Team-7 Progress Report

Title : Convergence of IoT, machine learning, and advanced analytics to address critical challenges in forecasting, anomaly detection, and cybersecurity

Project Overview

Our project focuses on analyzing sensor data from multiple IoT devices deployed in different environments. The goal is to identify characteristic patterns in sensor readings, detect seasonality trends, and predict environmental variations using deep learning techniques. Additionally, we address cybersecurity concerns by implementing anomaly detection to identify potential cyberattacks. The project utilizes two datasets to improve model generalization and accuracy. Leveraged the diagram we created in the 1st assignment.

Exploratory Data Analysis and Preprocessing

To prepare the data for analysis, we conducted extensive exploratory data analysis and preprocessing. We standardized the sensor data format, converted timestamps from Unix time to a readable format, and transformed boolean sensor values such as motion and light into numerical representations.

One of the key challenges was handling inconsistencies in sampling rates across different devices. To address this, we resampled the data to ensure proper alignment for time-series analysis. Feature selection was also a critical step, as we found strong correlations among CO, LPG, and smoke sensor readings.

To optimize efficiency, we removed LPG and smoke from the dataset, ensuring that only relevant information was used for modeling. Through visualization techniques, we identified environmental differences across devices, with some showing stable conditions while others exhibited fluctuations influenced by human activity. Autocorrelation plots confirmed daily seasonality patterns, particularly in temperature readings from Device 1c.

Cybersecurity Considerations and Anomaly Detection

Since IoT networks are susceptible to cyber threats, we integrated anomaly detection techniques to identify potential attacks such as data tampering, sensor spoofing, and denial-of-service (DoS) incidents. By analyzing irregularities in motion and temperature sensor readings, we implemented statistical anomaly detection methods, including standard deviation

analysis, to flag suspicious patterns. Additionally, machine learning-based classifiers are being tested to distinguish normal sensor behavior from potential cyberattacks.

Using Transformers for Predictive Modeling

For predictive modeling, we chose a Transformer-based architecture instead of traditional recurrent networks like LSTMs. Transformers are more effective for time-series forecasting due to their ability to capture long-range dependencies through self-attention mechanisms. This approach allows for better scalability and performance compared to recurrent models, which suffer from vanishing gradient issues.

The Transformer model consists of an embedding layer to convert sensor data into learnable representations, positional encoding to maintain temporal relationships, and multi-head self-attention to focus on the most relevant time steps. The feedforward layers process these extracted features for forecasting environmental conditions. By leveraging Transformers, we aim to improve seasonality detection, correlation learning, and real-time forecasting of environmental variations.

Challenges and Solutions

During implementation, we faced challenges related to dataset size and Transformer hyperparameter tuning. The large dataset, with sensor readings collected at different frequencies, required optimization techniques such as removing redundant features, applying data compression, and implementing batch processing for efficient memory management. Tuning Transformer hyperparameters, including attention heads, hidden layer size, and learning rate, was another challenge. We addressed this through empirical testing to achieve a balance between model accuracy and computational efficiency.

Next Steps

Moving forward, we will finalize Transformer model training and evaluate its performance in predicting IoT sensor trends. Additionally, we will refine the anomaly detection pipeline for identifying cyber threats and deploy the trained model for real-time IoT monitoring. This project highlights how deep learning, IoT analytics, and cybersecurity work together to enhance smart environments and improve sensor-based automation.