHANMOHAMMADI, S.; GAO, T.; ARASHPOUR, M. Predicting safety accident costs in construction projects using ensemble data-driven models**. Journal of Construction Engineering and Management**, v. 150, n. 4, p. 04024015, v. 150, n. 7, 2024. DOI: https://doi.org/10.1061/JCEMD4.COENG-14567

[11] CHENG, M.-Y.; VU, Q.-T.; DESSALEGN, M.; CHEN, J.-H. Time-dependent rebar price prediction for procurement decision-making using deep learning. **Engineering, Construction and Architectural Management**, v. 32, n. 12, p. 7938, 2025. DOI: https://doi.org/10.1108/ECAM-03-2024-0324

[12] ALZARA, M.; GIHAD, N.; ABDOU, H.; SOLTAN, A.; HILAL, A.; EHAB, A. Macroeconomic-aware forecasting of construction costs in developing countries: using gated recurrent unit and long short-term memory deep learning framework. **PLOS ONE**, v. 20, n. 10, p. e0301234, 2025. DOI: https://doi.org/10.1371/journal.pone.0333189

[13] MHADY, A. A.; GURGUN, A. P.; BUDAYAN, C.; KOC, K. New hybrid models integrating the firefly optimization algorithm with the artificial neural networks and adaptive neuro-fuzzy inference systems to improve estimation at completion. **Journal of Construction Engineering and Management**, v. 151, n.11, 2025. DOI: https://doi.org/10.1061/JCEMD4.COENG-16487