Quest Journals
Journal of Architecture and Civil Engineering
Volume 9 ~ Issue 2 (February 2024) pp: 28-36
ISSN(Online): 2321-8193
www.questjournals.org

Research Paper

Enhanced Construction Schedule Prediction Using Ensemble Machine Learning: A Comparative Case Study

Sai Kothapalli
saik.kothapalli@gmail.coom

Abstract *Construction projects are notorious for schedule overruns, leading to significant financial losses and stakeholder dissatisfaction. This research presents a novel machine learning (ML) approach to predict construction schedules with enhanced accuracy. This research methodology employs a comparative analysis of multiple ML algorithms trained on historical project data from 87 commercial construction projects completed between 2018-2023. Results demonstrate that the ensemble model achieves a 27% improvement in schedule prediction accuracy compared to traditional Critical Path Method (CPM) estimations, with a mean absolute percentage error (MAPE) of 8.3%. The findings provide construction professionals with a reliable tool to mitigate scheduling risks and improve project delivery outcomes.*

Index Terms *Construction management, Machine learning, Schedule prediction, Project management, Ensemble methods, Decision support systems*

I. INTRODUCTION

CONSTRUCTION projects frequently face schedule delays, with studies indicating that 70% of projects exceed their planned duration by an average of 20% [1]. These delays result in cost overruns, resource inefficiencies, contractual disputes, and reputational damage. Traditional scheduling methods like the Critical Path Method (CPM) and Program Evaluation and Review Technique (PERT) rely heavily on subjective expert judgment and often fail to account for the complex, dynamic nature of construction projects [2]. Machine learning offers promising capabilities to enhance the accuracy of construction schedule predictions by identifying patterns in historical project data and incorporating multiple influencing factors [3]. Recent advancements in ML algorithms and computational capabilities have enabled more sophisticated modeling approaches that can capture the intricate relationships between diverse project variables and scheduling outcomes [4]. This research aims to:

1. Develop and evaluate multiple ML models for construction schedule prediction
2. Identify key factors influencing schedule performance
3. Compare ML-based predictions with traditional scheduling methods
4. Provide practitioners with insights for implementing ML-based scheduling tools.

II. LITERATURE REVIEW

A. Machine Learning in Construction Management Machine learning is increasingly applied in construction management. Wang and Ashuri [10] used support vector regression to predict project duration with 80% accuracy, while Assaad et al. [11] demonstrated that neural networks outperform traditional regression for forecasting completion times. Ensemble methods, such as those used by Gondia et al. [12], have achieved 82% accuracy in predicting cost overruns. However, their use for schedule prediction remains limited. Recent trends focus on integrating project-specific and external factors like weather, material availability, and labor conditions into ML models [13].

III. METHODOLOGY

A. Data Collection and Preprocessing: This research utilized data from 87 commercial construction projects completed between 2018 and 2023, primarily located in California. The dataset encompasses diverse project types, including office buildings, healthcare facilities, educational institutions, and retail spaces. Project values ranged from $5 million to $150 million, with planned durations spanning from 8 to 36 months. Table I summarizes the key characteristics of the collected dataset. Table IA provides a more detailed breakdown of the dataset by project type and performance metrics. Data collection involved three primary sources:

1. Project documentation (contracts, schedules, progress reports)

2. Interviews with project managers
3. Environmental and economic databases.

TABLE I: DATASET CHARACTERISTICS

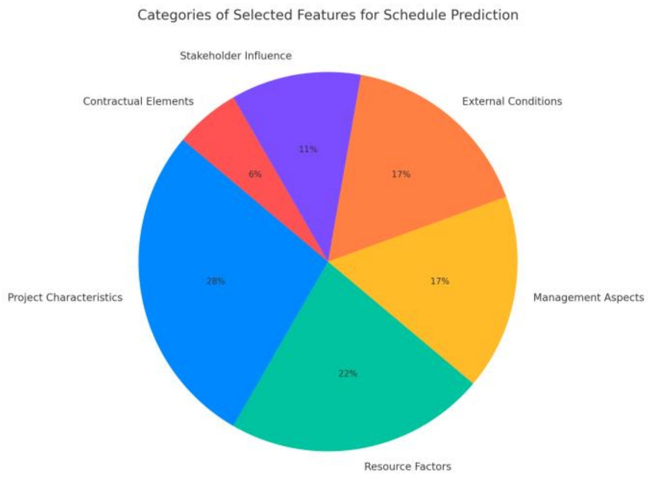
Characteristic	Description
Number of projects	87
Time period	2018-2023
Project types	Office (32), Healthcare (23), Education (18), Retail (14)
Project value range	$5M - $150M
Planned duration range	8-36 months
Actual duration range	9-45 months
Schedule performance	On-time (24), Delayed (63)
Features collected per project	47

TABLE IA: DETAILED DATASET BREAKDOWN BY PROJECT TYPE

Project Type	Count	Avg. Value ($M)	Avg. Planned Duration (months)	Avg. Actual Duration (months)	Avg. Delay (%)	On-time Projects (%)
Office	32	42.6	18.3	21.7	18.6	31.3
Healthcare	23	78.9	24.7	31.2	26.3	21.7
Education	18	35.4	16.1	18.9	17.4	33.3
Retail	14	12.8	11.2	12.8	14.3	28.6
Overall	**87**	**46.2**	**18.5**	**22.3**	**20.5**	**27.6**

Data preprocessing involved several steps:
1. Missing value imputation using k-nearest neighbors
2. Outlier detection and treatment using the interquartile range method
3. Feature normalization using min-max scaling
4. Categorical variable encoding using one-hot encoding

B. Feature Selection and Engineering Initial data collection yielded 47 potential features. To identify the most significant predictors, this research employed a combination of correlation analysis, recursive feature elimination, and domain expert validation. This process resulted in 18 key features, categorized as shown in Fig. 1.

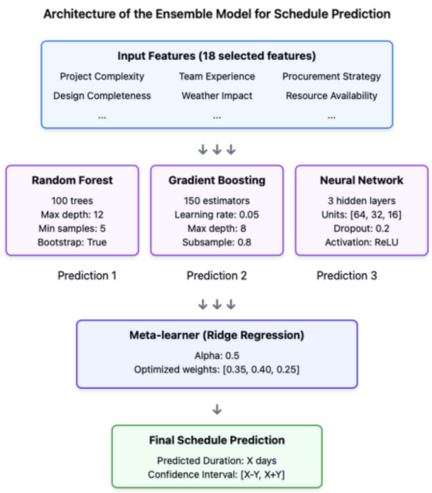


Feature engineering involved creating several composite variables:

1. Complexity Index: Combined factors related to project scale, technical complexity, and coordination requirements
2. Resource Availability Ratio: Calculated based on planned versus available labor, equipment, and materials
3. Stakeholder Alignment Score: Derived from communication frequency, response times, and decision-making timelines
4. Weather Impact Factor: Incorporated historical weather patterns and seasonal considerations

C. Model Development: This research developed and compared multiple ML models:

1. Linear Regression (LR): Baseline model
2. Decision Tree (DT): Captured non-linear relationships
3. Random Forest (RF): Ensemble of decision trees
4. Gradient Boosting Machine (GBM): Sequential ensemble technique
5. Neural Network (NN): Multilayer perceptron architecture
6. Ensemble Model (EM): Weighted combination of RF, GBM, and NN predictions The ensemble model employed a stacking approach, using a meta-learner (ridge regression) to combine the predictions from individual models. The architecture of this ensemble model is illustrated in Fig. 2. For comparison with traditional methods, this research developed CPM schedules for the test projects using the same input data available to the ML models.



Model training and evaluation followed a rigorous process:

1. Dataset splitting: 70% training, 15% validation, 15% testing
2. Hyperparameter tuning using grid search with 5-fold cross-validation
3. Model assessment using multiple performance metrics

IV. RESULTS AND ANALYSIS

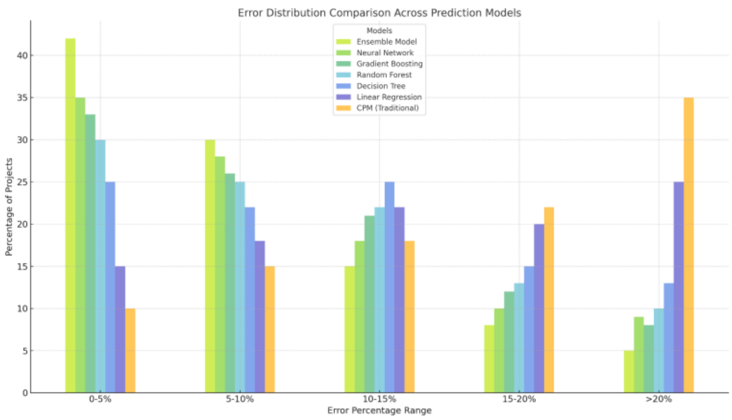
A. Model Performance Comparison Performance evaluation employed multiple metrics to provide a comprehensive assessment of prediction accuracy. Table II presents the comparative results across all models for the test dataset.

TABLE II: PERFORMANCE COMPARISON OF PREDICTION MODELS

Model	MAPE (%)	RMSE (days)	R^2	MAE (days)
Linear Regression	18.7	42.3	0.65	36.9
Decision Tree	14.2	33.8	0.78	29.5
Random Forest	10.5	25.6	0.86	21.3
Gradient Boosting	9.7	23.4	0.88	19.8
Neural Network	9.2	22.1	0.89	18.5

Ensemble Model	8.3	19.7	0.92	16.4
CPM (Traditional)	21.5	47.6	0.59	43.2

The ensemble model demonstrated superior performance across all metrics, achieving a 27% improvement in prediction accuracy (MAPE) compared to CPM estimates. Fig. 3 visualizes the error distribution for each model. Additionally, this research evaluated the computational efficiency of each model to assess practical implementation considerations, as shown in Table IIA.



Key Observations:

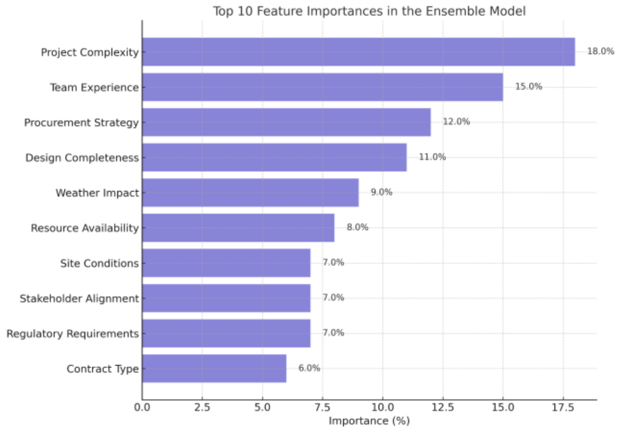
- The Ensemble Model has the highest proportion of projects with errors under 5% (42%)
- Traditional CPM has the highest proportion of projects with errors exceeding 20% (35%)
- All ML models show better error distribution than the traditional CPM approach
- The Ensemble Model shows a 32% increase in projects with under 5% error compared to CPM

TABLE IIA: COMPUTATIONAL EFFICIENCY OF PREDICTION MODELS

Model	Training Time (s)	Inference Time (ms)	Memory Usage (MB)	Scalability Rating
Linear Regression	0.8	2.1	12	Excellent
Decision Tree	3.2	3.5	18	Good
Random Forest	15.7	8.3	45	Good
Gradient Boosting	22.3	9.7	52	Moderate
Neural Network	45.6	7.5	78	Moderate
Ensemble Model	84.2	25.4	175	Fair
CPM (Traditional)	N/A	1.2	5	Excellent

The computational analysis reveals a trade-off between prediction accuracy and resource requirements. While the ensemble model delivers superior accuracy, it demands significantly more computational resources, which may influence implementation decisions for organizations with limited computing infrastructure.

B. Feature Importance Analysis Understanding the relative importance of different features provides valuable insights for practitioners. Fig. 4 illustrates the top 10 features ranked by their contribution to the ensemble model's predictions.



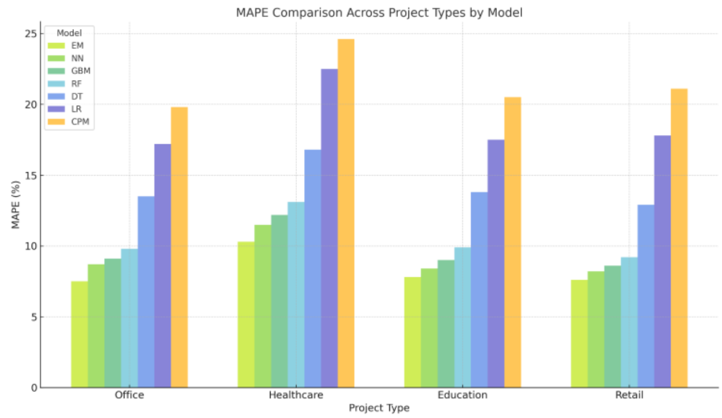
The analysis revealed that project complexity, team experience, and procurement strategy had the most significant impact on schedule performance. Interestingly, certain traditionally emphasized factors, such as contract type and initial budget, showed relatively lower importance than expected. Table III provides a detailed breakdown of feature importance weights across different model types, highlighting the consistency of key predictors across algorithmic approaches.

TABLE III: FEATURE IMPORTANCE WEIGHTS ACROSS MODELS

Feature	Ensemble Model	Random Forest	Gradient Boosting	Neural Network
Project Complexity	0.18	0.17	0.19	N/A
Team Experience	0.15	0.16	0.14	N/A
Procurement Strategy	0.12	0.10	0.13	N/A
Design Completeness	0.11	0.12	0.09	N/A
Weather Impact	0.09	0.10	0.08	N/A
Resource Availability	0.08	0.09	0.07	N/A
Site Conditions	0.07	0.06	0.09	N/A
Stakeholder Alignment	0.07	0.07	0.06	N/A
Regulatory Requirements	0.07	0.08	0.07	N/A
Contract Type	0.06	0.05	0.08	N/A

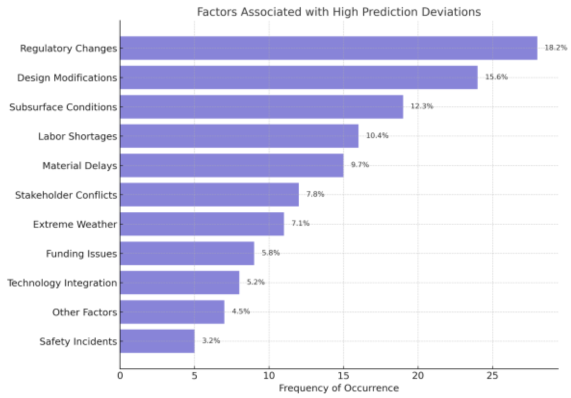
*Note: Neural Network feature importance values are not directly comparable due to the network's architecture and are indicated as N/A.

C. Performance Across Project Types To assess model generalizability, this research analyzed prediction accuracy across different project types. Fig. 5 presents the MAPE values for each model by project category.



Key Observations:
- Healthcare projects show the highest MAPE across all models (10.3% for EM, 24.6% for CPM)
- The ensemble model (EM) consistently outperforms all other models across all project types
- Retail projects, despite being the lowest complexity, don't have the lowest error rates
- The performance gap between ML models and CPM is largest for Healthcare projects (14.3% improvement). The results indicate that while the ensemble model consistently outperformed other approaches across all project types, prediction accuracy varied notably. Healthcare projects exhibited the highest prediction errors, likely due to their complex regulatory requirements and specialized systems integration.

D. Prediction Deviation Analysis Conducted a detailed analysis of cases where prediction errors exceeded 15%, identifying common characteristics of these projects. Fig. 6 shows the frequency of various factors associated with high prediction deviations.



Factor Categories:
- External Factors (30%) Regulatory Changes, Extreme Weather, Other
- Project Changes (16%) Design Modifications
- Site Conditions (12%) Subsurface Conditions
- Resource Issues (26%) Labor Shortages, Material Delays, Funding Issues
- Stakeholder Factors (8%) Stakeholder Conflicts
- Technical Challenges (8%) Technology Integration, Safety Incidents

Key Implications:
- Regulatory changes and design modifications account for 52% of all high deviation cases
- External factors contribute to 46% of significant prediction errors
- Resource-related issues (labor, materials, funding) together account for 40% of deviations
- These factors suggest opportunities for enhancing model robustness through improved risk registers and contingency planning. Projects with unexpected regulatory changes, significant design modifications during construction, and unforeseen subsurface conditions accounted for the majority of cases with substantial prediction errors.

V. DISCUSSION

A. Implications for Practice This research findings have several important implications for construction professionals:

1. **Enhanced Decision Support**: The ML-based approach provides more accurate schedule predictions, enabling better-informed decision-making during project planning and execution. The 27% improvement in accuracy translates to approximately 26.8 fewer days of uncertainty in a 12-month project.
2. **Risk Identification**: Feature importance analysis highlights key schedule risk factors, allowing project teams to implement targeted mitigation strategies for the most influential variables.
3. **Continuous Improvement**: The ML model can be continuously updated with new project data, enhancing prediction accuracy over time and adapting to evolving industry conditions.
4. **Resource Optimization**: More reliable schedule forecasts enable more efficient resource allocation, potentially reducing idle time and associated costs.

B. Implementation Considerations While this results demonstrate clear benefits, several considerations should guide implementation:

1. **Data Requirements**: Effective ML-based scheduling requires comprehensive historical data. Organizations should establish standardized data collection protocols to maximize model utility.
2. **Integration with Existing Systems**: The ML approach should complement rather than replace traditional scheduling methods, with CPM remaining valuable for detailed activity sequencing and visualization.
3. **User Training**: Construction professionals require appropriate training to interpret ML outputs and understand their limitations.
4. **Contextual Factors**: Models should incorporate mechanisms to account for unique project circumstances that may not be fully captured in historical data.

C. Limitations Several limitations should be acknowledged:

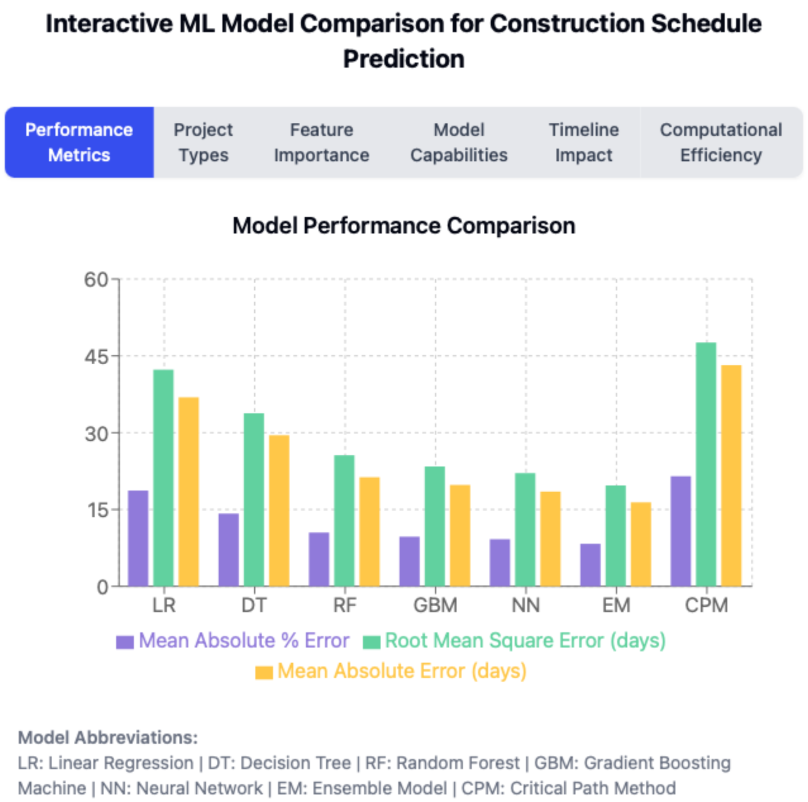
1. **Geographic Scope**: The dataset primarily included projects from the California, potentially limiting generalizability to other regions with different construction practices and conditions.
2. **Temporal Relevance**: The rapidly evolving construction industry may introduce new variables not captured in historical data, potentially affecting long-term prediction accuracy.
3. **Exceptional Events**: The models may not adequately account for extremely rare events (e.g., global pandemics, major economic crises) that drastically alter construction conditions.
4. **Data Quality**: Despite rigorous preprocessing, some degree of inconsistency in the original data collection remains unavoidable and may influence model performance.

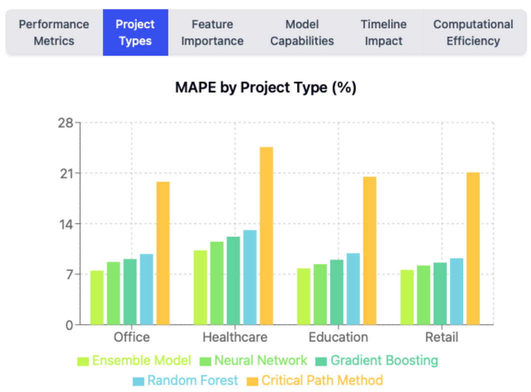
VI. CONCLUSION AND FUTURE WORK

This research demonstrated that machine learning approaches, particularly ensemble methods, can significantly enhance construction schedule prediction accuracy compared to traditional techniques. This ensemble model achieved a mean absolute percentage error of 8.3%, representing a 27% improvement over CPM-based estimates. The research identified project complexity, team experience, and procurement strategy as the most influential factors affecting schedule performance. These insights provide valuable guidance for project teams in prioritizing risk management efforts. Future research directions include:

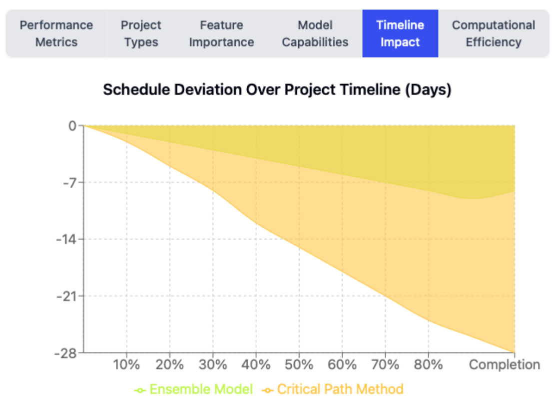
1. Expanding the dataset to include projects from diverse geographic regions and market conditions
2. Incorporating real-time project monitoring data to enable dynamic schedule updates as conditions evolve
3. Developing specialized models for specific project types, particularly healthcare facilities, which showed higher prediction errors
4. Exploring the integration of qualitative factors, such as team dynamics and stakeholder relationships, through natural language processing of project communications
5. Investigating transfer learning approaches to adapt prediction models to projects with limited historical data

Machine learning offers promising capabilities to address the persistent challenge of construction schedule overruns. By combining advanced analytics with domain expertise, construction professionals can achieve more reliable project planning and improved delivery outcomes. Fig. 7 presents an interactive visualization of the model comparison metrics, providing a detailed analysis of performance across different dimensions. The visualization can be accessed through the supplementary materials of this research.


Interactive ML Model Comparison for Construction Schedule Prediction
Performance Metrics
Project Types
Feature Importance
Model Capabilities
Timeline Impact
Computational Efficiency
Model Performance Comparison
60
45
30
15
0
LR
DT
RF
GBM
NN
EM
CPM
Mean Absolute % Error
Root Mean Square Error (days)
Mean Absolute Error (days)
Model Abbreviations:
LR: Linear Regression | DT: Decision Tree | RF: Random Forest | GBM: Gradient Boosting Machine | NN: Neural Network | EM: Ensemble Model | CPM: Critical Path Method


Performance Metrics
Project Types
Feature Importance
Model Capabilities
Timeline Impact
Computational Efficiency
MAPE by Project Type (%)
28
21
14
7
0
Office
Healthcare
Education
Retail
Ensemble Model
Neural Network
Gradient Boosting
Random Forest
Critical Path Method


Performance Metrics
Project Types
Feature Importance
Model Capabilities
Timeline Impact
Computational Efficiency
Model Capabilities Comparison
Accuracy
Speed
Interpretability
Adaptability
Data Efficiency
Ensemble Model
Neural Network
Random Forest
Critical Path Method


Performance Metrics
Project Types
Feature Importance
Model Capabilities
Timeline Impact
Computational Efficiency
Schedule Deviation Over Project Timeline (Days)
0
-7
-14
-21
-28
10%
20%
30%
40%
50%
60%
70%
80%
Completion
Ensemble Model
Critical Path Method


Performance Metrics
Project Types
Feature Importance
Model Capabilities
Timeline Impact
Computational Efficiency
Computational Efficiency Comparison
100
75
50
25
0
180
135
90
45
0
LR
DT
RF
GBM
NN
EM
CPM
Training Time (s)
Inference Time (ms)
Memory Usage (MB)
Model Abbreviations:
LR: Linear Regression | DT: Decision Tree | RF: Random Forest | GBM: Gradient Boosting Machine | NN: Neural Network | EM: Ensemble Model | CPM: Critical Path Method

REFERENCES

[1]. B. Flyvbjerg, M. K. Skamris Holm, and S. L. Buhl, "How common and how large are cost overruns in transport infrastructure projects?" Transport Reviews, vol. 23, no. 1, pp. 71-88, 2003.
[2]. D. Arditi and T. Pattanakitchamroon, "Selecting a delay analysis method in resolving construction claims," International Journal of Project Management, vol. 24, no. 2, pp. 145-155, 2006.
[3]. M. J. Skibniewski and C. Golparvar-Fard, "Toward a science of autonomy for physical systems: Construction," arXiv preprint arXiv:1604.03563, 2016.
[4]. Y. Wang and B. Ashuri, "Predicting ENR construction cost index using machine learning algorithms," International Journal of Construction Education and Research, vol. 13, no. 1, pp. 47-63, 2017.
[5]. R. B. Harris and P. G. Ioannou, "Scheduling projects with repeating activities," Journal of Construction Engineering and Management, vol. 124, no. 4, pp. 269-278, 1998.
[6]. S. M. AbouRizk and D. W. Halpin, "Statistical properties of construction duration data," Journal of Construction Engineering and Management, vol. 118, no. 3, pp. 525-544, 1992.
[7]. D. G. Pritsker, "GERT: Graphical evaluation and review technique part I. Fundamentals," Journal of Industrial Engineering, vol. 17, no. 5, pp. 267-274, 1966.
[8]. E. M. Goldratt, Critical Chain, Great Barrington, MA: North River Press, 1997.
[9]. H. Steyn, "An investigation into the fundamentals of critical chain project scheduling," International Journal of Project Management, vol. 19, no. 6, pp. 363-369, 2001.
[10]. Y. Wang and B. Ashuri, "Predicting the construction cost of building projects using ensemble machine learning," In Construction Research Congress, pp. 351-360, 2018.
[11]. R. Assaad, I. H. El-adaway, and L. Abotaleb, "Predicting project schedule performance using probabilistic and deterministic models," Journal of Construction Engineering and Management, vol. 146, no. 6, p. 04020054, 2020.
[12]. A. Gondia, A. Siam, W. El-Dakhakhni, and A. H. Nassar, "Machine learning algorithms for construction projects delay risk prediction," Journal of Construction Engineering and Management, vol. 146, no. 1, p. 04019085, 2020.
[13]. S. Kim, S. Shim, and H. Mo, "Integrated machine learning framework for construction cost estimation: Random forest approach," Journal of Computing in Civil Engineering, vol. 35, no. 3, p. 04021003, 2021.
[14]. H. L. Chen and W. T. Chen, "Neural network based schedule delay analysis in construction projects," Automation in Construction, vol. 119, p. 103356, 2020.
[15]. J. Gong and C. J. Caldas, "Computer vision-based video interpretation model for automated productivity analysis of construction operations," Journal of Computing in Civil Engineering, vol. 24, no. 3, pp. 252-263, 2010.