Machine Learning for IoT

Part-I: Data Analysis

Objectives

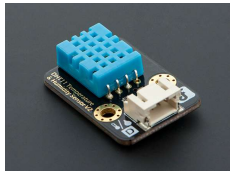
- Review visualization techniques to understand the structure of data
- Learn to identify useful trends to guide the training process
- Understand how to prepare and organize data for training

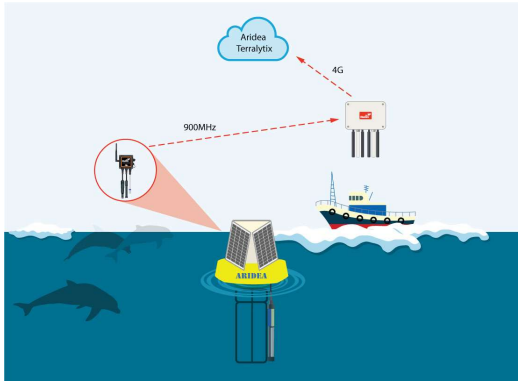
Contents

- Use Case I: Temperature & Humidity Forecasting
  - Data Visualization
  - Data Imputation
- Use Case II: Predictive Maintenance
  - Data Visualization
  - Data Cleaning
  - Data Preparation

Temperature & Humidity forecasting

- Applications
  - Precision Agriculture
  - Smart Buildings (HVAC)
  - Environmental protection
- Dataset:
  - One day records from DHT-11 sensor [Kaggle]
  - Challenge: Missing data due to network/sensors failures





Temperature & Humidity forecasting

- **Notebook:** Kaggle DHT-11

Predictive Maintenance

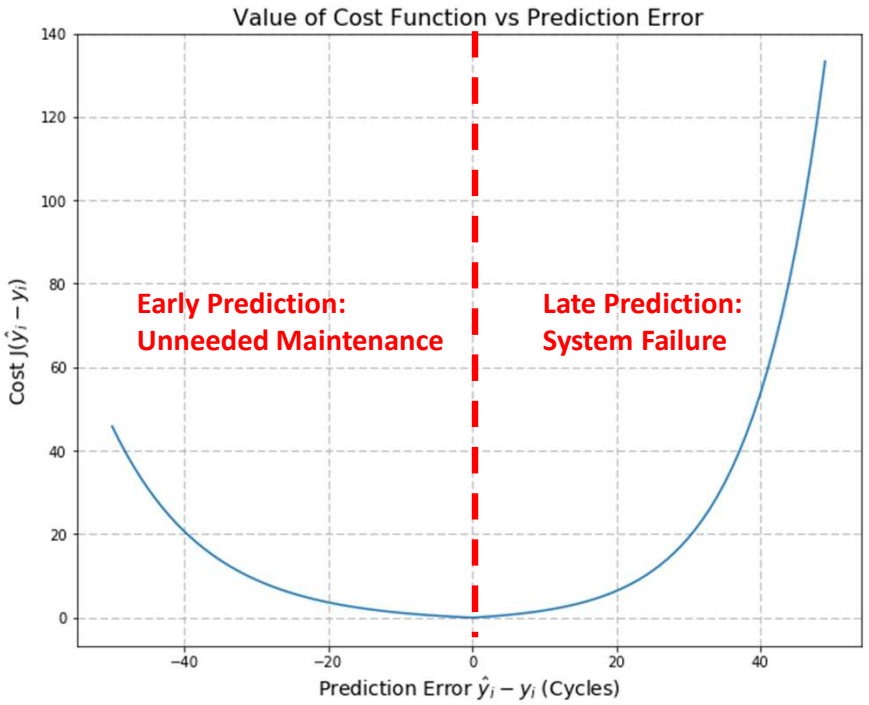
- Design systems able to identify issues before the equipment or the machine fails

Windmill Disaster

Why predictive maintenance?

- Reactive maintenance
  - Replace the machine after the failure or damage (e.g. battery car)
  - × Replacement is more expensive than fixing
  - × Failures can be dangerous
  - × Maintenance is also costly and dangerous
- Scheduled maintenance
  - Do maintenance at regular rate (e.g. change car's oil every 5000 miles)
  - × Unnecessary maintenance can be wasteful
  - × May not eliminate all failures
- Predictive maintenance
  - Forecast failures before they arise
  - Challenge: difficult to make accurate forecasts for complex equipment
  - ✓ Pros: Increase reliability, save costs, improve reputation

The cost of a misprediction



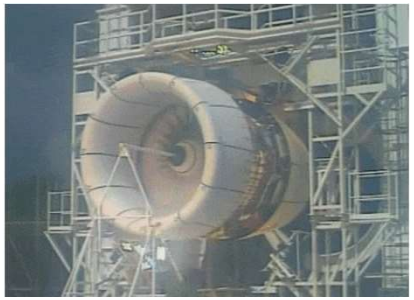
Towards predictive maintenance

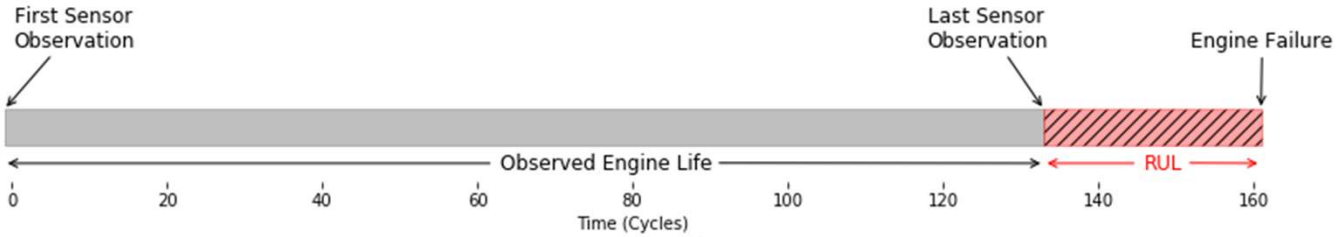
- How it works:
  - Monitor the system
  - Detect failure indicators
  - Identify components that need to be fixed
- Steps:
  - **Data Analysis**
  - Model Building
  - Model Validation
  - Model Deployment
  - Real-time Analytics integrated with maintenance

A case study

- Turbofan Engine Degradation Simulation Data Set
  - Sensor measurements collected on a fleet of aircraft engines
- Goal: Estimate the Remaining Useful Lifetime (RUL)
- How: Make predictions from the data collected by multiple sensors that monitor the turbofan engine
- Why we need ML:
  - Number of variables too high
  - System too complex to know the governing equations

  → We need a black-box model





Predictive Maintenance

- **Notebook:** Turbofan Engine Degradation Simulation