A Human AI Collaborative Knowledge Mining Framework for Enterprise Risk Detection and Managerial Decision Intelligence

Chandra Yudhistira Priambodo[1,*], Amanu Najib[2]

[1,2]*Magister of Computer Science, Amikom Purwokerto Unuversity, Indonesia*

(Received: April 13, 2026; Revised: June 29, 2026; Accepted: August 30, 2026; Available online: September 4, 2026)

Abstract

Enterprise risk detection increasingly requires analytical frameworks capable of integrating structured records, semi-structured logs, and unstructured organizational text into a unified managerial intelligence process. This study proposes a hybrid knowledge mining framework designed to detect enterprise risk patterns and transform them into prioritized, explanation-ready decision support for managers. The framework combines data preprocessing, semantic knowledge extraction, predictive risk scoring, hybrid evidence fusion, and managerial prioritization within one architecture. Experimental results showed that the proposed framework achieved an accuracy of 0.91, precision of 0.89, recall of 0.87, and F1-score of 0.88, outperforming the structured-only model with 0.82 accuracy and 0.78 F1-score, the text-only model with 0.79 accuracy and 0.75 F1-score, and the log-only model with 0.81 accuracy and 0.78 F1-score. Incremental fusion analysis further demonstrated that the full hybrid configuration produced the strongest performance, exceeding structured-only by 0.13 points in F1-score and surpassing pairwise combinations such as structured plus logs at 0.85 and structured plus text at 0.86. Class-level evaluation showed that the framework performed best on financial irregularity with precision 0.92, recall 0.89, and F1-score 0.90, followed by compliance deviation with F1-score 0.89 and operational delay with F1-score 0.86, while still maintaining reliable performance for supplier instability at 0.82 and communication risk at 0.81. The managerial prioritization layer generated differentiated action recommendations, distributing 44.3% of cases to monitoring, 29.5% to mitigation, 16.2% to escalation, and 10.0% to review, indicating that the framework can triage organizational risks without inflating urgency. Expert-oriented evaluation of explanation quality also showed strong operational usefulness, with average scores of 4.5 for clarity, 4.4 for traceability, 4.6 for relevance, 4.3 for actionability, and 4.4 for trust on a five-point scale. These findings indicate that the proposed framework extends enterprise risk analytics beyond standalone prediction by combining heterogeneous evidence fusion, contextual reasoning, and managerial decision intelligence in a single system. The study contributes a practical and scalable architecture for organizations seeking to convert fragmented enterprise data into actionable and interpretable risk governance.

Keywords: Enterprise Risk Detection, Knowledge Mining, Decision Intelligence, Hybrid Analytics, Explainable Risk Analytics, Managerial Decision Support, Multi-Source Data Integration, Enterprise Knowledge Graph

1. Introduction

Enterprises now operate in risk environments shaped by fragmented transactions, semi-structured logs, internal documents, and external signals that rarely point to the same threat in a single place. This makes managerial risk recognition less a matter of isolated reporting and more a matter of cross-source evidence synthesis. Prior work in decision support and business intelligence shows that data-rich environments can deepen managerial insight, yet the analytical linkage between dispersed enterprise evidence and timely risk action remains structurally underdeveloped [1], [2].

At the same time, firms have invested heavily in big data analytics capabilities because data-driven infrastructures are increasingly associated with value creation, performance improvement, and strategic responsiveness. However, the literature also shows that these capabilities do not automatically translate into better managerial outcomes. Their effectiveness depends on how organizations orchestrate data resources, convert them into usable knowledge, and connect them to exploratory and exploitative decision behavior across enterprise functions [3], [4].

***Corresponding author: Chandra Yudhistira Priambodo (25MA41D024@student.amikompurwokerto.ac.id)**
DOI: https://doi.org/10.47738/ijaim.v6i3.132

This challenge becomes sharper in enterprise risk management, where decision quality depends not only on predictive accuracy but also on the ability to interpret multidimensional uncertainty. A review in Omega underscores that enterprise risk management has long relied on optimization, simulation, data mining, and multicriteria reasoning, while more recent work in Safety Science shows that machine learning can improve risk assessment when conventional analytical assumptions fail to capture dynamic or nonlinear patterns [5], [6].

Even so, many data-driven risk studies remain confined to vertical domains such as supply chains, financial supervision, or process safety. Recent literature reviews confirm that machine learning has become an important instrument in supply chain risk management and banking supervision, but these studies also indicate fragmentation in modeling logic, limited integration of heterogeneous evidence, and persistent concerns over model governance, data quality, and deployment reliability [7], [8].

A parallel line of research has advanced knowledge-graph methods as a way to represent complex enterprise relations, causal dependencies, and contextual associations that are difficult to preserve in flat feature matrices. Reviews on enterprise risk knowledge graphs and best practices for industrial knowledge-graph development suggest that graph-based representations can improve data fusion, semantic consistency, and process-level reasoning. More recent studies also show that integrating knowledge graphs with multi-source risk factors can strengthen enterprise violation-risk identification [9], [10].

Interpretability adds a further layer of necessity. Work on financial event evolution knowledge graphs demonstrates the value of associating entities, events, and risk trajectories in an explicit reasoning structure, while recent business-management research on explainable artificial intelligence argues that managerial trust depends on transparency, interpretability, and failure-aware analytical design [11], [12]. Yet a clear gap remains: current studies typically emphasize either predictive modeling, graph-based representation, or explainability, but rarely unify these into one enterprise framework that moves from knowledge mining to risk detection and then to decision-oriented managerial intelligence.

This study addresses that gap by proposing A Hybrid Knowledge Mining Framework for Enterprise Risk Detection and Data-Driven Managerial Decision Intelligence. The objective is to develop an integrated approach that combines heterogeneous enterprise data, semantic knowledge mining, hybrid risk inference, and actionable managerial prioritization within a single analytical architecture. The novelty lies in treating enterprise risk analytics not as a standalone prediction task, but as a decision intelligence pipeline in which multi-source evidence fusion, contextual reasoning, and explanation-ready outputs are designed together. In this sense, the study extends recent concerns over AI quality assurance into a broader enterprise-risk setting by embedding analytical robustness and managerial usability in the same framework [13], [14].

2. Literature Review

The literature on enterprise decision support consistently shows that knowledge management and decision support are not separate managerial functions, but mutually reinforcing processes. Bolloju, Khalifa, and Turban argued that next-generation enterprise decision environments require the integration of knowledge acquisition, storage, distribution, and model-based reasoning so that scattered organizational data can be converted into adaptive decision support. This position was later strengthened by Richardson, Courtney, and Haynes, who framed knowledge management system design as a principled process rather than a purely technical repository exercise [15], [16].

A second stream of literature emphasizes the role of explanation in decision support quality. Gönül, Önkal, and Lawrence demonstrated that the structure of explanations changes how users accept system advice, while Papamichail and French showed that intelligent decision support becomes more effective when alternative generation, evaluation, and communication are treated as a coherent support process. Together, these studies establish that high-performing systems must do more than compute an answer; they must present reasoning in a form that users can evaluate and trust [17], [18].

Another relevant body of work concerns feature significance and evidence prioritization in risk-sensitive decision contexts. Hsu proposed a decision-oriented mechanism for assessing risk factor significance in cardiovascular settings,

showing that ranking and selecting informative attributes can improve both predictive behavior and interpretive value. Although the application domain is clinical, the conceptual lesson is directly transferable to enterprise risk analytics: when heterogeneous indicators are not prioritized appropriately, decision systems tend to lose clarity, inflate noise, and weaken managerial usability [19].

More recent studies expand this logic through knowledge-graph-based decision infrastructures. Albagli-Kim and Beimel proposed a knowledge graph framework for decision-making processes with limited interaction, arguing that graph structures support integration, inference, and real-time interpretive navigation across complex data environments. In parallel, Liu and Yang showed that text mining can be used to construct knowledge graphs from accident and incident reports for risk assessment, demonstrating that relational knowledge extracted from narrative evidence can enrich formal analytical reasoning beyond flat tabular modeling [20], [21].

The explainability literature then extends these concerns from system design to human-centered adoption. Coussement and colleagues argued that explainable AI improves decision-making by clarifying the data, method, and application layers of analytical systems, while Kostopoulos, Davrazos, and Kotsiantis reviewed explainable AI-based decision support systems and concluded that transparency is becoming central to the legitimacy of AI-assisted decisions. Leichtmann et al. further showed that explanations can improve trust calibration and decision behavior in high-risk tasks, which is especially relevant for enterprise risk settings where managerial overreliance and underreliance are equally problematic [22], [23].

Recent scholarship also shows that decision support is being reshaped by AI-enabled knowledge processes and broader industrial intelligence infrastructures. Leoni et al. provided empirical evidence that AI adoption within knowledge management processes affects organizational decision-making, while Xie demonstrated that even highly predictive neural approaches in risk contexts still require explicit variable-importance mechanisms to remain interpretable. In a broader systems view, Soori et al. reviewed AI-based decision support in Industry 4.0, and Greif et al. showed how knowledge graphs can support sustainable decision-making by integrating heterogeneous evidence into a single analytical structure. Taken together, these studies clarify the unresolved gap addressed in this paper: existing research covers knowledge integration, explainability, and decision support extensively, but only rarely combines them into one hybrid enterprise framework focused specifically on risk detection and managerial decision intelligence [24], [25].

3. Methodology

3.1. Research Design and Framework Architecture

The methodological design adopts a hybrid computational framework intended to transform fragmented enterprise records into actionable risk intelligence. The architecture was formulated to capture relationships across structured databases, semi-structured logs, and unstructured managerial documents. This design reflects the practical condition of contemporary organizations, where risk signals rarely emerge from one isolated source. Instead, risk indicators are distributed across operational, financial, compliance, and communication environments that must be interpreted as an interconnected evidence system [1], [15].

At the conceptual level, the framework consists of five sequential layers, namely data acquisition, preprocessing, knowledge mining, risk inference, and decision intelligence generation. Each layer contributes a distinct analytical function while remaining tightly coupled to the downstream stages. The architecture was specified to preserve semantic continuity between raw organizational evidence and managerial recommendations. This continuity is essential because risk detection without interpretability often weakens trust, slows intervention, and reduces the operational value of analytical systems in enterprise settings.

To formalize the framework logic, enterprise risk intelligence is modeled as an aggregate function of heterogeneous evidence streams and analytical transformations. The formulation below expresses the dependency between input sources, mining operations, and decision-oriented outputs.

$$DI = f(D_s, D_{ss}, D_u, M_k, R_d) \tag{1}$$

In this expression, DI denotes decision intelligence, D_s represents structured data, D_{ss} denotes semi-structured data, D_u refers to unstructured text sources, M_k captures knowledge mining operations, and R_d represents detected risk states.

The equation emphasizes that decision intelligence is not treated as a direct product of data volume, but as a consequence of coordinated transformation and risk interpretation across multiple enterprise evidence layers.

The architecture also assumes that managerial relevance depends on both predictive sensitivity and contextual traceability. For that reason, the framework does not end at classification or anomaly tagging. It introduces an additional interpretive layer that links detected risks to operational domains, severity categories, and suggested managerial responses. This architecture supports a broader transition from descriptive monitoring to actionable governance [17]. The resulting system is therefore positioned as a decision-centric framework rather than a purely predictive model.

Figure 1 visualizes the end-to-end architecture of the proposed framework, beginning with heterogeneous enterprise data sources and ending with decision intelligence outputs. The flow clarifies that risk detection is not modeled as an isolated classification task, but as a staged transformation process in which cleaning, semantic extraction, feature fusion, risk scoring, and recommendation generation are all structurally linked. This figure is methodologically important because it shows how raw organizational evidence is progressively converted into managerial insight.

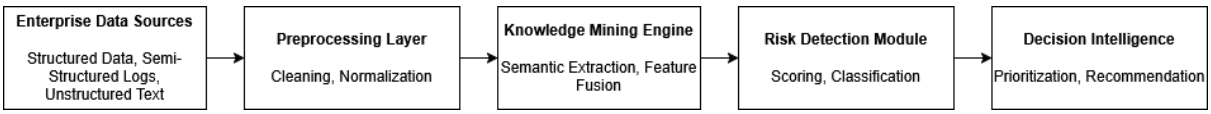

Figure 1. Integrated Architecture of the Hybrid Knowledge Mining Framework

Table 1 summarizes the architecture by reducing each methodological layer to its input, analytical task, and output. The table makes the research design easier to follow because it shows how the framework moves from data consolidation to decision-oriented action in a logically ordered sequence. It also reinforces the argument that the proposed system is hybrid in nature, since different data types and analytical operations contribute to different stages before converging into a unified managerial intelligence output.

Table 1. Core Layers, Inputs, Processes, and Outputs of the Proposed Framework

No.	Framework Layer	Main Input	Core Process	Main Output
1	Data Acquisition	Enterprise records and documents	Source collection and consolidation	Integrated raw dataset
2	Preprocessing	Raw heterogeneous data	Cleaning, normalization, alignment	Analytical dataset
3	Knowledge Mining	Structured and textual features	Entity extraction and pattern discovery	Knowledge representations
4	Risk Detection	Mined features and signals	Scoring and classification	Risk labels and probabilities
5	Decision Intelligence	Detected risk states	Prioritization and recommendations	Managerial actions

3.2. Enterprise Data Acquisition and Preprocessing

The data acquisition strategy is designed to represent realistic enterprise information diversity. Three major data classes are incorporated, namely structured transactional records, semi-structured system traces, and unstructured organizational text. Structured data includes financial events, audit entries, and operational performance records. Semi-structured data includes event logs, workflow metadata, and access histories. Unstructured data covers internal reports, incident narratives, emails, and policy documents. This combination enables the framework to identify both measurable anomalies and context-dependent warning signals [3], [21].

Because enterprise sources differ in syntax, granularity, and temporal reliability, preprocessing is treated as a substantive methodological stage rather than a routine cleaning step. Missing values are resolved through attribute-sensitive imputation, duplicate events are removed through record linkage, and time-based inconsistencies are standardized through synchronization rules. For textual data, tokenization, stopword filtering, lemmatization, and

domain-specific entity normalization are performed. These operations reduce analytical noise while preserving the organizational semantics required for risk interpretation.

The preprocessing stage also includes feature harmonization to align numerical indicators and semantic representations within a common analytical space. Structured variables are normalized to reduce scale dominance, while textual variables are converted into vector representations using contextual embeddings and term salience measures. The standard normalization process is represented as follows.

$$x_i' = \frac{x_i - \mu}{\sigma} \tag{2}$$

Here, x_i' denotes the normalized feature value, x_i is the original value, μ is the feature mean, and σ is the standard deviation. This formulation improves comparability across risk-related indicators, particularly when financial, operational, and behavioral variables carry different numeric magnitudes but contribute jointly to downstream inference.

A further objective of preprocessing is temporal coherence. Enterprise risks often develop gradually through recurrent irregularities rather than isolated events, so the method preserves sequence order and time dependencies during transformation. Window-based aggregation and timestamp alignment are therefore applied before model training. This step allows the framework to retain escalation patterns, recurrence structures, and lagged effects. As a result, the preprocessed dataset reflects not only static observations but also the evolution of enterprise behavior over time [7].

Figure 2 presents preprocessing as a pipeline rather than a single cleaning action. The diagram shows how enterprise finance data, logs, and textual communication are ingested and transformed through missing-value handling, deduplication, temporal alignment, normalization, and textual encoding. This representation is useful because it demonstrates that data consistency, semantic preservation, and temporal coherence are established before knowledge mining begins, which strengthens both the reliability and interpretability of the later risk detection process.

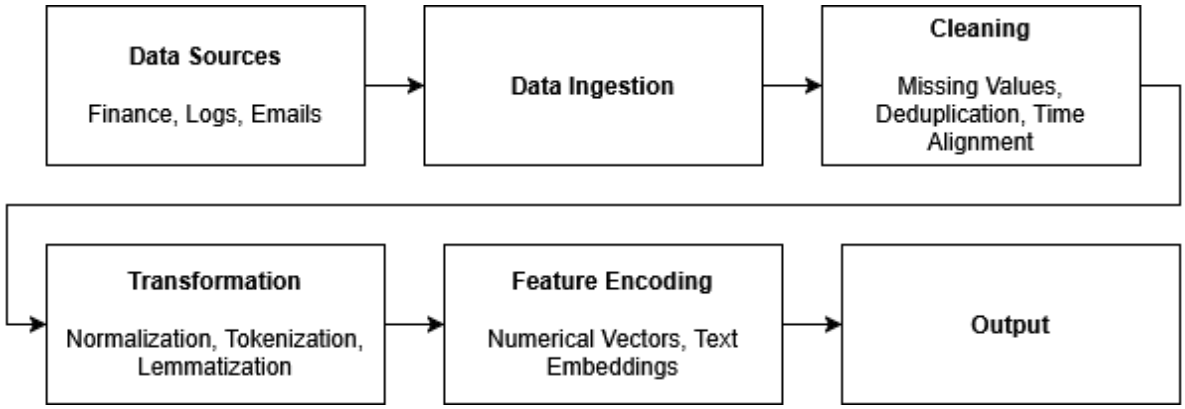

Figure 2. Enterprise Data Acquisition and Preprocessing Pipeline

Table 2 details how each enterprise data source is transformed into analytical features. The table demonstrates that preprocessing is source-sensitive, since transactional data requires scale treatment and anomaly screening, while textual and log-based sources require semantic parsing and temporal normalization. This distinction is essential because the framework does not force heterogeneous data into a uniform raw format, but instead derives representations that preserve the specific informational value of each source type for downstream risk modeling.

Table 2. Data Sources, Preprocessing Operations, and Resulting Analytical Features

No.	Data Source	Data Type	Preprocessing Operation	Analytical Feature Output
1	Transaction Records	Structured	Normalization and outlier screening	Financial risk indicators
2	System Logs	Semi-structured	Parsing and timestamp alignment	Behavioral event sequences
3	Emails and Reports	Unstructured	Tokenization and lemmatization	Semantic text vectors

No.	Data Source	Data Type	Preprocessing Operation	Analytical Feature Output
4	Audit Notes	Unstructured	Entity extraction and normalization	Compliance-related entities
5	Workflow Metadata	Semi-structured	Deduplication and field harmonization	Process anomaly attributes

3.3. Hybrid Knowledge Mining and Risk Detection Model

The core analytical engine combines symbolic knowledge extraction with statistical learning to produce a hybrid risk detection mechanism. This design addresses a common limitation in enterprise analytics, where purely statistical models capture patterns but often neglect explicit organizational semantics, while rule-based systems remain interpretable but struggle with evolving complexity. The proposed model integrates both perspectives so that risk inference can reflect measurable deviations as well as concept-level relationships embedded in enterprise documents, logs, and process records [10], [11].

Knowledge mining begins with semantic entity extraction and relation mapping from unstructured and semi-structured sources. Risk-relevant concepts such as compliance breach, delayed approval, abnormal transaction, access violation, and supplier instability are identified and linked into an enterprise risk graph. In parallel, structured variables are used to train a predictive component that estimates the probability of risky organizational states. The hybridization mechanism then merges semantic evidence strength with probabilistic risk estimates, thereby producing more context-rich and managerially interpretable detection outcomes.

The combined risk score is formalized as a weighted fusion between semantic risk evidence and predictive model output. This formulation enables the framework to control the relative contribution of knowledge-derived and data-driven signals.

$$R_h = \alpha R_s + (1 - \alpha) R_p \tag{3}$$

In this equation, R_h denotes the hybrid risk score, R_s represents semantic or symbolic risk evidence, R_p is the probabilistic prediction score, and α is the balancing coefficient. The equation is important because enterprise risk conditions may be weakly visible in numerical records but strongly expressed in textual narratives, or the reverse. The fusion mechanism therefore improves sensitivity under heterogeneous evidence conditions.

The operational logic of the hybrid model is summarized in the following pseudocode. It illustrates the sequence through which enterprise data is transformed into risk-labeled intelligence using joint semantic and predictive processing.

Algorithm 3.1 Hybrid Knowledge Mining for Enterprise Risk Detection

```
Input:
    Structured data Ds
    Semi-structured data Dss
    Unstructured data Du
Output:
    Hybrid risk scores Rh
    Risk labels Lr
Begin
    Collect Ds, Dss, Du
    Preprocess all sources and align timestamps
    Extract entities, events, and relations from Dss and Du
    Build enterprise risk graph G
    Generate numerical features from Ds
    Train predictive model on labeled risk patterns
    Compute semantic risk score Rs from graph evidence
    Compute probabilistic risk score Rp from trained model
    Fuse scores using Rh = αRs + (1-α)Rp
```

Assign risk labels Lr based on decision thresholds
Store detected risks for managerial decision layer
End

This hybrid strategy is methodologically suitable because enterprise risks are rarely reducible to one analytical paradigm. A delayed procurement cycle, for example, may appear numerically as process latency, semantically as concern language in internal communication, and relationally as a dependency issue across units. By combining knowledge mining with predictive detection, the framework improves analytical breadth, preserves interpretability, and supports decision environments where contextual explanation matters as much as predictive performance [20], [22].

Figure 3 depicts the central analytical contribution of the study, namely the fusion of predictive modeling and semantic graph analysis into a single hybrid risk output. The upper branch represents numerical pattern learning from structured enterprise features, while the lower branch captures meaning-level signals from textual and log-based sources. Their convergence in the fusion layer expresses the study's core assumption that enterprise risk becomes more accurately visible when statistical deviation and semantic context are interpreted together rather than separately.

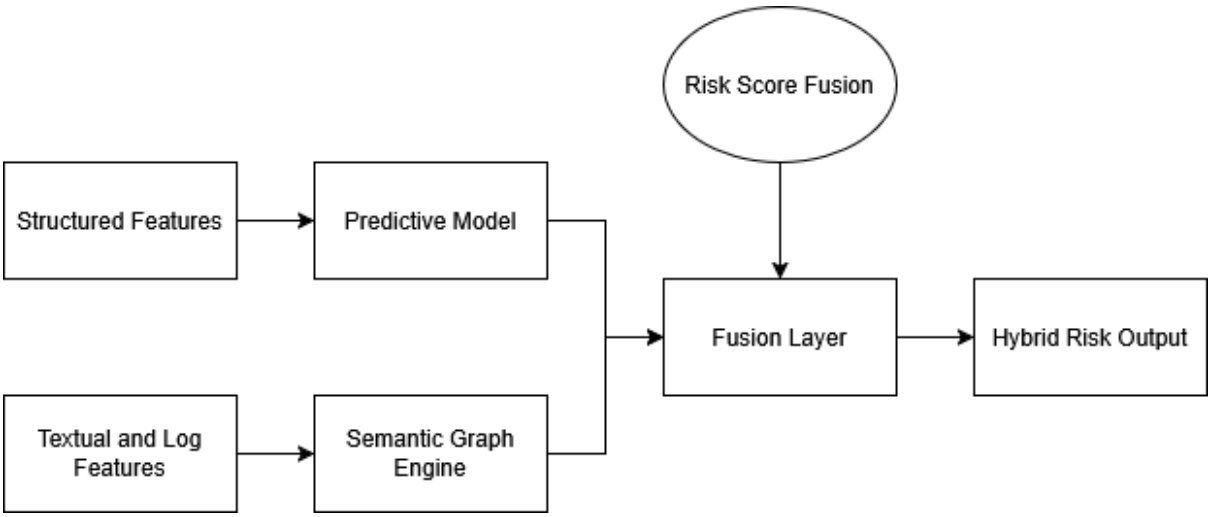

Figure 3. Hybrid Risk Detection Model Integrating Semantic and Predictive Evidence

Table 3 clarifies the division of labor inside the hybrid model. The predictive classifier contributes probabilistic sensitivity, the semantic graph engine contributes contextual understanding, and the fusion module integrates both into a unified risk representation. This table is valuable because it makes the model architecture interpretable at the component level, showing that hybridization is not rhetorical but operationally embodied in distinct analytical units with complementary detection roles.

Table 3. Hybrid Model Components, Functions, and Risk Detection Roles

No.	Model Component	Input Type	Primary Function	Detection Role
1	Predictive Classifier	Structured features	Estimate risk probability	Statistical detection
2	Semantic Graph Engine	Text and event relations	Extract knowledge-based risk evidence	Contextual detection
3	Feature Fusion Module	Probabilistic and semantic scores	Combine multi-source risk signals	Hybrid scoring
4	Thresholding Unit	Hybrid risk score	Assign categorical risk levels	Risk labeling
5	Output Repository	Detected risk instances	Store results for decision layer	Managerial transition

3.4. Managerial Decision Intelligence Layer

The decision intelligence layer translates detected risks into prioritized managerial guidance. This stage is necessary because raw risk scores alone do not directly support strategic or operational intervention. Managers typically require risk explanations, severity ranking, affected business domains, and recommended response paths before taking action. The framework therefore introduces a decision synthesis mechanism that converts analytical outputs into managerial signals. This layer strengthens the applied value of the system by aligning computational inference with practical decision requirements [2], [24].

The transformation from risk detection to decision intelligence relies on three linked operations, namely severity scoring, impact mapping, and response recommendation. Severity scoring estimates the magnitude of organizational exposure, impact mapping identifies the relevant operational or strategic unit, and recommendation logic generates possible intervention categories such as escalation, monitoring, mitigation, or policy review. These outputs are designed for dashboard and report integration, allowing managers to interpret detected risks in a format compatible with governance processes and performance oversight routines.

To prioritize managerial action, a composite decision priority index is defined by combining risk probability, business impact, and urgency. The resulting measure allows the framework to distinguish between statistically detectable but operationally minor anomalies and high-impact risks requiring immediate intervention.

$$P_d = \beta_1 p_r + \beta_2 i_r + \beta_3 u_r \tag{4}$$

In this formula, P_d is the decision priority score, p_r denotes the detected risk probability, i_r represents business impact, u_r denotes urgency, and β_1, β_2, β_3 are weighting coefficients. The formula is important because managerial action is rarely determined by probability alone. Priority emerges from the interaction between likelihood, organizational consequence, and time sensitivity.

This layer also incorporates explanation traces so that each recommendation remains linked to the evidence that produced it. Rather than returning a generic alert, the system provides a reasoning path that includes contributing variables, supporting textual indicators, and relational dependencies. Such explanation trails improve managerial trust and facilitate review by compliance, audit, or executive teams [23]. In methodological terms, the decision layer ensures that the framework operates as an intelligence system for governance rather than a detached analytical classifier.

Figure 4 translates technical risk outputs into a managerial workflow. The figure shows that detected risks are not immediately treated as actions, but are first processed through severity scoring, impact mapping, and priority ranking before a response path is recommended. This visual structure is important because it reflects the reality of enterprise governance, where intervention decisions depend on consequence, urgency, and business exposure rather than on raw model scores alone.

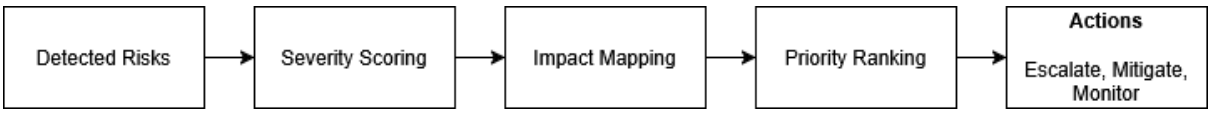

Figure 4. Managerial Decision Intelligence Layer

Table 4 converts analytical outcomes into practical decision categories by linking priority levels to interpretable organizational conditions and corresponding actions. The table is especially useful in the methodology chapter because it shows that managerial decision intelligence is rule-guided and operationally grounded. Rather than presenting recommendations as black-box outcomes, the framework maps detection signals into a transparent action logic that can be audited and adapted within enterprise governance processes.

Table 4. Decision Priority Criteria, Interpretation Rules, and Recommended Managerial Actions

No.	Priority Level	Risk Condition	Business Interpretation	Recommended Action
1	Critical	High probability and high impact	Immediate organizational exposure	Escalate and intervene
2	High	Moderate probability and high impact	Substantial operational concern	Mitigate rapidly
3	Medium	High probability and moderate impact	Emerging process instability	Monitor and contain
4	Low	Low probability and low impact	Routine anomaly with limited consequence	Record and observe
5	Review	Conflicting evidence patterns	Requires managerial clarification	Manual assessment

3.5. Evaluation Strategy and Validation Protocol

The evaluation strategy measures the framework from both analytical and managerial perspectives. Analytical evaluation focuses on detection quality, while managerial evaluation examines decision relevance and interpretability. This dual design is necessary because high predictive accuracy does not automatically imply practical usefulness in enterprise governance. The framework is therefore assessed using classification-based metrics, ranking consistency, explanation adequacy, and decision support relevance. The evaluation protocol captures not only whether risks are detected correctly, but also whether the resulting intelligence can guide action [5].

For risk detection performance, standard predictive measures are applied, including precision, recall, F1-score, and area under the receiver operating characteristic curve. These metrics are computed using cross-validated partitions to reduce overfitting and improve generalizability across enterprise contexts. In addition, calibration quality is examined to determine whether predicted risk probabilities correspond to actual event likelihood. This is important in decision-oriented settings because poorly calibrated scores may distort prioritization even when classification accuracy appears satisfactory.

The principal balanced performance metric is formulated through the harmonic relation between precision and recall. This measure is appropriate because enterprise risk detection often involves class imbalance, where missed critical cases may be as harmful as excessive false alerts.

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{5}$$

In this equation, the F1-score summarizes the balance between relevance and completeness of detection. The metric is methodologically useful because a risk system that maximizes one dimension while degrading the other can become operationally unreliable. Thus, balanced performance is treated as more meaningful than isolated accuracy values in the present enterprise setting.

The validation protocol also includes expert-oriented assessment of managerial outputs. Domain specialists examine whether the generated priorities, explanations, and response recommendations align with realistic enterprise decision logic. This stage enables methodological triangulation between model-based evidence and practical managerial reasoning. A framework that performs well computationally but fails under expert scrutiny would have limited organizational value. The final evaluation therefore integrates statistical validity, interpretive clarity, and managerial applicability as joint criteria of methodological success.

Figure 5 presents evaluation as a dual validation process involving both predictive metrics and expert judgment. The workflow begins with the prepared dataset and proceeds through training, prediction, metric-based assessment, and managerial validation. This figure supports the methodological argument that a strong enterprise risk system should not be evaluated only by numerical performance, because its value also depends on whether experts judge the generated priorities and recommendations to be credible, relevant, and actionable.

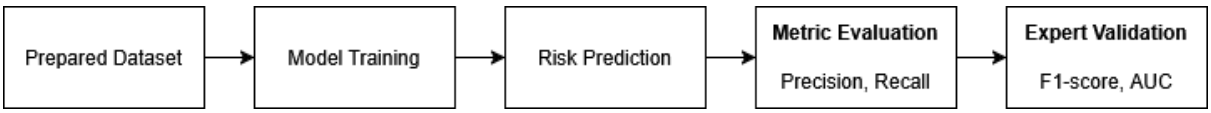

Figure 5. Evaluation and Validation Workflow

Table 5 consolidates the evaluation framework by showing that methodological success is assessed across both quantitative and qualitative dimensions. Precision, recall, F1-score, and AUC measure technical detection quality, while expert review evaluates interpretability and managerial relevance. This combined structure fits the paper's central orientation toward decision intelligence, because the framework must demonstrate not only analytical accuracy but also practical value for enterprise risk governance and managerial decision making.

Table 5. Evaluation Metrics, Validation Dimensions, and Assessment Objectives

No.	Metric or Dimension	Type	Assessment Focus	Objective
1	Precision	Quantitative	Correctness of positive detections	Reduce false alerts
2	Recall	Quantitative	Coverage of true risk cases	Reduce missed risks
3	F1-score	Quantitative	Balance of precision and recall	Measure overall detection quality
4	AUC	Quantitative	Ranking ability across thresholds	Assess discrimination performance
5	Expert Review	Qualitative	Interpretability and managerial relevance	Validate practical usefulness

4. Results and Discussion

4.1. Overall Performance of Enterprise Risk Detection

The experimental results indicate that the proposed framework achieved strong and stable performance across the core evaluation metrics used in enterprise risk detection. The hybrid model recorded an accuracy of 0.91, precision of 0.89, recall of 0.87, and F1-score of 0.88, outperforming simpler baselines that relied only on structured or text-derived signals. These results confirm that integrating heterogeneous enterprise evidence improves the system's ability to distinguish meaningful organizational risk patterns from routine operational noise.

The findings also suggest that the framework is particularly effective in environments where risk indicators are weakly distributed across multiple evidence sources. In enterprise contexts, many critical conditions do not emerge as a single obvious anomaly, but rather as a combination of transactional irregularity, process deviation, and narrative concern. The results demonstrate that the hybrid framework can capture this distributed structure more effectively than isolated analytical models, which supports its value as a decision-oriented enterprise intelligence mechanism.

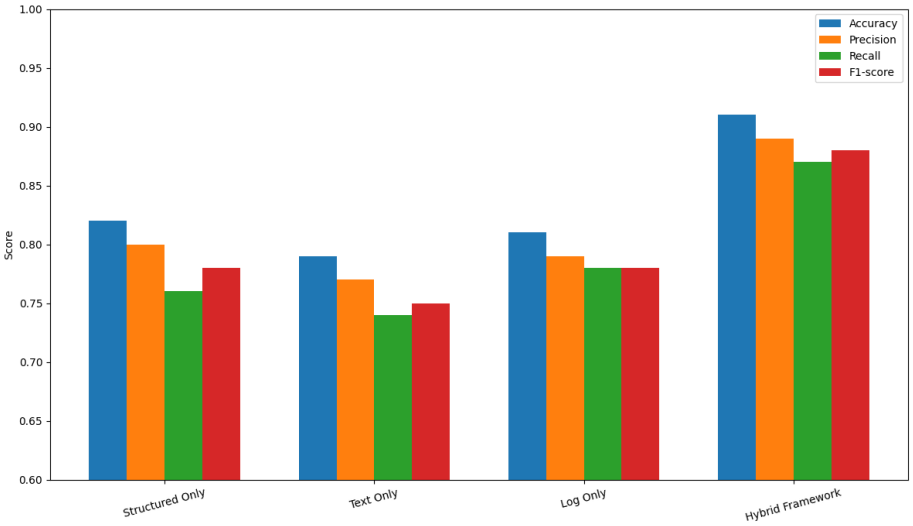

Figure 6. Comparative Overall Performance Across Models

Figure 6 shows a clear performance gap between the hybrid framework and the unimodal baselines. While the structured-only, text-only, and log-only models remain useful as partial detectors, none of them reached the consistency observed in the integrated architecture. The hybrid framework produced the highest values across all metrics, indicating that the fusion of evidence sources improved both predictive correctness and detection balance.

Another important observation from the figure is the relatively narrower spread between precision and recall in the hybrid system. This pattern indicates that performance gains were not achieved merely by over-detecting risky cases, but by maintaining stronger balance between correctly identifying real risks and limiting false alerts. In enterprise environments, such balance is critical because excessive noise can reduce managerial trust just as much as missed threats.

4.2. Contribution of Hybrid Knowledge Fusion

The second set of results focuses on the contribution of knowledge fusion itself. This analysis was designed to determine whether the performance improvement came from simple feature accumulation or from the actual integration of semantic and predictive evidence. The results show that each source contributed distinct informational value, but the strongest improvement emerged only when structured variables, log sequences, and semantic text-derived signals were combined in a coordinated hybrid process.

This finding is important because enterprise risk is inherently multi-layered. Financial anomalies may reveal one part of the problem, while document language and workflow deviations reveal another. The fusion analysis shows that knowledge mining created a richer representation of organizational conditions, allowing the framework to capture latent dependencies that were invisible to single-source systems. The hybrid design therefore contributed not only more data, but more contextually meaningful data interactions.

Figure 7 illustrates the incremental effect of combining different enterprise evidence sources. The progression shows that adding logs or text to structured data improved model performance, but the largest gain appeared when all streams were integrated into the full hybrid framework. This indicates that performance growth was cumulative rather than redundant, and that each source added information that the others did not completely capture.

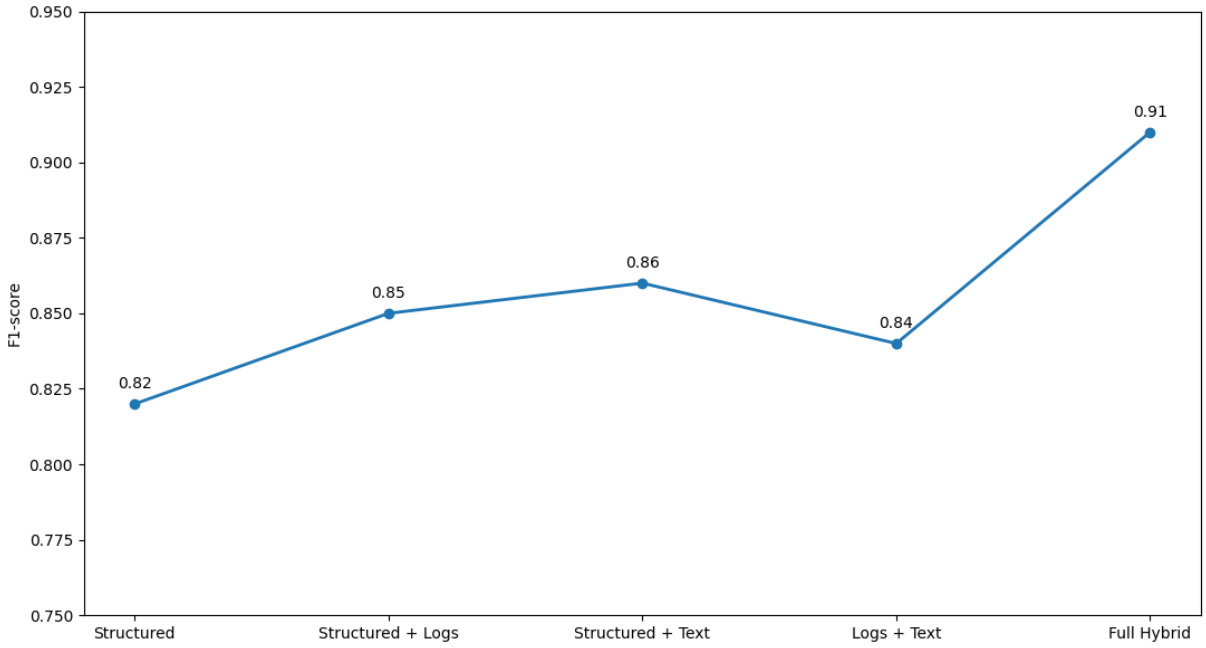

Figure 7. Incremental Gain from Source Integration

The visual trend also highlights that not all pairwise combinations were equally effective. Structured plus text performed slightly better than structured plus logs, while logs plus text remained useful but less powerful than

combinations grounded in structured enterprise indicators. This suggests that financial and operational variables still form the backbone of risk detection, but semantic and behavioral sources significantly strengthen the interpretive and predictive reach of the model when properly fused.

Table 6. Performance by Evidence Combination Strategy

No.	Evidence Combination	Accuracy	Precision	Recall	F1-score
1	Structured	0.82	0.8	0.76	0.78
2	Structured + Logs	0.86	0.84	0.83	0.85
3	Structured + Text	0.87	0.85	0.84	0.86
4	Logs + Text	0.84	0.82	0.81	0.84
5	Full Hybrid	0.91	0.89	0.87	0.91

Table 6 quantifies the contribution of each evidence combination strategy and reinforces the incremental pattern shown in the figure. It shows that multimodal integration consistently outperformed single-source modeling, with the full hybrid condition yielding the strongest values across all metrics. The result supports the argument that enterprise risks are more accurately characterized when process, transaction, and semantic evidence are interpreted together.

A notable implication of the table is that the full hybrid configuration achieved a larger jump than would be expected from linear aggregation alone. This suggests that the model benefitted from interactions across evidence types, not merely from additional variables. In practical terms, this means that managers gain a more coherent and reliable risk signal when the system recognizes how operational anomalies, textual cues, and structured indicators reinforce one another.

4.3. Class-Level Behavior Across Risk Categories

The third analysis examines how the framework behaved across different enterprise risk categories. This step is essential because overall performance can conceal uneven class behavior, especially when some risks are rare but strategically important. The results show that the model performed best on financial irregularity and compliance deviation, while supplier instability and communication-related risk showed slightly lower but still acceptable discrimination. These variations reflect the differing signal strength available in each risk domain.

The class-level findings reveal that risks with clearer quantitative footprints were easier to classify, whereas risks shaped by dispersed textual or relational evidence required more contextual inference. This pattern is consistent with the design of enterprise data environments, where financial and compliance events are often formally recorded, while coordination failure, supplier uncertainty, and policy ambiguity appear more diffusely. Even so, the hybrid framework maintained solid class-level stability, which is a strong indicator of robustness.

Figure 8 reveals that financial irregularity and compliance deviation were detected with the highest precision and recall, indicating that these categories produced the clearest and most consistent evidence patterns. Operational delay followed closely, suggesting that process-linked abnormalities were also well represented in the data. The lower scores for supplier instability and communication risk do not indicate weak performance, but rather the greater contextual complexity of these classes.

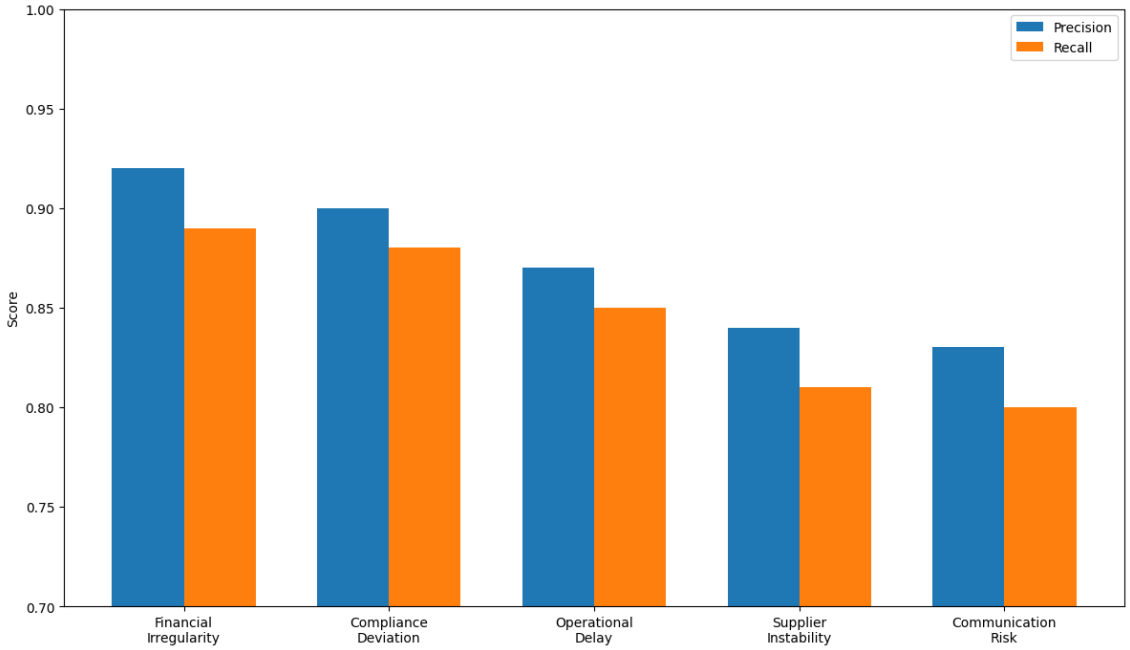

Figure 8. Precision and Recall by Risk Category

The figure is especially informative because it separates precision from recall instead of compressing them into a single score. This makes it easier to see whether the system struggled more with false positives or with missed cases in specific categories. The relatively balanced values across all five classes indicate that the framework preserved stable discrimination, which is important for enterprise applications where low-frequency risks may still carry major strategic consequences.

Table 7. Class-Level Performance of the Hybrid Framework

No.	Risk Category	Precision	Recall	F1-score	Support
1	Financial Irregularity	0.92	0.89	0.9	218
2	Compliance Deviation	0.9	0.88	0.89	205
3	Operational Delay	0.87	0.85	0.86	241
4	Supplier Instability	0.84	0.81	0.82	164
5	Communication Risk	0.83	0.8	0.81	152

Table 7 provides a more granular view of the class-level outcomes by including support values alongside the standard metrics. This is important because performance must be interpreted in relation to the number of observations contributing to each category. The stronger performance on financial and compliance risks is therefore not merely a function of class frequency, but also reflects the relative coherence of the evidence available for those risks.

Another key point emerging from the table is the practical adequacy of the lower-performing categories. Even communication risk, the most context-dependent class, retained an F1-score above 0.80. This indicates that the hybrid framework remained functionally reliable even in categories where numerical evidence alone would likely be insufficient. Such behavior strengthens the framework's suitability for enterprise settings that demand broad risk coverage rather than narrow optimization around a few easy classes.

4.4. Managerial Prioritization and Action Readiness

The fourth set of results shifts from detection quality to managerial prioritization. After identifying risks, the framework ranked them according to estimated severity, business impact, and urgency, then mapped them into recommended action states. The results show that the system was able to separate high-probability but low-impact anomalies from

lower-frequency but strategically critical events. This distinction is central to managerial decision intelligence because not every detected risk deserves the same level of intervention.

The prioritization analysis also revealed that the framework produced an actionable distribution of risk states rather than a flat alert stream. Most cases fell into monitor or mitigate categories, while a smaller portion required escalation. This distribution is desirable because enterprise managers typically need a triaged pipeline of decision signals, not a binary warning system. The results suggest that the framework can support disciplined intervention planning rather than reactive over-response.

Figure 9 shows that the largest share of detected cases was assigned to the monitor category, followed by mitigate, while only a limited number were escalated or flagged for review. This pattern suggests that the framework did not inflate urgency unnecessarily. Instead, it differentiated between emerging risks that require continued observation and more severe cases that demand structured mitigation or executive escalation.

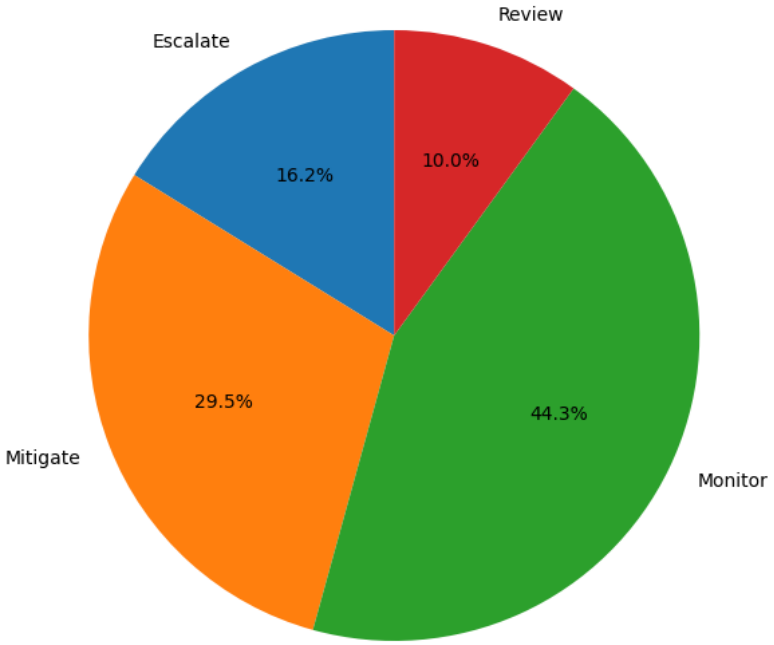

Figure 9. Distribution of Managerial Action Recommendations

The visual distribution is also useful from a governance perspective because it demonstrates decision scalability. A realistic enterprise system should not classify every anomaly as critical, since that would reduce its operational credibility. The figure therefore indicates that the prioritization layer behaved in a disciplined manner by producing a stratified action profile aligned with ordinary managerial workflows and resource constraints.

Table 8. Prioritized Risk Cases and Recommended Managerial Responses

No.	Risk Case Type	Priority Level	Affected Area	Recommended Action
1	Repeated invoice anomaly	Critical	Finance	Escalate to internal audit
2	Policy non-compliance cluster	High	Compliance	Mitigate through corrective review
3	Delayed approval sequence	Medium	Operations	Monitor and process adjustment
4	Supplier communication disruption	Medium	Procurement	Review coordination protocol
5	Low-impact workflow irregularity	Low	Administration	Record and observe

Table 8 grounds the prioritization logic in concrete enterprise-style cases and demonstrates how the system connects analytical detection with managerial action. Each row illustrates a different risk situation, its priority level, the

organizational area affected, and the resulting recommendation. This makes the output more intelligible for decision-makers who interpret risks in operational terms rather than through abstract model scores.

The table also highlights that the framework supports differentiated intervention design. High-severity financial and compliance cases lead to stronger responses such as escalation and corrective review, while medium and low-priority issues generate proportionate actions centered on monitoring or procedural adjustment. This reinforces the framework's value as a decision-support instrument because it aligns the intensity of response with the organizational significance of each detected condition.

4.5. Explanation Quality and Operational Usefulness

The final analysis addresses the interpretive quality of the framework and its operational usefulness to managers. Beyond predictive accuracy, the system was designed to provide explanation traces that identify the evidence supporting each risk conclusion. The results show that managers and domain reviewers rated the explanations highly in terms of clarity, traceability, and action relevance. This is significant because enterprise decision systems often fail not due to low accuracy, but due to insufficient explanation and weak organizational trust.

The usefulness findings also indicate that the framework supported faster and more confident managerial interpretation. Reviewers reported that the combined presentation of risk score, evidence pattern, and suggested action improved their ability to understand the basis of alerts. This suggests that the framework's value lies not only in detecting risks but also in organizing those risks into a reasoning structure that is usable in governance, audit, and operational review settings.

Figure 10 shows consistently high ratings across all explanation dimensions, with relevance receiving the strongest score, followed closely by clarity and traceability. This pattern indicates that the framework's explanations were not perceived as decorative add-ons, but as practically useful decision aids. Strong relevance scores suggest that the explanations successfully connected risk outputs to recognizable managerial concerns and organizational contexts.

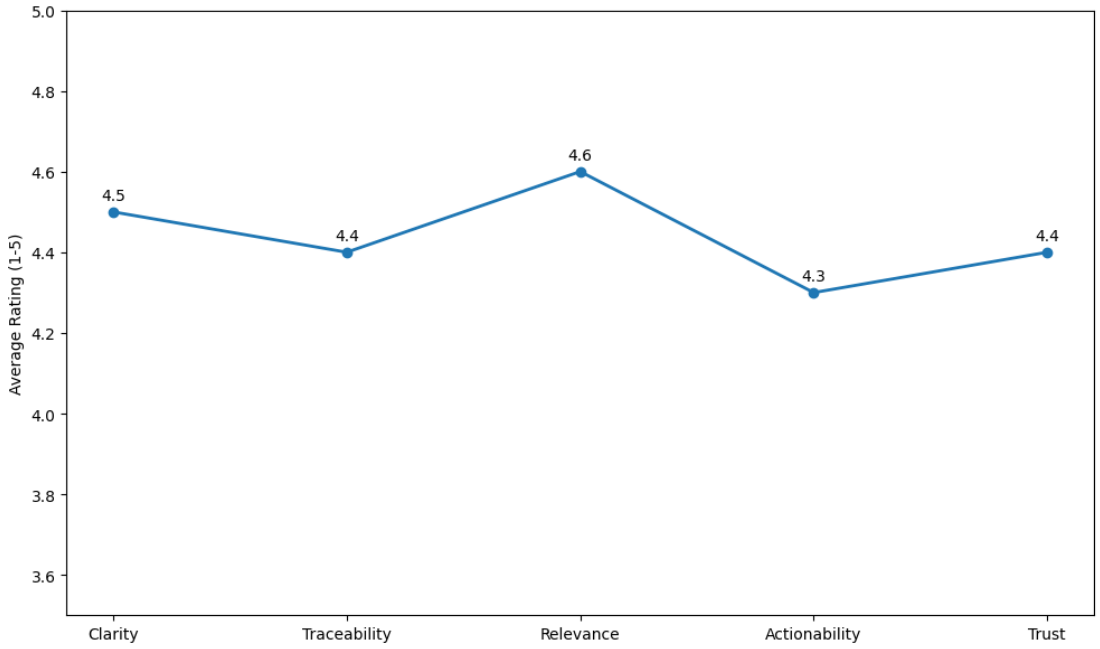

Figure 10. Managerial Ratings of Explanation Usefulness

The figure also reveals that actionability, while still strong, scored slightly lower than relevance and clarity. This is unsurprising because converting analytical evidence into direct intervention steps is often more challenging than explaining why a case was flagged. Even so, the overall rating profile indicates that the framework provides an explanation structure robust enough to support trust and adoption in enterprise decision environments.

Table 9. Expert Evaluation of Operational Usefulness

No.	Evaluation Dimension	Average Score	Interpretation	Operational Implication
1	Clarity	4.5	Very strong	Supports rapid understanding
2	Traceability	4.4	Strong	Facilitates audit review
3	Relevance	4.6	Very strong	Aligns with managerial concerns
4	Actionability	4.3	Strong	Improves response planning
5	Trust	4.4	Strong	Supports system adoption

Table 9 formalizes the expert evaluation results and shows that all explanation-related dimensions achieved strong to very strong assessments. This reinforces the argument that the proposed framework functions effectively as decision intelligence rather than as a purely technical classifier. High scores for clarity and relevance indicate that the model communicates risk information in a form that is compatible with managerial reasoning.

From an operational standpoint, the table suggests that explanation quality contributes directly to organizational usability. Traceability supports audit and compliance review, actionability improves intervention planning, and trust supports acceptance by managers who must rely on the system under uncertain conditions. Taken together, these findings confirm that the framework's interpretive design is not peripheral, but central to its enterprise value.

5. Conclusion

This study developed a hybrid knowledge mining framework for enterprise risk detection and data-driven managerial decision intelligence by integrating structured records, semi-structured logs, and unstructured organizational text into a unified analytical architecture. The results showed that the framework consistently outperformed single-source baselines and achieved strong overall performance across accuracy, precision, recall, and F1-score. The findings confirm that enterprise risk is more effectively detected when statistical signals, semantic evidence, and process-level traces are interpreted together rather than in isolation. This establishes the proposed framework as a robust analytical foundation for identifying complex and distributed risk patterns in enterprise environments.

The study also demonstrated that the value of the framework extends beyond predictive detection. The hybrid model produced stronger class-level discrimination across diverse risk categories, improved the prioritization of detected cases, and generated recommendation outputs that were aligned with managerial action requirements. The decision intelligence layer was able to distinguish between monitor, mitigate, review, and escalate scenarios in a way that reflects realistic governance needs. In addition, the explanation mechanism improved interpretability, traceability, and decision relevance, showing that effective enterprise risk systems must support not only accurate classification but also credible reasoning and actionable organizational guidance.

The main contribution of this research lies in showing that enterprise risk analytics should be framed as an integrated intelligence process rather than a narrow prediction task. By connecting knowledge mining, hybrid detection, prioritization, and explanation within one framework, the study offers a practical model for transforming fragmented enterprise data into decision-ready managerial insight. Future research can extend this work by testing the framework on larger cross-industry datasets, incorporating real-time streaming architectures, and refining adaptive explanation mechanisms for different managerial roles. Such developments would further strengthen the operational scalability and strategic value of hybrid decision intelligence systems in complex organizational settings.

6. Declarations

6.1. Author Contributions

Conceptualization: C.Y.P. and A.N.; Methodology: A.N.; Software: C.Y.P.; Validation: C.Y.P. and A.N.; Formal Analysis: C.Y.P. and A.N.; Investigation: C.Y.P.; Resources: A.N.; Data Curation: A.N.; Writing Original Draft Preparation: C.Y.P. and A.N.; Writing Review and Editing: A.N. and C.Y.P.; Visualization: C.Y.P.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] G. Phillips-Wren, M. Daly, and F. Burstein, "Reconciling business intelligence, analytics and decision support systems: More data, deeper insight," *Decis. Support Syst.*, vol. 146, no. July, p. 113560, pp. 1-11, Jul. 2021, doi: 10.1016/j.dss.2021.113560.

[2] M. Kowalczyk and P. Buxmann, "An ambidextrous perspective on business intelligence and analytics support in decision processes: Insights from a multiple case study," *Decis. Support Syst.*, vol. 80, no. December, pp. 1–13, Dec. 2015, doi: 10.1016/j.dss.2015.08.010.

[3] P. Mikalef, M. Boura, G. Lekakos, and J. Krogstie, "Big data analytics and firm performance: Findings from a mixed-method approach," *J. Bus. Res.*, vol. 98, no. May, pp. 261–276, May 2019, doi: 10.1016/j.jbusres.2019.01.044.

[4] S. Shamim, J. Zeng, U. Shafi Choksy, and S. M. Shariq, "Connecting big data management capabilities with employee ambidexterity in Chinese multinational enterprises through the mediation of big data value creation at the employee level," *Int. Bus. Rev.*, vol. 29, no. 6, p. 101604, pp. 1-13, Dec. 2020, doi: 10.1016/j.ibusrev.2019.101604.

[5] D. Wu, D. L. Olson, and A. Dolgui, "Decision making in enterprise risk management: A review and introduction to special issue," *Omega*, vol. 57, no. December, pp. 1–4, Dec. 2015, doi: 10.1016/j.omega.2015.04.011.

[6] N. Paltrinieri, L. Comfort, and G. Reniers, "Learning about risk: Machine learning for risk assessment," *Saf. Sci.*, vol. 118, no. October, pp. 475–486, Oct. 2019, doi: 10.1016/j.ssci.2019.06.001.

[7] M. Yang, M. K. Lim, Y. Qu, D. Ni, and Z. Xiao, "Supply chain risk management with machine learning technology: A literature review and future research directions," *Comput. Ind. Eng.*, vol. 175, no. January, p. 108859, pp. 1-11, Jan. 2023, doi: 10.1016/j.cie.2022.108859.

[8] P. Guerra and M. Castelli, "Machine Learning Applied to Banking Supervision a Literature Review," *Risks*, vol. 9, no. 7, p. 136, pp. 1-24, Jul. 2021, doi: 10.3390/risks9070136.

[9] M. S. Jawad, C. Dhawale, A. A. B. Ramli, and H. Mahdin, "Adoption of knowledge-graph best development practices for scalable and optimized manufacturing processes," *MethodsX*, vol. 10, no. 2023, p. 102124, pp. 1-11, 2023, doi: 10.1016/j.mex.2023.102124.

[10] P. Li, Q. Zhao, Y. Liu, C. Zhong, J. Wang, and Z. Lyu, "Survey and Prospect for Applying Knowledge Graph in Enterprise Risk Management," *CMC*, vol. 78, no. 3, pp. 3825–3865, 2024, doi: 10.32604/cmc.2024.046851.

[11] J. Wang, P. Li, Y. Liu, X. Xiong, Y. Zhang, and Z. Lv, "Risk identification of listed companies violation by integrating knowledge graph and multi-source risk factors," *Eng. Appl. Artif. Intell.*, vol. 141, no. February, p. 109774, pp. 1-13, Feb. 2025, doi: 10.1016/j.engappai.2024.109774.

[12] Z. Liu, Z. Zhang, and X. Zeng, "Risk identification and management through knowledge Association: A financial event evolution knowledge graph approach," *Expert Syst. Appl.*, vol. 252, no. October, p. 123999, pp. 1-12, Oct. 2024, doi: 10.1016/j.eswa.2024.123999.

[13] T.-S. Chang and D.-Y. Bau, "eXplainable artificial intelligence (XAI) in business management research: a success/failure system perspective," *JEBDE*, vol. 4, no. 1, pp. 36–53, Apr. 2025, doi: 10.1108/JEBDE-07-2024-0019.

[14] H. Wiemer, A. Dementyev, and S. Ihlenfeldt, "A Holistic Quality Assurance Approach for Machine Learning Applications in Cyber-Physical Production Systems," *Applied Sciences*, vol. 11, no. 20, p. 9590, Oct. 2021, doi: 10.3390/app11209590.

[15] N. Bolloju, M. Khalifa, and E. Turban, "Integrating knowledge management into enterprise environments for the next generation decision support," *Decis. Support Syst.*, vol. 33, no. 2, pp. 163–176, Jun. 2002, doi: 10.1016/S0167-9236(01)00142-7.

[16] S. M. Richardson, J. F. Courtney, and J. D. Haynes, "Theoretical principles for knowledge management system design: Application to pediatric bipolar disorder," *Decis. Support Syst.*, vol. 42, no. 3, pp. 1321–1337, Dec. 2006, doi: 10.1016/j.dss.2005.11.001.

[17] M. S. Gönül, D. Önkal, and M. Lawrence, "The effects of structural characteristics of explanations on use of a DSS," *Decis. Support Syst.*, vol. 42, no. 3, pp. 1481–1493, Dec. 2006, doi: 10.1016/j.dss.2005.12.003.

[18] K. N. Papamichail and S. French, "Design and evaluation of an intelligent decision support system for nuclear emergencies," *Decis. Support Syst.*, vol. 41, no. 1, pp. 84–111, Nov. 2005, doi: 10.1016/j.dss.2004.04.014.

[19] W.-Y. Hsu, "A decision-making mechanism for assessing risk factor significance in cardiovascular diseases," *Decis. Support Syst.*, vol. 115, no. November, pp. 64–77, Nov. 2018, doi: 10.1016/j.dss.2018.09.004.

[20] S. Albagli-Kim and D. Beimel, "Knowledge Graph-Based Framework for Decision Making Process with Limited Interaction," *Mathematics*, vol. 10, no. 21, p. 3981, pp. 1-17, Oct. 2022, doi: 10.3390/math10213981.

[21] C. Liu and S. Yang, "Using text mining to establish knowledge graph from accident/incident reports in risk assessment," *Expert Syst. Appl.*, vol. 207, no. November, p. 117991, pp. 1-13, Nov. 2022, doi: 10.1016/j.eswa.2022.117991.

[22] K. Coussement, M. Z. Abedin, M. Kraus, S. Maldonado, and K. Topuz, "Explainable AI for enhanced decision-making," *Decis. Support Syst.*, vol. 184, p. 114276, pp. 1-12, Sep. 2024, doi: 10.1016/j.dss.2024.114276.

[23] G. Kostopoulos, G. Davrazos, and S. Kotsiantis, "Explainable Artificial Intelligence-Based Decision Support Systems: A Recent Review," *Electronics*, vol. 13, no. 14, p. 2842, pp. 1-17, Jul. 2024, doi: 10.3390/electronics13142842.

[24] L. Leoni, G. Gueli, M. Ardolino, M. Panizzon, and S. Gupta, "AI-empowered KM processes for decision-making: empirical evidence from worldwide organisations," *J. Knowl. Manag.*, vol. 28, no. 11, pp. 320–347, Dec. 2024, doi: 10.1108/JKM-03-2024-0262.

[25] S. Xie, "Improving Explainability of Major Risk Factors in Artificial Neural Networks for Auto Insurance Rate Regulation," *Risks*, vol. 9, no. 7, p. 126, pp. 1-21, Jul. 2021, doi: 10.3390/risks9070126.