Article | Received 14 May 2024; Accepted 30 August 2024; Published 4 September 2024
https://doi.org/10.55092/sc20240009

Can ChatGPT assist in cost analysis and bid pricing in construction estimating? A pilot study using a bridge rehabilitation project

Alireza Ghasemi and Fei Dai *

Wadsworth Department of Civil and Environmental Engineering, West Virginia University, Morgantown, WV 26506-6103, USA

* Correspondence author; E-mail: fei.dai@mail.wvu.edu.

Abstract: As the large language model Generative Pre-trained Transformer 4 (GPT-4) recently came into being and has attracted much attention, this study examined its efficacy in analyzing the cost of work items and estimating bid prices in construction estimating. This study utilized a rehabilitation project for the Beaver Dam Road Bridge in Pennsylvania, USA as a case study. The authors integrated ChatGPT-4 to handle bid pricing for five specific work items: concrete and formwork, reinforcement, structure backfill, membrane waterproofing system installation, and borrow excavation. Prior knowledge regarding production rates, labor hourly rates, equipment rates, and material rates was used as input. Prompts and instructions were established for interactive execution of the cost estimation. The model's outputs were compared with the ground truth and the bids from three bidders available at Pennsylvania Department of Transportation (PennDOT)'s website. The comparative analysis revealed that GPT-4 holds the potential for construction estimating with reasonable accuracy. However, it is also essential to recognize the consistency and reliability issues that may exist, which would affect ChatGPT's performance in new scenarios.

Keywords: construction management; cost analysis; ChatGPT; bid pricing

1. Introduction

Cost estimating plays a pivotal role in the construction industry, influencing project planning, budgeting, and overall profitability. Traditional methods, reliant on manual procedures, are prone to be tedious and time-consuming, posing a challenge in the industry where efficiency and speed are key competitive factors [1]. In these methods, it requires comprehending project requirements, calculating the needed materials, labor, and equipment, sourcing current market prices, and estimating total costs with contingencies. Contractors often use spreadsheets, calculators, and historical data for these manual calculations. While being practiced for long time, these methods are susceptible to human errors, which can cause discrepancies between estimated and actual costs, potentially jeopardizing financial stability. Therefore, there's been more use of technologies such as data analytics and artificial intelligence [2]. The advent of Generative Pre-Trained Transformers (GPT), one of Natural Language Processing (NLP) techniques, introduces a shift toward automating and enhancing the cost-critical process. GPT allows machines to understand and interpret human language, transforming unstructured data into actionable insights [3]. This capability is vital in extracting relevant information from project documents and specifications [4] and thereby automating the cost estimating process with its deep learning capabilities. This study made an attempt to explore the potential of ChatGPT in assisting in construction estimating. The objective is to evaluate how ChatGPT-4 can analyze construction related documents and estimate bid prices in an efficient manner. By doing so, this study seeks to address the industry's challenges related to estimating inefficiencies and project delays.

2. Literature review

2.1. State of practice in construction estimating

In the construction industry, cost estimating is a process used to estimate the costs of a project before it begins. This process involves forecasting the costs of materials, labor, equipment, and other resources required to complete the construction project within a specified scope, quality, and timeframe. Construction cost estimating is essential for accurate budgeting and financial planning, ensuring that the projects remain within their economic constraints. It provides a crucial foundation for decision-making in project management, helping to control costs and efficiently schedule resources. Inaccurate estimates, whether they are overestimates or underestimates, can lead to significant deviations between the budgeted and actual costs, potentially impacting financial stability [5].

Currently, the construction cost estimating process that is based on traditional methods which involves several steps, each reliant on the estimator's experience and expertise. Estimators review project plans, specifications, and other documents to comprehend the full extent and identify all the materials, labor, and equipment needed for the project. This process largely relies on the manual interpretation and calculation for each project element by contractors using spreadsheets, calculators, reference books, and historical data to calculate costs [6]. Estimating software is often used to facilitate this process. Using specialized software programs such as STACK and Bluebeam that provide templates, databases, and built-in formulas, contractors may automate the estimating process in steps such as material takeoff and measurements. To gather the material rates, estimators reach out to multiple suppliers to obtain quotes for materials such as concrete, steel, lumber, and other construction materials to ensure that they have competitive pricing. Estimators review price lists and catalogs provided by suppliers, which contain information about the cost of materials, including bulk discounts, seasonal variations, and delivery charges. Regularly updating price lists is essential to account for market fluctuations and ensure that estimates reflect current market conditions. For labor rates, estimators obtain current labor rates, which may vary based on factors such as geographic location, union agreements, and the skill level required for different tasks. There is a need for consulting with labor unions, reviewing wage agreements, and analyzing historical labor cost data. Another component is the cost of renting equipment (e.g., cranes, excavators, loaders). Estimators gather rental rates from equipment rental companies, considering factors like rental duration, transportation costs, and availability. Once the quantities of materials, labor, and equipment are calculated and current market prices are obtained, estimators estimate each construction item's unit prices and total costs. The unit price for each item is determined by dividing the total cost by the quantity. For example, the unit price of concrete might be calculated by dividing the total cost of concrete (including material cost, delivery, and any additional charges) by the volume of concrete required. Estimators compile the unit prices and quantities for all project items to calculate the total cost. However, traditional methods that sum up construction cost variables fall short of providing accurate estimates due to complex patterns of distribution, numerous dependencies, and their variable nature over time [7]. Addressing these complexities, recent advancements have leveraged artificial intelligence (AI) techniques to enhance the accuracy of construction cost estimations. Unlike the conventional rule-based methods, which are inadequate for managing the intricate relationships between project features and cost outcomes, AI offers a level of cognitive processing akin to human intelligence. This technology has proven effective in tackling the challenges in construction cost estimating [8].

2.2. State of research in artificial intelligence-assisted construction estimating

Researchers have led efforts to integrate AI in construction estimating and quantity take-off in order to increase efficiency, reduce cost, and improve accuracy. Tang *et al.* [9] introduced a method that integrates Hidden Markov Models with structured labeling rules to automate information extraction from construction work descriptions. This advancement promises to streamline quantity takeoff (QTO) and cost estimation processes, reducing manual labor and improving accuracy. Karan *et al.* [10] explored NLP and image processing to streamline QTO from 2D construction drawings. Their method, demonstrated via a residential building case study, integrated NLP for text analysis, image processing for object measurement, and machine learning to automate analysis and enhance efficiency. Ambrule *et al.* [11] developed a model to estimate pre-design costs of building projects, specifically focusing on reinforced concrete buildings during the early stage of design. This model also identifies key factors influencing the overall cost of buildings, particularly for construction projects in Philippines. Al-Tawal *et al.* [12] conducted a study to Artificial Neural Network (ANN) to enhance cost estimation accuracy in early phases of building design. Utilizing data from 104 construction projects in Jordan, they developed three ANN models corresponding to concept, schematic, and detailed design stages using different design factors. Yi *et al.* [13] introduced an AI-driven model for improving construction cost estimation and control using Deep Neural Networks (DNN) based on real-world project data. Wang *et al.* [14] investigated the impact of economic factors on construction cost estimation, using DNN and SHapley Additive exPlanations (SHAP) for analysis. They focused on 98 public school projects in Hong Kong. The findings revealed that the economic factors affect estimation accuracy and are more influential than project-specific parameters. In another research, Wang *et al.* [15] created a method for predicting construction costs by utilizing a gray back-propagation neural network (BPNN) model, and tested the model using real-world engineering cost data from city of Zhengzhou, China. Their results showed that the gray BPNN model has a promising prediction accuracy.

In exploring the integration of AI in construction estimating, challenges have been identified that could hinder the effective use and adoption of AI. Key challenges among these are the quality and availability of data, which are critical for training reliable AI models amenable to various construction projects. Moreover, AI's complexity and lack of transparency can reduce trust among users.

2.3. Generative pre-trained transformers

Generative Pre-Trained Transformer (GPT) is a language model developed by OpenAI that uses reinforcement learning to generate human-like text. GPT is a machine learning model that can generate text in response to a prompt, without the need for pre-defined rules or decision trees. GPT models are trained using vast amounts of unstructured text data, enabling them to generate language almost indistinguishable from human-generated text. The series started with GPT-1 in June 2018 and has since evolved to GPT-2, GPT-3, and GPT-3.5 [16]. The latest addition, GPT-4, was launched in March 2023 and demonstrates significant advancements in generating coherent and understandable text. One of the main advantages of GPT models is their capacity to produce language that is cohesive, fluent, and nearly indistinguishable from text produced by humans. They can produce answers to open-ended questions, making them an important tool for natural language communication [17]. The transformer network creates coherent and fluent output while the attention mechanism enables the model to focus on pertinent portions of the input text [18].

ChatGPT-4 introduces a novel feature termed 'GPTs', enabling users to develop tailored versions of ChatGPT. These custom models can be trained with specific datasets and instructions provided by the user. This innovation allows for the creation of specialized ChatGPT instances, each designed for a distinct function or task.

2.4. Recent advances in applying generative pre-trained transformer (gpt) in civil engineering

Many researchers are currently investigating the application of ChatGPT in different disciplines such as finance, health, education, and biotechnology. However, little published information is available on the application of ChatGPT in civil engineering yet there is significant potential [19]. In civil engineering, Prieto *et al.* [20] investigated the application of ChatGPT in automating construction project scheduling, demonstrating its potential to streamline repetitive tasks and receive positive feedback from users on both interaction quality and output coherence. Zhang *et al.* [21] investigated how the ChatGPT-4 tool can help with energy management in buildings by doing tasks like predicting energy needs, finding faults, and spotting unusual patterns. They found that GPT-4 is quite useful in these tasks, making it easier to use data mining techniques in real-world situations. Aladağ [22] examined the potential ChatGPT in improving risk management in construction projects. His research, using performance indicators and expert evaluations, indicated a moderate effectiveness of ChatGPT, with strengths in risk response and monitoring but limitations in risk identification and analysis. In a study conducted by Naser *et al.* [23], the ability of advanced chatbots was explored, specifically ChatGPT-4 and Google's Bard, to pass the Fundamentals of Engineering (FE) and Principles and Practice of Engineering (PE) exams. They assessed the chatbots on typical civil and environmental engineering questions in these exams, focusing on the relevance, accuracy, and clarity of the responses. The findings revealed that the February 2023 edition of ChatGPT-4 scored 70.9% on the FE exam and 46.2% on the PE exam, while Bard scored 39.2% and 41% respectively. The October 2023 edition of ChatGPT-4 showed improved results, scoring over 70% in both exams. Ray *et al.* [24] explored how ChatGPT and Bard, impact water research by posing 50 questions on water treatment and 50 questions on water harvesting to them. They found these chatbots effective in various water treatment fields, including conventional and advanced techniques, membrane technology, and seawater desalination. Rane *et al.* [25] delved into the integration of ChatGPT and Bard, within the realms of architectural design and engineering.

ChatGPT has shown proficiency in mathematical problem-solving across a range of complexities and tasks. ChatGPT-4 and its predecessors have demonstrated a capacity to solve a variety of mathematical problems, from basic arithmetic to more complex tasks [26]. For example, OpenAI's evaluations showed GPT-3.5 scored in the 70th percentile and GPT-4.0 in the 89th percentile on Math SAT tests, showcasing their competency in mathematical reasoning [27]. Research by Plevris *et al.* [28] assessed GPT’s performance in solving problems of varying difficulty, finding reliable results for straightforward calculations and basic logical tasks. This aligns with studies by Dao and Le [29] and Wardat *et al.* [30] which confirmed GPT's problem-solving skills across multiple mathematical domains, although its performance can vary significantly with task complexity and specific mathematical areas. In terms of mathematical modeling, Blomhøj and Jensen [31] emphasized that modeling competencies involve understanding real-world problems, simplifying and structuring them, mathematizing the real model, and interpreting and validating the solutions. ChatGPT's ability to engage in these activities highlights its potential in solving text-based tasks into mathematical problems [32].

3. Problem statement and objective

The construction industry that currently relies on traditional cost-estimating methods is diverging from the imperative for precision, promptness, and effectiveness in project cost estimating. These methods, which depend on manual data handling, are tedious and time-consuming and therefore error-prone and inefficient. While AI, notably GPT such as ChatGPT-4, has demonstrated transformative potential across various sectors, its application remains underexplored in enhancing the efficiency of construction cost estimating.

The objective of this study is to evaluate the applicability of ChatGPT-4 in construction cost estimating using a case study of a real-world project. This study will examine accuracy and reliability of applying ChatGPT-4 in estimating costs and determining bid prices for specific construction items. By comparing the AI-generated and actual bid data, this study seeks to establish an understanding of ChatGPT-4's potential as a tool for enhancing the speed and efficiency of cost estimating in construction.

4. Case project

To assess ChatGPT-4's potential in aiding in construction cost estimating and bid pricing, a real project was selected. This project was to rehabilitate Beaver Dam Road bridge, located in Lawrence County, Pennsylvania that started on 07/26/2021 (award date) through 12/08/2021 (completion date). As seen in Figure 1, the project entailed a rehabilitation of the bridge including removal of a portion of the existing bridge, reinforcement bar work, concrete work and formwork, structural backfill and improvements in drainage systems, and various other construction activities. The details of the scope of work for this project are provided on the Engineering and Construction Management System (ECMS) website [33] (*i.e.*, project's ID: 29541), which lists 86 distinct items, each described with its respective units and bid quantities. The ECMS provides information regarding the lowest, second, and third lowest bidders for each listed item.

Figure 1. Images of Beaver Dam Road Bridge project, Lawrence County, Pennsylvania, ID: 29541, including, (a) satellite image, (b) before rehabilitation, (c) after rehabilitation, (d) demolition, (e) reinforcement bars, (f) erosion control, (g) structural backfill, (h) concrete work, and (j) saw cutting.

Figure 2 provides a screenshot of the ECMS' first page, in which it can be seen one of the items selected for this research. The item, described as 'STRUCTURE BACKFILL COARSE AGGREGATE NO. 57' and identified by item number 0205-0297, is presented with detailed bid information. This includes the unit bid price of \$42.00, the total bid price of \$3444.00, and the bid quantity of 82.00 Cubic Yard.

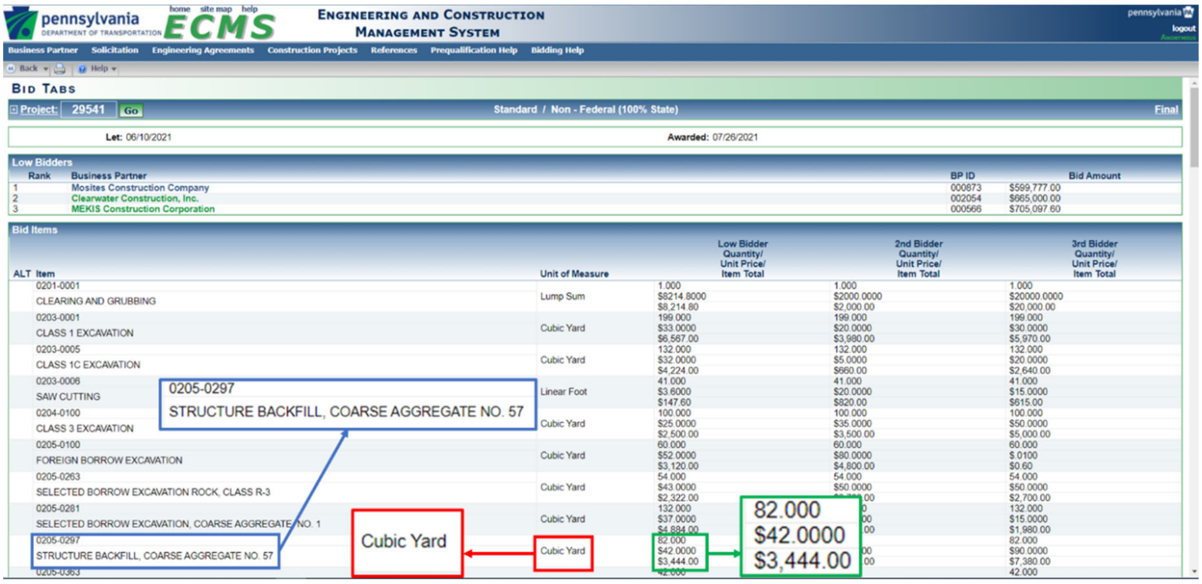

Figure 2. The bid prices from the lowest, second, and third bidders for Beaver Dam Road bridge.

5. Design and implementation

A case study was designed in which ChatGPT-4 was examined to perform cost estimating and bid pricing for five items of the selected project namely, concrete work, reinforcement work, structure backfill coarse aggregate No. 57, membrane waterproofing system installation, and borrow excavation. This selection among 86 distinct items was made to cover a broad spectrum of construction activities. Nevertheless, it is important to acknowledge that in this project, as in many others, the multitude of construction items makes it challenging to provide an exhaustive cost estimation for the entire project. Quantities of these five work items in terms of cubic yard (*i.e.*, concrete work, formwork and structural backfill), pound (*i.e.*, reinforcement work), square yard (*i.e.*, membrane waterproofing system installation), and ton (*i.e.*, borrow excavation) were obtained through calculation by an industry expert. These quantities were used as input for use of ChatGPT-4 through developing a knowledge base and instructions that allows ChatGPT-4 for an interactive process in cost estimating and bid pricing. Evaluation was conducted to examine the performance of ChatGPT-4. The following details the procedure and methods used to perform the case study.

5.1. Development of interactive process for use of ChatGPT-4

Figure 3 illustrates the process of utilizing ChatGPT-4 for cost estimating. The process begins with defining the project scope, including quantity takeoff calculated by industry experts and bid quantity determined by ECMS. ChatGPT-4 then accesses its knowledge base, which contains information on production rates, labor wages, equipment costs, and material prices. After the user verifies the details provided by ChatGPT-4, such as the required crews, necessary equipment, and material rates, ChatGPT-4 calculates the costs for the project item. This calculation includes crew hours, unit rates, and total costs, and at the end estimates the bid price for the item.

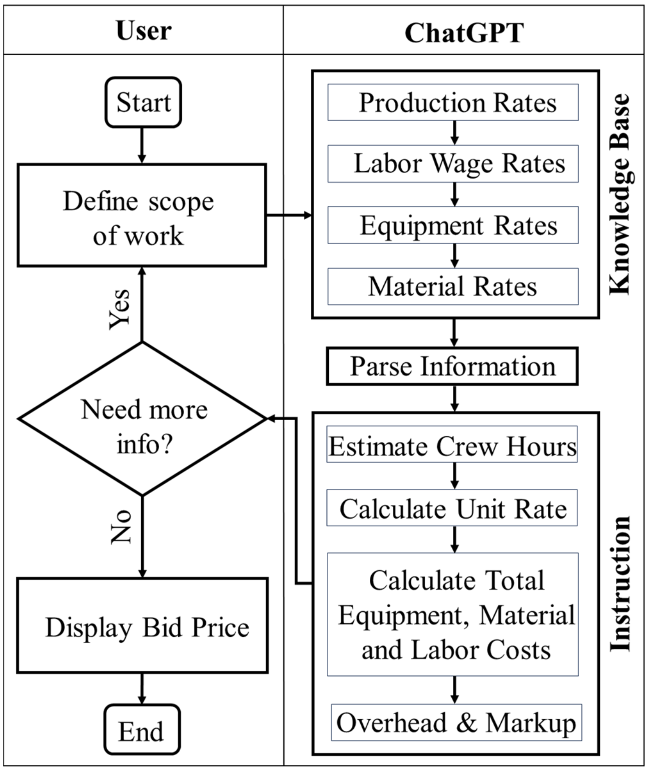

Figure 3. Workflow of integration of ChatGPT-4 with cost estimating.

In this study, a knowledge base was developed to enhance ChatGPT-4's performance. This knowledge base integrates information related to the scope of work for the project, encompassing production rates, hourly labor wage rates, equipment rates, and material rates for five construction items. To do so, the relevant data gathered from different sources were uploaded into the knowledge base in configure tab in ChatGPT-4.

To ensure ChatGPT-4's performance in construction cost estimating and bid pricing, detailed prompts and instructions were implemented. We employed a standard construction estimating method for estimating the Total Bid Price and Unit Bid Price, by referencing major construction books such as "Construction Management - Fourth Edition" by Daniel W. Halpin [34]. First, the unit prices for each construction item were collected. In this study, the unit prices were determined by collecting rates for equipment and material rates helped by the Construction Association of Western Pennsylvania (CAWP) , as well as labor wages derived from the Bureau of Labor Law Compliance Prevailing Wages Project Rates [35].

Second, specifying the production rates for each construction item, which sourced from a well-known and credible contractor. These rates include the number and type of personal required (*i.e.*, crew composition) and equipment for each specific item. Next, the total costs for each construction item were computed by summing the labor, material, and equipment costs while considering any material wastage and taxes. The following step-by-step instructions were established to navigate ChatGPT-4 through the cost estimating and bid pricing process.

Step 1. Input Request: "*Prompt ChatGPT-4 to ask the user for the bid quantity and the user's calculated quantity take-off, specifying the item.*"

Step 2. Display Relevant Details: "*For the user-specified item, present detailed information from three pre-uploaded tables within the GPT's database. This includes the production rate, hourly labor costs, equipment rates, and the materials required for the specified item.*"

Step 3. User Confirmation: "*Instruct ChatGPT-4 to seek the user's consent to display the information by posing a straightforward question, such as 'Do you agree with this information?' Based on the user's response, ChatGPT-4 should offer options to add, delete, or modify any of the details provided.*"

Step 4. Crew Hours Estimation: "*Instruct ChatGPT-4 to estimate the Crew Hour for the specified item by accessing the provided data, using Equation 1. The instruction specifies not to round up to the nearest hour to maintain precision.*"

$$Crew\,Hour = \frac{Takeoff\ Quantity}{Production\ Rate} \tag{1}$$

Step 5. Labor Cost Calculation: "*Calculate the total labor cost for specified item based on the production rate. The AI is directed to refer to the hourly labor wage rate table for the unit rate, using Equation 2.*"

$$Unit\,Rate = \left(Hourly\,Rate + Fringe\,Benefits + 0.25 \times Hourly\ rate\right) \times Crew\ Hour \tag{2}$$

Step 6. Equipment Cost Determination: "*Direct ChatGPT-4 to the equipment rate table to select the necessary equipment for each item and to calculate the total equipment cost, using Equation 3, incorporating the equipment rate and tax.*"

$$Equipment\ Cost = No.\ Equipment \times Unit\ rate \times Taxes \tag{3}$$

Step 7. Material Cost Computation: "*Identify the material rate, wastage, and tax for each item from the Required Material table and to compute the total material cost, using Equation 4, ensuring unit consistency throughout the process.*"

$$Material\ Cost = Quantity\ of\ Material \times Unit\ rate \times Taxes \tag{4}$$

Step 8. Cost Summary Presentation: "*Display the Total Labor, Total Equipment, and Total Material Cost in a structured format. Please seek the user's consent to display the* *information by posing a straightforward question, such as 'Do you agree with this information?' Based on the user's response, ChatGPT-4 should offer options to add, delete, or modify any of the details provided. Prompt ChatGPT-4 to ask the user for the Markup and Overhead percentage for following calculations."*

Step 9. Final Costing Table: *"Present the cost estimations in a detailed table format, including columns for Total Labor Cost, Total Equipment Cost, Total Material Cost, Total Cost (Equation 5), Overhead, Markup, Unit Cost (Equation 6), Total Bid Cost (Equation 7), Total Bid Price (Equation 8), and Unit Bid Price (Equation 9). This structured presentation facilitates easy interpretation of the cost components and the final estimations."*

$$Total\ Cost = Total\ Laborer\ Cost + Total\ Equipment\ Cost + Total\ Material\ Cost \tag{5}$$

$$Unit\ Cost = \frac{Total\ Cost}{Takeoff\ Quantity} \tag{6}$$

$$Total\ Bid\ Cost = Unit\ Cost \times Bid\ Quantity \tag{7}$$

$$\begin{aligned} Total\ Bid\ Price = {} & Total\ Bid\ Cost + (Total\ Bid\ Cost \times 8\%\ Markup) \\ & + (Total\ Bid\ Cost \times 10\%\ Overhead) \end{aligned} \tag{8}$$

$$Unit\ Bid\ Price = \frac{Total\ Bid\ Price}{Bid\ Quantity} \tag{9}$$

5.2. Implementation

To evaluate ChatGPT-4's performance in cost estimating and bid pricing, ten trials for each project item were considered. The inputs for each item were identical, and ChatGPT-4 was asked to estimate the cost of the specific item and bid pricing. Parameters used for evaluation include accuracy, relevance, and consistency. Measuring accuracy was achieved by comparing the project cost estimates and bid prices generated by ChatGPT-4 to the ground truth by manual calculation and bid prices from the three lowest bidders from ECMS. The relevance was evaluated by examining the alignment of ChatGPT-4's responses with the project's specific scope of work. Consistency was evaluated by examining the uniformity of ChatGPT-4's outputs with repetition using the same inputs.

6. Results

6.1. Input preparation results

After implementation of the manual quantity take-off process by the industry expert, the authors have derived a dataset that will be utilized as an input for ChatGPT-4's application. Table 1 demonstrates the results of this process.

In Table 1, it can be seen that there are differences between the takeoff quantity calculated by industry expert and the bid quantity presented by ECMS. For instance, for concrete work and formwork, the takeoff quantity was 2 cubic yards, a modest underestimation when compared to the bid quantity of 3 cubic yards. The estimation for reinforcement work was similar, with a quantity takeoff of 20,107 pounds close to the bid quantity of 20,115 pounds. On the other hand, the structure backfills presented a greater discrepancy, where the estimated 58.36 cubic yards fell short of the bid's 82 cubic yards. For the membrane waterproofing system installation, the industry expert's estimate was 25.34 square yards, again lower than the bid's 36 square yards. Lastly, for borrow excavation, the takeoff quantity was 11.18 tons which is lower than bid quantity of 42 ton.

Table 1. Summary of the quantity take-off and bid quantity of five construction items.

Item	Takeoff Quantity	Bid Quantity	Unit
Concrete work & formwork	2	3	CY
Reinforcement work	20,107	20,115	lb.
Structure backfills	58.36	82	CY
Membrane waterproofing system installation	25.34	36	SY
Borrow excavation	11.18	42	Ton

It should be noted that the process of quantity takeoff involves interpreting blueprints and specifications, which can vary based on the individual's experience, expertise, and judgment. Individuals have different methodologies in their calculations, leading to variations in the estimated quantities compared to the bid quantities. These differences are natural and expected. Despite the variations between the quantity takeoffs conducted by the industry expert and the bid quantities, they will not impact the conduct of this research, as they were used only inputs for ChatGPT-4 to perform the estimation.

6.2. Uploaded documents as knowledge base for ChatGPT

In this study, the information uploaded to ChatGPT-4 pertains to the project's scope of work, which includes production rates, hourly labor wages, equipment rates, and material costs.

The production rates for five bid items were sourced from a well-known contractor, which include the number and type of personnel required (*i.e.*, crew composition) for each specific item. The production rates also include the type and rates of equipment required to support the crew in meeting the productivity benchmarks. For example, 'Structure backfills' requires one laborer and two operators to meet the production rate of 75 CY and one Cat D- 6 dozer and one vibratory roller. Table 2 provides information for the production rates, crew composition and equipment needed for the five construction items.

The labor wage rates incorporated into the knowledge base in ChatGPT-4 are derived from the Bureau of Labor Law Compliance Prevailing Wages Project Rates. Table 3 presents the hourly labor wage rates, including fringe benefits and insurance & tax considerations, for the crews involved in five construction items. For instance, the labor foreman receives an hourly wage of $28.06, with additional fringe benefits of $24.80 and insurance & tax considerations amounting to $7.02.

Table 2. Production rates.

Item	Production Rates	Crews	Equipment
Concrete work	10 CY/5Men	1 Laborer Foreman 3 Laborers 1 Cement Finisher	5-kw portable generator Concrete Truck Pump Pickup Truck
Formwork	30 SF/6Men	1 Carpenter Foreman 3 Carpenters 1 Laborer 1 Operator	5 kw portable generator Rough Terrain Crane Pickup Truck
Reinforcement work	300 LB/5Men	1 Laborer Foreman 3 Laborers 1 Operator	Rough Terrain Crane Pickup Truck
Structure backfills	75 CY/3Men	1 Laborer 2 Operators	Cat D-6 Vibratory Roller
Membrane waterproofing system installation	4 SY/3Men	1 Laborer Foreman 2 Laborers	Pickup Truck
Borrow excavation	56 ton/3Men	2 Laborers 1 Operator	Cat 320/336 Excavator Dump Truck Pickup Truck

Table 3. Hourly labor wage rates for five construction items.

Item	Crews	Hourly Labor Wage Rate		
		Hourly Rate	Fringe Benefits	Insurance & Tax
Concrete work	1 Laborer Foreman	$28.06	$24.80	$7.02
	3 Laborers	$27.06	$24.80	$6.77
	1 Cement Finisher	$32.84	$22.60	$8.21
Formwork	1 Carpenter Foreman	$37.12	$19.32	$9.28
	3 Carpenters	$36.12	$19.32	$9.03
	1 Laborer	$26.90	$24.80	$6.77
	1 Operator	$33.89	$22.73	$8.47
Reinforcement work	1 Laborer Foreman	$28.06	$24.80	$7.02
	3 Laborers	$27.06	$24.80	$6.77
	1 Operator	$33.89	$22.73	$8.47
Structure backfills	1 Laborer	$27.06	$24.80	$6.77
	2 Operators	$33.89	$22.73	$8.47
Membrane waterproofing system installation	1 Laborer Foreman	$28.06	$24.80	$7.02
	2 Laborers	$27.06	$24.80	$6.77
Borrow excavation	2 Laborers	$27.06	$24.80	$6.77
	1 Operator	$33.89	$22.73	$8.47

Table 4 outlines the equipment rates for the necessary machinery and vehicles used in the five construction items alongside the 6% tax. They were collected from the CAWP, which provides comprehensive data on rental costs and operational expenses for various types of construction equipment. For instance, under Concrete work, one 5-kW portable generator, one concrete truck pump, and one pickup truck are listed with an hourly rate of $7.00, $90.00, and $8.50, respectively.

Table 4. Equipment rates for five construction items.

Item	Equipment	Rate (US$/hr.)	Tax (%)
Concrete work	5-kw portable generator	7.00	6
	Concrete Truck Pump	90.00	6
	Pickup Truck	8.50	6
Formwork	5 kw portable generator	7.00	6
	Rough Terrain Crane	75.00	6
	Pickup Truck	8.5	6
Reinforcement work	Rough Terrain Crane	75.00	6
	Pickup Truck	8.50	6
Structure backfills	Cat D-6	70.00	6
	Vibratory Roller	40.00	6
Membrane waterproofing system installation	Pickup Truck	8.50	6
Borrow excavation	Cat 320/336 Excavator	65.00	6
	Dump Truck	100	6
	Pickup Truck	8.5	6

Table 5 illustrates the required materials for each construction item, including material rates, wastage percentages, and tax considerations which was collected from CAWP. For instance, the cost is $128.50 per cubic yard for material concrete class A in concrete work, and $20 per ton for R-3 rock in borrow excavation work.

Table 5. Required materials for five construction items.

Item	Material rate	Rate	Wastage (%)	Tax(%)
Concrete work	Concrete Class A	$128.50/CY	5	6
Formwork	Concrete Curing Material	$1/SF	5	6
	General Form Lumber	$2/SF	5	6
Reinforcement work	Grade 60, Epoxy Coated Bars	$0.70/LB	5	6
Structure backfills	Coarse Aggregate No.57	$20/Ton	5	6
Membrane waterproofing system installation	Waterproofing Membrane Materials	$23.40/SY	5	6
Borrow excavation	R-3 Rock	$20/Ton	5	6

6.3. Case study results

We collected data from ChatGPT-4 on several cost-related parameters, namely, total labor cost, total material cost, total equipment cost, total cost, total bid price, and unit bid price. Detailed results about these experiments are provided in Appendix A. The average results of ChatGPT-4's performance, alongside the ground truth data, are presented in Table 6. In this research, the term "ground truth" is designated to the cost calculations derived from industry expert's analyses.

In Table 6, it can be seen that for 'Concrete work & Formwork,' the model showed a good level of accuracy, with a MAE of just 25.19 for Total Unit Price and an MAPE of 1.64%. In the 'Reinforcement work', the performance was mostly consistent, except for the Total Material Cost where the model had a higher MAE of 674.24 and a MAPE of 4.51%. In 'Structure backfills', ChatGPT-4 posed a challenge for estimating the Total Labor Cost, where the MAE reached 16.66 and the MAPE was 11.29%. For this item, The Total Labor Cost estimation varies across the ten trials, from as low as $105.78 to as high as $169.37 (see Appendix A).

Table 6. ChatGPT-4's overall performance for ten trials for each construction items.

Item	Description	Ground truth	Mean	Std	Range	MAE	MAPE
Concrete work & Formwork	Total Labor Cost	1557.37	1571.82	34.75	106.14	23.62	1.51
	Total Equip. Cost	398.89	395.22	5.66	18.17	4.84	1.21
	Total Material Cost	641.3	639.55	17.92	61.02	14.16	2.20
	Total Cost	2597.55	2624.99	50.90	193.18	44.45	1.71
	Total Bid Price	4597.67	4646.24	90.10	341.93	78.68	1.71
	Unit Bid Price	1532.56	1547.14	30.60	113.98	25.19	1.64
Reinforcement work	Total Labor Cost	20,164.64	20,142.69	71.66	241.34	35.18	0.17
	Total Equip. Cost	5932.24	5867.92	122.99	304.33	66.51	1.12
	Total Material Cost	14,919.39	15,593.16	243.13	882.29	674.24	4.51
	Total Cost	41,016.27	41,603.77	318.19	1046.31	650.09	1.58
	Total Bid Price	48,418.85	49,117.96	389.94	1305.88	777.88	1.60
	Unit Bid Price	2.41	2.43	0.018	0.06	0.033	1.38
Structure backfills	Total Labor Cost	147.74	139.82	20.32	63.59	16.66	11.29
	Total Equip. Cost	90.73	90.21	1.67	5.67	0.68	0.75
	Total Material Cost	1855.85	1832.65	39.15	125.13	42.65	2.29
	Total Cost	2094.31	2062.69	35.44	133.35	43.68	2.08
	Total Bid Price	3472.34	3311.25	321.99	1146.45	180.95	5.21
	Unit Bid Price	42.35	40.38	3.92	13.98	2.20	5.21
Membrane waterproofing system installation	Total Labor Cost	1118.29	1128.56	22.97	68.65	13.24	1.18
	Total Equip. Cost	57.08	57.06	0.098	0.36	0.062	0.10
	Total Material Cost	628.53	659.34	2.63	8.26	30.81	4.90
	Total Cost	1803.9	1844.85	24.12	74.23	40.95	2.27
	Total Bid Price	3024.06	2635.82	435.02	1214.56	479.60	15.85
	Unit Bid Price	84.00	71.85	12.76	33.73	14.68	17.48

Table 6. *Cont.*

Item	Description	Ground truth	Mean	Std	Range	MAE	MAPE
Borrow excavation	Total Labor Cost	39.01	39.01	10.11	33.97	6.495	16.64
	Total Equip. Cost	39.16	38.45	1.04	3.17	0.94	2.42
	Total Material Cost	237.02	242.51	6.56	14.91	6.52	2.75
	Total Cost	315.19	320.00	12.62	41.45	11.38	3.61
	Total Bid Price	1397.22	1298.41	350.25	1148.47	182.33	13.04
	Unit Bid Price	33.27	30.91	8.34	27.35	4.34	13.04

In the item, 'Membrane waterproofing system installation', the model faced challenges in estimating the Total Bid Price and the Unit Bid Price. The MAE for these estimations was high, with the Total Bid Price experiencing an MAE of 479.60, and for the Unit Bid Price, an MAE of 14.68. Furthermore, the MAPE for these categories reached 15.85% and 17.48% respectively. However, the model demonstrated a more reliable performance in estimating the total cost, where the MAPE was low at 2.27%. Similarly, the 'Borrow excavation' item showed variability in the Total Bid Price and Unit Bid Price, where the MAE was recorded at 182.33 and 4.34, respectively. Both exhibited a MAPE of 13.04%. These discrepancies can be attributed to a miscalculation in the Total Labor Cost, especially notable in the 10th trial (See Appendix A). In this trial, ChatGPT-4 calculated the Total Labor Cost to be $13.04, whereas the ground truth indicated it should have been $39.01.

7. Discussion

7.1. Impact of knowledge base and instruction in ChatGPT-4's performance

At first, the research began with an effort to create custom GPTs by having a conversation with ChatGPT-4. The first prompt is:

"*I want to create a GPTs for construction estimating and bid pricing*".

After a couple of back-and-forth conversions between ChatGPT-4 and the authors, all questions were answered. These questions were:

"*Great choice! How about we name it Construct Estimator? Does that work for you, or do you have another name in mind?*"

"*Here's the initial profile picture for Construct Estimator. Do you like it, or would you like any changes?*"

"*When Construct Estimator provides estimates, how detailed should the breakdown be? For example, should it include categories like materials, labor, machinery, and overheads, or do you prefer a more general overview?"*

"How would you like Construct Estimator to communicate? Should it use technical jargon familiar to construction professionals, or prefer simpler language to be accessible to a broader audience, including those new to the field?"

"Try out Construct Estimator in the preview on the right. Feel free to ask for any refinements or additional features you'd like to see. How can I further tailor this GPT to your needs?"

In the first initial attempt, the authors focused on concrete work, compiling all pertinent data such as hourly labor wages, equipment rates, and material requirements. However, the outcomes were found to be inaccurate, inconsistent, and not relevant. Following several attempts to refine the scope of work, the authors decided to revise their approach due to the original process being overly time-consuming and complex. Additionally, interactions with ChatGPT-4 occasionally led to circular discussions, hindering the ChatGPT-4's ability to function effectively. Thus, the authors moved to the configuration section of the GPTs and developed a specialized version named 'Construction Estimator'.

In the Construction Estimator's knowledge base, a PDF file has been uploaded to enhance the conversation between GPTs and the user. This document is divided into three primary sections: the production rate of five construction items, hourly laborer wage rates (*i.e.*, hourly rate, fringe benefit, insurance, and tax), equipment rates, inclusive of tax, and a comprehensive inventory of materials, along with their costs, waste percentages, and taxes. After uploading information and establishing instructions into the Construction Estimator, the application's performance was improved. It should be noted that in this study, human interaction with ChatGPT-4 is primarily in the process of inputting takeoff quantities and bid quantities into the system. The cost-related component estimating process is then computed by ChatGPT-4 and configured with prompts to ensure consistency. We followed standardized procedures for quantity takeoff, unit pricing, and cost component calculations to minimize subjective bias based on well-established industry practices and reference materials. The data used for labor wages, equipment rates, and material costs were also sourced from reliable and authoritative sources. Throughout the process, ChatGPT-4 prompted the user for confirmation and allowed adjustments to the information provided, ensuring any discrepancies or errors could be identified and corrected.

7.2. Accuracy and consistency

In terms of accuracy, the overall accuracy was around 99%. However, the model showed some variations in the responses which might also indicate inherent limitations in ChatGPT-4's ability to apply constant numerical estimation processes across different trials. These limitations could stem from the model's training data or its algorithmic structure, which might not always align perfectly with the demands of construction cost estimation. ChatGPT-4 has exhibited the potential to represent fragmented information as human-understandable sentences. However, ChatGPT-4 is very sensitive to prompts and some changes to the prompts may affect the querying result quality.

In ten trials where ChatGPT-4 was tasked with providing identical input for estimating the cost for each construction item, it was observed that the model did not always present the expected numerical outputs (see Appendix A). Instead, it sometimes presented only the conceptual framework and calculation methodology without the quantitative data (*i.e.*, mentioned with N/A in Appendix A). This inconsistency was observed once for tasks concrete work and formwork, reinforcement bars, structural backfill, and the installation of membrane waterproofing systems, and it occurred twice for borrow excavation. Even though the model corrected its errors after the authors pointed out the issue, the authors decided to keep these errors in the analysis to show some inconsistency in ten experiments. This inconsistency can firstly be attributed to an understanding of context and instructions. ChatGPT-4's responses reflect its understanding of the task's context and instructions. If the model interprets the request as seeking a methodological explanation rather than specific numerical results, it might respond with a conceptual framework or calculation methodology.

7.3. Reliability

In the established procedure, a feedback mechanism in Step 8 was implemented (*i.e.*, human-in-the-loop) to refine ChatGPT-4's calculations. In this step, after ChatGPT-4 generates the initial cost estimates, the model presents the results to the user in a structured format, prompting the user to review and confirm the information. The user is asked a straightforward question like, "Do you agree with this information?" This interaction allows the user to engage in the process, providing an opportunity to add, delete, or modify details before the final calculations are made. Once these modifications are completed, ChatGPT-4 recalculates the total costs, including the revised labor, equipment, and material costs, to ensure that the final estimates are accurate. Despite these adjustments, in the experiments, we still identified some considerable differences between ChatGPT-4's results and the ground truth such as the 10th trial for Total Labor Cost for Borrow Excavation Work. As a result, further investigation is still needed to enable this ChatGPT-4 entailed technique to become more consistent and reliable in estimating construction costs.

7.4. Comparison with real-world bids

In the realm of construction project management, the Total Bid Package is a critical aspect that encapsulates the comprehensive cost a contractor proposes to complete a project. While individual task costs—such as labor, materials, and equipment—are essential components of the overall bid, focusing on the Total Bid Package offers an advantage for comparative analysis and decision-making. The Total Bid Package provides a holistic view of the project's cost, enabling stakeholders like those from PennDOT, to make informed decisions based on the financial commitment required. This is useful in scenarios where decision-makers need to assess and compare the financial viability of multiple bids quickly. In this research, we calculate the Total Bid Package for ChatGPT-4 by summing up the average Total Bid Prices of each construction item, which ChatGPT-4 has already estimated. Table 7 presents the details of how well the model performs compared to ground truth data and bids from the three lowest bidders. Acknowledging that each contractor's bid varies due to unique specifications and rates, comparing ChatGPT-4's estimates with the three lowest bidders offers insights that show how ChatGPT-4's estimates match up with actual bids. By looking at the range of bids, we see where ChatGPT-4 stands in the market, helping us understand its capability.

Table 7. Total costs for five construction items (ChatGPT-4's, ground truth, and the first, second, and third lowest bidders from PennDOT).

Item	ChatGPT-4	Ground Truth	Lowest Bidder	2nd Lowest Bidder	3rd Lowest Bidder
Concrete work & Formwork	4646.24	4597.67	9300.00	13,500.00	7500.00
Reinforcement Bar	49,117.96	48,418.85	36,207.00	40,230.00	60,345.00
Structural Backfill	3311.25	3472.34	3444.00	4920.00	7380.00
Membrane Waterproofing System Installation	2635.82	3024.06	2628.00	1008.00	1440.00
Borrow Excavation	1298.41	1397.22	3444.00	4920.00	380.00
Total Bid Packages	61,009.68	60,910.14	55,023.00	64,578.00	84,045.00

By looking at Table 6 and Figure 4, it can be seen that the average Total Bid Package calculated by ChatGPT-4 amounts to \$61,009.68, which is close to the Ground Truth total of \$60,910.14 (*i.e.*, accuracy: 99.84%). The minor difference of just \$99.54 between ChatGPT- 4's estimation and the ground truth underscores the ChatGPT-4's precision in aggregating the total costs across different construction items. When we position ChatGPT-4's Total Bid Package in the landscape of actual bids from contractors, it stands notably as the second-lowest bid, surpassing the lowest bidder's total of \$55,023.00. This standing is primarily influenced by the variance observed in the Reinforcement Bar work estimates, where ChatGPT-4's estimate is \$49,117.96 compared to the lowest bidder's \$36,207.00. While ChatGPT-4 adheres to conventional estimation approaches, the lowest bidder might have employed specific methodologies, including unique production rates, composition crew and equipment costs, to arrive at their notably lower figure.

The adoption of AI-driven tools like ChatGPT-4 in the construction industry has the potential to significantly impact workforce skills and process automation. As AI tools become more integrated into construction practices, there might be a shift in the required skill sets, with an increasing demand for professionals who are not only knowledgeable in construction but also proficient in AI and data management. This shift could lead to changes in training and education to equip the workforce with the necessary skills to manage and interact with AI tools effectively. Furthermore, the potential for AI tools to automate various construction processes, such as cost estimating, scheduling, and resource management, could enhance efficiency across the industry. However, this also raises questions about the evolving role of human expertise and the balance between automated processes and traditional construction practices. As the industry adapts to these changes, it will be crucial to consider both the opportunities and challenges that come with the adoption of advanced AI technologies like ChatGPT-4.

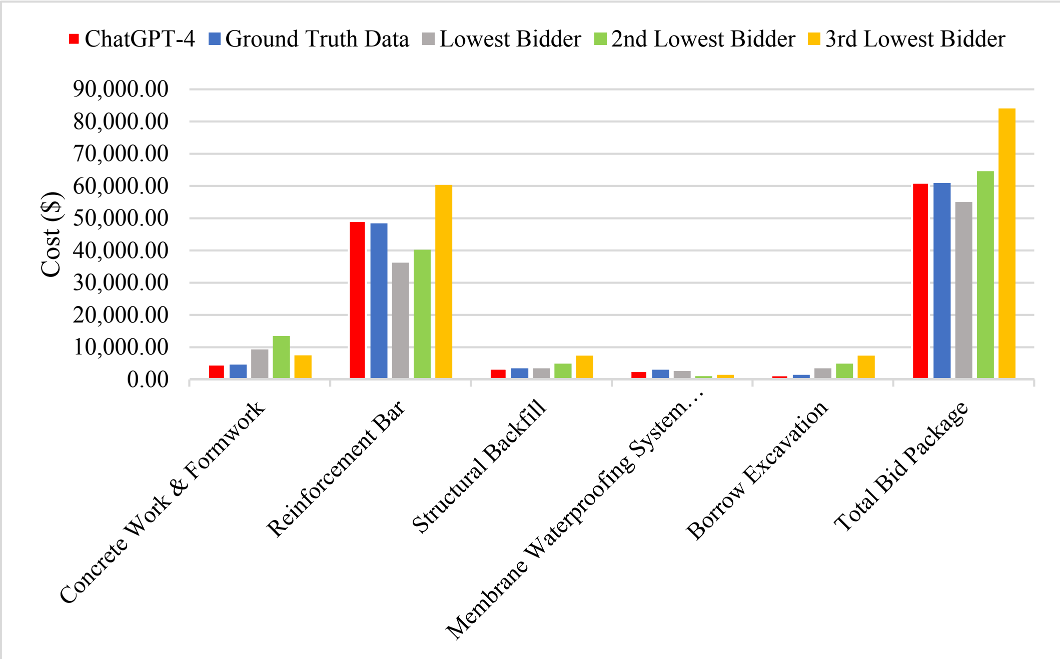

Figure 4. Comparison of ChatGPT-4's performance with ground truth data and three lowest bidders.

7.5. Limitations

One limitation identified in this study is the dependency of ChatGPT-4 on the structure of the data in the knowledge base to generate accurate estimations. Initially, when the data is sparse, ChatGPT-4's ability to provide reliable insights is constrained. Furthermore, the scope of work in this study is limited. We selected five construction items from the project's 86 distinct items, aiming to encompass a variety of construction activities, units, and crew compositions. However, expanding the range of construction activities would enhance the generalizability of our findings. It is tried to focus on these representative items to provide meaningful insights. Future research will aim to include a broader range of construction activities to further enhance the generalizability of the results. Additionally, the process of construction cost estimating in this study only considered one set of production rates, labor, equipment and material consumption rates for each construction item. Different construction methods can impact the quantities of crew compositions, equipment and materials required for a project. For instance, choosing between traditional and advanced modular formwork methods can result in varying amounts of concrete, reinforcement, and formwork materials. To account for these variations, our method can be adapted to include different sets of production rates, labor, equipment, and material consumption rates tailored to specific construction methods. This can be achieved by creating separate knowledge bases for each construction method within the ChatGPT-4 framework. By doing so, we can input the appropriate method-specific data into ChatGPT-4, ensuring that the cost estimation and bid pricing reflect the actual material quantities needed for each construction item. This adaptability enhances the accuracy and reliability of our cost estimates.

8. Conclusions

This research conducted a case study to evaluate ChatGPT-4 using the five construction items involved in a real construction project. The findings provide a preliminary understanding of the potential of applying ChatGPT-4 in assisting in construction cost estimating and bid pricing. It has been found that while ChatGPT-4 may offer benefits in efficiency and accuracy, its performance is contingent upon data quality and model advances. This insight pointed to the need for a continuous improvement in this area, particularly the potential integration of ChatGPT-4, domain knowledge, and data sources to extend its applicability across various estimating tasks.

Acknowledgments

This work was sponsored by a grant from the Center for Integrated Asset Management for Multimodal Transportation Infrastructure Systems (CIAMTIS), a US Department of Transportation University Transportation Center, under federal grant number 69A3551847103. The authors are grateful for the support. Any opinions, findings, conclusions, and recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the CIAMTIS.

Conflicts of interests

The authors declare no conflict of interest.

Authors' contribution

Alireza Ghasemi: Investigation, Methodology, Visualization, Validation, Writing - Original Draft. Fei Dai: Conceptualization, Supervision, Writing - Review & Editing, Funding acquisition.

References

[1] Gambatese J, Hallowell M. Factors that influence the development and diffusion of technical innovations in the construction industry. *Constr. Manag. Econ.* 2011, (29):507–517.

[2] Saka A., Taiwo R., Saka N., Salami B. A., Ajayi S., *et al.* GPT models in construction industry: Opportunities, limitations, and a use case validation. *Dev. Built Environ.* 2023, 100300.

[3] Chowdhary K. R. and Chowdhary K. R. Natural language processing. *Found. Artif. Intell.* 2020, 603–649.

[4] Moon S., Lee G., Chi S., and Oh H. Automated construction specification review with named entity recognition using natural language processing. *J. Constr. Eng. Manag.* 2021, (147):04020147.

[5] Tayefeh Hashemi S., Ebadati O. M., and Kaur H. Cost estimation and prediction in construction projects: A systematic review on machine learning techniques. *Discov. Appl. Sci.* 2020, (2):1703.

[6] Monteiro A. and Martins J. P. A survey on modeling guidelines for quantity takeoff- oriented BIM-based design. *Autom. Constr.* 2013, (35):238–253.

[7] Saeidlou S. and Ghadiminia N. A construction cost estimation framework using DNN and validation unit. *Build. Res. Inf.* 2023, 1–11.

[8] Dang-Trinh N., Duc-Thang P., Nguyen-Ngoc Cuong T., and Duc-Hoc T. Machine learning models for estimating preliminary factory construction cost: case study in Southern Vietnam. *Int. J. Constr. Manag.* 2023, (23):2879–2887.

[9] Tang S., Liu H., Almatared M., Abudayyeh O., Lei Z.*, et al.* Towards Automated Construction Quantity Take-Off: An Integrated Approach to Information Extraction from Work Descriptions. *Buildings*. 2022, (12):354.

[10] Karan E., Mansoob V. K., Khodabandelu A., Asgari S., Mohammadpour A.*, et al.* Using Artificial Intelligence to Automate the Quantity Takeoff Process. In *Proceedings of the International Conference on Software Business Engineering,* Amsterdam, The Netherlands, 2021, pp. 13–14.

[11] Ambrule V. R. and Bhirud A. N. Use of artificial neural network for pre design cost estimation of building projects. *Int. J. Recent Innov. Trends Comput. Commun.* 2017, (5):173–176.

[12] Al-Tawal D. R., Arafah M., and Sweis G. J. A model utilizing the artificial neural network in cost estimation of construction projects in Jordan. *Eng. Constr. Archit. Manag.* 2021, (28):2466–2488.

[13] Yi Z. and Luo X. Construction Cost Estimation Model and Dynamic Management Control Analysis Based on Artificial Intelligence. *Iran. J. Sci. Technol. Trans. Civ. Eng.* 2023, 1–12.

[14] Wang R., Asghari V., Cheung C. M., Hsu S.-C., and Lee C.-J. Assessing effects of economic factors on construction cost estimation using deep neural networks. *Autom. Constr.* 2022, (134):104080.

[15] Wang B., Yuan J., and Ghafoor K. Z. Research on Construction Cost Estimation Based on Artificial Intelligence Technology. *Scal. Comput. Pract. Exp.* 2021, (22):93–104.

[16] Radford A., Wu J., Child R., Luan D., Amodei D.*, et al.* Language models are unsupervised multitask learners. *OpenAI blog*. 2019, (1):9.

[17] Neelakantan A., Xu T., Puri R., Radford A., Han J. M.*, et al.* Text and code embeddings by contrastive pre-training. *arXiv* 2022*,* arXiv:2201.10005.

[18] Zhang H., Song H., Li S., Zhou M., and Song D. A survey of controllable text generation using transformer-based pre-trained language models. *ACM Comput. Surv.* 2023, (56):1–37.

[19] Aluga M. Application of CHATGPT in civil engineering. *East Afr. J. Eng.* 2023, (6):104–112.
[20] Prieto S. A., Mengiste E. T., and García De Soto B. Investigating the use of ChatGPT for the scheduling of construction projects. *Buildings*. 2023, (13):857.
[21] Zhang C., Lu J., and Zhao Y. Generative pre-trained transformers (GPT)-based automated data mining for building energy management: Advantages, limitations and the future. *Energy Built Environ.* 2024, (5):143–169.
[22] Aladağ H. Assessing the Accuracy of ChatGPT Use for Risk Management in Construction Projects. *Sustainability*. 2023, (15):16071.
[23] Naser M. Z., Ross B., Ogle J., Kodur V., Hawileh R., *et al.* Evaluating the Performance of Artificial Intelligence Chatbots and Large Language Models in the FE and PE Structural Exams. *Pract. Period. Struct*. 2024, (29):02524001.
[24] Ray S. S., Peddinti P. R. T., Verma R. K., Puppala H., Kim B., *et al.* Leveraging ChatGPT and Bard: What does it convey for water treatment/desalination and harvesting sectors? *Desalination*. 2024, (570):117085.
[25] Rane N., Choudhary S., and Rane J. Integrating ChatGPT, Bard, and leading-edge generative artificial intelligence in architectural design and engineering: applications, framework, and challenges. *SSRN Electron. J.* 2023.
[26] Helfrich-Schkarbanenko A., *Mathematik und ChatGPT: Ein Rendezvous am Fuße der technologischen Singularität*, ed: Springer, 2023, pp. 213–230.
[27] Achiam J., Adler S., Agarwal S., Ahmad L., Akkaya I., *et al.* Gpt-4 technical report. *arXiv* 2023, arXiv:2303.08774.
[28] Plevris V., Papazafeiropoulos G., and Jiménez Rios A. Chatbots put to the test in math and logic problems: a comparison and assessment of ChatGPT-3.5, ChatGPT-4, and Google bard. *AI*. 2023, (4):949–969.
[29] Dao X and Le N. Investigating the effectiveness of chatgpt in mathematical reasoning and problem solving: Evidence from the vietnamese national high school graduation examination. *arXiv* 2023, arXiv:2306.06331.
[30] Wardat Y., Tashtoush M. A., Alali R., and Jarrah A. M. ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia J. Math. Sci. Technol. Educ.* 2023, (19):em2286.
[31] Blomhøj M. and Jensen T. H. Developing mathematical modelling competence: Conceptual clarification and educational planning. *Teach. Math. Its Appl.* 2003, (22):123–139.
[32] Zong M. and Krishnamachari B. Solving math word problems concerning systems of equations with gpt-3. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Washington, DC, USA, February 7–14, 2023, pp. 15972–15979.
[33] Penndot. Engineering and Construction Management System. 2024. Available: https://www.ecms.penndot.pa.gov/ECMS/ (accessed on 1 May 2024).
[34] Halpin D. W., *Construction management*, John Wiley & Sons, 2010.

[35] Penndoli. The Bureau of Labor Law Compliance. 2024. Available: https://www.dli.pa.gov/Individuals/Labor-Management-Relations/llc/prevailing-wage/Pages/default.aspx (accessed on 20 April 2024).

Appendix A.

Table A1. ChatGPT-4’s performance in cost estimation for five construction items across ten trials.

Item	No. Attempt	Total Labor Cost ($)	Total Equipment Cost ($)	Total Material Cost ($)	Total Cost ($)	Total Bid Price ($)	Unit Bid Price
Concrete Work & Formwork	1	1631.00	382.85	672.32	2686.68	4755.42	1585.14
	2	1636.89	388.97	664.32	2714.18	4804.10	1601.37
	3	1555.01	400.92	636.84	2618.77	4635.22	1545.07
	4	1555.02	401.02	639.05	2621.09	4639.33	1546.44
	5	1530.75	393.89	611.30	2521.00	4462.17	1487.39
	6	1554.04	396.06	635.46	2611.56	4622.46	1540.82
	7	1560.76	400.03	643.14	2629.93	4654.98	1551.66
	8	1550.76	396.05	631.30	2604.11	4609.27	1536.42
	9	N/A	N/A	N/A	N/A	N/A	N/A
	10	1572.22	397.22	622.22	2617.66	4633.26	1530.02
Reinforcement Bar	1	20,185.28	5632.08	14,917.28	40,734.64	48,064.38	2.39 $/lb.
	2	20,157.76	5936.41	15,665.43	41,759.60	49,370.26	2.45 $/lb.
	3	20,163.93	5932.21	15,663.36	41,759.50	49,276.21	2.45 $/lb.
	4	20,170.64	5932.22	15,629.36	41,732.22	49,132.90	2.44 $/lb.
	5	N/A	N/A	N/A	N/A	N/A	N/A
	6	20,177.44	5932.47	15,665.37	41,775.28	49,370.26	2.45 $/lb.
	7	20,166.02	5934.21	15,665.36	41,765.59	49,306.65	2.45 $/lb.
	8	20,183.37	5932.21	15,665.37	41,780.95	49,370.26	2.45 $/lb.
	9	19,943.94	5935.76	15,667.36	41,547.06	49,037.84	2.43 $/lb.
	10	20,135.85	5643.72	15,799.57	41,579.14	49,132.90	2.44 $/lb.
Structural Backfill	1	169.37	91.16	1818.58	2079.11	3447.56	42.04 $/CY
	2	137.33	90.77	1818.75	2046.85	2415.28	29.45 $/CY
	3	147.64	85.49	1818.26	2051.39	3402.08	41.49 $/CY
	4	105.78	90.83	1818.64	2015.25	3342.09	40.76 $/CY
	5	162.18	90.75	1818.73	2071.66	3434.98	41.89 $/CY
	6	147.76	90.60	1820.15	2058.51	3411.76	41.61 $/CY
	7	122.86	90.72	1818.64	2032.22	3369.18	41.09 $/CY
	8	114.49	90.72	1943.39	2148.60	3561.73	43.43 $/CY
	9	151.00	90.93	1818.73	2060.66	3416.59	41.67 $/CY
	10	N/A	N/A	N/A	N/A	N/A	N/A

Table A1. *Cont.*

Item	No. Attempt	Total Labor Cost ($)	Total Equipment Cost ($)	Total Material Cost ($)	Total Cost ($)	Total Bid Price ($)	Unit Bid Price
Membrane Waterproofing System Installation	1	1119.02	57.12	653.87	1830.01	2929.32	81.37 $/SY
	2	1174.94	57.08	661.08	1893.09	2233.85	62.05 $/SY
	3	1117.65	57.07	661.07	1835.80	2608.48	60.17 $/SY
	4	1118.19	56.93	659.96	1835.08	2165.40	60.15 $/SY
	5	1118.02	57.08	659.96	1834.06	2164.19	60.12 $/SY
	6	1106.29	57.23	655.34	1818.86	3378.75	93.85 $/SY
	7	1118.24	57.08	660.05	1835.37	2938.21	81.62 $/SY
	8	1166.72	57.08	660.65	1884.45	2223.66	61.77 $/SY
	9	N/A	N/A	N/A	N/A	N/A	N/A
	10	1117.98	56.87	662.13	1836.98	3080.52	85.57 $/SY
Borrow Excavation	1	39.02	39.16	248.49	326.67	1448.64	34.49 $/ton
	2	42.48	39.12	235.57	317.17	1430.94	34.07 $/ton
	3	45.36	37.03	247.98	330.37	1536.73	36.59 $/ton
	4	47.01	38.21	236.45	321.67	1435.56	34.18 $/ton
	5	39.79	37.91	250.18	327.88	1454.09	34.62 $/ton
	6	N/A	N/A	N/A	N/A	N/A	N/A
	7	42.94	40.13	235.27	318.34	1411.47	33.61
	8	42.46	36.96	249.62	329.04	388.26	9.24
	9	N/A	N/A	N/A	N/A	N/A	N/A
	10	13.04	39.11	236.77	288.92	1281.62	30.52