Team Crouching Tigers

Predictive Maintenance on IOT devices

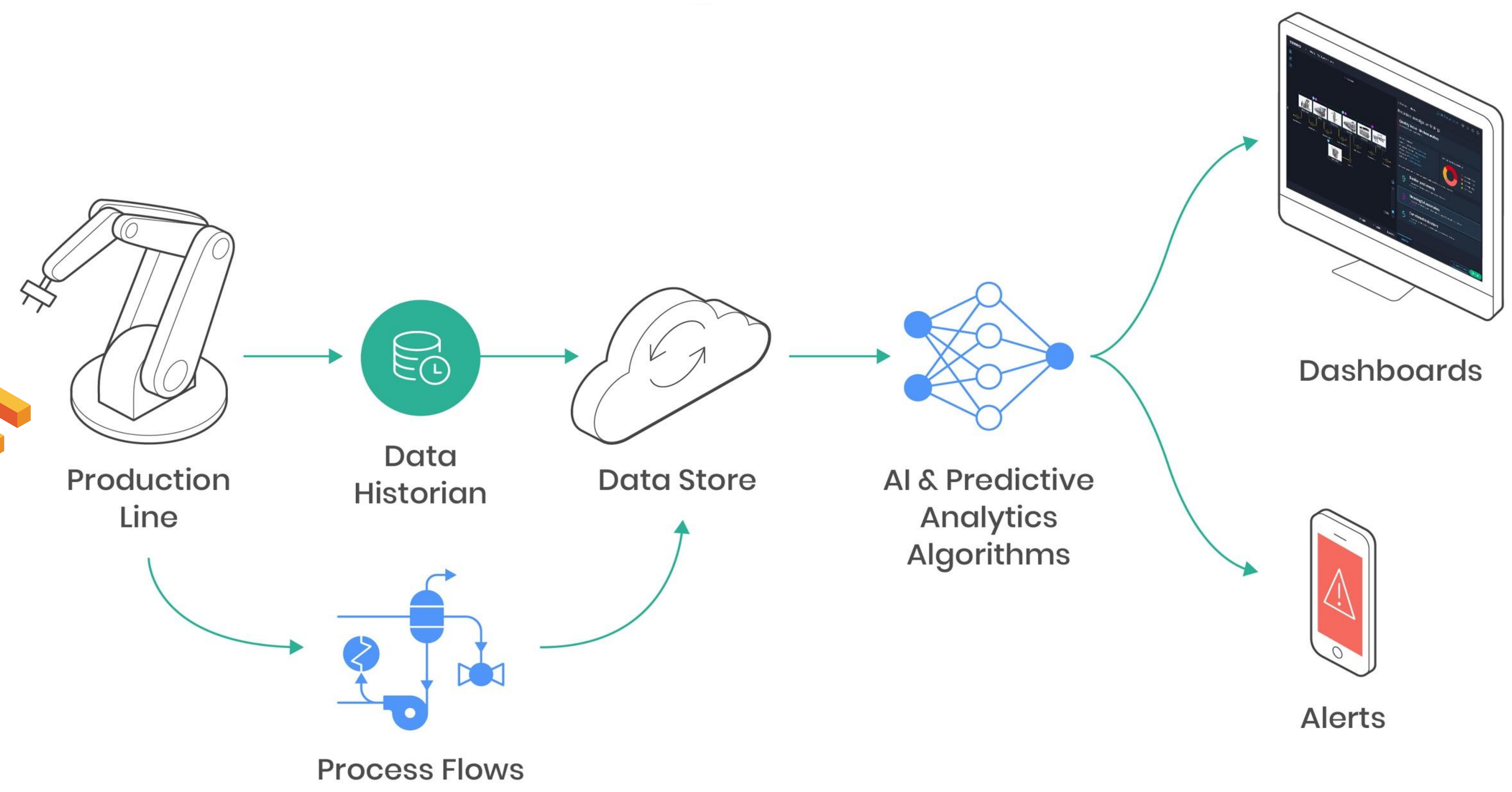
Chuangxin Xia, Risheng Wang, Yifan Li

SIMON FRASER UNIVERSITY

INTRODUCTION AND OBJECTIVE

Architecture

- Raw data collection from IoT devices
- ETL, EDA and feature selection using **pySpark** and **Pandas**
- Model training, Anomaly Detection using **Tensorflow**, **Sklearn**
- Model serving on **Google Cloud Machine Learning Engine**
- Model status server running **Node.js**
- Device real time monitoring and prediction using **React.js**



Our goal is to achieve an end-to-end IoT devices maintenance solution, from fields to web console.

MODEL ARCHITECTURE

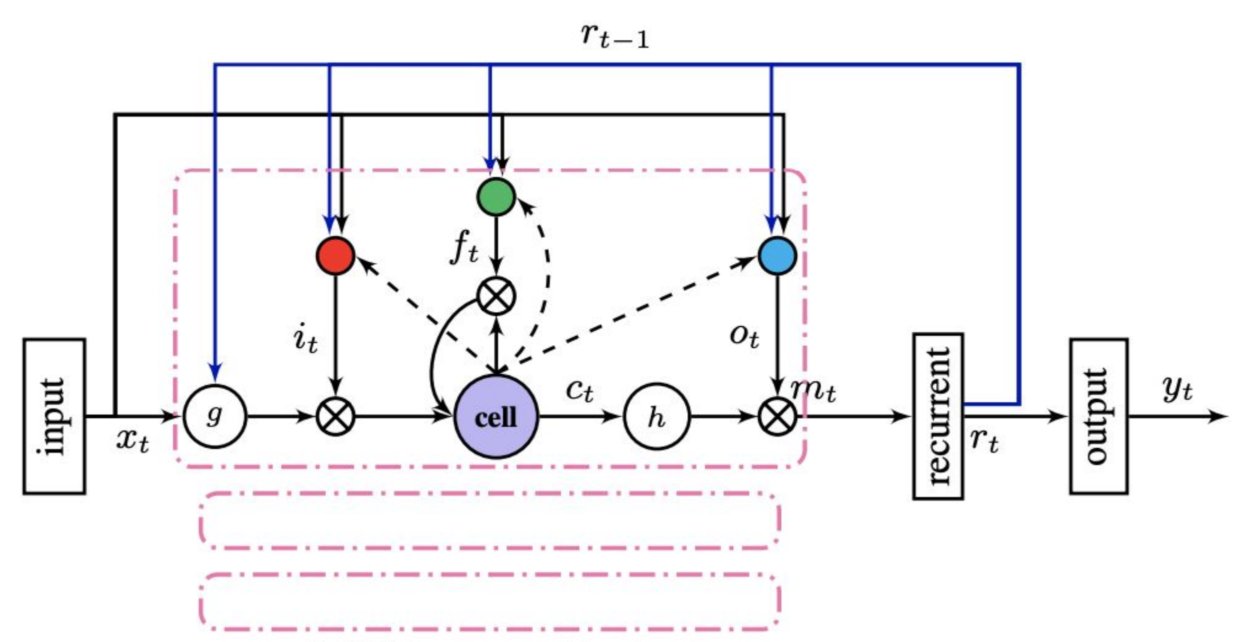
LSTM: Long Short-Term Memory RNN

The LSTM is consist of units called memory blocks in the recurrent hidden layer, which contains cells with self-connections storing the temporal state of the network in addition to special multiplicative units called gates to control the flow of information. Each memory block contains an input gate, an output gate, and forget gates.

- Input Gate: controls the flows of input activations into memory cell
- Output Gate: controls the output flow of cell activations into the rest of the network
- Forget Gate: scales the internal state of the cell before adding it to the cell through the self-recurrent connection of cell

The network computes a sequence to sequence mapping from inputs X to outputs Y by calculating network activations using following equations

$$i_t = \sigma(W_{ix}x_t + W_{im}m_{t-1} + W_{ic}c_{t-1} + b_i) \quad (1)$$
$$f_t = \sigma(W_{fx}x_t + W_{fm}m_{t-1} + W_{fc}c_{t-1} + b_f) \quad (2)$$
$$c_t = f_t \odot c_{t-1} + i_t \odot g(W_{cx}x_t + W_{cm}m_{t-1} + b_c) \quad (3)$$
$$o_t = \sigma(W_{ox}x_t + W_{om}m_{t-1} + W_{oc}c_t + b_o) \quad (4)$$
$$m_t = o_t \odot h(c_t) \quad (5)$$
$$y_t = \phi(W_{ym}m_t + b_y) \quad (6)$$



LSTM memory blocks

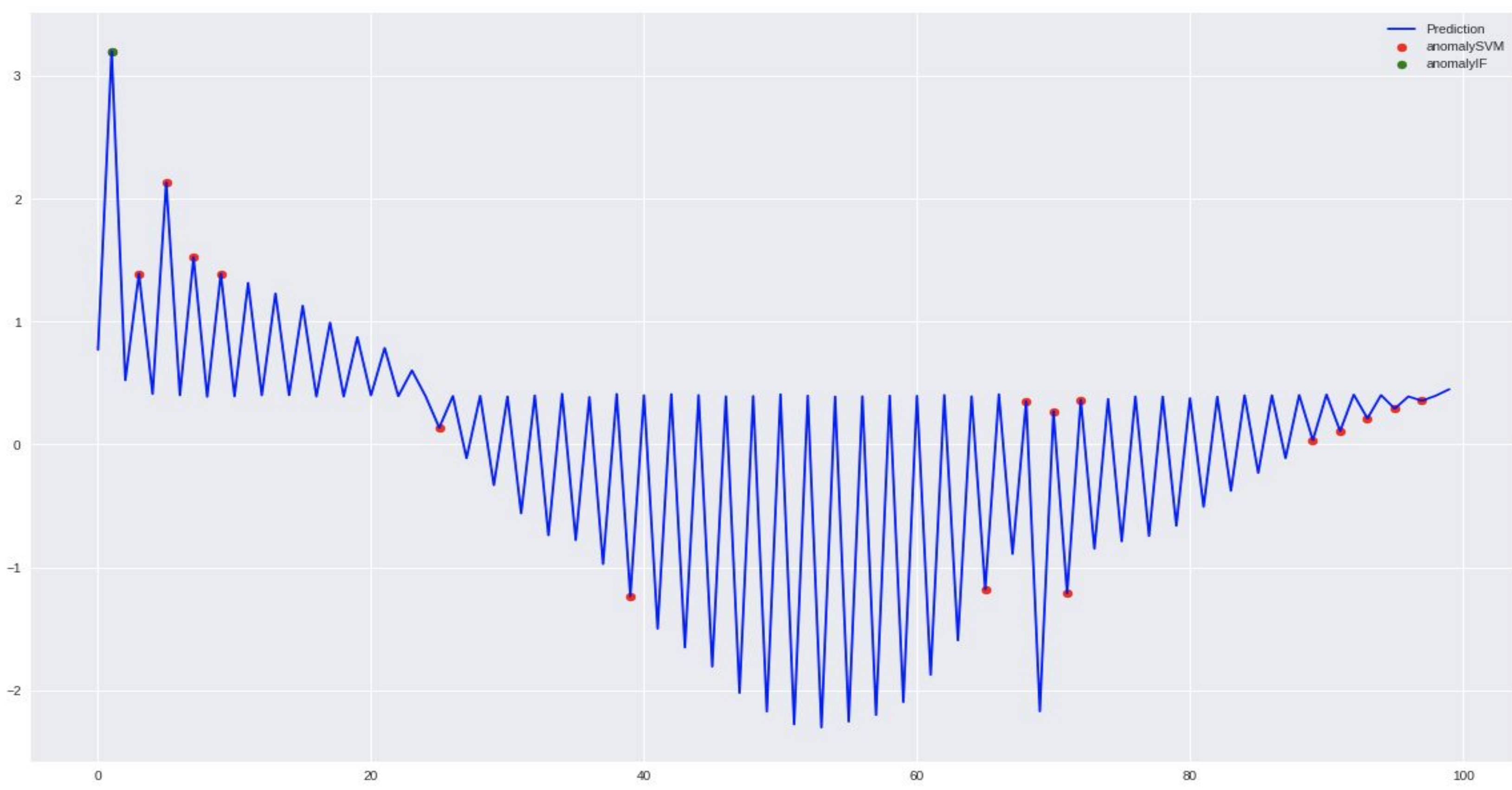
Anomaly Detection

The One-Class SVM for anomaly detection separates all the data points from hyperplane to the origin and maximizes the distance from the hyperplane to the origin point and is conduct to optimize the following problem:

$$\frac{1}{2}\|w\|^2 + \frac{1}{vn}\sum_{i=1}^{n}\varepsilon_i - p$$

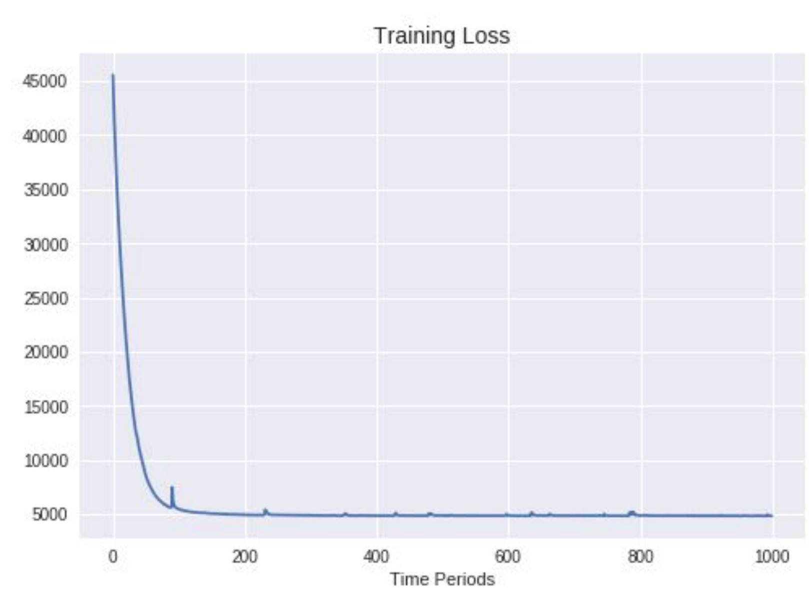
Isolation Tree explicitly identifies anomalies instead of profiling normal data points and is built on the basis of decision trees. The anomaly scores are required for decision making. In case of Isolation Forest, it is shown as: where h(x) is the path length of data points and c(n) is the average path length of unsuccessful search in Binary Tree Search Tree and n is the number of external nodes.

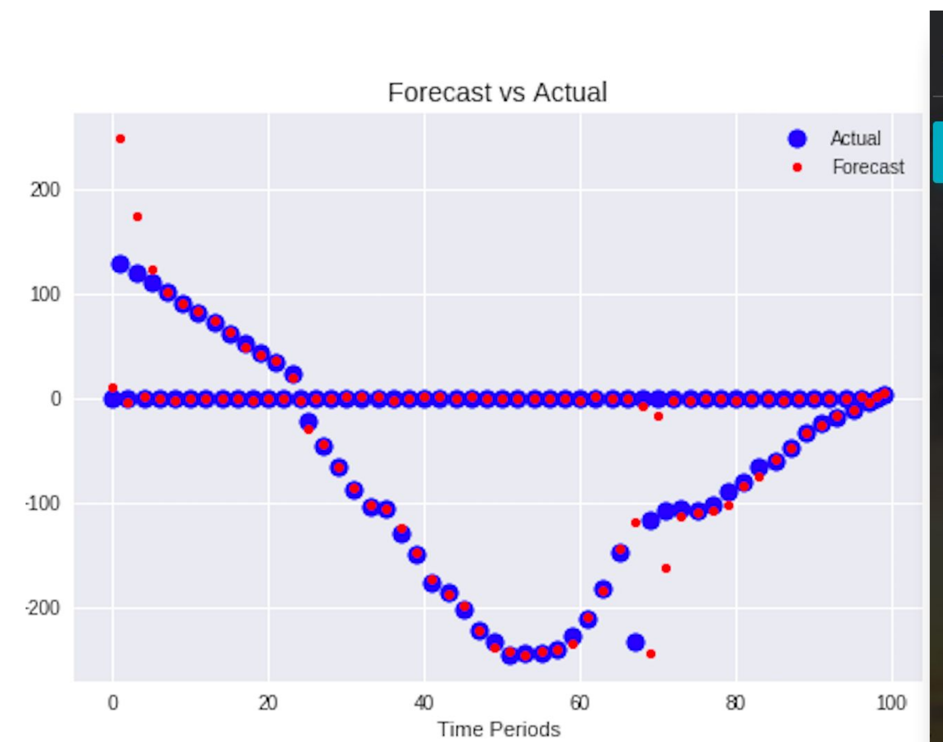
$$s(x|n) = 2^{-\left(\frac{E(h(x))}{c(n)}\right)}$$



EXPERIMENTS AND OUTCOMES

Training loss with 1 LSTM layer, learning rate 0.01, 1000 epochs





Predicted feature values vs actual feature values