1.Introduction

* Predictive maintenance (PdM) using AI and IIoT is becoming crucial in manufacturing to reduce costs from unplanned failures and downtime.
* This study develops a ML-based PdM approach to predict remaining useful life (RUL) of production lines using real-world IoT sensor data.
* The problem is treated as a regression task rather than classification to provide early warnings before failures occur.

2.Literature Review

* Reviews statistical and data-driven AI approaches to PdM. Statistical methods include ARIMA models, while data-driven methods use ML and DL.
* Highlights the importance of expert knowledge in feature engineering and data interpretation for PdM.
* Notes challenges like data resilience, censorship, and computational requirements that impact model choice.

3.Methods & Analysis

3.1 Data Collection and Preprocessing The study used two datasets collected over a year from a real-world consumer goods production line:

1. IoT sensor data with 101 features (50 sensor readings, 50 timestamps) and 8,668,431 instances recorded every 3-6 seconds. An additional "status" feature indicated if the line was running (0) or not (non-zero).
2. Unplanned stop data with 34 columns and 6787 rows, including failure type, time, duration, etc.

Data preprocessing involved several steps:

* Min-max normalization of sensor values
* Filling missing values (~1%) with feature medians
* Converting local timestamps to Unix time
* Generating Remaining Useful Life (RUL) values by mapping sensor readings to the next stop event
* Creating an alternative "Stops-removed" dataset by excluding readings during stoppage periods
* Selecting a subset of 18 expert-identified features relevant to the failure type
* Applying Principal Component Analysis (PCA) for dimensionality reduction (not used in final models)

The preprocessed datasets were split into train (70%) and test (30%) sets, with 5-fold cross-validation for hyperparameter tuning.

Correlation analysis revealed several highly correlated sensor features, likely due to physical proximity. A "Correlations-removed" dataset was created by dropping these features, resulting in 11 remaining attributes.

3.2 Prediction Models Several machine learning algorithms were evaluated for predicting RUL:

* Support Vector Regression (SVR) with RBF kernel
* Multilayer Perceptron (MLP) with ReLU activation, Adam optimizer, 50 hidden nodes
* Random Forest (RF) with 51 trees (tuned)
* XGBoost (XGB) with max\_depth=5, learning\_rate=0.3, n\_estimators=100 (tuned)

Unsupervised learning techniques were also applied:

* K-means clustering to identify underlying data patterns. The elbow method suggested 3 optimal clusters in the "Stops-removed" dataset. Cluster labels were added as a new feature.
* Autoencoders (AE) to learn normal sensor behavior. Reconstruction error from the tuned AE (256 batch size, 3 hidden layers, 100 epochs) was included as a feature.

1. Results Model performance was evaluated using R-squared, MAE, MAPE, and RMSE. Key findings:

* RF consistently outperformed other models across all datasets, with R2 > 0.99. XGB was second best.
* Removing stop periods improved results, especially for MLP. Clustering further enhanced performance.
* Median filtering of predictions reduced false positives in the new TP-FP evaluation scheme.
* The "Stops-removed" RF model achieved the highest precision of 0.923 for a 1-hour prediction window.
* Autoencoders moderately boosted scores for all models by capturing deviation from normal sensor patterns.

A new validation set from the following month revealed lower accuracies, prompting a revised evaluation approach. The final TP-FP method used a sliding window to compare predicted RUL to actual failures. An alert was considered a true positive (TP) if a stoppage occurred within the specified time horizon.

This TP-FP framework, combined with median filtering and RF modeling, successfully predicted 42% of real production line failures in the validation period.

1. Discussion and Conclusion The study demonstrates the effectiveness of machine learning, particularly ensemble methods like Random Forest, for predictive maintenance using IoT sensor data. Careful data preprocessing, feature selection, and results interpretation are crucial for real-world impact.

DEFINITIONS:

1. Predictive Maintenance (PdM):
   1. A proactive maintenance strategy that uses data analysis to predict when equipment is likely to fail, allowing maintenance to be scheduled in advance to prevent unexpected breakdowns and minimize downtime.
2. Internet of Things (IoT):
   1. A network of connected devices, sensors, and software that enables data collection, exchange, and analysis across machines and systems. IoT is a key enabler for predictive maintenance in Industry 4.0.
3. Random Forest (RF):
   1. An ensemble learning method that constructs multiple decision trees and outputs the mean or mode of the individual trees' predictions. RF is known for its robustness to overfitting and ability to handle high-dimensional data.
4. XGBoost (XGB):
   1. An optimized gradient boosting library that implements machine learning algorithms under the Gradient Boosting framework. XGBoost is known for its speed and performance in many supervised learning tasks.
5. Multilayer Perceptron (MLP):
   1. A class of feedforward artificial neural networks consisting of an input layer, hidden layers, and an output layer. MLPs can learn non-linear relationships and are used for classification and regression tasks.
6. Support Vector Regression (SVR):
   1. A regression algorithm that tries to fit the best line (or hyperplane) within a predefined or threshold error value. SVR can handle non-linear relationships by transforming data into a higher-dimensional space.
7. K-means Clustering:
   1. An unsupervised learning algorithm that partitions n observations into k clusters, where each observation belongs to the cluster with the nearest mean. K-means is used to identify underlying patterns or groupings in data.
8. Autoencoder (AE):
   1. A type of neural network that learns to compress and encode data, then reconstructs the data back from the reduced encoded representation. Autoencoders can be used for dimensionality reduction, feature learning, and anomaly detection.
9. Principal Component Analysis (PCA):
   1. A statistical technique used to reduce the dimensionality of a dataset while retaining most of the important information. PCA identifies the principal components, which are linear combinations of the original variables that capture the maximum variance.
10. R-squared (R2):
    1. A statistical measure that represents the proportion of variance in the dependent variable that is predictable from the independent variable(s). R-squared ranges from 0 to 1, with higher values indicating better model fit.
11. Mean Absolute Error (MAE):
    1. A measure of the average magnitude of errors in a set of predictions, without considering their direction. MAE is calculated as the average of the absolute differences between the predicted and actual values.
12. Root Mean Squared Error (RMSE):
    1. A quadratic scoring rule that measures the average magnitude of the error. RMSE gives a relatively high weight to large errors, making it more sensitive to outliers compared to MAE.