Supervised Operational Change Point Detection using Ensemble Long-Short Term Memory in a Multicomponent Industrial System

Ashit Gupta, Vishnu Masampally, Vishal Jadhav, Anirudh Deodhar,
Venkataramana Runkana
Research & Innovation, Tata Consultancy Services Limited

SAMI 2021 Conference
Submission – 33 , Intelligent Manufacturing Systems

January 2021

TCS: Research & Innovation for Manufacturing & Engineering

20+ years of experience in physics-based and data-based modeling, optimization

Multidisciplinary Team
Mechanical, Chemical, Metallurgical, Biochemical and Aerospace Engineers, Material Scientists, Surface/Inorganic Chemists

Modeling & Simulation expertise
FEA, CFD, Molecular Modeling, Kinetics, DEM, PBM; Statistical, machine learning, deep learning and Reinforcement Learning techniques supported by plant/lab data

Building and deploying Digital Twins for process industries

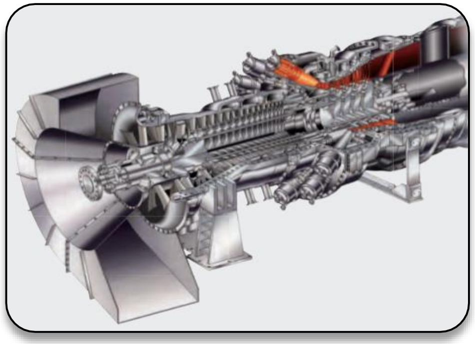
Dynamic process optimization


November 20, 2016 1:40 am JST
Tata, Japanese partners to use AI at coal-fired power plants

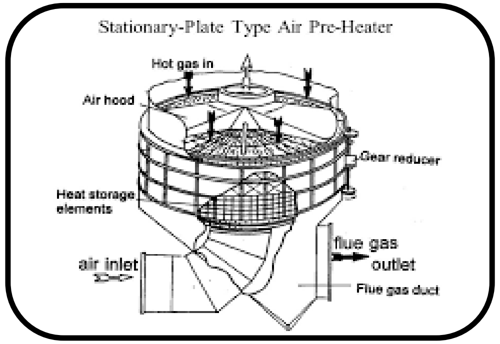
Boilers: commissioning and performance optimization

Blast Furnace: Prediction of silicon content of hot metal

Predictive Maintenance



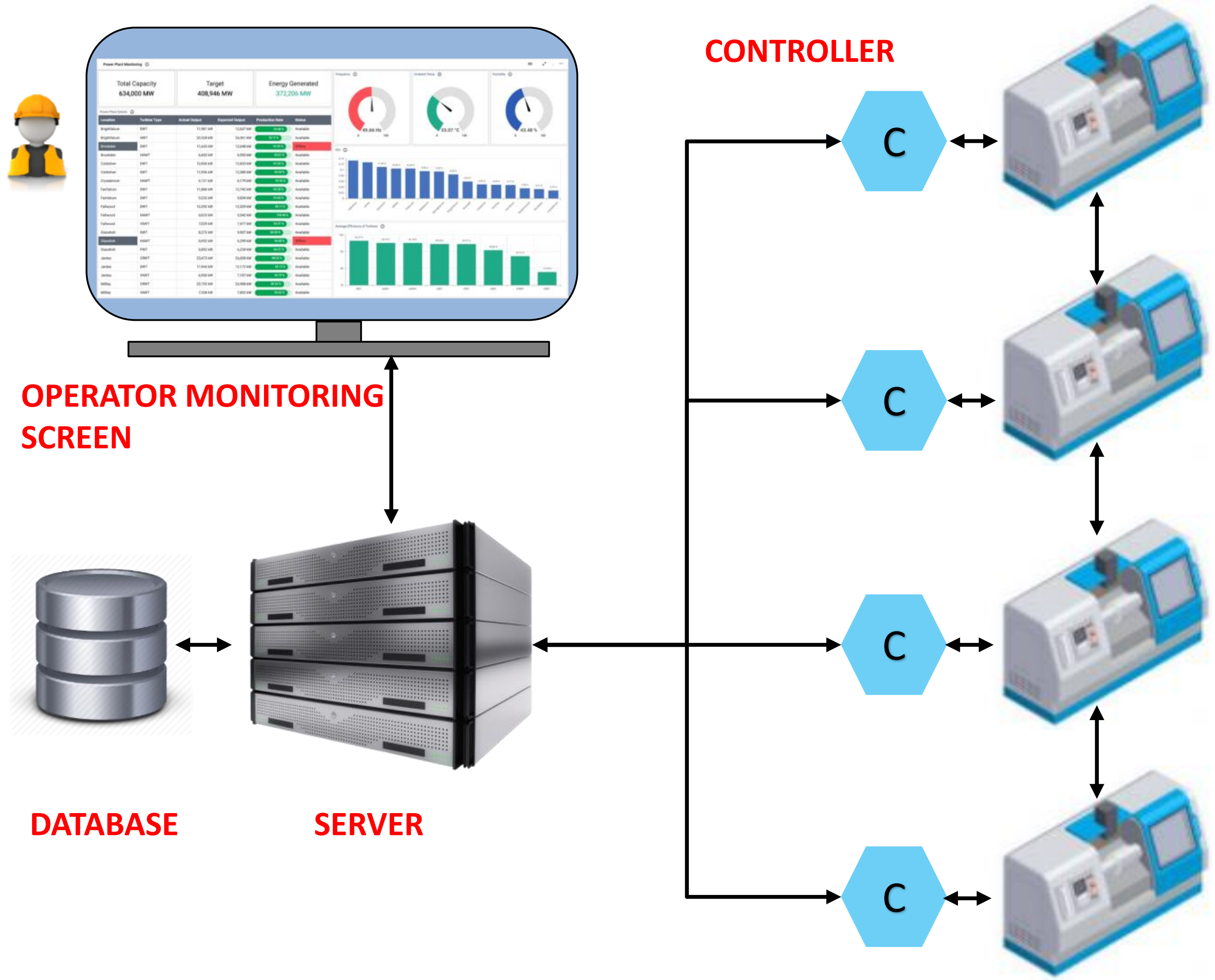
Gas Turbines: Anomaly detection and diagnosis



Air Pre-Heaters: Deposition and Clogging Prediction

Source: ambiente.usach.cl , asia.nikki.com, slideplayer.com, tatasteelbsl.co.in

Introduction



Why Early Detection?

- Take Preventive and Corrective Measures.
- Avoid Hazards.
- Avoid unplanned shutdowns/maintenance.
- Avoid loss in revenue, productivity & reputation.

Challenges faced by Operator

- Too much information from multiple Sensors of each component.
- Unspecified dependency among components.
- Operational change point not explicitly apparent.

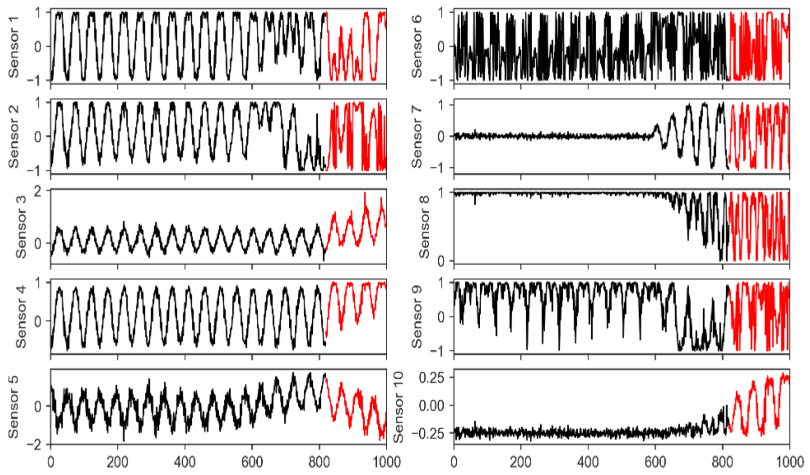
Goals

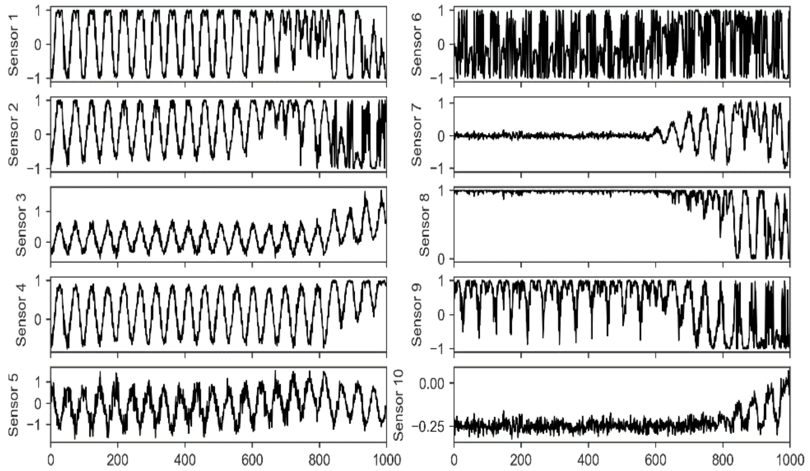
- Understand the behavior of plurality of sensors.
- Identify the onset of abnormal condition.

TATA CONSULTANCY SERVICES

depositphotos.com, iconfinder.com,itcreations.com, www.boldbi.com, www.ibm.com, http://clipart-library.com/

SYSTEM DESCRIPTION



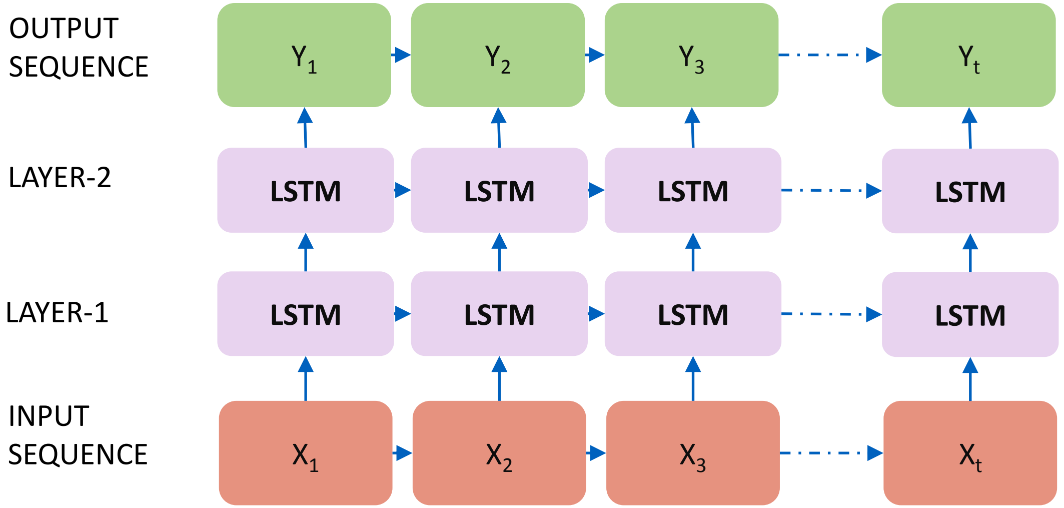


- System Description
  - 4 identical interconnected components of each system..
  - 10 sensors of each component.
  - The components either fail before mission time (1000 atu) or remain under normal operation.
- Descriptive Analysis
  - Normal behavior indicated by black line and abnormal behavior indicated by red line.
  - Change point depends upon multiple sensors and is not intuitively identifiable.
  - 39 unique sequences of onset of abnormal condition are obtained from the given 200 systems showing weak interdependency

Time bracket (atu)	number of times component fails under a particular time bracket			
	Component 1	Component 2	Component 3	Component 4
[800, 850)	9	10	6	7
[850, 900)	25	21	28	37
[900, 950)	42	42	48	35
[950, 1000)	30	25	21	25
No Failure	94	102	97	96

atu- arbitrary time units

METHOD-1, LSTM-IC



LSTM-IC Architecture

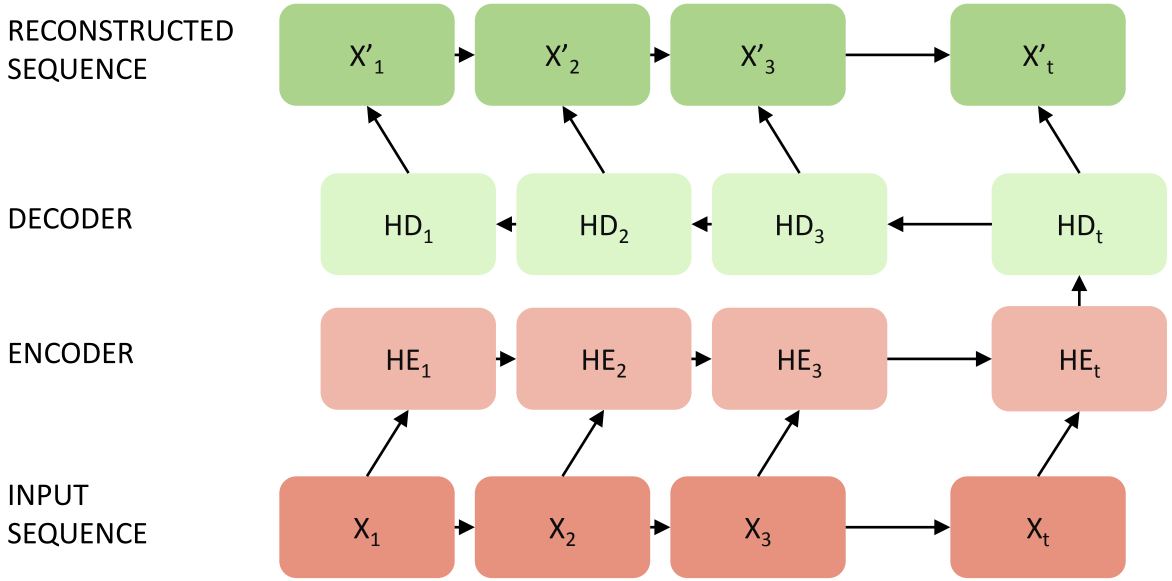
- LSTM many-to-many network with each input have an output label
- Component's change point is considered independent of each other.
- Selected sensors of a component are used as input.
- Suitable algorithm for weak to no interdependency between components

Timeliness error* of LSTM-IC

Method	Train Error	Test Error	Total Error
LSTM-IC (10 sensors)	0.0394	0.08	0.0512
LSTM-IC (stationary sensors)	0.0081	0.021	0.0117

* Described in Appendix

METHOD-2, LSTM-ED



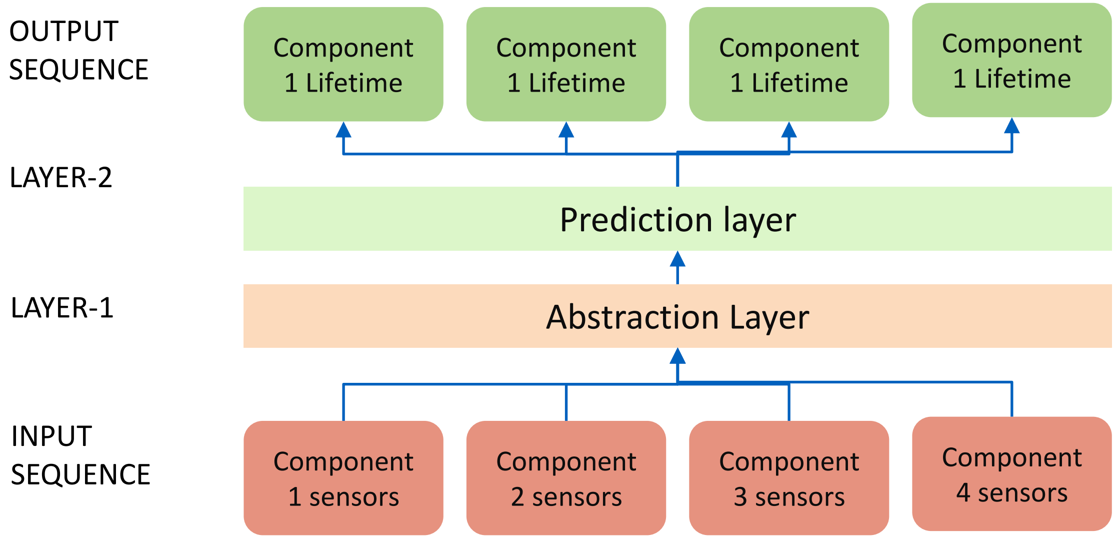
LSTM-ED Architecture

- LSTM based Encoder Decoder
- Component's change point is considered independent of each other.
- Data obtained before the change point of all the sensors is reconstructed.

Timeliness error table

Method	Train Error	Test Error	Total Error
LSTM (10 sensors)	0.0394	0.08	0.0512
LSTM (stationary sensors)	0.0081	0.021	0.0117
LSTM-ED	0.4423	0.417	0.435

METHOD-3, LSTM-MDA



The timeliness error of LSTM-IC is the least in all the 3 methods used.

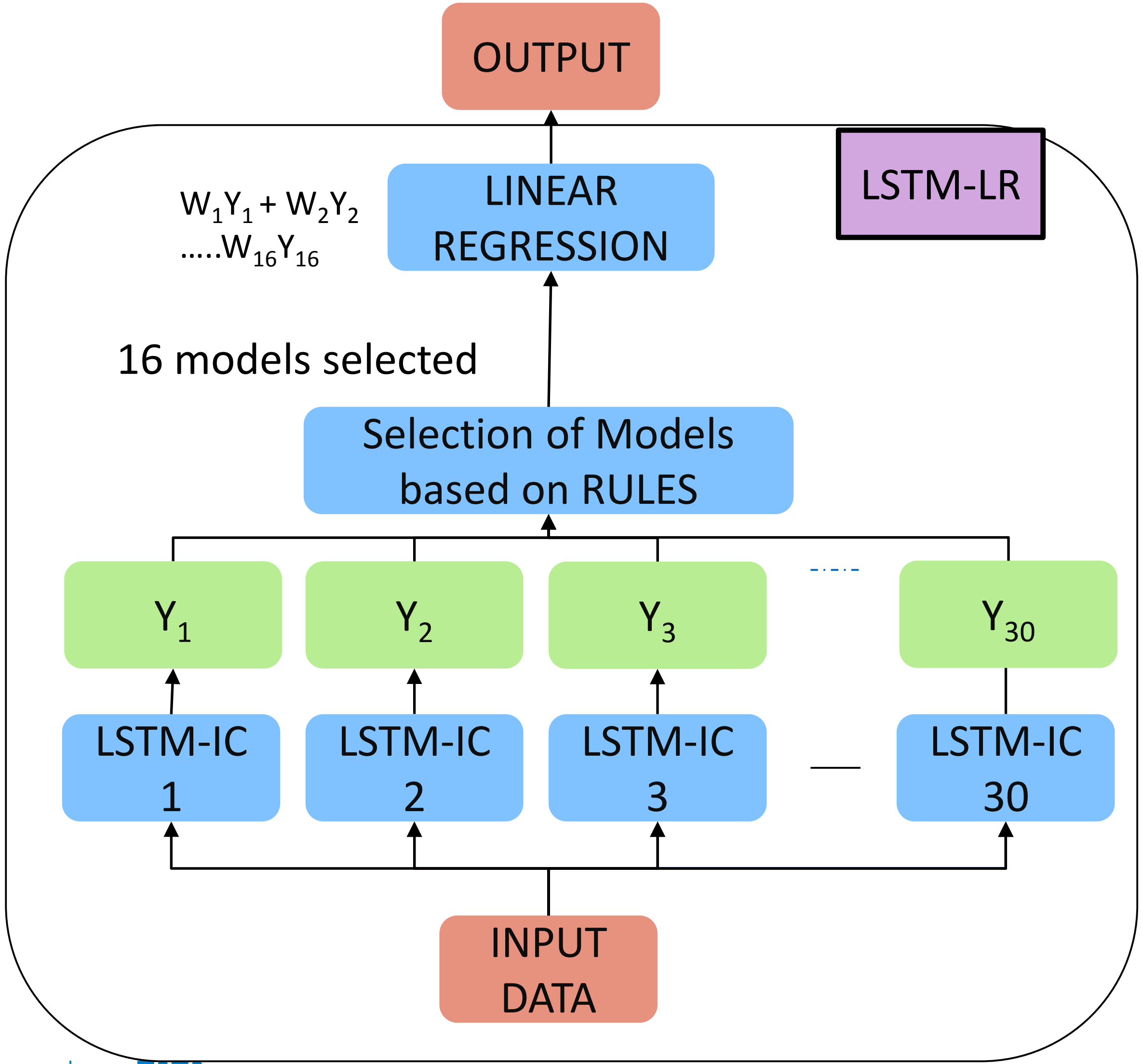
LSTM-MDA Architecture

- Component's change point is considered dependent of each other.
- Selected sensors of all components are fed to the network.
- Suitable for high interdependency between components

Timeliness error table

Method	Train Error	Test Error	Total Error
LSTM (10 sensors)	0.0394	0.08	0.0512
LSTM (stationary sensors)	0.0081	0.021	0.0117
LSTM-ED	0.4423	0.417	0.435
LSTM-MDA	0.0585	0.2385	0.1089

LSTM- Ensemble Model



Why Ensemble?

- Stand alone LSTM-IC model had missed and false alarms contributing to high timeliness error.
- Stand alone models have high variance between train and test error
- An ensemble model on top of LSTM-IC is built to minimize the timeliness error

RULE-1 (Missed Alarms)

- Models unable to identify change point at least once are rejected

RULE-2 (False Alarms)

- All the models considered for ensemble must predict an alarm, else no alarm.

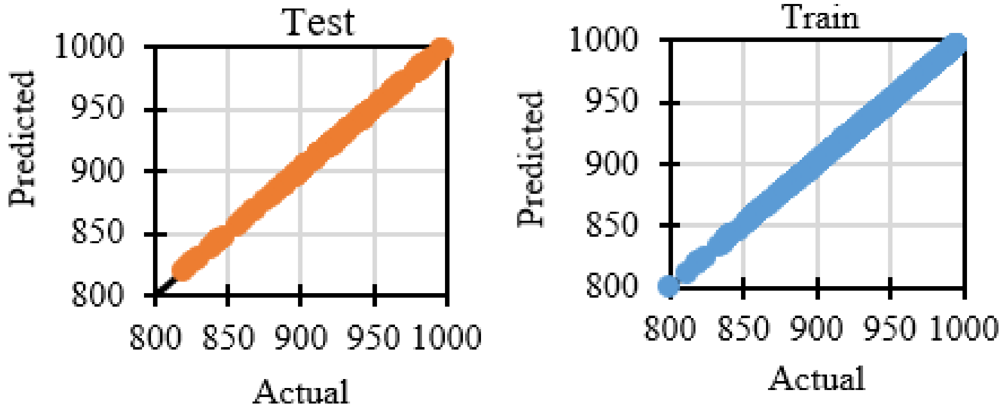
ENSEMBLING TECHNIQUES

- Median (LSTM-median)
- Linear Regression (LSTM-LR)

Results

- Timeliness error for all models with a train to test split of systems as 72:28.
- Error for LSTM-IC is lower than LSTM-MDA indicating weak interdependency between components.
- Use of only Stationary sensors reduces error for LSTM-IC by 73%.
- LSTM-median ensemble model decreases the error over LSTM-IC model by 34%.
- LSTM-LR model has the minimum error of 0.0086 on the test data

Parity plots of actual time of change point vs predicted time of change point for LSTM-LR model

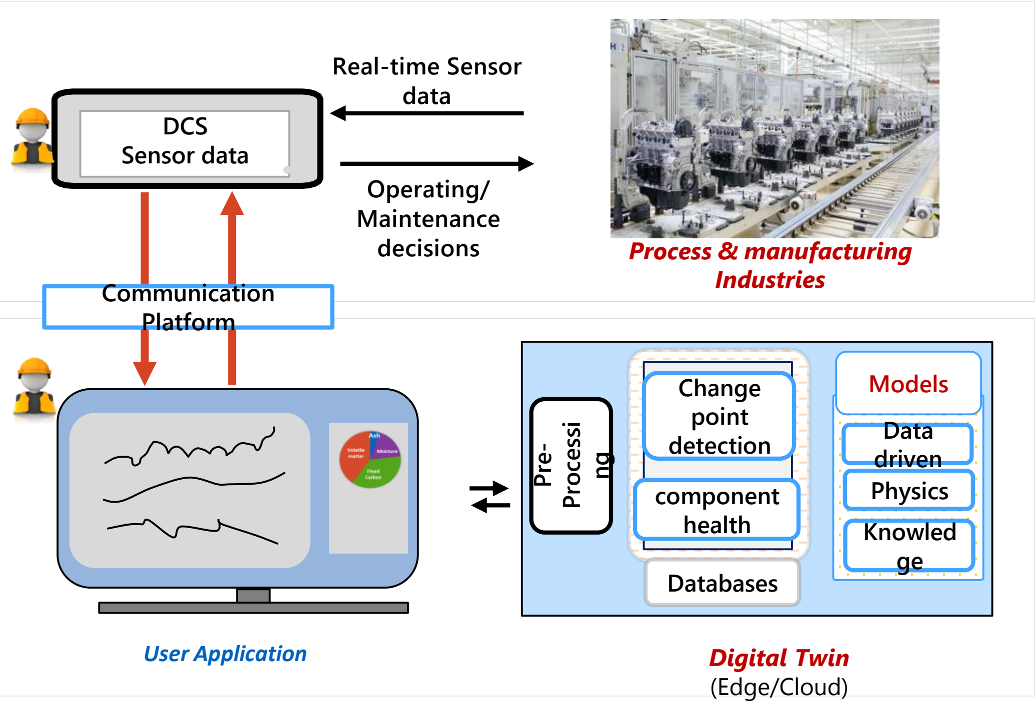


Method	Train Error	Test Error	Total Error
LSTM (10 sensors)	0.0394	0.08	0.0512
LSTM (stationary sensors)	0.0081	0.021	0.0117
LSTM-ED	0.4423	0.417	0.435
LSTM-MDA	0.0585	0.2385	0.1089
LSTM- median	0.0085	0.0137	0.0099
LSTM-LR	0.0084	0.0086	0.0085

Conclusion / Future Work

Conclusions

- ✓ Comparison of 3 deep learning methods is done for identifying the change point.
- ✓ LSTM-IC method with stationarity of sensors turns out to have the least error.
- ✓ Linear Regression ensemble on top of LSTM-IC output provides the least error with least variance between train and test.



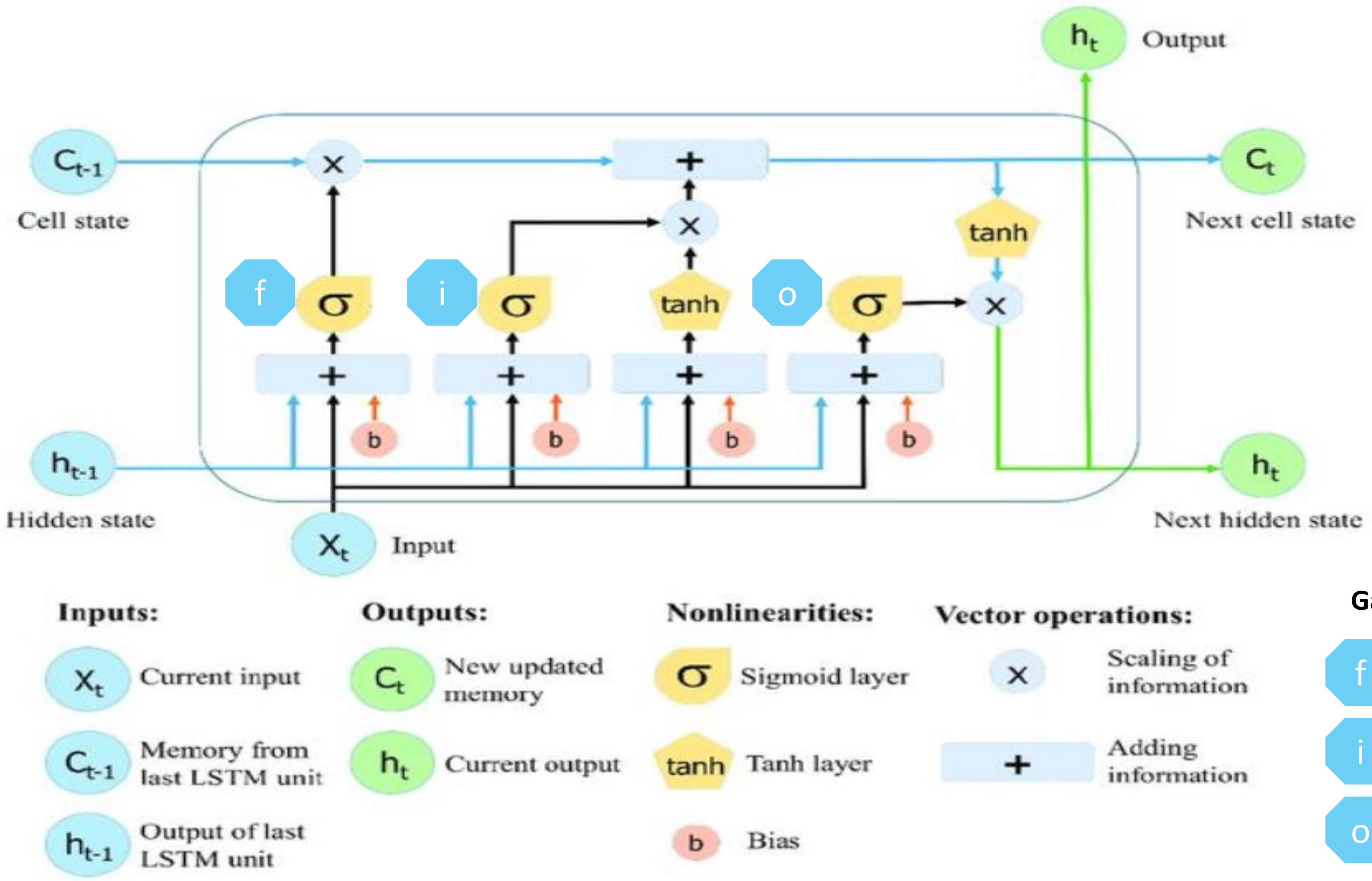
Solution as a DIGITAL TWIN

- ✓ Real Time detection of change point/anomaly of industrial equipment
- ✓ Ensure health management of equipment and reduce unplanned shutdowns.
- ✓ Take preventive and corrective measures.
- ✓ Avoid loss in revenue, productivity & reputation

THANK YOU

- APPENDIX

Long Short-Term Memory

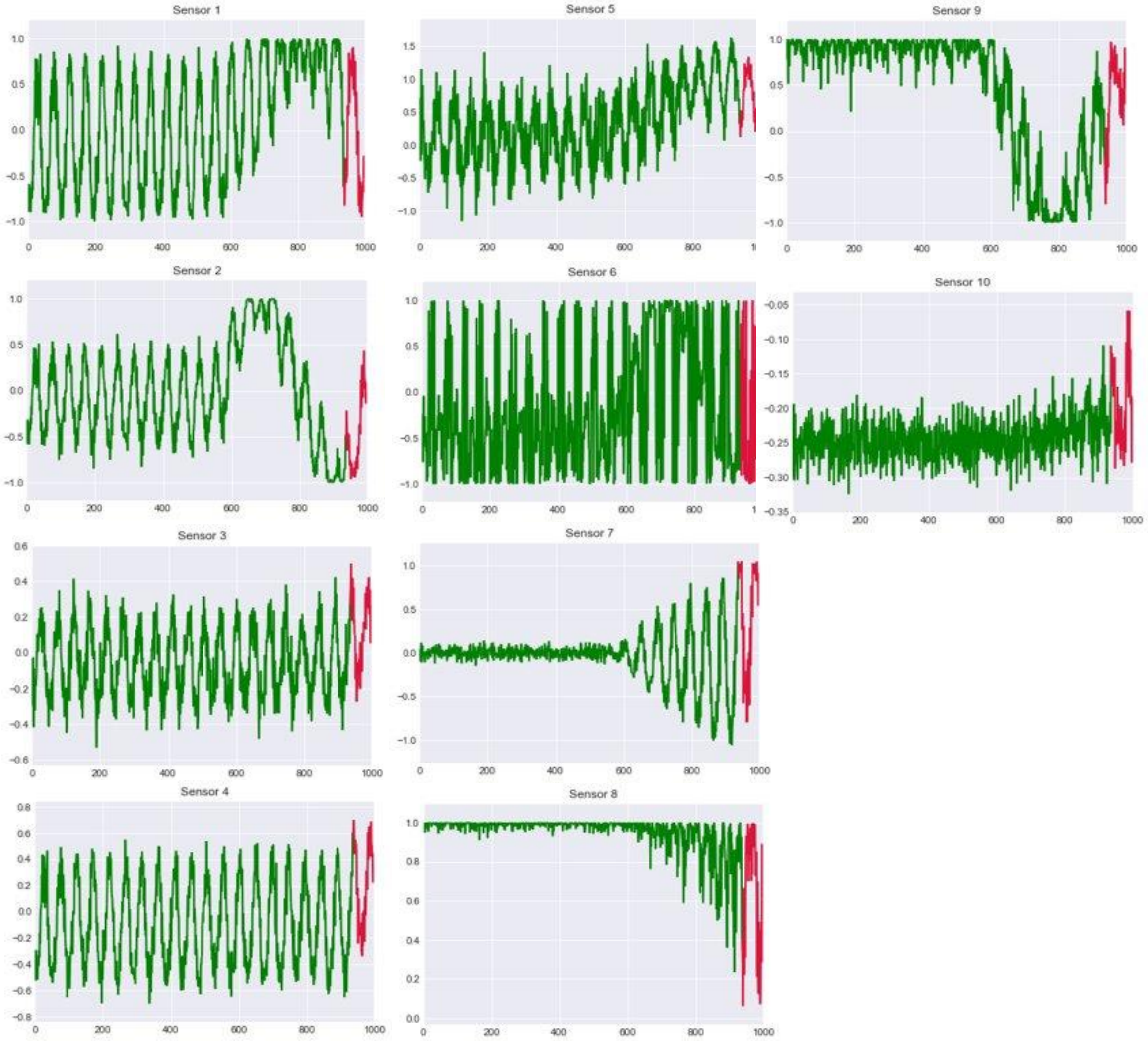


Gate classification

- f Forget gate
- i Input gate
- o Output gate

https://www.researchgate.net/publication/334268507_Application_of_Long_Short-Term_Memory_LSTM_Neural_Network_for_Flood_Forecasting/figures?lo=1

System Description



- System Description
  - 4 identical interconnected components.
  - 10 sensors of each component.
  - Normal operation indicated by ‘0’ and abnormal by ‘1’.
  - The components either fail before mission time (Tm) or remain under normal operation.

- Timeliness Error / Performance Metric

$$\Delta^{j,m} = \begin{cases} \tau^{j,m} - \hat{\tau}^{j,m} & \tau^{j,m} \neq \text{NaN}, \hat{\tau}^{j,m} \neq \text{NaN} \\ 0 & \tau^{j,m} = \text{NaN}, \hat{\tau}^{j,m} = \text{NaN} \\ k_{false} & \tau^{j,m} = \text{NaN}, \hat{\tau}^{j,m} \neq \text{NaN} \\ -k_{missed} & \tau^{j,m} \neq \text{NaN}, \hat{\tau}^{j,m} = \text{NaN} \end{cases} \quad j = 1,..,J; m = 1, \dots, M$$

$$A = \frac{1}{4M_{test}} \sum_{m=1}^{M_{test}} \sum_{j=1}^{4} \varphi(\Delta^{j,m}) \in [0,1] \qquad \begin{aligned} b_1 &= 1/\left(1 - e^{-T/a_1}\right) \\ b_2 &= 1/\left(1 - e^{-T/a_2}\right) \end{aligned}$$

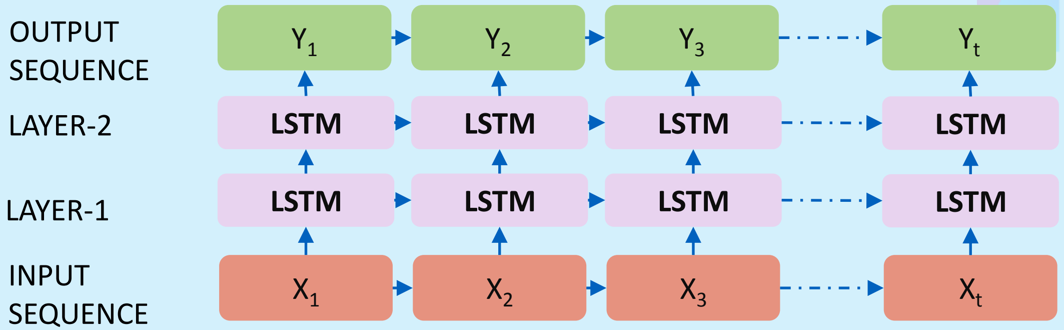
$$\varphi(\Delta^{j,m}) = \begin{cases} 1 & \Delta^{j,m} < -T \\ \left(1 - e^{\Delta^{j,m}/a_1}\right) b_1 & -T \leq \Delta^{j,m} < 0 \\ \left(1 - e^{-\Delta^{j,m}/a_2}\right) b_2 & 0 \leq \Delta^{j,m} \leq T \\ 1 & \Delta^{j,m} > T \end{cases} \quad j = 1,..,J; m = 1, \dots, M$$

Methodology: Long Short-Term Memory

Method-1

LSTM-IC Architecture

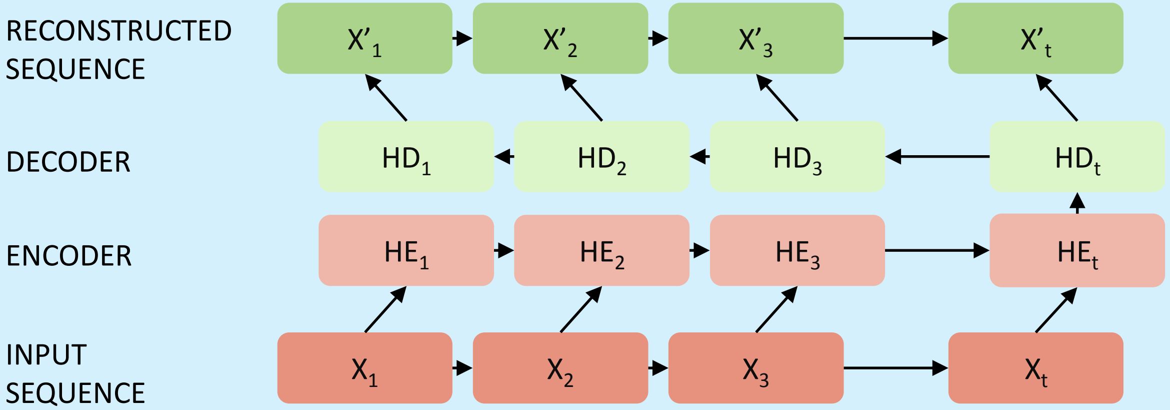
- Component's change point is considered independent of each other.
- Selected sensors of a component are used as input.
- Suitable algorithm for weak to no interdependency between components



Method-2

LSTM-ED Architecture

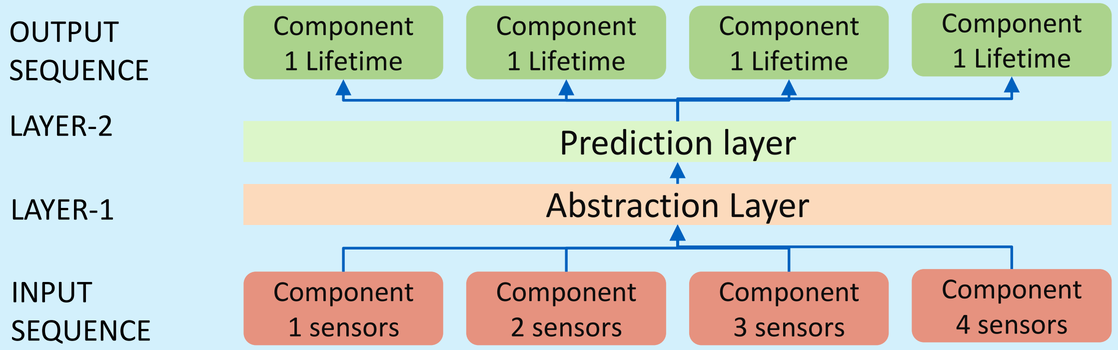
- Component's change point is considered independent of each other.
- Data obtained before the change point of all the sensors is reconstructed.



Method-3

LSTM-MDA Architecture

- Component's change point is considered dependent of each other.
- Selected sensors of all components are fed to the network.
- Suitable for high interdependency between components



Method	Model Names	Train Error	Test Error	Total Error
LSTM (10 sensors)	Model 1	0.0394	0.08	0.0512
LSTM (stationary sensors)	Model 2	0.0081	0.021	0.0117
LSTM-ED	Model 3	0.4423	0.417	0.435
LSTM-MDA	Model 4	0.0585	0.2385	0.1089
LSTM- median	Model 5	0.0085	0.0137	0.0099
LSTM-LR	Model 6	0.0084	0.0086	0.0085