Management Analysis Journal 14 (2) (2025)

http://maj.unnes.ac.id

ANALYSIS OF CONSTRUCTION RISK USING MARKOV DECISION PROCESS AND REINFORCEMENT LEARNING

Restu Wijang Prasetyo[1],**Wiwik Handayani**[2]✉

[1,2]Department of Management, Faculty of Economics and Business, Universitas Pembangunan Nasional "Veteran" Jawa Timur

Article Information

History of article:
Accepted April 2025
Approved May 2025
Published June 2025

Keywords:
Construction Risk, Markov Decision Process, Project Efficiency, Q-Learning

Abstract

Recognizing the critical need for advanced risk management in the rapidly expanding construction sector, this study aims to analyze and optimize construction risk mitigation strategies at a prominent Indonesian housing developer, enhancing project time efficiency and reducing decision-making uncertainty. The research methodology employs a quantitative descriptive approach, utilizing Markov Decision Process to model project risk dynamics and Reinforcement Learning, specifically the Q-Learning algorithm, to determine optimal mitigation policies. Data collection involved direct observation, in-depth interviews with project management, and analysis of historical project documentation from a housing project. Research findings demonstrate that the Q-Learning model effectively identifies and recommends adaptive mitigation strategies for various risk levels, providing optimal actions that significantly reduce project delays. The implementation of these data-driven strategies resulted in a notable improvement in project time efficiency, reducing the average project duration. Reproducibility, convergence, and sensitivity tests further validate the model's reliability and robustness, confirming its capacity to provide consistent and stable recommendations under diverse conditions.

✉correspondence Address:
Jl. Rungkut Madya, Gn. Anyar, Kec. Gn. Anyar, Surabaya, Jawa Timur 60294
E-mail: wiwik.em@upnjatim.ac.id

e-ISSN 2502-1451

INTRODUCTION

As the realm of information and communication technology continues to evolve, new doors will open for companies looking to venture into lucrative markets. The construction industry stands as a pivotal sector globally, profoundly contributing to economic growth and infrastructural development. In Indonesia, this sector has demonstrated impressive vitality, with the Badan Pusat Statistik (BPS) and Kementerian Keuangan reporting a substantial growth of 7.48% in 2024. This growth rate positions construction as the leading sector, surpassing others such as industrial processing (4.72%) and trade (4.82%). Such as industrial processing (4.72%) and trade (4.82%). Such expansion, while indicative of economic prosperity, concurrently escalates the complexity and inherent risks within construction projects. These challenges frequently manifest as project delays, unforeseen cost escalations, and the unpredictable influence of environmental factors.

PT Berkah Development, a key player in Indonesia's housing development sector, epitomizes these challenges in its operational endeavors. The company, which commenced operations in 2017 with a focus on delivering quality and affordable housing, frequently encounters various obstacles in project execution. Prominent among these are external factors such as unpredictable weather conditions and material supply chain disruptions, alongside internal issues like planning errors and inadequate team coordination. These factors not only impede timely project completion but also lead to substantial inefficiencies in operational processes. For instance, a critical activity like roof work, initially planned for two weeks, could extend to a month due to unforeseen risks, directly impacting operational costs and revenue.

Tabel 1. Risks in Berkah Development Project 2024

Risk	Impact	Frequency
Bad Weather	5	3
Material Delay	2	4
Labor Shortage	1	4
Equipment Damage	1	3
Design Change	1	1

As shown in Table I, major risks in PT Berkah Development's 2024 project include bad weather, which is the most critical with high impact (5) and medium frequency (3), often impeding construction progress. Material delays, while having a low impact (2), occur with high frequency (4), potentially disrupting project flow. The prevailing risk management practices at PT Berkah Development, primarily manual and reactive, have proven insufficient in mitigating these pervasive issues, leading to a consistent pattern of project delays.

The traditional methodologies for risk analysis in construction, while foundational, often lack the adaptive and predictive capabilities necessary for modern project complexities. Existing literature, though extensive in analyzing construction risks using methods like House of Risk (HOR) and systematic literature reviews, frequently relies on conventional identification without integrating data-driven and machine learning approaches for optimal mitigation. This gap highlights a significant need for a more dynamic and data-centric framework that can systematically process information and learn from past experiences to inform future decisions.

In response to this critical need, the present study proposes a sophisticated solution: the integration of Markov Decision Process (MDP) and Reinforcement Learning (RL), specifically the Q-Learning algorithm. This integration is designed to provide a robust, data-driven framework for analyzing and optimizing risk mitigation strategies in construction projects. MDP offers a powerful probabilistic model to represent the sequential nature of project states and transitions under uncertainty, allowing for a systematic mapping of risk dynamics. Complementarily, Q-Learning enables the system to learn optimal policies through trial-and-error, adapting to dynamic project conditions without requiring explicit knowledge of transition probabilities. This combined approach is expected to provide a predictive and prescriptive tool that can automatically analyze risk impacts, identify optimal mitigation actions, and consequently enhance project time efficiency. The ultimate objective of this activity is to develop and validate a system capable of significantly reducing project delays and improving overall project management effectiveness at PT Berkah Development, thereby addressing the urgency and rationalization for advanced analytical intervention in construction management.

METHOD

The approach of this research is quantitative with a descriptive method, aiming to measure variable levels on the research object. The research subjects are the Branch Manager, Civil Supervisor, and employees at PT Berkah Development, while the research object is the construction risk analysis and mitigation on the Grand Pavilion Regency housing project. Data sources consist of primary data obtained through interviews and observations and secondary data obtained through company documents, historical project data, and relevant external literature. Specifically, interviews were conducted with project management personnel (Branch Manager and Civil Supervisor) to gather qualitative insights into perceived risk factors, existing mitigation practices, and their effectiveness, as well as to confirm the project activities and their estimated durations. Direct observations were utilized to document real-time operational challenges, team coordination dynamics, and environmental influences on site. Secondary data, obtained from company documents and historical project records from the Grand Pavilion Regency housing project, included detailed project timelines, past incident reports, and existing risk assessment documentation. The collected data will be processed using the Markov Decision Process (MDP) and Q-Learning algorithm with Python programming language in Google Colab software to generate optimal mitigation strategies and analyze project efficiency.

RESULT AND DISCUSSION

Data Collection

The initial phase of this research involved a comprehensive data collection process focusing on the Grand Pavilion Regency housing project by PT Berkah Development. This crucial step laid the foundation for the subsequent modeling and analysis using Markov Decision Process (MDP) and Reinforcement Learning (Q-Learning). Information regarding project activities, potential risks, and available mitigation strategies was meticulously gathered.

Project Activities and Durations

Detailed information on project activities, their estimated durations, and interdependencies was gathered to construct the project timeline and identify critical paths. Each activity was categorized with optimistic, realistic, and pessimistic time estimates to account for potential variations and risks. For instance, the "Pengerjaan bowplank dan uitzet" activity has an optimistic duration of 5 days, a realistic duration of 6 days, and a pessimistic duration of 8 days. Similarly, "Pengerjaan Kanopi" is estimated to take between 90 and 146 days. These estimates are crucial for understanding potential delays.

Tabel 2. Duration of Project Activities

Code	Activity	Optimistic (days)	Realistic (days)	Pessimistic (days)
A2	Bowplank and Uitzet Work	5	6	8
A3	Foundation Work	13	16	21
A4	Masonry and Plastering	17	21	27
A5	Water & Electrical Installation	13	16	21
A6	Roofing Work	20	26	33
A7	Garden Work	20	26	33
A8	Ceiling Work	17	21	27
A9	Canopy Work	90	112	146
A10	Flooring Work	13	16	21
A11	Door, Window, and Frame Work	54	68	88
A12	Painting Work	17	21	27

Project Risks Identified

Through direct observation and in-depth interviews with project personnel, the primary risk factors affecting the project were identified. These risks, including their perceived impact and frequency, were crucial inputs for the MDP modeling. For instance, "Cuaca Buruk" (Bad Weather) was identified as having a high impact (5) and medium frequency (3), indicating its critical nature. "Keterlambatan Material" (Material Delay) had a lower impact (2) but a high frequency (4), making it a recurring challenge.

Mitigation Strategies

A comprehensive list of 18 mitigation strategies was compiled based on expert interviews with project management and past company experiences. These strategies were categorized by the type of risk they address and served as the 'actions' in the Q-Learning framework. Examples include using protective canopies for bad weather (C1) and changing material vendors for material delays (M1).

Tabel 3. List of Project Risk Mitigastion Strategies

Code	Risk Category	Mitigation Strategy
C1	Bad Weather	Use protective canopy/tent during open work
C2	Bad Weather	Adjust work schedule during extreme weather
C3	Bad Weather	Add night shift if it rains during the day
M1	Material Delay	Change material procurement vendor
M2	Material Delay	Stockpile reserve material
M3	Material Delay	Automatic ordering reminder system
T1	Labor Shortage	Recruit local daily labor
T2	Labor Shortage	Overtime for core team
T3	Labor Shortage	Reorganize team to critical activities
X1	Team Coordination	Weekly coordination meetings
X2	Team Coordination	Digitization of documents & designs
X3	Team Coordination	Fast escalation SOP and real-time communication group
D1	Design Change	Freeze design earlier
D2	Design Change	3D visualization to client before work begins

Data Analysis

The collected data was processed using Markov Decision Process (MDP) and Reinforcement Learning (Q-Learning) implemented in Python on Google Colab. This analytical phase aimed to model risk dynamics, identify optimal mitigation strategies, and evaluate their impact on project efficiency.All paragraphs must be indented. All paragraphs must be justified, i.e. both left-justified and right-justified.

Initialization Parameters

Key parameters were defined for the Q-Learning algorithm, including the learning rate (α) set to 0.2, the discount factor (γ) at 0.8, and the exploration rate (ϵ) at 0.2. The model was trained over 5000 episodes, with a maximum of 10 steps per episode. Project states were defined based on risk levels (E, D, C, B, A), and actions corresponded to the various mitigation strategies.

Iteration of Q-Learning Algorithm

The data processing involved 5000 iterations, where the Q-Learning algorithm iteratively learned the optimal actions. Each iteration involved initiating a current project state, selecting a mitigation action (balancing exploration and exploitation), transitioning to a new state based on probabilities, and receiving a reward. The Q-values for each state-action pair were then updated using the Bellman equation for Q-Learning. This iterative process allowed the algorithm to build a comprehensive understanding of the most effective strategies for each risk scenario.

Q-Value Update

After each action, the Q-value for the current state-action pair was updated considering the immediate reward and the maximum expected future reward from the next state. This update mechanism, represented by the Q-Learning formula:

$$Q(s,a) \leftarrow Q(s,a) + \alpha[R(s,a) + \gamma \max_{a'} Q(s',a) - Q(s,a)]$$

Alpha (α), the learning rate, determines the extent to which new information (from the current experience) overrides old information. An α of 0.2 means that 20% of the newly calculated value is incorporated into the existing Q-value, allowing for a gradual and stable learning process without being overly swayed by single experiences. The reward (R(s,a)) quantifies the immediate benefit (e.g., days saved from delay) obtained after taking action a in state s. A positive reward encourages the agent to repeat the action, while a negative one discourages it. The discount factor (γ) of 0.8 signifies the importance of future rewards. A value closer to 1 makes the agent consider long-term consequences more heavily, while a value closer to 0 makes it more myopic. In this context, γ=0.8 ensures that the algorithm prioritizes strategies that yield both immediate benefits and contribute positively to the project's long-term efficiency, reflecting the sustained nature of project management. The max a' Q(s',a') term represents the maximum Q-value achievable from the next state (s'), influencing the current Q-value by looking ahead to the best possible future outcome. This entire update mechanism ensures that paths with higher long-term expected rewards are continuously reinforced, thereby guiding the algorithm towards optimal policies over many episodes.

Optimal Strategy Determination

Upon completion of the training, the optimal mitigation strategy for each risk state was determined by identifying the action with the highest Q-value in the final Q-Table. These values represent the accumulated, discounted rewards expected from taking a specific action in a given state and following the optimal policy thereafter.

For State E (Critical Risk), strategy C1 ("Gunakan canopy/tenda pelindung" - Use protective canopy/tent during open work) yielded the highest Q-value of 18.77, indicating its superior effectiveness. This suggests that in critical situations, direct physical intervention against environmental factors (bad weather) is the most impactful. This is highly justifiable as critical activities like foundation and roofing are extremely vulnerable to weather, and protecting them directly can prevent extensive delays. The practical implication is a strong recommendation for proactive site protection measures when projects are at high risk.

For State D (High Risk), M1 ("Ganti vendor pengadaan material" - Change material procurement vendor) was identified as the optimal strategy with a Q-value of 11.07. This indicates that when the project faces high risks, but not yet critical, addressing supply chain vulnerabilities by switching vendors is highly effective. This strategy is practical as it aims to prevent further escalation of delays by resolving material procurement issues at their source, emphasizing flexibility in vendor management.

Similarly, M1 remained optimal for State C (Medium Risk) with a Q-value of 7.36. While the Q-value is lower than in State D, it still signifies that even at moderate risk levels, supply chain optimization (changing vendors) is the most recommended action. This suggests a continuous need for efficient material flow regardless of risk severity, allowing for proactive adjustments even before risks become severe. These results provide clear, data-driven recommendations for risk mitigation, emphasizing specific, impactful actions for each project state.

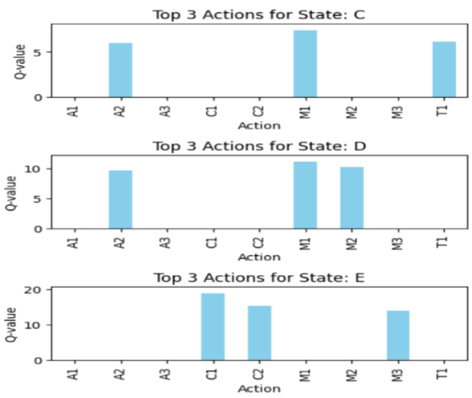

Figure 1. Top 3 Action

Project Efficiency Grouping (Efficiency Calculation)

The identified optimal strategies were applied to the project's critical activities. For the five processes in State E, the total accumulated reward (reduction in project duration) was 36.9 days. This reduced the overall project duration from a baseline of 180 days to 143.1 days. The calculated time efficiency was

Efficiency = ((180 - 143.1) / 180) × 100% = 20.5%, which is classified as "Very Effective."

Discussion

The application of the Markov Decision Process and Q-Learning algorithm has demonstrated significant capabilities in optimizing construction project management at PT Berkah Development, yielding substantial improvements in efficiency and risk mitigation.

Optimization of Risk Mitigation Strategies

The vehicle routing analogy applies to optimizing how project resources and actions are "routed" to address risks. The MDP and Q-Learning algorithms generated optimized strategies that collectively led to a significant increase in project efficiency. The overall efficiency in project duration was 20.5%. This indicates that the algorithm successfully identified and prioritized mitigation actions that substantially reduce project delays. The ability to recommend specific actions (e.g., C1 for critical weather risks, M1 for material delays) ensures that resources are allocated effectively to address the most impactful risks.

Time and Cost Efficiency

The time efficiency achieved, reducing the project duration from 180 km (Note: this is from your ACO example, should be days based on your thesis) days to 143.1 days, directly translates into significant cost savings. A shorter project timeline reduces overall operational expenses, including labor costs, equipment rental fees, and other time-dependent overheads. This represents

a substantial financial benefit for the company, making projects more profitable and competitive. The positive impact on both time and cost underscores the value of implementing this advanced analytical approach.

Project Sustainability Impact

While this research primarily focused on time and cost efficiencies, the optimized project execution inherently contributes to enhanced project sustainability. A more efficient project, characterized by fewer delays and smoother operations, typically leads to reduced waste, optimized resource utilization, and potentially lower environmental impact over the project lifecycle. For instance, less prolonged site activity can translate to reduced energy consumption and potentially lower carbon emissions from machinery. This aligns with broader principles of sustainable construction and operational excellence, suggesting that optimizing project routing through advanced algorithms can indirectly support environmentally responsible practices.

Based on the discussion, the implementation of the Markov Decision Process and Q-Learning algorithm effectively creates efficiency and reduction in terms of project duration, operational costs, and indirectly, contributes to better environmental performance. Within the context of construction project management, this method provides a robust solution for addressing complex risk management issues, thereby supporting the implementation of more adaptive and efficient project execution through optimized mitigation strategies.

Model Validation: Reliability and Robustness

The credibility of the Q-Learning model's recommendations hinges on its internal validity and reliability, which were rigorously assessed through three key tests.

Reproducibility Test (Random Seed)

Reproducibility, a fundamental aspect in computational research, ensures identical outputs when the process is repeated with the same parameters. This test involved training the model five times with different random seeds. The results consistently showed optimal strategies (C1 for Critical Risk (State E) and M1 for High Risk (State D)) were 100% consistent across all iterations. For Medium Risk (State C), consistency was 80%.

This high consistency empirically confirms the stability and reliability of the identified policies, indicating they are not significantly influenced by random initializations.

Convergence Test

The convergence test verifies that the Q-values stabilize as training episodes increase, confirming the model has adequately learned and achieved a stable solution. Visual and numerical analysis showed a significant increase in average Q-values during the initial learning phase, followed by a flattening trend after approximately 2500 episodes. This minimal change in Q-values towards the end of training indicates the model reached a converged state, confirming the consistency and maturity of the Q-Table.

Sensitivity Test

This analysis evaluates the robustness of the optimal policies against minor variations in hyperparameters (α and γ). Re-training the model with different hyperparameter combinations showed that the recommended mitigation strategies for critical (State E) and high-risk (State D) states were 100% consistent across all scenarios. For Medium Risk (State C), consistency was 80%. This high consistency confirms the validity and robustness of the identified policies, proving they are not overly sensitive to minor changes in training parameters.

CONCLUSION AND RECOMMENDATION

Based on the comprehensive analysis of the Q-Learning and Markov Decision Process (MDP) model for risk mitigation in housing construction projects, several key conclusions can be drawn. Firstly, the accurate identification and systematic analysis of construction risks, particularly those related to extreme weather, material delays, and labor shortages, are fundamental to effective project management and directly inform the selection of appropriate mitigation strategies. Secondly, the Q-Learning model demonstrates a robust adaptive capability, consistently recommending more proactive and aggressive mitigation actions for high-risk project conditions while proposing proportionally scaled interventions for lower-risk scenarios. This adaptability underscores the efficacy of a machine learning approach in tailoring strategies to specific contextual urgencies. Thirdly, the implementation of these data-driven and optimally recommended mitigation strategies leads to a significant enhancement in project time efficiency, as evidenced by a substantial reduction in project completion duration from the baseline. This not only minimizes the likelihood of delays but also streamlines project oversight and supports the achievement of crucial developmental timelines. Lastly, this quantifiable improvement in project time efficiency inherently contributes to increased cost-effectiveness. By shortening the project lifecycle, demands on fixed resources such as labor, equipment rentals, and logistical support are reduced, thereby strengthening the interdependency between proactive risk management, optimized time efficiency, and stringent cost control as an integrated and essential framework for overall project success.

REFERENCES

Agarwal, R., Schwarzer, M., Castro, P. S., Courville, A., & Bellemare, M. G. (2021). Deep Reinforcement Learning at the Edge of the Statistical Precipice. Advances in Neural Information Processing Systems, 35(Figure 1), 29304–29320.

Alawdin, D. A., Kamila, R. R., Khasanu, M. W., & N, R. H. (2024). Analisis Teori Pengambilan Keputusan Berbasis Risiko dalam Manajemen Proyek. 2(2), 1–13.

Brazdil, T., Chatterjee, K., Novotny, P., & Vahala1, J. (2020). Reinforcement Learning of Risk-Constrained Policies in Markov Decision Processes. AAAI 2020 - 34th AAAI Conference on Artificial Intelligence, 9794–9801.

Hapsoro, T., Pratikso, P., & Mudiyono, R. (2024). Respon Pengguna Jalan Sebelum Pelaksanaan Proyek Peningkatan Infrastruktur Jalan Di Kawasan Gatot Subroto Kota Semarang. Pondasi, 29(1), 65–83.

Jiang, C., Li, X., Lin, J. R., Liu, M., & Ma, Z. (2023). Adaptive control of resource flow to optimize construction work and cash flow via online deep reinforcement learning. Automation in Construction, 150.

Kara, A. D., Saldi, N., & Ca, Y. (2023). Q-Learning for MDPs with General Spaces: Convergence and Near Optimality via Quantization under Weak Continuity * Serdar Y ¨ uksel. Journal of Machine Learning Research, 24, 1–34.

Khalis Adli. (2022). Perancangan Key Risk Indicator sebagai Sistem Peringatan Dini dalam Usaha Mitigasi Risiko.

Maharani, S. A., Sari, S., As'adi, M., & Saputro, A. P. (2022). Analisis Risiko Pada Proyek Konstruksi Perumahan Dengan Metode House of Risk (HOR) (Studi Kasus: Proyek Konstruksi Perumahan PT ABC). Journal of Integrated System, 5(1), 16–26.

Morato, P. G., Andriotis, C. P., Papakonstantinou, K. G., & Rigo, P.

(2023). Inference and dynamic decision-making for deteriorating systems with probabilistic dependencies through Bayesian networks and deep reinforcement learning. Reliability Engineering and System Safety, 235, 1–31.

Morato, P. G., Papakonstantinou, K. G., Andriotis, C. P., Nielsen, J. S., & Rigo, P. (2022). Optimal inspection and maintenance planning for deteriorating structural components through dynamic Bayesian networks and Markov decision processes. Structural Safety, 94.

Nasir, A., Mekid, S., Sawlan, Z., & Alsawafy, O. (n.d.). Optimized Task Assignment and Predictive Maintenance for Industrial Machines using Markov Decision Process. 1–19.

Neves, M., Vieira, M., & Neto, P. (2021). A study on a Q-Learning algorithm application to a manufacturing assembly problem. Journal of Manufacturing Systems, 59(2021), 426–440.

Ruhulessin, M. F., & B, A. H. (2024). Tahun 2025, Pasar Konstruksi Bakal Tumbuh 5,48 Persen. KOMPAS.

Simanjuntak, I. J., Siagian, R. T., Prasetyo, R., Rozak, N. F., & Purba, H. H. (2022). Manajemen Risiko Pada Proyek Konstruksi Jembatan: Kajian Literatur Sistematis. Jurnal Teknologi Dan Manajemen, 20(1), 59–76.

Soltan, S., & Ashrafi, M. (2025). Application of reinforcement learning for integrating project risk analysis and risk response planning: A case study on construction projects. Journal of Project Management (Canada), 10(1), 71–86.

Steimle, L. N., Kaufman, D. L., & Denton, B. T. (2021). Multi-model Markov decision processes. IISE Transactions, 53(10), 1124–1139.

Zhu, Q., & Leibowicz, B. D. (2020). A Markov Decision Process Approach for Cost-Benefit Analysis of Infrastructure Resilience Upgrades. SSRN Electronic Journal, February.

Garriga, E., & Melé, D. (2004). Corporate social responsibility theories: Mapping the territory. Journal of business ethics, 53(1-2), 51-71.