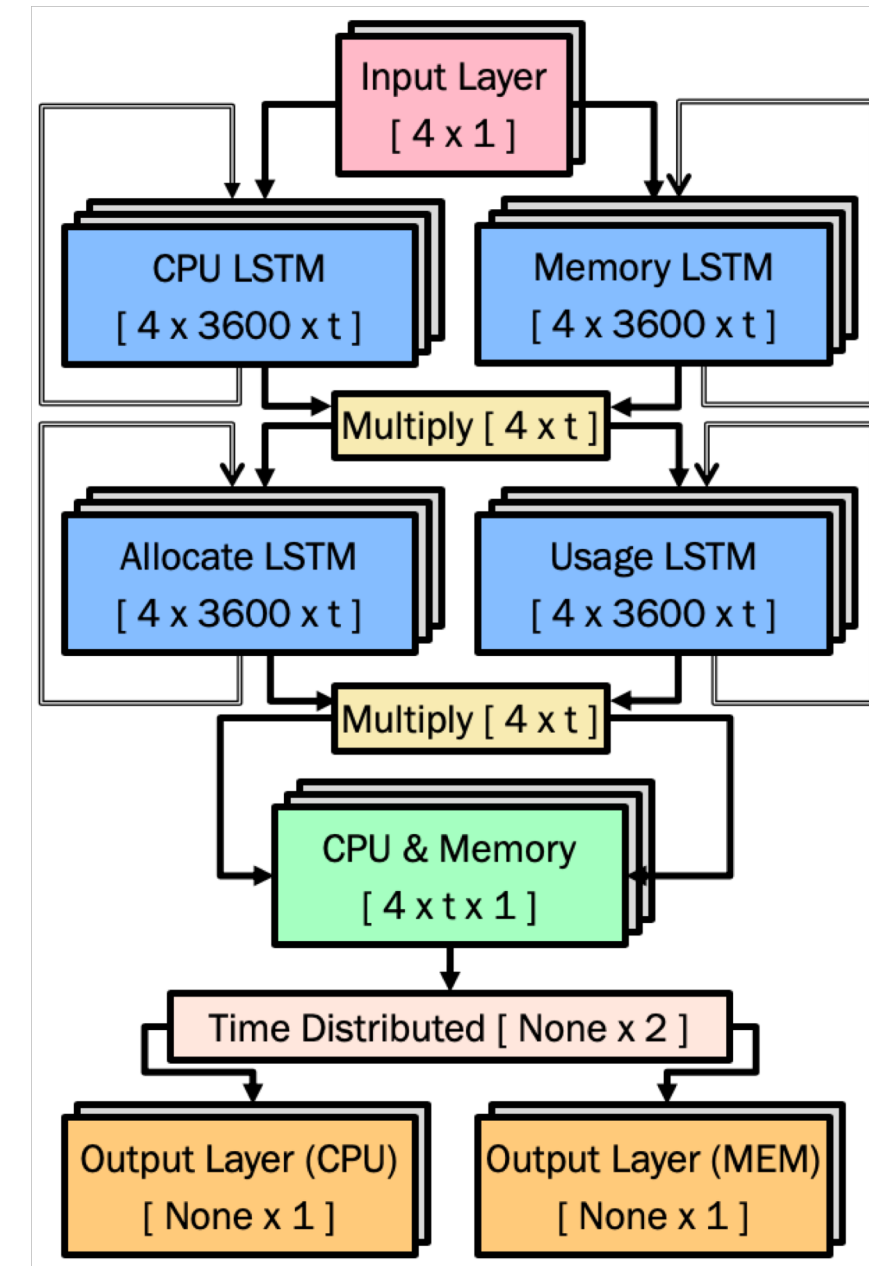# Towards Optimal Resource Utilization in Data Centers using Long Short-Term Memory

Kundjanasith Thonglek, Kohei Ichikawa, Keichi Takahashi, Chawanat Nakasan, Hajimu Iida

- <u>Goal</u>: Improve resource utilization in data centers

- <u>Methodology</u>:

  - Analyze Google's cluster usage dataset using PCA and select the important features
  - Design and implement DNN model based on LSTM to predict optimal resource allocation

- <u>Evaluation</u>:

  - Simulated resource usage and allocated resource
  - Improved resource utilization rate for 17.98%

20 minutes   **40 minutes**   **60 minutes**

CPU (%) — actual usage / actual allocation / predicted allocation

Memory (%) — actual usage / actual allocation / predicted allocation

DNN model based on LSTM:

- Input Layer [4 x 1]
- CPU LSTM [4 x 3600 x t]
- Memory LSTM [4 x 3600 x t]
- Multiply [4 x t]
- Allocate LSTM [4 x 3600 x t]
- Usage LSTM [4 x 3600 x t]
- Multiply [4 x t]
- CPU & Memory [4 x t x 1]
- Time Distributed [None x 2]
- Output Layer (CPU) [None x 1]
- Output Layer (MEM) [None x 1]

DNN model based on LSTM to predict optimal resource allocation