IERG4998R Final Year Project 1

Machine Learning and Big Data Analytics for Load Profiling in Smart Power Grids

By

Chu Kai Ming

1155094746

A FINAL YEAR PROJECT REPORT
SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF BACHELOR OF INFORMATION ENGINEERING
DEPARTMENT OF INFORMATION ENGINEERING
THE CHINESE UNIVERSITY OF HONG KONG

December, 2019

Abstract

This is a project verifying the application of machine learning and big data analytics for load profiling in smart power grids. Dynamic pricing is a solution of resolving the problem of ineffectiveness in electricity retail market. By applying the Q-Learning algorithm of reinforcement learning, we can merge the usage of machine learning with big data analysis to estimate several predictions to a more effective wholesale price. According to the basic law of demand in economics, effective dynamic pricing strategy could find the better balance point between demand and supply. Applying these techniques for maximizing the profit of the electricity retailer with consideration of customers to minimize their own cost.

This report covers the basic algorithm and explanation of mathematical equations with results of the analysis on a specific day of a customer. The ultimate goal of the project is to discover the most effective dynamic prices to create a highly efficient economical model of electricity market.

Table of Contents

1. Introduction

1.1 Background

In the 21st century, environmental issues are one of the most discussed issues as the effect of climate change and shortage of energy are being more critical on the earth. According to United Nations, there are 17 sustainable development goals to be achieved before the year of 2030. Two of the goals are “Goal 7: Affordable and Clean Energy” and “Goal 13: Climate Action” [1-3]. To achieve these two goals, one of the solutions is enhance the efficiency of electricity usage.

Through machine learning and big data analytics, the electricity supplier could set up a more efficient wholesale price for the retailer to achieve a higher profit, while the consumers could use reasonable and affordable retail price to maintain their quality of life. When the price of a product is effective in an economical model, the demand and supply of this product will be efficient [13]. By applying machine learning and big data analysis, we can adjust the price according to our collected data to achieve a balanced and efficient energy usage on the market. An effective dynamic pricing solution could also prevent overload of power girds.

1.2 Reinforcement Learning

Reinforcement learning is one of the machine learning methods. It involved an agent and an environment. The agent(program) would give actions without any rules or regulations and the environment would return rewards or states to the agent. There is no example for the agent to follow. The agent needs to achieve the best reward by trying and learning different actions to approach a better reward. The following diagram is the process of the reinforcement learning model.

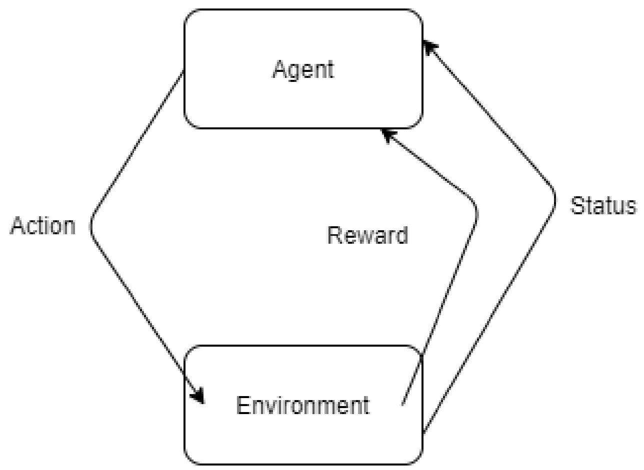


Fig. 1 Reinforcement Learning Model

Q-learning algorithm [10] is selected to be the algorithm of this project. The reason of using Q-learning is the model-free characteristic of the algorithm and there are pre-programmed python libraries and open-source resources [11] as references.

2. Methodology

2.1 Mathematics Representations of the electricity market

The usage of electricity in the city could be divided into two main components, critical load and curtailable load.

2.1.1 Critical load

Critical load [4] is the essential usage of electricity which has the highest priority and less flexibility to the price changes. For instance, the electricity usage in hospitals, agriculture, data centers, etc. The critical load usually would not be affected by price of electricity.

$$e_{t,n}^{critic} = E_{t,n}^{critic}$$

Equation for Critical Load

$e_{t,n}^{critic}$ *is the electricity consumption for critical load.*

$E_{t,n}^{critic}$ *is the electricity demand for critical load.*

*The **t** are the time slots.*
*The **n** is customers n.*

2.1.2 Curtailable load

Curtailable load [5] is electricity usage with more flexibility which would affect by the retail price of electricity. The curtailable load includes electricity usage of non-critical and non-emergency services such as home appliances, entertainments.

$$e_{t,n}^{curt} = E_{t,n}^{curt} \cdot \left(1 + \xi_t \cdot \frac{\lambda_{t,n} - \pi_t}{\pi_t}\right)$$

Equation for Curtailable load

$e_{t,n}^{curt}$ *is the electricity consumption for curtailable load.*

$E_{t,n}^{curt}$ *is the electricity demand for curtailable load.*

ξ_t *is the elasticity coefficient.*

$\lambda_{t,n}$ *is the retail price of electricity.*

π_t *is the wholesale price of electricity.*

*The **t** are the time slots.*
*The **n** is customers n.*

2.2 The goals of this project

The main focuses of this paper will be discovering the effect of dynamic pricing on curtailable load usage of electricity. Using reinforcement learning to predict the most efficient wholesale price and maximize the profit of the service provider while considering the customers' goal of minimizing their own costs. There is a **dissatisfaction cost** [6] representing the situation of the retail price of electricity dissatisfy the customer and decrease the usage of the curtailable load. This leads to the decrease of demand in the market.

$$\min \sum_{t=1}^{T} [\lambda_{t,n} \cdot (e_{t,n}^{curt} + e_{t,n}^{critic}) + \varphi_{t,n}]$$

*Function for the customers' **cost minimizing** goal [7]*

$\varphi_{t,n}$ *is dissatisfaction cost defined as* $\varphi_{t,n} = \frac{\alpha_n}{2}(E_{t,n}^{curt} - e_{t,n}^{curt})^2 + \beta_n(E_{t,n}^{curt} - e_{t,n}^{curt})$

where $\alpha_n > 0$ *and* $\beta_n > 0$

$$\max \sum_{n=1}^{N} \sum_{t=1}^{T} (\lambda_{t,n} - \pi_t) \cdot (e_{t,n}^{curt} + e_{t,n}^{critic})$$

*Function for service providers' **profit maximizing** goal [8]*

The function of the service provider required to consider the usage of both curtailable load and critical load and maximize the profit with the dynamic pricing.

2.3 Application of Reinforcement Learning

The application of reinforcement learning is to apply Q-Learning algorithm to study and estimate the best retail price that offer by service provider(agent) and interact with the response from market(environment).

$$\max \sum_{n=1}^{N} \sum_{t=1}^{T} [\rho \cdot (\lambda_{t,n} - \pi_t) \cdot e_{t,n} - (1-\rho) \cdot (\lambda_{t,n} \cdot e_{t,n} + \varphi_{t,n})]$$

Function of the electricity market [10]

The determinant of the machine learning program is ρ in the function which is the weighting factor constant between and including 0 and 1. The reinforcement learning model is shown in the diagram below.

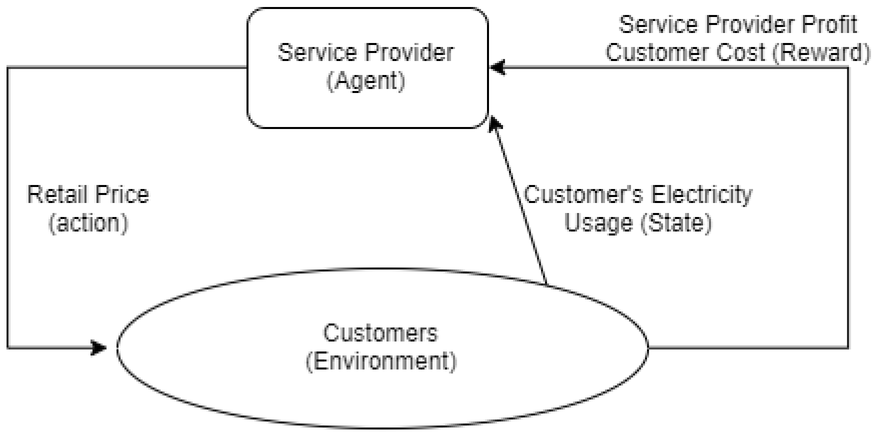


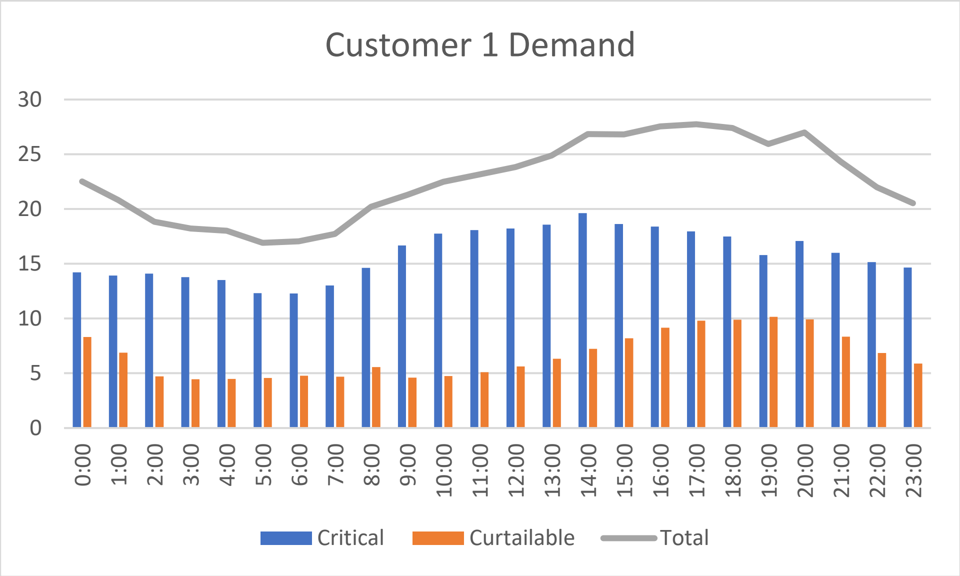
Fig.2 Reinforcement Learning Model (Electricity Retail Market)

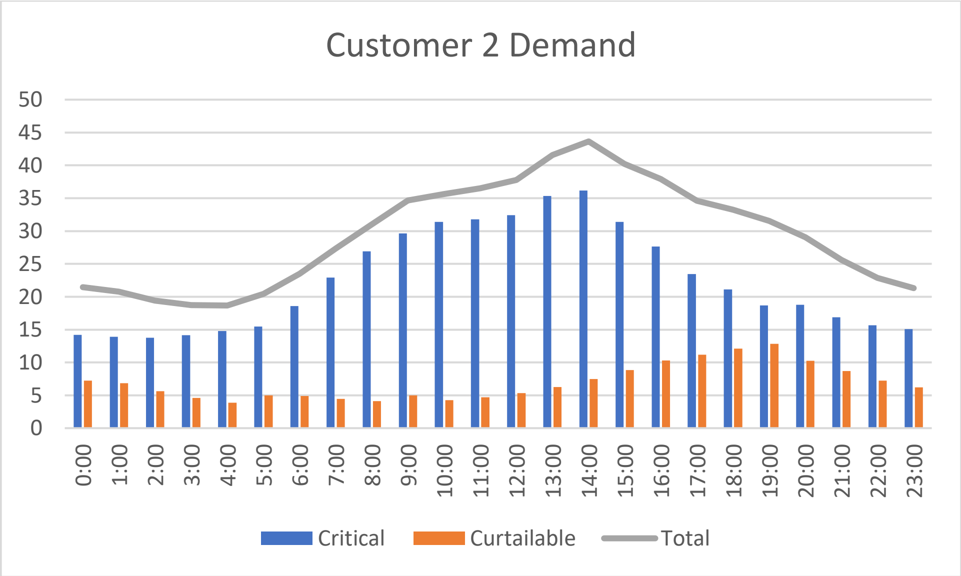
The process is a loop until the reward has been maximized. The service provider (Agent) will decide the dynamic retail price (action) without any policies or regulations. The customers (Environment) will response with customer's electricity usage (State) which involve the behavior of demand the customer. Service provider profit and customer cost (Reward) will also be send back to the agent, the machine learning program will analysis the collected responses from environment and keep looping the process until the reward cannot be increased by dynamic pricing of retail price.

3. Data sets

3.1 Demand of Customers

The demand data sets of customers were extracted from the San Diego Gas & Electric. We obtain 2 cases of customer demand which include the customer's electricity usage of critical load and curtailable load.





3.2 Wholesale Price

The data sets of electricity wholesale price were extracted from the website of Commonwealth Edison (ComEd) [12]. 31st October,2019(Thursday) was randomly chosen to be one of the wholesale price data sets.

Fig3. The Wholesale Price from other thesis [11]

Time	Wholesale Price
0:00	1.90
1:00	1.80
2:00	1.70
3:00	1.60
4:00	1.60
5:00	1.90
6:00	2.00
7:00	2.20
8:00	2.30
9:00	2.60
10:00	2.90
11:00	3.30
12:00	3.40
13:00	3.60
14:00	3.50
15:00	5.40
16:00	5.50
17:00	5.00
18:00	4.00
19:00	3.20
20:00	3.10
21:00	3.60
22:00	3.00
23:00	2.60

Fig4: The Wholesale Price on 2019/10/31

Time	Wholesale Price
0:00	2.20
1:00	2.00
2:00	2.00
3:00	2.00
4:00	2.20
5:00	2.20
6:00	2.40
7:00	3.20
8:00	2.40
9:00	3.10
10:00	2.50
11:00	2.90
12:00	2.60
13:00	2.70
14:00	2.60
15:00	2.60
16:00	2.70
17:00	2.50
18:00	2.50
19:00	4.80
20:00	1.90
21:00	2.20
22:00	2.20
23:00	2.20

4. Simulation Results

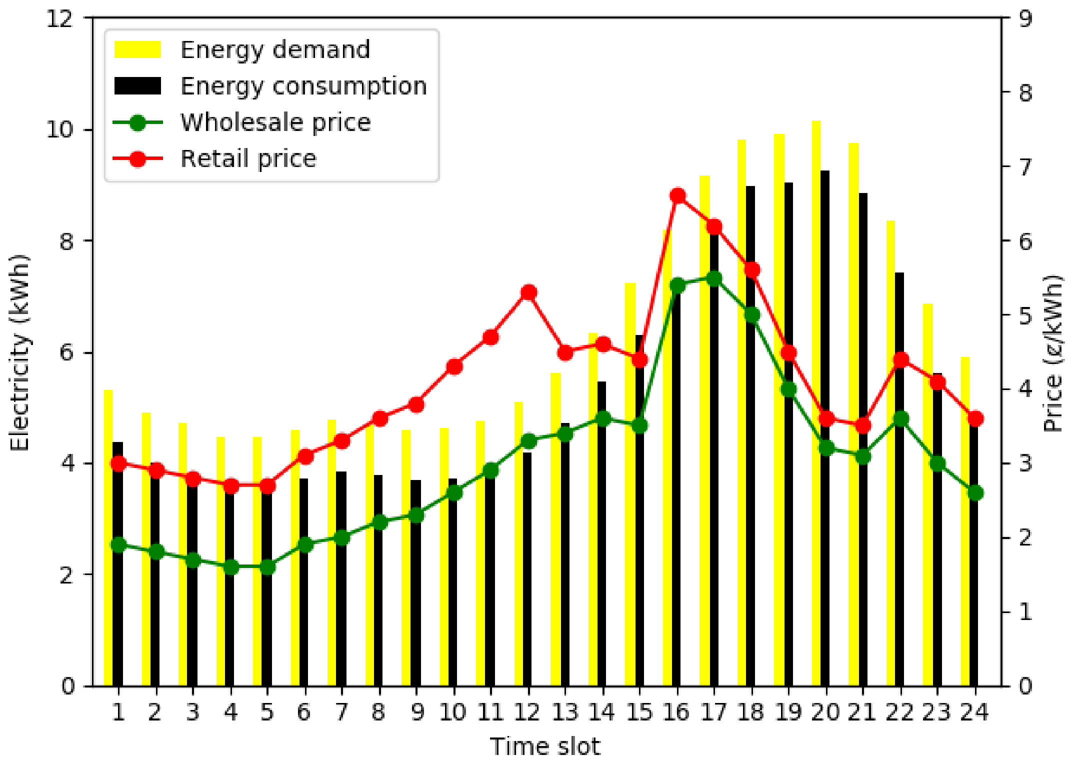


Fig5: Demand of Customer 1 with wholesale price on 2017/06/22

This simulation result is applying the data sets of demand of customer 1 and the wholesale price on the date of 2017/06/22. There were 7722 iterations executed by the algorithm with a reward of 140.791.

Log Example:

iteration 2,420; delta: 12.609796400058826; reward: -578.7350372110142...

[3.0, 2.9, 2.8, 3.2, 2.8, 3.3, 3.4, 4.7, 5.1, 6.7, 5.5, 6.1, 4.9, 5.0, 4.7, 7.1, 6.6, 5.9, 4.6, 3.6, 3.5, 4.7, 4.1, 3.6]

iteration 6,340; delta: 19.908668400822876; reward: -23.620014279520035...

[3.0, 2.9, 2.8, 2.7, 2.7, 3.8, 3.4, 3.7, 4.0, 4.5, 4.9, 5.5, 4.6, 4.7, 4.4, 6.7, 6.3, 5.7, 4.5, 3.6, 3.5, 4.4, 4.1, 3.6]

finished at iteration 7,722, with a delta of 0.08698304846473093...

[3.0, 2.9, 2.8, 2.7, 2.7, 3.1, 3.3, 3.6, 3.8, 4.3, 4.7, 5.3, 4.5, 4.6, 4.4, 6.6, 6.2, 5.6, 4.5, 3.6, 3.5, 4.4, 4.1, 3.6]

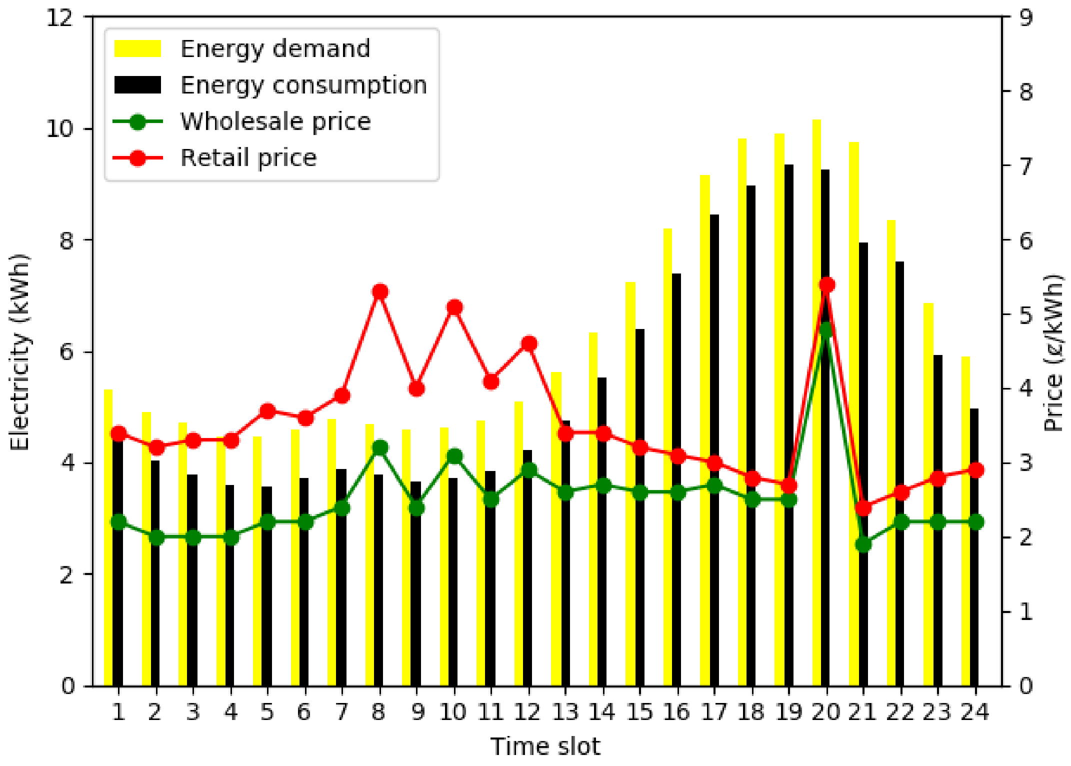


Fig6: Demand of Customer 1 with wholesale price on 2019/10/31

We changed the date of the wholesale price and apply the same customer's demand, the simulation result is applying the data sets of demand of customer 1 and the wholesale price on the date of 2019/10/31. There were 4495 iterations executed by the algorithm with a reward of 82.508.

Log Example:

iteration 140; delta: 1584.853143268501; reward: -1120.370179322612...

[3.4, 4.1, 3.4, 4.4, 5.0, 4.0, 4.3, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4]

iteration 2,320; delta: 146.3314864790537; reward: -600.3205840650633...

[3.4, 3.2, 3.3, 3.3, 4.8, 3.8, 4.1, 6.9, 4.3, 8.2, 4.8, 5.3, 3.8, 3.8, 3.5, 3.3, 3.0, 2.8, 2.7, 5.6, 2.4, 2.6, 2.8, 2.9]

iteration 3,460; delta: 19.837736568584546; reward: 37.72501374399319...

[3.4, 3.2, 3.3, 3.3, 3.7, 3.6, 3.9, 5.3, 4.0, 5.1, 4.1, 4.6, 3.4, 3.4, 3.2, 3.1, 3.0, 2.8, 2.7, 5.6, 2.4, 2.6, 2.4, 2.4]

finished at iteration 4,495, with a delta of 0.08856486200011204...

[3.4, 3.2, 3.3, 3.3, 3.7, 3.6, 3.9, 5.3, 4.0, 5.1, 4.1, 4.6, 3.4, 3.4, 3.2, 3.1, 3.0, 2.8, 2.7, 5.4, 2.4, 2.6, 2.8, 2.9]

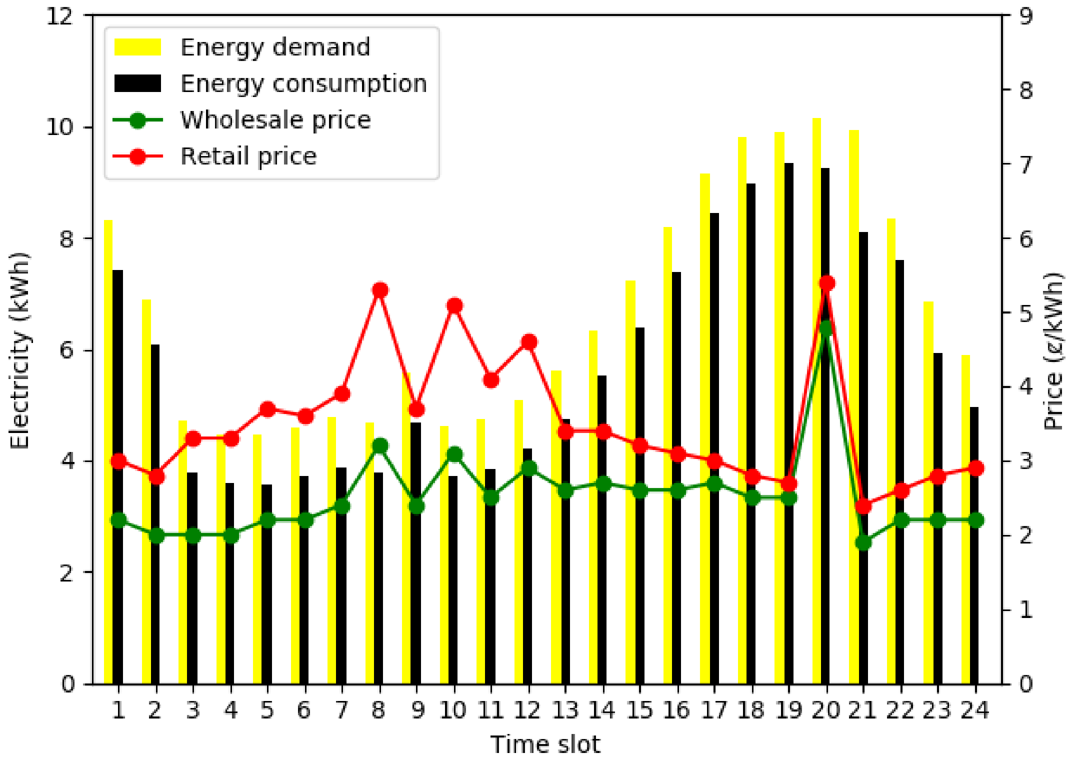


Fig6: Demand of Customer 2 with wholesale price on 2019/10/31

This simulation result is applying the data sets of demand of customer 2 and the wholesale price on the date of 2019/10/31. There were 4603 iterations executed by the algorithm with a reward of 69.083.

Log Example:

iteration 1,920; delta: 295.1449049616415; reward: -965.0873860873628...

[3.0, 2.8, 3.3, 3.3, 4.8, 3.8, 4.1, 6.9, 3.9, 6.9, 4.5, 7.0, 3.8, 3.8, 3.5, 3.3, 3.0, 2.8, 2.7, 2.4, 2.4, 2.6, 2.8, 2.4]

iteration 2,900; delta: 154.82114028733804; reward: -350.216862028958...

[3.0, 2.8, 3.3, 3.3, 3.7, 3.6, 3.9, 5.3, 3.7, 5.1, 4.1, 4.6, 3.4, 3.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.5, 2.8, 2.9]

iteration 2,920; delta: 119.55089335367234; reward: 22.675189843829568...

[3.0, 2.8, 3.3, 3.3, 3.7, 3.6, 3.9, 5.3, 3.7, 5.1, 4.1, 4.6, 3.4, 3.4, 3.2, 3.1, 3.0, 2.8, 2.7, 5.6, 2.4, 2.4, 2.4, 2.4]

finished at iteration 4,603, with a delta of 0.056826742523526264...

[3.0, 2.8, 3.3, 3.3, 3.7, 3.6, 3.9, 5.3, 3.7, 5.1, 4.1, 4.6, 3.4, 3.4, 3.2, 3.1, 3.0, 2.8, 2.7, 5.4, 2.4, 2.6, 2.8, 2.9]

5. Conclusions

Application of machine learning and big data analysis is a new trend of data analysis. The method could discover a more effective dynamic retail price for the retailer.

In this project, applying reinforcement learning in dynamic pricing of smart power girds is one of the usages of the above techniques. With the usage of the above techniques, people are able to eliminate the ineffectiveness caused by information asymmetry in the electricity retail market.

Information asymmetry is a classic economical issue in most of the product markets. The issue has been improved due to the popular use of internet. In the trend of machine learning and big data analysis, the issue would be improved more. The overall effectiveness of electricity market and other markets would be increased by the techniques of machine learning and big data analysis.

6. Future Directions

6.1 Automatic daily demand tracing with Python

Currently, our program requires datasets input manually or using csv files to input the data from different websites. This is a troublesome process and it increase the usage difficulty of the reinforcement learning algorithm. It would be more ideal if the program can extract datasets online automatically and update the parameters with a script or a program. With the libraries on the internet, it is not difficult to create a web crawler using python to extract the websites' data automatically. It is also possible to write a script without the libraries to extract these datasets from the websites. We can change the updated parameters in the Q-learning algorithm with the web crawler script. Python Pandas and Selenium could be used as the libraries to create the crawler.

6.2 Adjustments for higher adaptability

The program was only tested with the data sets from the above websites. It is desired to be adapted to work under data sets from other companies with adjustments on the parameters. If the program is able to adapt and process more data sets and keep learning with the experience and statistics, it would be more useful and has more variety usage in dynamic pricing of smart power grids from around the world.

6.3 Develop a software with user-interface

If the above two future directions have been achieved, a non-programmer user-friendly version of the program could be produced. It would be desired if this program could be distributed into different clients.

7. References

[1]: United Nations Sustainable Development Goals, https://www.un.org/sustainabledevelopment/sustainable-development-goals/

[2]: Ensure access to affordable, reliable, sustainable and modern energy, https://www.un.org/sustainabledevelopment/energy/

[3]: Goal 13: Take urgent action to combat climate change and its impacts, https://www.un.org/sustainabledevelopment/climate-change/

[4-9]: “A Dynamic pricing demand response algorithm for smart grid: Reinforcement learning approach”, Renzhi Lu, Seung Ho Hong , Xiongfeng Zhang, Department of Electronic Engineering, Hanyang University, Republic of Korea

[10]: “Q-Learning”, ”Machine Learning”, Christopher J.C.H.Watkins & Peter Dayan, May 1992, Kluwer Academic Publisher, 279-292

[11] Open Source Q-Learning Algorithm for dynamic pricing of smart grids : https://github.com/hekimgil/grid_pricing

[12]: Commonwealth Edison Company electricity price: https://hourlypricing.comed.com/live-prices/

[13]: Demand, supply, and the market mechanism ,Richard E. Hattwick, Joel W. Sailors, Bernard G. Brown, Englewood Cliffs, N.J. : Prentice Hall 1971

[14]: San Diego Gas & Electric https://www.sdge.com/

[15]: California ISO: http://www.caiso.com/TodaysOutlook/Pages/default.aspx