



International Journal of Innovative Research in Computer and Communication Engineering

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





Data Governance-Driven Trust Scoring Framework and Its Impact on Machine Learning Performance

Mr. Tirth Raval, Dr. Manisha Bharati

M.Sc. Data Science Student, Department of Technology, Savitribai Phule Pune University, Pune, Maharashtra, India

Guide, Department of Technology, Savitribai Phule Pune University, Pune, Maharashtra, India

ABSTRACT: The reliability of machine learning systems depends heavily on data trustworthiness. Traditional models assume equal reliability across observations despite missing values, duplicates, invalid entries, and label corruption. This research proposes a Data Governance-Driven Trust Scoring Framework integrated with Trust-Aware Adaptive Learning (TAAL) to evaluate and improve predictive robustness under degraded data conditions. Experimental evaluation using Random Forest and Logistic Regression demonstrated measurable performance degradation under low-trust data and modest but consistent improvements through trust-aware adaptive weighting.

KEYWORDS: Data Governance, Machine Learning, Trust Scoring, Data Quality, Adaptive Learning

I. INTRODUCTION

Machine learning has emerged as a foundational technology across modern industries, enabling intelligent decision-making in domains such as finance, healthcare, manufacturing, customer analytics, cybersecurity, and enterprise automation. The effectiveness of machine learning models depends significantly on the quality, integrity, and trustworthiness of the data used for training and evaluation [1]. Even highly sophisticated algorithms may produce unreliable predictions when trained on degraded, incomplete, inconsistent, or corrupted datasets.

In practical enterprise environments, data quality challenges are common and often unavoidable. Missing values, duplicate records, invalid entries, noisy labels, stale information, and inconsistent metadata can substantially impact predictive model performance [1]. Conventional machine learning workflows generally assume that all training observations are equally reliable, which creates a disconnect between real-world data uncertainty and model learning behavior.

Parallel to this, the field of data governance has evolved to address concerns related to data ownership, stewardship, metadata visibility, trust management, compliance, lineage, and overall data reliability [2]. Organizations increasingly invest in governance frameworks to ensure that data assets remain trustworthy, transparent, and suitable for analytical consumption [2]. Despite the growing importance of data governance, its trust-centric principles are rarely incorporated directly into machine learning training mechanisms.

This creates a meaningful research gap at the intersection of data governance and machine learning. While extensive research exists in areas such as data quality management, robust machine learning, noisy label handling, and adaptive weighting strategies, relatively limited work explicitly connects governance-inspired trust assessment with predictive learning workflows [5].

This research addresses that gap by proposing a **Data Governance-Driven Trust Scoring Framework** to quantify dataset reliability using governance-oriented trust dimensions including completeness, uniqueness, validity, consistency, timeliness, and lineage. The framework is further extended into a **Trust-Aware Adaptive Learning (TAAL)** approach, where trust-aware weighting is introduced into model training to improve robustness under imperfect data conditions.

To evaluate the proposed approach, controlled synthetic data corruption was introduced into a benchmark classification dataset to simulate realistic data quality degradation scenarios [6]. Comparative experimentation was then conducted



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

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

using Random Forest and Logistic Regression models across clean, corrupted, and trust-aware learning conditions. The objective was not only to assess the relationship between dataset trust degradation and predictive performance decline, but also to investigate whether trust-aware adaptive learning could offer measurable robustness improvements.

The significance of this work lies in demonstrating that data trust need not remain solely a passive governance metric. Instead, trust can serve as an active computational signal that influences machine learning decision-making. By bridging governance concepts with adaptive learning strategies, this research contributes a practical proof-of-concept for trust-aware intelligent systems.

II. LITERATURE SURVEY

The reliability of machine learning systems has long been recognized as strongly dependent on the quality of underlying data. Foundational research by Wang and Strong established that data quality extends beyond simple accuracy and includes multiple dimensions such as completeness, consistency, validity, and interpretability, all of which influence downstream analytical effectiveness [1]. Their work laid important conceptual foundations for understanding data trust from a broader information quality perspective.

As organizations increasingly rely on data-driven decision-making, data governance has emerged as a critical discipline for ensuring trust, accountability, stewardship, compliance, and metadata visibility across enterprise data ecosystems. Khatri and Brown emphasized that effective data governance involves structured control mechanisms that improve confidence in data assets through ownership, policies, lineage visibility, and quality assurance [2]. These governance principles are especially relevant in modern cloud-based data platforms where data trust directly influences business intelligence and operational analytics.

In machine learning research, data quality degradation has consistently been shown to negatively affect predictive performance. Missing values, noisy labels, duplicate observations, outliers, and inconsistent data representations can distort learned patterns, reduce generalization capability, and introduce biased decision boundaries. Numerous approaches have therefore focused on improving robustness under imperfect data conditions through preprocessing, noise filtering, feature engineering, imputation, and model regularization.

Label noise handling has received particular attention in classification research [5]. Frénay and Verleysen extensively studied classification under noisy labels, demonstrating that corrupted supervision can significantly reduce classifier reliability [5]. Their work explored adaptive strategies to mitigate learning degradation when label trust is uncertain. Similarly, robust machine learning methods have explored weighting mechanisms that assign differential importance to observations based on estimated reliability.

Adaptive weighting strategies are especially relevant to this research. Traditional machine learning assumes equal contribution from all observations, whereas weighted learning allows selective emphasis on higher-confidence samples. Such approaches have shown benefits in noisy environments, imbalance scenarios, and uncertain supervision contexts. However, most weighting methods derive confidence from statistical heuristics or model behavior rather than governance-inspired trust assessment.

Random Forest and Logistic Regression remain widely used baseline models in applied machine learning due to their interpretability, robustness, and established performance across structured classification tasks. Random Forest, introduced by Breiman, provides ensemble-based robustness through aggregated decision trees [3], while Logistic Regression offers a simpler probabilistic linear baseline suitable for comparative evaluation. These models are particularly appropriate for benchmarking trust-aware learning effects under controlled corruption scenarios.

Despite substantial prior work in data quality assessment, governance frameworks, and robust machine learning, a notable research gap remains. Existing studies generally address governance trust assessment and machine learning robustness as separate domains. Limited research explicitly integrates governance-derived trust metrics into adaptive machine learning workflows as an active learning signal.

This research addresses that gap by combining governance-inspired trust scoring with adaptive machine learning weighting through the proposed Trust-Aware Adaptive Learning (TAAL) framework. Unlike conventional robustness

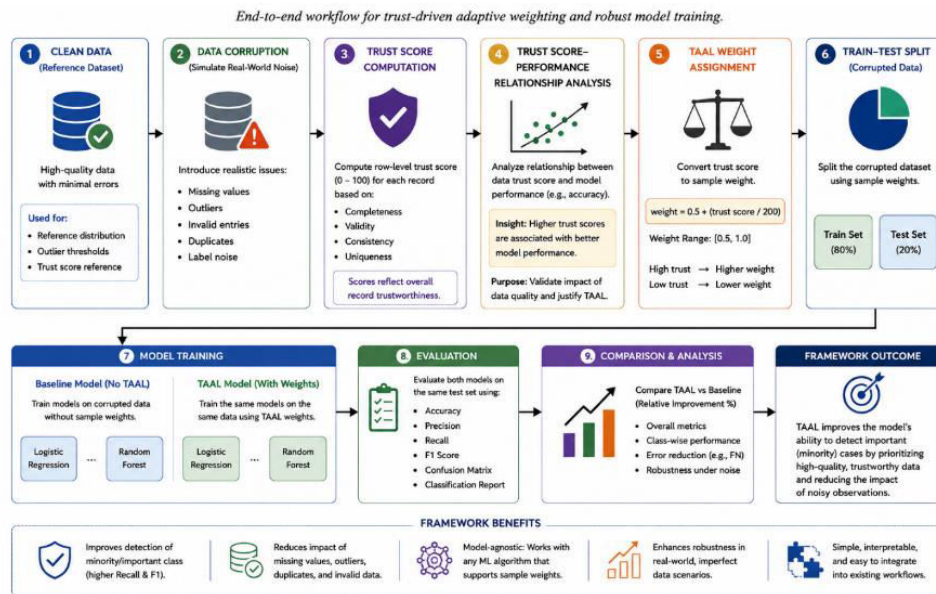


International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

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

methods that rely purely on statistical heuristics, this approach introduces governance-aligned trust assessment as a structured computational mechanism for influencing model learning behavior.

III. METHODOLOGY / APPROACH



A. Research Framework Overview

This research proposes a governance-inspired trust evaluation and adaptive machine learning framework for analyzing the impact of dataset reliability on predictive performance. The methodology combines data quality degradation simulation, trust score computation, baseline model evaluation, and trust-aware adaptive learning.

The experimental workflow consists of the following stages:

1. Benchmark dataset selection
2. Controlled synthetic data corruption
3. Governance-driven dataset trust score computation
4. Baseline machine learning model training and evaluation
5. Trust-Aware Adaptive Learning (TAAL) implementation
6. Comparative performance analysis

The overall objective is to evaluate whether governance-inspired trust assessment can function as both a dataset quality indicator and an active computational signal in machine learning workflows.

B. Dataset Selection

The **Adult Income dataset** from the UCI Machine Learning Repository was selected for experimentation due to its widespread use in supervised binary classification research [6]. The dataset contains structured demographic and socioeconomic attributes used to predict whether an individual's annual income exceeds a predefined threshold.

Representative attributes include:

- age
- workclass
- education
- marital status
- occupation
- relationship
- race



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

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

- sex
- capital gain
- capital loss
- hours per week
- native country

The target variable is binary income classification.

This dataset was selected because:

- it is a standardized benchmark dataset
- it supports reproducible experimentation
- it contains structured features suitable for trust quality analysis
- it enables controlled corruption simulation

C. Synthetic Data Corruption Strategy

To simulate realistic data quality degradation commonly observed in enterprise environments, controlled corruption was introduced into the benchmark dataset.

The corruption strategy included:

1) Missing Values

Randomly selected values were replaced with null entries to simulate incomplete data scenarios.

This directly impacts:

- completeness
- downstream model reliability

2) Duplicate Records

Duplicate observations were artificially introduced.

This simulates:

- redundant ingestion
- repeated transactional data
- poor deduplication governance

This impacts:

- uniqueness
- learning distribution fairness

3) Invalid Entries

Categorical fields were injected with invalid or unexpected values.

Examples include:

- non-domain categorical values
- corrupted labels
- schema violations

This impacts:

- validity
- model interpretability

4) Outliers

Extreme anomalous numerical values were introduced into selected features.

This simulates:

- erroneous ingestion
- sensor anomalies
- abnormal business records



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

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

This impacts:

- consistency
- model stability

5) Label Corruption

Target label noise was introduced by modifying a subset of classification labels.

This simulates:

- annotation error
- human labeling mistakes
- downstream decision contamination

This directly impacts:

- classifier learning behavior

D. Governance-Driven Trust Scoring Framework

A composite dataset trust score was designed using governance-inspired quality dimensions.

The trust framework included six dimensions:

Completeness

Measures the proportion of available non-missing values within the dataset. Higher missingness results in lower completeness trust.

Uniqueness

Measures duplicate record prevalence. Higher duplication reduces uniqueness trust.

Validity

Measures conformity to expected domain constraints. Invalid categorical entries reduce validity trust.

Consistency

Measures logical coherence and structured conformity across records. Outliers and corruption reduce consistency trust.

Timeliness

Represents data freshness.

Since benchmark datasets lack operational timestamp metadata, timeliness was represented using a fixed governance proxy score.

Lineage

Represents data origin traceability and transformation visibility.

As benchmark datasets do not provide enterprise lineage metadata, lineage was represented using a fixed governance proxy score.

E. Trust Weighting Model

The final dataset trust score was computed using weighted aggregation:

$$Trust = \sum_{i=1}^n (Dimension_i \times Weight_i)$$

Weight distribution:

Dimension	Weight
Completeness	30%
Uniqueness	20%



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

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

Validity	20%
Consistency	15%
Timeliness	10%
Lineage	5%

This weighting emphasizes dimensions most directly relevant to machine learning reliability.

F. Baseline Machine Learning Evaluation

Two supervised classification models were selected:

Random Forest

Chosen for:

- ensemble robustness
- strong structured data performance
- resilience to feature variability

Logistic Regression

Chosen for:

- interpretability
- probabilistic linear benchmarking
- comparative baseline simplicity

Dataset splitting, preprocessing, model training, and evaluation were implemented using the Scikit-learn machine learning framework [4]. Evaluation metrics included:

- Accuracy
- Precision
- Recall
- F1 Score
- Confusion Matrix Analysis

These metrics provide balanced performance assessment beyond overall accuracy.

G. Trust-Aware Adaptive Learning (TAAL)

The central methodological contribution of this research is the proposed **Trust-Aware Adaptive Learning (TAAL)** framework.

Unlike conventional training where all observations contribute equally, TAAL introduces trust-aware weighting at the row level.

Each record was assigned a trust-based reliability score derived from operational data quality indicators.

Adaptive weighting was computed as:

$$Weight = 0.5 + \frac{Trust}{200}$$

This ensures:

- low-trust observations retain reduced influence
- high-trust observations contribute proportionally more
- no observation is completely discarded

This design preserves dataset diversity while reducing the impact of degraded observations.

H. Comparative Experimental Design

Three experimental conditions were evaluated:

Scenario 1: Clean Dataset

Baseline evaluation using original benchmark data.



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

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

Scenario 2: Corrupted Dataset

Evaluation after synthetic trust degradation.

Scenario 3: TAAL Framework

Evaluation using trust-aware adaptive weighting under corrupted conditions.

Comparative analysis across these scenarios enabled measurement of:

- trust degradation impact
- predictive performance decline
- trust-aware robustness recovery

IV. RESULTS & DISCUSSION

A. Dataset Trust Score Analysis

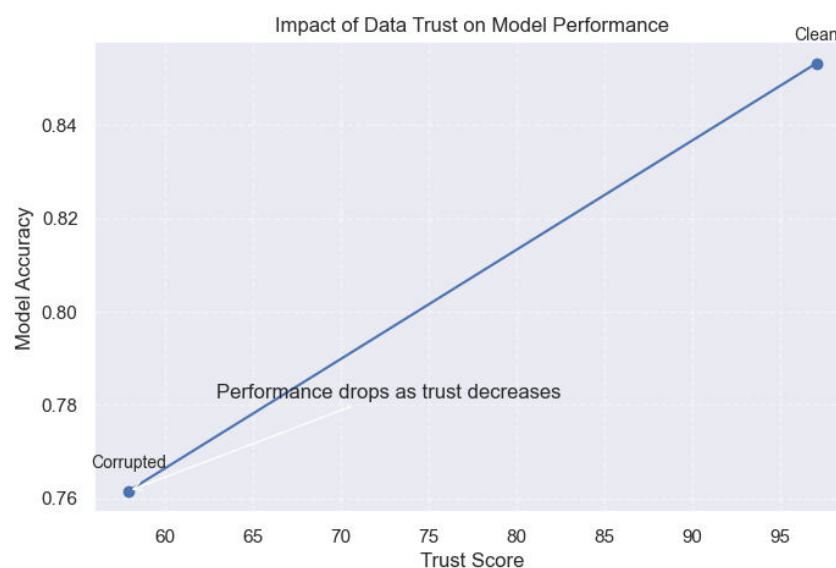
The first objective of this research was to evaluate the impact of data quality degradation on overall dataset trustworthiness.

Using the governance-driven trust scoring framework, trust scores were computed for both the clean benchmark dataset and the synthetically corrupted dataset.

The clean dataset achieved an overall trust score of approximately **97.1**, indicating high reliability, strong structural integrity, and minimal quality degradation. Following synthetic corruption, the trust score dropped significantly to approximately **57.9**, demonstrating substantial trust deterioration due to introduced missing values, duplicate observations, invalid entries, outliers, and label corruption.

This measurable trust reduction confirms that the proposed framework effectively captures the impact of data quality degradation through governance-oriented trust dimensions.

The decline in trust score also establishes the foundation for evaluating whether predictive machine learning performance is similarly affected.



B. Baseline Machine Learning Performance

Baseline experiments were conducted using both Random Forest and Logistic Regression classifiers under clean and corrupted data conditions.



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

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

Random Forest Performance

Dataset Condition	Accuracy	Precision	Recall	F1 Score
Clean Dataset	85.32%	80.40%	51.72%	62.95%
Corrupted Dataset	76.15%	72.11%	36.48%	48.45%

The Random Forest model demonstrated strong performance under clean data conditions, achieving an accuracy of **85.32%**.

Following trust degradation, performance declined significantly across all metrics:

- Accuracy decreased by approximately **9.17 percentage points**
- Precision declined
- Recall showed substantial deterioration
- F1 Score reduced noticeably

The most pronounced decline occurred in recall, suggesting reduced effectiveness in identifying positive class observations.

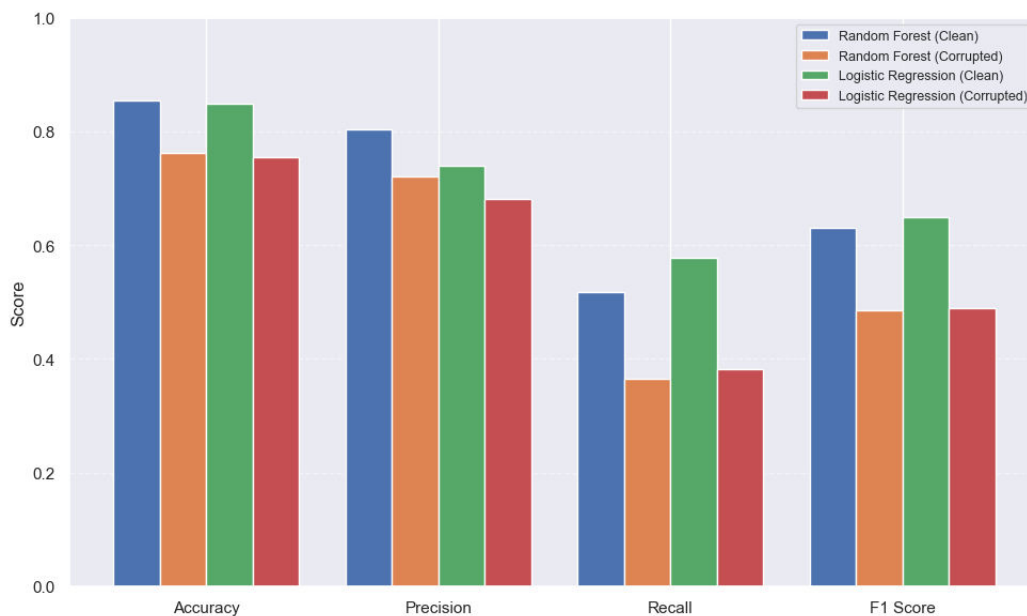
Logistic Regression Performance

Dataset Condition	Accuracy	Precision	Recall	F1 Score
Clean Dataset	84.90%	73.91%	57.75%	64.84%
Corrupted Dataset	75.52%	68.09%	38.22%	48.96%

A similar degradation pattern was observed for Logistic Regression.

Performance under corrupted data declined across all evaluation metrics, reinforcing the conclusion that reduced dataset trust negatively affects predictive reliability regardless of model architecture.

The results demonstrate a clear association between governance-driven trust degradation and machine learning performance decline.



C. Confusion Matrix Analysis

Confusion matrix analysis was performed to better understand classification behavior under degraded conditions.



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

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

Baseline Random Forest (Corrupted Dataset)

	Predicted Negative	Predicted Positive
Actual Negative	7296	487
Actual Positive	2192	1259

The confusion matrix indicates that while the model maintained strong negative class identification, performance on the positive class was weaker.

Notably:

- **2192 false negatives** were observed
- only **1259 true positives** were correctly detected

This suggests that degraded data disproportionately impacted positive class sensitivity.

Because false negatives represent missed positive observations, this reduction in recall is practically important.

D. Trust-Aware Adaptive Learning (TAAL) Performance

The proposed Trust-Aware Adaptive Learning framework introduced row-level adaptive weighting to reduce the influence of degraded observations during model training.

Comparative performance improvements are shown below.

Random Forest (TAAL vs Baseline Corrupted)

Metric	Baseline	TAAL	Relative Improvement
Accuracy	76.15%	76.25%	+0.13%
Precision	72.11%	72.46%	+0.49%
Recall	36.48%	36.60%	+0.32%
F1 Score	48.45%	48.63%	+0.38%

The TAAL framework produced measurable improvements across all evaluation metrics.

Although absolute gains were modest, improvements were consistently positive.

Logistic Regression (TAAL vs Baseline Corrupted)

Metric	Baseline	TAAL	Relative Improvement
Accuracy	75.52%	75.56%	+0.05%
Precision	68.09%	68.22%	+0.18%
Recall	38.22%	38.25%	+0.08%
F1 Score	48.96%	49.02%	+0.11%

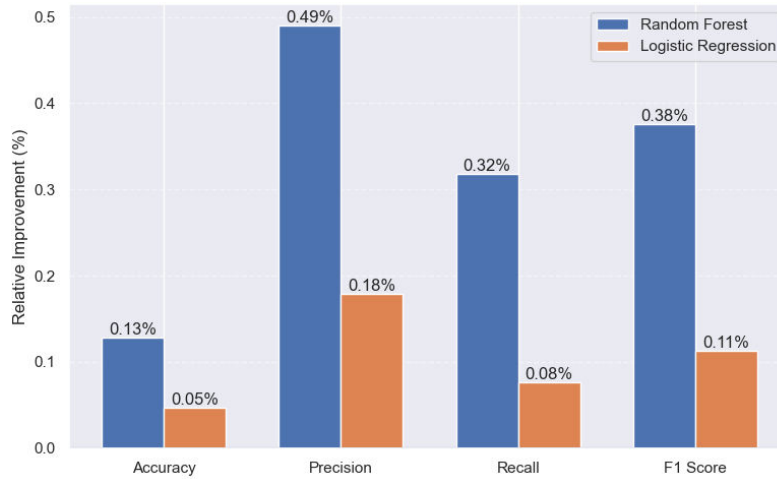
Logistic Regression also showed consistent positive improvements, although smaller than Random Forest.

This indicates that trust-aware adaptive weighting can provide architecture-independent robustness benefits.



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

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



E. TAAL Confusion Matrix Interpretation

TAAL-based confusion matrix analysis showed measurable behavioral improvement.

TAAL Random Forest

	Predicted Negative	Predicted Positive
Actual Negative	7303	480
Actual Positive	2188	1263

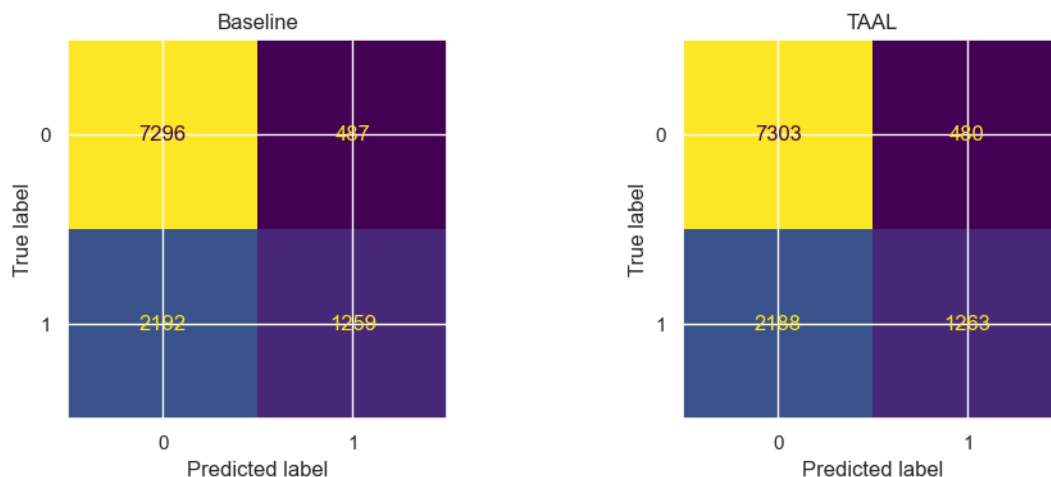
Compared with baseline:

- False positives reduced from 487 to 480
- False negatives reduced from 2192 to 2188
- True positives increased from 1259 to 1263

These changes explain the observed improvements in:

- precision
- recall
- F1 score

Although incremental, the improvement demonstrates that trust-aware weighting improves decision sensitivity without requiring architectural model redesign.





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

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

F. Discussion

The results validate the central hypothesis of this research: **data trust significantly influences machine learning performance.**

Three key findings emerged:

1. Trust degradation reduces predictive reliability

The transition from high-trust clean data to degraded corrupted data consistently reduced model performance.

2. Positive class sensitivity is especially vulnerable

Recall suffered more strongly than other metrics, indicating increased difficulty in correctly identifying positive observations under low-trust conditions.

3. Trust-aware adaptive learning improves robustness

TAAL produced consistent improvements by introducing governance-inspired trust as an active learning signal.

Unlike conventional preprocessing methods that remove suspect observations entirely, TAAL preserves all observations while proportionally reducing the influence of low-trust samples.

This makes the framework computationally lightweight, interpretable, and practically relevant.

However, limitations remain:

- single benchmark dataset
- synthetic corruption rather than natural enterprise degradation
- fixed proxy values for timeliness and lineage
- single train-test evaluation without repeated statistical validation

Despite these constraints, the findings establish a meaningful proof-of-concept at the intersection of data governance and machine learning robustness.

V. CONCLUSION

This research investigated the relationship between dataset trustworthiness and machine learning performance through a governance-inspired trust evaluation framework.

The experimental findings demonstrated that data quality degradation significantly reduces predictive reliability across machine learning models. The governance-driven trust scoring framework successfully quantified measurable trust deterioration, with the corrupted dataset showing a substantial decline compared with the clean benchmark dataset. This performance degradation was consistently reflected across evaluation metrics including accuracy, precision, recall, and F1 score.

To address this challenge, the proposed **Trust-Aware Adaptive Learning (TAAL)** framework introduced trust-informed adaptive weighting into the training process. Experimental evaluation showed modest but consistently positive improvements across both Random Forest and Logistic Regression models. Confusion matrix analysis further confirmed reduced classification errors and improved positive class sensitivity.

A key contribution of this work lies in demonstrating that data trust can function not only as a passive governance quality indicator but also as an active computational signal for improving machine learning robustness. By integrating governance-oriented trust assessment with adaptive learning behavior, this research establishes a practical proof-of-concept at the intersection of data governance and intelligent analytics.

Although limitations exist, including reliance on synthetic corruption, benchmark datasets, proxy governance metadata, and single train-test evaluation, the proposed framework provides a meaningful foundation for future trust-aware machine learning research.

Future work may extend this approach through:

- evaluation across multiple datasets and domains
- integration with real enterprise metadata platforms



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

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

- dynamic trust optimization strategies
- broader model experimentation
- statistical validation through repeated cross-validation

Overall, this study demonstrates that governance-driven trust assessment can meaningfully contribute to more robust and interpretable machine learning systems.

REFERENCES

- [1] Wang, R. Y., & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12(4), 5–33.
- [2] Khatri, V., & Brown, C. V. (2010). Designing Data Governance. *Communications of the ACM*, 53(1), 148–152.
- [3] Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32.
- [4] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- [5] Frénay, B., & Verleysen, M. (2014). Classification in the Presence of Label Noise: A Survey. *IEEE Transactions on Neural Networks and Learning Systems*, 25(5), 845–869.
- [6] UCI Machine Learning Repository. Adult Income Dataset.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



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