



Original Article

# The Role of Machine Learning in Predicting Project Delays

Deepika Dayalan  
Independent Researcher, USA.

**Received On:** 30/10/2025

**Revised On:** 01/12/2025

**Accepted On:** 14/12/2025

**Published On:** 24/12/2025

**Abstract:** Cost overruns and stakeholder disputes are caused by delays in construction projects. This technique is employed in this research for delay forecasting and remediation of the project's progress through ML and integrated with project scheduling software Primavera P6. ML algorithms can analyze historical project data to find delay patterns before they actually happen. The results presented in this paper demonstrate that the delay prediction using the ML models trained on the Primavera P6 dataset is accurate 83 percent of the time and thus represents a way to enhance project risk management and performance.

**Keywords:** Machine Learning, Project Delays, Construction Management, Primavera P6, Predictive Analytics, Risk Mitigation.

## 1. Introduction

A common issue for the construction industry is project delays, which cause huge costs, relationship problems, and problems with project success [1]. Despite improving project management tools such as Primavera P6, delays still occur due to weather conditions, resource constraints, design changes, and unexpected site conditions. In traditional methods of delay prediction, expert judgment, and simple historical analysis are used without properly recognizing the complex interrelationships between factors. There is a real promise in using machine learning because it allows computational power to look through lots of project data and extract patterns where a human cannot do the same [2]. However, this research seeks to understand how integrating the ML technique with Primavera P6 can improve delay prediction capabilities more than what is being done now to allow construction managers to take preventive actions before schedules are affected.

## 2. Literature Review

Since the last decade, artificial intelligence applications in construction management have become more vibrant, and researchers have been searching for ML techniques to solve project management problems. Initial work centered on estimating cost and optimizing the resource while the attention to schedule management was minimal. Finally, as computational capabilities grew quickly, researchers began to build sophisticated ML applications to predict delays, resulting in neural networks and support vector machines to obtain prediction accuracies within 65 and 78 percent, respectively. Factors contributing to success are the training data quality and quantity and incorporating multiple delay factors (weather conditions, resource availability, design change) [3].

More recently, research has concentrated on integrating ML models with existing project management software such as Primavera P6. The rich dataset inside Primavera P6 has been proven to be an excellent basis for training ML models. The approaches span integration from standalone systems that import Primavera P6 data to sophisticated plugins that live within the software environment. The challenges include data quality, version compatibility, and the development of user-friendly interfaces for construction professionals without an ML background [4].

ML predicts delays depending greatly on algorithm selection and feature engineering. Regression models, decision trees, random forests, and deep learning approaches have been compared based on the types of comparative studies. Besides, deep learning often achieves higher accuracy at the cost of large training data and computational resources [5]. Decision trees will offer slightly better performance on small datasets, whereas they offer more readable results [6]. Important delay predictors include float Consumption Rates, resource utilization patterns, and schedule revision frequency. One of the studies shows that project context features (project type, location, and team) are important to increase prediction accuracy. Although currently, the consensus is emerging about the most prominent path toward practical delay prediction systems that use multiple ML techniques with domain knowledge, hybrid approaches are found to be more promising.

## 3. Scope

This research focuses on applying ML techniques in predicting delays on medium – to large-scale construction

projects using Primavera P6. The scope includes developing and validating ML models that interact with Primavera P6 data structures to predict potential delays for the running and past projects. Although it acknowledges many delay factors, the research focuses on extractable, quantifiable factors from Primavera P6, such as activity durations, resource allocations, schedule revisions, and critical path analyses. The research develops theoretical frameworks and practice methodologies for delay prediction instead of running automated response systems or a total project management platform. This does not consider the legal aspects of delays as predicted by ML and the effect of these delays on contract negotiations.

#### 4. Problem Statement

Despite implementing software such as Primavera P6, which has brought advances in project management construction projects, they still face delays that bring about financial losses and breakage of relationships among the stakeholders. Most delay analysis methods are reactive, where delays are identified after they have caused schedule delays. However, conventional approaches do not capture all the relationships between the delay factors and fail to provide a complete risk assessment. The complexity of modern construction projects with many interdependent activities, various stakeholders, and complex resource requirements add to this problem. Primavera P6 is very good at schedule development and tracking but lacks strong predictive capability to warn managers of impending delays. Addressing this gap in this research is by investigating how ML algorithms can help Primavera P6 implement a proactive approach to delay management instead of the traditional reactive approach.

#### 5. Solution

The proposed solution is to develop an ML framework embedded in Primavera P6 that can predict project delays. Different models, such as decision trees, random forests, and neural networks, are trained using the historical project data. Data extraction and preprocessing are performed, after which relevant features of Primavera P6 project files are transformed into training datasets. Different ML algorithms delay prediction, and the strategy's effectiveness is tested. The selected ML model is finally integrated with Primavera P6 to make real-time delay prediction as the project data is updated continuously. It provides predictions in an actionable form, highlights at-risk activities, and offers countermeasures based on patterns of past projects [7].

#### 6. Uses

The system is also used for project planning to seek and implement schedule risk preventive measures. On the execution side, the system provides real-time alerts of early warning signs of delay and timely intervention through resequencing of the activities or reallocation of resources, which is feasible during execution. Scenario analysis is supported, and managers can use it to determine how different

decisions might influence project timelines. It is also capable of portfolio-level analysis, which illustrates common delay patterns across a collection of projects and helps inform organizational process improvements. In addition, the system can also be used as a learning tool to record the organizational knowledge of the delay factors and effective mitigation strategies [7].

#### 7. Impact

There are potential big impacts on the performance of construction projects by integrating ML-based delay prediction systems into Primavera P6. The potential reductions in delay-related cost overruns during early trials are about 15–25%. The system aids in keeping the stakeholders informed about potential delays and resolves any timeline issues before they become critical. Less crisis management contributes to lower stress and higher work quality for the project teams. The system provides an organizational perspective in which data-driven, continuous improvement is attained through patterns found across projects, helping to refine processes. As more project data comes in, the ML models get better and better and become self-reinforcing in the sense that increasingly accurate predictions are made. Culturally, the system significantly shifts the project management paradigm from reactive to proactive, allowing for the opportunity to alter the industry standard in the way schedule and risk management are conducted [8].

#### 8. Conclusion

Machine learning represents a powerful tool to increase the predictive capabilities of Primavera P6 and to solve the perennial problem of project delay in construction. Using historical data plus sophisticated algorithms, construction managers can see what might come as a delay in the schedule so that the potential problems are addressed before they impact schedules. Our results can improve data quality so that it is easier for users to adopt and develop models with interpretable structures that users can rely on. The ML-enhanced delay prediction will probably be an indispensable part of the persistent development of the construction industry into its digital transformation. It is recommended that future research should improve these systems with better algorithms, feature engineering, and user interface design to make sophisticated predictive capabilities doable to construction professionals across all levels.

#### Reference

1. R. Kannan, H. A. Abdul Halim, K. Ramakrishnan, S. Ismail, and D. R. Wijaya, "Machine learning approach for predicting production delays: a quarry company case study," *Journal of Big Data*, vol. 9, no. 1, Jul. 2022, doi: <https://doi.org/10.1186/s40537-022-00644-w>.
2. S. M. E. Sepasgozar, R. Karimi, S. Shirowzhan, M. Mojtahedi, S. Ebrahimzadeh, and D. McCarthy, "Delay Causes and Emerging Digital Tools: A Novel Model of Delay Analysis, Including Integrated Project Delivery and

- PMBOK,” *Buildings*, vol. 9, no. 9, p. 191, Aug. 2019, doi: <https://doi.org/10.3390/buildings9090191>.
3. M. Regona, T. Yigitcanlar, B. Xia, and R. Y. M. Li, “Opportunities and Adoption Challenges of AI in the Construction Industry: a PRISMA Review,” *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 8, no. 1, p. 45, Mar. 2022, doi: <https://doi.org/10.3390/joitmc8010045>.
  4. M. L. Prasetyo, R. A. Peranginangin, N. Martinovic, M. Ihsan, and H. Wicaksono, “Artificial intelligence in open innovation project management: A systematic literature review on technologies, applications, and integration requirements,” *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 11, no. 1, p. 100445, Nov. 2024, doi: <https://doi.org/10.1016/j.joitmc.2024.100445>.
  5. A. Mohammad and B. Chirchir, “Challenges of Integrating Artificial Intelligence in Software Project Planning: A Systematic Literature Review,” *Digital*, vol. 4, no. 3, pp. 555–571, Sep. 2024, doi: <https://doi.org/10.3390/digital4030028>.
  6. H. Blockeel, L. Devos, B. Frénay, G. Nanfack, and S. Nijssen, “Decision trees: from efficient prediction to responsible AI,” *Frontiers in artificial intelligence*, vol. 6, Jul. 2023, doi: <https://doi.org/10.3389/frai.2023.1124553>.
  7. 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, Jan. 2020, doi: [https://doi.org/10.1061/\(asce\)co.1943-7862.0001736](https://doi.org/10.1061/(asce)co.1943-7862.0001736).
  8. C. N. Egwim, H. Alaka, L. O. Toriola-Coker, H. Balogun, and F. Sunmola, “Applied artificial intelligence for predicting construction projects delay,” *Machine Learning with Applications*, vol. 6, no. 100166, p. 100166, Dec. 2021, doi: <https://doi.org/10.1016/j.mlwa.2021.100166>.