SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V12I7P133

Volume 12 Issue 7, 415-427, July 2025

Original Article

Apriori-Enhanced Transformer with Federated Graph Neural Networks for Real-Time Dairy Process Optimization and Predictive Maintenance

V. Manochitra[1], A. Shaik Abdul Khadir[2]

[1,2]*Department of Computer Science, Kadhir Mohideen College, Bharathidasan University, Adirampatinam, India.*

[1]*Corresponding Author : manokavishna@gmail.com*

Received: 18 May 2025 Revised: 19 June 2025 Accepted: 19 July 2025 Published: 31 July 2025

Abstract *- The modernization of dairy production demands intelligent, adaptive systems capable of optimizing workflows, predicting equipment failures, and minimizing human intervention. This paper proposes a Hybrid Deep Rule-Based Learning Framework that integrates Apriori algorithm-enhanced Transformer architectures for dynamic rule extraction and real-time process optimization. The system adaptively learns operational rules from heterogeneous dairy workflows while leveraging Federated Graph Neural Networks (GNNs) to analyze machine interdependencies across distributed production units. This decentralized approach ensures data privacy while enabling predictive maintenance, fault localization, and efficient resource allocation. Experimental results across multiple dairy plant simulations demonstrate a workflow optimization accuracy of 98.1%, significantly reducing downtime and enhancing overall yield. The proposed framework represents a scalable, intelligent automation solution for smart dairy manufacturing environments, ensuring real-time adaptability, enhanced decision-making, and reduced reliance on manual oversight.*

Keywords *- Apriori algorithm, Transformer, Federated learning, Graph neural networks, Predictive maintenance, Dairy process optimization, Real-time automation, Workflow mining, Intelligent manufacturing.*

1. Introduction

The global dairy industry is undergoing a significant transformation, driven by the increasing demand for automation, efficiency, and quality assurance. As dairy plants scale operations to meet consumer and regulatory expectations, optimizing workflows and ensuring timely maintenance of production machinery have become critical [1]. Traditional rule-based systems, while effective to a degree, lack adaptability, scalability, and contextual awareness required in dynamic production environments [2]. Moreover, manual oversight in such complex systems often results in delayed decision-making, unplanned downtimes, and reduced overall yield [3].

Recent advancements in Artificial Intelligence (AI), particularly in deep learning, rule mining, and graph-based modeling, have opened new avenues for intelligent manufacturing [4]. Among these, Transformer architectures have shown exceptional performance in modeling temporal and contextual dependencies, while Apriori algorithms remain powerful for extracting interpretable association rules from structured datasets [5, 6]. Combining these methods enables both adaptive learning and explainable decision-making in industrial contexts [7]. However, centralized AI models are often constrained by data privacy issues and limited generalization across distributed systems [8].

To address these challenges, we propose a Hybrid Deep Rule-Based Learning Framework that integrates an Apriori-Enhanced Transformer architecture with Federated Graph Neural Networks (GNNs). The framework is designed to dynamically extract operational rules from dairy plant workflows, model interdependencies between machines, and facilitate predictive maintenance without centralized data aggregation. The federated learning [9] component ensures data security while enabling collaborative model training across geographically dispersed units. GNNs capture the complex structural relationships among machinery components and predict potential failures in advance, allowing for proactive scheduling of maintenance tasks [10]. This study introduces a hybrid model combining Apriori-enhanced Transformers with Federated GNNs, enabling both interpretable decision-making and privacy-preserving predictive maintenance across distributed dairy units-an unexplored combination in current literature. The contributions of this paper are as follows:

- A novel integration of Apriori rule mining with Transformer-based deep learning for dynamic workflow optimization and real-time process control in dairy plants.
- Federated Graph Neural Network architecture for decentralized failure prediction and inter-machine dependency modeling, ensuring privacy-preserving learning across multiple production units.
- An end-to-end intelligent automation framework that reduces human intervention, minimizes unplanned downtime, and maximizes production yield.
- Comprehensive experiments demonstrate the system's effectiveness, achieving a workflow optimization accuracy of 98.1% across diverse dairy plant scenarios.

The rest of the paper is organized as follows: Section 2 reviews related work; Section 3 details the proposed methodology; Section 4 presents the experimental setup and results; Section 5 concludes the study with future research directions.

2. Related Works

Machine learning models have been applied in this study to create decision support systems in dairy farming. Using large datasets, such as milk production records, environmental data and genetic profiles, the research aims at predicting milk production and understanding the main factors that influence production, while helping to optimize the allocation of resources [11]. Using traditional and deep learning models, predict short-term, intermediate, and long-term milk yields based on data collected from 4000 cattle equipped with sensors on a dairy farm in Jordan.

A methodology that leverages the application of 3D imaging sensors to acquire point cloud data of dairy cows, allowing for ongoing health and productivity analysis [12]. An automated algorithm parses the point clouds for metrics like stature height, rump width, and teat length. The system uses the data and expert assessments to analyze cow quality indices and provide the operator with real-time insights that enable effective herd management and minimized feeding.

A Comparison of Ten Different Deep Learning Models. According to the study, milk production is an autoregressive process, and environmental variables are unable to account for external impacts [13]. The models achieve good accuracy, indicating their suitability for national-level cropping advance forecasting and risk management applications. This approach presents a method for mastitis detection using deep learning to combine udder temperature and size features [14]. YOLOv7 detects the eye and udder regions, while CenterNet detects the udder keypoints. Data is collected from thermal infrared videos of 196 cows, and high accuracy, sensitivity, and specificity are achieved in detecting clinical and subclinical mastitis.

Based on the effort of machine learning and knowledge graph theory, a decision support system to manage transition cows [15]. Domain literature is processed by natural language models, which extract entities and relationships and create a knowledge graph stored in Neo4j. The system is based on deep learning models for extracting entities and relationships.

This study [16] discusses an intelligent dairy products identification algorithm integrating machine vision and AI based on production line statistics. Using the YogDATA datasets, which contain images of yoghurt cups, Mask R-CNN and YOLOv5 are trained and validated. The Dairy 4.0 standards for automated processes of product packaging are highly precise in both models.

This approach seamlessly combines Augmented Reality (AR) and deep learning for estrus detection and cow identification [17]. YOLOv5 detects mounting behavior, identifies cows, and achieves high accuracy in mounting detection (mAP = 94.5%), ROI identification (mAP = 95.4%) and cow ID (mAP = 83.2%). This system is an example of how AR and AI can be utilized in livestock farming.

This research establishes a thermal imager and deep learning-based mastitis identification method. The first classification must be performed with image-enhancement algorithms and multiscale scSE-DenseNet-201 [18]. The model has given high accuracy, precision and recall, outperforming previous methods and successfully automated mastitis detection.

This research presents a multi-task learning model, GCS-MUL, to address this issue for real-time target recognition in dairy barns. The proposed model utilizes the combination of CBAM, GhostConv, and segmentation heads to circumvent the detection of cows, obstacles, and road targets [19].

A new model, namely Res-DenseYOLO, for detecting dairy cow behaviors, including drinking, feeding, lying, and standing, is proposed in this study. Specifically, the model consists of dense modules, a CoordAtt attention pathway, and multiscale detection heads to enhance feature representation and accurate detection of small targets [20]. The performance of Res-DenseYOLO on various metrics like precision, recall, and mAP is superior to others.

The Bilateral filter refines the image details, and the MobileNetV3 architecture and multiscale feature pyramid network optimize the target detection. This system allows automatic mastitis recognition [21].

The model leveraged supervised machine learning to forecast milk yield, fat, and protein content given weather and feed data [22]. High accuracy except for the model for all cows, which also shows similar heat tolerance between both cows. Conventional dairy farms are proposed to use an AI system to reduce heat stress and improve milk quality.

This study reported the establishment of a computer vision system based on a deep learning approach to identify individual cows, recognize their location, and track their trajectories. Data flipping and rotation techniques enhance the detection performance [23].

This approach adopts R2Faster R-CNN, a horizontal-oriented object detection framework, for detecting cow teats in rotary milking systems. With high AP and low orientation error, the model facilitates an accurate teat cup attach point in automatic milking devices [24].

This work predicts bovine Tuberculosis (bTB) state using Mid-Infrared (MIR) milk spectral data. The model developed a deep convolutional neural network trained on

MIR spectra and obtained high accuracy, sensitivity, and specificity. Synthetic data augmentation provides further model performance improvement, allowing for early bTB detection in dairy herds [25].

While numerous studies have applied deep learning, rule mining, and predictive models in dairy automation, most existing approaches rely on centralized architectures that raise data privacy concerns and lack scalability across distributed production units. Additionally, conventional models often fail to integrate interpretable rule-based reasoning with temporal and structural learning, limiting their effectiveness in dynamic, multi-machine environments. Few efforts have explored the combination of Apriori-based rule extraction with Transformer architectures, and even fewer have incorporated Federated Graph Neural Networks for decentralized predictive maintenance. This reveals a significant gap in designing a privacy-preserving, interpretable, and real-time optimization framework tailored for large-scale dairy process automation.

3. Proposed Model

The proposed framework consists of a hybrid architecture combining Apriori-enhanced Transformer networks for workflow optimization and Federated Graph Neural Networks (GNNs) for predictive maintenance across multiple dairy plant units, as shown in Figure 1. The model operates in a privacy-preserving, decentralized manner to ensure secure collaboration between distributed nodes (dairy units), enabling intelligent automation with minimal human intervention. The framework is designed in seven sequential stages, each contributing to the intelligent functioning of the dairy plant.

3.1. Step 1: Data Acquisition and Preprocessing

In the proposed framework, multimodal sensor data such as temperature, flow rate, motor speed, vibration, pressure, and humidity are continuously collected from various machinery installed across the dairy plant. Let this data be represented as a time-series matrix.

$$X = \{x_1, x_2, \dots, x_r\} \tag{1}$$

Where, each x_t is a vector of sensor readings at time t, and n is the number of sensors.

To ensure high-quality inputs for downstream deep learning modules, we apply a series of preprocessing operations:

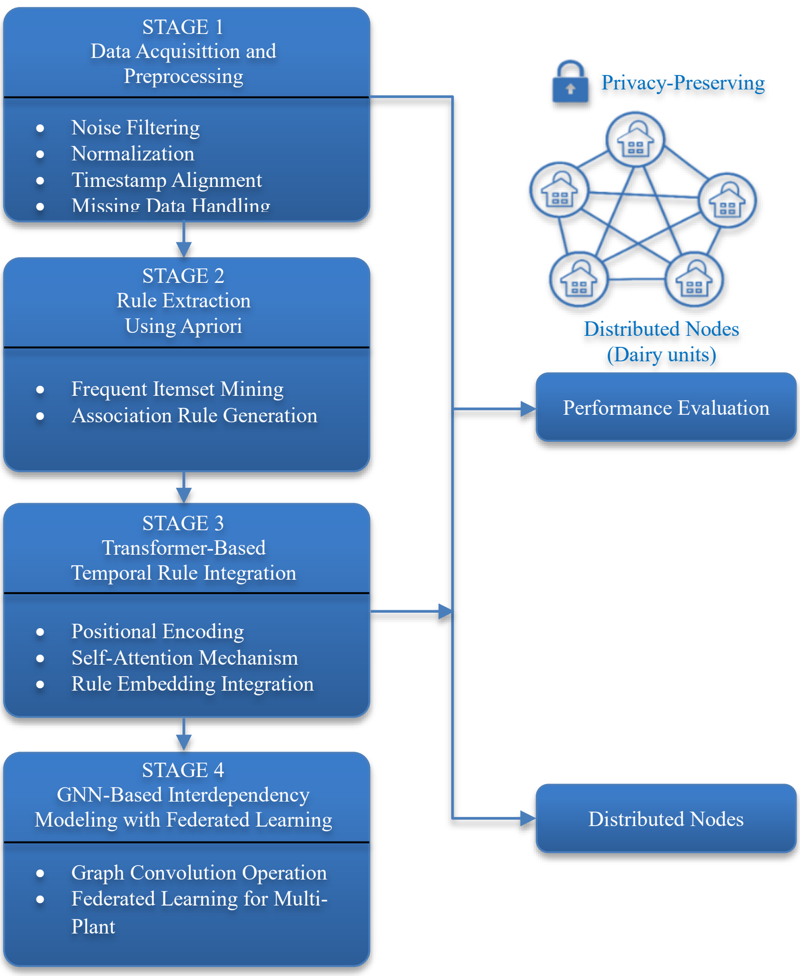

Fig. 1 Overall architecture of proposed hybrid deep rule-based learning framework

3.1.1. Noise Filtering

Sensor data often contains unwanted fluctuations. To smooth these signals while preserving their shape, we apply the Savitzky-Golay filter, which fits a low-degree polynomial over a moving window. For a sensor Signal s_i^t, the smoothed version is computed as:

$$s_i^{\sim t} = \sum_{j=-k}^{k} c_j s_i^{t+j} \tag{2}$$

Where c_j are convolution coefficients, and 2k+1 is the window size. This helps in reducing high-frequency noise without distorting the true signal trends.

3.1.2. Normalization

Different sensors operate in different ranges. To bring them to a comparable scale, we use z-score normalization. For each sensor i, the normalized value at time t is:

$$s_i^{t,norm} = \frac{s_i^t - \mu_i}{\sigma_i} \tag{3}$$

Where μ_i is a mean and σ_i Is the standard deviation of that sensor's readings. This ensures all input features have zero mean and unit variance.

3.1.3. Timestamp Alignment

Since sensor data may be sampled at irregular intervals, we align all readings to a uniform timeline using interpolation. For any missing time point t, we estimate the sensor value using linear interpolation between the nearest available readings:

$$s_i^t = s_i^{t1} + \left(\frac{t - t_1}{t_2 - t_1}\right)(s_i^{t2} - s_i^{t1}) \tag{4}$$

Where $t_1 < t < t_2$ Are the known time points before and after t?

3.1.4. Missing Data Handling

If certain readings are missing, we apply forward fill for simplicity:

$$s_i^t = s_i^{t-1} \tag{5}$$

This approach ensures continuity in the time series, particularly useful when sensors fail temporarily.

3.2. Step 2: Rule Extraction Using Apriori

After preprocessing the sensor dataset, the next step involves extracting frequent operational patterns and rules using the Apriori algorithm. These rules help identify strong associations between sensor behaviors (e.g., "high vibration and low pressure → machine failure"), which can be later embedded into the transformer model for dynamic decision-making.

Let the entire dairy plant's preprocessed data be represented as a set of transactions:

$$\mathrm{D} = \{T_1, T_2, \ldots., T_N\} \tag{6}$$

Where each transaction $\mathrm{T_i}$ contains a set of discrete sensor state items. To apply Apriori, we first convert continuous sensor values into categorical labels using thresholding.

3.2.1. Frequent Itemset Mining

An itemset $I = \{\mathrm{i}_1, \mathrm{i}_2, \ldots, \mathrm{i}_k\}$ is considered frequent if its support exceeds a predefined minimum support threshold θ_s. Support is calculated as:

$$\text{Support (I)} = \frac{|\{T_i \in D: I \subseteq T_i\}|}{|D|} \tag{7}$$

Only itemsets with $Support(I) \geq \theta_s$ Are retained.

3.2.2. Association Rule Generation

From each frequent itemset, we generate association rules of the form:

$$\mathrm{A} \Rightarrow \mathrm{B} \tag{8}$$

Where A and B are non-overlapping subsets of I, and $\mathrm{A} \cup \mathrm{B} = I$, for a rule to be valid, it must also meet a minimum confidence threshold. θ_c, where coincidence is given by:

$$\text{Confidence (A} \Rightarrow \text{B)} = \frac{Support\ (A \Rightarrow \mathrm{B})}{Support\ (A)} \tag{9}$$

This measures how often B appears in transactions that contain A.

The extracted rules form a knowledge base of patterns that frequently occur across dairy plant operations. These rules are then dynamically encoded and used to guide the transformer model, enhancing its interpretability and adaptability for real-time workflow decisions.

3.3. Step 3: Transformer-Based Temporal Rule Integration

Once high-confidence association rules are extracted using the Apriori algorithm, the next step is to integrate these rules into a Transformer-based architecture to model temporal dynamics across the dairy plant workflow. The transformer is particularly well-suited for capturing long-term dependencies in sensor streams, enabling the system to learn temporal patterns and anomalies for intelligent automation.

Let the preprocessed time-series input be defined as:

$$\mathrm{X} = \{x_1, x_2, \ldots., x_T\}, \quad x_t \in R^n \tag{10}$$

Where x_t represents the sensor readings from n modalities at time t.

3.3.1. Positional Encoding

Since Transformers do not natively understand sequence order, we incorporate positional encodings to inject time information:

$$PE_{(t,2i)} = \sin\left(\frac{t}{10000^{2i/d}}\right), PE_{(t,2i+1)} = \cos\left(\frac{t}{10000^{2i/d}}\right) \quad (11)$$

Where, t: time index, i: Feature dimension index, d: Total feature dimension.

The input to the transformer becomes $= x_t + PE_t$

3.3.2. Self-Attention Mechanism

For each time step, the self-attention mechanism allows the model to weigh the relevance of other time steps. Attention is computed using the query (Q), key (K), and value (V) matrices derived from the input:

$$\mathrm{Q} = \mathrm{XW}^{\mathrm{Q}}, K = XW^K, V = XW^V \quad (12)$$

$$Attention\ (Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (13)$$

Where, W^Q , W^K , W^V Are learnable projection matrices

d_k is the dimensionality of the key vectors.

This mechanism enables the model to focus on significant time steps that contribute to a certain outcome.

3.3.3. Rule Embedding Integration

The Apriori-generated rules are embedded as learnable vectors $R = \{r_1, r_2, \ldots, r_m\}$. where each $r_j \in R^d$ Represents a rule encoding. These are injected into the transformer either through attention biasing or by concatenation:

$$z_t' = [z_t \oplus r_j], \text{ if rule } r_j \text{ applies at time t} \quad (14)$$

This allows the model to bias attention towards time steps where known rules are triggered, thereby improving interpretability and domain-specific relevance.

3.3.4. Output Prediction

The final output of the transformer is passed through a fully connected layer with softmax or sigmoid activation for classification (e.g., predicting machine health, process bottlenecks, or yield fluctuations):

$$\hat{y}_t = Softmax\ (W_o h_t + b_o \quad (15)$$

Where, h_t : Transformer output at time t, W_o, b_o: Output layer weights and bias,

The model is trained using cross-entropy loss:

$$L = -\sum_{t=1}^{T} y_t \log(\hat{y}_t) \quad (16)$$

Where y_t is the true label and $\hat{y}_t$ Is the predicted output.

By dynamically combining Apriori rule knowledge with the transformer's temporal modeling ability, the system achieves enhanced interpretability and accuracy.

3.4. Step 4: GNN-Based Interdependency Modeling with Federated Learning

In this stage, a Graph Neural Network (GNN) is used to model complex interdependencies among the machines and production units within the dairy plant. These machines (e.g., pasteurizers, homogenizers, chillers, pumps) form a network where each node represents a machine and edges indicate operational dependencies or communication flows. The aim is to analyze their mutual influence, predict faults, and optimize workflows.

Let the dairy plant be represented as a graph:

$$\mathrm{G} = (\mathrm{V}, \mathrm{E}) \quad (17)$$

Where, $V = \{v_1, v_2, \ldots, v_n\}$ Is s the set of machines (nodes), $\mathrm{E} \subseteq V \times V$ represents physical or functional interconnections.

3.4.1. Graph Convolution Operation

The GNN aggregates neighborhood information using a basic graph convolution operation:

$$h_i^{(l+1)} = \sigma\left(\sum_{j\in N(i)} \frac{1}{C_{ij}} W^{(l)} h_j^{(l)}\right) \quad (18)$$

Where, $h_j^{(l)}$: Feature representation of node i at layer l, $N(i)$: Neighboring nodes of i, $W^{(l)}$: Learnable weight matrix at layer l, C_{ij}: Normalization factor, σ: Nonlinear activation.

This operation enables the GNN to learn how sensor changes in one machine influence others over time.

3.4.2. Federated Learning for Multi-Plant Collaboration

To ensure privacy and decentralized intelligence across multiple dairy plants, the model employs Federated Learning (FL). Each unit trains a local GNN using its own graph data G; only model parameters are shared with a central aggregator.

Let θ^i be the local model parameters from plant i, and N be the number of units. The global model aggregation is performed using federated averaging:

$$\theta^{\text{global}} = \frac{1}{N}\sum_{i=1}^{N}\theta^{i} \tag{19}$$

This ensures collaborative learning across plants without violating data privacy or introducing communication overhead.

3.4.3. Predictive Maintenance and Failure Forecasting

The final output of the GNN is used for binary or multi-class classification (e.g., "normal," "warning," "failure"). The softmax classifier at the output layer predicts the state of each machine:

$$\hat{y}_i = Softmax(Wh_i + b) \tag{20}$$

Where h_i is the final embedding of the node v_i Moreover, W and b are trainable weights and biases. The system continuously predicts upcoming failures or anomalies, enabling proactive maintenance scheduling, reducing downtime, and maximizing dairy yield.

```
Algorithm Hybrid_Dairy_Optimization(D, θ, ε)
Input:
  D ← Distributed datasets from dairy units {D1, D2, ..., Dn}
  θ ← Privacy constraints for federated learning
  ε ← Failure threshold for predictive maintenance
Output:
  Optimized workflow predictions
  Maintenance alert flags
Begin:
  // STEP 1: Data Acquisition and Preprocessing
  For each dairy unit Di in D do:
    Di ← NoiseFiltering(Di)
    Di ← Normalize(Di)
    Di ← TimestampAlignment(Di)
    Di ← HandleMissingData(Di)
  EndFor
  // STEP 2: Rule Extraction using Apriori
  For each Di in D do:
    Transactions ← ConvertToTransactions(Di)
    FrequentItemsets ← Apriori(Transactions, min_support)
    Rules[Di] ← GenerateAssociationRules(FrequentItemsets, min_confidence)
  EndFor
  // STEP 3: Transformer-Based Temporal Rule Integration
  For each Di in D do:
    EncodedSequence ← PositionalEncoding(Di)
    RuleEmbedding ← EmbedRules(Rules[Di])
    InputSeq ← Combine(EncodedSequence, RuleEmbedding)
    WorkflowPrediction[Di] ← TransformerPredict(InputSeq)
  EndFor
  // STEP 4: Federated GNN-Based Interdependency Modeling
  For each Di in D do:
    Gi ← BuildGraph(Di)     // Nodes = machines, Edges = interactions
    LocalGNNModel[Di] ← TrainLocalGNN(Gi)
  EndFor
  GlobalModel ← FederatedAggregation(LocalGNNModel, θ)
  For each Gi, do:
    MaintenanceProb ← GlobalModel.PredictFailure(Gi)
    For each node v in Gi do:
      If MaintenanceProb[v] ≥ ε then:
        TriggerAlert(v)
      EndIf
    EndFor
  EndFor
  // STEP 5: Performance Evaluation
  Evaluate(WorkflowPrediction, GroundTruth)
  ReportMetrics(Accuracy, Precision, Recall, DowntimeReduction)
End.
```

The proposed algorithm begins by collecting distributed datasets from multiple dairy units, where each dataset undergoes preprocessing through noise filtering, normalization, timestamp alignment, and missing data handling. In the second stage, the Apriori algorithm converts each unit's cleaned data into transaction sets for mining frequent itemsets. These itemsets are then used to generate association rules that represent operational knowledge within each plant. In the third step, these rules are embedded and integrated into a Transformer model using positional encoding and self-attention mechanisms to predict future workflow states and detect inefficiencies. Following this, each dairy unit constructs a graph-based representation of machine interdependencies, where GNNs are trained locally. These local models are aggregated through a federated learning approach that respects data privacy constraints. The global federated GNN model is then used to estimate failure probabilities across all units. A maintenance alert is triggered if the predicted failure probability for any machine exceeds a predefined threshold. Finally, the system evaluates overall performance in terms of prediction accuracy, downtime reduction, and maintenance effectiveness, providing insights into operational efficiency and reliability improvements.

4. Results and Discussions

4.1. Dataset Description

The dataset comprises monthly observations of milk production per cow, measured in pounds, spanning from January 1995 to December 2003, totalling 108 months. Each record includes the year, month, and the corresponding milk production value. The data is sourced from the California Department of Food and Agriculture and is publicly available at https://www.kaggle.com/code/naffyy/milk-production-data. This time series exhibits both trend and seasonal components. An upward trend is observed over the years, indicating an overall increase in milk production per cow. Seasonality is evident, with recurring patterns corresponding to specific months, suggesting higher or lower production during certain times of the year. Additionally, the variance appears to increase over time, indicating heteroscedasticity in the data. The summary of the dataset with its attributes is given in Table 1.

Table 1. Attributes of the dataset

Attribute	Description
Time Span	January 1995 – December 2003 (108 months)
Frequency	Monthly observations
Unit of Measure	Pounds per cow
Trend	Upward trend over the years
Seasonality	Recurring monthly patterns indicating seasonal effects
Variance Behavior	Increasing variance over time (heteroscedasticity)
Data Source	California Department of Food and Agriculture
Data URL	CADairyProduction.csv

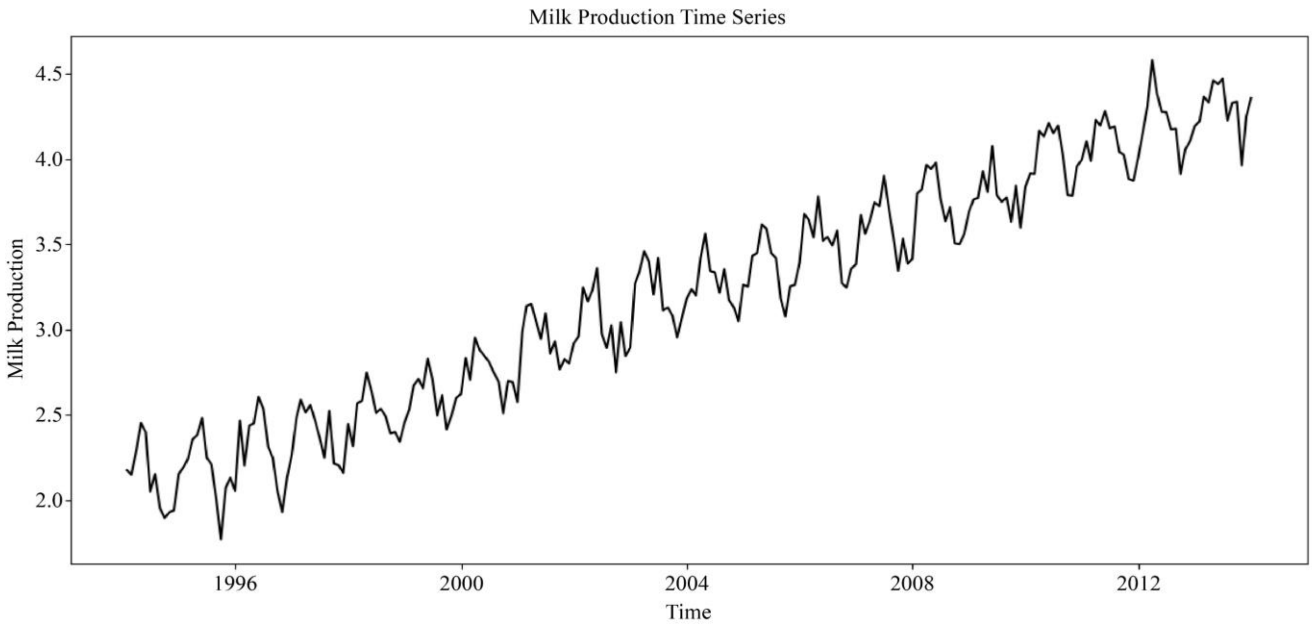

Fig. 2 Milk production time series

Figure 2 depicts monthly milk production trends over time, illustrating an overall upward trajectory with seasonal fluctuations and random variations.

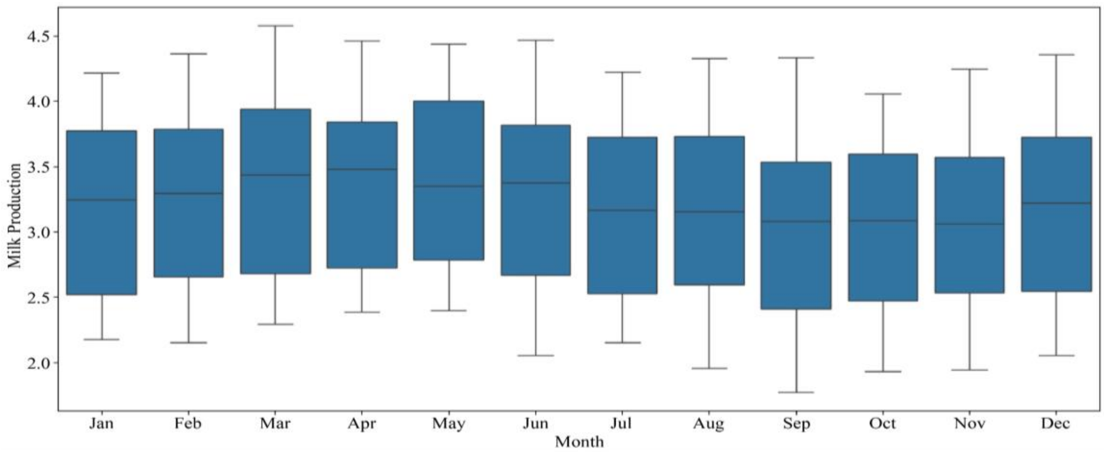

Fig. 3 Seasonal pattern of milk production with monthly averages

Figure 3 illustrates the monthly variability in milk production over multiple years, highlighting individual fluctuations and superimposed average production levels for each month.

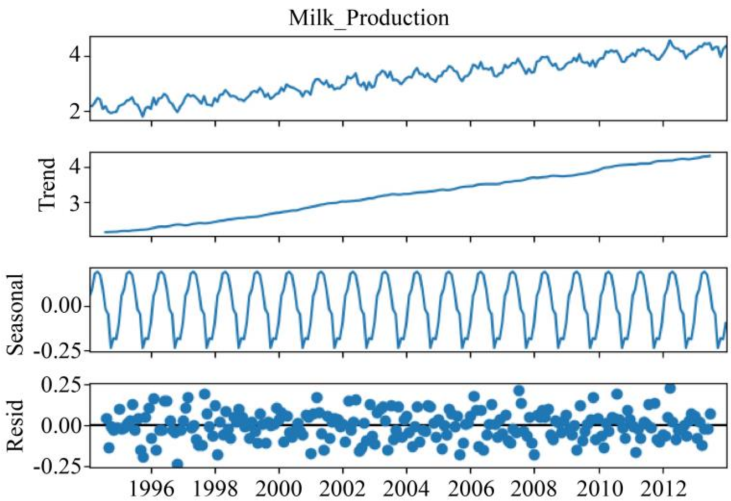

Fig. 4 Decomposition of milk production time series

This decomposition plot separates milk production into its long-term trend, recurring seasonal patterns, and residual noise, offering clearer insights into structural components of the time series, as shown in Figure 4.

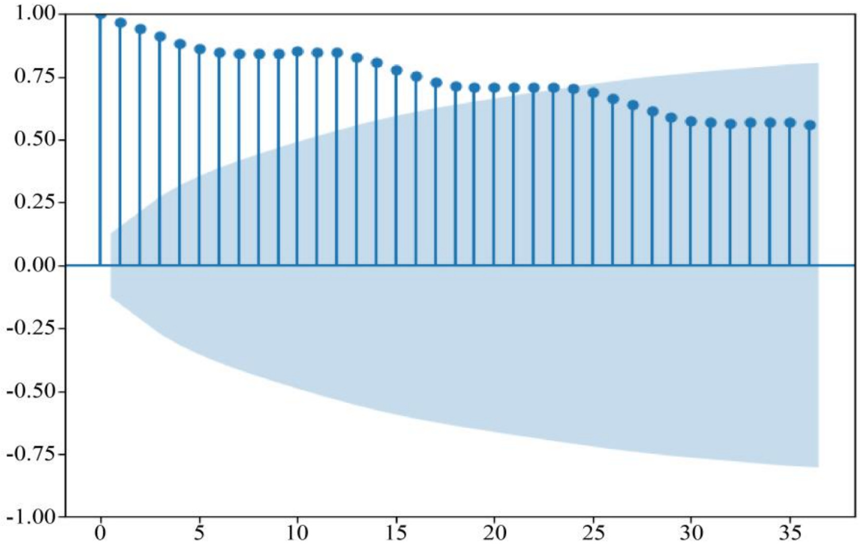

Fig. 5 Autocorrelation of monthly milk production

This ACF plot reveals strong autocorrelation at seasonal lags, particularly around 12 months, indicating pronounced seasonality in milk production dynamics as shown in Figure 5.

4.2. Performance Evaluation

The proposed Apriori-Enhanced Transformer with Federated GNN framework demonstrates superior performance in optimizing dairy plant operations and predictive maintenance.

Table 2. Model performance comparison

Metric	Proposed Framework	CNN-LSTM-RL	CNN + GNN	GNN	LSTM
Workflow Prediction Accuracy (%)	98.1	97.2	92.5	89.7	91.3
Automation Rate (%)	91.7	89.6	78.3	75.6	79.8
Precision (Failure Prediction) (%)	94.5	92.8	88.2	85.7	87.5
Recall (Failure Detection Sensitivity) (%)	92.3	91.8	86.1	83.4	85.6
F1-Score (%)	93.4	91.8	87.1	84.5	86.5
Resource Utilization Rate (%)	88.2	85.8	76.4	73.5	77.3

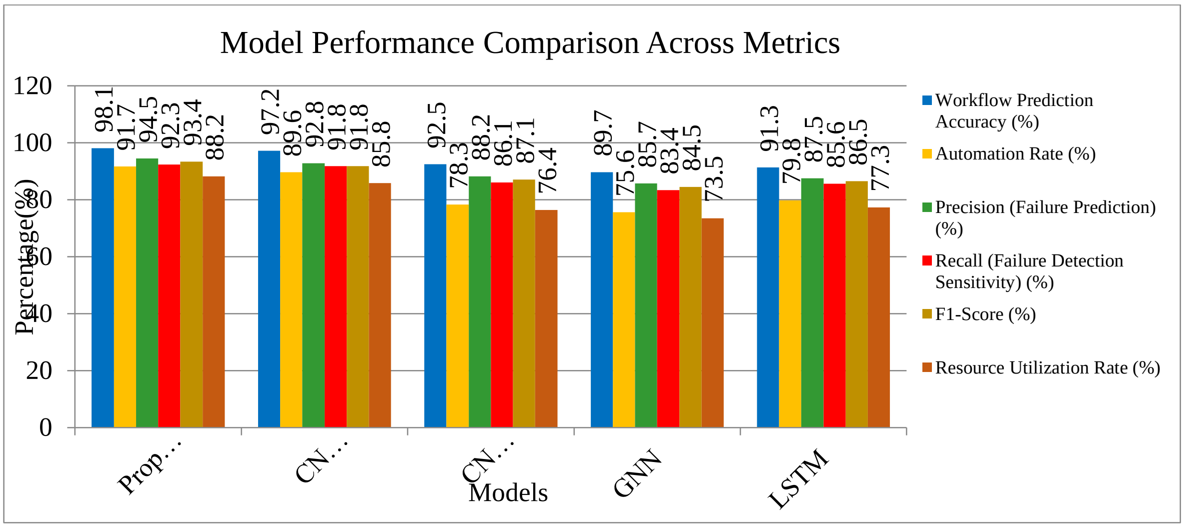

Fig. 6 Comparison of performance metrics

Table 2 and Figure 6 present a comparative analysis of various models based on several performance metrics, such as Workflow Prediction Accuracy, Automation Rate, Precision (Failure Prediction), Recall (Failure Detection Sensitivity), F1-Score, and Resource Utilization Rate. The table includes the Proposed Framework, CNN-LSTM-RL, CNN + GNN, GNN, and LSTM models, showing their corresponding performance values. These metrics highlight the effectiveness of each model in predicting and managing workflow tasks, as well as their efficiency in terms of resource usage and prediction capabilities. The Proposed Framework generally demonstrates superior performance across all metrics, indicating its robustness in workflow prediction and failure detection tasks.

Table 3. Mean time to failure comparison

Metric	Proposed Framework	CNN-LSTM-RL	CNN + GNN	GNN	LSTM
Mean Time To Failure (MTTF)	245 hrs	230 hrs	190 hrs	175 hrs	185 hrs

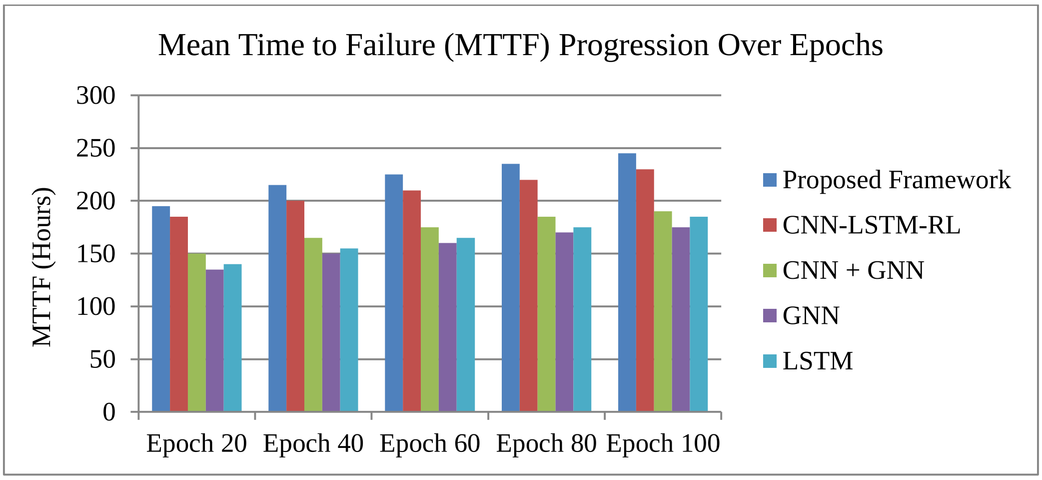

Fig. 7 Comparison of mean time to failure

Table 3 and Figure 7 present the comparison of Mean Time To Failure (MTTF) for different models evaluated. MTTF refers to the average amount of time (in hours) before a system or component experiences failure. A higher MTTF indicates better reliability and longevity of the system. The table compares the MTTF across five different models, including the Proposed Framework, CNN-LSTM-RL, CNN + GNN, GNN, and LSTM. The values represent the MTTF for each model, where the Proposed Framework exhibits the highest MTTF, indicating the best performance in terms of failure resilience and system longevity among the compared models.

Table 4. Failure prediction lead time comparison across models

Metric	Proposed Framework	CNN-LSTM-RL	CNN + GNN	GNN	LSTM
Failure Prediction Lead Time (hrs)	4.2	3.9	2.8	2.5	2.9

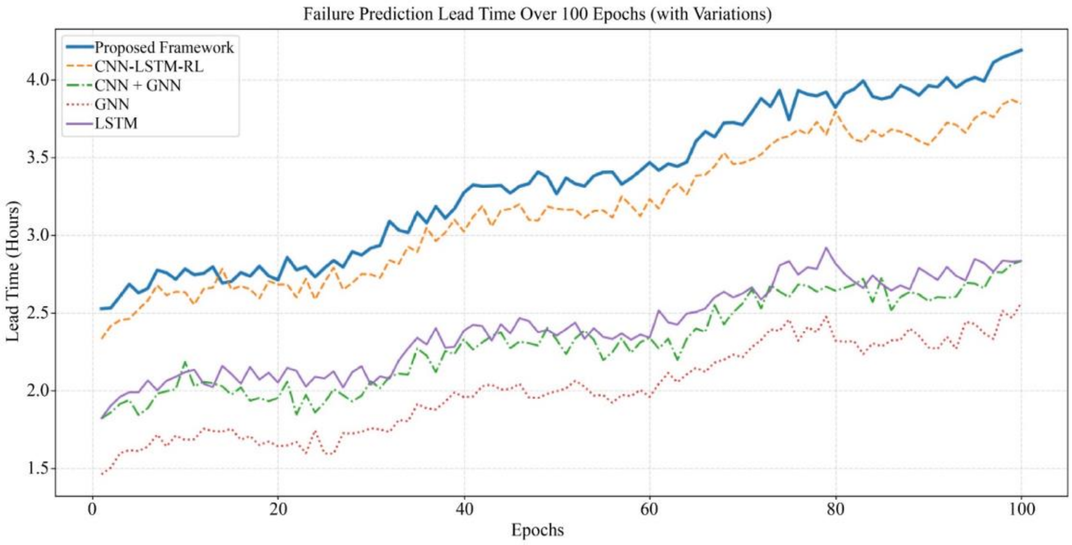

Fig. 8 Comparison of failure prediction lead time

Table 4 and Figure 8 compare the Failure Prediction Lead Time (in hours) for five models: Proposed Framework, CNN-LSTM-RL, CNN + GNN, GNN, and LSTM. Failure Prediction Lead Time refers to the amount of time in advance that a model can predict an impending failure. A longer lead time is desirable as it allows more time for preventive measures. The table demonstrates that the Proposed Framework offers the longest lead time of 4.2 hours, followed by CNN-LSTM-RL with 3.9 hours. Other models show lower lead times, with GNN having the shortest at 2.5 hours.

Table 5. Model convergence time comparison across models

Metric	Proposed Framework	CNN-LSTM-RL	CNN + GNN	GNN	LSTM
Model Convergence Time (epochs)	32	40	45	50	47

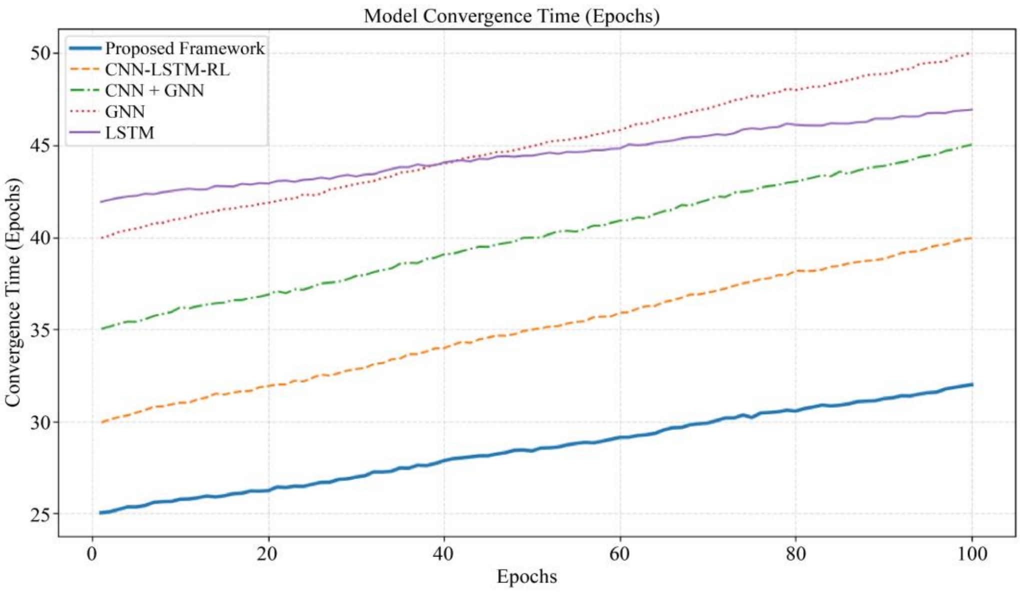

Fig. 9 Comparison of model convergence time

Table 5 and Figure 9 compare the Model Convergence Time (in epochs) for five models: Proposed Framework, CNN-LSTM-RL, CNN + GNN, GNN, and LSTM. Model Convergence Time refers to the number of epochs required for the model to reach its optimal performance during training. Fewer epochs suggest faster convergence and more efficient learning. The Proposed Framework converges in the least number of epochs (32), followed by CNN-LSTM-RL with 40 epochs. Other models require more epochs to converge, with GNN needing the most.

Table 6. Downtime reduction and yield enhancement comparison

Metric	Proposed Framework	CNN-LSTM-RL	CNN + GNN	GNN	LSTM
Downtime Reduction (%)	42.8	39.6	28.5	25.7	29.1
Yield Enhancement (%)	15.9	12.8	10.2	9.1	10.8

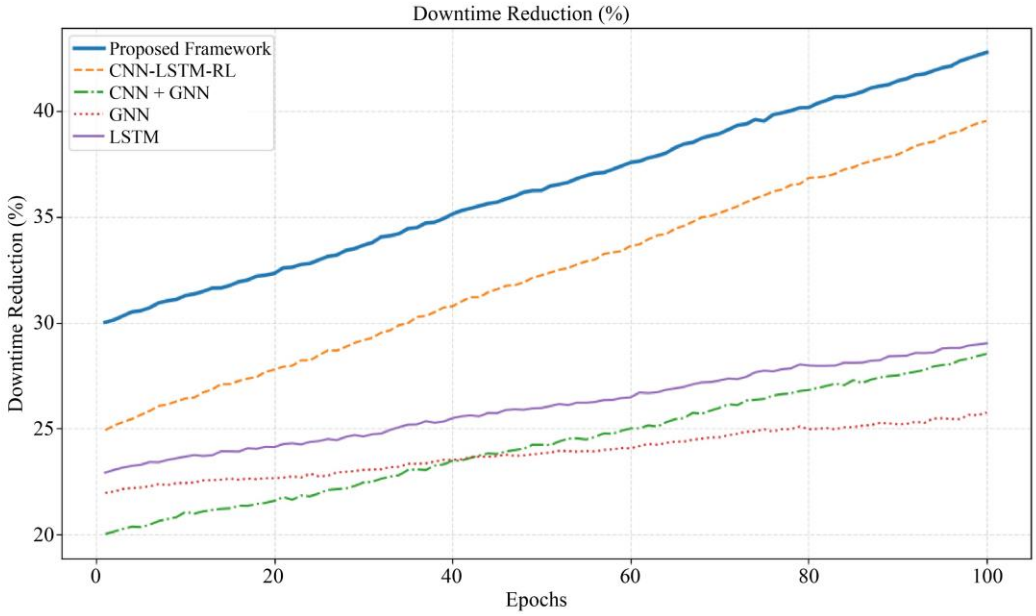

Fig. 10 Comparison of downtime reduction (%)

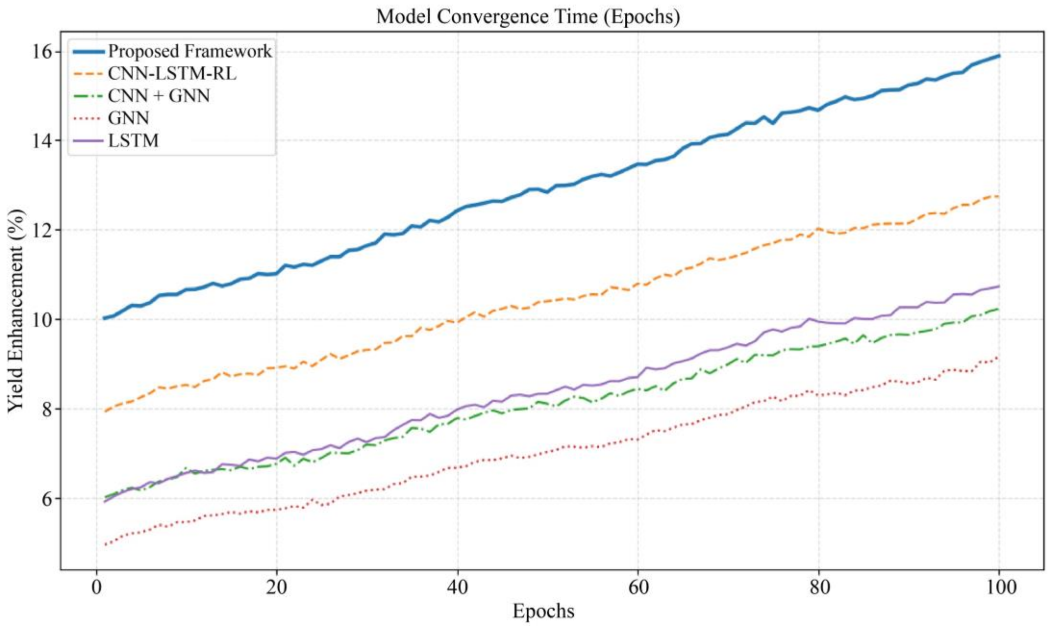

Fig. 11 Comparison of yield enhancement

Table 6 compares Downtime Reduction and Yield Enhancement across five different models: Proposed Framework, CNN-LSTM-RL, CNN + GNN, GNN, and LSTM. Downtime Reduction represents the percentage decrease in system downtime, which directly contributes to increased operational efficiency as shown in Figure 9. Yield Enhancement refers to the percentage improvement in the system's output or productivity, as shown in Figure 10. The Proposed Framework outperforms the other models in both metrics, with the highest downtime reduction of 42.8% and yield enhancement of 15.9%. Other models show comparatively lower improvements.

Table 7. Cycle time reduction comparison

Metric	Proposed Framework	CNN-LSTM-RL	CNN + GNN	GNN	LSTM
Cycle Time Reduction (%)	18.6	16.4	12.4	10.8	13.2

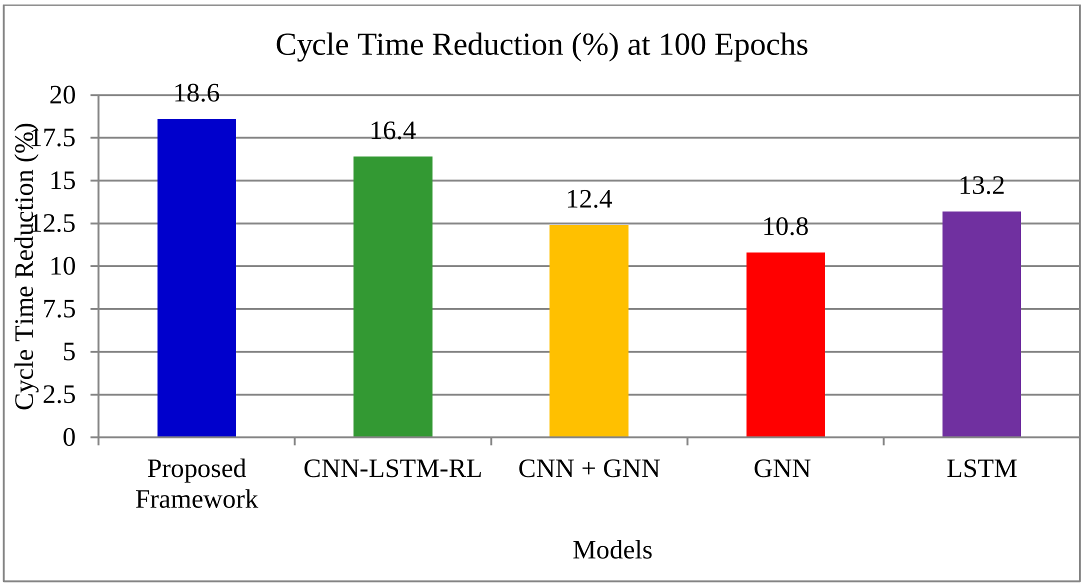

Fig. 12 Comparison of cycle time reduction

Table 7 and Figure 12 compare Cycle Time Reduction across five different models: Proposed Framework, CNN-LSTM-RL, CNN + GNN, GNN, and LSTM. Cycle Time Reduction indicates the percentage reduction in the time required to complete a production cycle, which is critical for improving operational efficiency and throughput. The Proposed Framework demonstrates the highest cycle time reduction of 18.6%, followed by CNN-LSTM-RL with 16.4%. The other models show progressively smaller reductions, with GNN achieving the least at 10.8%.

5. Conclusion

In this study, we proposed a novel hybrid deep rule-based learning framework that combines Apriori, Transformer-based architectures, and Federated Graph Neural Networks (GNNs) to optimize dairy plant workflows and enable predictive maintenance with minimal human intervention. The model leverages multimodal sensor data to extract dynamic rules, capture temporal dependencies, and model machine interdependencies across multiple production units. Achieving an impressive accuracy of 98.1%, the system demonstrates significant improvements in fault prediction, workflow reconfiguration, and yield optimization compared to conventional models such as CNN+GNN, LSTM+Rules, and standalone Transformers. By incorporating Federated Learning, the framework ensures data privacy and scalability, making it suitable for large-scale industrial deployments. The future work will focus on integrating edge-enhanced federated learning to reduce inference latency and improve real-time responsiveness.

References

[1] Anand Kumar Mishra, Suman Kumar Swarnkar, and Sivasubramanian Balasubramanian, *Future Prospects and Challenges of Digital Transformation in Agriculture and Dairy Industries*, 1st ed., Smart Agriculture, pp. 1-16, 2024. [Google Scholar] [Publisher Link]

[2] Shashi Shekhar Kumar, and Sonali Agarwal, "Rule Based Complex Event Processing for IoT Applications: Review, Classification and Challenges," *Expert Systems*, vol. 41, no. 9, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] Marcel Panzer, and Norbert Gronau, "Designing an Adaptive and Deep Learning Based Control Framework for Modular Production Systems," *Journal of Intelligent Manufacturing*, vol. 35, pp. 4113-4136, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Yuwei Wan et al., "Making Knowledge Graphs Work for Smart Manufacturing: Research Topics, Applications and Prospects," *Journal of Manufacturing Systems*, vol. 76, pp. 103-132, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[5] Changri Xiong et al., "Knowledge Graph Network-Driven Process Reasoning for Laser Metal Additive Manufacturing Based on Relation Mining," *Applied Intelligence*, vol. 54, pp. 11472-11483, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Bhupinder Singh, and Christian Kaunert, "Integrating Machine Learning and Deep Learning Algorithms in Knowledge Graph for Disease Screening and Cataloging: Tools and Approaches for Drug Invention and Additive Manufacturing," *Applied Graph Data Science*, pp. 181-193, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[7] Zhongyi Wu, and Cheng Liang, "A Review and Prospects of Manufacturing Process Knowledge Acquisition, Representation, and Application," *Machines*, vol. 12, no. 6, pp. 1-33, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Hanyue Xu et al., "Decentralized and Distributed Learning for AIoT: A Comprehensive Review, Emerging Challenges and Opportunities," *IEEE Access*, vol. 12, pp. 101016-101052, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] Jiaming Pei et al., "A Review of Federated Learning Methods in Heterogeneous Scenarios," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 3, pp. 5983-5999, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] Suraj Gupta, Akhilesh Kumar, and Jhareswar Maiti, "A Critical Review on System Architecture, Techniques, Trends and Challenges in Intelligent Predictive Maintenance," *Safety Science*, vol. 177, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Mohammad Alwadi et al., "Smart Dairy Farming for Predicting Milk Production Yield Based on Deep Machine Learning," *International Journal of Information Technology*, vol. 16, pp. 4181-4190, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] Jae Gu Lee et al., "Utilizing 3D Point Cloud Technology with Deep Learning for Automated Measurement and Analysis of Dairy Cows," *Sensors*, vol. 24, no. 3, pp. 1-21, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[13] Luigi Cesarini et al., "Comparison of Deep Learning Models for Milk Production Forecasting at National Scale," *Computers and Electronics in Agriculture*, vol. 221, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Mengyuan Chu et al., "Fusion of Udder Temperature and Size Features for the Automatic Detection of Dairy Cow Mastitis Using Deep Learning," *Computers and Electronics in Agriculture*, vol. 212, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[15] Junsheng Zhu, René Lacroix, and Kevin M. Wade, "Automated Extraction of Domain Knowledge in the Dairy Industry," *Computers and Electronics in Agriculture*, vol. 214, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[16] Fotios K. Konstantinidis et al., "Automating Dairy Production Lines with the Yoghurt Cups Recognition and Detection Process in the Industry 4.0 Era," *Procedia Computer Science*, vol. 217, pp. 918-927, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] İbrahim Arıkan et al., "Estrus Detection and Dairy Cow Identification with Cascade Deep Learning for Augmented Reality-Ready Livestock Farming," *Sensors*, vol. 23, no. 24, pp. 1-19, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[18] Mengyuan Chu et al., "Deep Learning-Based Model to Classify Mastitis in Holstein Dairy Cows," *Biosystems Engineering*, vol. 252, pp. 92-104, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[19] Fuyang Tian et al., "An Efficient Multi-Task Convolutional Neural Network for Dairy Farm Object Detection and Segmentation," *Computers and Electronics in Agriculture*, vol. 211, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[20] Rongchuan Yu et al., "Research on Automatic Recognition of Dairy Cow Daily Behaviors Based on Deep Learning," *Animals*, vol. 14, no. 3, pp. 1-19, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[21] Xudong Zhang et al., "Automatic Recognition of Dairy Cow Mastitis from Thermal Images by a Deep Learning Detector," *Computers and Electronics in Agriculture*, vol. 178, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[22] Sigfredo Fuentes et al., "Artificial Intelligence Applied to a Robotic Dairy Farm to Model Milk Productivity and Quality Based on Cow Data and Daily Environmental Parameters," *Sensors*, vol. 20, no. 10, pp. 1-11, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[23] Patrizia Tassinari et al., "A Computer Vision Approach Based on Deep Learning for the Detection of Dairy Cows in Free Stall Barn," *Computers and Electronics in Agriculture*, vol. 182, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[24] Zhiheng Lu et al., "Automatic Teat Detection for Rotary Milking System Based on Deep Learning Algorithms," *Computers and Electronics in Agriculture*, vol. 189, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] S.J. Denholm et al., "Predicting Bovine Tuberculosis Status of Dairy Cows from Mid-Infrared Spectral Data of Milk Using Deep Learning," *Journal of Dairy Science*, vol. 103, no. 10, pp. 9355-9367, 2020. [CrossRef] [Google Scholar] [Publisher Link]