SI670: Project Proposal

Machine Learning Application in PUE optimization problem with small server room

Lee, Yung-Chun
Department of Statistics
yuncglee@umich.edu

Wu, Amber
Department of Economics
wuca@umich.edu

October 2019

1 Introduction

The number of Data center has grown exponentially every since we entered the data era, Google has published its results in successfully reducing the total power usage by increasing the PUE (Power usage effectiveness) by using Machine Learning. In this study, we are interested in if we can use the similar approach and apply their method onto a smaller and more common setup - server room. Server room are fairly common for medium even a small size company to maintain there data or provide online services. We aim to find out given the environment parameters such as temperature, humidity, and CO_2 level in the server room with proper power usage monitor the overall power usage can be optimized through machine learning algorithm.

2 Data Source

Our data is provided by the courtesy of WebEnv IoT, a company founded in 2000 in Tainan, Taiwan that first started as a manufacturer to design and produce network equipment for server room environmental monitoring, and over the last decade, it dedicated to IT systems integration services and network devices solutions under the information security environment.

The data is collected in each of the controller installed in the server room (depending on the customer demand and budgets different model and spec will be used)and reported to the company server every 5 minutes over 2 years from Oct/01/2017 - Sep/31/2019 as long as the connection is not interrupted or whenever a special event is trigger such as authorized personnel entered the server room, abnormal temperature detected in the server room. The data is a time series that consists of the following key variables:

Variable	Data Type	Description
controller Model	categorical	controller model of which the data is collected
controller Serial number	categorical	controller unique serial number
location	categorical	server room which the controller is installed
timestamp	datetime	time recorded, in GMT +8 time zone
controller temperature	float	recording from the built-in temperature censor, °C
controller humidity	float	recording from the built-in humidity censor, 0-100%
AI #1-12	numerical	the current/ voltage reading of a device
AO #1-12	numerical	the output node
AC status #1-n	binary	on/off status of the air conditioner
AC temperature #1-n	float	built-in value from air conditioner
AC humidity #1-n	float	built-in value from air conditioner
fan status #1-n	binary	on/off status of the heat fan

3 Methodology

We attempt to use 3-layer neural net work from the Google PUE paper from the literature.

4 Time line

The tentative time table as follow:

Week	Date	Tentative schedule
1	10/22-10/27	extracting data from database, consult engineers and setup expectation of improvement
2	10/28-11/03	Data cleaning ,preprocessing and literature review
3	11/04-11/10	Exploratory analysis, Data quality testing, neural net week1- layer and structural building
4	11/11-11/17	neural net week2- model training
5	11/18-11/24	neural net week3- tuning and testing performance
6	11/25-12/01	simulation
7	12/02 -	poster session and paper writeups