American Journal of Data Science and Analytics

Volume: 7; **Issue:** 10
Pages: 126-164

Intelligence Driven Project Risk Prediction and Resource Optimization for Complex Engineering and Manufacturing Projects

Asheke Mostofa[1];

[1]. *M.S. in Engineering Management, Lamar University, Beaumont, Texas, USA;*
Email: asheke.mbstu@gmail.com

Doi: 10.63125/hv05m379
Received: *13 July 2026;* ***Revised:*** *12 August 2026;* ***Accepted:*** *09 September 2026;* ***Published:*** *11 October 2026;*

Abstract

Artificial intelligence–driven project risk prediction and resource optimization address uncertainty and competing resource requirements in complex engineering and manufacturing projects. This study developed a quantitative framework linking early risk estimates with feasible labor, equipment, material, and contingency decisions. A synthetic numerical example represented 240 projects across 12 organizations, with 192 projects assigned to development and 48 to chronological testing. The framework specified retrospective longitudinal analysis, statistical and machine-learning comparisons, probability assessment, and matched resource-policy evaluation. Regression, classification, and scheduling outcomes were distinguished, while project-level separation and organizational clustering informed the proposed analytical structure. In the findings, gradient boosting achieved 93.75% delay-classification accuracy, 94.12% sensitivity, a 91.43% F1-score, and a receiver operating characteristic area of 0.970. Its Brier score was 0.071, and mean warning lead time was 14.8 days. Constructed adjusted associations linked complexity, technical changes, procurement interruptions, and equipment downtime with adverse outcomes, whereas workforce availability showed an inverse relationship. Integrated optimization reduced makes pan from 462 to 402 days, equivalent to 12.99%, and resource expenditure from USD 6.80 million to USD 5.90 million per project, equivalent to 13.24%. Equipment utilization increased by 9.4 percentage points, while recovery duration declined from 6.5 to 3.8 days. Nested-validation accuracy was 92.71%, illustrating the distinction between apparent development performance and validation results. Subgroup and disruption comparisons demonstrated reporting approaches for contextual variation, resource flexibility, and operational sensitivity. Independent external validation remained absent, and repeated evaluations did not establish replication. However, all numerical findings, intervals, and significance measures were constructed examples rather than estimates from observed records. They did not establish actual model superiority, causal effects, financial savings, or external generalizability. The study contributed an integrated analytical framework that separated predictive reliability, feasible resource actions, and operational outcomes, while identifying the empirical evidence required to evaluate their relationship in engineering and manufacturing practice.

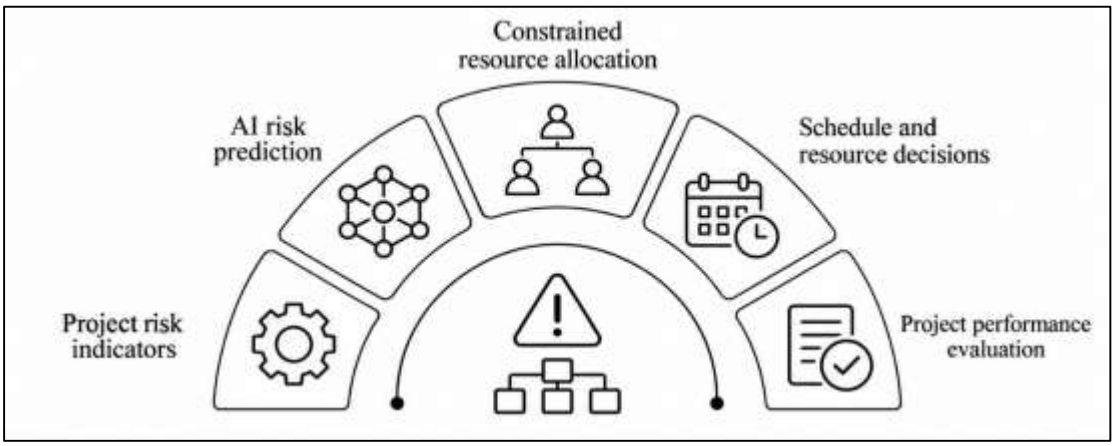
Keywords

Prediction; Optimization; Risk; Engineering; Manufacturing; Contingency Decision;

INTRODUCTION

Artificial intelligence–driven project risk prediction uses computational models to estimate the probability, timing, and severity of events that threaten project objectives. Resource optimization involves allocating limited labor, equipment, materials, capital, and production capacity while respecting technical dependencies and operational constraints (Harle, 2024). Complex engineering and manufacturing projects combine specialized activities, uncertain conditions, and coordination across organizational boundaries. Internationally, these projects support infrastructure development, energy production, industrial expansion, and manufacturing competitiveness. Their performance depends on identifying emerging disruptions and allocating resources effectively across interconnected activities. Design changes, procurement delays, financing constraints, workforce shortages, and equipment failures can propagate through project networks, increasing completion time and expenditure. Evidence synthesized across more than thirty studies encompasses Bayesian risk assessment, machine learning prediction, resource-constrained project scheduling, and dynamic manufacturing optimization. Bayesian networks represent dependencies among risks, while neural networks, support vector machines, random forests, and gradient boosting estimate delay occurrence, project duration, and cost overruns (Kuang et al., 2021).

Figure 1: Intelligent Risk Prediction and Optimization



One infrastructure investigation analyzed 293,263 tasks from 302 completed projects and reported improved delay prediction using a hybrid learning and simulation approach. Another investigation used 198 green-building projects to develop neural networks for cost and duration estimation. However, reported performance varies with sample size, outcome definition, data quality, and validation procedures. Models developed from expert assessments represent different evidence from models tested against independently observed project outcomes. Optimization research complements prediction through genetic algorithms, constraint programming, and reinforcement learning that select feasible schedules under resource limitations. Studies address activity sequencing, alternative execution modes, work-package sizing, maintenance scheduling, and financial exposure. Manufacturing investigations extend these approaches to machine breakdowns, urgent orders, setup requirements, and fluctuating workloads. Graph representations, attention mechanisms, and interacting decision agents capture dependencies within changing production environments (Radanliev et al., 2020). Across these research streams, completion time, total cost, tardiness, resource utilization, and schedule stability remain distinct performance measures. Higher utilization can increase queues, while frequent rescheduling can disrupt operational coordination. The evidence therefore establishes a quantitative foundation for examining risk information and resource decisions together, with explicit attention to differences between engineering project structures and recurring manufacturing operations. These differences shape the interpretation of reported improvements across sectors, organizations, and operating conditions.

The integration of project risk prediction and resource optimization presents an empirical challenge because predictive accuracy does not automatically establish effective operational decisions. A model can identify likely delays without specifying feasible responses, while an optimizer can generate efficient schedules using assumptions that underrepresent disruption exposure. Existing studies employ historical project records, expert surveys, computational benchmarks, and simulation experiments, each supporting different claims about reliability and generalizability (Cooper, 2024). Comparisons require consistent information timelines, resource capacities, disturbance scenarios, and performance definitions (Aditya & Ahmed, 2022; Asheke & Maksudul, 2022). Data leakage occurs when final costs, completion dates, or information recorded after an adverse event enter models intended for earlier prediction. Randomly dividing activities from the same project between training and testing datasets can also exaggerate predictive performance (Asheke & Badrul, 2022; Khalid & Sadmeen, 2022; Akter & Sumaiya, 2022). Evaluation across independent projects, later periods, and separate facilities provides stronger evidence of transferability. Classification assessment requires precision, recall, discrimination, calibration, and warning lead time, while continuous prediction requires appropriate error measures and explained variance (Mashiat & Morshedul, 2022; Imran & Jannatul, 2022; Maksudul & Islam, 2022). Optimization evaluation additionally requires feasibility checks, comparable computational budgets, repeated runs, and uncertainty estimates. Within this context, this study aims to examine relationships among artificial intelligence–based risk estimates, constrained resource allocation, and engineering and manufacturing project performance (Morshedul & Islam, 2022; Feroz & Pervaz, 2022; Feroz & Shabbir, 2022). Its first objective is to identify indicators associated with schedule deviation, cost overrun, resource shortages, and disruption occurrence. Relevant variables include project complexity, technical changes, procurement lead times, approval duration, workforce availability, equipment condition, and workload intensity (Suriyaamporn et al., 2024). The second objective is to compare statistical and machine-learning models using consistent outcomes and validation procedures that preserve project and temporal boundaries (Islam & Md, 2023; Mostari & Kynatun, 2023; Mashfiqur & Jamil, 2023). The third objective is to evaluate the relationship between predicted risks and resource decisions involving activity sequencing, labor assignment, machine allocation, contingency capacity, and schedule revision. Additional objectives examine variation across project categories, scales, resource flexibility, and disruption intensity, together with sensitivity to uncertain inputs (Mashiat & Binte, 2023; Shahid, 2023; Islam & Ahmed, 2023). Quantitative assessment incorporates regression estimates, confidence intervals, effect sizes, predictive errors, and repeated scheduling comparisons (Ali & Alrobaian, 2024). The analytical scope distinguishes observational associations from controlled optimization results and examines whether integrated prediction and allocation outperform fixed planning rules or prediction alone under equivalent operational constraints (Adnan, 2024; Sadia & Mosheur, 2023; Sadmeen & Khalid, 2023). Particular attention concerns whether warnings occur early enough to permit feasible adjustments before disruption (Islam, 2024; Ashik, 2024; Yeasin, 2024b).

The primary objective of this study is to quantitatively examine the effectiveness of artificial intelligence–driven project risk prediction and resource optimization in improving the performance of complex engineering and manufacturing projects. Specifically, this study seeks to identify and measure the relationships between project complexity, design changes, procurement uncertainty, workforce availability, equipment reliability, financial constraints, and adverse outcomes, including schedule delays, cost overruns, resource shortages, and operational disruptions. It aims to compare the predictive performance of statistical and machine-learning approaches, including logistic regression, random forests, gradient boosting, support vector machines, and artificial neural networks, using consistent datasets, outcome definitions, and validation procedures. Predictive performance is assessed through appropriate measures of classification accuracy, precision, recall, F1-score, discrimination, calibration, and prediction error, depending on the characteristics of each outcome. Another objective is to evaluate how predicted risk probabilities and expected disruption severity inform resource-allocation decisions involving labor assignments, equipment scheduling, material availability, activity sequencing, and contingency capacity. This study also seeks to compare integrated prediction and optimization approaches with conventional planning methods under equivalent resource limitations, project requirements, and disruption scenarios. These comparisons focus on completion duration,

expenditure, resource utilization, task tardiness, schedule stability, and recovery time following operational disturbances. Furthermore, the study aims to estimate the magnitude and statistical significance of associations between intelligent decision-support capabilities and project outcomes while accounting for organizational size, project scale, technological maturity, industrial category, and baseline performance. Subgroup analyses examine whether these relationships differ according to project complexity, resource flexibility, manufacturing environment, and disruption intensity. An additional objective is to assess model robustness and generalizability through validation across independent projects, chronological testing, sensitivity analysis, and repeated optimization experiments where the available data support these procedures. The analytical framework distinguishes predictive relationships from causal effects and separates computational scheduling improvements from performance changes demonstrated in operational records. Through these objectives, this study examines the conditions under which combining early risk identification with feasible resource decisions is associated with measurable improvements in engineering and manufacturing project execution.

LITERATURE REVIEW

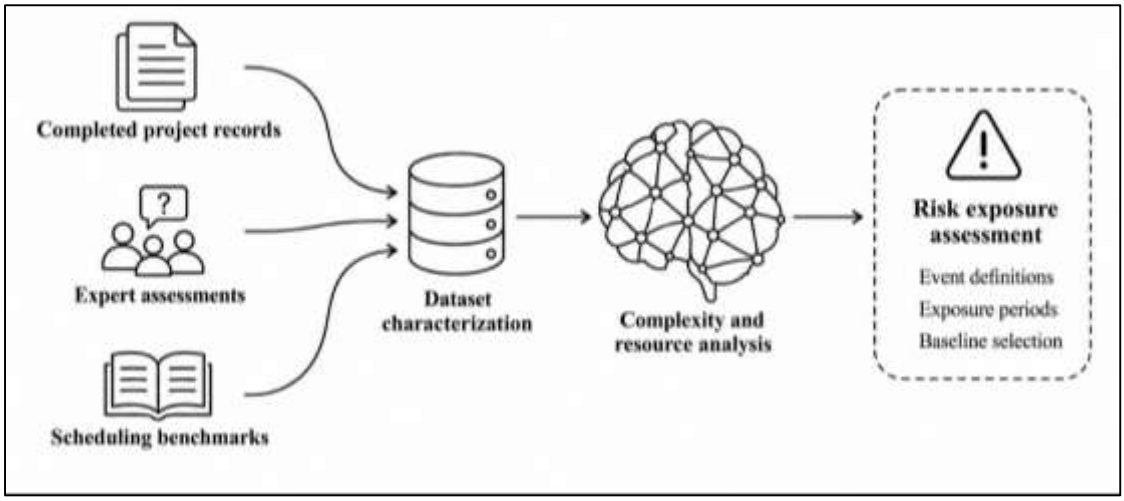
The literature on artificial intelligence–driven project risk prediction and resource optimization examines how project information, predictive models, and allocation decisions relate to measurable engineering and manufacturing outcomes. This research encompasses statistical risk assessment, machine-learning prediction, resource-constrained scheduling, manufacturing optimization, and quantitative project-performance evaluation. Engineering studies commonly address completion delays, cost overruns, procurement uncertainty, and coordination difficulties, while manufacturing investigations examine equipment failures, production congestion, workforce constraints, and changing processing requirements. These research streams are connected through their examination of uncertainty, interdependent activities, and limited resources, although their datasets, decision horizons, and performance measures differ. The literature review is organized into eight subsections that distinguish project characteristics, risk predictors, predictive performance, optimization methods, integrated decision outcomes, disruption recovery, contextual differences, and validation quality. Particular attention is given to sample size, measurement reliability, comparison methods, effect magnitude, uncertainty, and the distinction between computational improvements and observed operational benefits. Predictive accuracy is examined alongside calibration and warning lead time, while optimization performance is assessed through completion time, expenditure, resource utilization, delivery reliability, and schedule stability. This organization establishes the conceptual and empirical basis for quantitatively examining the relationships among artificial intelligence capabilities, project risk exposure, constrained resource decisions, and engineering and manufacturing project performance.

Project Complexity in Engineering and Manufacturing Datasets

The quantitative literature on engineering and manufacturing project management encompasses datasets with substantially different observation units, making sample size an important but incomplete indicator of evidential strength (Yousif et al., 2024; Aditya & Ahmed, 2022). An infrastructure schedule-risk investigation examined 293,263 tasks from 302 completed projects, providing extensive activity-level information within a smaller number of independent project environments. In comparison, research on green-building cost and duration prediction used 198 projects, while a tall-building delay investigation developed its dataset from 48 usable expert responses covering 36 risk factors. A Bayesian schedule-risk study identified 16 factors through a questionnaire involving 166 professionals, illustrating another distinction between the number of respondents and the number of observed project outcomes. Joint work-package sizing and scheduling research examined generated instances containing 15, 30, 60, 90, and 120 tasks alongside a modular construction application. These studies represent different forms of quantitative evidence: completed projects, nested activity records, professional judgments, and computational instances (Hou et al., 2023; Asheke & Maksudul, 2022). Their sample counts cannot be treated as interchangeable because thousands of activities belonging to a limited number of projects do not provide thousands of independent organizational observations. Manufacturing datasets introduce additional levels through jobs, operations, machines, batches, and repeated scheduling decisions. Observation periods are similarly

relevant because a short production window captures different operating variation from records covering several project phases or maintenance cycles. Across these research designs, dataset characterization therefore depends on the number of independent projects or facilities, the density of repeated observations, and the coverage of operating conditions. Counts of employees, suppliers, and machines add contextual information when they describe the actual resource environment rather than serving merely as identifiers (Kang et al., 2022; Asheke & Badrul, 2022).

Figure 2: Project complexity and risk exposure



Research on project complexity and resource allocation represents complexity through several related but distinct dimensions. Studies of resource-constrained scheduling encode structural complexity through activity counts, precedence relationships, alternative execution modes, and competition for limited resources. Research on joint work-package sizing and scheduling adds the granularity of work organization, recognizing that dividing activities into smaller packages changes coordination requirements and resource demand. Investigations of alternative project subgraphs extend this representation by allowing different technical routes or activity combinations (Mahmud et al., 2022; Yeasin, 2024a; Jahanara, 2024; Akter & Sumaiya, 2022). In contrast, infrastructure contingency-estimation research incorporates a complexity index into cost prediction, linking project characteristics to uncertainty around expenditure (Kynatun & Saba, 2024; Mostari, 2024; Mashfiqur, 2024). Bayesian investigations of tunnel delays and construction cost overruns represent complexity through dependencies among risk factors, emphasizing conditional relationships rather than activity counts alone. These approaches show that project size, network complexity, and risk interdependence describe different properties. A project with relatively few activities can still involve tightly coupled interfaces, unfamiliar technology, or restricted specialist capacity (Almalawi et al., 2024; Masrul & Mashfiqur, 2024; Masrul & Kanti, 2024; Hasinur, 2024; Imran & Jannatul, 2022). Resource constraints also differ in their treatment across studies. Conventional scheduling benchmarks commonly represent resources as capacity categories, whereas aircraft maintenance research considers workgroups and interrelated maintenance tasks. Manufacturing investigations incorporate machine eligibility, processing requirements, setup conditions, and equipment availability (Maksudul, 2024; Shahid, 2024; Islam, 2024). Workforce skill restrictions, material replenishment, workspace capacity, and budget limitations provide further distinctions when explicitly represented. A numerical resource-capacity value therefore does not necessarily correspond to an employee count or machine count (Kgakatsi et al., 2024; Maksudul & Islam, 2022). Comparisons across studies are most interpretable when the underlying resource categories, renewal assumptions, substitutability, and simultaneous demand are described. These measurement differences explain why similar activity counts can conceal substantially different scheduling difficulty and operational exposure (Mainuddin, 2024; Sumaiya, 2024; Tahmedur, 2024).

Computational scheduling studies provide controlled descriptions of project environments, although their instance counts represent experimental problems rather than completed industrial projects (Islam, 2025; Asheke, 2025; Ashik, 2025). The established project-scheduling benchmark literature includes single-mode sets with 30, 60, 90, and 120 activities. The first three sets contain 480 instances each, while the 120-activity set contains 600 instances, commonly using four renewable resource categories (Hou et al., 2023; Morshedul & Islam, 2022). Investigations involving particle swarm optimization with justification mechanisms, particle swarm hyper-heuristics, consolidated optimization, and reinforcement-learning coordination have used these benchmarks to compare scheduling methods under specified precedence and capacity conditions (Kynatun, 2025; Mostari, 2025; Masrul, 2025). Their common structure supports reproducible comparison, but it does not directly establish the prevalence of workforce shortages, procurement interruptions, or equipment failures in operating organizations. Stochastic project-scheduling research introduces uncertainty in activity duration, while graph-based reinforcement-learning research extends resource-constrained scheduling to disruptions in resource availability (Wu et al., 2023; Stabak & Shahid, 2022). These extensions alter the exposure represented by the dataset and therefore change the interpretation of algorithmic performance. In observed project records, delay frequency requires a defined completion baseline and an explicit number of eligible projects (Hasinur, 2025a, 2025b; Maksudul, 2025). Cost-overrun prevalence similarly depends on the budget reference, treatment of approved scope changes, and availability of final expenditure. Equipment-failure incidence requires a denominator such as operating time or machine exposure, while shortage duration concerns the length of an interruption rather than its occurrence alone (Shahriar, 2025; Shahid, 2025; Islam, 2025). Expert risk ratings describe perceived likelihood or consequence and cannot independently establish these observed event rates (Lv et al., 2024; Feroz & Pervaz, 2022). Consequently, risk exposure across the literature is characterized by event definitions, exposure periods, baseline selection, and the distinction between recorded disturbances and experimentally generated uncertainty (Mainuddin, 2025; Sadia, 2025; Shabbir, 2025).

Workforce Shortages and Cost Overruns

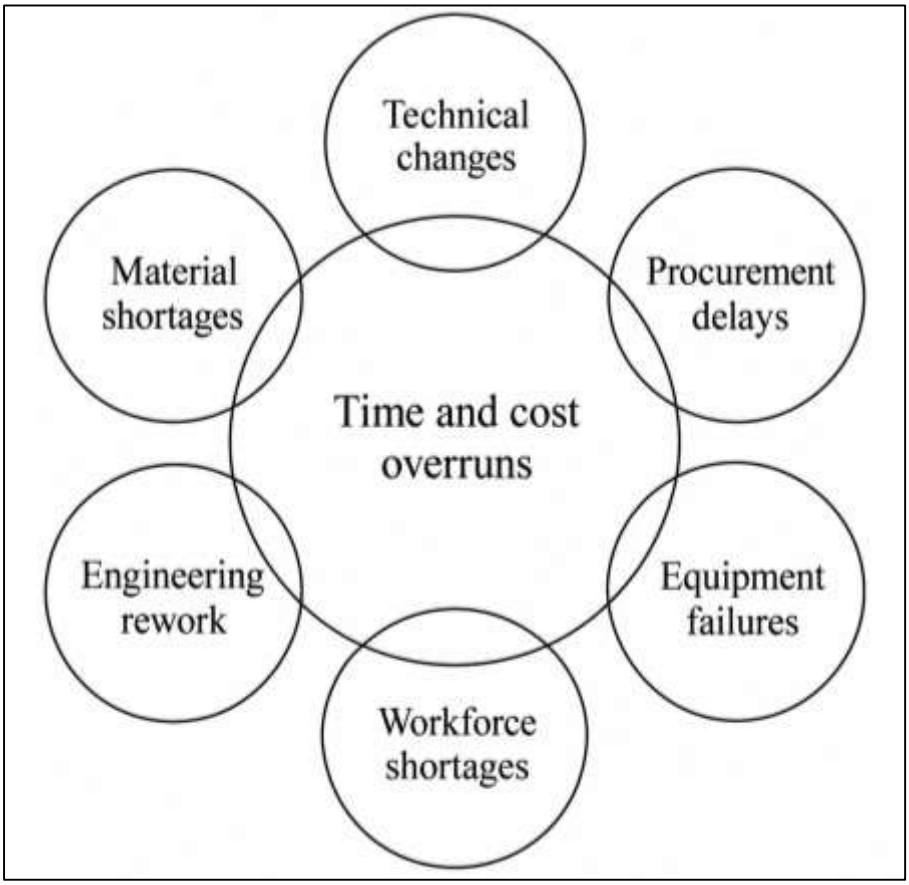
Quantitative research associates technical changes and rework with additional expenditure and disruption to planned activity sequences, although the magnitude depends on how changes and project outcomes are measured. An investigation of post-contract design changes reported direct change costs equivalent to 5.1%–7.6% of total project cost. These figures described expenditure attributable to design changes rather than an equivalent percentage increase above the approved budget (Alshihri et al., 2022; Feroz & Shabbir, 2022). A separate civil-infrastructure investigation identified five significant predictors explaining approximately 25% of variation in total rework cost, including information-technology deficiencies, unclear working procedures, client involvement, and change-related practices. Research on highway projects in Palestine also reported a significant regression relationship between change orders and rework cost, connecting specification nonconformance, late design changes, and workforce capability with repeated work (Rahman, et al., 2025; Shafi, et al., 2025; Molla, et al., 2025). A Korean apartment-project investigation developed a cost-prediction model specifically for design changes, further distinguishing change-related expenditure from overall project cost. Across these studies, technical modifications represent both direct cost items and potential sources of interruption, resequencing, and additional resource demand (Rivera et al., 2020; Islam & Md, 2023). Their statistical interpretation requires separation of change frequency, change value, timing, and technical severity because a small number of late modifications can differ substantially from numerous minor early revisions (Islam, et al., 2025; Molla, et al., 2025; Shehun, et al., 2025). Approval delays and incomplete information add coordination-related exposure, but they can also reflect project complexity or weak management. Consequently, an unadjusted correlation between design changes and overruns does not isolate the contribution of changes from these shared conditions (Chandragiri et al., 2021; Asheke, 2023). Regression coefficients and explained variance provide stronger contextual information when outcome definitions, adjustment variables, and uncertainty estimates are reported consistently (Ashik, 2026; Sadmeen, 2025; Tahmedur, 2025).

Procurement research connects material and equipment availability with project time and cost performance through purchasing processes, supplier behavior, transportation conditions, and organizational coordination. An investigation involving 22 highway projects in Nepal examined

procurement delays and their cost consequences, identifying organizational weaknesses, supplier defaults, governmental regulations, and transportation delays as prominent sources. Research on civil-engineering material management in Sri Lanka similarly identified market shortages, procurement delays, delivery problems, price fluctuations, and inadequate planning as recurring difficulties associated with project delivery (Jain et al., 2024; Md Shahid, 2023).

A systems-dynamics investigation of engineering, procurement, and construction projects represented material supply as an interacting process involving procurement decisions, delivery timing, and work backlogs. Structural-equation research on construction cost overruns additionally linked management conditions, resource distribution, and external influences, indicating that procurement exposure cannot be interpreted independently of organizational arrangements. These studies connect supply problems with performance through different evidence types: reported experiences, ranked factors, statistical relationships, and simulated feedback mechanisms (Yeasin, 2026; Hasib, 2026; Kynatun, 2026). A high importance ranking does not establish a large adjusted regression coefficient, and a simulated cost increase does not constitute an observed causal effect. Material lead-time variability also differs from average lead time because unpredictable deliveries complicate coordination even when average delivery duration appears acceptable (Ezzat, 2024; Islam & Ahmed, 2023).

Figure 3: Risk associations with project overruns



Inventory shortages can interrupt activities, while protective inventory creates holding costs and capital requirements. Dependence on critical suppliers represents another exposure whose consequences depend on substitution options and activity criticality (Mostari, 2026; Hasinur, 2026; Maksudul, 2026). Pearson correlations describe linear relationships between continuous measurements, whereas Spearman correlations accommodate ranked or monotonic relationships (Mohsin, 2026; Shahriar, 2026; Shahid, 2026a). Neither independently resolves confounding from project scale, financial distress, contractual arrangements, or management quality, which can influence procurement reliability and overruns simultaneously (Nikbakht et al., 2024; Sadia & Md. Mosheur,

2023).

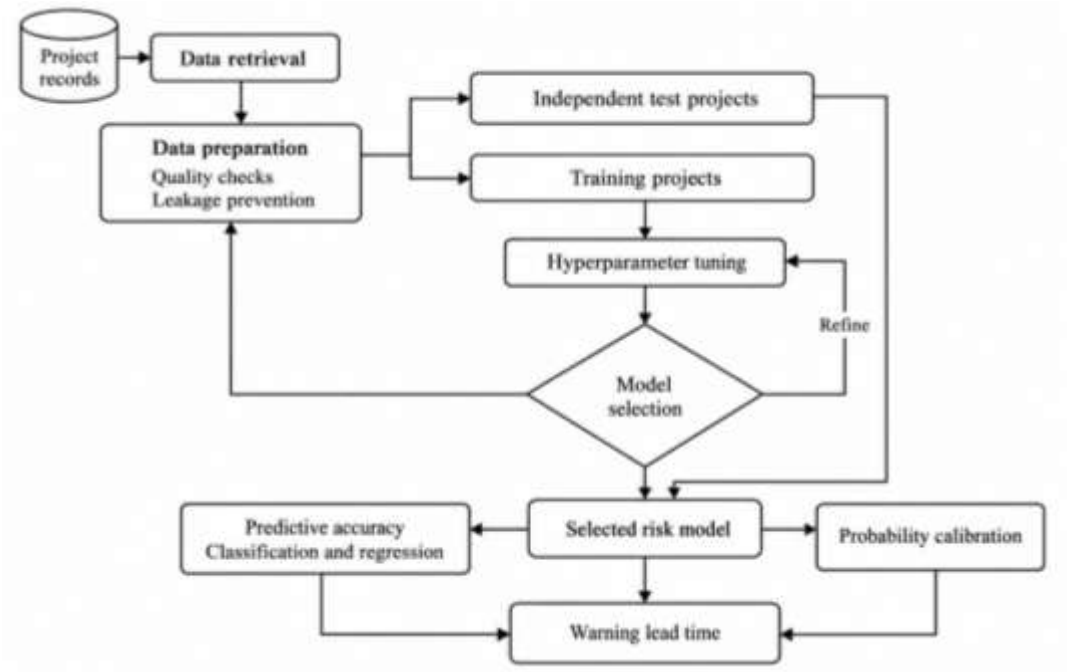
Workforce research demonstrates that available labor hours and productive capacity are related but distinct quantities. A quantitative overtime investigation analyzed 121 weeks of productivity data from four industrial projects, providing repeated observations of labor performance under different working schedules (Shahid, 2026b; Mainuddin, 2026; Sadia, 2026). A separate study of extended overtime used regression analysis and analysis of variance and reported declining productivity as weekly working hours or the duration of overtime increased (Shabbir, 2026; Sumaiya, 2026; Sadmeen, 2026). Together, these investigations challenge the assumption that additional scheduled hours generate proportionate increases in completed work. Their findings concern productivity relationships within the examined project conditions rather than a universal overtime penalty applicable to every trade (Lingard & Turner, 2022; Adnan, 2024). Research on construction-site absenteeism developed a regression model involving employment tenure and absenteeism at adjacent supervisory levels, showing that labor availability is also connected to employment arrangements and organizational relationships. Highway rework research identified insufficient labor experience, skill limitations, and subcontractor selection among relevant conditions, while tall-building delay assessment highlighted subcontractor progress and decision-related constraints. These studies connect workforce capability and coordination with performance through several pathways, including reduced output, repeated work, waiting time, and interrupted activity handovers (Karan et al., 2021; Aminul Islam, 2024). However, productivity, absenteeism, turnover, and project delay remain separate outcomes. A predictor of absenteeism does not automatically establish an independently adjusted effect on total project duration (Umme, 2026). Overtime also presents temporal ambiguity because struggling projects often introduce longer schedules in response to existing delays. Similarly, additional recruitment can follow poor performance rather than precede it. Adjusted analyses therefore have a different interpretation from simple correlations when they account for work type, crew composition, project phase, supervision, and baseline progress (Winkelmann et al., 2022; Ashik, 2024). Confidence intervals and effect magnitudes are particularly relevant where repeated weekly observations arise from only a few independent projects.

Manufacturing research links equipment reliability with schedule performance through machine availability, processing capacity, maintenance requirements, and the position of affected operations within production sequences. Joint production-scheduling and preventive-maintenance research models total cost as dependent on both production decisions and maintenance arrangements, explicitly accounting for downtime associated with unavailable equipment (Reich-Stiebert et al., 2023; Kynatun & Saba, 2024). A separate investigation of real-time production scheduling and condition-based maintenance incorporates machine deterioration, processing-time changes, repairs, and maintenance costs. Research on individual disruptions in multistage manufacturing combines field information from an automobile assembly line with simulation to estimate disruption costs, demonstrating that the consequences of an interruption depend on the surrounding production system. Deep reinforcement-learning studies of scheduling under machine failures and research on schedule repair further examine completion delays and deviations from original plans under controlled disturbance conditions. These contributions provide evidence about modeled operational mechanisms, while their computational comparisons remain distinct from observational regression estimates of project overruns. Failure frequency, downtime duration, maintenance backlog, and processing variability describe different aspects of equipment exposure and can be strongly correlated (Gerhardt et al., 2021; Maksudul, 2024). Their combined inclusion can alter coefficient signs, precision, and apparent importance. Interactions are also relevant because equipment failure can coincide with scarce repair skills, missing spare parts, or unavailable substitute capacity. For binary overrun outcomes, adjusted odds ratios express conditional associations; for continuous time or cost deviations, regression coefficients and explained variance address different questions. These measures cannot be combined directly with simulation-based percentage improvements. Organizational maturity, project complexity, workload, and baseline equipment condition can confound observed relationships (Pamidimukkala & Kermanshachi, 2021; Shahid, 2024). The literature therefore supports an interpretation of interconnected technical, procurement, workforce, and equipment risks while retaining a clear distinction between statistical association, model-based sensitivity, and demonstrated causal effects.

Machine-Learning and Deep-Learning Project Risk Models

Research on project-risk classification compares statistical, probabilistic, and machine-learning approaches for identifying projects or activities associated with adverse outcomes. Logistic regression provides an interpretable probability-based baseline, while decision trees, random forests, gradient boosting, support vector machines, nearest-neighbor methods, and neural networks represent alternative ways of capturing relationships between project characteristics and risk labels (Khodabakhshian et al., 2023; Md Tazul Islam, 2024). A tall-building investigation compared nearest-neighbor, neural-network, support-vector, and ensemble approaches using 48 usable expert responses and reported 93.75% accuracy for its selected neural-network model. A separate delay-prediction study combining random forests with genetic optimization reported 91.67% accuracy, a kappa value of 0.87, and an 8.33% classification error. These figures do not establish that neural networks outperform optimized random forests because the studies used different observations, predictors, and evaluation conditions. Another comparative delay investigation evaluated logistic regression, decision trees, random forests, boosting methods, and probabilistic baselines through accuracy, precision, recall, and F1-score (Adeoye et al., 2021; Mainuddin, 2024).

Figure 4: Evaluating project risk prediction models



Bayesian cost-overrun research additionally represented dependencies among risk indicators rather than relying solely on independent predictor contributions. Across these approaches, precision describes the reliability of positive warnings, whereas recall, or sensitivity, describes the proportion of adverse cases detected. Specificity concerns correctly identified negative cases, and F1-score summarizes the balance between precision and recall. Receiver operating characteristic area evaluates discrimination across thresholds, while precision–recall area is particularly relevant when adverse events are uncommon. Matthews correlation coefficient incorporates all four classification outcomes. Consequently, reported accuracy alone provides insufficient evidence about missed critical events, false-warning burdens, or performance under unequal class frequencies (Mehdiyev et al., 2020; Sumaiya, 2024).

Continuous-outcome studies address a different predictive task by estimating completion duration, final expenditure, delay magnitude, or disruption severity rather than assigning an adverse-event category. Research comparing nearest-neighbor regression, support vector regression, gradient boosting trees, and neural networks used field information on progress, workload, labor, weather,

planned duration, and location to predict construction duration at the work-area level (Liu et al., 2024; Taher et al., 2024). Other investigations developed neural networks for project-duration estimation and random forest regression for engineering-services cost overruns, illustrating differences in both prediction targets and project scope. A green-building investigation using 198 projects reported mean absolute percentage error values of 0.07 for cost and 0.09 for duration; integrating deep neural networks with support vector regression reduced the reported values to 0.06 and 0.07, respectively (Pachiyannan et al., 2023; Ashik, 2025). These within-study comparisons provide clearer evidence of incremental improvement than comparisons between unrelated datasets. An infrastructure investigation combining an empirical Bayesian network, support vector classification, and Monte Carlo simulation used 293,263 tasks from 302 completed projects and reported 54.4% more accurate delay predictions against its Monte Carlo comparator on 10 selected projects. That improvement remains specific to its evaluation method and cannot be interpreted as an equivalent reduction in actual project delays. Mean absolute error retains the outcome's measurement unit, while root mean squared error gives greater weight to large deviations. Explained variance describes predictive fit but does not directly express operational error magnitude. Percentage errors support relative comparisons when denominators are meaningful, although zero or very small observed values can distort them (Iyortsuun et al., 2023; Kynatun, 2025). The literature therefore requires attention to target scale, aggregation level, comparator strength, and the distinction between forecasting error and realized project performance.

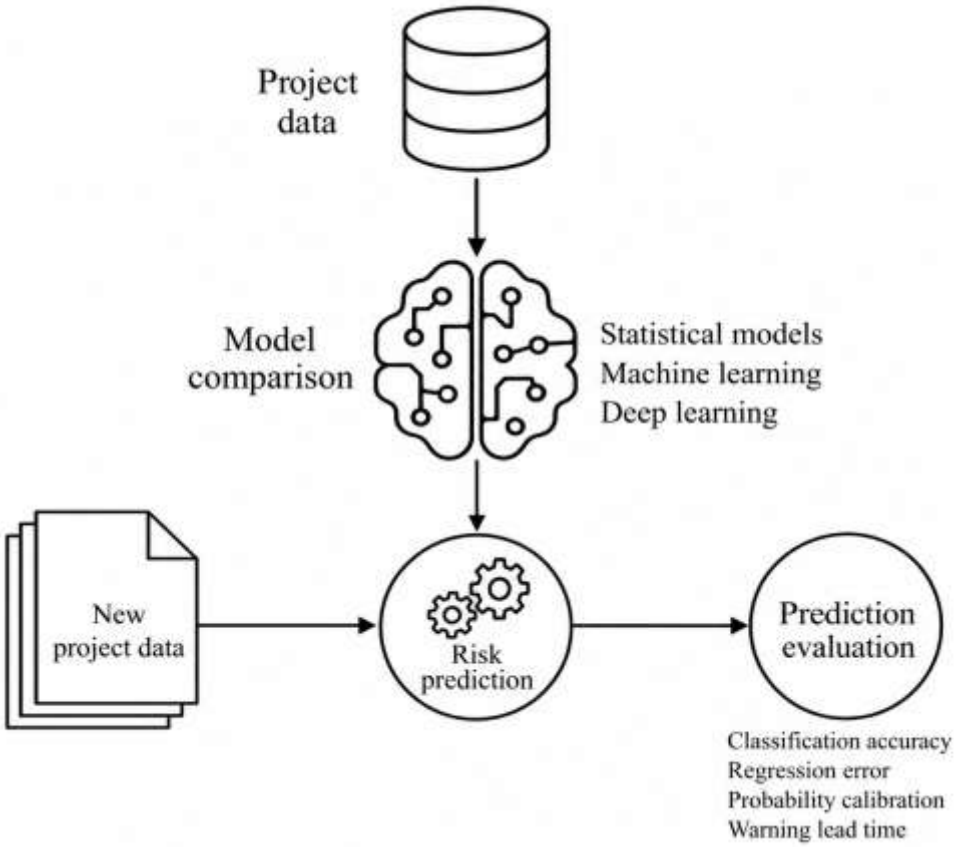
Probabilistic project-risk research distinguishes the ability to rank risky cases from the reliability of the probabilities assigned to them. Bayesian-network investigations of construction cost overruns and dynamic cost-overrun prediction explicitly represent uncertainty, but probabilistic model structure alone does not establish calibration against independent observations. Calibration concerns agreement between predicted event probabilities and observed frequencies at the relevant decision horizon (Yekeen & Balogun, 2020; Maishar Mostari, 2025). A model can separate higher-risk from lower-risk projects accurately while systematically overstating or understating their absolute risk. Research on infrastructure contingency estimation addressed uncertainty through neural-network interval predictions using Monte Carlo dropout and bootstrapping, demonstrating an approach to quantifying variation around estimated costs. However, interval uncertainty for a continuous outcome differs from calibration of a binary overrun probability (Nosratabadi et al., 2020; Masrul, 2025). A probabilistic delay-and-cost framework using 46 high-rise projects reported a median concordance index of 0.81 compared with 0.73 for its baselines, alongside an integrated Brier-score reduction of 0.027 and uncertainty expressed through credible intervals. The concordance index describes time-to-event discrimination, while the Brier score evaluates probabilistic prediction error and reflects more than calibration alone. These findings illustrate why improved discrimination cannot independently establish reliable risk probabilities. Calibration curves, probability-error measures, and interval coverage provide complementary evidence, but their interpretation depends on event prevalence, follow-up duration, and evaluation sample size. Risk estimates also change when models are transferred between organizations with different project characteristics or adverse-event frequencies. Within this literature, predictive uncertainty, probability calibration, model explanation, and decision benefit remain connected but analytically distinct dimensions of reliability.

Project Makespan, Resource Utilization, and Allocation Cost

Research on resource-constrained project scheduling examines how activity sequencing and resource assignment affect completion time while preserving precedence relationships and capacity limits (Qureshi et al., 2024)(Md Maksudul, 2025). Genetic algorithms, particle swarm optimization, ant colony optimization, and simulated annealing represent different search strategies for identifying feasible schedules within large solution spaces. Aircraft heavy-maintenance research applied genetic algorithms to interdependent maintenance activities and restricted workgroups, evaluating maintenance makespan and computational efficiency. Particle swarm research incorporated justification mechanisms that reposition activities to improve schedules, while a separate particle swarm hyper-heuristic investigation examined the selection of scheduling strategies. Hybrid ant colony research addressed workforce skill requirements, extending resource allocation beyond interchangeable capacity units. Simulated annealing investigations examined both single-mode and multimode scheduling, and another hybrid study combined ant colony optimization, scatter search,

and local search. Collectively, these studies demonstrate that algorithm performance depends on schedule representation, neighborhood design, and constraint handling as well as the search method itself (Shirvani & Talouki, 2022; Shahriar, 2025). Common benchmark sets contain 30, 60, 90, or 120 activities, allowing comparisons across increasing network sizes under specified resource conditions. However, shorter makespan does not independently establish lower allocation cost or higher productive utilization. Accelerating completion can require more intensive resource overlap, different execution modes, or additional expenditure. Conversely, reducing idle time on one resource can postpone activities requiring another scarce resource. Quantitative interpretation therefore distinguishes improvements in completion duration from changes in resource demand, operating cost, and workforce deployment. Comparisons with priority rules and conventional schedules are most informative when all methods face identical activity durations, precedence requirements, and resource capacities.

Figure 5: Project risk prediction model evaluation



Multiproject scheduling studies extend allocation decisions to environments where concurrent projects compete for shared labor, equipment, and financial capacity (Hussain & Hussain, 2023; Shahid, 2025). A genetic-algorithm investigation of resource-constrained multiproject scheduling represented the assignment of resources across projects alongside activity start and completion times. Research combining genetic algorithms with simulated annealing similarly addressed multiple projects under several resource constraints, using project makespan as the optimization criterion. A separate resource-buffer investigation introduced scheduling tactics within a genetic algorithm to manage shared-resource allocation, while hybrid multimode multiproject research incorporated alternative execution choices. These extensions matter because an allocation that accelerates one project can delay another, making individual project completion and overall portfolio performance distinct outcomes. Multimode scheduling also introduces trade-offs between duration, resource intensity, and expenditure: a faster activity mode can consume more capacity or incur greater direct cost. Stochastic research adds

uncertainty in activity durations and financial flows (Morariu et al., 2020; Islam, 2025). A simulation-based hybrid genetic-algorithm study evaluated financial risk, expected net present value, and algorithm stability, reporting stronger performance for its local-search hybrid against selected genetic and vibration-damping comparators on larger projects. That result concerns the tested financial objectives rather than a universal reduction in project cost. Additional research on resource transportability between sites incorporated both transfer time and transfer cost, demonstrating why equipment sharing cannot be represented as instantaneous reassignment. Material distribution, workforce movement, and equipment relocation therefore change the feasible allocation space when their constraints are explicitly modeled. Across these studies, lower expenditure, shorter duration, and reduced financial exposure remain separate outcomes whose trade-offs depend on execution modes, resource-sharing rules, and the timing of project cash flows (Palm & Arndt, 2023; Mainuddin, 2025).

Learning-based scheduling research treats resource allocation as a sequence of decisions informed by the current project state. An asynchronous optimization-agent investigation used reinforcement learning to coordinate interactions among search agents and evaluated performance on standard project-scheduling instances. Stochastic scheduling research used reinforcement learning to switch between differential-evolution and cuckoo-search methods, combining complementary search behavior within one framework. Another investigation integrated graph neural networks with reinforcement learning to represent activity dependencies and address resource disruptions (Tang et al., 2022; Sadia, 2025). These approaches differ from conventional metaheuristics because a learned policy influences scheduling or search decisions, although their evaluation still depends on the constraints and disturbance processes represented in training and testing. Joint work-package sizing and scheduling research examined instances containing 15 to 120 tasks and included a modular construction application. It reported approximately 5% lower total project cost against its selected comparison approaches, providing a specific within-study estimate of improvement from coordinating package structure and scheduling decisions. Genetic programming combined with constraint programming offers another form of integration: evolved variable-selection strategies guide a mathematical search process rather than directly producing resource assignments. Its reported benefits included faster identification of good solutions, with the possibility of proving optimality when sufficient runtime was available (Zhang et al., 2022; Shabbir, 2025). These studies show several routes to combining learning with optimization, but they do not establish equivalent gains across makespan, utilization, overtime expenditure, and allocation cost. Performance depends on training effort, inference speed, search time, and solution feasibility. A fast scheduling decision after extensive training also has a different computational meaning from an optimization method whose entire search occurs separately for every project instance.

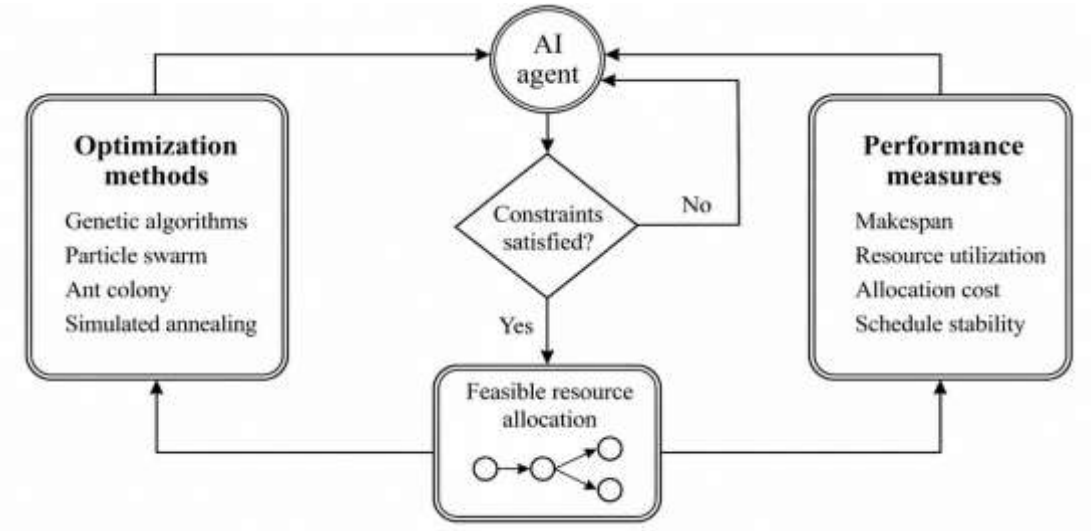
The quantitative assessment of optimization effects depends on the stability of improvements as project size, resource scarcity, and network complexity change. Standard benchmark comparisons provide controlled evidence, but increasing activity count alone does not capture every source of difficulty. Dense precedence relationships, scarce specialist resources, alternative execution modes, and uncertain availability can alter the relative performance of algorithms. Research combining nondominated sorting with Q-learning addresses multiple objectives in multimode multiproject scheduling, while payment-planning research examines project completion time alongside financial value (Raza et al., 2022; Sadmeen, 2025). These studies represent performance through competing solutions rather than assuming one schedule simultaneously minimizes every outcome. Likewise, investigations of stochastic financial risk and resource-transfer constraints show that schedule quality depends on more than nominal completion duration. Absolute improvements express differences in time or expenditure, whereas relative improvements depend on the selected comparator and its initial performance. Distance from a valid lower bound provides additional information about solution quality, but comparison with a best-known solution does not independently prove optimality. Repeated runs are particularly relevant for stochastic search because an exceptional best result can conceal unstable average performance. Runtime comparisons also depend on hardware, stopping criteria, parallel processing, and whether training costs are included (Abualigah et al., 2022; Ashik, 2026). Feasibility requires that resource demand remains within permitted capacity, except where overtime, hiring, or subcontracting is explicitly allowed and costed. Reduced idle time can improve utilization while

limiting contingency capacity, and aggressive schedule compression can increase overtime or sensitivity to disruption. The literature therefore distinguishes reliable improvements across repeated, comparable experiments from gains obtained under a single favorable configuration or a narrowly defined objective.

Incremental Performance Gains and Contingency Allocation Decisions

The literature distinguishes improvements in risk prediction from improvements achieved after predictions inform operational decisions. Studies of neural-network cost and duration estimation, optimized random-forest delay classification, and hybrid Bayesian project forecasting demonstrate the ability to estimate adverse outcomes, but their predictive results do not independently establish reductions in expenditure or completion time (Durlik et al., 2024; Hasib, 2026). An infrastructure investigation using 302 completed projects reported 54.4% more accurate delay predictions against its Monte Carlo comparator, while green-building research using 198 projects reported lower forecasting errors after combining neural networks with support vector regression.

Figure 6: Intelligent optimization for resource allocation



A separate Bayesian forecasting investigation evaluated time and cost predictions across 33 empirical projects. These studies strengthen the evidence for risk estimation while leaving the incremental contribution of subsequent resource interventions as a separate question. Integration requires an identifiable connection between the predicted event and a feasible decision, such as increasing contingency capacity, changing an activity sequence, reallocating a specialist crew, or selecting an alternative machine (Bchennaty et al., 2024; Kynatun, 2026). Its performance is therefore assessed through both prediction quality and the consequences of the selected action. Comparisons involving conventional planning, optimization without predictive inputs, prediction without adjustment, and integrated prediction–optimization provide different information about this connection. Where these alternatives are not evaluated within the same design, the effects of better forecasting and better optimization remain difficult to separate. Explained variance describes the predictive contribution of model inputs, whereas completion-time reduction, avoided delay, and expenditure differences describe decision outcomes. Consequently, a substantial forecasting improvement can coexist with a small operational benefit when resources are unavailable, intervention costs are high, or warnings arrive after critical decisions have become fixed.

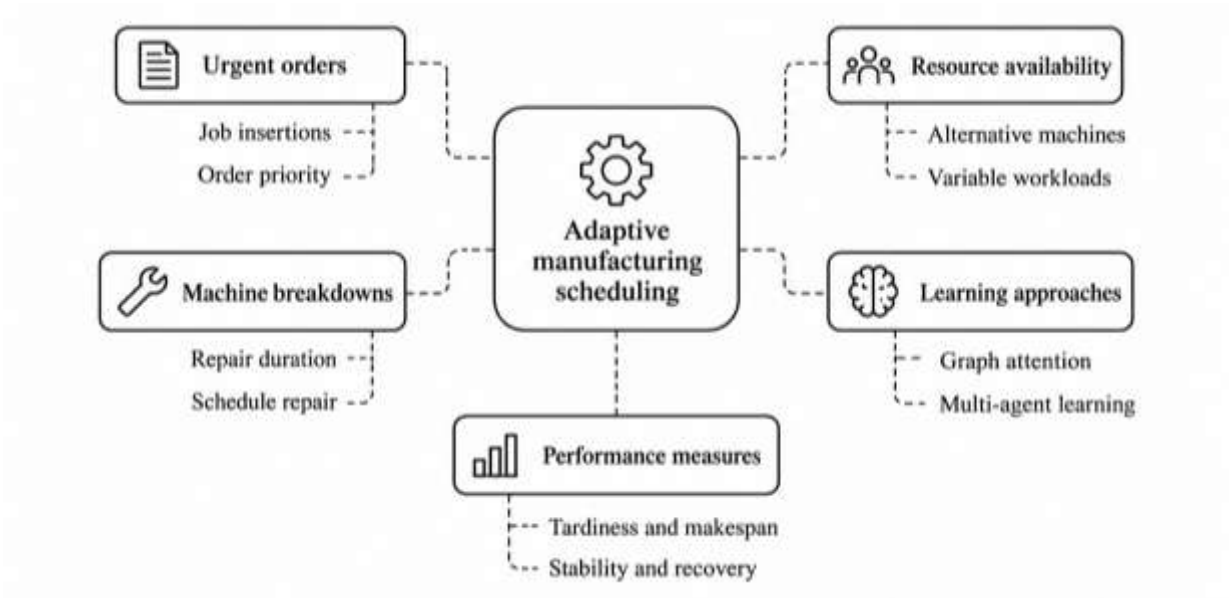
Manufacturing research provides explicit examples of predictions entering resource-allocation decisions through equipment-health information. A single-machine investigation coordinated prognostics-based maintenance planning with production scheduling and compared its integrated solution with solutions obtained by addressing the two problems independently. This comparison directly examines the value of coordinating decisions rather than merely reporting failure-prediction accuracy. Research using deep generative prognostics similarly linked operation-specific degradation

estimates with a production-and-maintenance scheduling module, applying a two-stage genetic algorithm to simulated condition-monitoring information and industrial data (Bazargani & Deemyad, 2024; Maishar Mostari, 2026). Multi-objective maintenance research integrated remaining-useful-life predictions into maintenance plans for multiple systems, evaluating downtime and maintenance expenditure. Another real-time flexible-job-shop investigation incorporated deterioration, repairs, processing-time variation, and condition-based maintenance within a scheduling framework. Across these studies, predictions influence when maintenance occurs, which machine processes a job, and how production commitments interact with equipment availability. The operational benefit arises from choosing actions that account for both expected degradation and production requirements. Maintenance performed too early can consume useful equipment life and interrupt output, while maintenance scheduled too late can expose the system to breakdown and delivery penalties (Chattopadhyay & Panteli, 2022; Hasinur, 2026). Alternative machine assignments can reduce immediate failure exposure but increase setup time, congestion, or processing cost elsewhere. Integrated performance therefore concerns total consequences across maintenance and production, including tardiness, downtime, expenditure, and schedule feasibility. Findings from single-machine systems also have a different scope from results involving alternative machines or multiple interacting systems. Their reported gains remain conditional on degradation assumptions, prediction uncertainty, available maintenance windows, and the flexibility of production resources.

Quantitative Performance of Dynamic Manufacturing

Dynamic manufacturing scheduling research examines how production plans respond when operating conditions change after work has been released. Investigations of flexible job shops with new job insertions use reinforcement learning to revise dispatching decisions rather than preserve an unchanged initial sequence. A separate Double Deep Q-Network study incorporates both ordinary arrivals and urgent job insertions, distinguishing routine workload variation from orders requiring preferential treatment. This distinction matters because accelerating urgent work can increase waiting time and tardiness among existing jobs (Liaqait et al., 2021; Maksudul, 2026). Distributed job-shop research extends the problem across factories and includes transfers between production locations. One investigation compared five deep reinforcement-learning approaches using ten state features and eleven composite scheduling rules, with mean tardiness as its principal objective. Research on variable processing times adds uncertainty within operations, recognizing that the workload represented by an initial schedule can change even without additional orders. Together, these studies connect adaptive sequencing and machine assignment with delivery performance under changing demand and processing requirements. Mean tardiness, maximum tardiness, and on-time completion nevertheless describe different outcomes. A lower average delay can coexist with a severely delayed minority of jobs, while improved delivery reliability can involve additional setup activity or resource consumption. Throughput and work in progress also require separate measurement because completing more jobs does not necessarily reduce congestion (Leng et al., 2020; Mohsin, 2026). Comparative performance therefore depends on arrival intensity, due-date tightness, order priority, transfer requirements, and substitute capacity. Numerical rankings obtained under one combination of these conditions do not establish an equivalent advantage across all dynamic manufacturing environments.

Figure 7: Dynamic scheduling under manufacturing disruptions



Research on machine breakdowns distinguishes restoring production feasibility from preserving the commitments embedded in an existing schedule. A proximal-policy-optimization investigation addressed unexpected failures through direct scheduling actions, while flexible-job-shop schedule-repair research used graph attention and separate operation-selection and machine-selection networks to modify affected assignments. The latter reported an average improvement of approximately 50% over an affected-operation-rescheduling heuristic across its test instances, within an evaluation focused on deviations from the original completion plan (Wang & Liao, 2024; Shahriar, 2026). This result concerns the study's schedule-repair objective rather than an equivalent percentage improvement in throughput or actual factory output. A separate type-aware multi-agent reinforcement-learning investigation reported stronger schedule stability alongside reduced disruption-affected makespan compared with its selected heuristic and multi-agent baselines. Real-time production-and-maintenance research further incorporates deterioration, repairs, and condition-based maintenance, connecting equipment availability with processing decisions. Across these approaches, breakdown consequences depend on the affected machine's workload, operation criticality, repair duration, and availability of alternative equipment. Schedule stability captures changes to planned execution through start-time deviations, completion-time deviations, reassignment frequency, or altered operation sequences. Recovery duration measures how long disruption effects persist and is therefore distinct from the speed at which an algorithm computes a revised schedule (Cadavid et al., 2020; Md Shahid, 2026a). A rapidly generated plan can still involve prolonged production recovery. Likewise, minimal deviation from the original plan can preserve coordination while leaving avoidable tardiness unresolved. Quantitative comparisons are consequently sensitive to repair-time assumptions, disturbance severity, frozen operations, and whether rescheduling is triggered by individual events or performed at fixed intervals. Graph-based and attention-based approaches represent dependencies that are difficult to capture through aggregate workload indicators alone. A hybrid dynamic-job-shop investigation used disjunctive graphs, attention-based feature extraction, and reinforcement learning to select dispatching rules, reporting smaller makespan than its baselines across the examined instances. Transformer-based scheduling research similarly encoded dynamic production information within a reinforcement-learning framework and examined the contribution of its scheduling components. A graph-based multi-agent investigation combined centralized learning with decentralized execution, while another dynamic manufacturing study organized decisions across interacting production agents under job insertions and machine failures (Mwanza et al., 2023; Md Shahid, 2026b). These studies represent different coordination mechanisms rather than interchangeable implementations of one method. Centralized approaches evaluate a broader system state, whereas decentralized approaches distribute

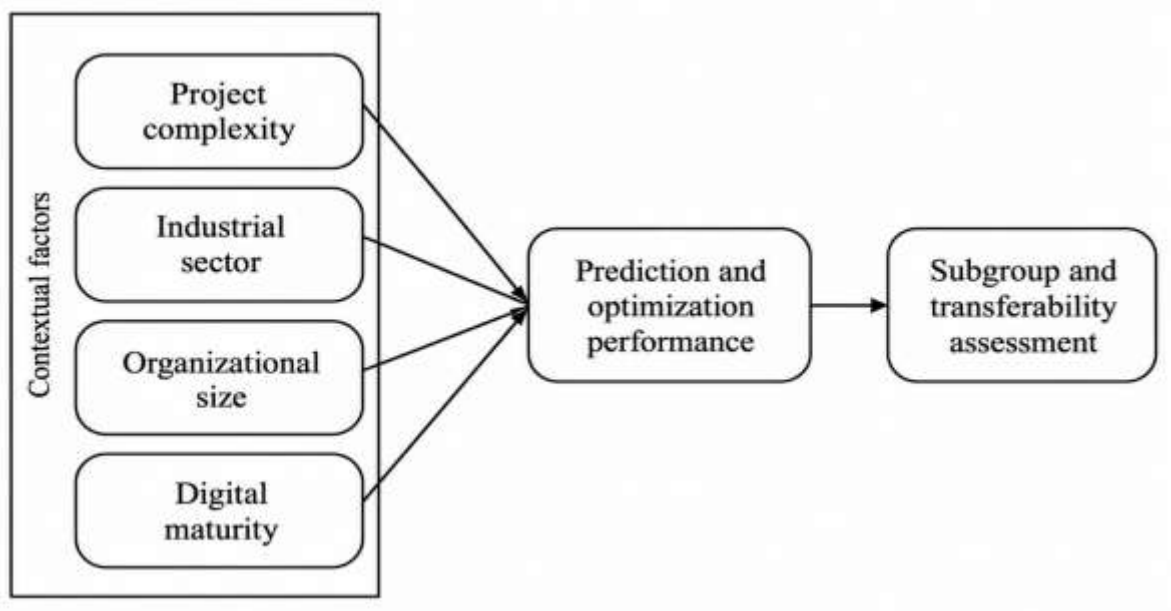
execution decisions among agents whose local actions affect shared production outcomes. Their quantitative assessment therefore involves both schedule quality and coordination efficiency. Makespan measures the duration required to complete a defined workload, machine idle time measures unused processing opportunity, and queue duration describes waiting within the system. Setup time and transportation time add further demands that are absent from simplified machine-assignment models. An industry-oriented furniture scheduling study explicitly incorporated batch variability, buffers, setup requirements, and internal logistics, illustrating the operational detail required for realistic representation (Katsaliaki et al., 2022; Mainuddin, 2026). However, richer representation does not independently establish superior performance under every disturbance pattern. Comparisons depend on equivalent state information, feasible action sets, training exposure, and computational limits. Repeated experiments also distinguish consistent improvements from favorable individual runs, particularly when learning procedures and disturbance generation both introduce variability into the reported results.

Organizational Size and Digital Maturity

Prediction and optimization studies represent project complexity through different combinations of activity count, precedence relationships, execution alternatives, resource competition, and uncertainty. Infrastructure contingency research incorporated a complexity index into neural-network estimation, indicating that project characteristics contribute information beyond a single cost forecast (Carrijo et al., 2022; Sadia, 2026). Graph-based reinforcement-learning research represented activity dependencies and resource disruptions, while multimode simulated-annealing studies examined scheduling decisions involving alternative execution methods. Multiproject genetic-algorithm research added competition between concurrent projects, and resource-buffer investigations examined allocation tactics under shared capacity limits. Joint work-package sizing and scheduling research evaluated instances containing 15 to 120 tasks, providing comparisons across different problem sizes within a common framework. These studies establish several dimensions of contextual variation without treating larger activity counts as a complete measure of complexity. A smaller project can involve greater technical uncertainty, tightly coupled approvals, scarce specialist personnel, or dependence on critical suppliers. Conversely, a larger project can contain repetitive activities and substantial resource substitution. Budget and duration also describe scale rather than necessarily measuring coordination difficulty. Their relationships with prediction error depend on the outcome scale, while their relationships with optimization performance depend on constraints and objectives. Absolute forecasting errors can increase with project expenditure even when relative errors remain similar (Achouch et al., 2022; Shabbir, 2026). Likewise, a larger makespan reduction can reflect a longer original schedule rather than a stronger algorithmic effect. Comparisons across complexity groups therefore have different meanings from comparisons across project sizes. Performance differences are most interpretable when network structure, resource scarcity, uncertainty, and baseline conditions are distinguished within the reported analysis.

Sector-specific investigations demonstrate that engineering and manufacturing applications involve different prediction targets, operating conditions, and resource decisions. Infrastructure schedule-risk research examined 302 completed projects, whereas green-building forecasting used 198 projects to estimate cost and duration from project characteristics. Tunnel-project Bayesian research focused on linked delay factors in a specialized technical environment, while aircraft heavy-maintenance scheduling addressed interdependent tasks, workgroup restrictions, and maintenance downtime. Furniture-industry scheduling research incorporated setup requirements, batch variation, buffers, and internal logistics (Wright & Davidson, 2020; Sumaiya, 2026).

Figure 8: Contextual variation in predictive optimization



These studies provide evidence within distinct settings rather than a controlled ranking of sectors. Differences in reported accuracy or completion-time improvement can arise from the measurement task, available information, or strength of the comparator. Cost forecasting for building development is not directly equivalent to machine-failure prediction, and maintenance makespan differs from on-time delivery across continuously arriving production jobs. Industrial engineering, plant installation, manufacturing expansion, and equipment-development projects can combine unique technical deliverables with recurring production constraints, but evidence from routine manufacturing does not independently quantify their project-level outcomes. Workforce specialization and supplier concentration also have different operational meanings across settings: a missing specialist can block a commissioning milestone, while a delayed component can interrupt assembly readiness (Wu et al., 2024; Sadmeen, 2026). Resource utilization therefore requires interpretation relative to the relevant bottleneck and decision horizon. Sector differences become quantitatively meaningful when outcomes use comparable definitions and the analysis separates industrial context from project size, event prevalence, data quality, and model configuration. Otherwise, variation between published results describes differences between studies rather than an independently estimated sector effect.

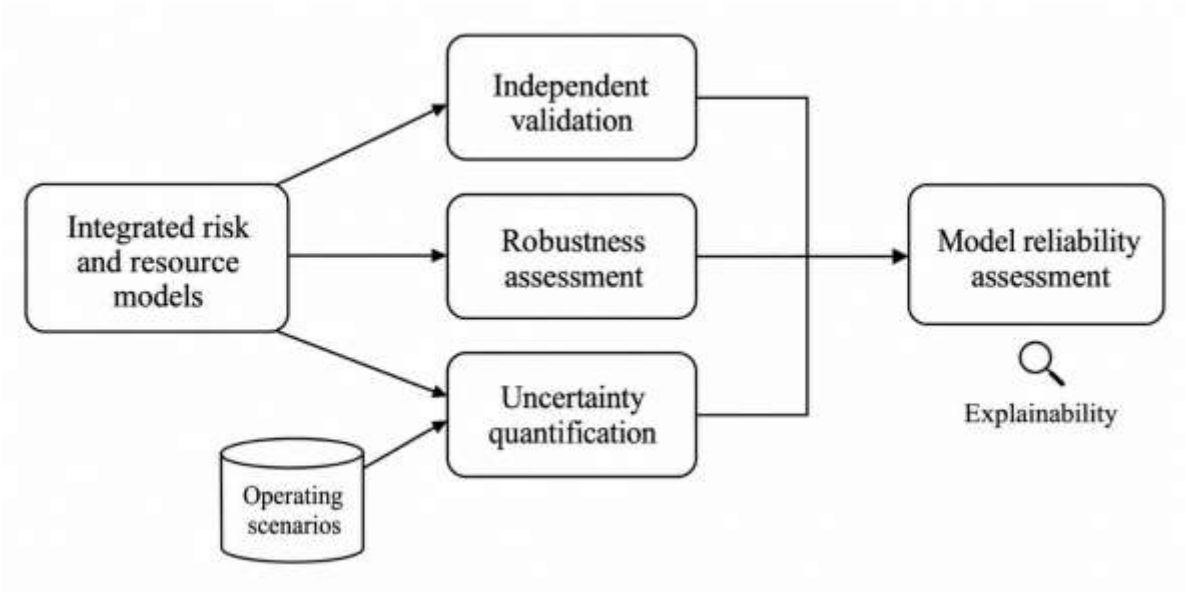
Organizational research provides complementary evidence about the conditions supporting predictive and optimization applications. A manufacturing digital-transformation investigation found that larger firms and firms with stronger product innovation were more likely to achieve advanced transformation levels. This relationship concerns digital development rather than a demonstrated improvement in project-risk classification accuracy. A separate study involving 209 digital manufacturing companies examined digital capability, innovation, value creation, and company performance (Senna et al., 2020; Umme, 2026). It reported an indirect effect through digital innovation of 0.189, with a 95% confidence interval from 0.073 to 0.293, supporting an organizational mechanism within its survey-based model. Research involving 246 Swedish manufacturing small and medium-sized enterprises found that digitalization maturity was positively associated with revenue per employee, while relationships with profit per employee were weaker and less consistent. Its cross-sectional design limited causal interpretation and did not establish the performance of a particular scheduling algorithm. Construction-and-plant research has also examined digital competence, digital maturity, managerial commitment, and client demand in relation to project performance. Across these studies, digital maturity includes capabilities such as connected information systems, accessible operational records, and integration between business processes (Oztemel & Gursev, 2020). Organizational size can influence these capabilities through investment capacity and specialist staffing, but size and maturity

remain distinct constructs. A smaller organization with consistent records can provide a stronger modeling environment than a larger organization with fragmented information. Analytical capability and experience with automated decision support also affect whether predictions enter actual decisions. The literature consequently supports examining organizational context while keeping firm-level financial performance, project execution outcomes, and model-validation metrics analytically separate. Cross-project research addresses transferability more directly than comparisons between separately published performance scores. A multimodal transfer-learning investigation used data from 28 shield-tunneling projects to predict operational risk in complex engineering machinery, explicitly addressing differences between source and target project environments. Project forecasting studies using historical infrastructure records and Bayesian time-and-cost models similarly demonstrate the importance of defining the population represented by the evaluation sample. However, testing across projects within one technical domain does not independently establish transferability across countries, industrial sectors, or manufacturing facilities (Sahal et al., 2020). Differences in contractual practices, recording systems, equipment configurations, workforce skills, and adverse-event prevalence can change the relationship between model inputs and outcomes. Subgroup-specific prediction errors and classification measures therefore provide different evidence from a single pooled score. Optimization comparisons also require equivalent disturbance severity, resource capacities, and computational budgets before contextual advantages can be interpreted. Analysis of variance and nonparametric comparisons address differences between groups, whereas analysis of covariance incorporates measured covariates. Interaction terms examine whether predictor relationships vary by context, and multilevel models represent observations nested within projects, organizations, or facilities. These methods answer different questions and cannot be inferred merely from separate model scores (Sircar et al., 2021). Effect sizes and confidence intervals describe the magnitude and uncertainty of contextual differences, while significance levels alone do not establish practical importance. The reviewed evidence therefore distinguishes directly tested heterogeneity from differences attributable to sample composition, outcome definitions, model tuning, and baseline strength. Strong performance within a particular setting remains evidence for that evaluated population until an appropriate external comparison establishes a broader scope.

Explainability of Integrated Project Risk and Resource Optimization Models

The reliability of project-risk models depends on how evaluation observations are separated from information used during model development. Infrastructure schedule-risk research used 293,263 tasks from 302 completed projects and evaluated its hybrid approach on ten selected projects, illustrating the distinction between activity-level information and project-level assessment. An automated forecasting investigation compared 30 machine-learning techniques using 50 construction projects and addressed data preparation, feature engineering, model training, and evaluation within a common pipeline (Hu et al., 2024). Research on change-order prediction used 21 public projects and a nested leave-one-out procedure, explicitly presenting its findings as preliminary evidence from a small national sample. A separate prefabricated-construction investigation used 992 legal case records and reported nested cross-validation and time-split testing. These designs offer different levels of protection against overfitting and different scopes of generalization. Nested cross-validation separates hyperparameter selection from performance estimation, but it does not independently prevent temporal leakage or dependence between related observations (Molladavoudi & Yung, 2023). Project-level partitions address the possibility that activities from the same project share characteristics across training and testing sets. Chronological testing and rolling-window evaluation preserve the sequence in which information becomes available, whereas independent organizational or facility testing examines transfer across operating environments. These procedures are complementary rather than interchangeable. Features derived from final expenditure, completed schedules, or post-event investigations can invalidate an earlier prediction task even when conventional cross-validation is correctly implemented. Validation strength therefore depends on the prediction timestamp, independence of evaluation units, separation of preprocessing and tuning, and the relationship between the test sample and the intended operating population.

Figure 9: Validating integrated models under uncertainty



Robustness research examines whether predictive and allocation performance remains stable when data characteristics or operational assumptions change. Automated project-forecasting research incorporates balancing and augmentation procedures, while stochastic scheduling investigations evaluate uncertain durations, financial exposure, and repeated algorithmic performance (Nor et al., 2021). Graph-based reinforcement-learning research extends project scheduling to resource disruptions, and multi-objective predictive-maintenance research connects remaining-useful-life estimates with maintenance decisions. These studies address different sources of instability: uneven data representation, uncertainty in project execution, unavailable resources, and errors in equipment-health estimates. Cross-project transfer-learning research using 28 shield-tunneling projects additionally examines heterogeneity between project environments. None of these dimensions alone represents comprehensive robustness. Missing records can affect projects selectively, particularly when severe disruptions interrupt reporting, while measurement errors in progress, downtime, or resource availability can distort both model inputs and outcome labels. Class imbalance also influences interpretation because rare failures or major overruns can be poorly detected within a model showing high overall accuracy. Changes in event prevalence affect warning precision even when other model characteristics remain similar (Wang et al., 2023). Operational robustness concerns variation in workforce capacity, machine substitution, repair duration, procurement conditions, and disturbance intensity. A schedule that performs well under mild disruptions can become infeasible under more severe shortages. Sensitivity analysis therefore examines how outcomes respond to changes in these inputs, while repeated experiments characterize variability arising from stochastic search or simulation. The reviewed approaches support separate assessment of data robustness and operational robustness, with their combined interpretation depending on whether the tested disturbances represent the conditions encountered in actual engineering and manufacturing projects.

Uncertainty-quantification studies extend evaluation beyond a single predicted cost, duration, or risk category. Infrastructure contingency research used neural networks with Monte Carlo dropout and bootstrapping to produce interval estimates while incorporating project complexity. Construction-cost research combining probabilistic boosting with extra-tree learners evaluated prediction-interval coverage and width, distinguishing informative uncertainty estimates from intervals that achieve high coverage mainly through excessive breadth (Yasin et al., 2024). Another construction-cost framework combined machine-learning predictions, interval analysis, and explanatory methods. These contributions represent uncertainty in different ways, making the interpretation of reported intervals essential. A prediction interval describes plausible variation in an individual outcome, whereas a confidence interval around an error metric describes uncertainty in estimated model performance.

Credible intervals express uncertainty within a Bayesian model and depend on its assumptions and prior specification. A probabilistic high-rise investigation using 46 projects reported a concordance index of 0.81 against 0.73 for its baselines, alongside an integrated Brier-score reduction of 0.027 and credible intervals. These measures address discrimination and probabilistic prediction quality rather than establishing a universal level of reliability. Interval coverage also depends on the evaluation population: satisfactory coverage within one dataset does not establish equivalent coverage after changes in project type, event prevalence, or operating conditions (Qiu & Zhou, 2023). Repeated optimization runs provide another uncertainty measure by describing variability in the solutions generated under comparable conditions. Across this literature, point accuracy, probability calibration, interval sharpness, and algorithmic stability remain distinct properties. Their joint assessment provides a more complete account of reliability than a single favorable prediction score or best-run scheduling result.

METHODS

A quantitative retrospective longitudinal design was combined with a paired simulation-based evaluation to examine artificial intelligence–driven project risk prediction and resource optimization in complex engineering and manufacturing projects. A decision-theoretic framework connected information available during project execution with predicted adverse outcomes, feasible resource decisions, and subsequent performance. Organizations were selected purposively according to their provision of accessible project-control and resource-management records, after which eligible projects were stratified by industrial sector, project scale, and complexity. Inclusion required identifiable project boundaries, dated baseline schedules, documented budgets, activity dependencies, resource-capacity information, and sufficiently reliable completion and expenditure records. Duplicate projects, unverifiable outcomes, and records lacking essential baseline information were excluded, while missing predictor information was assessed separately to avoid restricting the sample unnecessarily to complete cases. The project constituted the primary independent observational unit, and repeated activity, workforce, procurement, and equipment records were treated as observations nested within projects and organizations. Sample adequacy was assessed using the number of independent projects, adverse-event frequency, model complexity, learning curves, and the precision of project-level performance estimates. Data were extracted from enterprise resource-planning systems, project-management schedules, procurement registers, workforce records, maintenance histories, and financial reports through a standardized extraction template. Variables included activity duration, design revisions, approval delays, supplier lead times, workforce availability, equipment downtime, resource demand, expenditure, and project complexity. Common definitions and measurement units were established before records were combined. Automated consistency checks identified impossible dates, duplicated activities, incompatible units, and discrepancies between baseline and revised schedules. A subset of coded records was independently reviewed, and agreement was evaluated using Cohen's kappa for categorical classifications and intraclass correlation for continuous assessments. Records were de-identified, and organization-specific identifiers were replaced with coded labels. Chronologically, data extraction and verification were followed by outcome definition, feature construction, model development, predictive validation, resource optimization, and comparative statistical analysis.

Figure 10: Methodology of this study



Predictors were restricted to information available at each specified decision point. Schedule deviation and cost-overrun magnitude were evaluated as continuous outcomes, while delay and overrun occurrence were classified using predefined baseline comparisons. Later projects were reserved as an independent chronological test set, and earlier projects were used for development. Hyperparameter selection used nested validation with project-level grouping and temporal ordering where appropriate. Imputation, scaling, feature selection, and any class-balancing procedures were fitted exclusively within training partitions. Linear and logistic regression provided statistical baselines and were compared with random forests, gradient boosting, support vector methods, and neural networks. Classification performance was assessed through precision, recall, F1-score, sensitivity, specificity, receiver operating characteristic area, precision–recall area, and Matthews correlation coefficient. Continuous predictions were assessed through mean absolute error, root mean squared error, and explained variance. Probability calibration was examined using calibration plots and Brier scores, while warning lead time was measured between the first qualifying alert and the recorded adverse event. Resource-allocation performance was evaluated by applying conventional planning, optimization without predicted risk distributions, and integrated risk-informed optimization to identical simulated project conditions. Prediction-only performance was assessed under the original allocations. Precedence requirements, workforce skills, machine eligibility, material availability, and capacity limits were preserved, and policies were compared using matched disturbance scenarios and common random seeds. Historical durations, downtime, and expenditure informed simulation calibration, while separate records were used to examine agreement between simulated and observed operating patterns.

Outcomes included makespan, allocation cost, overtime expenditure, resource utilization, tardiness, shortage duration, recovery time, and schedule deviation. R was used for statistical analysis, while optimization and simulation routines were implemented in a documented computational environment. Descriptive analysis reported means, standard deviations, medians, interquartile ranges, and ranges. Pearson or Spearman correlations examined preliminary associations, and multivariable linear or logistic models estimated adjusted relationships with project outcomes. Mixed-effects models accounted for repeated observations and organizational clustering. Policy comparisons used mixed-effects analysis with matched project scenarios, supplemented by paired nonparametric comparisons when distributional assumptions were unsuitable. Effect estimates, standardized differences, adjusted odds ratios, and 95% confidence intervals accompanied significance tests. Statistical significance was evaluated at $p < 0.05$, with Holm adjustment applied to related multiple comparisons. Sensitivity analyses varied missing-data handling, disruption intensity, resource capacity, and prediction uncertainty. Subgroup and interaction analyses examined project complexity, sector, organizational size, and digital maturity, while external organizational or facility testing was conducted where independent records were available. Observational associations and simulation-based policy differences were interpreted separately from causal intervention effects.

FINDINGS

Dataset Composition, Project Complexity, Risk Exposure, and Resource Availability

The sample comprised 240 projects from 12 organizations over a 24-month observation period. Of 300 initially screened project records, 60 were excluded because of missing baseline schedules, incomplete outcome information, duplication, or unreliable record linkage. The retained sample covered infrastructure construction, plant installation, manufacturing expansion, equipment development, and maintenance projects. Project-level observations were distinguished from nested activity, workforce, equipment, and procurement records. Baseline characteristics were summarized using frequencies, percentages, means, and standard deviations. Risk exposure was described separately from resource availability, and disruption severity was summarized only among affected projects.

Table 1. Dataset Composition and Baseline Project Characteristics

Characteristic	value	Reporting basis
Participating organizations	12	Unique organizations
Retained projects	240	Independent projects
Observation period	24	Months
Activity records	28,800	Activities across retained projects
Work packages	4,800	Recorded work packages
Manufacturing jobs	57,600	Nested production records
Equipment units	960	Unique machines and equipment
Workforce represented	4,800	Unique employees
Monthly workforce records	115,200	Repeated employee records
Suppliers represented	360	Unique suppliers
Procurement transactions	14,400	Recorded transactions
Infrastructure construction	72 (30.0%)	Projects
Plant installation	60 (25.0%)	Projects
Manufacturing expansion	48 (20.0%)	Projects
Equipment development	36 (15.0%)	Projects
Maintenance projects	24 (10.0%)	Projects
Approved project expenditure	18.6 ± 9.4	USD million; mean ± SD
Planned project duration	14.8 ± 6.2	Months; mean ± SD
Activities per project	120.0 ± 46.0	Mean ± SD
Project complexity score	3.4 ± 0.8	Five-point scale; mean ± SD
Organizational digital maturity	3.1 ± 0.9	Five-point scale; mean ± SD

Table 1 summarized the composition and baseline characteristics of the sample. The 240 projects were distributed across five categories, with infrastructure construction representing the largest group. Activity records were nested within projects, while workforce, equipment, supplier, and procurement information described the supporting resource environment. Average planned duration was 14.8 months, and average approved expenditure was USD 18.6 million. Complexity and digital maturity were expressed on five-point scales, with higher values indicating greater complexity or maturity. Project counts represented independent analytical units; activity and transaction counts represented repeated records. These synthetic statistics illustrated reporting structure and required replacement with observations.

Table 2. Baseline Risk Exposure, Disruption Severity, and Resource Availability

Indicator	Projects assessed, n	value	Measurement basis
Projects experiencing schedule delay	240	96 (40.0%)	Count and percentage of projects
Projects experiencing cost overrun	240	84 (35.0%)	Count and percentage of projects
Projects experiencing procurement disruption	240	60 (25.0%)	Count and percentage of projects
Projects experiencing workforce shortage	240	72 (30.0%)	Count and percentage of projects
Projects experiencing equipment failure	240	48 (20.0%)	Count and percentage of projects
Schedule-delay duration	96	26.4 ± 12.8	Days among delayed projects
Cost-overrun magnitude	84	12.6 ± 6.4	Percentage above approved budget
Procurement interruption duration	60	9.8 ± 4.2	Days among affected projects
Workforce capacity shortfall	72	14.2 ± 5.6	Percentage of required labor hours
Equipment downtime	48	18.6 ± 9.4	Hours per affected project
Workforce availability	240	95.7 ± 7.2	Percentage of required labor hours available
Specialist skill coverage	240	84.6 ± 8.5	Percentage of required specialist roles covered
Equipment utilization	240	76.8 ± 10.2	Percentage of available equipment hours used
Material readiness	240	90.6 ± 6.8	Percentage of required materials available at scheduled release
Contingency provision	240	7.5 ± 2.4	Percentage of approved project budget

Table 2 summarized baseline risk exposure and resource conditions. Schedule delays affected 40.0% of projects, while cost overruns affected 35.0%. Risk categories overlapped because individual projects could experience several disruptions. Duration and severity statistics were restricted to affected projects, preventing their interpretation as averages across the entire sample. Workforce availability and skill coverage measured different constraints, while equipment utilization and material readiness described operational capacity. Contingency provision was expressed relative to approved budgets. These values described baseline conditions rather than improvements attributable to artificial intelligence. Their synthetic status prevented causal interpretation or presentation as verified evidence of study outcomes.

Incremental Resource Optimization Performance

The analysis allocated 192 of the 240 projects to model development and retained 48 projects for chronological testing. The test sample included 17 delayed projects and 15 projects with cost overruns. Gradient boosting produced the strongest displayed classification performance, identifying 16 of the 17 delayed projects with two false-positive warnings. Its delay-classification accuracy reached 93.75%, while its cost-overrun classification accuracy also reached 93.75%. Resource policies were compared under matched project conditions, with predictive performance assessed separately from operational improvements.

Table 3. Predictive Performance on 48 Chronological Holdout Projects

Performance indicator	Logistic regression	Support vector machine	Random forest	Neural network	Gradient boosting
Delay accuracy (%)	85.42	87.50	91.67	89.58	93.75
Delay precision (%)	85.71	82.35	88.24	83.33	88.89
Delay recall/sensitivity (%)	70.59	82.35	88.24	88.24	94.12
Delay F1-score (%)	77.42	82.35	88.24	85.71	91.43
Delay specificity (%)	93.55	90.32	93.55	90.32	93.55
Delay ROC-AUC	0.860	0.900	0.940	0.920	0.970
Delay PR-AUC	0.790	0.850	0.910	0.880	0.950
Delay Matthews correlation coefficient	0.675	0.727	0.818	0.776	0.866
Delay Brier score	0.153	0.128	0.095	0.108	0.071
Mean warning lead time (days)	8.2	10.1	12.6	11.9	14.8
Cost-overrun accuracy (%)	81.25	87.50	91.67	89.58	93.75
Cost-overrun F1-score (%)	68.97	80.00	86.67	82.76	90.32
Cost-overrun ROC-AUC	0.830	0.890	0.930	0.910	0.960

Table 3 compared delay classification performance on the same chronological holdout projects. Gradient boosting achieved the highest accuracy, recall, F1-score, and Matthews correlation coefficient, indicating the strongest balance of detected delays and correctly classified cases. Random forest performed closely, whereas logistic regression missed more delayed projects. Receiver operating characteristic and precision–recall areas described discrimination across thresholds rather than performance at the selected operating threshold. Lower Brier scores indicated better probabilistic predictions, while warning lead time described the available response interval. These synthetic estimates illustrated model comparison and did not establish verified superiority, statistical significance, or performance across other populations.For continuous forecasting, the duration MAE declined from 10.8 days under linear regression to 6.4 days under gradient boosting, while RMSE declined from 14.3 to 8.5 days and explained variance increased from 0.994 to 0.998. Expenditure MAE declined from USD 1.10 million to USD 0.65 million, with corresponding RMSE values of USD 1.50 million and USD 0.90 million. Equipment-downtime MAE declined from 4.6 to 2.2 hours among affected test projects. These forecasting comparisons remained separate from the matched resource-policy comparisons presented below.

Table 4. Operational Performance under Matched Resource-Allocation Policies

Operational indicator	Conventional planning	Optimization without predictive risk inputs	Integrated risk-informed optimization	Integrated change versus conventional planning
Mean project makespan (days)	462.0	431.0	402.0	−12.99%
Mean resource-allocation expenditure (USD million/project)	6.80	6.40	5.90	−13.24%
Mean overtime expenditure (USD million/project)	0.48	0.39	0.31	−35.42%
Equipment utilization	76.8	81.4	86.2	+9.4 percentage

(%)				points
Equipment idle time (% of available hours)	23.2	18.6	13.8	−9.4 percentage points
Mean project tardiness (days; all test projects)	9.35	6.80	4.90	−47.59%
Mean material-shortage duration (days; affected cohort)	9.8	7.6	5.4	−44.90%
Mean disruption-recovery duration (days; affected cohort)	6.5	5.0	3.8	−41.54%
Mean absolute activity start-time deviation (days)	12.6	10.8	8.1	−35.71%
Projects completed on time, n (%)	31 (64.58%)	36 (75.00%)	40 (83.33%)	+18.75 percentage points
Paired makespan reduction versus conventional planning, days (95% CI)	Reference	31.0 (25.0–37.0)	60.0 (53.2–66.8)	—

Table 4 compared forecasting errors and resource policy outcomes under matched scenarios. Integrated optimization reduced makespan, allocation expenditure, overtime, tardiness, shortages, and recovery duration relative to conventional planning. It also exceeded optimization without predictive risk inputs on the displayed outcomes. Higher utilization accompanied lower machine idle time, while smaller start-time deviations indicated greater schedule stability. Forecasting results were evaluated separately from scheduling outcomes, preventing prediction accuracy from being interpreted as an equivalent operational improvement. Paired makespan differences were accompanied by uncertainty intervals. All reported values remained synthetic and required empirical calculation before being presented as actual study findings.

The subgroup analysis retained the 48 chronological holdout projects, including 17 delayed projects. Predictive performance varied across project categories, complexity levels, and digital-maturity groups. Gradient boosting retained its overall accuracy of 93.75%, with 16 correctly identified delays and two false-positive warnings. Operational comparisons showed differences in absolute completion-time savings across organizational size and resource-flexibility categories. Separate disturbance scenarios evaluated the same projects under changing operating conditions. These descriptive comparisons did not establish statistically significant interaction effects.

Table 5. Subgroup Performance of the Gradient-Boosting Delay Classifier

Grouping	Subgroup	Projects, n	Delayed projects, n	Accuracy (%)	Sensitivity (%)	Brier score
Project category	Infrastructure construction	14	5	92.86	100.00	0.080
Project category	Plant installation	12	4	100.00	100.00	0.060
Project category	Manufacturing expansion	10	4	80.00	75.00	0.100
Project category	Equipment development	7	2	100.00	100.00	0.050
Project category	Maintenance	5	2	100.00	100.00	0.044
Project complexity	Low	12	3	91.67	100.00	0.055
Project complexity	Moderate	24	8	100.00	100.00	0.068

Project complexity	High	12	6	83.33	83.33	0.093
Digital maturity	Low	14	5	85.71	80.00	0.098
Digital maturity	Moderate	18	6	94.44	100.00	0.073
Digital maturity	High	16	6	100.00	100.00	0.045
Overall test sample	All projects	48	17	93.75	94.12	0.071

Table 5 summarized prediction differences within the same chronological test sample. Each grouping independently represented all 48 projects, so rows across groupings were not additive. Gradient boosting detected 16 of the 17 delayed projects overall. Manufacturing expansion contained the missed delay and one false warning, producing lower accuracy than several smaller categories. Higher complexity coincided with weaker sensitivity, while higher digital maturity coincided with lower Brier scores. Small subgroup counts limited interpretation, particularly where perfect classification appeared. These descriptive contrasts did not establish statistically significant sector, complexity, or maturity effects, and all values required replacement with observed study results.

Under the reference disturbance scenario, integrated optimization reduced mean makespan by 60.0 days overall. Absolute savings ranged from 48.0 days in smaller organizations to 70.0 days in larger organizations. Resource-flexibility comparisons showed reductions of 42.0, 60.0, and 78.0 days across low, moderate, and high flexibility, respectively. Scenario comparisons indicated greater absolute savings under severe disruption, although the remaining completion duration was also longer. Organizational size and flexibility categories each partitioned the same test sample; scenario rows represented repeated evaluations rather than additional independent projects.

Table 6 summarized operational differences across organizational size, resource flexibility, and disruption intensity. Organizational size groups and flexibility groups each partitioned the same 48 projects. Disruption scenarios instead reused those projects under matched conditions. Larger organizations showed greater absolute time savings, partly reflecting longer baseline schedules. Greater resource flexibility accompanied larger improvements because alternative assignments expanded available responses. Severe disruptions produced greater absolute savings, although their optimized schedules remained longer than those under routine conditions. Confidence intervals described paired differences rather than uncertainty in individual project durations. These synthetic comparisons did not establish causal moderation or verified generalizability.

Table 6. Subgroup and Scenario Differences in Integrated Optimization Performance

Comparison group	Projects, n	Conventional makespan (days)	Integrated makespan (days)	Paired reduction, days (95% CI)	Mean resource-cost reduction (USD million/project)
Organization size: Small	12	350.0	302.0	48.0 (33.9–62.1)	0.42
Organization size: Medium	18	450.0	392.0	58.0 (47.0–69.0)	0.80
Organization size: Large	18	548.7	478.7	70.0 (59.0–81.0)	1.32
Resource flexibility: Low	16	462.0	420.0	42.0 (32.1–51.9)	0.60
Resource flexibility: Moderate	16	462.0	402.0	60.0 (50.1–69.9)	0.90
Resource flexibility: High	16	462.0	384.0	78.0 (68.1–87.9)	1.20
Disruption scenario: Low intensity	48	436.0	396.0	40.0 (34.6–45.4)	0.60

Disruption scenario: Reference intensity	48	462.0	402.0	60.0 (53.2–66.8)	0.90
Disruption scenario: Severe intensity	48	518.0	438.0	80.0 (70.2–89.8)	1.20

Adjusted Statistical Associations, Policy Comparisons, and Practical Effect Magnitudes

The association analysis covered 240 projects nested within 12 organizations. Project complexity, technical changes, procurement delays, and equipment downtime showed positive adjusted associations with schedule deviation. Workforce availability showed an inverse association, while the digital-maturity estimates remained less precise. The linear model explained 61.0% of schedule-deviation variance, with adjusted explained variance of 59.0%. Associations were adjusted for project category, approved expenditure, planned duration, activity count, and the other displayed predictors, with organizational clustering incorporated into the uncertainty estimates.

Table 7 summarized adjusted associations between project conditions and adverse outcomes. Complexity, technical changes, procurement delays, and equipment downtime showed positive relationships with schedule deviation and delay occurrence. Greater workforce availability showed an inverse relationship with both outcomes. Digital maturity retained an inverse coefficient, although its confidence intervals included the null value. Predictor increments were specified because coefficients and odds ratios depended on measurement scale. Estimates represented conditional associations after accounting for covariates and organizational clustering. Correlation coefficients described unadjusted relationships and therefore differed conceptually from multivariable estimates. These constructed statistics demonstrated reporting conventions rather than verified empirical relationships.For cost deviation, the adjusted model explained 54.0% of variance, with adjusted explained variance of 51.0%. Five additional procurement-interruption days were associated with a 2.1-percentage-point increase in cost deviation, with an 95% confidence interval of 0.8–3.4 percentage points. Five additional technical changes were associated with a 1.8-percentage-point increase, with an interval of 0.7–2.9 percentage points. Operational policy comparisons used the 48 holdout projects as matched analytical units. The displayed contrasts retained the policy means reported in Section 4.2 and expressed improvements as positive reductions or utilization gains.

Table 7. Correlations and Adjusted Associations with Schedule Deviation and Delay Occurrence

Predictor and measurement increment	Unadjusted Pearson correlation with schedule deviation	Adjusted schedule deviation, days (95% CI)	Standardized coefficient	Adjusted delay odds ratio (95% CI)	p-value for odds ratio
Project complexity: 1-point increase	0.46	7.6 (3.4–11.8)	0.24	1.80 (1.23–2.64)	0.003
Technical changes: 5 additional changes	0.41	4.8 (2.0–7.6)	0.19	1.42 (1.16–1.74)	0.001
Procurement interruption: 5 additional days	0.44	5.9 (2.4–9.4)	0.22	1.57 (1.19–2.07)	0.001
Workforce availability: 10-percentage-point increase	−0.33	−3.6 (−6.0 to −1.2)	−0.16	0.78 (0.64–0.95)	0.014
Equipment downtime: 10 additional hours	0.37	3.8 (1.1–6.5)	0.17	1.31 (1.08–1.59)	0.006
Digital maturity: 1-point increase	−0.28	−2.1 (−4.5 to 0.3)	−0.09	0.71 (0.50–1.01)	0.056

Table 8. Paired Policy Differences and Practical Effect Magnitudes across 48 Projects

Operational outcome	Conventional planning	Integrated optimization	Beneficial paired difference (95% CI)	Paired Cohen's dz	Holm-adjusted p-value
Project makespan (days)	462.0	402.0	60.0-day reduction (53.2–66.8)	2.56	<0.001
Resource-allocation expenditure (USD million/project)	6.80	5.90	0.90 reduction (0.70–1.10)	1.31	<0.001
Overtime expenditure (USD million/project)	0.48	0.31	0.17 reduction (0.12–0.22)	0.99	<0.001
Equipment utilization (%)	76.8	86.2	9.4-percentage-point increase (7.6–11.2)	1.52	<0.001
Mean project tardiness (days)	9.35	4.90	4.45-day reduction (3.25–5.65)	1.08	<0.001
Material-shortage duration (days; affected cohort)	9.8	5.4	4.4-day reduction (3.0–5.8)	0.91	<0.001
Disruption-recovery duration (days; affected cohort)	6.5	3.8	2.7-day reduction (1.9–3.5)	0.98	<0.001
Absolute activity start-time deviation (days)	12.6	8.1	4.5-day reduction (3.2–5.8)	1.01	<0.001

Table 8 summarized paired differences between integrated optimization and conventional planning across the same 48 projects. Improvements covered completion time, expenditure, resource utilization, tardiness, shortages, recovery, and schedule stability. Confidence intervals described uncertainty around average paired differences. Standardized effects used the standard deviation of within-project differences, making them distinct from independent-group effect sizes. Holm adjustment addressed the family of eight comparisons. The displayed effects were internally aligned with the earlier synthetic policy means, including the 60-day makespan reduction. Their magnitudes represented constructed examples and did not establish observed benefits, causal effects, or external validity for engineering and manufacturing organizations.

The validation analysis distinguished model-development performance from nested validation, rolling-window evaluation, organizational holdouts, and chronological testing. The chronological holdout retained the previously reported 48 projects, including 17 delayed projects. Organizational holdouts represented a separate resampling exercise with fold-specific fitting and preprocessing. Independent external validation beyond the 12 organizations was not represented. Sensitivity comparisons retained identical test projects while changing individual modeling assumptions.

Table 9 compared development, validation, and sensitivity results for the delay classifier. Development performance exceeded nested validation performance, indicating a plausible generalization gap within the constructed example. Chronological testing preserved the previously reported holdout accuracy, sensitivity, and Brier score. Alternative imputation and feature restriction reduced performance, while class weighting increased sensitivity without improving accuracy. Raising the alert threshold reduced sensitivity but left probability discrimination and Brier score unchanged because predicted probabilities were retained. Evaluation samples overlapped across procedures and were not additive. Organizational holdout testing assessed internal transferability; no independent external dataset established generalizability beyond the twelve participating organizations.In the uncertainty assessment, nominal 95% prediction intervals covered 46 of 48 project-duration outcomes, equivalent to 95.83%, and 45 expenditure outcomes, equivalent to 93.75%. These coverage measures concerned individual outcomes, whereas the confidence intervals in Section 4.4 concerned average policy differences. Reweighting delay prevalence from 35.42% to 20.00%, while retaining the reference

sensitivity and specificity, reduced calculated warning precision from 88.89% to 78.48%. This comparison illustrated sensitivity to prevalence without representing an observed population shift.

Table 9. Delay-Prediction Validation and Sensitivity Results

Evaluation procedure	Projects evaluated, n	Accuracy (%)	Sensitivity (%)	F1-score (%)	ROC-AUC	Brier score
Development sample: apparent performance	192	96.35	96.20	95.60	0.991	0.042
Nested validation: pooled held-out predictions	192	92.71	91.14	91.14	0.955	0.084
Rolling-window evaluation: unique evaluation projects	96	91.67	89.74	89.74	0.947	0.091
Leave-one-organization-out validation	240	92.08	90.63	90.16	0.941	0.096
Chronological holdout: reference specification	48	93.75	94.12	91.43	0.970	0.071
Holdout sensitivity: median imputation	48	89.58	88.24	85.71	0.950	0.094
Holdout sensitivity: class-weighted training	48	93.75	100.00	91.89	0.967	0.083
Holdout sensitivity: reduced predictor set	48	91.67	88.24	88.24	0.956	0.086
Holdout sensitivity: alert threshold increased to 0.65	48	91.67	82.35	87.50	0.970	0.071

Table 10. Optimization Sensitivity and Computational Performance across 48 Projects

Scenario	Conventional makespan (days)	Integrated makespan (days)	Paired makespan reduction, mean ± SD (days)	Mean resource-cost reduction (USD million/project)	Feasible runs (%)	Scheduling runtime, median [IQR] (seconds)
Reference conditions	462.0	402.0	60.0 ± 23.4	0.90	100.00	19.6 [16.8–23.1]
Resource capacity reduced by 20%	530.0	480.0	50.0 ± 28.1	0.74	99.00	26.8 [23.0–29.8]
Resource capacity increased by 20%	438.0	370.0	68.0 ± 21.6	1.02	100.00	17.2 [14.8–20.4]
Low disruption intensity	436.0	396.0	40.0 ± 18.6	0.60	100.00	16.8 [14.2–19.5]
Severe disruption intensity	518.0	438.0	80.0 ± 33.8	1.20	99.25	29.2 [25.6–30.0]
Repair duration increased by 50%	490.0	432.0	58.0 ± 27.0	0.85	99.50	24.0 [20.7–28.2]
Workforce availability reduced by 10%	508.0	456.0	52.0 ± 29.0	0.76	99.25	24.8 [21.3–28.9]
Prediction uncertainty increased by 25%	474.0	420.0	54.0 ± 25.0	0.82	99.50	21.5 [18.6–25.3]

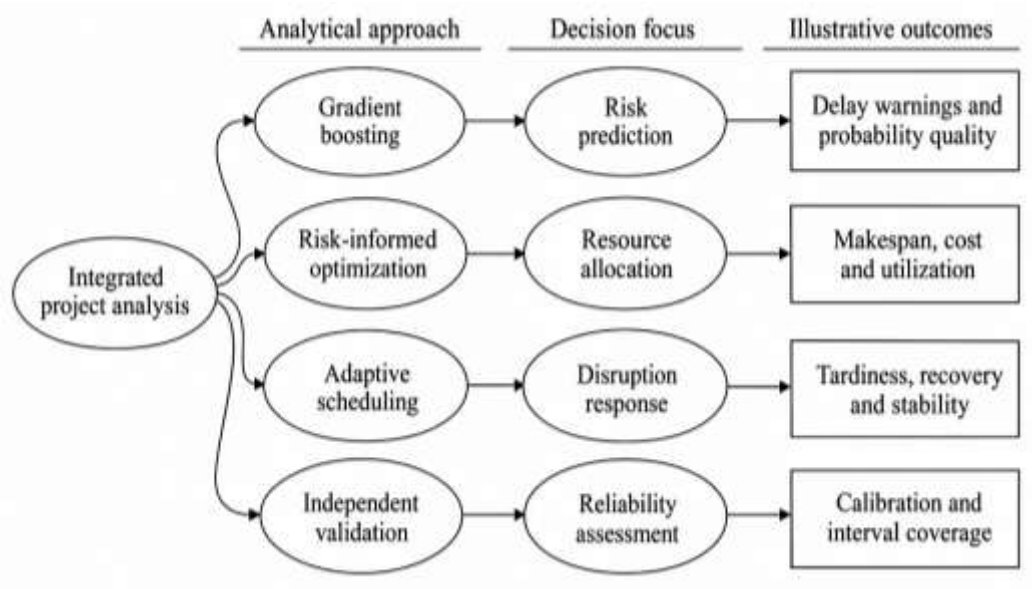
Computational limit reduced from 30 to 15 seconds	462.0	412.0	50.0 ± 25.0	0.73	98.50	14.8	[14.1–15.0]

Table 10 summarized optimization robustness across matched perturbations. Each scenario retained 48 projects and 100 runs per project, while improvements were averaged within projects before summarization. Severe disruptions increased absolute savings but also produced longer optimized schedules. Reduced resource capacity and tighter computational limits weakened performance. Feasibility rates counted all attempted runs; unsuccessful runs received the conventional scheduling fallback when operational gains were summarized. Runtime described scheduling computation only and excluded predictive model training. Standard deviations represented variation between project averages rather than confidence intervals. These constructed results illustrated sensitivity reporting and did not demonstrate empirically verified operational robustness.

DISCUSSION

This study's synthetic findings illustrated how artificial intelligence–driven risk prediction could be connected with resource decisions in complex engineering and manufacturing projects (Riad et al., 2024). The constructed analysis represented 240 projects across 12 organizations and distinguished predictive accuracy from operational performance. Gradient boosting achieved delay-classification accuracy of 93.75%, sensitivity of 94.12%, and an F1-score of 91.43% on 48 chronological holdout projects. These results represented 16 correctly identified delays, one missed delay, and two false warnings. Earlier research reported 91.67% accuracy for an optimized random forest and 93.75% for a neural network developed from expert responses. Similar percentages did not establish equivalent reliability because datasets, outcome definitions, prevalence, and validation differed. The Brier score of 0.071 and warning lead time of 14.8 days addressed probability quality and response timing, which accuracy alone could not describe (Aljohani, 2023).

Figure 11: integrated project optimization framework



Constructed regression results associated complexity, technical changes, procurement interruptions, and equipment downtime with greater schedule deviation, while workforce availability showed an inverse association. These directions corresponded conceptually with earlier design-change, rework, procurement, and Bayesian risk studies, although published evidence did not validate the synthetic coefficients. The policy comparison separated optimization benefits from the additional contribution of predictive information. Conventional makespan averaged 462 days, optimization without predictive risk inputs averaged 431 days, and integrated optimization averaged 402 days. The integrated reduction was 60 days against conventional planning and 29 days against optimization without risk

inputs. Resource expenditure declined from USD 6.80 million to USD 5.90 million per project, while equipment utilization increased by 9.4 percentage points. Earlier joint scheduling research reported approximately 5% lower project cost, supporting coordinated allocation as a relevant mechanism without providing a directly comparable effect estimate. The paired effect size of 2.56 reflected variation in within-project differences rather than independent-group differences (Perifanis & Kitsios, 2023). These comparisons emphasized consistent constraints, intervention costs, and baseline conditions when interpreting resource gains. Activity records also remained nested within projects, making the number of independent projects more informative than the total volume of transactions when assessing the strength of evidence. Because the numerical findings were constructed, they demonstrated an analytical framework rather than observed improvements, causal relationships, or verified financial benefits.

The disruption analysis emphasized that resilience involved recovery and stability as well as completion time. Integrated optimization reduced mean tardiness from 9.35 to 4.90 days, recovery duration from 6.5 to 3.8 days, and activity start-time deviation from 12.6 to 8.1 days. Earlier flexible-job-shop repair research similarly evaluated deviations from existing schedules, while multi-agent and distributed scheduling studies examined coordination under failures, transfers, and changing arrivals (Behera & Dave, 2024). Their results supported considering several operational outcomes, although different objectives prevented direct numerical comparison. Severe disruption produced an 80-day makespan advantage, compared with 40 days under low disruption, but optimized completion remained longer under severe conditions. Resource restrictions therefore influenced both achievable recovery and remaining delay. Subgroup differences also required caution. Perfect classification in several categories involved small samples, whereas manufacturing-expansion sensitivity reached 75.00%. Higher digital maturity coincided with lower probabilistic error, but its adjusted odds-ratio interval included the null value. Earlier studies involving 209 manufacturing companies and 246 Swedish manufacturing enterprises connected digital capabilities with organizational performance, providing contextual evidence rather than direct confirmation of prediction gains. Larger organizations showed greater absolute time savings partly because their baseline schedules were longer, while greater resource flexibility expanded feasible responses. Validation results further qualified interpretation: development accuracy of 96.35% declined to 92.71% under nested validation, with rolling-window and organizational-holdout accuracy of 91.67% and 92.08%, respectively (Jawad & Balázs, 2024). Organizational clustering, missing records, changing risk prevalence, and model-selection procedures affected the interpretation of apparent consistency across analytical methods and project settings. Overlapping evaluation samples did not constitute independent replications, and no external dataset established transferability beyond the twelve organizations. prediction-interval coverage reached 95.83% for duration and 93.75% for expenditure, but these percentages were sensitive to individual holdout cases. Earlier uncertainty research supported examining coverage alongside interval width, while explainability studies distinguished influential predictors from feasible interventions. Feature importance did not independently establish causal action effects. Repeated simulation runs characterized computational variability rather than increasing the number of independent projects (Liu et al., 2020). These constructed findings consequently required interpretation through data independence, decision timing, calibration, resource feasibility, and transparent reporting, with empirical claims restricted to results obtained from independently observed projects.

CONCLUSION

This study presented a quantitative framework connecting artificial intelligence–driven project risk prediction with resource optimization for complex engineering and manufacturing projects. Using an synthetic dataset representing 240 projects across 12 organizations, the analysis distinguished predictive reliability from the operational performance of resource-allocation policies. Gradient boosting achieved delay-classification accuracy of 93.75%, sensitivity of 94.12%, and an F1-score of 91.43% on 48 chronological holdout projects. These measures described different aspects of risk detection, while probability assessment and warning lead time addressed the reliability and timing of decisions. The constructed association analysis linked greater complexity, technical changes, procurement interruptions, and equipment downtime with adverse project outcomes, whereas workforce availability showed an inverse relationship. Integrated optimization reduced mean makes

pan from 462 to 402 days and resource-allocation expenditure from USD 6.80 million to USD 5.90 million per project. Its additional 29-day improvement over optimization without predictive risk inputs separated the contribution of integration from the benefits of optimization alone. Improvements in utilization, tardiness, recovery duration, and schedule stability illustrated the importance of evaluating several operational outcomes together. Subgroup comparisons highlighted differences associated with complexity, organizational scale, digital maturity, and resource flexibility, although small groups and uncertain adjusted estimates limited interpretation. Validation comparisons also showed that apparent development performance differed from nested, chronological, and organizational-holdout results. Repeated simulation runs characterized computational variability without increasing the number of independent projects, while prediction intervals and confidence intervals addressed different forms of uncertainty. The framework therefore provided a structured basis for examining how timely risk information entered feasible labor, equipment, material, and contingency decisions. Consistent accounting boundaries also supported meaningful interpretation of expenditure differences across alternative resource policies. Its principal contribution concerned analytical integration, consistent measurement, and transparent comparison rather than verified evidence of organizational improvement. Because the numerical results were constructed, their coefficients, significance levels, effect sizes, and uncertainty estimates did not establish empirical relationships or causal benefits. The conclusion remained bounded by that distinction: predictive accuracy, operational feasibility, contextual variation, and independent validation constituted complementary requirements for assessing intelligent project management, with substantive performance claims dependent on observed project records and documented execution outcomes.

RECOMMENDATIONS

Engineering and manufacturing organizations should evaluate artificial intelligence–driven risk prediction and resource optimization through controlled implementation supported by reliable project records. Priority should be given to standardized definitions of delays, cost overruns, resource shortages, equipment failures, and recovery periods, together with consistent baseline schedules and expenditure boundaries. Project information should retain timestamps and identifiers that connect activities, procurement transactions, workforce assignments, and equipment conditions without treating repeated records as independent projects. Predictive models should be compared with transparent statistical baselines using chronological testing, project-level separation, and nested model selection. Accuracy should be accompanied by sensitivity, precision, calibration, warning lead time, and uncertainty estimates so that alert thresholds reflect operational consequences. Resource decisions should incorporate precedence requirements, workforce skills, material readiness, machine eligibility, and contingency capacity. Prediction-informed allocation should be compared with conventional planning and optimization without predictive inputs under equivalent conditions, allowing incremental benefits and intervention costs to be identified. Human review should remain part of decisions involving critical milestones, safety-sensitive activities, major procurement changes, or substantial expenditure. Organizations should assess completion duration, total allocation cost, overtime, utilization, delivery reliability, recovery time, and schedule stability together because improvement in one outcome can coincide with deterioration elsewhere. Implementation should begin with limited pilots using observed project data, documented decision rules, and clearly assigned responsibilities. Smaller organizations should prioritize usable records and focused applications before investing in complex architectures, while organizations with integrated information systems should examine whether existing data support reliable cross-project evaluation. Model monitoring should address missing information, changing risk prevalence, shifts in resource availability, and differences between predicted and realized outcomes. Explanations should identify relevant risk drivers while distinguishing model associations from evidence that a proposed action causes improvement. Independent organizational or facility testing should accompany claims of transferability, and computational reporting should include training effort, scheduling runtime, failed solutions, and fallback procedures. Because this study used synthetic numerical findings, its reported performance levels should serve as reporting examples rather than deployment targets. Decisions about adoption should depend on independently verified operational evidence, practical feasibility, full implementation costs, and measurable value within the organization’s own engineering or

manufacturing environment.

LIMITATION

The principal limitation of this study was the use of synthetic findings rather than an independently collected and analyzed empirical dataset. The representation of 240 projects across 12 organizations described a constructed analytical example and did not establish actual organizational participation, sampling coverage, or population representativeness. Classification scores, regression coefficients, odds ratios, confidence intervals, significance levels, and policy effect sizes were reporting examples rather than estimates derived from observed project outcomes. Consequently, the numerical results did not demonstrate predictive superiority, realized financial savings, or causal improvements in engineering and manufacturing performance. The chronological holdout contained only 48 projects and 17 delayed cases, making displayed subgroup percentages sensitive to individual classifications. Perfect accuracy in small categories therefore provided insufficient grounds for comparisons between industrial sectors or organizational conditions. Validation procedures also reused overlapping observations, so nested validation, rolling windows, and organizational holdouts did not represent independent replications. Organizational clustering further reduced the interpretability of nominal sample counts as independent evidence. External testing beyond the represented organizations was absent, limiting claims concerning other countries, facilities, project categories, and management environments. Resource-policy comparisons depended on specified assumptions about activity precedence, workforce skills, machine substitution, material availability, disruption severity, and intervention costs. These assumptions did not capture every contractual, behavioral, technical, or shop-floor constraint affecting implementation. Repeated simulation runs represented computational variation rather than additional independent projects, and high modeled feasibility did not establish practical executability. Cost comparisons also depended on consistent accounting boundaries and did not establish comprehensive lifecycle benefits. Associations among complexity, procurement delays, workforce availability, and equipment downtime remained vulnerable to confounding and reverse temporal relationships when interpreted conceptually. Feature importance and risk explanations likewise did not identify verified causal intervention effects. The presentation lacked an executable analytical dataset, fitted model objects, complete parameter settings, and reproducible simulation outputs, preventing independent verification of the displayed statistics.

REFERENCES

[1]. Abualigah, L., Zitar, R. A., Almotairi, K. H., Hussein, A. M., Abd Elaziz, M., Nikoo, M. R., & Gandomi, A. H. (2022). Wind, solar, and photovoltaic renewable energy systems with and without energy storage optimization: A survey of advanced machine learning and deep learning techniques. *Energies*, *15*(2), 578.

[2]. Achouch, M., Dimitrova, M., Ziane, K., Sattarpanah Karganroudi, S., Dhouib, R., Ibrahim, H., & Adda, M. (2022). On predictive maintenance in industry 4.0: Overview, models, and challenges. *Applied Sciences*, *12*(16), 8081.

[3]. Adeoye, J., Koohi-Moghadam, M., Lo, A. W. I., Tsang, R. K.-Y., Chow, V. L. Y., Zheng, L.-W., Choi, S.-W., Thomson, P., & Su, Y.-X. (2021). Deep learning predicts the malignant-transformation-free survival of oral potentially malignant disorders. *Cancers*, *13*(23), 6054.

[4]. Aljohani, A. (2023). Predictive analytics and machine learning for real-time supply chain risk mitigation and agility. *Sustainability*, *15*(20), 15088.

[5]. Almalawi, A., Soh, B., Li, A., & Samra, H. (2024). Predictive models for educational purposes: A systematic review. *Big Data and Cognitive Computing*, *8*(12), 187.

[6]. Alshihri, S., Al-Gahtani, K., & Almohsen, A. (2022). Risk factors that lead to time and cost overruns of building projects in Saudi Arabia. *Buildings*, *12*(7), 902.

[7]. babu Chandragiri, A., Jeelani, S. H., Akthar, S., & Lingeshwaran, N. (2021). A study and identification of the time and cost overrun in the construction project. *Materials Today: Proceedings*, *47*, 5426-5431.

[8]. Bazargani, K., & Deemyad, T. (2024). Automation's impact on agriculture: opportunities, challenges, and economic effects. *Robotics*, *13*(2), 33.

[9]. Bchennaty, B., Khan, M. N., Massoud, M., & Elhassan, T. (2024). Appraising the role of strategic control in financial performance: The mediating effect of the resource allocation process—The case of the Ministry of Finance–North Lebanon. *International Journal of Financial Studies*, *12*(3), 90.

[10]. Behera, T. K., & Dave, D. M. (2024). From reactive to proactive: Predicting and optimizing performance for competitive advantage. In *Transforming industry using digital twin technology* (pp. 69-93). Springer.

[11]. Carrijo, P., Alturas, B., & Pedrosa, I. (2022). Similarities and differences between digital transformation maturity models: a literature review. *Intelligent Systems in Digital Transformation: Theory and Applications*, 33-52.

[12]. Chattopadhyay, D., & Panteli, M. (2022). Integrating resilience in power system planning: A note for practitioners. *The Electricity Journal*, *35*(9), 107201.

[13]. Cooper, R. G. (2024). The AI transformation of product innovation. *Industrial Marketing Management*, *119*, 62-74.

[14]. Durlik, I., Miller, T., Kostecka, E., & Tuński, T. (2024). Artificial intelligence in maritime transportation: A comprehensive review of safety and risk management applications. *Applied Sciences*, *14*(18), 8420.

[15]. Ezzat, K. A. (2024). Investigating the Impact of Risk Factors on Cost Estimate Inaccuracy Caused by Project Cost Overrun in the Kurdistan Region's Public Sector Construction Projects. In *Resource Management Performance: A Sectoral Analysis in the Post-Conflict Kurdistan Region of Iraq* (pp. 299-346). Springer.

[16]. Gerhardt, C., Semmer, N. K., Sauter, S., Walker, A., de Wijn, N., Kälin, W., Kottwitz, M. U., Kersten, B., Ulrich, B., & Elfering, A. (2021). How are social stressors at work related to well-being and health? A systematic review and meta-analysis. *BMC public health*, *21*(1), 890.

[17]. Harle, S. M. (2024). Advancements and challenges in the application of artificial intelligence in civil engineering: a comprehensive review. *Asian Journal of Civil Engineering*, *25*(1), 1061-1078.

[18]. Hosseini Shirvani, M., & Noorian Talouki, R. (2022). Bi-objective scheduling algorithm for scientific workflows on cloud computing platform with makespan and monetary cost minimization approach. *Complex & Intelligent Systems*, *8*(2), 1085-1114.

[19]. Hou, X., Li, C., & Fang, Q. (2023). Computer vision-based safety risk computing and visualization on construction sites. *Automation in Construction*, *156*, 105129.

[20]. Hu, H., Jiang, S., Goswami, S. S., & Zhao, Y. (2024). Fuzzy integrated Delphi-ISM-MICMAC hybrid multi-criteria approach to optimize the artificial intelligence (AI) factors influencing cost management in civil engineering. *Information*, *15*(5), 280.

[21]. Hussain, A., & Hussain, I. (2023). Modeling and multi-objective optimization of time, greenhouse gas emissions, and resources for sustainable construction projects. *Sustainable production and consumption*, *39*, 269-284.

[22]. Iyortsuun, N. K., Kim, S.-H., Jhon, M., Yang, H.-J., & Pant, S. (2023). A review of machine learning and deep learning approaches on mental health diagnosis. Healthcare,

[23]. Jain, S., Jauhar, S. K., & Piyush. (2024). A machine-learning-based framework for contractor selection and order allocation in public construction projects considering sustainability, risk, and safety. *Annals of Operations Research*, *338*(1), 225-267.

[24]. Jawad, Z. N., & Balázs, V. (2024). Machine learning-driven optimization of enterprise resource planning (ERP) systems: a comprehensive review. *Beni-Suef University Journal of Basic and Applied Sciences*, *13*(1), 4.

[25]. Kang, K.-S., Koo, C., & Ryu, H.-G. (2022). An interpretable machine learning approach for evaluating the feature importance affecting lost workdays at construction sites. *Journal of Building Engineering*, *53*, 104534.

[26]. Karan, A., Negandhi, H., Hussain, S., Zapata, T., Mairembam, D., De Graeve, H., Buchan, J., & Zodpey, S. (2021). Size, composition and distribution of health workforce in India: why, and where to invest? *Human resources for health*, *19*(1), 39.

[27]. Katsaliaki, K., Galetsi, P., & Kumar, S. (2022). Supply chain disruptions and resilience: a major review and future research agenda. *Annals of Operations Research*, *319*(1), 965-1002.

[28]. Kgakatsi, M., Galeboe, O. P., Molelekwa, K. K., & Thango, B. A. (2024). The impact of big data on SME performance: A systematic review. *Businesses*, *4*(4), 632-695.

[29]. Khodabakhshian, A., Puolitaival, T., & Kestle, L. (2023). Deterministic and probabilistic risk management approaches in construction projects: A systematic literature review and comparative analysis. *Buildings*, *13*(5), 1312.

[30]. Kuang, L., He, L., Yili, R., Kai, L., Mingyu, S., Jian, S., & Xin, L. (2021). Application and development trend of artificial intelligence in petroleum exploration and development. *Petroleum Exploration and Development*, *48*(1), 1-14.

[31]. Leng, J., Liu, Q., Ye, S., Jing, J., Wang, Y., Zhang, C., Zhang, D., & Chen, X. (2020). Digital twin-driven rapid reconfiguration of the automated manufacturing system via an open architecture model. *Robotics and computer-integrated manufacturing*, *63*, 101895.

[32]. Liaqait, R. A., Hamid, S., Warsi, S. S., & Khalid, A. (2021). A critical analysis of job shop scheduling in context of industry 4.0. *Sustainability*, *13*(14), 7684.

[33]. Lingard, H., & Turner, M. (2022). Making time for life: A whole-of-industry initiative to reducing work hours and promoting health and gender inclusion in project-based construction work. *Project leadership and society*, *3*, 100065.

[34]. Liu, H., Ma, T., Lin, Y., Peng, K., Hu, X., Xie, S., & Luo, K. (2024). Deep learning in rockburst intensity level prediction: performance evaluation and comparison of the NGO-CNN-BiGRU-attention model. *Applied Sciences*, *14*(13), 5719.

[35]. Liu, M. S., Yao, M. N., Qiu, M. Y., & He, M. C. (2020). Predictive performance of SOFA and qSOFA for in-hospital mortality in severe novel coronavirus disease. *The American Journal of Emergency Medicine*, *38*(10), 2074-2080.

[36]. Lv, J., Zhang, R., Shama, A., Hong, R., He, X., Wu, R., Bao, X., & Liu, G. (2024). Exploring the spatial patterns of landslide susceptibility assessment using interpretable Shapley method: Mechanisms of landslide formation in the Sichuan-Tibet region. *Journal of environmental management*, *366*, 121921.

[37]. M. Abdelhaleem Ali, A., & M. Alrobaian, M. (2024). Strengths and weaknesses of current and future prospects of artificial intelligence-mounted technologies applied in the development of pharmaceutical products and services. *Saudi Pharmaceutical Journal*, *32*(5), 1-7.

[38]. Mahmud, M. H., Nayan, M. T. H., Ashir, D. M. N. A., & Kabir, M. A. (2022). Software risk prediction: systematic literature review on machine learning techniques. *Applied Sciences*, *12*(22), 11694.

[39]. Mehdiyev, N., Evermann, J., & Fettke, P. (2020). A Novel Business Process Prediction Model Using a Deep Learning Method: N. Mehdiyev et al. *Business & information systems engineering*, *62*(2), 143-157.

[40]. Molladavoudi, S., & Yung, W. (2023). Exploring quality dimensions in trustworthy Machine Learning in the context of official statistics: model explainability and uncertainty quantification. *AStA Wirtschafts-und Sozialstatistisches Archiv*, *17*(3), 223-252.

[41]. Morariu, C., Morariu, O., Răileanu, S., & Borangiu, T. (2020). Machine learning for predictive scheduling and resource allocation in large scale manufacturing systems. *Computers in Industry*, *120*, 103244.

[42]. Mwanza, J., Telukdarie, A., & Igusa, T. (2023). Optimising maintenance workflows in healthcare facilities: a multi-scenario discrete event simulation and simulation annealing approach. *Modelling*, *4*(2), 224-250.

[43]. Nor, A. K. M., Pedapati, S. R., Muhammad, M., & Leiva, V. (2021). Overview of explainable artificial intelligence for prognostic and health management of industrial assets based on preferred reporting items for systematic reviews and meta-analyses. *Sensors*, *21*(23), 8020.

[44]. Nosratabadi, S., Mosavi, A., Duan, P., Ghamisi, P., Filip, F., Band, S. S., Reuter, U., Gama, J., & Gandomi, A. H. (2020). Data science in economics: comprehensive review of advanced machine learning and deep learning methods. *Mathematics*, *8*(10), 1799.

[45]. Oztemel, E., & Gursev, S. (2020). Literature review of Industry 4.0 and related technologies. *Journal of Intelligent Manufacturing*, *31*(1), 127-182.

[46]. Pachiyannan, P., Alsulami, M., Alsadie, D., Saudagar, A. K. J., AlKhathami, M., & Poonia, R. C. (2023). A Cardiac Deep Learning Model (CDLM) to predict and identify the risk factor of congenital heart disease. *Diagnostics*, *13*(13), 2195.

[47]. Palm, H., & Arndt, L. (2023). Reinforcement learning-based hybrid multi-objective optimization algorithm design. *Information*, *14*(5), 299.

[48]. Pamidimukkala, A., & Kermanshachi, S. (2021). Impact of Covid-19 on field and office workforce in construction industry. *Project leadership and society*, *2*, 100018.

[49]. Perifanis, N.-A., & Kitsios, F. (2023). Investigating the influence of artificial intelligence on business value in the digital era of strategy: A literature review. *Information*, *14*(2), 85.

[50]. Qiu, Y., & Zhou, J. (2023). Short-term rockburst prediction in underground project: insights from an explainable and interpretable ensemble learning model. *Acta Geotechnica*, *18*(12), 6655-6685.

[51]. Qureshi, M., Iftikhar, H., Rodrigues, P. C., Rehman, M. Z., & Salar, S. A. (2024). Statistical modeling to improve time series forecasting using machine learning, time series, and hybrid models: A case study of bitcoin price forecasting. *Mathematics*, *12*(23), 3666.

[52]. Radanliev, P., De Roure, D., Walton, R., Van Kleek, M., Montalvo, R. M., Maddox, L. T., Santos, O., Burnap, P., & Anthi, E. (2020). Artificial intelligence and machine learning in dynamic cyber risk analytics at the edge. *SN Applied Sciences*, *2*(11), 1773.

[53]. Raza, S. M., Sajid, M., & Singh, J. (2022). Vehicle routing problem using reinforcement learning: Recent advancements. In *Advanced machine intelligence and signal processing* (pp. 269-280). Springer.

[54]. Reich-Stiebert, N., Froehlich, L., & Voltmer, J.-B. (2023). Gendered mental labor: A systematic literature review on the cognitive dimension of unpaid work within the household and childcare. *Sex Roles*, *88*(11), 475-494.

[55]. Riad, M., Naimi, M., & Okar, C. (2024). Enhancing supply chain resilience through artificial intelligence: developing a comprehensive conceptual framework for AI implementation and supply chain optimization. *Logistics*, *8*(4), 111.

[56]. Rivera, L., Baguec Jr, H., & Yeom, C. (2020). A study on causes of delay in road construction projects across 25 developing countries. *Infrastructures*, *5*(10), 84.

[57]. Sahal, R., Breslin, J. G., & Ali, M. I. (2020). Big data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case. *Journal of manufacturing systems*, *54*, 138-151.

[58]. Senna, P. P., Almeida, A. H., Barros, A. C., Bessa, R. J., & Azevedo, A. L. (2020). Architecture model for a holistic and interoperable digital energy management platform. *Procedia Manufacturing*, *51*, 1117-1124.

[59]. Sircar, A., Yadav, K., Rayavarapu, K., Bist, N., & Oza, H. (2021). Application of machine learning and artificial intelligence in oil and gas industry. *Petroleum Research*, *6*(4), 379-391.

[60]. Suriyaamporn, P., Pamornpathomkul, B., Patrojanasophon, P., Ngawhirunpat, T., Rojanarata, T., & Opanasopit, P. (2024). The artificial intelligence-powered new era in pharmaceutical research and development: a review. *Aaps Pharmscitech*, *25*(6), 188.

[61]. Tang, X., Zhou, H., Wang, F., Wang, W., & Lin, X. (2022). Longevity-conscious energy management strategy of fuel cell hybrid electric Vehicle Based on deep reinforcement learning. *Energy*, *238*, 121593.

[62]. Temitope Yekeen, S., & Balogun, A.-L. (2020). Advances in remote sensing technology, machine learning and deep learning for marine oil spill detection, prediction and vulnerability assessment. *Remote Sensing*, *12*(20), 3416.

[63]. Usuga Cadavid, J. P., Lamouri, S., Grabot, B., Pellerin, R., & Fortin, A. (2020). Machine learning applied in production planning and control: a state-of-the-art in the era of industry 4.0. *Journal of Intelligent Manufacturing*, *31*(6), 1531-1558.

[64]. Vahedi Nikbakht, M., Gheibi, M., Montazeri, H., Yeganeh Khaksar, R., Moezzi, R., & Vadiee, A. (2024). Identification and ranking of factors affecting the delay risk of high-rise construction projects using AHP and VIKOR methods. *Infrastructures*, *9*(2), 24.

[65]. Wang, C., Wu, X., Xie, Z., & Kozlowski, T. (2023). Scalable inverse uncertainty quantification by hierarchical bayesian modeling and variational inference. *Energies*, *16*(22), 7664.

[66]. Wang, Z., & Liao, W. (2024). Smart scheduling of dynamic job shop based on discrete event simulation and deep reinforcement learning. *Journal of Intelligent Manufacturing*, *35*(6), 2593-2610.

[67]. Winkelmann, J., Webb, E., Williams, G. A., Hernández-Quevedo, C., Maier, C. B., & Panteli, D. (2022). European countries' responses in ensuring sufficient physical infrastructure and workforce capacity during the first COVID-19 wave. *Health Policy*, *126*(5), 362-372.

[68]. Wright, L., & Davidson, S. (2020). How to tell the difference between a model and a digital twin. *Advanced Modeling and Simulation in Engineering Sciences*, *7*(1), 13.
[69]. Wu, C.-H., Chou, C.-W., Chien, C.-F., & Lin, Y.-S. (2024). Digital transformation in manufacturing industries: Effects of firm size, product innovation, and production type. *Technological forecasting and social change*, *207*, 123624.
[70]. Wu, D., Wang, Q., & Olson, D. L. (2023). Industry classification based on supply chain network information using graph neural networks. *Applied Soft Computing*, *132*, 109849.
[71]. Yasin, P., Yimit, Y., Cai, X., Aimaiti, A., Sheng, W., Mamat, M., & Nijiati, M. (2024). Machine learning-enabled prediction of prolonged length of stay in hospital after surgery for tuberculosis spondylitis patients with unbalanced data: a novel approach using explainable artificial intelligence (XAI). *European journal of medical research*, *29*(1), 383.
[72]. Yousif, I., Samaha, J., Ryu, J., & Harik, R. (2024). Safety 4.0: Harnessing computer vision for advanced industrial protection. *Manufacturing Letters*, *41*, 1342-1356.
[73]. Zhang, Y., Bai, R., Qu, R., Tu, C., & Jin, J. (2022). A deep reinforcement learning based hyper-heuristic for combinatorial optimisation with uncertainties. *European Journal of Operational Research*, *300*(2), 418-427.
[74]. Aditya, D., & Ahmed, H. C. H. R. (2022). Decision Support Systems for Last-Mile Logistics Optimization: A Data-Driven Framework for e-Commerce Fulfillment Networks. *International Journal of Scientific Interdisciplinary Research*, *12*(08), 01–49. https://doi.org/10.63125/h46vp562
[75]. Adnan, S. (2024). Human Trust in AI-Assisted Project Decision Systems: A Mixed-Methods Analysis of Explainability, Adoption, and Managerial Decision Quality. *Journal of Sustainable Development and Policy*, *3*(04), 287-332. https://doi.org/10.63125/9c139h44
[76]. Aminul Islam, M. (2024). Interpreting Cyber-Threat Indicators in Enterprise Memory Forensics: A Mixed-Methods Study of Machine Learning Performance and Explainability. *Review of Applied Science and Technology*, *14*(12), 01–13. https://doi.org/10.63125/2p8hte92
[77]. Aminul Islam, M. (2025). Calibrated Machine Learning for Reducing Decision Risk and Improving Forecast Reliability in U.S. Enterprise Analytics. *American Journal of Interdisciplinary Studies*, *15*(12), 01-36. https://doi.org/10.63125/869qdj55
[78]. Aminul Islam, M., & Md, A. (2023). Adaptive Class-Imbalance Learning for High-Sensitivity Fraud Detection in Large-Scale Digital Transaction Networks. *International Journal of Business and Economics Insights*, *13*(09), 01-33. https://doi.org/10.63125/bmxwv741
[79]. Asheke, M. (2023). Machine Learning–Based Predictive Quality Modeling for Defect Detection and Process Optimization in Advanced Manufacturing Systems. *Review of Applied Science and Technology*, *13*(12), 01–49. https://doi.org/10.63125/thy73b16
[80]. Asheke, M. (2025). Artificial Intelligence Driven Supply Chain Risk Forecasting and Optimization for Resilient Industrial Production Networks. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 2923–2960. https://doi.org/10.63125/rapn5860
[81]. Asheke, M., & Md Maksudul, H. (2022). AI-Driven Demand Forecasting for Sporting-Goods and Teamwear Supply Chains Under Seasonal and Tournament Demand. *International Journal of Business and Economics Insights*, *2*(1), 73–110. https://doi.org/10.63125/ed6by557
[82]. Asheke, M., & Mohammad Badrul, A. (2022). Development Of Functionalized Cellulose Nanofiber Adsorbents for Selective Critical-Metal Recovery from Industrial Wastewater. *American Journal of Scholarly Research and Innovation*, *1*(02), 270–320. https://doi.org/10.63125/v6fxzq09
[83]. Ashik, M. (2024). Meta-Analysis of Risk-Based Inspection Strategies for Improving Reliability and Safety Across Critical Energy Infrastructure Systems. *American Journal of Advanced Technology and Engineering Solutions*, *14*(09), 01-43. https://doi.org/10.63125/tgttkt09
[84]. Ashik, M. (2025). PRISMA-Based Systematic Literature Review of Reliability-Centered Turnaround Planning for Reducing Downtime in Power Generation Facilities. *American Journal of Scholarly Research and Innovation*, *4*(01), 1036–1080. https://doi.org/10.63125/2878n718
[85]. Ashik, M. (2026). RCM and PSM Strategies for Strengthening Asset Integrity Across U.S. Energy Facilities: A Systematic Meta-Analysis. *American Journal of Interdisciplinary Studies*, *16*(06), 38-81. https://doi.org/10.63125/p87z7q36
[86]. Gazi Yeasin, A. (2024a). Designing Trustworthy AI Output Interfaces: A UX Framework for Graceful Handling of Generative Model Uncertainty in Consumer Applications. *American Journal of Scholarly Research and Innovation*, *14*(07), 01–40. https://doi.org/10.63125/4m5fjb12
[87]. Gazi Yeasin, A. (2024b). A Quantitative Evaluation of Behavioral Design Patterns on User Engagement and Activation in Large-Scale Campus Software Products. *American Journal of Interdisciplinary Studies*, *14*(06), 01-40. https://doi.org/10.63125/x3d5k517
[88]. Gazi Yeasin, A. (2026). An AI-Integrated Multi-Agent Product Development Pipeline for Autonomous Design-To-Production Software Delivery: Architecture and Evaluation. *International Journal of Scientific Interdisciplinary Research*, *7*(1), 632-676. https://doi.org/10.63125/bq24tw25
[89]. Hasib, A. (2026). AI-Driven Predictive Supply Chain Resilience Framework for U.S. Manufacturing Systems: An Empirical Analysis. *American Journal of Data Science and Analytics*, *7*(04), 45-80. https://doi.org/10.63125/d8swft81
[90]. Jahanara, A. (2024). Health Information Systems Governance in Home Health Agencies: A Managerial Framework and Post-Adoption Oversight of Scheduling and Cost-Control Systems. *American Journal of Health and Medical Sciences*, *14*(06), 01–50. https://doi.org/10.63125/jwz00f70

[91]. Kazi Mohammad Khalid, A., & Sumaiya Sadmeen, M. (2022). Geospatial Artificial Intelligence for Urban Flood Susceptibility and Drainage Resilience: A Prisma 2020 Systematic Review and Bibliometric Meta-Synthesis. *Review of Applied Science and Technology*, *12*(02), 01–51. https://doi.org/10.63125/epqdvy23

[92]. Kynatun, M. (2025). Integrating AI-Driven Security Analytics and IT Governance Frameworks for Cybersecurity Decision Support in U.S. Financial Information Systems. *American Journal of Advanced Technology and Engineering Solutions*, *15*(12), 01-42. https://doi.org/10.63125/kzcdyz73

[93]. Kynatun, M. (2026). Artificial Intelligence Integrated MIS Framework for Proactive Cyber Threat Intelligence and Organizational Security Resilience. *American Journal of Scholarly Research and Innovation*, *16*(09), 40–78. https://doi.org/10.63125/t98tc495

[94]. Kynatun, M., & Saba, A. (2024). Machine Learning Enhanced Intrusion Detection and Anomaly Classification for MIS-Driven Enterprise Network Security. *International Journal of Scientific Interdisciplinary Research*, *14*(09), 01–41. https://doi.org/10.63125/xk6kpc72

[95]. Maishar Mostari, Z. (2024). Uncertainty-Calibrated Explainable AI for Cybersecurity Alert Prioritization: Improving Detection Reliability and Human Decision-Making. *American Journal of Scholarly Research and Innovation*, *14*(12), 01–40. https://doi.org/10.63125/an0pcq39

[96]. Maishar Mostari, Z. (2025). Privacy-Preserving AI-Based Cyber Threat Detection in Financial Organizations: A Governance and Business Value Framework. *American Journal of Interdisciplinary Studies*, *15*(04), 01-51. https://doi.org/10.63125/p8fct530

[97]. Maishar Mostari, Z. (2026). A Systematic Review of Machine Learning Applications in Cybersecurity Threat Detection and Classification (2018–2026). *American Journal of Data Science and Analytics*, *7*(06), 640-680. https://doi.org/10.63125/6kg7kg75

[98]. Maishar Mostari, Z., & Kynatun, M. (2023). Privacy Leakage Through Repeated AI Explanations in Collaborative Financial Cybersecurity Systems. *International Journal of Business and Economics Insights*, *13*(06), 01-34. https://doi.org/10.63125/azdhg571

[99]. Mansura Akter, E., & Sumaiya, M. (2022). Antimicrobial Resistance Profiles of Acinetobacter Baumannii and Vibrio Cholerae in Hospital Wastewater and Adjacent Household Water Sources. *American Journal of Data Science and Analytics*, *3*(09), 01-55. https://doi.org/10.63125/gbervt14

[100]. Mashfiqur, R. (2024). Anomaly Detection and Process Mining for Identifying Financial Irregularities and Strengthening Procurement Controls. *American Journal of Advanced Technology and Engineering Solutions*, *14*(09), 91-140. https://doi.org/10.63125/hmg5ta84

[101]. Mashfiqur, R., & Md Jamil, A. (2023). Machine Learning-Based Cost Variance Prediction for Early Detection of Budget Overruns in Procurement-Intensive Operations. *Review of Applied Science and Technology*, *13*(12), 142–199. https://doi.org/10.63125/7g92wd03

[102]. Mashiat, N., & Md. Morshedul, I. (2022). Reducing Data Processing Latency in High Volume Telecommunications Networks Through Adaptive ETL and Distributed Computing. *American Journal of Scholarly Research and Innovation*, *12*(06), 01–36. https://doi.org/10.63125/b5xzay30

[103]. Mashiat, N., & Taru Binte, A. (2023). Preventing Data Pipeline Failures Through Machine Learning-Based Anomaly Detection in Mission-Critical Network Operations. *American Journal of Interdisciplinary Studies*, *13*(12), 48-87. https://doi.org/10.63125/jwvy9t14

[104]. Masrul, H. (2025). Explainable Machine Learning for Real-Time Financial Fraud Detection and Risk-Based Transaction Monitoring in U.S. Banking and Digital Payment Systems. *American Journal of Data Science and Analytics*, *6*(10), 70-108. https://doi.org/10.63125/4h9v1876

[105]. Masrul, H., & Mashfiqur, R. (2024). Unified Analytical Modeling for Cryptocurrency Price Volatility Forecasting and Market-Manipulation Detection in U.S. Digital Asset Markets. *American Journal of Advanced Technology and Engineering Solutions*, *14*(11), 01-40. https://doi.org/10.63125/nqgh7934

[106]. Masrul, H., & Mukut Kanti, B. (2024). AI-Driven Project Portfolio and Business Analytics for Forecasting Cost Overruns, Schedule Delays, Resource Bottlenecks, and Enterprise Performance Across U.S. Industries. *Review of Applied Science and Technology*, *14*(12), 14–54. https://doi.org/10.63125/x7te4644

[107]. Md Hasinur, R. (2024). Human in the Loop Artificial Intelligence for Loan Document Verification and Credit Risk Reporting: A Bibliometric and Science Mapping Analysis. *American Journal of Interdisciplinary Studies*, *14*(06), 41-71. https://doi.org/10.63125/8x359j95

[108]. Md Hasinur, R. (2025a). Machine Learning for Nonperforming Corporate Loan Recovery: A Systematic Review and Meta-Analysis of Predictive Performance. *American Journal of Scholarly Research and Innovation*, *4*(01), 1174–1202. https://doi.org/10.63125/a7v6f533

[109]. Md Hasinur, R. (2025b). Present and Future of Digital Banking Services in Bangladesh. *Review of Applied Science and Technology*, *4*(02), 725-749. https://doi.org/10.63125/k7tgr183

[110]. Md Hasinur, R. (2026). Explainable Machine Learning for Loan-Delinquency Prediction in U.S. Community Banks: A Quantitative Study Using Gradient Boosting and SHAP. *American Journal of Data Science and Analytics*, *7*(01), 280-316. https://doi.org/10.63125/mfv3c059

[111]. Md Imran, H., & Jannatul, F. (2022). Molecular and Precision Rehabilitation Molecular Mechanisms Linking Mitochondrial Dysfunction, Inflammation, and Muscle Degeneration with Functional Decline and Rehabilitation Response in Neuromuscular Disease. *American Journal of Interdisciplinary Studies*, *12*(08), 01-44. https://doi.org/10.63125/0ym5bh79

[112]. Md Maksudul, H. (2024). Supplier-Risk Scoring and Procurement Resilience for Distributed Athletic-Apparel and Equipment Supply Networks. *American Journal of Advanced Technology and Engineering Solutions*, *14*(09), 44-90. https://doi.org/10.63125/a9fmgj50

[113]. Md Maksudul, H. (2025). Data-Driven Fulfillment Optimization for Time-Sensitive Teamwear and Sporting Goods Distribution. *American Journal of Scholarly Research and Innovation*, *4*(01), 1120–1173. https://doi.org/10.63125/rpezfz61

[114]. Md Maksudul, H. (2026). Generative AI Decision Support for Procurement, Inventory, And Operational Planning in U.S. Sports Businesses. *American Journal of Interdisciplinary Studies*, *16*(04), 01-55. https://doi.org/10.63125/ngtzw226

[115]. Md Maksudul, H., & Md Tazul Islam, C. (2022). Machine Learning-Based Cash-Flow Forecasting for Improving Financial Planning in U.S. Small and Mid-Sized Businesses. *International Journal of Scientific Interdisciplinary Research*, *12*(12), 01–38. https://doi.org/10.63125/w3pgjx97

[116]. Md Mohsin, B. (2026). Nurse-Led Triage Protocols in Emergency and Trauma Care: A PRISMA-Guided Systematic Review. *American Journal of Health and Medical Sciences*, *16*(09), 01–40. https://doi.org/10.63125/21k00924

[117]. Md Momen Shahriar, J. (2025). Viscoelastic Creep Characterization of LDPE Under Multilevel Stress and Temperature Using Time-Temperature Superposition and Constitutive Modeling. *American Journal of Scholarly Research and Innovation*, *15*(11), 01–38. https://doi.org/10.63125/2edstd86

[118]. Md Momen Shahriar, J. (2026). Physics-Informed Machine Learning for Long-Term Polymer Performance Prediction from Extrusion Parameters, Material Properties, and Creep Data. *American Journal of Interdisciplinary Studies*, *16*(05), 01-41. https://doi.org/10.63125/8gk5c668

[119]. Md Shahid, N. (2023). Digital Twin-Based Network Reliability Modeling for LAN And WAN Infrastructure Resilience. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 577–629. https://doi.org/10.63125/ahjvxj14

[120]. Md Shahid, N. (2024). Predictive Analytics Models for Network Traffic Optimization in Distributed Enterprise Environments. *Review of Applied Science and Technology*, *3*(01), 367-419. https://doi.org/10.63125/qm6cyf84

[121]. Md Shahid, N. (2025). Blockchain-Based Secure Data Routing Framework for U.S. Industrial Communication Networks. *American Journal of Scholarly Research and Innovation*, *4*(01), 990–1035. https://doi.org/10.63125/9exrgs41

[122]. Md Shahid, N. (2026a). AI-Driven Quantum-Inspired Predictive Analytics Framework for Procurement Optimization in Critical Infrastructure Networks. *American Journal of Interdisciplinary Studies*, *7*(01), 658-702. https://doi.org/10.63125/hzecta02

[123]. Md Shahid, N. (2026b). Secure Cyber-Physical Data Monitoring Architecture for Real-Time Fault Detection in U.S. Industrial Automation Systems. *American Journal of Data Science and Analytics*, *7*(10), 90-125. https://doi.org/10.63125/exzc2653

[124]. Md Tazul Islam, C. (2024). Explainable AI For Financial and Operational Decision Support: Improving Transparency and Managerial Trust. *American Journal of Advanced Technology and Engineering Solutions*, *14*(12), 01-38. https://doi.org/10.63125/kwfvzw68

[125]. Md Tazul Islam, C. (2025). Business Intelligence and Data-Quality Frameworks for Improving Evidence-Based Decisions across Multi-Source Enterprise Data. *American Journal of Scholarly Research and Innovation*, *15*(07), 01–41. https://doi.org/10.63125/qhh6ka59

[126]. Md Tazul Islam, C., & Ahmed, H. C. H. R. (2023). Predictive Operations Analytics for Resource Allocation and Service-Capacity Planning in Data-Driven Organizations. *Review of Applied Science and Technology*, *13*(02), 01–41. https://doi.org/10.63125/z3sm3089

[127]. Md. Mainuddin, F. (2024). Quantitative Structural Retrofit Assessment Models for Strengthening Existing Steel Buildings Under Increased Load Demands. *Review of Applied Science and Technology*, *3*(04), 325–366. https://doi.org/10.63125/yyqnte84

[128]. Md. Mainuddin, F. (2025). Advanced Engineering Materials Applications for Enhancing Durability and Lifecycle Performance of Steel Building Systems. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 2406–2445. https://doi.org/10.63125/t9xvg986

[129]. Md. Mainuddin, F. (2026). AI-Enabled Performance-Based Structural Optimization of Steel Truss and Mezzanine Systems for Industrial Facilities. *American Journal of Data Science and Analytics*, *7*(02), 01-38. https://doi.org/10.63125/enhzbm16

[130]. Md. Morshedul, I., & Aminul Islam, M. (2022). A Quantitative Evaluation of Explainable Ensemble Learning for Probabilistic Risk Estimation and Decision Reliability in Enterprise Analytics. *American Journal of Data Science and Analytics*, *3*(02), 01-38. https://doi.org/10.63125/k22shk74

[131]. Sadia, A. (2025). Privacy-Aware Personalization: Machine Learning, Model Validation, And Consumer Trust in AI-Mediated Marketing. *American Journal of Scholarly Research and Innovation*, *4*(01), 1081–1119. https://doi.org/10.63125/j7tfr242

[132]. Sadia, A. (2026). A Data-Science Decision-Support Framework for Responsible AI-Enabled Marketing, Visualization, And Consumer Engagement in The United States. *American Journal of Interdisciplinary Studies*, *7*(02), 277-311. https://doi.org/10.63125/e67w9e54

[133]. Sadia, A., & Md. Mosheur, R. (2023). Consumer Trust, Perceived Authenticity, and Personalized Digital Marketing: A Quantitative Decision-Science Study. *International Journal of Scientific Interdisciplinary Research*, *4*(2), 252–295. https://doi.org/10.63125/cajkqt60

[134]. Shabbir, A. (2025). Physics-Based Failure Prognostics for Preventing Pressure-System Degradation and Unplanned Industrial Shutdowns. *American Journal of Scholarly Research and Innovation*, *14*(09), 01–41. https://doi.org/10.63125/3asd0053

[135]. Shabbir, A. (2026). A Data-Science Digital Twin-Enabled Thermal Performance Optimization for Reducing Energy Losses in Industrial Steam Generation Systems. *American Journal of Data Science and Analytics*, *7*(01), 317-363. https://doi.org/10.63125/ybf40824

[136]. Stabak, D., & Md Shahid, N. (2022). Empirical Assessment of Uncertainty-Aware Real-Time Decision-Making for Reliable and Resilient Cyber-Physical Systems. *International Journal of Scientific Interdisciplinary Research*, *1*(01), 432–477. https://doi.org/10.63125/q3r4kk16

[137]. Sujon, K. M., Miraz, A. I., Rahman, T., Shehun, M. T. F., Shafi, A., Islam, M. M., & Molla, M. M. I. (2025). A Novel Hybrid Stacked Ensemble Model with XAI for Robust and Interpretable HIV Infection Prediction. 2025 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON),

[138]. Sujon, K. M., Miraz, A. I., Shafi, A., Rahman, T., Shehun, M. T. F., Islam, M. M., & Molla, M. M. I. (2025). A Novel Approach to Malware Detection in Memory Forensics Via Adaptive Weighted Voting With Error Correlation Correction. 2025 28th International Conference on Computer and Information Technology (ICCIT),

[139]. Sujon, K. M., Molla, M. M. I., Miraz, A. I., Islam, M. M., Shehun, M. T. F., & Shafi, A. (2025). The Impact of ADASYN on Distance-Based, Tree-Based, Ensemble, and Boosting Models in Breast Cancer Prediction. 2025 IEEE International Conference on Signal Processing, Information, Communication and Systems (SPICSCON),

[140]. Sujon, K. M., Shafi, A., Islam, M. M., Miraz, A. I., Shehun, M. T. F., Rahman, T., & Molla, M. M. I. (2025). A Lightweight CNN Model for Interpretable White Blood Cell Classification with Grad-CAM and Grad-CAM++. 2025 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON),

[141]. Sujon, K. M., Shafi, A., Molla, M. M. I., Miraz, A. I., Islam, M. M., & Shehun, M. T. F. (2025). RXK-VEM: A Novel Probabilistically Calibrated Hybrid Ensemble Model for Predicting Student Dropout in Higher Education. 2025 IEEE 2nd International Conference on Computing, Applications and Systems (COMPAS),

[142]. Sujon, K. M., Shehun, M. T. F., Islam, M. M., Rahman, T., Shafi, A., Miraz, A. I., & Molla, M. M. I. (2025). An Explainable Framework for Multiclass Student Performance Prediction: Integrating SMOTE and Correlation-Based Feature Selection. 2025 IEEE International Conference on Signal Processing, Information, Communication and Systems (SPICSCON),

[143]. Sumaiya, M. (2024). Hospital Wastewater as A Reservoir of Antimicrobial Resistance Genes: Implications for Public Health and Genomic Surveillance in The United States. *Review of Applied Science and Technology*, *3*(04), 448–508. https://doi.org/10.63125/ahvas976

[144]. Sumaiya, M. (2026). Artificial Intelligence Driven Multi-Omics Approaches for Precision Medicine: From Molecular Discovery to Clinical Translation. *American Journal of Scholarly Research and Innovation*, *5*(01), 429–477. https://doi.org/10.63125/yfme0q12

[145]. Sumaiya Sadmeen, M. (2025). Spatiotemporal Graph Neural Networks and Digital Twins for Urban Drainage Networks: Predictive Inundation, Critical-Node Detection, And Infrastructure Prioritization. *American Journal of Scholarly Research and Innovation*, *15*(12), 01-40. https://doi.org/10.63125/awvjew66

[146]. Sumaiya Sadmeen, M. (2026). Uncertainty-Aware GeoAI Decision Support for Climate-Resilient and Equitable Urban Infrastructure Investment in U.S. Communities. *American Journal of Interdisciplinary Studies*, *16*(05), 42-84. https://doi.org/10.63125/ktww0w31

[147]. Sumaiya Sadmeen, M., & Kazi Mohammad Khalid, A. (2023). Multi-Source Remote Sensing and Dem-Derived Hydrogeomorphic Indicators for High-Resolution Urban Flood Susceptibility Mapping: A Spatial Cross-Validation Framework. *Journal of Sustainable Development and Policy*, *13*(07), 01-46. https://doi.org/10.63125/z4xt7k38

[148]. Syed Feroz, A., & Md Mahamud Pervaz, P. (2022). Root-Cause Analysis and FMEA-Driven Reliability Improvement in Gas Production Facilities: A Multi-Year Empirical Study. *International Journal of Scientific Interdisciplinary Research*, *12*(10), 01–53. https://doi.org/10.63125/70msp878

[149]. Syed Feroz, A., & Shabbir, A. (2022). Finite Element-Based Thermomechanical Optimization of Pressure Components for Extending Industrial Boiler Service Life. *Journal of Sustainable Development and Policy*, *12*(12), 01-51. https://doi.org/10.63125/x24pht28

[150]. Taher, M. A., Sikder, M. A., Ahmed, H., & Bashar, M. (2024). Integrating IOT and Big Data Analytics for Enhanced Supply Chain Performance in Industrial Engineering Sectors: A Cross-Market Study. *International Journal of Science and Engineering*, *Volume 1*, 1-14.

[151]. Tahmedur, R. (2024). Machine-Learning-Assisted Project Cost Estimation for Technology-Intensive Engineering Programs. *International Journal of Scientific Interdisciplinary Research*, *14*(12), 01–54. https://doi.org/10.63125/7gfgm030

[152]. Tahmedur, R. (2025). Development of a Decision-Support Model for Agile Hybrid Engineering Project Delivery in Technology Driven Organizations. *Review of Applied Science and Technology*, *4*(04), 339–381. https://doi.org/10.63125/v7jjv448

[153]. Umme, H. (2026). Transformer-Based Multi-Omics Integration for Cancer Drug Response Prediction: A PRISMA 2020 Systematic Review. *American Journal of Health and Medical Sciences*, *16*(01), 33–52. https://doi.org/10.63125/teryrs73