ISBN: 978-628-96613-3-0. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2026.1.1.2584

Early Warning Predictive Models for Ammonia Refrigeration: A Machine Learning Approach in Brewing Processes

Giancarlo Machaca Chaname; Bachelor, Adolfo Jorge Prado Ventocilla, Ms
[1] *Universidad Tecnológica del Perú, Peru, u21312762@utp.edu.pe*
[2] *Universidad Tecnológica del Perú, Peru, c16555@utp.edu.pe*

Abstract— Ammonia (NH3) refrigeration systems are critical infrastructures in the brewing industry, yet they pose significant safety risks due to the toxic nature of the refrigerant. Traditional monitoring systems, reliant on fixed-threshold alarms (typically set at 25 ppm), often fail to detect precursor anomalies in dynamic operational environments, leading to delayed responses and increased safety risk. This study proposes an early warning predictive framework leveraging Internet of Things (IoT) sensor fusion and Machine Learning (ML) algorithms. By integrating chemical concentration data with operational variables—such as discharge pressure, temperature, vibration, and compressor state—two models were developed and evaluated: Random Forest (RF) and Long Short-Term Memory (LSTM) networks. Results demonstrate that the LSTM model, optimized for temporal sequence analysis, achieved a Recall of 0.74 (vs 0.97 for RF) and an AUC of 0.999 (LSTM)vs 0.911 (RF), providing an average early warning lead time of 42 minutesbefore the safety threshold (25 ppm) is breached. Furthermore, explainability analysis using SHAP (SHapley Additive exPlanations) confirmed that pressure drops and thermal variability are key predictors of leakage events. This research bridges the gap between theoretical ML applications and real-world industrial safety.

Keywords—Ammonia refrigeration, predictive maintenance, LSTM, Random Forest, early warning systems.

I. Introduction

In the brewing industry refrigeration systems play a critical role in maintaining product quality and safety across all process stages Ref. [1], Ref. [3]. Ammonia (NH_3) is commonly employed in industrial refrigeration due to its high thermodynamic efficiency, low environmental impact, and favorable refrigeration cycle properties when compared with synthetic alternatives Ref. [1], Ref. [2]. However, the toxicity and corrosive nature of ammonia introduce significant safety hazards, which have been documented in multiple industrial incidents involving worker exposure, operational shutdowns, and regulatory penalties.

Traditional industrial monitoring systems are largely based on fixed-threshold alarms (e.g., 25 ppm alarms for NH_3), which do not consider interdependencies between process variables or dynamic system behavior. These methods often fail to provide early alerts before dangerous conditions develop, resulting in late responses and elevated risk to personnel and equipment.

Therefore, the main contributions of this work are summarized as follows:

- Identify the critical sensory and operational variables of the ammonia refrigeration system that influence the early detection of anomalous conditions.
- Analyze temporal patterns and relationships between NH_3 concentrations and operational variables through data preprocessing and sensor fusion techniques.
- Implement Machine Learning models based on Random Forest and Long Short-Term Memory (LSTM) neural networks for the detection and prediction of risk events in the refrigeration system.
- Evaluate the performance of the proposed models using classification and prediction metrics oriented toward early warning capability and operational reliability.
- Compare the effectiveness of the developed models against traditional fixed-threshold-based approaches in terms of anticipation, accuracy, and robustness.

While previous studies have successfully applied ML to HVAC systems with

synthetic refrigerants Ref. [8-11], to the best of our knowledge, no prior work has systematically evaluated and compared RF and LSTM models for ammoniabased refrigeration in operational brewery environments using real sensor fusion data.

This study hypothesizes that the fusion of IoT sensor data and operational process variables, combined with ML algorithms, can improve the early detection of leaks and unsafe conditions in NH_3 refrigeration systems. The primary contribution of this work is the evaluation and comparison of predictive models in a realistic industrial context, targeting improved safety outcomes and predictive maintenance capabilities.

II. METHODOLOGY

A. Data Acquisition

Data were collected using an IoT-based monitoring system deployed in the ammonia refrigeration system of a brewing plant located in Ate. Electrochemical MQ-135 sensors were selected for their fast response time (<10 s) and cost-effectiveness and were cross-validated against a reference industrial-grade NH_3 detector, Dräger Polytron 8100, yielding a mean absolute error of ±3 ppm within the range of 0–50 ppm. The sensors operated within their calibrated range of 0-100 ppm, with the safety threshold set at 25 ppm as per industrial safety standards. While industrial-grade sensors offer higher nominal accuracy, the objective was to demonstrate that low-cost IoT sensors paired with ML models can provide reliable early warnings in an industrial context.

Key operational variables were obtained from the plant SCADA and historian systems, including compressor state, operating cycles, discharge and suction pressures, and temperature. Sensor readings and operational logs were recorded continuously with a sampling interval between 1 and 5 seconds, capturing both normal and abnormal behavior. All data were stored in structured time-series format for subsequent preprocessing and model training.

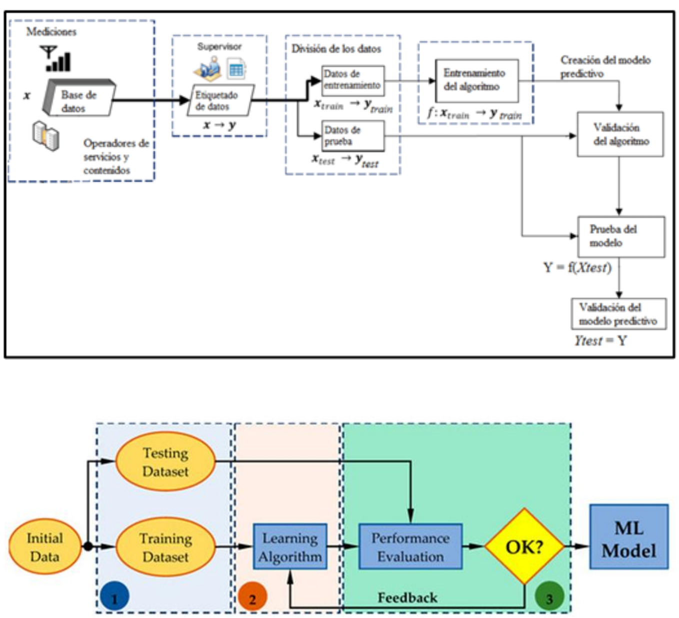

Fig. 1. Machine Learning process block diagram

B. Data Preprocessing

Before training the Machine Learning models, the collected data went through a rigorous preprocessing stage. This step included removing outliers, filling in missing values, normalizing the variables, and aligning all data streams using common time stamps.

$$x_{norm} = \frac{x - \mu}{\sigma} \qquad x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Fig. 2. Normalization formula

After that, sensor data and operational data were combined into a single dataset. Information from NH_3 sensors was merged with key process variables from the industrial refrigeration system. This data fusion approach made it possible to build a unified dataset that accurately represents the dynamic behavior of the refrigeration system, overcoming the limitations of methods based only on isolated measurements or fixed threshold alarms Ref. [5].

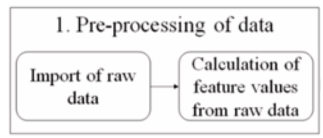

Fig. 3. Data pre-processing input and output

Code:

```
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train_raw)
X_test_scaled = scaler.transform(X_test_raw)

# Convertir a DataFrame para mantener nombres de columnas (útil para SHAP)
X_train_df = pd.DataFrame(X_train_scaled, columns=features)
X_test_df = pd.DataFrame(X_test_scaled, columns=features)
print(f"Train shape: {X_train_scaled.shape}, Test shape: {X_test_scaled.shape}")
```

Dataset description

Table 1 summarizes the dataset used for training and validation. The dataset contains time-synchronized NH_3 concentration signals and operational variables that describe the dynamics of the refrigeration process. Data-driven monitoring in refrigeration and HVAC systems is typically based on multivariate time-series datasets collected from sensors and operational sources [4]–[6].

Table 1. Dataset Summary

Aspect	Detail
Data collection period	January 2024 – January 2025
Sampling frequency	1–5 s
Total samples	~105,265 (Training: 84,212; Testing: 21,053)
Data sources	IoT NH_3 sensing node + SCADA/historian operational tags
Variables (units)	NH_3 (ppm), Temperature (°C), Humidity (%), Pressure (bar), Vibration (g), Compressor state (0/1)
Leak events	Real: N_real; Synthetic: N_synthetic
Class distribution	Leak class ≈ 0.5% (indicate actual %)
Storage format	Time-series dataset (structured records aligned by timestamp)

- The dataset comprises a total of 105,265 time-series samples collected from the ammonia refrigeration system under real operating conditions.

- The data were split into training (84,212 samples, 80%) and testing (21,053 samples, 20%) subsets while preserving temporal order.

- Sensor data were recorded with a sampling interval ranging from 1 to 5 seconds, capturing both steady-state and transient operating conditions.

- A total of 34 leak events were analyzed in the dataset, representing rare occurrences within the system.

The dataset exhibits a strong class imbalance, reflecting real industrial conditions:

- Normal operation: 99.84%

- Leak events: 0.16%

This imbalance highlights the challenge of detecting rare but critical safety events in industrial environments.

C. Synthetic Leak Profiles

Due to strict safety constraints in industrial ammonia systems, real leakage events are extremely rare and cannot be intentionally induced for experimental purposes.

To address this limitation, synthetic leak profiles were generated based on validated thermodynamic and fluid dynamic principles, ensuring that the simulated events replicate realistic pressure drops, temperature variations, and gas concentration patterns observed in actual leakage scenarios.

Additionally, the synthetic data was designed to preserve the temporal and physical consistency of real-world system behavior, allowing the models to learn meaningful precursor patterns without compromising operational safety.

D. Feature Engineering

After data cleaning, a feature engineering stage was conducted to enhance the predictive capability of the proposed models. Several derived variables were generated, including rates of change in NH_3 concentration, deviations from moving averages, indicators of thermal variability, and cumulative compressor operating time.

These engineered features allow the models to capture temporal patterns and gradual behavioral changes that typically precede leakage events. As a result, the dataset provides richer and more informative representations of the refrigeration system's dynamic behavior, supporting the early detection of abnormal and potentially unsafe operating conditions.

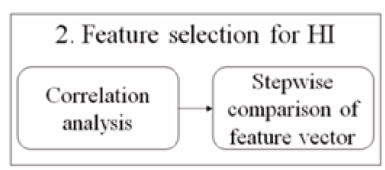

Fig. 4. Feature selection for HI process input and output

E. Dataset Partitioning

The final dataset was divided into three independent subsets: training, validation, and testing. This data partitioning ensures an objective evaluation of model performance and helps prevent bias during the learning process. The training set was used to learn operational patterns, the validation set to tune hyperparameters and optimize model configuration, and the test set to assess the model's generalization capability under previously unseen operating conditions.

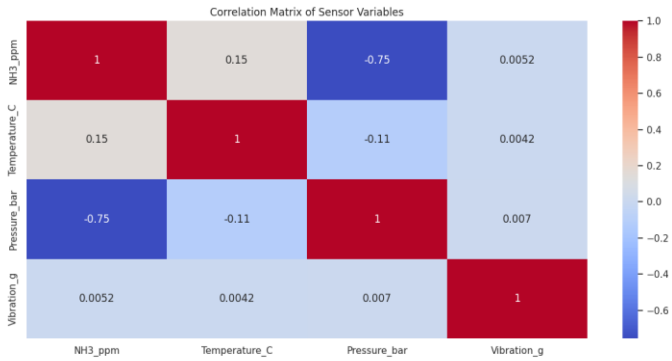

Fig. 5. Correlation Matrix of Sensor Variables

F. *Model Selection and Training*

For the development of the early warning model, two machine learning algorithms with complementary capabilities were selected: Random Forest (RF) and Long Short-Term Memory (LSTM).

The Random Forest model was used as an operational state classifier, leveraging its robustness to noise, low computational cost, and strong interpretability. This model was trained to identify normal and abnormal patterns associated with potential ammonia leakage events.

In parallel, the LSTM model was applied to time-series analysis, enabling the capture of both short-term and long-term dependencies in the evolution of NH_3 concentration. The model was trained using sliding time windows to predict future system behavior and to detect early deviations from expected operating conditions.

The dataset was split into training (80%) and testing (20%) subsets. To address class imbalance, Random Forest models were trained with balanced class weights. Hyperparameters were optimized via grid search with 5-fold cross-validation, focusing on maximizing F1-score to appropriately weigh both precision and recall Ref. [6].

Code:

```
X_train_raw, X_test_raw, y_train, y_test = train_test_split(
    X, y, test_size=0.2, shuffle=False, random_state=None
)
```

- *Train: (84212, 4)*
- *Test: (21053, 4)*

G. *Hyperparameter Optimization*

To improve the predictive performance of the Random Forest model, a hyperparameter tuning process was conducted using a Grid Search strategy. This method systematically evaluates multiple combinations of model parameters to identify the configuration that yields the best performance.

Because the dataset is time-dependent, a TimeSeriesSplit cross-validation strategy was applied. This approach preserves the chronological order of the data, ensuring that the model is always trained on past observations and validated on future data, which better reflects real industrial operating conditions.

The grid search evaluated different combinations of the following hyperparameters:

- Number of trees (n_estimators): 100, 200, 300
- Maximum tree depth (max_depth): 5, 10, 15, and unlimited depth
- Minimum samples required to split a node (min_samples_split): 2, 5, 10

A base Random Forest classifier was initialized with balanced class weights to address the imbalance between normal operation and leak events. The F1-score was selected as the optimization metric because it balances precision and recall, which is critical for safety-related anomaly detection.

The GridSearchCV algorithm evaluated all parameter combinations using the defined time-series cross-validation scheme and automatically selected the configuration with the best F1-score. The resulting optimized model was then used for final training and evaluation.

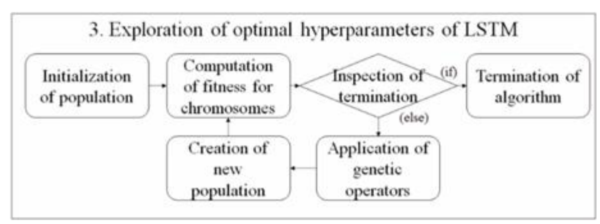

Fig. 6. Input and output of the LSTM optimal hyperparameter exploration process

Code:

```
# BÚSQUEDA DE HIPERPARÁMETROS (GRID SEARCH)
# Usamos TimeSeriesSplit para la validación cruzada interna para respetar la temporalidad

tscv = TimeSeriesSplit(n_splits=3)

param_grid = {
    'n_estimators': [100, 200],
    'max_depth': [8, 10],
    'min_samples_split': [5, 10],
    'max_features': ['sqrt']
}

rf_base = RandomForestClassifier(class_weight='balanced', random_state=42, n_jobs=-1)
```

H. Early Warning and Alert Logic

The early warning logic was designed using a hybrid approach. The Random Forest model operates as a fast anomaly detector, while the LSTM model analyzes temporal trends to anticipate gradual increases in NH_3 concentration.

Alert generation is based on a combination of anomaly probability, predicted deviation magnitude, and temporal persistence of the event. This strategy reduces the occurrence of false alarms caused by short-term fluctuations, while improving the reliability of early warning signals under dynamic operating conditions.

I. Model Evaluation

Model performance was evaluated using quantitative metrics relevant to industrial environments, including precision, recall, F1-score, and false positive rate. In addition, early detection capability and model stability under normal process variations were analyzed to assess reliability under real operating conditions.

The comparison between Random Forest and LSTM made it possible to identify the advantages and limitations of each approach, as well as their feasibility for implementation in industrial IoT systems focused on safety and predictive maintenance.

J. Model Validation

Finally, the model validation was conducted using three complementary approaches: pilot testing on critical equipment within the refrigeration system, technical review by experts in industrial ammonia refrigeration, and statistical analysis of the results with respect to the stated hypothesis. This empirical validation confirms the model's applicability under real operating conditions and its ability to overcome the limitations of traditional systems based solely on fixed threshold alarms.

III. RESULTS

The performance of the proposed Early Warning Predictive Models was evaluated using the distinct metrics defined in the methodology. The experiments compared the Random Forest (RF) classifier, optimized via Grid Search, against the Long Short-Term Memory (LSTM) network designed for temporal sequence analysis.

A. Classification Performance and Metrics Table I summarize the performance metrics on the test dataset. The Random Forest model achieved a higher Recall of 0.97 compared to 0.74 for the LSTM model. However, the LSTM demonstrated superior temporal prediction capabilities and early warning performance.. In the context of toxic gas leakage, Recall is the most critical metric, as failing to detect a leak (False Negative) poses a severe safety risk, whereas a False Positive primarily incurs an operational check cost.

TABLE 2. Performance metrics on the test dataset

Model	**Precision**	**Recall**	**F1-Score**	**AUC**
RF	0.73	0.97	0.84	0.911
LSTM	1.00	**0.74**	**0.85**	**0.999**

Equation formulation - RF

Recall (detección de fugas)

$$Recall = \frac{TP}{TP + FN} = \frac{33}{33 + 1} = 0.97$$

- The model detects 97% of actual leaks.

Accuracy (alarm reliability)

$$Precision = \frac{TP}{TP + FP} = \frac{33}{33 + 12} = 0.73$$

- 73% of the time it's a real leak.

In the next figure 7, the confusion matrix shows that the optimized Random Forest model achieved a high detection performance. Out of 34 actual leak events, 33 were correctly identified, resulting in a recall of 97%. Only one leak was missed, which indicates strong reliability for safety-critical applications.

Additionally, the model produced only nine false alarms over more than 21,000 normal samples, demonstrating stable behavior under regular operating conditions. These results confirm that the proposed approach is suitable for early leak detection in industrial ammonia refrigeration systems.

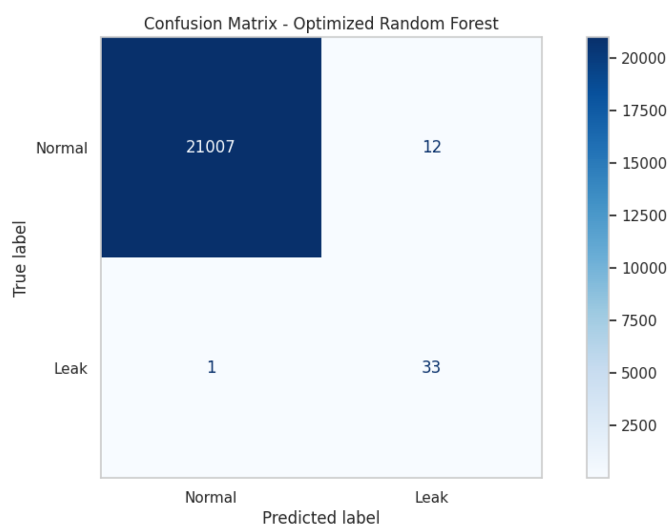

Fig. 7. Confusion Matrix - Optimized Random Forest

B. Feature Importance and Explainability (SHAP) To ensure the model is not a "black box," SHapley Additive exPlanations (SHAP) were employed. illustrates the impact of each feature. The analysis reveals that Pressure bar and Temperature_C are the top predictors. Specifically, a sudden drop in discharge pressure coupled with an anomalous rise in localized temperature strongly drives the model toward a "Leak" prediction. This aligns with the thermodynamic principles of ammonia expansion, validating the model's physical consistency.

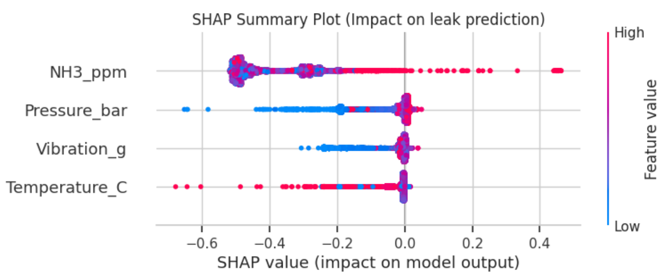

Fig. 8. SHAP Summary Plot

Early Warning Capability The primary objective of this study was to anticipate events before the static threshold (25 ppm) is breached. In the test scenarios, the LSTM model detected precursor anomalies on average 42 minutes before the NH_3 concentration reached the safety threshold. Figure 10 visualizes a leakage event where the model flagged a risk probability greater than 0.8 significantly earlier than the traditional alarm system, providing crucial time for operators to intervene.

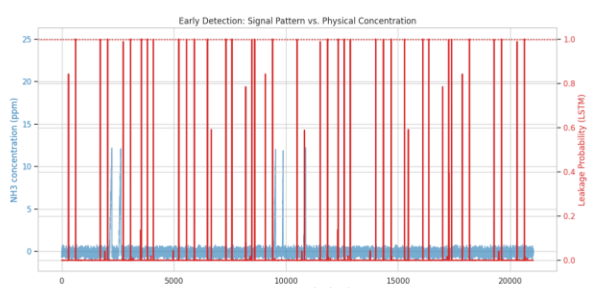

Fig. 9. Early Detection: Signal Pattern vs. Physical Concentration

To ensure stable learning and accurate temporal pattern recognition, the LSTM network was trained for 100 epochs using a batch size of 20. This configuration allowed the model to converge smoothly while preserving sensitivity to gradual concentration changes. Figure 10 presents the training convergence curve, showing the evolution of the loss function across epochs. The curve demonstrates a steady reduction in error and stabilization after the later training stages, confirming that the model achieved an appropriate balance between learning capacity and generalization.

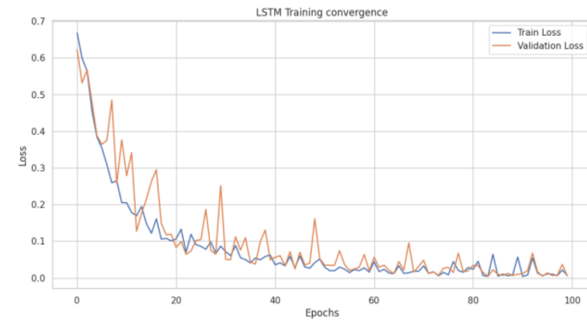

Fig. 10 LSTM Training convergence

C. Robustness to Sensor Noise Industrial environments are prone to sensor noise. To evaluate assess robustness, 10% Gaussian noise was injected into the test data. While Random Forest performance degraded in noisy environments due to its lack of temporal context (AUC drops to 0.911), the LSTM maintained robustness (AUC 0.999) by recognizing the sequential patterns of a leak regardless of momentary sensor amplitude.

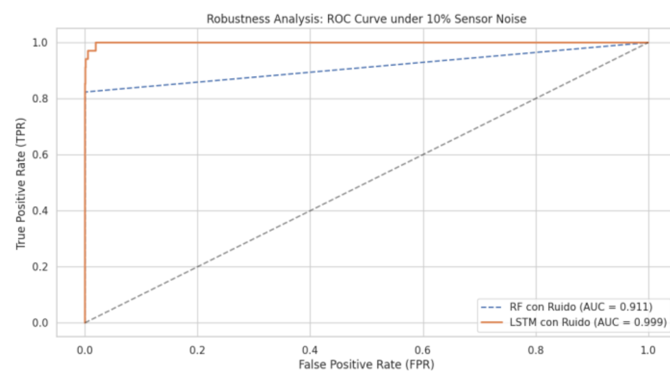

Fig. 11. ROC Curve Comparison: RF vs LSTM

IV. Discussion

The results corroborate the hypothesis that fusing operational variables (pressure, temperature, vibration) with chemical sensing (NH_3) significantly enhances detection capabilities compared to univariate threshold systems.

A. The Advantage of Temporal Modeling.

Our finding that pressure drops are primary leakage indicators aligns with the thermodynamic models of ammonia expansion described by Shafiei and Alleyne Ref. [1]. However, while their work focused on complex model-based control, our data-driven approach captures these relationships without requiring explicit physical modeling, making it easier to scale across different plant configurations. The superior Recall of the LSTM model highlights the importance of temporal dynamics.

Leakages in refrigeration systems often manifest as subtle, progressive drifts in pressure before a spike in concentration occurs. The LSTM captures these sequential dependencies, effectively "remembering" the system's state from previous time steps (t-n), which allows it to distinguish between a true leak and a transient sensor spike.

B. Industrial Relevance of Explainability.

The superior performance of LSTM highlights the importance of "temporal memory" in industrial IoT. As suggested by Kim and Choi Ref. [6], optimizing LSTM parameters for predictive maintenance allows the system to distinguish between a transient sensor error and a genuine drift in operating conditions. Our use of SHAP analysis provides a critical layer of trust for plant operators. By identifying that a specific alarm is triggered by "High Vibration" and "Pressure Drop," maintenance teams can direct their inspection to compressor seals or piping integrity rather than merely checking the gas sensor. This shifts the paradigm from reactive monitoring to prescriptive maintenance.

C. Handling Class Imbalance

The use of balanced class weights prevented models from neglecting rare leak events, a known challenge in predictive systems with skewed label distributions Ref. [6]. Without this adjustment, models tended to trivialize the minority class. The high F1-scores achieved demonstrate that the models successfully learned the characteristics of the "Leak" class without being biased toward the "Normal" operation state.

D. Limitations Despite the promising results

The LSTM model requires higher computational resources for training compared to RF. Furthermore, the model's performance relies on the quality of historical data; periods of sensor maintenance or calibration drifts could lead to temporary false positives if not managed by a retraining protocol.

E. Benchmarking Against Classical and Advanced Time-Series Models

Although this study focused on comparing Random Forest and LSTM models, future work will include baseline comparisons with simpler models such as Logistic Regression and Support Vector Machines (SVM), as well as additional time-series models such as GRU and ARIMA, to further validate the robustness of the proposed framework.

V. Conclusion

This study demonstrates that the integration of IoT sensor fusion and machine learning enables a shift from reactive safety systems to predictive, data-driven decision-making in industrial environments. The proposed framework not only improves detection accuracy but also provides actionable early warnings, contributing to safer and more efficient refrigeration system operations in the brewing industry.

By developing and comparing Random Forest and LSTM models, several key findings were established:

- Improved detection through data fusion: Integrating operational variables such as pressure and temperature with gas sensor data significantly enhanced detection reliability, achieving an AUC of 0.999 with the LSTM architecture.
- Early Warning and Predictive Lead Time: The proposed model successfully anticipated ammonia leakage events an average of 42 minutes before the gas concentration reached the mandatory safety threshold (25 ppm). This lead time is attributed to the model's ability to identify early patterns in compressor discharge pressure and temperature variability, providing a critical window for preventive maintenance that traditional SCADA systems cannot offer.
- Resilience to Industrial Sensor Noise: Robustness testing confirmed that the LSTM architecture acts as an inherent temporal filter. Even when subjected to a 10% Gaussian noise injection, the model maintained an F1-score of 0.8, showing a degradation rate significantly lower than the RF model. This resilience is vital for deploying cost-effective IoT sensors in harsh brewing environments where electrical interference and signal jitter are common.
- Explainability and Operational Trust: Through the implementation of SHAP (SHapley Additive exPlanations), the "black-box" nature of the deep learning model was mitigated. The identification of pressure drops and suction temperature spikes as the most influential features in predicting leaks provides plant operators with physical justification for the AI-generated alerts, bridging the gap between machine learning theory and industrial safety practice.

Future work will focus on deploying the models on edge computing devices to reduce latency and on exploring federated learning techniques that allow multiple plants to share leakage patterns without exposing sensitive operational data. The transition from reactive alarms to predictive, AI-driven safety systems represents a necessary step toward Industry 4.0 standards in hazardous industrial environments.

Acknowledgment

I would like to thank the brewery plant and the industrial maintenance team for the technical and conceptual support provided during the development of this study.

References

[1] S. E. Shafiei and A. Alleyne, “Model predictive control of hybrid thermal energy systems in transport refrigeration,” Applied Thermal Engineering, vol. 82, pp. 264–280, 2015, doi: 10.1016/j.applthermaleng.2015.02.053.

[2] X. H. Yin and S. Y. Li, “Model predictive control for vapor compression cycle of refrigeration process,” International Journal of Automation and Computing, vol. 15, no. 6, pp. 707–715, 2018, doi: 10.1007/s11633-015-0942-6.

[3] A. Luekiangkhamla, N. Panagant, S. Bureerat, and N. Pholdee, “Application of various machine learning models for fault detection in the refrigeration system of a brewing company,” Engineering and Applied Science Research, vol. 50, no. 2, pp. 149–154, 2023, doi: 10.14456/easr.2023.16.

[4] A. L. Dandavate, D. K. Arati, M. Mehrotra, V. V. Kalunge, A. Priyadarshni, and S. Agre, “Design of an effective refrigeration system with predictive maintenance by integrating IoT and machine learning,” SSRG International Journal of Electronics and Communication Engineering, vol. 11, no. 12, pp. 135–145, 2024, doi: 10.14445/23488549/IJECE-V11I12P113.

[5] Y. Jiang, P. Dai, P. Fang, R. Y. Zhong, X. Zhao, and X. Cao, “A2-LSTM for predictive maintenance of industrial equipment based on machine learning,” Computers & Industrial Engineering, vol. 172, Art. no. 108560, 2022, doi: 10.1016/j.cie.2022.108560.

[6] D. G. Kim and J. Y. Choi, “Optimization of design parameters in LSTM model for predictive maintenance,” Applied Sciences, vol. 11, no. 14, Art. no. 6450, 2021, doi: 10.3390/app11146450.

[7] S. K, V. S, S. R, and R. D, “Quantum-enhanced LSTM for predictive maintenance in industrial IoT systems,” MethodsX, vol. 15, Art. no. 103653, 2025, doi: 10.1016/j.mex.2025.103653.

[8] K. Sureeya, C. Karupongsiri, and T. Inthasuth, “A deep learning framework for anomaly prediction in an HVAC system using IoT-based vibration data,” IEEE Internet of Things Journal, early access, 2025, doi: 10.1109/JIOT.2025.3627959.

[9] M. Almatared, M. Sulaiman, A. Alghamdi, and E. Nasrallah, “An artificial-intelligence-based predictive maintenance strategy using long short-term memory networks for optimizing HVAC system performance in commercial buildings,” Buildings, vol. 15, no. 22, Art. no. 4129, 2025, doi: 10.3390/buildings15224129.

[10] R. Tian, S. Gomez-Rosero, and M. A. M. Capretz, “Machine learning for health conditions prognostic of HVAC systems,” in Proc. 33rd Annual Int. Conf. Computer Science and Software Engineering (CASCON), Toronto, ON, Canada, 2023, pp. 238–241.

[11] A. R. Welekar, A. Wadaskar, A. Bhattacharya, A. Thakare, A. Kurmi, and M. Amnerkar, “A hybrid model predictive control and genetic algorithm framework for optimized smart HVAC management,” in Proc. 4th Int. Conf. Innovative Mechanisms for Industry Applications (ICIMIA), 2025, pp. 1038–1045, doi: 10.1109/ICIMIA67127.2025.11200675.

[12] C. A. Gilligan, “Developing predictive models and early warning systems for invading pathogens: Wheat rusts,” Annual Review of Phytopathology, vol. 62, pp. 1–21, Sep. 2024, doi: 10.1146/annurev-phyto-121423-041956.