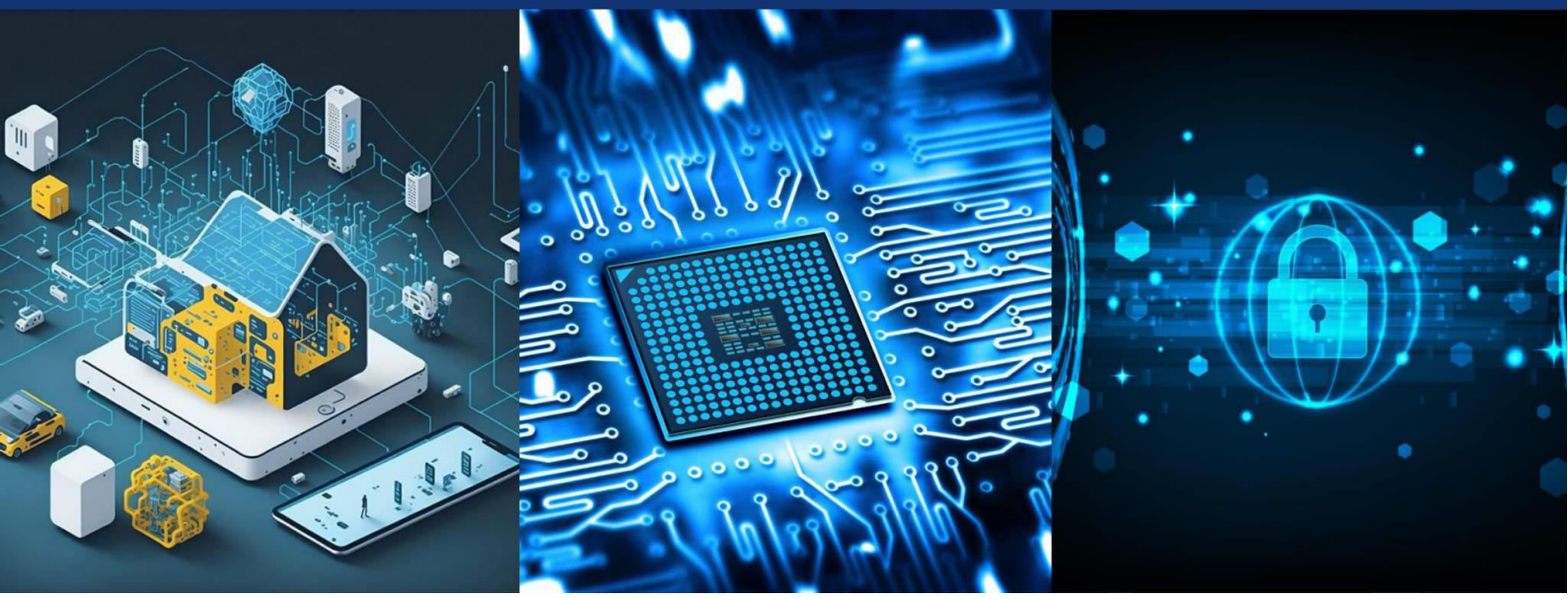
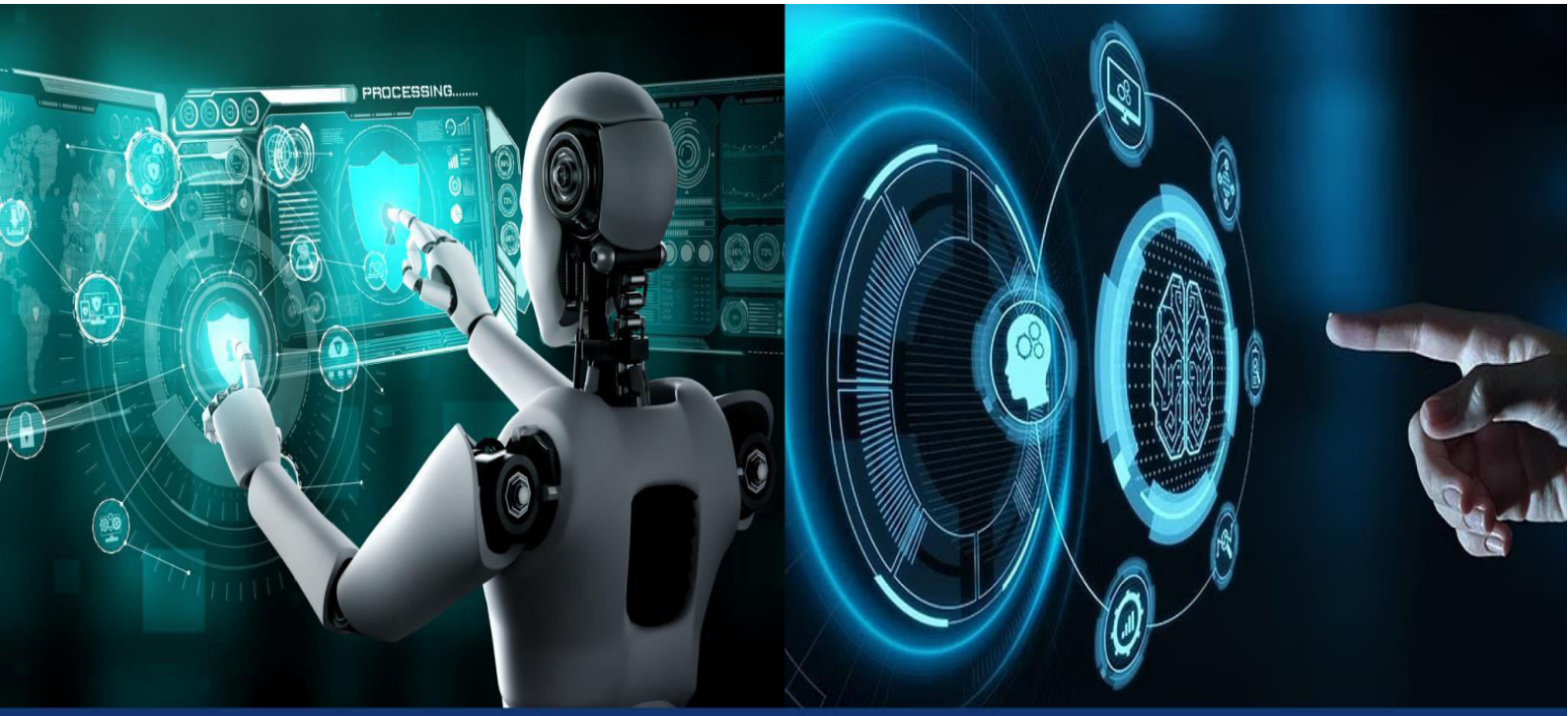
ISSN(O): 2320-9801
ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.771 **Volume 13, Issue 7, July 2025**

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Machine Learning-Based Anomaly Detection for Inbound Payroll and Benefits Data Feeds

Ujwal Dyasani

Lead Software Engineer Integrations, USA

ABSTRACT: Inbound payroll and benefits data feeds carry some of the highest-stakes data in an enterprise human capital management environment: an undetected anomaly, whether a data entry error, a systemic feed malfunction, or fraudulent manipulation, can directly translate into an incorrect paycheck, an improper benefits deduction, or a compliance exposure before any human reviewer has the opportunity to intervene. Static, threshold-based validation, while necessary, is structurally limited to detecting anomalies that violate a predefined rule, leaving a substantial category of subtler, statistically unusual but rule-compliant anomalies undetected. This article develops an ensemble machine learning framework for anomaly detection in inbound payroll and benefits data feeds, combining three complementary detection methods, statistical process control, isolation forest models, and autoencoder neural network reconstruction error, into a unified anomaly scoring engine. Drawing on established statistical process control theory, unsupervised anomaly detection literature, and neural network reconstruction-based detection methods, the article presents a radial pipeline architecture, an illustrative three-detector ensemble agreement analysis, a calendar-style anomaly rate heatmap across pay groups and pay periods, and a severity-based alert escalation and routing framework. The findings indicate that an ensemble combining these three detection methods achieves substantially higher illustrative precision and recall than any single method alone, particularly for anomalies that are individually subtle across any one detection lens but jointly identifiable when multiple lenses agree. The article further addresses severity classification, alert routing to appropriate payroll, benefits, and compliance stakeholders, and governance practices for sustaining detection accuracy as pay populations and benefit plan structures evolve over time.

KEYWORDS: anomaly detection; machine learning; payroll data integrity; benefits administration; statistical process control; isolation forest; autoencoder; ensemble detection; data quality; human capital management.

I. INTRODUCTION

Payroll and benefits data occupies a uniquely consequential position within enterprise human capital management: unlike most operational data categories, an undetected error in a payroll or benefits feed does not merely create downstream reporting inconvenience, it can directly produce an incorrect payment to an employee, an improper benefits deduction, or a compliance filing inaccuracy, often before any human has the opportunity to review the affected record. Organizations have long relied on threshold-based validation rules, confirming that a compensation value falls within an approved range, or that a benefits election matches an eligible plan option, to guard against the most obvious classes of error. These rules remain necessary, but they share a structural limitation: they can only detect violations of a rule that has been explicitly predefined, leaving undetected any anomaly that is statistically unusual without technically violating a known rule (Chandola, Banerjee, and Kumar, 2009).

Machine learning-based anomaly detection offers a complementary detection paradigm, identifying records or patterns that deviate from learned normal behavior even when no explicit rule has been violated. This article develops an ensemble framework combining three distinct detection methodologies, statistical process control, isolation forest models, and autoencoder neural network reconstruction error, each bringing a different mathematical lens to the anomaly detection problem, into a unified scoring system for inbound payroll and benefits data feeds. The research questions guiding this study are as follows.

- RQ1: What complementary strengths and blind spots characterize statistical process control, isolation forest, and autoencoder reconstruction detection methods when applied to payroll and benefits data specifically?
- RQ2: How should an ensemble scoring architecture combine these three methods to improve illustrative detection performance relative to any single method alone?
- RQ3: What anomaly rate patterns emerge across different pay groups and pay periods, and what do these patterns imply for detection tuning priorities?

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- RQ4: What severity classification and escalation routing framework appropriately directs detected anomalies to the correct payroll, benefits, and compliance stakeholders?

The remainder of this article proceeds as follows. Section 2 reviews related literature on anomaly detection methodology. Section 3 examines the specific limitations of threshold-based payroll validation. Section 4 details the research methodology. Sections 5 through 9 present the core detection framework, including the radial pipeline architecture, the three detection layers individually, and ensemble score fusion. Sections 10 through 13 present illustrative pattern and performance analyses. Sections 14 and 15 address governance and pitfalls. Sections 16 through 19 discuss limitations, recommendations, future research, and conclusions.

II. BACKGROUND AND RELATED WORK

2.1 Statistical Process Control Theory

Statistical process control, originating in manufacturing quality management, establishes control limits around an expected process distribution, typically the trailing mean plus or minus a multiple of the standard deviation, and flags observations falling outside those limits as candidates for investigation (Montgomery, 2019). This technique transfers naturally to monitoring aggregate payroll metrics, such as total gross pay per pay period, where a sudden, statistically improbable deviation from the established trailing pattern is itself informative regardless of whether any individual record violates a specific business rule.

2.2 Unsupervised Anomaly Detection and Isolation Forests

The isolation forest algorithm, introduced by Liu, Ting, and Zhou, provides an efficient unsupervised method for identifying anomalous records within a multivariate dataset, based on the insight that anomalous points are, on average, easier to isolate through random recursive partitioning than typical points, requiring fewer partitioning steps to separate from the rest of the data (Liu, Ting, and Zhou, 2008). This method's principal advantage for payroll data is its ability to detect anomalies defined by unusual combinations of multiple fields simultaneously, rather than requiring each field to be evaluated independently against its own threshold.

2.3 Autoencoder Neural Networks for Anomaly Detection

Autoencoder neural networks, trained to reconstruct their own input after passing it through a reduced-dimensionality bottleneck layer, provide a third detection lens: because the network is trained predominantly on normal data, it reconstructs normal records with low error, while anomalous records, which do not conform to the patterns the network learned, produce elevated reconstruction error, a signal directly usable as an anomaly score (Goodfellow, Bengio, and Courville, 2016; Sakurada and Yairi, 2014).

2.4 Ensemble Methods in Anomaly Detection

Anomaly detection literature has increasingly emphasized ensemble approaches, combining multiple detection methods to improve robustness, on the basis that different detection methods exhibit different, only partially overlapping blind spots, such that requiring agreement across multiple methods, or fusing their scores into a composite measure, generally improves detection reliability relative to any single method applied alone (Aggarwal, 2017; Chandola, Banerjee, and Kumar, 2009).

2.5 Positioning of This Study

While statistical process control theory, isolation forest methodology, autoencoder-based detection, and general ensemble anomaly detection literature each address components of the problem independently, comparatively little published material synthesizes these three specific methods into a unified framework tailored to the payroll and benefits data context specifically. This article's contribution is that synthesis, together with illustrative comparative performance analysis and a governance framework.

III. THE LIMITS OF THRESHOLD-BASED PAYROLL VALIDATION

Before presenting the ensemble framework, it is useful to characterize specifically which anomaly categories threshold-based validation rules can and cannot detect.

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Table.1. Anomaly categories and their detectability by static threshold-based validation rules

Anomaly Category	Detectable by Static Threshold Rules?	Illustrative Example
Out-of-Range Field Value	Yes	A compensation value exceeding an absolute maximum defined by policy
Missing Required Field	Yes	A benefits enrollment record missing a required plan code
Unusual Multi-Field Combination	No, generally undetected	A compensation value within range individually, but unusual given the specific job profile and tenure combination
Statistically Improbable Aggregate Shift	No, generally undetected	A pay period's total gross pay deviating sharply from the established trailing pattern without any single record violating a threshold
Subtle Multivariate Drift	No, generally undetected	A gradual shift in the joint distribution of several fields that individually remain within normal range

The three undetected categories in Table 1 directly motivate the three machine learning detection layers developed in Sections 6 through 8: statistical process control targets aggregate shifts, isolation forests target unusual multi-field combinations, and autoencoder reconstruction targets subtle multivariate drift, illustrating that the choice of these three specific methods is directly responsive to the specific detection gap threshold rules leave open.

IV. RESEARCH METHODOLOGY

This study employs a design framework and illustrative comparative analysis methodology, synthesizing statistical process control theory (Montgomery, 2019), isolation forest methodology (Liu, Ting, and Zhou, 2008), autoencoder-based anomaly detection literature (Sakurada and Yairi, 2014; Goodfellow, Bengio, and Courville, 2016), and ensemble anomaly detection principles (Aggarwal, 2017) into the radial pipeline architecture presented in Section 5. The comparative agreement, pattern, and performance figures presented throughout Sections 9 through 13 use illustrative figures constructed to demonstrate expected directional relationships consistent with the underlying literature, rather than measurements audited against a specific production payroll dataset.

Table.2. Summary of research methodology components

Methodology Component	Method	Purpose
Threshold Limitation Analysis	Synthesis of anomaly detection taxonomy literature	Establish specific detection gaps the framework should target (Section 3)
Ensemble Architecture Development	Synthesis of SPC, isolation forest, and autoencoder methodology	Establish the three-layer detection framework (Sections 6-8)
Comparative Illustration Construction	Directional estimate construction consistent with established literature	Illustrate expected agreement, pattern, and performance differences
Governance Framework Development	Synthesis of model maintenance and alert management practice	Establish sustained detection accuracy practices (Section 14)

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. RADIAL ANOMALY DETECTION PIPELINE ARCHITECTURE

This article proposes a radial pipeline architecture centered on a unified anomaly score engine, with ingestion, standardization, three parallel detection layers, ensemble fusion, severity classification, alert routing, and an analyst feedback loop arranged as surrounding nodes.

FIGURE 1: RADIAL ANOMALY DETECTION PIPELINE ARCHITECTURE

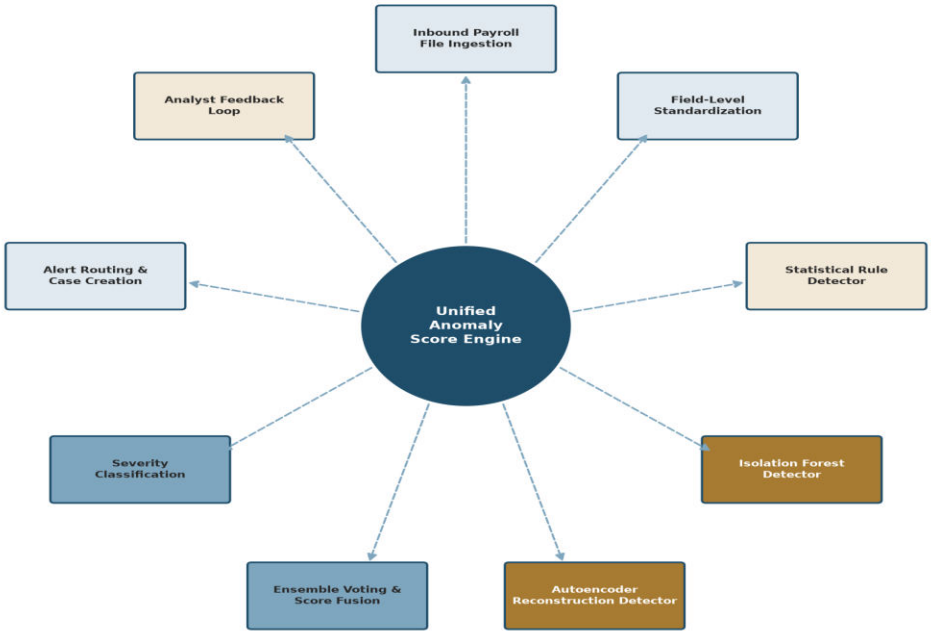


Figure.1. Radial anomaly detection pipeline architecture, centered on a unified anomaly score engine coordinating ingestion, three parallel detection layers, ensemble fusion, and alert routing.

Table.3. Radial pipeline nodes, primary function, and downstream connections

Pipeline Node	Primary Function	Feeds Into
Inbound Payroll File Ingestion	Receive raw payroll and benefits feed files from source systems	Field-Level Standardization
Field-Level Standardization	Normalize field formats and units before detection	All three detection layers
Statistical Rule Detector	Apply SPC-based aggregate and field-level control limits	Ensemble Voting and Score Fusion
Isolation Forest Detector	Score records for multivariate outlier characteristics	Ensemble Voting and Score Fusion
Autoencoder Reconstruction Detector	Score records by reconstruction error against learned normal patterns	Ensemble Voting and Score Fusion
Ensemble Voting and Score Fusion	Combine the three detector outputs into a unified anomaly score	Severity Classification
Severity Classification	Assign a severity tier to each	Alert Routing and Case Creation

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

	flagged anomaly	
Alert Routing and Case Creation	Route flagged anomalies to the appropriate stakeholder team	Analyst Feedback Loop
Analyst Feedback Loop	Capture reviewer determinations to inform ongoing model tuning	Field-Level Standardization (retraining)

The analyst feedback loop's connection back into the standardization stage, rather than terminating the pipeline at alert routing, reflects a deliberate design choice: reviewer determinations about whether a flagged anomaly was genuine or a false positive constitute valuable labeled data for ongoing model retraining, a governance practice addressed further in Section 14.

VI. STATISTICAL PROCESS CONTROL DETECTION LAYER

The statistical process control layer establishes control limits around trailing aggregate and field-level metrics, flagging observations that fall outside those limits as candidates for investigation.

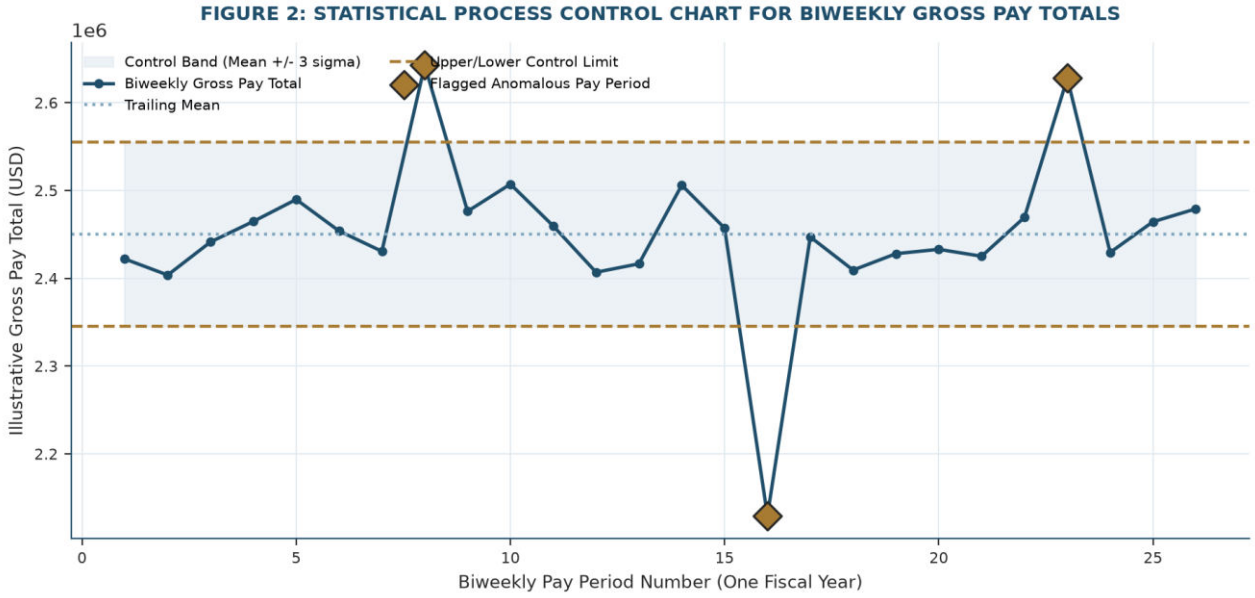


Figure.2. Statistical process control chart for biweekly gross pay totals across one fiscal year, with control bands and three flagged anomalous pay periods.

Table.4. Representative statistical process control monitored metrics and detected anomaly types.

Monitored Metric	Illustrative Control Limit Basis	Representative Detected Anomaly
Total Gross Pay per Pay Period	Trailing 12-period mean plus/minus 3 standard deviations	An unexpected spike from a bulk retroactive adjustment batch
Record Count per Pay Group per Period	Trailing mean plus/minus 3 standard deviations	A sharp drop suggesting a partial file transmission failure
Average Deduction Amount per Benefit Plan	Trailing mean plus/minus 3 standard deviations	A systemic deduction calculation error following a plan configuration change
New Enrollment Count per Period	Trailing mean plus/minus 3 standard deviations	An unusually high enrollment spike suggesting a duplicate file submission

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The SPC layer's principal strength, as illustrated in Figure 2, is its interpretability: a flagged anomaly can be explained directly in terms of how far a specific metric deviated from its established trailing pattern, which is particularly valuable when communicating a flagged issue to a non-technical payroll or benefits stakeholder during the escalation process described in Section 12.

VII. ISOLATION FOREST DETECTION LAYER

The isolation forest layer scores individual records for multivariate outlier characteristics, identifying records whose combination of field values is unusual even when no individual field value would itself trigger a threshold rule or an SPC control limit breach.

Table.5. Isolation forest feature categories and representative multivariate patterns detected.

Feature Category	Representative Fields	Multivariate Pattern Detected
Compensation Structure	Base pay, bonus amount, overtime hours, pay rate change frequency	An unusual combination of high overtime hours paired with a recent pay rate change
Benefits Enrollment Pattern	Plan type, dependent count, election change frequency, enrollment date proximity to hire date	A benefits election pattern inconsistent with the employee's tenure and family status profile
Organizational Context	Job profile, department, location, supervisory organization	A compensation value that is unremarkable in isolation but unusual for the specific job profile and location combination
Temporal Pattern	Days since last pay rate change, days since last benefits election change	A cluster of near-simultaneous changes across multiple fields, unusual relative to typical change patterns

The isolation forest layer's contribution is most distinct from the SPC layer specifically in the organizational context and temporal pattern feature categories in Table 5, where the anomaly is defined by an unusual combination across several fields jointly, a pattern SPC's single-metric control-limit approach is not designed to capture.

VIII. AUTOENCODER RECONSTRUCTION DETECTION LAYER

The autoencoder layer trains a neural network to reconstruct normal payroll and benefits records after compressing them through a reduced-dimensionality bottleneck, using elevated reconstruction error at scoring time as the anomaly signal.

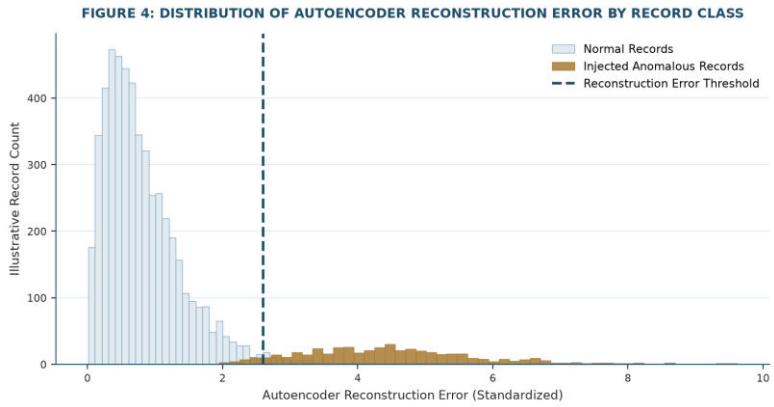


Figure.3. Distribution of autoencoder reconstruction error by record class, showing separation between normal records and injected anomalous records relative to the detection threshold

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Table.6. Autoencoder design considerations, recommended approaches, and rationale

Autoencoder Design Consideration	Recommended Approach	Rationale
Training Data Composition	Train predominantly on records confirmed normal through prior review cycles	Ensures the network learns genuinely normal patterns rather than patterns contaminated by undetected anomalies
Bottleneck Dimensionality	Sufficiently reduced to force genuine pattern learning, not so reduced that normal variation is lost	Balances reconstruction fidelity for normal records against sensitivity to anomalous records
Reconstruction Error Threshold	Calibrated against the separation point between normal and known anomalous error distributions, illustrated in Figure 3	Directly determines the precision/recall trade-off for this detection layer
Retraining Cadence	Periodic retraining incorporating recently confirmed normal records	Prevents the network's definition of normal from becoming stale as legitimate patterns evolve

The distributions illustrated in Figure 3 show meaningful but incomplete separation between normal and anomalous records, a realistic characteristic of reconstruction-based detection: some anomalous records produce only modestly elevated reconstruction error, falling near the threshold, which is precisely the population of cases where ensemble agreement with the other two detection layers, discussed in Section 9, becomes most valuable.

IX. ENSEMBLE AGREEMENT AND SCORE FUSION

This section examines how the three detection layers agree and disagree across a representative flagged population, illustrating the ensemble agreement structure that the score fusion stage operates upon.

FIGURE 3: THREE-DETECTOR ENSEMBLE AGREEMENT (ILLUSTRATIVE OVERLAP COUNTS)

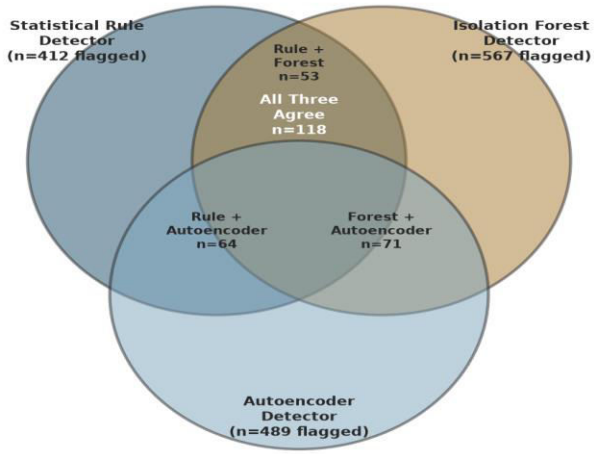


Figure.4. Three-detector ensemble agreement, showing illustrative overlap counts among the statistical rule detector, isolation forest detector, and autoencoder detector

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Table.7. Ensemble agreement patterns, illustrative counts, and recommended score fusion treatment.

Agreement Pattern	Illustrative Count	Recommended Fusion Treatment
All Three Detectors Agree	118	Highest-confidence tier; route directly to high or critical severity
Exactly Two Detectors Agree	188 (sum of three pairwise-only overlaps)	Moderate-confidence tier; route to standard review queue
Exactly One Detector Flags	Remainder of each detector's total (approximately 1,062 combined)	Lower-confidence tier; apply weighted scoring rather than automatic escalation

The score fusion approach implied by Table 7 treats detector agreement as itself informative, consistent with the general ensemble anomaly detection principle that agreement across methodologically distinct detectors substantially reduces the likelihood that a flagged anomaly reflects a false positive specific to one detector's particular blind spot (Aggarwal, 2017). Records flagged by only one detector are not discarded outright, but are weighted according to that detector's individual precision characteristics rather than automatically escalated to a high severity tier.

X. ANOMALY RATE PATTERNS ACROSS PAY GROUPS AND PERIODS

This section presents an illustrative calendar-style view of anomaly rate across seven pay groups and twenty-four pay periods spanning one fiscal year, surfacing patterns relevant to detection tuning priorities.

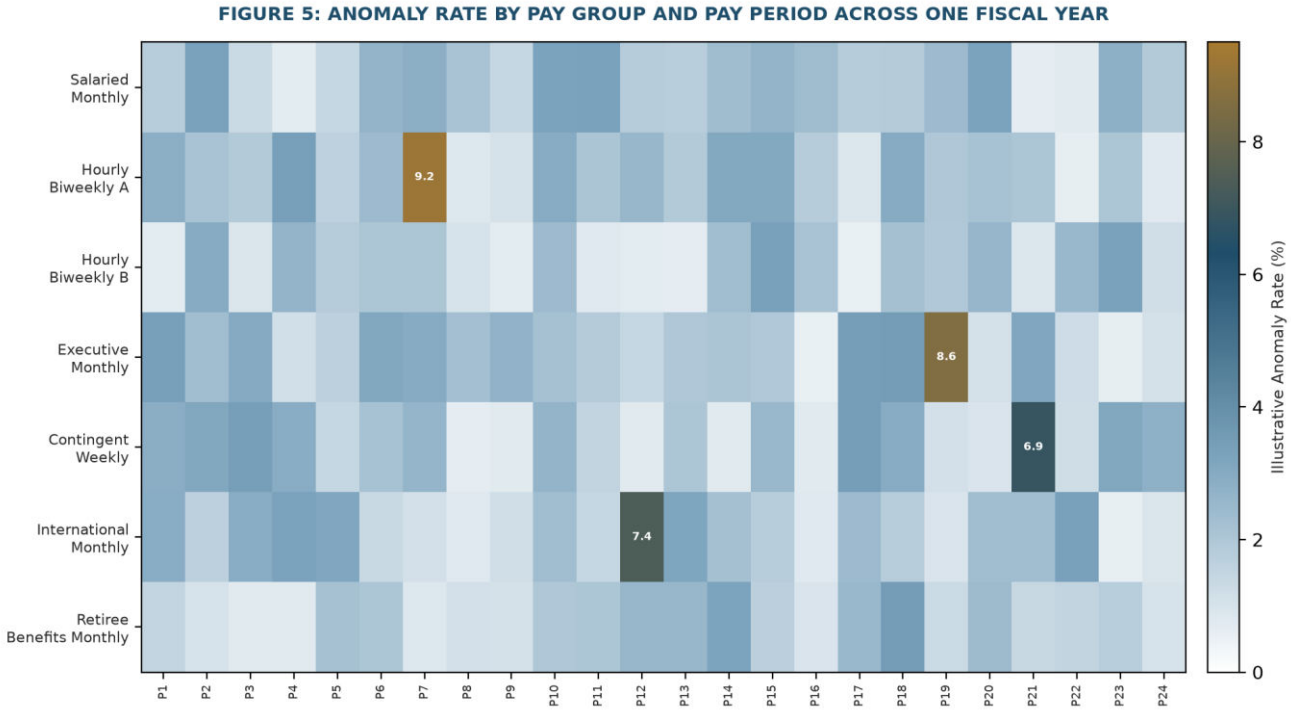


Figure.5. Anomaly rate by pay group and pay period across one fiscal year, with several concentrated hot spots corresponding to specific pay group and period combinations.

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Table.8. Illustrative anomaly hot spots identified from the calendar-style heatmap and their likely root causes.

Hot Spot	Pay Group	Approximate Period	Illustrative Likely Root Cause
Hot Spot 1	Hourly Biweekly A	Period 7	Bulk retroactive wage adjustment following a rate change
Hot Spot 2	Executive Monthly	Period 19	Annual equity compensation vesting event
Hot Spot 3	International Monthly	Period 12	Currency conversion rate update affecting multiple records simultaneously
Hot Spot 4	Contingent Weekly	Period 21	Seasonal contingent workforce ramp-up period

The illustrative hot spots documented in Table 8 share a common characteristic worth noting for detection tuning: each corresponds to a legitimate, expected business event, a rate adjustment, a vesting event, a currency update, a seasonal ramp-up, rather than a genuine data quality problem. This underscores the importance of incorporating known business calendar events into the detection framework's context, so that expected periodic events do not repeatedly trigger false-positive alerts, a governance consideration addressed further in Section 14.

XI. ALERT SEVERITY CLASSIFICATION AND VOLUME TRENDS

This section presents an illustrative twelve-month view of alert volume by severity tier, demonstrating how the ensemble framework's output translates into an operational alert stream over time.

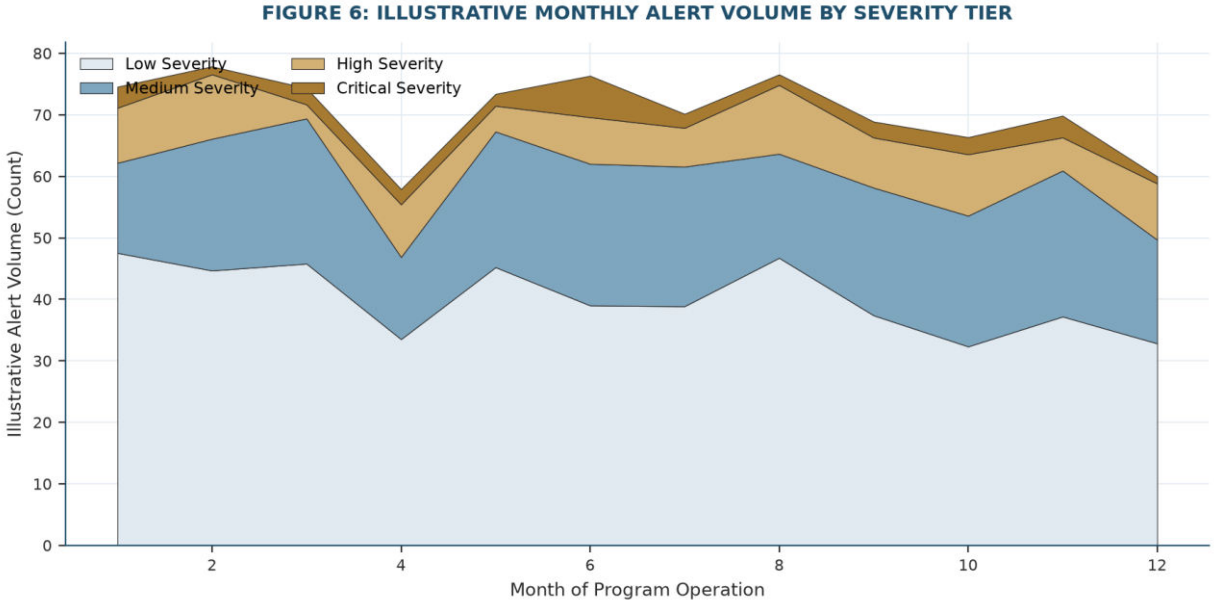


Figure.6. Illustrative monthly alert volume by severity tier over a twelve-month period, including a critical-severity spike in month six.

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Table.9. Alert severity tier definitions and typical response expectations

Severity Tier	Illustrative Definition	Typical Response Expectation
Low Severity	Single detector flag, low ensemble confidence, limited potential financial impact	Reviewed in routine batch cycle, no immediate action required
Medium Severity	Two-detector agreement or moderate potential financial impact	Reviewed within one business day
High Severity	Three-detector agreement or significant potential financial impact	Reviewed within four business hours
Critical Severity	Three-detector agreement combined with regulatory or large-scale financial exposure	Immediate review; may trigger the cross-functional incident review described in Section 12

The critical severity spike visible in month six of Figure 6 illustrates the kind of event the cross-functional incident review pathway, detailed in Section 12, is specifically designed to handle: a sudden, concentrated cluster of high-confidence anomalies warranting coordinated attention across payroll, benefits, and compliance stakeholders simultaneously, rather than routine, independent handling by a single team.

XII. ALERT ESCALATION AND OWNERSHIP ROUTING

This section presents the escalation and ownership routing framework directing flagged anomalies to the appropriate stakeholder team based on anomaly category and severity.

FIGURE 7: ALERT ESCALATION AND OWNERSHIP ROUTING TREE

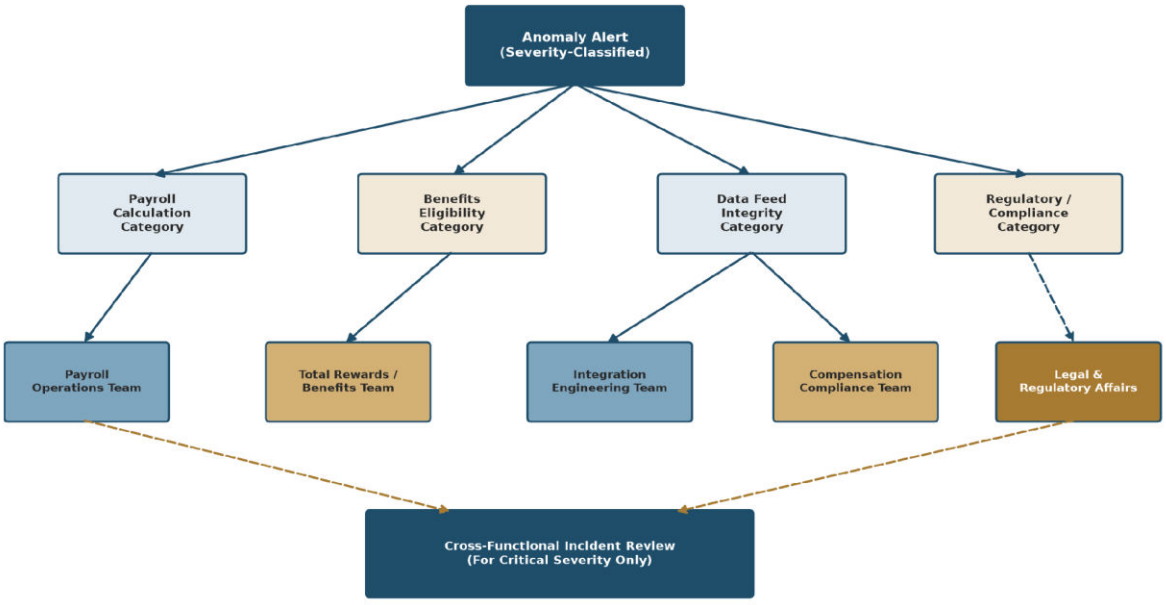


Figure.7. Alert escalation and ownership routing tree, spanning four anomaly categories, five owning teams, and a cross-functional incident review pathway for critical severity alerts

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Table.10. Anomaly category, primary owning team, and cross-functional escalation triggers.

Anomaly Category	Primary Owning Team	Escalation Trigger to Cross-Functional Review
Payroll Calculation Category	Payroll Operations Team	Critical severity with confirmed multi-employee financial impact
Benefits Eligibility Category	Total Rewards / Benefits Team	Critical severity with confirmed plan-wide eligibility miscalculation
Data Feed Integrity Category	Integration Engineering Team	Critical severity indicating a systemic upstream feed failure
Regulatory / Compliance Category	Compensation Compliance Team and Legal and Regulatory Affairs	Any critical severity alert in this category, given inherent regulatory exposure

The regulatory and compliance category in Table 10 is treated distinctly from the other three categories: because any critical severity alert in this category carries inherent regulatory exposure regardless of confirmed financial impact, it triggers the cross-functional review pathway automatically, rather than requiring confirmed impact as a precondition, reflecting the asymmetric cost of delayed regulatory response relative to the cost of an unnecessary review meeting.

XIII. COMPARATIVE DETECTION PERFORMANCE ACROSS METHODS

This section presents an illustrative comparative precision and recall analysis across five detection approaches, from static threshold rules through the full three-detector ensemble.

FIGURE 8: ILLUSTRATIVE PRECISION AND RECALL BY DETECTION METHOD

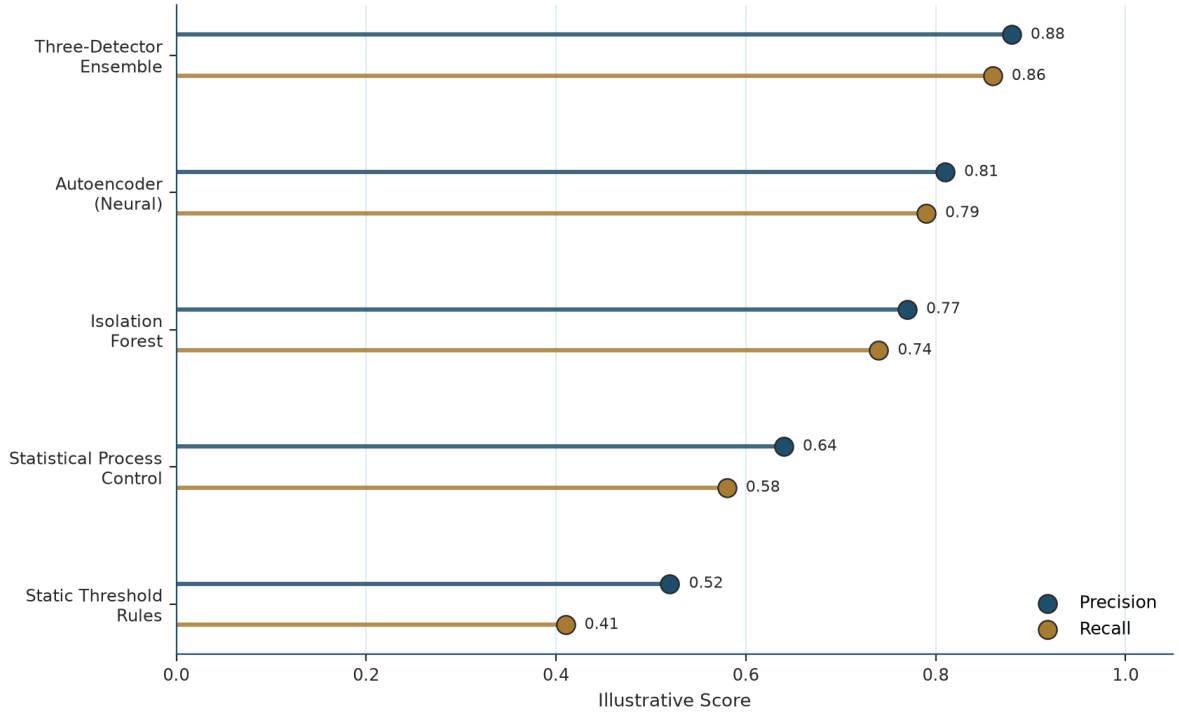


Figure .8.Illustrative precision and recall by detection method, showing progressive improvement from static threshold rules through the three-detector ensemble.

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Table.11. Illustrative precision, recall, and F1 score by detection method.

Detection Method	Illustrative Precision	Illustrative Recall	Illustrative F1 Score
Static Threshold Rules	0.52	0.41	0.46
Statistical Process Control	0.64	0.58	0.61
Isolation Forest	0.77	0.74	0.75
Autoencoder (Neural)	0.81	0.79	0.80
Three-Detector Ensemble	0.88	0.86	0.87

The progressive improvement documented in Table 11, from static threshold rules through the full ensemble, is directionally consistent with the anomaly category analysis in Section 3: each additional method addresses a category of anomaly the preceding methods could not, and the ensemble's combined performance exceeds any individual method precisely because it inherits the detection coverage of all three underlying approaches simultaneously.

XIV. GOVERNANCE FOR SUSTAINED DETECTION ACCURACY

Table.12. Recommended governance practices for sustaining detection accuracy after initial deployment.

Governance Practice	Purpose	Recommended Cadence
Business Calendar Integration	Incorporate known events (vesting, rate changes, seasonal ramp-up) into detection context, per Section 10	Updated per business calendar cycle
SPC Control Limit Refresh	Update trailing mean and standard deviation baselines to reflect legitimate pattern changes	Quarterly, or upon known structural change
Isolation Forest and Autoencoder Retraining	Retrain both models on recently confirmed normal data, incorporating analyst feedback loop determinations	Quarterly at minimum; triggered earlier upon detected drift
Ensemble Weight Recalibration	Adjust fusion weighting if one detector's relative precision shifts materially over time	Semi-annual review
Escalation Routing Accuracy Audit	Confirm alerts are reaching the correct owning team per the routing framework in Section 12	Quarterly

XV. COMMON PITFALLS IN PAYROLL ANOMALY DETECTION

Table.13. Recurring pitfalls in payroll and benefits anomaly detection and recommended corrections.

Pitfall	Description	Recommended Correction
Relying on a Single Detection Method	Organization deploys only threshold rules or only one machine learning method	Implement the full three-detector ensemble per Sections 6-9
Ignoring Known Business Calendar Events	Legitimate periodic events repeatedly trigger false-positive alerts	Integrate business calendar context per Section 10 and Section 14
Static Detection Models	Models are trained once and never revisited as pay populations evolve	Apply the retraining governance in Section 14
Uniform Escalation Regardless of Category	All alerts routed to a single generic queue regardless of anomaly category	Apply the category-specific routing framework in Section 12
Treating Single-Detector Flags as Equally Confident as Ensemble Agreement	All flagged records treated with equal urgency regardless of agreement pattern	Apply the tiered fusion treatment in Section 9

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

XVI. LIMITATIONS OF THE STUDY

Several limitations should be considered when interpreting this article's findings. First, the comparative agreement, pattern, and performance figures presented throughout Sections 9 through 13 are illustrative estimates constructed to demonstrate expected directional relationships consistent with established anomaly detection literature, rather than measurements audited against a specific production payroll dataset. Second, this study assumes sufficient historical data volume to train the isolation forest and autoencoder models reliably, a condition that may not hold for smaller organizations with limited payroll population size. Third, the severity classification and routing framework in Sections 11 and 12 reflects a generalized structure that individual organizations should adapt to their specific organizational structure and risk tolerance rather than adopting verbatim. Fourth, this article does not address the specific regulatory and data privacy considerations involved in processing payroll data for model training in depth, focusing instead on the detection methodology itself.

Table.14. Summary of study limitations and suggested mitigations for future research

Limitation	Potential Effect on Findings	Suggested Mitigation for Future Studies
Illustrative, not audited, comparative figures	Absolute figures may not generalize across organizations or datasets	Controlled empirical study across multiple production payroll environments
Assumes sufficient historical training data	Framework may not be immediately applicable to smaller payroll populations	Extension study addressing cold-start approaches for limited historical data
Generalized severity and routing structure	Organization-specific adaptation required before direct application	Comparative study across multiple organizational routing structures
Limited treatment of data privacy considerations	Regulatory and privacy implications not fully addressed	Dedicated study addressing data governance requirements for model training on payroll data

XVII. RECOMMENDATIONS

- Implement all three detection layers, statistical process control, isolation forest, and autoencoder reconstruction, rather than relying on threshold rules or a single machine learning method alone.
- Treat detector agreement as a first-class signal in score fusion, applying differentiated confidence tiers rather than uniform treatment of all flagged records.
- Integrate known business calendar events into the detection context to reduce false positives from legitimate periodic events, per Section 10.
- Establish differentiated severity classification and category-specific escalation routing, rather than a single generic alert queue.
- Apply automatic cross-functional escalation for any critical severity alert in the regulatory and compliance category, given the asymmetric cost of delayed regulatory response.
- Establish scheduled retraining and control limit refresh cadences from initial deployment, incorporating analyst feedback loop determinations as labeled training data.
- Adapt the severity and routing framework to organization-specific structure rather than adopting the illustrative structure in this article verbatim.

XVIII. FUTURE RESEARCH DIRECTIONS

This study suggests several directions for further research. First, an empirical study measuring actual detection precision, recall, and ensemble agreement patterns across one or more production payroll and benefits environments would strengthen the illustrative findings presented throughout this article with statistically validated figures. Second, research into cold-start detection approaches suitable for organizations with limited historical payroll data volume could extend this framework's applicability beyond the sufficient-data assumption addressed in Section 16. Third, comparative research examining additional detection methods, including graph-based detection capturing relationships

between related employee records, or time-series-specific deep learning architectures, could clarify whether meaningful performance gains remain available beyond the three-detector ensemble examined here. Finally, research into the specific data governance and privacy safeguards required when training machine learning models on payroll and benefits data would directly address the limitation noted in Section 16 and would provide valuable guidance for organizations navigating the intersection of data protection requirements and predictive detection capability.

XIX. CONCLUSION

This article has presented an ensemble machine learning framework for anomaly detection in inbound payroll and benefits data feeds, combining statistical process control, isolation forest models, and autoencoder neural network reconstruction into a unified radial pipeline architecture. The comparative analysis in Section 13 indicates that this three-detector ensemble achieves substantially stronger illustrative precision and recall than any single method alone, or than static threshold rules, directly addressing the specific categories of anomaly, aggregate shifts, multivariate outlier combinations, and subtle multivariate drift, that threshold-based validation structurally cannot detect. At the same time, the ensemble agreement analysis in Section 9 and the anomaly rate pattern analysis in Section 10 underscore that effective deployment depends on more than detection methodology alone: differentiated confidence tiers based on detector agreement, business calendar context to reduce false positives from legitimate periodic events, and a category-specific escalation routing framework are each necessary to translate raw detection output into an operationally usable alert stream. Organizations processing high-stakes inbound payroll and benefits data are well positioned to substantially improve their anomaly detection posture by adopting this ensemble framework, provided that framework is supported by the governance practices addressing retraining, control limit refresh, and routing accuracy that Section 14 identifies as necessary to sustain detection performance over time.

REFERENCES

1. Aggarwal, C. C. (2017). Outlier analysis (2nd ed.). Springer.
2. Chandola, V., Banerjee, A., and Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), 1-58.
3. Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep learning. MIT Press.
4. Liu, F. T., Ting, K. M., and Zhou, Z. H. (2008). Isolation forest. Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, 413-422.
5. Montgomery, D. C. (2019). Introduction to statistical quality control (8th ed.). Wiley.
6. Sakurada, M., and Yairi, T. (2014). Anomaly detection using autoencoders with nonlinear dimensionality reduction. Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis, 4-11.

Appendix A: Glossary of Terms

Table.A-1. Glossary of key terms used throughout this article.

Term	Definition
Anomaly Detection	The identification of records or patterns that deviate from expected or learned normal behavior
Statistical Process Control (SPC)	A quality management methodology using control limits around an expected distribution to flag anomalous observations
Isolation Forest	An unsupervised anomaly detection algorithm based on the relative ease of isolating anomalous points through random partitioning
Autoencoder	A neural network trained to reconstruct its own input through a reduced-dimensionality bottleneck, used for anomaly detection via reconstruction error
Reconstruction Error	The difference between an autoencoder's reconstructed output and its original input, used as an anomaly signal
Ensemble Detection	An approach combining multiple detection methods to improve overall detection robustness and reliability

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Control Limit	A statistical threshold, typically based on trailing mean and standard deviation, used to flag anomalous observations in SPC
Severity Classification	The process of assigning a priority tier to a detected anomaly based on confidence and potential impact
False Positive (Anomaly Detection)	A record incorrectly flagged as anomalous when it is in fact a legitimate, normal record
Score Fusion	The process of combining outputs from multiple detection methods into a single unified anomaly score

Appendix B: Supplementary Data Tables

This appendix provides additional illustrative data tables supporting the analysis in the main body of the article, including an extended detection layer parameter reference and a consolidated alert response time summary.

B.1 Extended Detection Layer Parameter Reference

Table.B-1. Extended detection layer parameter reference with illustrative values and adjustment effects.

Detection Layer	Key Parameter	Illustrative Value	Effect of Adjustment
Statistical Process Control	Control limit multiplier	3 standard deviations	Lower values increase sensitivity and false positive rate
Isolation Forest	Number of trees	200	Higher values improve stability at increased compute cost
Isolation Forest	Contamination parameter	0.02 (2 percent expected anomaly rate)	Directly affects the model's decision threshold
Autoencoder	Bottleneck dimension	8 (from 32 input features)	Lower values force more aggressive compression, increasing sensitivity to subtle drift
Autoencoder	Reconstruction error threshold	2.6 (standardized)	Directly determines the precision/recall trade-off for this layer

B.2 Illustrative Alert Response Time Summary

Table.B-2. Illustrative alert response time summary by severity tier

Severity Tier	Illustrative Mean Time to Initial Review	Illustrative Mean Time to Resolution
Low Severity	3.2 business days	6.5 business days
Medium Severity	0.8 business days	2.1 business days
High Severity	3.1 hours	1.2 business days
Critical Severity	0.6 hours	8.4 hours

© 2025 IJIRCCE | Volume 13, Issue 7, July 2025| DOI:10.15680/IJIRCCE.2025.1307046

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

B.3 Illustrative Model Retraining Outcomes Over One Year

Table.B-3. Illustrative model retraining outcomes over a one-year monitoring period

Retraining Cycle	Illustrative Ensemble F1 Score After Retraining	Notable Drift Addressed
Cycle 1 (Initial Deployment)	0.87	N/A (initial deployment)
Cycle 2 (Quarter 2)	0.85	Moderate drift following a benefits plan restructuring
Cycle 3 (Quarter 3)	0.86	Minor drift; primarily scheduled retraining
Cycle 4 (Quarter 4)	0.89	Adjusted for a new pay group introduced mid-year

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details