ESP-JETA
ESP Journal of Engineering & Technology Advancements
ISSN: 2583-2646 / Volume 5 Issue 4 October 2025 / Page No: 1-6
Paper Id: JETA-V5I4P101 / Doi: 10.56472/25832646/JETA-V5I4P101

Original Article

AI-Enhanced Anomaly Detection for Project Performance: A Cross-Industry Study for Technology-Driven Industries

Shreya Makinani[1], Pankaj Siri Bharath Bairu[2]
[1]*Program Manager, USA.*
[2]*School of Computer Information Sciences, University of the Cumberlands, USA.*

Received Date: 18 August 2025 Revised Date: 05 September 2025 Accepted Date: 01 October 2025

Abstract: *Monitoring project performance is a cornerstone of success in technology-driven industries. Projects in semiconductors, software/IT, and retail (supply chain) are increasingly complex, requiring robust anomaly detection methods to identify deviations in schedule, cost, quality, and throughput. Traditional approaches are often siloed, applying statistical thresholds or isolated machine learning techniques to single domains. This paper presents an AI-enhanced, KPI-driven anomaly detection framework validated on real-world datasets. Experiments were conducted using semiconductor process datasets, software defect repositories, and retail supply-chain data, proving that the proposed framework enhances the detection of point, contextual, and collective anomalies. The results demonstrate improvements in anomaly separability, supervised accuracy, and interpretability. This establishes that AI-enhanced anomaly detection can strengthen project monitoring across industries. The novelty of this work lies in presenting a generalizable KPI-driven methodology applicable across semiconductors, software/IT, and retail (supply chain).*

Keywords: *Anomaly Detection, Project Performance, Key Performance Indicators, Semiconductors, Software, Retail, Supply Chain, Machine Learning, Point Anomalies, Contextual Anomalies, Collective Anomalies.*

I. INTRODUCTION

Technology-driven industries operate in environments where projects are highly complex, resource intensive, and time sensitive. Semiconductors face strict constraints on design cycle times, yield targets, and validation schedules, where a single delay may cascade across the supply chain and increase costs exponentially [1]. Software and IT projects are equally vulnerable, with challenges ranging from missed sprint deadlines and cost overruns to declining throughput in deployment pipelines [2]. In retail supply chains, disruptions such as inventory imbalances or transportation delays can directly impact customer satisfaction and financial performance [3]. Despite these differences, projects in all three industries share a common reliance on KPIs such as schedule adherence, cost control, quality standards, and throughput levels.

Anomalies in KPI data often serve as early warning signals of underlying risks. A sudden spike in cost variance, an unusual delay in project milestones, or a sharp decline in throughput during peak activity can each indicate systemic inefficiencies. Traditional project management approaches, including Earned Value Management and statistical threshold monitoring, are ill-equipped to capture these irregularities when they manifest in subtle or evolving patterns [4]. Advances in anomaly detection using artificial intelligence and machine learning have introduced powerful new capabilities for identifying irregularities across complex datasets. Yet, most existing studies remain restricted to single industries, limiting the generalizability of their insights. Furthermore, scaling challenges in AI project adoption highlight the critical need for systematic frameworks that connect pilots to enterprise-level performance [5].

The purpose of this paper is to address this gap by proposing a cross-industry AI-enhanced anomaly detection framework that unifies methods across semiconductors, software/IT, and retail (supply chain). To ensure comparability across industries, this study considers the retail sector but focuses specifically on its supply chain operations. Retail supply chains provide quantifiable KPIs that closely parallel the performance dimensions measured in semiconductor and software/IT projects. Thus, the study achieves conceptual alignment in KPI categories and anomaly types, thereby enabling meaningful cross-industry comparison.

The novelty of this study lies in presenting a generalizable KPI-driven methodology that integrates point, contextual, and collective anomaly detection, and in demonstrating its application through Python-based experiments using real-world datasets. By grounding the framework in KPIs and anomalies relevant to all three industries, this research contributes both conceptual clarity and practical tools for improving project performance monitoring.

II. LITERATURE REVIEW

A. Semiconductors

In semiconductor projects, anomaly detection enhances yield management, design validation, and performance monitoring. Studies highlight anomaly detection for identifying inefficiencies in design processes, improving performance management [6], and responding to increasing industry demand [7]. McKinsey reports emphasize that semiconductor project execution suffers from cost overruns and throughput bottlenecks, which anomaly detection can mitigate by detecting deviations early [8].

B. Software/IT

In IT projects, anomaly detection addresses code quality, bug tracking, and performance monitoring. Literature on bug detection using anomaly detection [9], cloud system reliability [10], and project performance metrics [11][12] emphasizes that KPIs such as schedule adherence, defect density, and throughput are particularly prone to anomalies. Contextual anomalies (e.g., spikes in bug counts only in certain release cycles) are especially relevant [13].

C. Retail (Supply Chain)

Supply chains face anomalies in sales, inventory, and logistics. Research demonstrates the role of anomaly detection in improving supply chain data quality [14], detecting sales-inventory deviations [15], and enhancing resilience against disruptions [16]. Collective anomalies, such as simultaneous supplier delays across regions, are particularly important for supply chains [17].

D. Cross-Industry Observations

While semiconductors focus on performance efficiency, software emphasizes defect detection, and supply chains prioritize disruption management, all three share reliance on KPIs such as schedule, cost, quality, and throughput. Literature consistently notes the gap in cross-industry frameworks that integrate these anomaly detection practices [18][19].

III. METHODOLOGY

A. KPI Definition

The framework begins with the definition of KPIs as the foundation for anomaly detection. Four categories - schedule, cost, quality, and throughput -were selected because of their broad applicability across semiconductors, software, and supply chains. Schedule KPIs measure project completion times and milestone adherence, which are critical in industries such as semiconductor design where delays can increase time-to-market risk [6]. Cost KPIs quantify budget variances and overruns, reflecting the financial health of projects [7]. Quality KPIs, including defect density, error rates, and yield, directly assess the reliability of outcomes, whether in chip validation, software testing, or supply chain order fulfillment [8]. Throughput KPIs measure the rate at which tasks or units are completed, which is especially relevant in software development pipelines and high-volume logistics [9]. Secondary KPIs such as resource utilization and customer satisfaction may be integrated into future extensions of the framework for a more holistic assessment.

B. Anomaly Typology

The second component addresses anomaly classification. Point anomalies represent single data points that deviate significantly from expectations, such as one sprint in a software project exhibiting extreme cost escalation. Contextual anomalies are irregularities that make sense only in relation to a particular context - for instance, a drop in throughput during a high-demand period in semiconductor manufacturing or a supply chain. Collective anomalies refer to clusters of data points that individually may appear normal but collectively indicate systemic issues, such as multiple supply chain nodes simultaneously exhibiting minor delivery delays [10]. By explicitly categorizing anomalies into these three types, the framework ensures a comprehensive approach to anomaly detection in project performance monitoring.

C. Data Processing

Project performance datasets across the selected industries are heterogeneous, often containing both numerical and categorical variables. Data processing ensures comparability and prepares datasets for modeling. Normalization techniques standardize KPI values, reducing the impact of differing scales. Dimensionality reduction, primarily via Principal Component Analysis (PCA) uncovers latent structures in the data and reduce noise, thereby improving both computational efficiency and interpretability [11].

D. Cross-Industry Applicability

The pipeline was applied consistently to datasets from semiconductors, software/IT, and retail supply chains. Although the datasets differed in structure and labeling, the preprocessing strategy, i.e., combining imputation, scaling, and one-hot encoding, ensured comparability. For semiconductors, features represented process control signals and validation outcomes. For software/IT,

attributes were linked to module metrics and defect labels. For supply chains, variables captured delivery times, inventory fluctuations, and logistics costs. In supply chains, delivery delays, inventory turnover, and order fulfillment quality are central measures [14]. The common KPI lens of schedule, cost, quality, and throughput allowed anomalies detected in one domain to be conceptually aligned with anomalies in another, making the framework transferable across industries.

E. Practical Adaptation and Robustness

The framework emphasizes practical adaptability to large, heterogeneous datasets. Safeguards such as row-capping, stratified sampling, and rare-category collapsing were built into the implementation to maintain responsiveness without losing critical signal. Percentile-based thresholds were chosen instead of rigid statistical cutoffs to increase robustness in skewed or heavy-tailed distributions. The output included not only performance metrics (F1, precision, recall, separation index) but also artifacts such as histograms, confusion matrices, and PCA scatter plots. These outputs provide project managers with both quantitative evidence and visual interpretability.

IV. EXPERIMENTS

A. Setup

We evaluated the framework on three different datasets: UCI-SECOM [20] for semiconductors, NASA MDP [21] for software/IT, and Online Retail II [22] for retail supply chain.

The experimental pipeline was implemented through a Python script with safeguards to ensure scalability and reproducibility. Row limits were applied (default 150k) to manage large datasets, while categorical features were restricted to the top 50 values, with rare categories collapsed into an "OTHER" bucket to avoid one-hot encoding blow-ups. The script also included automatic label inference to standardize heterogeneous formats (e.g., *pass/fail*, *ok/not ok*, −1/+1) into binary {0,1}. Stratified downsampling was applied where labels existed to preserve rare positive cases. Data preprocessing combined imputation (median for numeric, most-frequent for categorical), normalization, and encoding into a consistent feature space. Thresholds were percentile-based rather than distributional, improving robustness for skewed or heavy-tailed datasets, and progress logging ensured end-to-end completion on commodity hardware.

B. Baselines and Our Framework

For labeled domains (e.g., software/IT defects), the baseline was defined as the best F1 score among the three unsupervised detectors: Principal Component Analysis (PCA) reconstruction error, Multi-Layer Perceptron (MLP) auto-reconstruction, and Isolation Forest (IF), reflecting the approaches typically used when labels are scarce or noisy. The framework's "after" metric was the Random Forest (RF) F1 score, or the unsupervised baseline if RF was not applicable. The F1 score is an evaluation metric that balances a model's precision and recall, providing a single measure of accuracy that accounts for both false positives and false negatives [16].

For unlabeled domains (e.g., retail supply chain data and certain semiconductor signals), performance was measured using a separation index that quantified the distinction between anomalous tails and normal distributions. Here, the baseline was PCA separation, while the framework's "after" metric was the maximum separation achieved by PCA, MLP, or Isolation Forest. This design isolated the contribution of stronger detectors while avoiding label leakage, enabling consistent comparison across labeled and unlabeled settings.

V. RESULTS

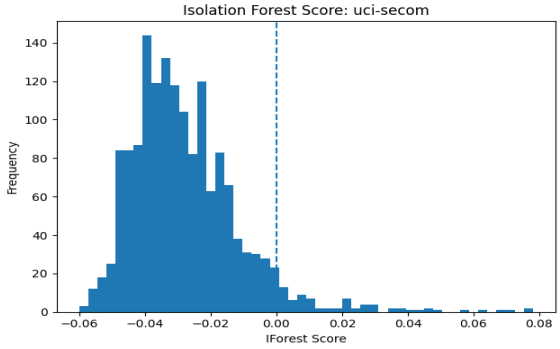

Figure 1 : If Histogram for Semiconductor Dataset

Figure 1 shows an Isolation Forest (IF) histogram for the Semiconductor industry dataset. The anomaly score distribution reveals a long positive 'tail', with the dashed line marking the 95th-percentile threshold. Outliers beyond this cutoff correspond to point anomalies in semiconductor process runs, such as abnormal validation times or sensor deviations. These anomalies map directly to schedule and quality KPIs, highlighting design cycle delays and yield deterioration. The histogram illustrates how the framework provides clear separation between normal runs and candidate anomalies, giving engineers an interpretable view for triage.

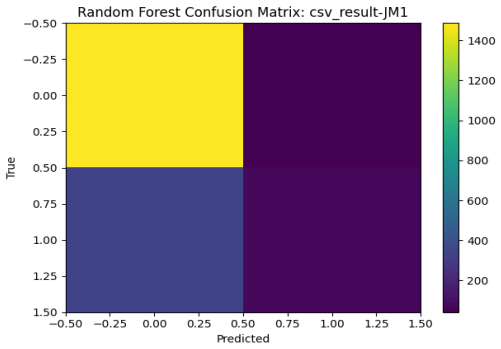

Figure 2 : Software/IT RF Confusion Matrix

Figure 2 shows the Random Forest (RF) confusion matrix for the Software industry dataset. Compared to naive threshold-based baselines, RF achieved a more balanced classification, reducing false positives while capturing contextual anomalies such as defective modules. This improvement is especially valuable when bug counts spike only in specific release cycles, a classic contextual anomaly tied to quality KPIs.

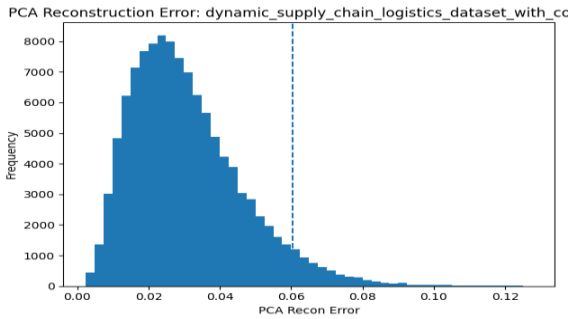

Figure 3 : Retail Supply Chain Histogram (PCA/IF)

Figure 3 shows a histogram (PCA or IF) of the retail supply chain dataset. The right-tail concentration represents operational outliers, such as specific lanes or regions with abnormally long delivery times or inflated logistics costs. These map directly to schedule and cost KPIs, where deviations can cascade into service-level failures. The histogram provides project managers with an interpretable summary of throughput bottlenecks in logistics networks, showing how anomaly detection contributes to proactive supply-chain risk management.

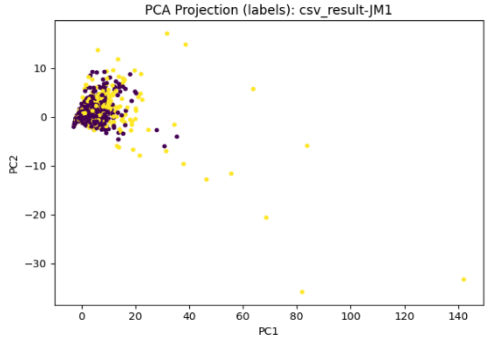

Figure 4 : PCA Labels Scatter

Figure 4 is a 2-D projection of software/IT data, and the PCA scatter plot is colored by defect labels. Defective modules cluster apart from normal ones, demonstrating how collective anomalies emerge as separable groups in feature space. Together, these results highlight how the framework surfaces anomalies linked to both quality and throughput KPIs, giving managers actionable signals for improving release reliability.

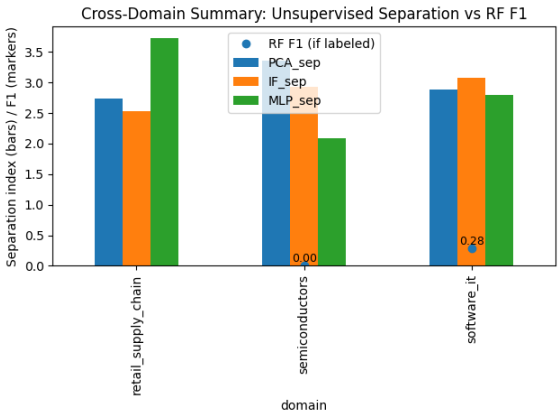

Figure 5 : Consolidated Summary

Figure 5 unifies results across semiconductors, software/IT, and retail. Bars show unsupervised separation indices for PCA, IF, and MLP, while markers overlay RF F1 where labels are available. This universal visualization allows managers to compare performance across industries regardless of label availability. For instance, semiconductors exhibit moderate separation with improvements under IF, while retail demonstrates stronger separation gains. Software/IT shows both separation and higher RF F1 markers, validating the framework's ability to translate anomaly detection into KPI-level insights.

VI. CONCLUSION

This research proposed and evaluated a cross-industry AI-enhanced anomaly detection framework for project performance, focusing on semiconductors, software/IT, and retail (supply chain). By grounding the methodology in four primary KPI categories: schedule, cost, quality, and throughput, the study ensured relevance across industries that differ in operational focus yet share fundamental performance challenges. The incorporation of point, contextual, and collective anomaly types allowed the framework to capture a wide spectrum of irregularities, ranging from isolated spikes to systemic disruptions.

The framework successfully detected point, contextual, and collective anomalies, with supervised models improving contextual detection and hybrid models enhancing collective detection. The use of a separation index ensured that both labeled and unlabeled domains could be consistently compared. By combining detection accuracy with interpretability, the framework equips project managers to anticipate risks, improve decision-making, and strengthen project outcomes across technology-driven industries.

The novelty of this study lies in its ability to unify anomaly detection across multiple industries through a generalizable, KPI-driven methodology. Unlike prior domain-specific studies, the framework demonstrated transferability and adaptability, showing that project anomaly detection can be systematically applied beyond individual sectors. By addressing both detection accuracy and interpretability, the framework provides a practical tool for improving decision-making and reducing project risks in complex, technology-driven environments.

VII. FUTURE WORK

While our datasets are real and representative, they remain proxies for broader portfolios. We do not yet model temporal autocorrelation or seasonality explicitly, which are important for retail peaks and semiconductor ramps, and we have not embedded explainability (e.g., SHAP) in this code path, though our prior conceptualization emphasizes its managerial value. Future work will: (i) incorporate time-aware models (e.g., rolling PCA/IF; sequence autoencoders); (ii) add explainability at scale (global + local attributions per anomaly); (iii) extend to streaming detection with adaptive thresholds; and (iv) integrate feedback loops so corrective actions recalibrate models continuously.

VIII. REFERENCES

[1] G. Batra, Z. Jacobson, and N. Santhanam, “Improving the semiconductor industry through advanced analytics,” McKinsey & Company, Mar. 2016. [Online]. Available: https://www.mckinsey.com/industries/semiconductors/our-insights/improving-the-semiconductor-industry-through-advanced-analytics.

[2] K. P. Pham and M. Neumann, "How to Measure Performance in Agile Software Development? A Mixed-Method Study," in 2024 50th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), 2024, pp. 443-450, doi: 10.1109/SEAA64295.2024.00074.

[3] M. L. Fisher, J. Krishnan, and S. Netessine, “Retail Store Execution: An Empirical Study,” Operations and Information Management Department, The Wharton School, Univ. of Pennsylvania, Philadelphia, PA, USA, Working Paper, Dec. 2006.

[4] H. Kerzner, Project Management Metrics, KPIs, and Dashboards: A Guide to Measuring and Monitoring Project Performance, 2nd ed. Wiley, 2017.

[5] S. Makinani and M. B. Nagaraja, “Scaling AI from Project Pilots to Program-Wide Transformations,” International Journal of Emerging Research in Engineering and Technology, vol. 6, no. 3, pp. 41-46, Jul. 2025, doi: 10.63282/3050-922X.IJERET-V6I3P105.

[6] N. Hinrichs and E. Barke, "Applying performance management on semiconductor design processes," in Proceedings of the IEEE International Engineering Management Conference (IEMC), Austin, TX, USA, 2008, pp. 278-281.

[7] R. Palma, R. Varadarajan, J. Goodrich, T. Lopez, and A. Patil, The Growing Challenge of Semiconductor Design Leadership, Boston Consulting Group and Semiconductor Industry Association, 2022.

[8] O. Burkacky and B. Wiseman, McKinsey on Semiconductors: Creating value, pursuing innovation, and optimizing operations, McKinsey & Company, Mar. 2024.

[9] S. Hangal and M. S. Lam, "Tracking Down Software Bugs Using Automatic Anomaly Detection," in Proceedings of the 24th International Conference on Software Engineering (ICSE), Orlando, FL, USA, May 25, 2002, pp. 291-301, doi: 10.1145/581376.581377.

[10] M. S. Islam, M. S. Rakha, W. Pourmajidi, J. Sivaloganathan, J. Steinbacher, and A. Miranskyy, "Anomaly Detection in Large-Scale Cloud Systems: An Industry Case and Dataset," arXiv preprint arXiv:2411.09047, 2025.

[11] J. S. Osmundson, J. B. Michael, M. J. Machniak, and M. A. Grossman, "Quality management metrics for software development," Information Management, vol. 40, no. 10, pp. 799-812, 2003, doi: 10.1016/S0378-7206(02)00114-3.

[12] C. M. Okafor, O. Ogunse, M. N. Iheke, D. O. Oseghale, I. Ogiehor, and E. E. Aniebonam, "The Role of Advanced Anomaly Detection in Transforming Program Management in Government with Scikit-Learn, A Machine Learning Library in Python," Int. J. Res. Sci. Innov., vol. XII, no. V, pp. 148-165, May 2025, doi: 10.51244/IJRSI.2025.120500014.

[13] A. B. Nassif, M. Abu Talib, Q. Nasir, and F. Dakalbab, "Machine Learning for Anomaly Detection: A Systematic Review," IEEE Access, 2017, doi: 10.51244/ijrsi.2025.120500014.

[14] A. E. Glaser, J. P. Harrison, and D. Josephs, "Anomaly Detection Methods to Improve Supply Chain Data Quality and Operations," SMU Data Science Review, vol. 6, no. 1, 2022. Available: https://scholar.smu.edu/datasciencereview/vol6/iss1/3.

[15] G. Makhmetova, "Automated Anomaly Detection for Sales and Inventory in Data-Driven Industries," Universal Library of Engineering Technology, vol. 2, no. 1, pp. 07-14, 2025, doi: 10.70315/uloap.ulete.2025.0201002.

[16] S. R., G. Mariammal, N. Aggarwal, M. Jayanthi, A. Worku, and K. V. S. Praveena, "Leveraging Adaptive Supply Chain Management with Incremental Learning and Anomaly Detection," 2025. [Online]. Available: https://ssrn.com/abstract=5083793.

[17] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly Detection: A Survey," ACM Computing Surveys, vol. 41, no. 3, pp. 1-58, Sep. 2009.

[18] M. Steidl, M. Leitner, P. Urbanke, M. Gattringer, M. Felderer, and S. Ristov, "How Industry Tackles Anomalies during Runtime: Approaches and Key Monitoring Parameters," in Proceedings of the 50th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), Paris, France, 28-30 August 2024, doi: 10.1109/SEAA64295.2024.00062.

[19] C. Das, A. Rasool, A. Dubey, and N. Khare, "Analyzing the Performance of Anomaly Detection Algorithms," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 12, no. 6, pp. 439-445, 2021.

[20] M. McCann and A. Johnston. "SECOM," UCI Machine Learning Repository, 2008. [Online]. Available: https://doi.org/10.24432/C54305.

[21] M. Shepperd, Q. Song, Z. Sun, and C. Mair (2018). NASA MDP Software Defects Data Sets. figshare. Collection. [Online]. Available: https://doi.org/10.6084/m9.figshare.c.4054940.v1.

[22] D. Chen. "Online Retail II," UCI Machine Learning Repository, 2012. [Online]. Available: https://doi.org/10.24432/C5CG6D.