RECURRENT NEURAL NETWORK BASED AIR POLLUTION FORECAST

HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

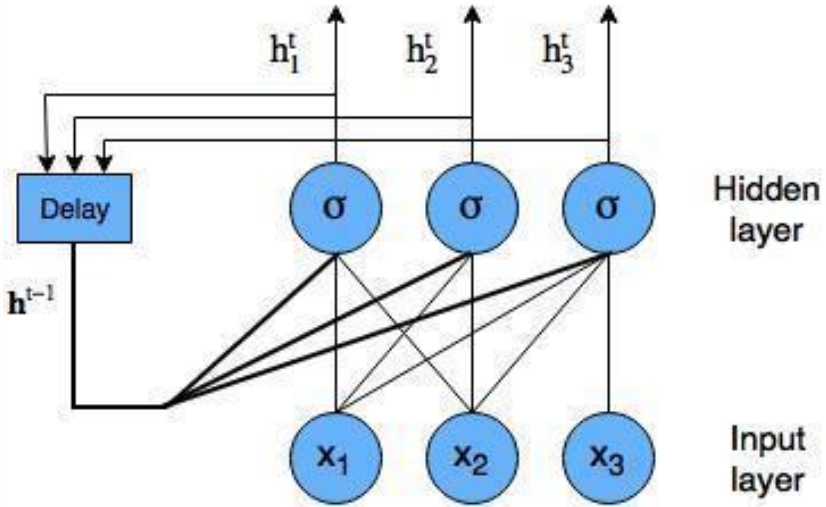
HIEN VU - SPARC Lab

Contents

RNNs and Applications

❑ Recurrent Neural Network (RNN)

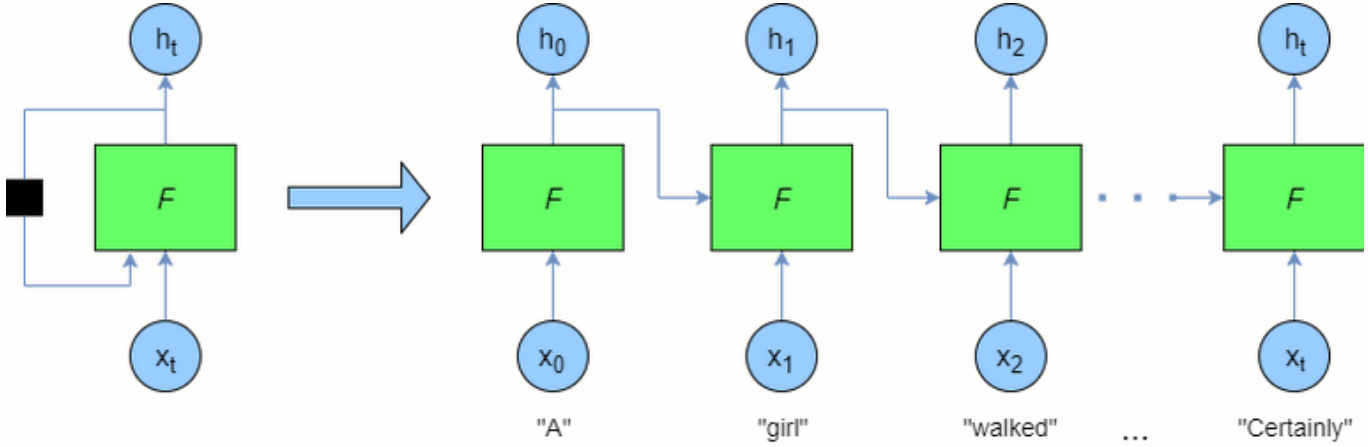
- A neural network that attempts to ***model time or sequence dependent behavior*** – such as language, stock prices, electricity demand and so on
- This is performed by feeding back the output of a neural network layer at time *t* to the input of the same network layer at time *t + 1*.



Recurrent Neural Network

❑ Network Model

- Recurrent neural networks are “unrolled” programmatically during training and prediction:



Recurrent Neural Network

❑ Vanishing Gradient Problem

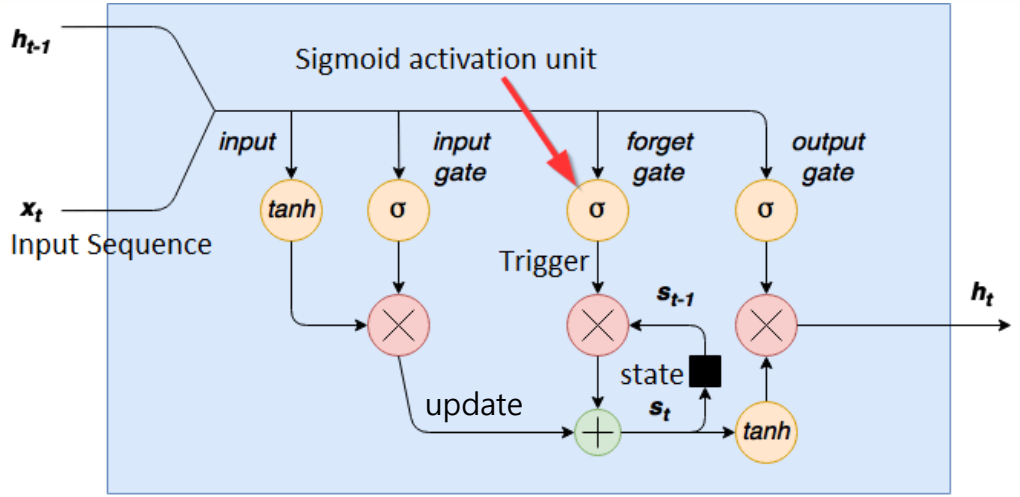
- When a small gradient or weight (values less than 1) are multiplied many times through the multiple time steps, and the gradients shrink asymptotically to zero.
- This means the weights of those earlier layers won't be changed significantly and therefore the network won't learn long-term dependencies.

Long Short-Term Memory (LSTM)

❑ Solution for Vanishing Gradient Problem

- An ***LSTM network*** is a recurrent neural network that is trained by using ***Backpropagation Through Time*** and over come ***Vanishing Gradient Problem***.
- Instead of neurons, LSTMs have ***memory blocks*** that are connected through layers.
- A block contains gates that manage the block's state and output.

An LSTM block



Long Short-Term Memory (LSTM)

- Input Gate is squashed between -1 and 1 using a *tanh* activation function

$$g = tanh(b^g + x_t U^g + h_{t-1} V^g)$$

$$i = б(b^i + x_t U^i + h_{t-1} V^i)$$

- Forget gate and state loop

$$f = б(b^f + x_t U^f + h_{t-1} V^f)$$

$$s_t = s_{t-1} \circ f + g \circ i$$

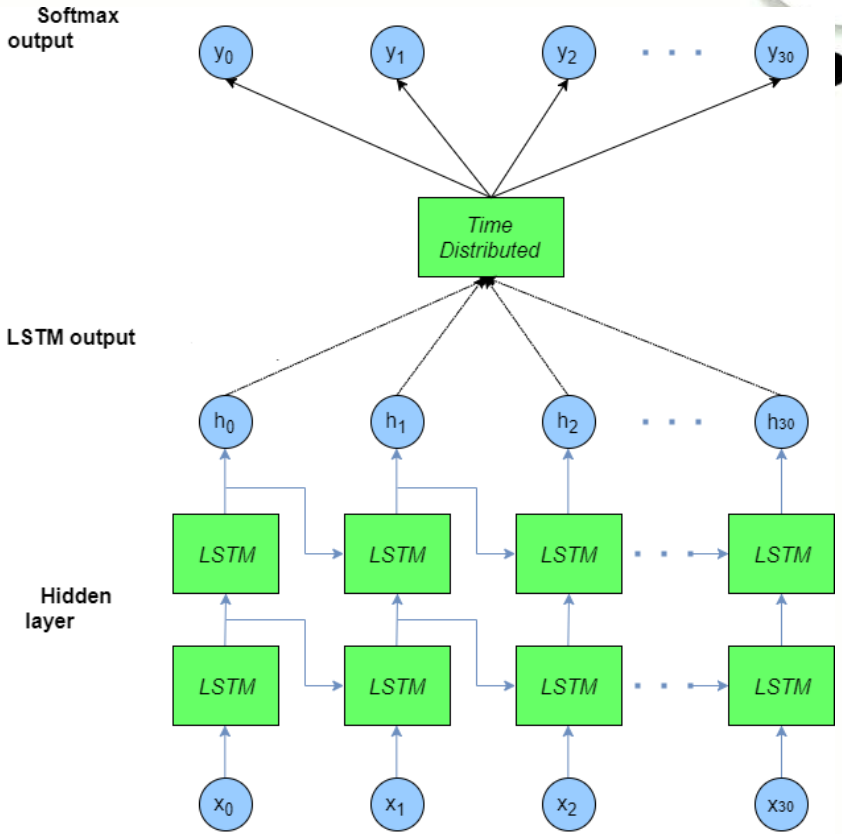
- Output gate

$$o = б(b^o + x_t U^o + h_{t-1} V^o)$$

$$h_t = tanh(s_t \circ o)$$

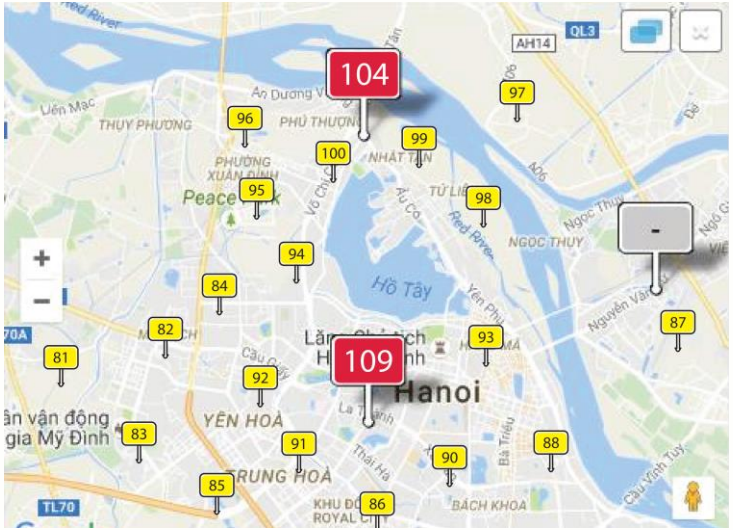
The Keras LSTM Architecture

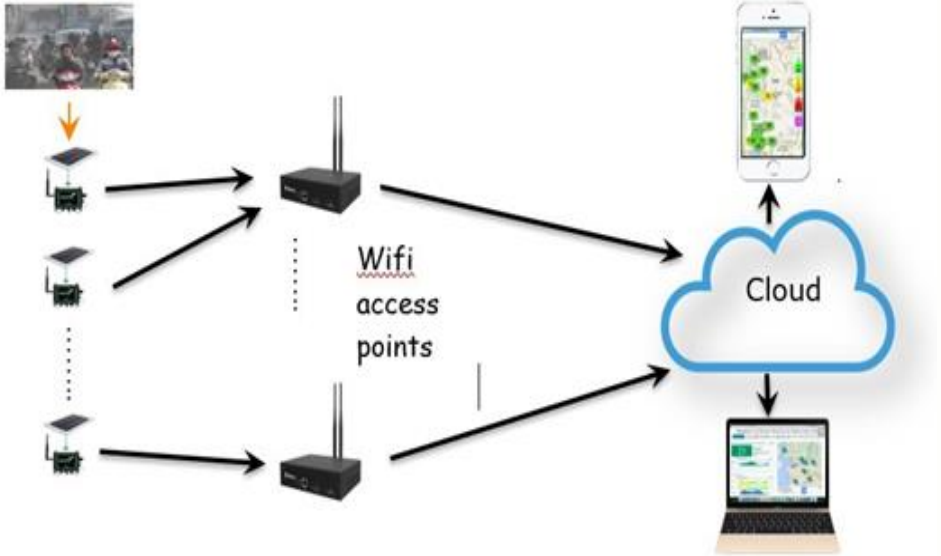
- The input shape of the data is batch size, number of time steps, hidden size
- The output is compared to the training *y* data for each batch.
- The error and gradient back propagation is performed from each comparison in Keras.



Air Pollution Forecast System

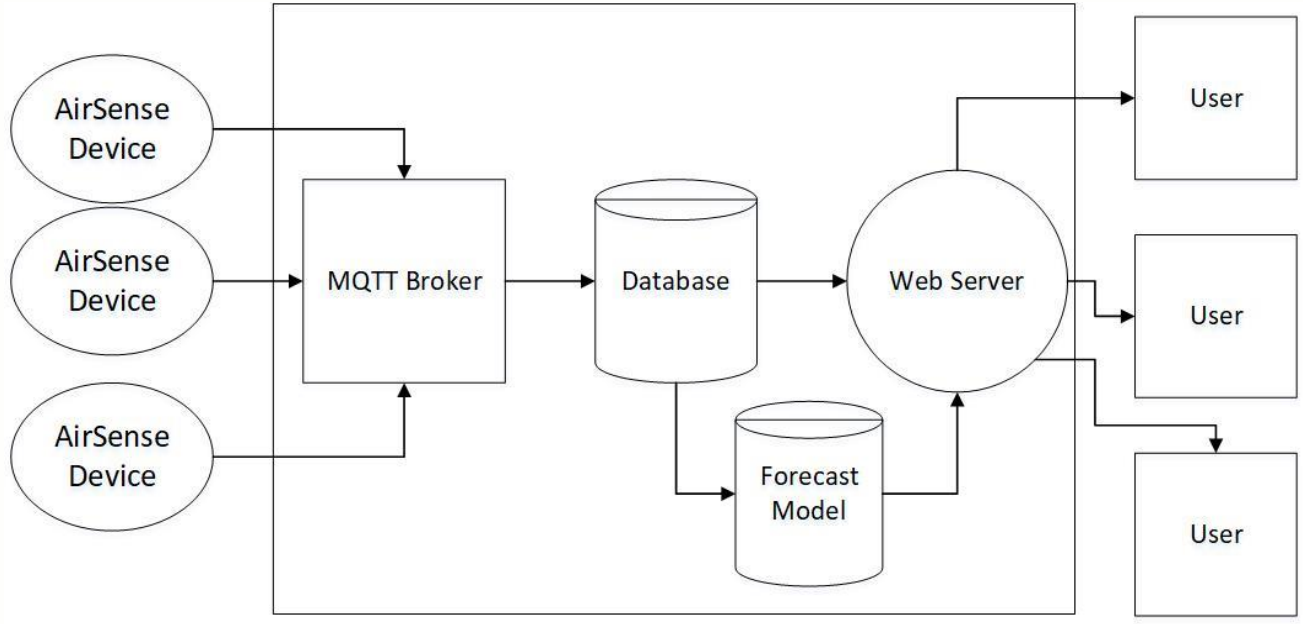
❑ Data Collecting System





Air Pollution Forecast System

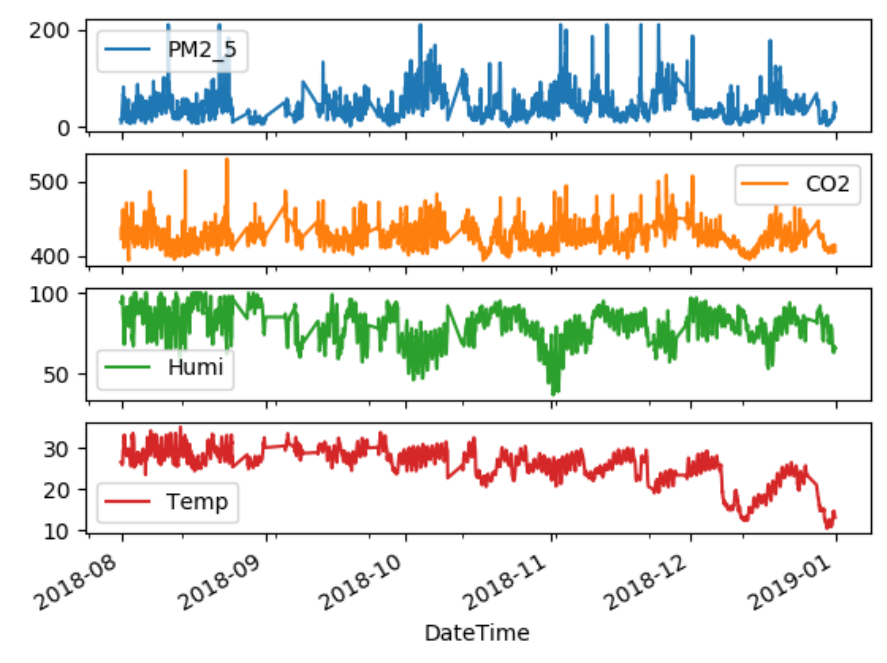
❑ System Model



Data Processing

❑ Hanoi Air Pollution Data

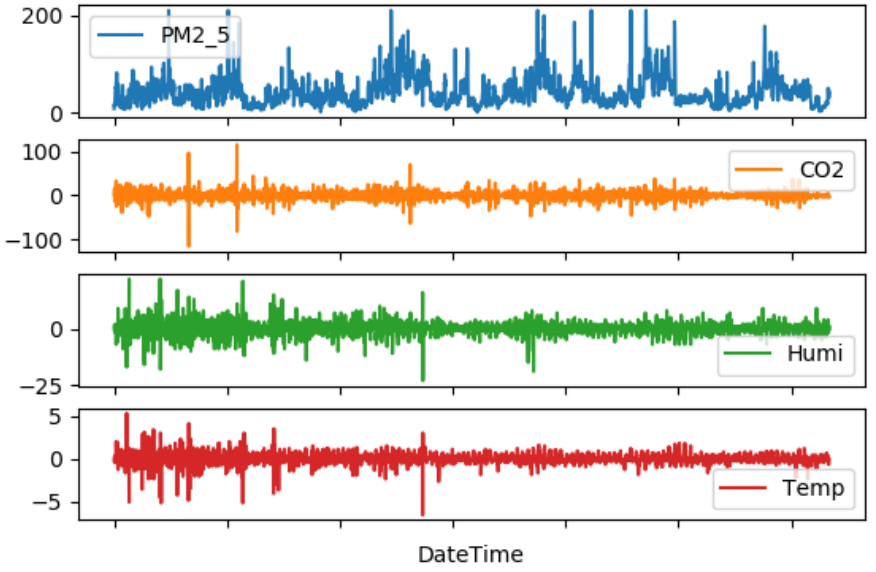
- Data from a specific node in 5 months of observation
- The recorded data: PM2.5, CO2, Humility and Temperature
- From more than 1M samples, the data was averaged hourly



Data & Training Model

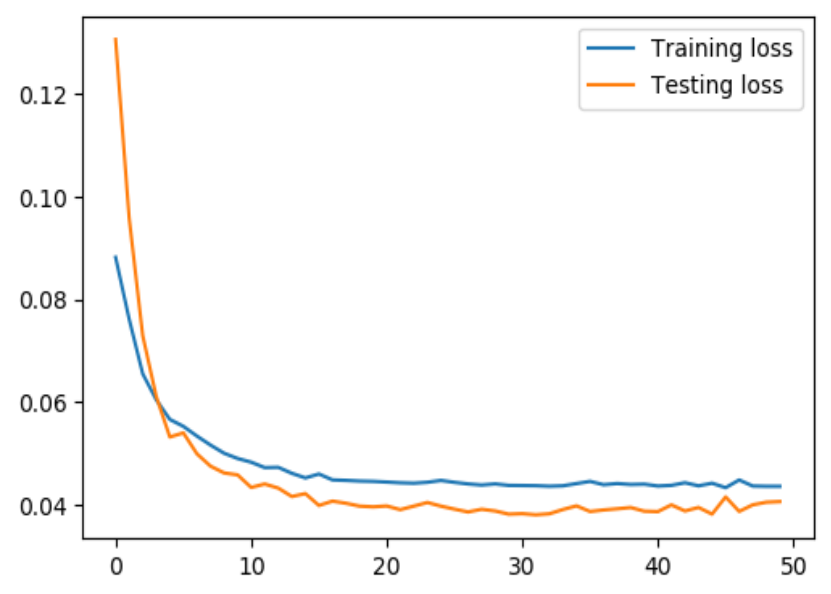
❑ Training Profile

- **Input:** Differentiated data from data set, 50% for training, 50% for testing
- **An RNN:**
  - Consists of 100 layers
  - Trained over 50 epochs
  - Mini-Batch Gradient Descent: 32
  - Uses 4 input variables for each time steps
  - Predicts PM2.5 concentration over 24 time steps ahead

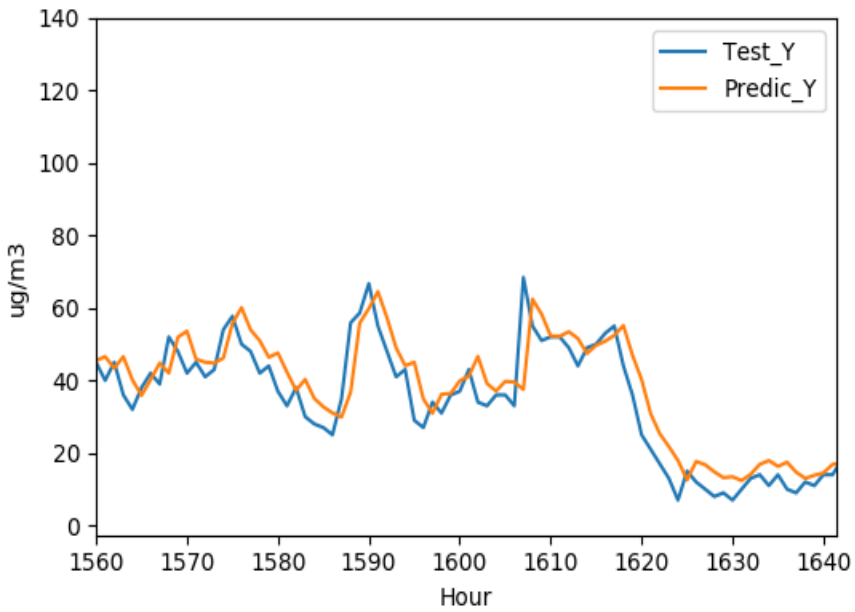


Results

❑ Training Performance



❑ Forecast Results for 24h



Conclusion

❑ Predicting Air Pollution by LSTM

- Recurrent Neural Network has showed a good performance in predicting time-series problems
- The Air Pollution can be predicted with high accuracy by using LSTM
- The system needs to take into account more variables to improve the long-term predictability of LSTM in this application

Thank you!