A Two-Phase AI-Driven Approach to Automated Construction Planning Using Small Language Models for Activity Sequencing and Missing Task Prediction

Akarsth Kumar Singh[1], Aritra Pal[2] and Shang-Hsien Hsieh[1]

[1]Department of Civil Engineering, National Taiwan University, Taiwan
[2]Department of Engineering, University of Cambridge, United Kingdom
akarsth@caece.net , ap2425@cam.ac.uk , shhsieh@ntu.edu.tw

Abstract

This research presents a two-phase AI-driven approach to enhance automated construction planning and scheduling, addressing common challenges such as project delays and budget overruns in the Architecture, Engineering, and Construction (AEC) industry, often caused by inefficient project management. Traditional AI methods rely heavily on static programming and manually annotated datasets, leading to inefficiencies. The proposed methodology uses prompt engineering to systematically extract and structure information from historical project schedules, creating datasets for training domain-specific Small Language Models (SLMs). In the first phase, fine-tuning SLMs, particularly mistral-7B, achieved 83.69% accuracy, outperforming traditional approaches. In the second phase, another fine-tuned SLM identified activity patterns, optimized sequencing, and predicted missing tasks, achieving a mean similarity score of 81.43%. The results demonstrate that structured prompt engineering significantly improves model accuracy and efficiency, reduces manual efforts, and addresses computational inefficiencies. This scalable, secure, and efficient approach sets a new standard in automated construction planning.

Keywords

Prompt Engineering; Domain-Specific Instruct Fine-Tuning; Small Language Models; Construction Planning and Scheduling; Activity Sequencing

1 Introduction

Construction delays and cost overruns are pervasive challenges in the AEC industry, significantly impacting the successful completion of projects [5]. Planning and scheduling are at the core of these issues and pivotal in determining project timelines and budgets. A deficiency in planning and scheduling processes can lead to delays, misallocated resources, and, ultimately, increased costs [7]. Inaccurate estimations, inadequate risk assessments, and failure to account for the complexities and interdependencies of construction activities often result in unrealistic and unachievable schedules. This lack of detailed and precise planning leads to inefficiencies in resource utilization, conflicts in task scheduling, and disruptions in the workflow, all of which contribute to extended project durations and escalated costs.

1.1 Challenges in AEC Project Planning and Scheduling

In the AEC industry, project planning and scheduling face significant challenges, notably in learning from documents, leveraging historical schedules, addressing fragmentation and interoperability issues, mitigating human errors in plan generation and updates, and implementing automation [4]. The intricate process of extracting actionable insights (key construction information and activity sequences) from extensive project documentation and past schedules is compounded by the diverse formats and standards, leading to fragmentation that hinders seamless data integration and interoperability across different project management tools. Furthermore, the reliance on manual input for generating and updating project plans introduces a high potential for human error, affecting the accuracy and reliability of schedules. These challenges highlight the need for advanced solutions to automate processes, improving efficiency and accuracy in project planning. Using sophisticated algorithms and machine learning to analyze historical data and automate schedule generation can greatly reduce these issues, leading to more streamlined and error-resistant project management in the AEC sector.

1.2 Need for Automated Planning and Scheduling

Automation in AEC project planning and scheduling is vital for optimizing project management practices [8].

Over the past three decades, researchers have explored various methods to reduce manual work and improve precision, including case-based reasoning, knowledge-based systems, model-based approaches, meta-heuristic algorithms, expert systems, and learning-based techniques [4]. However, challenges in data integration, platform interoperability, and adapting automation to project-specific needs persist. The complex and dynamic nature of construction projects demands more flexible solutions. Generative AI, with advanced Natural Language Understanding (NLU) and Natural Language Generation (NLG), presents a promising alternative [6], enhancing efficiency by interpreting complex documents and generating detailed schedules from large datasets. This marks a significant improvement in AEC automation in planning and scheduling.

2 Related Works

In the AEC industry, Building Information Modeling (BIM) and Natural Language Processing (NLP) have significantly enhanced project management by improving communication and information efficiency [11]. BIM revolutionizes the industry with digital building representations, facilitating automatic schedule generation and optimization, thereby reducing reliance on planners' expertise and using various data types like spatial and material information. Meanwhile, NLP, a sub-field of AI, focuses on computer interactions with human language through syntactic, semantic, and ontological analysis, accelerating information extraction from documents and interpreting construction processes. NLP's applications in the AEC sector include extracting data from BIM for safety compliance, analyzing contract documents, and automating Construction Schedule Management (CSM) [11], thereby improving schedule generation, validation, and accuracy.

2.1 Current Practices in Construction Planning and Scheduling

Recent efforts to automate construction scheduling have focused on leveraging BIM, with early methods combining BIM data with genetic algorithms [8] and incorporating graph-based indexing and AI for task scheduling [12], though BIM models often lack the granularity to capture project management aspects such as design approvals and procurement. More recent approaches employ NLP and deep learning to extract construction knowledge from historical schedules. Amer and Golparvar-Fard [2] introduced dynamic templates by learning from past schedules, while subsequent studies applied deep-learning-based NLP for sequence evaluation. Hong et al. [9] explored Graph Convolutional Networks to identify activity types and sequences.

2.2 Role of Language Models in the AEC Industry

LMs are revolutionizing planning and scheduling in the AEC industry by analyzing complex project data and identifying critical patterns for effective management [3]. They extract essential task details from project documents and past schedules [10], improving accuracy and reducing human error. Domain-specific training is essential [14], enabling LMs to handle construction-specific terminology and enhance project management applications. This training helps LMs identify tasks, understand resource needs, and predict challenges effectively. Integrating LMs into AEC workflows supports data-driven decisions, resource optimization, and delay reduction through historical data and predictive analytics [2][9]. Their performance improves with continued learning from real-world projects, promoting efficient project delivery.

This study addresses critical challenges in construction planning, particularly the extraction of structured information from historical project schedules and the generation of optimized activity sequences, including the prediction of missing tasks. To tackle these challenges, this research proposes a two-phase AI-driven framework integrating prompt engineering techniques with domain-specific fine-tuned SLMs. The paper is systematically organized, beginning with an introduction that outlines current industry challenges and emphasizes the need for automation in planning practices. It then reviews relevant literature, details the proposed methodology, presents experimental validations and analyses, and concludes by summarizing key findings with recommendations for future research directions.

3 Methodology

Figure 1 illustrates the framework proposed in this study for construction project planning using Generative AI. The approach leverages historical project schedules as the primary input. At its core is the extraction of key information, including detailed descriptions of the work breakdown structure, specific activities, tasks, and the network of predecessor-successor relationships. It captures the complex interplay of critical dependencies in construction planning. This study adopts a two-phase methodology that applies Generative AI and fine-tuned SLMs to enhance construction project planning and scheduling.

In Phase 1, prompt engineering is applied to extract and organize key information from existing and historical project schedules, resulting in a structured dataset. This dataset is then used to fine-tune an SLM for domain-specific information extraction in construction planning.

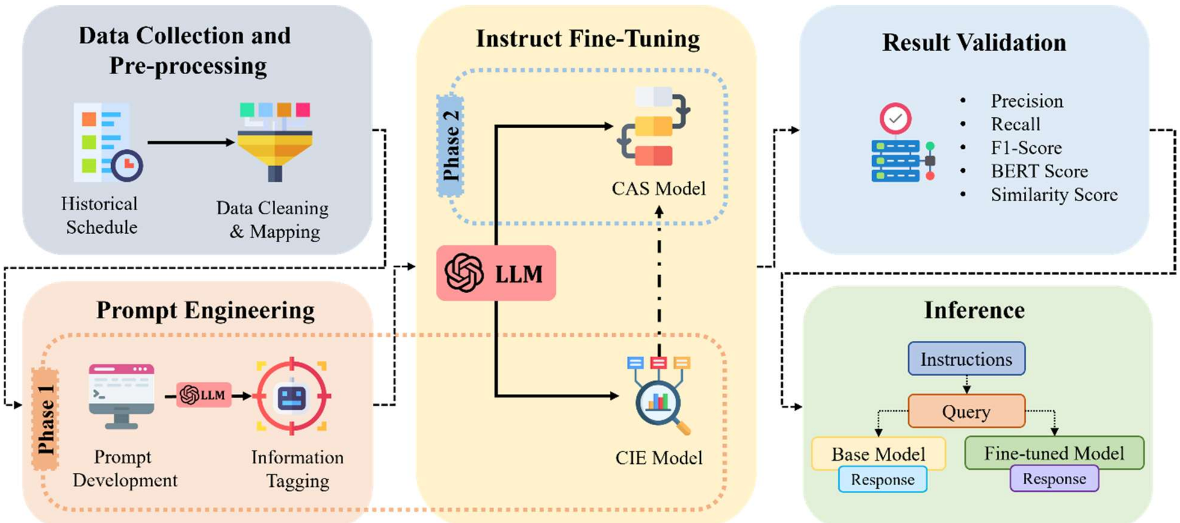

Figure 1. Overview of Proposed Methodology

In Phase 2, a separate SLM is trained on the organized data to learn activity patterns, enabling it to generate optimized activity sequences and identify missing tasks. The fine-tuned SLMs performance is evaluated against a benchmark model to assess its accuracy and reliability. Additionally, local deployment of the model ensures enhanced data privacy and security. This two-phase approach effectively reduces manual errors, accelerates processing, and improves the efficiency and accuracy of construction planning and scheduling.

3.1 Construction Activity Formalization for Data Preparation

In real-world construction projects, planning and scheduling have unique characteristics but share common structures, formats, and dependencies due to similar construction elements. Extracting information from historical schedules is challenging due to individualized approaches by schedulers, who use distinct terminologies based on specific project requirements [1]. For example, terms like “formwork” and “shuttering” serve the same function but differ in terminology. Work breakdown structures also vary; a cast-in-situ project may describe a superstructure task as “Construction-Super Structure-First Floor-Column-Reinforcement” under labels like “EPC Type-Work Item-Location-Object-Task,” while a firefighting task may be labeled as “Construction-MEPF (Fire Fighting work)-Firefighting pipes-Fix or Installation-First Floor” under “EPC Type-Work Item-Object-Task-Location.” This variation underscores the lack of uniformity in construction labels.

This study presents a structured approach to formalize construction activity definitions to improve data preparation for downstream processing. It consolidates details from diverse work breakdown structures and maps them to standardized labels “EPC Type, Work Items, Location, Object, and Task,” emphasizing Location, Object, and Task. By harmonizing inconsistent terminologies and structures, this method improves the uniformity and usability of construction schedule data, enhancing the efficiency and accuracy of data analysis and model training.

3.2 Prompt Engineering and Key Construction Information Extraction

Prompt engineering is a vital technique for advanced models like GPT, involving the crafting of specific queries to guide the model in generating relevant outputs. It enhances precision and fine-tuning capabilities, enabling refined interaction and alignment with specific objectives. In the AEC industry, prompt engineering is a powerful tool for streamlining data retrieval [14], particularly when processing historical scheduling data. It automates the extraction and tagging of construction labels from past schedules, transforming a labor-intensive task into an efficient process. Effective prompt design requires understanding both model mechanics and domain-specific knowledge. In this study, prompts are strategically designed to capture relevant data and the complexities of construction project management. Prompt engineering, by targeting tasks, objects, and locations, ensures precise data extraction and categorization, forming a core part of our methodology.

3.3 Instruct Fine-Tuning SLMs for Construction Planning

In this study, we apply instruct fine-tuning to SLMs, tailoring these advanced AI systems to meet the unique requirements of the construction industry. Although LMs are inherently powerful, their generalist training needs more depth for specialized tasks [13]. Our methodology strategically guides these models with carefully crafted prompts from prompt engineering data, enabling them to produce outputs that accurately align with the nuanced demands of construction management. This fine-tuning enhances the LMs proficiency in interpreting task-specific data [6], significantly improving their performance in complex construction tasks. The models are trained to parse detailed construction data, providing insights directly applicable to various aspects of construction management, such as extracting key information, generating activity sequences, and identifying missing activities. By refining these models to understand and process the unique aspects of construction scheduling and task interrelations, we transform SLMs into essential tools for delivering targeted, domain-specific solutions. This approach positions instruct fine-tuning at the core of our methodology, leveraging SLMs capabilities to advance construction project management.

4 Experimental Setup

4.1 Data Collection

In our approach to tagging construction constituents within comprehensive activities, we collected structured historical schedule data from two residential building projects, selecting 1500 activities for thorough analysis. To understand activity dependencies and sequences, we also compiled 1375 pairs of predecessor and successor activities, covering structural work, finishing, MEP (Mechanical, Electrical, and Plumbing), and commissioning domains. Using Python libraries, we efficiently extracted and parsed information from .xer files, the metadata files of Primavera, a widely used project management software in construction. This systematic organization prepared the data for subsequent analysis and model training in our study.

4.2 Materials and Methods

In this research, dataset preparation and model fine-tuning were conducted using Large Language Models (LLMs) via the OpenAI API and high-performance computational resources. Models were chosen for their domain-specific fine-tuning suitability, computational efficiency, and open-source availability. Initial evaluations were performed using the OpenAI API, while fine-tuning was done on Mistral-7B, Gemma-2-9B, and LLaMA-3.1-8B. Training utilized an NVIDIA RTX A6000 GPU with 48 GB VRAM, 50 GB RAM, and 8 virtual CPUs.

4.3 Prompt Template Preparation

Prompt templates are essential for effectively guiding and fine-tuning pre-trained models. In our approach, we use OpenAI's structured response prompt template with various instruction formats—zero-shot, one-shot, few-shot, chain-of-thought (CoT), and CoT with few-shot prompting. These formats help the model interpret and respond accurately by providing varying levels of contextual guidance. For fine-tuning, we adopt the Alpaca dataset format, illustrated in Figure 2, which standardizes instruction, user input, and assistant response. This ensures model outputs align with intended objectives. Fine-tuning is performed using open-source SLMs, allowing customization for domain-specific tasks. Combining OpenAI's prompt templates with the Alpaca format enhances the performance and adaptability of pre-trained models. This structured, context-rich data preparation enables the model to handle complex, domain-specific tasks with improved precision and efficiency.

Prompt Template for Fine-Tuning an Open-Source Model

```
### Instruction:
CIE - {instruction: "Detailed Instruction for Information Extraction"}
CAS - {instruction: "Detailed Instruction of Activity Sequence Prediction"}

### Input:
CIE - {input: "Activity Description" }
CAS - {input: "Activity_i, Activity_n" }

### Response:
CIE - {response: ""epc_type": "extract", "work_item": "extract", "location":
"extract", "object": "extract", "task": "extract""}
CAS - {response: "Predicted_Activities"}
```

Figure 2. Overview of Prompt Template

4.4 Prompt Template for Key Construction Information Extraction (CIE)

This template is intricately designed to facilitate the extraction of construction constituents from activities. It includes explicit instructions that detail the specific constituents needed for each activity, like EPC Type, Work Item, Location, Object, and Task. Structured input is provided to ensure that all aspects of a construction activity are precisely captured and categorized. This categorized data is then used to fine-tune the Key Construction Information Extraction model. Initially, the template provides clear instructions for the information extraction task. Following this, the user input section allows for entering specific details regarding the construction activity. The model then processes this input and generates responses in specified categories. This structured interaction ensures the models accurate understanding and categorization of essential construction information, thereby enhancing information extraction accuracy.

4.5 Prompt Template for Construction Activity Sequencing (CAS)

For sequencing construction activities, the instruction section clearly outlines the sequencing task, guiding users to input detailed descriptions of construction activities. Users then provide specific sequences. The model processes this information and responds with a logically organized sequence of activities, capturing the necessary progressions and dependencies. This template is pivotal in accurately forecasting construction activity flows, which is crucial for effective project planning and execution.

Building upon the dataset preparation and fine-tuning processes described, the subsequent section presents a comprehensive evaluation of the selected language models. The results highlight their effectiveness in domain-specific tasks, particularly in information extraction and activity sequence generation, demonstrating their applicability and performance within the context of automated construction planning and scheduling.

5 Results and Discussion

In this research, LLM was first used to extract activity constituents for dataset preparation. The model's performance was evaluated using various instruction formats with a base LLM to identify the most effective strategy. The best-performing format was then used to prepare the dataset for fine-tuning, requiring only minor manual corrections to ensure accuracy and quality. The finalized dataset was split into 80% for training and 20% for testing. This structured approach ensured effective model learning and reliable performance evaluation on unseen test data.

5.1 Phase 1: Performance of Key Construction Information Extraction (CIE)

Figure 3 shows the impact of different prompting strategies on the GPT-4o-mini-base model's performance in extracting activity constituent information. The CoT Few-shot approach achieved the highest accuracy, proving that combining step-by-step reasoning with examples greatly enhances the model's predictive ability. Similarly, the CoT and Few-shot methods also improved accuracy, highlighting the importance of structured reasoning. In contrast, the Zero-shot strategy led to more non-predictions and hallucinations, indicating that lacking contextual guidance reduces the model's effectiveness. Overall, Figure 3 underscores that structured prompting with reasoning and examples are key to improving model performance in specialized tasks.

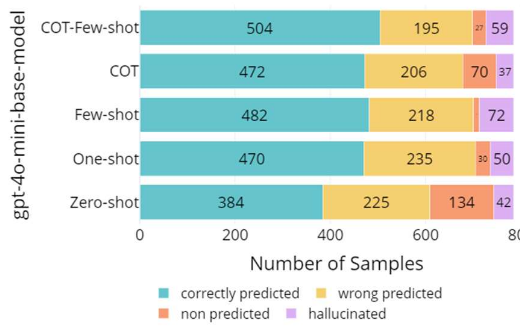

Figure 3. Performance Analysis of the CIE Model by Structured Prompt Techniques

The performance of fine-tuned SLMs, including mistral 7B, gemma 2 9B, and llama 3 8B, was evaluated for information extraction tasks. As shown in Figure 4, the mistral 7B model achieved the highest accuracy with the most correct predictions, followed closely by gemma 2 9B. In contrast, llama 3 8B exhibited lower accuracy with a higher number of non-predictions and hallucinated outputs. These models were fine-tuned using a consistent configuration across all models to ensure a fair comparison.

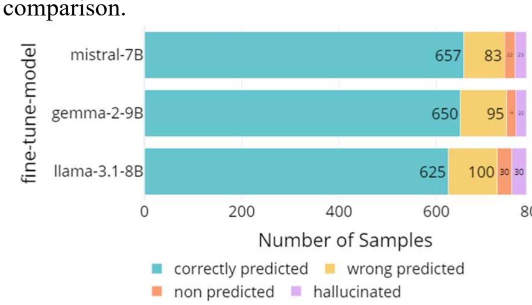

Figure 4. Performance Analysis of the CIE Model by Fine-tune model

Specifically, Low-Rank Adaptation (LoRA) adaptation was applied with a rank of 16, targeting key projection layers (e.g., q_proj, k_proj, v_proj) and optimized settings such as lora_alpha of 16 and lora_dropout set to 0. Training was conducted with a batch size of 2, gradient accumulation steps of 4, a learning rate of 2e-4.

To quantify model performance, all models were tested using these identical parameters, ensuring consistent evaluation. However, these parameters can be adjusted in future experiments to further enhance model performance.

Table 2 presents a performance comparison between fine-tuned open-source models and closed-source base models for information extraction tasks. Among the fine-tuned models, mistral 7B achieved the highest accuracy with 83.69% correct prediction, followed closely by

Table 2: Performance Comparison of Fine-Tuned Open-Source Models and Closed-Source Base Models for Information Extraction

Model Type	Model/Approach	Performance Gap to Best (%)	Correctly Predicted (%)
Fine Tune Model (Open Source)	mistral-7B	0.00	83.69
	gemma-2-9B	-0.89	82.80
	llama-3.1-8B	-4.07	79.61
Base Model (Closed Source)	gpt-4o-mini (CoT-Few-shot)	-19.49	64.20
	gpt-4o-mini (Few-shot)	-22.29	61.40
	gpt-4o-mini (CoT)	-23.56	60.12
	gpt-4o-mini (One-shot)	-23.82	59.87
	gpt-4o-mini (Zero-shot)	-34.77	48.91

gemma 2 9B and llama 3 1 8B, with slight performance gaps. In contrast, the closed-source gpt-4o-mini model, evaluated across various prompting strategies, showed noticeably lower performance. The CoT Few-shot approach led the closed-source models with 64.20% accuracy, while the Zero-shot strategy had the lowest accuracy at 48.91%. The performance gap to the best-performing model was smallest for mistral 7B and significantly larger for the base gpt-4o-mini models, especially with simpler prompting methods. These results highlight the superior performance of fine-tuned open-source models over closed-source base models in information extraction tasks.

5.2 Phase 2: Performance of Construction Activity Sequences (CAS)

As shown in Figure 5, the evaluation of construction activity sequence generation using BERTScores metrics Precision, Recall, and F1-score primarily focused on the performance of fine-tuned SLMs, specifically mistral-7B, gemma-2-9B, and llama-3.1-8B. Among these models, mistral-7B demonstrated the most balanced and consistent performance across all evaluation metrics, indicating its relative effectiveness in accurately generating and predicting construction activity sequences. In comparison, gemma-2-9B and llama-3.1-8B exhibited lower performance, suggesting that these models require further fine-tuning and optimization to improve their sequence generation capabilities. To provide a benchmark, the gpt-4o-mini model was included in the evaluation and achieved the highest scores across all metrics, with a precision of 0.89, recall of 0.87, and an F1-score of 0.88. While these results highlight the strong performance of larger models in complex sequence generations, this study primarily focused on assessing and improving fine-tuned SLMs for their computational efficiency and scalability.

The similarity scores of generated construction activity sequences were evaluated to compare the performance of fine-tuned SLMs and the larger gpt-4o-mini model. As shown in Figure 6, the gpt-4o-mini model achieved the highest mean similarity score of 91.59%, highlighting its superior ability to generate sequences closely aligned with the reference data. Among the fine-tuned models, mistral-7B performed the best with a mean similarity of 81.43%, while llama-3.1-8B and gemma-2-9B recorded mean similarity scores of 78.20% and 77.68%, respectively.

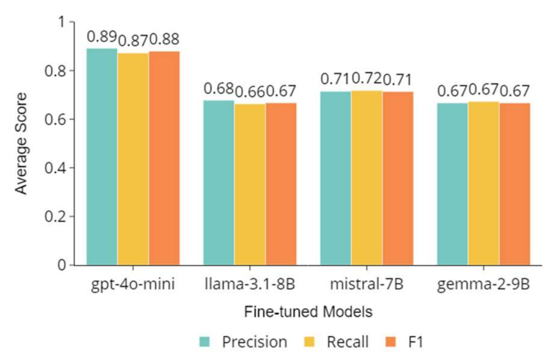

Figure 5. Evaluation of CAS Model: BERTScores (Precision, Recall, and F1) Analysis

These results demonstrate that while the fine-tuned SLMs performed reasonably well, they still lag the gpt-4o-mini in generating accurate and coherent construction activity sequences. However, the performance of the fine-tuned models, particularly mistral-7B, suggests that SLMs have strong potential for improvement.

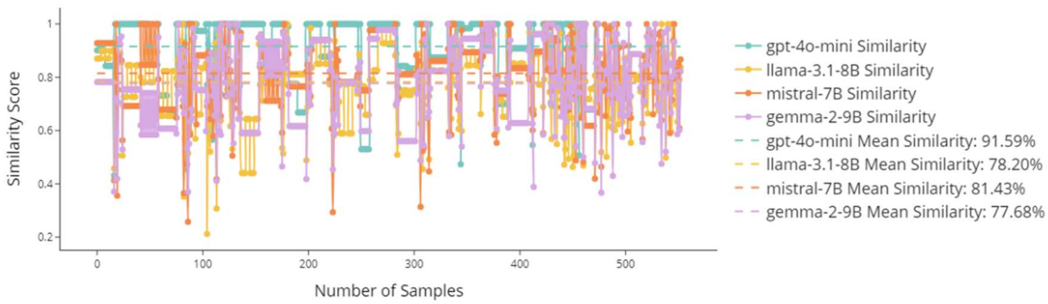

Figure 6. Evaluation of CAS Model: BERT Similarity Scores

5.3 Key Findings and Lesson Learned

In comparison to previous studies [1][9], which heavily rely on human-labeled schedule data and models such as LSTM, GNN, or BERT for learning activity patterns [9][2][3], this research adopts a more scalable and efficient approach using Small Language Models (SLMs) fine-tuned on minimally labeled data. Although pre-trained LLMs such as gpt-4o-mini were fine-tuned independently within this study and used as performance benchmarks, the primary emphasis remains on open-source SLMs. Among these, mistral-7B exhibited strong adaptability and effectiveness in handling construction planning tasks, offering a domain-aligned and practical alternative.

A key lesson learned is that targeted improvements in fine-tuning strategies such as advanced hyperparameter optimization and task-specific adaptation significantly enhance SLM performance. Emphasizing object and task identification during training allowed the models to infer contextual elements like location, improving the accuracy and relevance of generated sequences. The iterative process of hyperparameter tuning and model refinement proved crucial in balancing performance and computational efficiency.

These findings highlight the value of domain-specific data preparation, continuous model enhancement, and the combined potential of large and small models. With further optimization, fine-tuned SLMs can serve as effective, resource-efficient solutions for large-scale construction activity sequence generation.

5.4 Inferences of CIE and CAS model

Figure 7 showcases the optimal fine-tuned CIE model, mistral-7B, illustrating the differences in responses to user queries between the base model assistant and the fine-tuned model. Similarly, Figure 8 highlights the superior CAS fine-tuned model, mistral-7B, emphasizing the discrepancies in responses between the base and fine-tuned models. In both figures, errors are marked in red, while correct and newly identified sequences previously missed by the base model assistant are indicated in green. This comparison demonstrates the enhanced accuracy and predictive capabilities of the fine-tuned mistral-7B model.

System:
<Instrcution_predicting_comprehensive_construction_activity_constituents>

User: "First Floor Column Shuttering Work"
Base Model Assistant: "EPC Type: Not specified, Work Item Domain: Construction, Location: First Floor, Object: Column, Task: Shuttering Work"
Fine-Tuned Model Assistant: first floor: location, column: object, shuttering: task, construction: epc type, structural: work item domain

Figure 7. Inference Outcomes Demonstrated by the CIE Model

System:
<Instrcution_predicting_sequence_and_missing_activity>

Example: Same construction work item domain structural
User: "first floor column reinforcement,..., first floor slab formwork"
Base Model Assistant: "first floor column reinforcement, first floor column pouring, first floor beam reinforcement, first floor beam pouring, first floor slab formwork"
Fine-Tuned Model Assistant: first floor column reinforcement, first floor column formwork, first floor column pouring, first floor slab formwork

Example: Different construction work item domain finishing to mep
User: User: "Ceiling Plastering,..., HVAC testing"
Base Model: Based on the typical order of construction activities, ………..could include:……………......"Ceiling Plastering, Wiring installation, Insulation installation, Drywall installation, Painting, Trim installation, Flooring installation, HVAC testing"
Fine-Tuned Model Assistant: Ceiling Plastering, Ceiling Priming, Ceiling Sanding, Ceiling Painting, HVAC Duct Installation, HVAC Duct Insulation, HVAC Duct Testing, HVAC Testing

Figure 8. Inference Outcomes Demonstrated by the CAS Model

6 Conclusion

This research presents a two-phase AI-driven approach to enhance automated construction project planning using SLMs for activity sequencing and missing task prediction. In the first phase, prompt engineering extracts and organizes key information from historical project schedules to produce structured outputs. Due to inconsistencies and lower accuracy, an SLM was fine-tuned with domain-specific data to improve extraction. In the second phase, another SLM, trained on the refined dataset, learns activity patterns to generate optimized sequences and identify missing tasks. The gpt-4o-mini model served as a benchmark and achieved the highest accuracy. However, fine-tuned models like mistral-7B delivered promising results with notable advantages in efficiency, scalability, and security.

A key advantage of SLMs is their enhanced data privacy and security. Once fine-tuned, these models can be deployed and used locally without relying on external servers or cloud-based services, ensuring sensitive construction data remain secure and protected. This makes SLMs particularly suitable for industries like construction, where data confidentiality is critical.

Future work will address unstandardized schedule data, generate synthetic data for diverse projects, and enhance reasoning in models for better understanding of scheduling logic. Expanding datasets, adopting open-source models, and integrating with BIM will further improve performance and efficiency in AEC project management.

Acknowledgements

The authors thank the National Science and Technology Council, Taiwan, for supporting this research through research grant Nos. NSTC 112-2221-E-002-078-MY3.

References

[1] Amer, F., & Golparvar-Fard, M. (2019). "Formalizing Construction Sequencing Knowledge and Mining Company-Specific Best Practices from Past Project Schedules." Proceedings of ASCE International Conference on Computing in Civil Engineering, ASCE, Atlanta, GA, 215-223.

[2] Amer, F., & Golparvar-Fard, M. (2021). "Modeling Dynamic Construction Work Template from Existing Scheduling Records via Sequential Machine Learning." Advanced Engineering Informatics, 47, 101198.

[3] Amer, F., Jung, Y., & Golparvar-Fard, M. (2021). "Transformer Machine Learning Language Model for Auto-alignment of Long-term and short-term Plans in Construction." Automation in Construction, 132.

[4] Amer, F., Koh, H. Y., & Golparvar-Fard, M. (2021). "Automated Methods and Systems for Construction Planning and Scheduling: Critical Review of Three Decades of Research." Journal of Construction Engineering and Management, 147(7), 03121002.

[5] Changali, S., Mohammad, A., & Van Nieuwland, M. (2015). "The Construction Productivity Imperative." McKinsey Quarterly, 1-10.

[6] Chen, H., Zhang, Y., Zhang, Q., Yang, H., Hu, X., Ma, X., Yanggong, Y., & Zhao, J. (2023). "Maybe Only 0.5% Data is Needed: A Preliminary Exploration of Low Training Data Instruction Tuning." arXiv, cs.AI, 2305.09246.

[7] Doloi, H. (2013). "Cost Overruns and Failure in Project Management: Understanding the Roles of Key Stakeholders in Construction Projects." Journal of Construction Engineering and Management, 139(3), 267-279.

[8] Faghihi, V., Nejat, A., Reinschmidt, K. F., & Kang, J. H. (2015). "Automation in Construction Scheduling: A Review of the Literature." International Journal of Advanced Manufacturing Technology, 81, 1845-1856.

[9] Hong, Y., Xie, H., Hovhannisyan, V., & Brilakis, I. (2022). "A Graph-based Approach for Unpacking Construction Sequence Analysis to Evaluate Schedules." Advanced Engineering Informatics, 52, 101625.

[10] Pal, A., Lin, J. J., & Hsieh, S.H. (2023). "Schedule-Driven Analytics of 3D Point Clouds for Automated Construction Progress Monitoring." Proceedings of ASCE International Conference on Computing in Civil Engineering 2023, ASCE, Reston, VA, 412.

[11] Singh, A. K., Pal, A., Kumar, P., Lin, J. J., & Hsieh, S. H. (2023). "Prospects of Integrating BIM and NLP for Automatic Construction Schedule Management." Proceedings of the International Symposium of Automation and Robotics in Construction, Madras, TN, India, Vol. 40, 238-245.

[12] Singh, J., & Anumba, C. J. (2022). "Real-Time Pipe System Installation Schedule Generation and Optimization Using Artificial Intelligence and Heuristic Techniques." Journal of Information Technology in Construction (ITcon), 27, 173-190

[13] Zhang, S., Dong, L., Li, X., Zhang, S., Sun, X., Wang, S., Li, J., Hu, R., Zhang, T., Wu, F., & Wang, G. (2023). "Instruction Tuning for Large Language Models: A Survey." arXiv preprint arXiv:2308.10792.

[14] Zheng, J., & Fischer, M. (2023). "Dynamic Prompt-based Virtual Assistant Framework for BIM Information Search." Automation in Construction, 155,105067.