

# Machine Learning: An Overview

**Mohd Izhan Mohd Yusoff**

Telekom Malaysia Berhad, Kuala Lumpur, Malaysia

Email: mohdizhanmohdyusoff@gmail.com

**How to cite this paper:** Yusoff, M.I.M. (2024) Machine Learning: An Overview. *Open Journal of Modelling and Simulation*, 12, 89-99. <https://doi.org/10.4236/ojmsi.2024.123006>

**Received:** June 1, 2024

**Accepted:** June 23, 2024

**Published:** June 26, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Machine learning is an Artificial Intelligence (or AI) application, an idea that came into being by giving machines access to data and letting them learn by themselves. AI has been making headlines, especially since ChatGPT was introduced. Malaysia has taken many significant steps to embrace and integrate the technology into various sectors. These include encouraging large companies to build AI infrastructure, creating AI training opportunities (for example, the local media reported Microsoft and Google plan to invest USD 2.2 billion and USD 2 billion, respectively, in the said activities), and, as part of AI Talent Roadmap 2024-2030, establishing AI faculty in one of its public universities (*i.e.*, “Universiti Teknologi Malaysia”) leading the way in the integration and teaching of AI throughout the country. This article introduces several products developed by the author (for the energy and transportation industries) and recommends their improvement by incorporating Machine learning.

## Keywords

Machine Learning, Artificial Intelligence, SA<sup>2</sup>VING, Non-Intrusive Load Monitoring, Transportation Pricing System, System Dynamics, Dynamic Pricing

## 1. Introduction

Machine learning is a current application of AI based on the notion that we should just give machines access to data and let them learn for themselves. It is an AI application that enables systems to learn and improve from experience, without being explicitly programmed automatically. It focuses on developing computer programs to access and use data for self-discovery. The learning process begins with observations (such as direct experience or instruction) or looking for patterns in the data to make better decisions in the future, based on the given examples. The primary aim is to allow computers to learn automatically, without

human intervention or assistance, and adjust actions accordingly. Machine learning is an AI discipline and the key driver behind the advances of narrow AI in recent years. It is a collection of tools and methods that enable computers to learn from observations, data, and examples, to improve their performance. It does not require explicit step-by-step explanations on how to execute a task, as is necessary in traditional programming. In essence, machine learning is based on statistical models for making predictions. During the learning process, the parameters of the statistical model are optimally adapted according to the training data provided. The system learns by experience. It can then make predictions for unseen events/data, based on the given training data. For a model to generalize well for those unseen events, it is essential to define a model well suited to the problem statement and to have training data available in the necessary quantity and quality.

*Supervised Machine Learning* is a technique in which we teach or train the machine using well-labeled data. The Supervised Machine Learning models are trained with labeled data sets, which allow the models to learn and grow over time. It also means providing an algorithm with fully labeled training data. Labeled training data indicates that every training record also has the associated answer attached to it. Therefore, we have already provided clear guidance to the data interpretation algorithm. The algorithm will then construct and test a model against the given labels. For each iteration, an error function will validate whether the prediction the model came up with aligns with the label. The model will then be adjusted slightly with each iteration until it performs well on the training data. There are two types of supervised learning. The critical thing to remember is that classification separates the data, while regression fits the data. Classification is the algorithm that predicts the categorical response value for a given observation/input. It will predict a discrete value. A prerequisite is that we can segment the response values into distinct classes. A typical example is the classification of handwritten letters and numbers, as used by the post office to interpret an address written on a letter. Regression is an algorithm that predicts a numerical continuous response value for a given observation/input. It will predict a continuous value. The prediction of house prices based on their location, size, age, etc. is an example.

*Unsupervised Machine Learning* involves training through unlabeled data, allowing the model to act on that information without guidance. In Unsupervised Machine Learning, a program looks for patterns in unlabelled data. Where it is difficult or too expensive to obtain enough labeled training data, unsupervised methods become relevant. With unsupervised learning, we do not provide the algorithm with labeled training data. Instead, we want the algorithm to find a way to classify/separate the data. The following vital categories exist: clustering is an algorithm that tries to create clusters of similar data points. For example, one can imagine that a machine can classify vehicles independently when the algorithm receives only the front view of millions of cars. A BMW looks different

from a Ford or Audi; therefore, the algorithm should easily be able to cluster the pictures into different categories (in this case, car brands). Anomaly Detection is an Algorithm that tries to find outliers/anomalies in a given data set. The identification of fraudulent behavior in bank transactions is an example. If a person puts through many transactions of relatively small amounts while traveling between Europe and the US, a USD 10,000 transaction from Nigeria in between might evoke suspicion. Association rules are algorithms that identify relationships and dependencies within data sets. An online store recommender engine, where correlated products can be grouped based on similar features and then suggested to the consumer, is an example.

*Semi-Supervised Machine Learning:* When labeling massive amounts of training data is too expensive, a mix of supervised and supervised machine learning comes into play. The process starts with a person labeling a small amount of data (for example, CT scans, which are interpreted and labeled by an expert radiologist). The most famous example of semi-supervised learning is a special kind of neural network called GAN (General Adversarial Network). It uses labeled training data to generate new data (through the generator network). It sends this data to another neural network (discriminator), which needs to identify whether the data is fake or part of the training data. The generator and discriminator networks jointly improve in a positive feedback loop. The generator will improve at creating convincing counterfeit data, and the discriminator will be better at separating fake data from training data. A great use case for a GAN is the creation of artificial images that mimic authentic images. A GAN-generated portrait was recently sold at Christie's for a ridiculously high price.

*Reinforcement Learning* is a part of Machine Learning, where an agent is put in an environment and learns to behave in this environment by performing specific actions and observing the rewards it receives therefrom. Reinforcement learning trains machines through trial and error to take the best action by establishing a reward system. It assumes that positive rewards can enforce an optimal behavior or action for a favorable action. The basic setup for reinforcement learning comprises an agent that interacts with its environment. Based on the positive/negative feedback the agent receives from interactions in a given situation, a particular behavior is rewarded, and favorable actions are therefore reinforced. In many iterations, the agent will be trained, and its performance for achieving the overall goal will be improved. For each action, the agent must decide whether they want to explore their environment (exploration) to search for potentially higher rewards (the risk of failing is also higher), compared to the actions that have already proven successful (exploitation) [1] [2] [3] [4].

The products highlighted in the following section were designed and developed by the author (using Java language). Some of their names (plus actual features and figures) have not been revealed to ensure secrecy. Improvement through the employment of machine learning is highlighted in the Recommendation section.

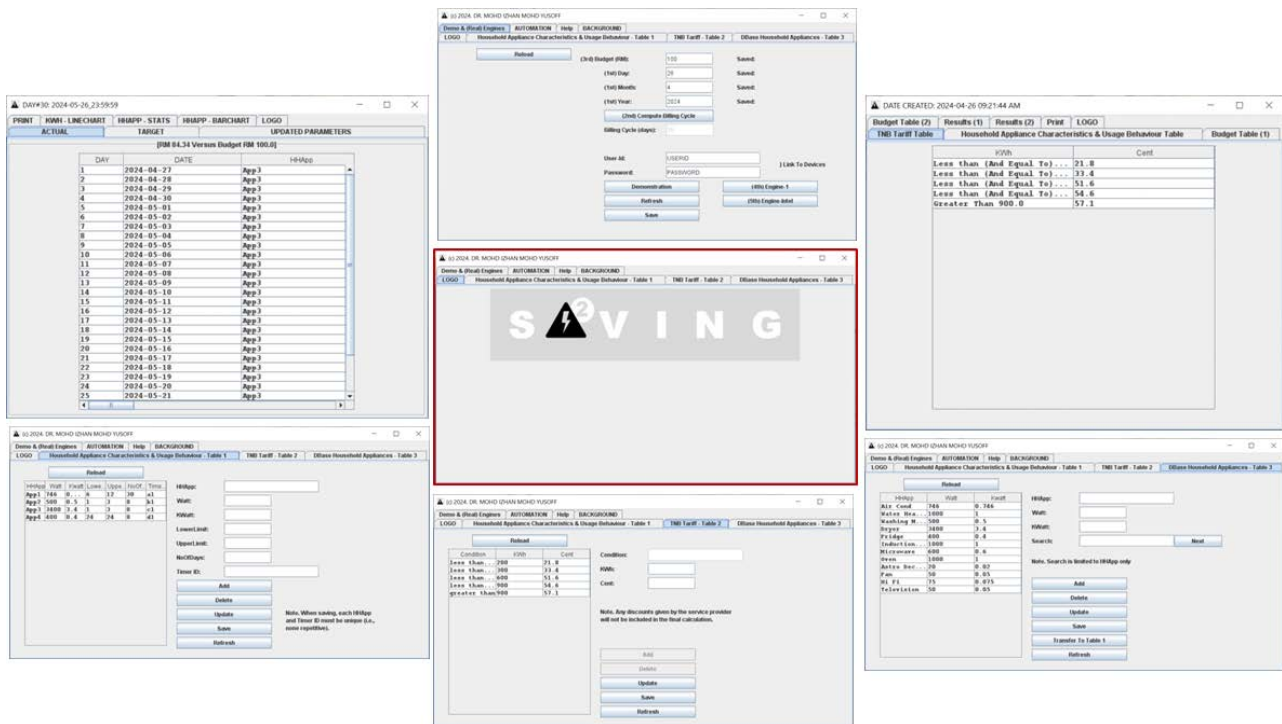
## 2. Products

### 2.1. SA<sup>2</sup>VING

Household Electrical Appliances Voltage Monitoring System (SA<sup>2</sup>VING in short—refer to **Figure 1**; apart from being developed using Java language, it is still in the development and testing stage), will inform the users which (listed) household electrical appliances to switch on and for how long, if their monthly electric budget is, say MYR 500. This could lead to significant cost savings, similar to customers who choose 100 Mbps internet speed, which costs MYR 135 monthly if their budget is below MYR 200.

**Figure 2** shows the execution (or setting up) of the system in a typical household, which involves using timer devices (or smart plugs). The timer devices are connected wirelessly to SA<sup>2</sup>VING installed on a desktop, laptop or mobile phone; they can offer information about household electrical appliances, especially the time and date they are switched on and off. Some of the timers (or smart plugs) are equipped with JSON (a text-based, human-readable data interchange format) and REST API (an application programming interface that conforms to the design principles of the representational state transfer architectural style).

In addition, the said system will also continuously monitor the usage of (listed) household electrical appliances and suggest the best way to meet the monthly budget. **Figure 3(a)** and **Figure 3(b)** show the system’s results produced through a simulation to test the process for continuous monitoring (an important component of SA<sup>2</sup>VING Intel, short form for Intelligence), as shown



**Figure 1.** SA<sup>2</sup>VING and its (non-exhaustive) components.

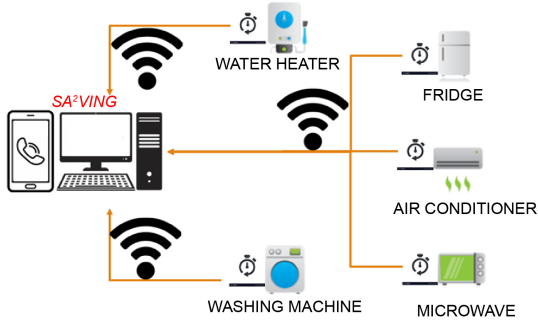
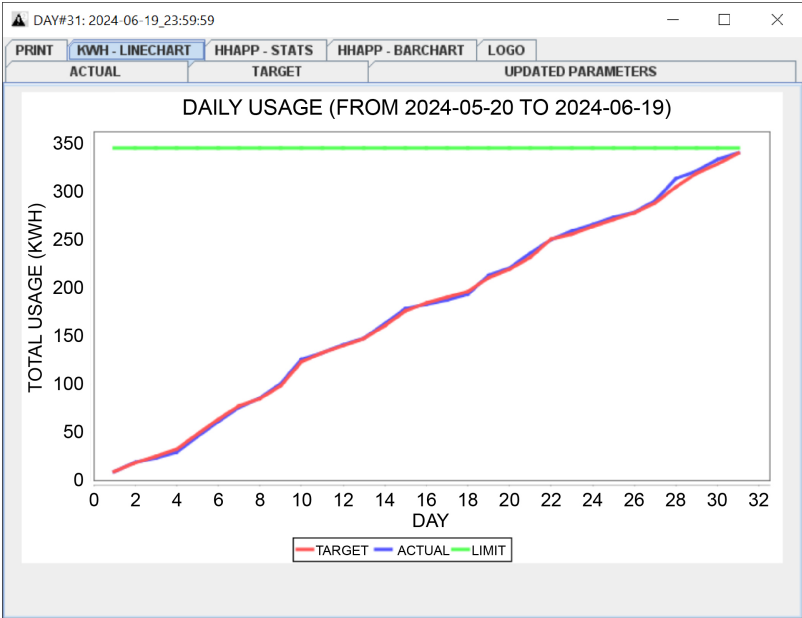


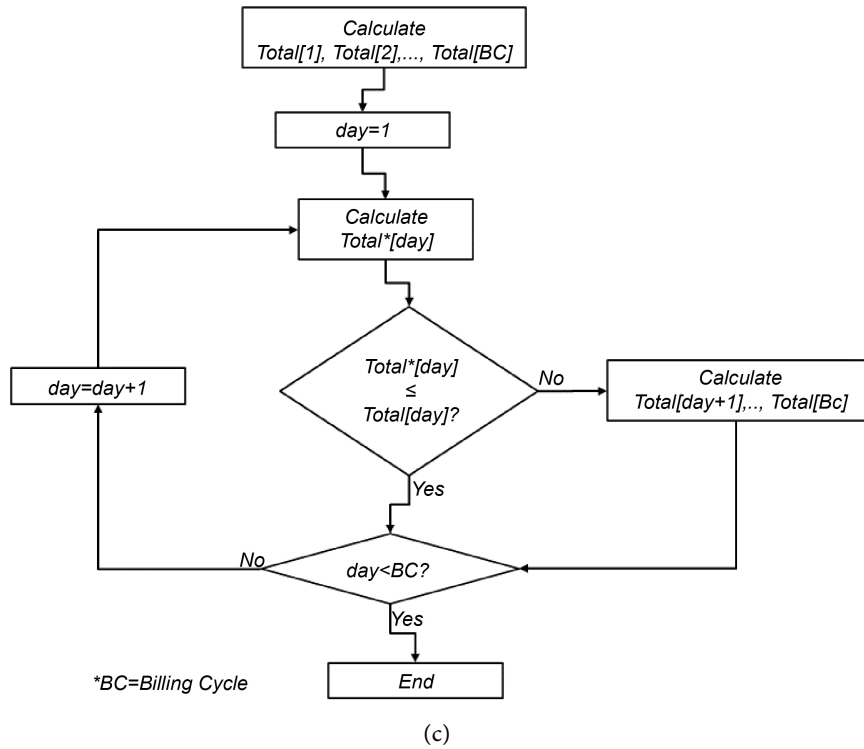
Figure 2. SA<sup>2</sup>VING and its set-up.



(a)



(b)



**Figure 3.** (a) The line chart shows the (cumulative) target (colored red) and actual (colored blue) total daily usage in KWH. The green line represents the limit for total daily usage. Notably, actual achieves the limit near the end of the billing cycle and (b) the bar chart shows the number of hours each household appliance is operational (or switched on). The bar is labeled “Alarm” (colored red) if the actual hours for a particular day (and household electrical appliance) exceed the target and “Normal” (blue) if it is the opposite (or vice versa). Note that 58 percent of displayed bars are labeled “Alarm” and (c) the flow diagram shows the process for continuous monitoring.

in **Figure 3(c)** flow diagram. The simulation was performed before the setting up stage and, apart from the random number generator, used information the user entered in SA<sup>2</sup>VING Table 1’s cells for a particular household appliance, namely the “lower limit,” “upper limit,” and “number, no. of days.” The “frequency” of using a household appliance is represented by a minimum (“lower limit”) and maximum (“upper limit”) number of hours per day and the number of days (“number, no. of days”) per month. For example, suppose the user entered air conditioning for 6 and 9 hours for the “lower limit” and “upper limit” cells, respectively, and eight days for the “number, no. of days” cell. In that case, the system will use a random number generator to choose (to create some realism, higher than, equal to, or lower than) eight days between 1 and the “billing cycle” and, for each selected day, to determine the number of hours between 6 and 9 hours. Notably, the selected number of hours may or may not comply with the continuous monitoring process’s suggestions, thus creating some realism. When inputting the Budget MYR, the user needs to ensure it falls within the range of the Lower Limit MYR and Upper Limit MYR, where the range is available in SA<sup>2</sup>VING Demo & (Real) Engines, and they (*i.e.*, Lower and Upper Limit

MYR are derived from the user's inputs and tariff in SA<sup>2</sup>VING Tables 1 and 2, respectively. The Upper Limit MYR should be greater than (or equal to) the maximum monthly electricity bill.

With the increase of electrical tariffs in some countries (e.g., Singapore [5]), SA<sup>2</sup>VING is relevant to government and private sector workers who choose to work from home (as well as senior citizens or retirees living on a stringent budget). Hospitality would also benefit from the said system (*i.e.*, by modifying timer devices to include control for switching on and off); for example, hotels could offer different room rates to customers, depending on their daily electrical usage preference. SA<sup>2</sup>VING could be used in any country with minor changes, especially in the electricity tariff, as displayed in SA<sup>2</sup>VING Table 2.

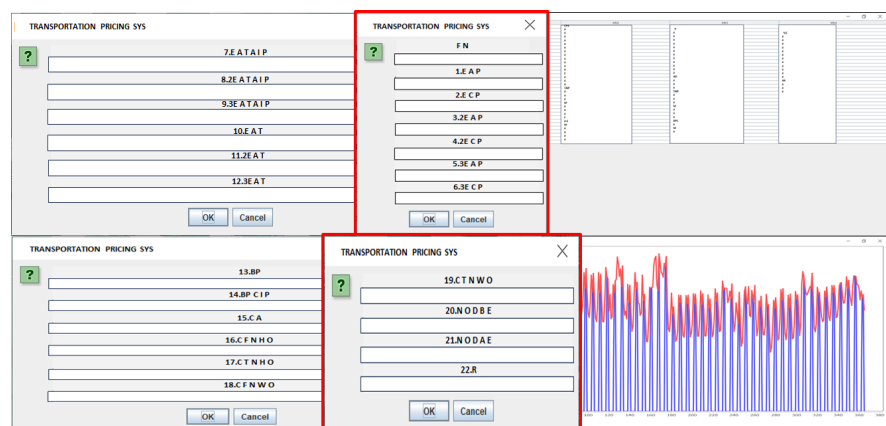
The development of SA<sup>2</sup>VING resulted from extensive research on, for example, Malaysia's electricity service provider, "Energy Usage Calculator" [6], a tool that calculates energy usage for a given household appliance.

## 2.2. Transportation Pricing System

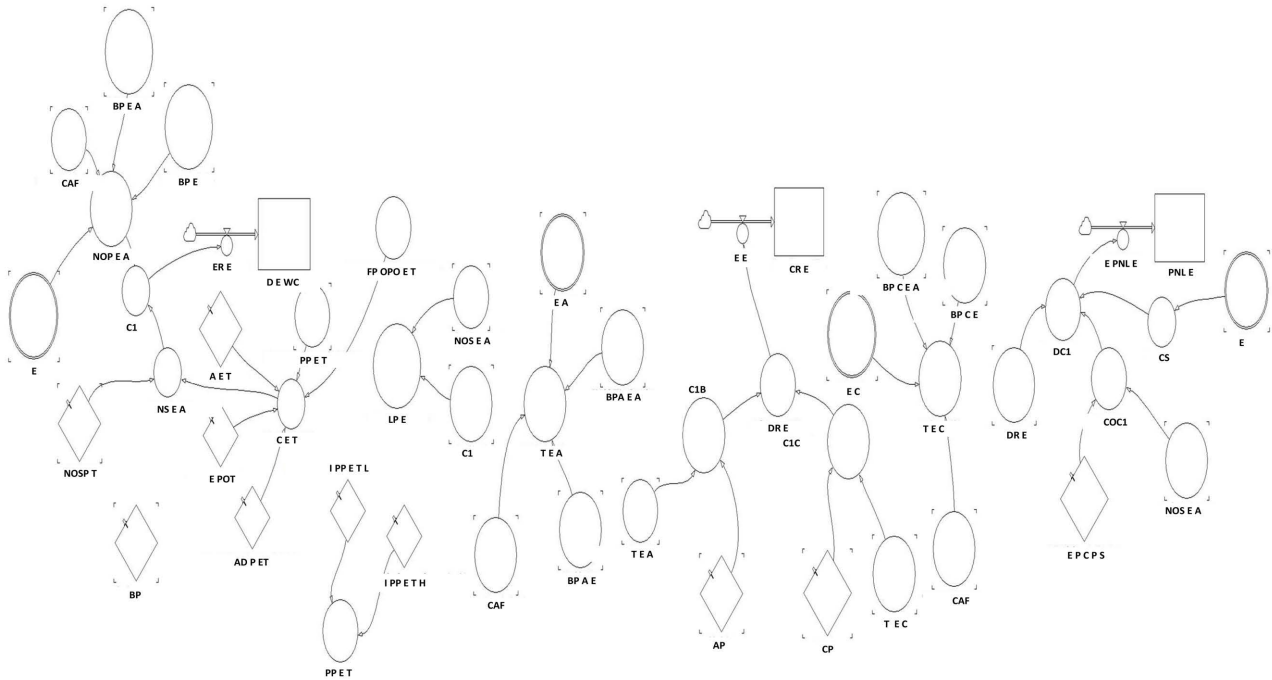
The "Transportation Pricing System" (name changed to ensure secrecy), as shown in **Figure 4**, was developed for a transportation company located in Malaysia using Java language based on the System Dynamics models, [7] [8] created using Powersim software (as exemplified in **Figure 5**), thus lowering the overall development cost.

The Transportation Pricing System is a strategic tool that enables the company to determine the most favorable pricing for its customers while safeguarding the company's financial health. To achieve this, System Dynamics models were integrated into the system. These models leverage the company's strategic and daily performance data from the previous year, which serve as the default or reference values in the Transportation Pricing System.

One of the critical features of the Transportation Pricing System is its user-friendliness. Users can explore several "what-if" scenarios by changing default values such as child and adult fare, number of vehicles, cost per seat, the total



**Figure 4.** Transportation pricing system and its components. The actual labels are not revealed to ensure confidentiality.

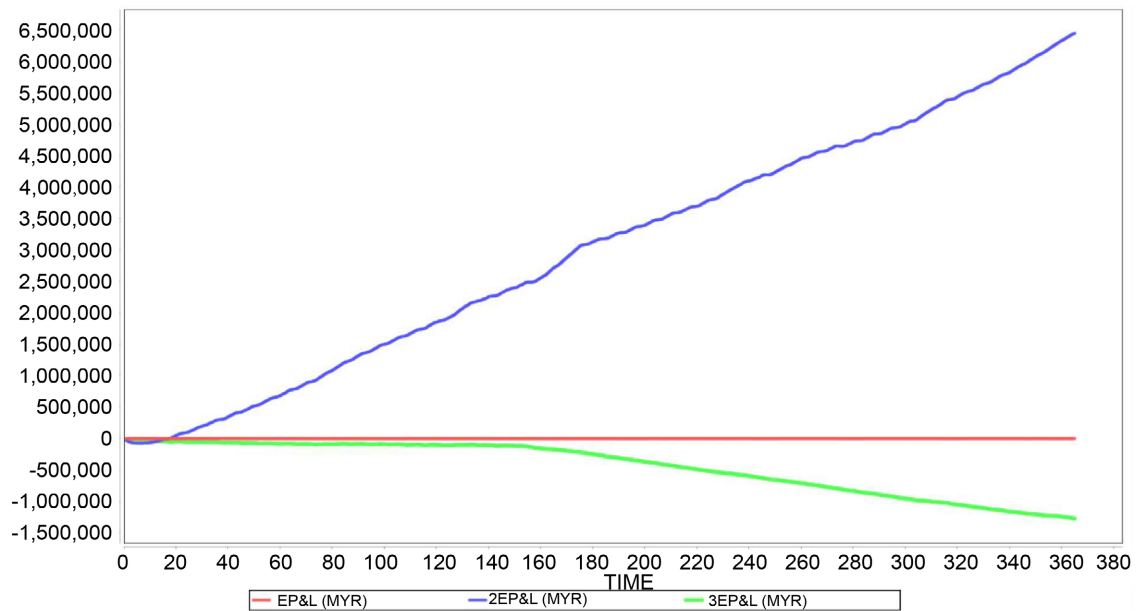


**Figure 5.** One of the system dynamics models is displayed. The actual labels are not revealed to ensure confidentiality.

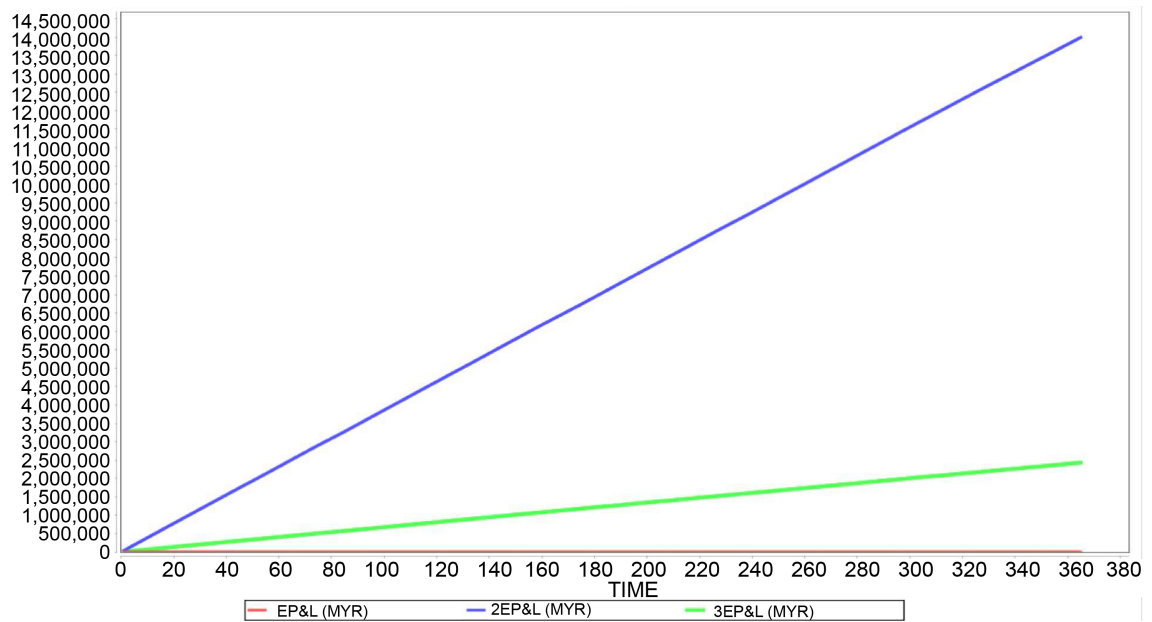
number of seats per vehicle and the performance of all vehicles in percentage throughout the year, on weekends or holidays. In other words, the users are allowed to alter or change the outcomes of the previous year's financial performance, without worrying about making mistakes or taking unnecessary (or uncalculated) risks. The system then produces comprehensive results in graphs, making it easy to understand and analyze the impact of different “what-if” scenarios, as exemplified in **Figure 6(a)** and **Figure 6(b)**. The bottom-up approach (*i.e.*, per route basis) is highly recommended. The user must perform many scenarios before achieving the desired outcome, and the system is equipped with an expert forecasting system capable of predicting the future demand for alternative transportation.

### 3. Recommendation

SA<sup>2</sup>VING can be further improved by including Machine Learning's Non-Intrusive Load Monitoring (NILM) [9]. NILM uses Factorial Hidden Markov Models [10] (an extension of Hidden Markov Models [11]) and Expectation Maximization [12] (to find the Factorial Hidden Markov Models' final parameters), can determine which appliances are switched on (and their energy consumption) by analyzing the building's electricity usage “behavior”. The NILM's features replace the requirement for user input of household appliances in SA<sup>2</sup>VING Table 1. The weakness of NILM is its inability to deliver cost savings and applications at scale because it has to work with systems that collect and process data frequently (preferably every second or minute). Most systems in the market today collect and process data every five minutes, half an hour, one hour, or one day [13]. The



(a)



(b)

**Figure 6.** (a) Profit and Loss (MYR) graph for E, 2E, and 3E, using route X-Y when Business-As-Usual (BAU) is displayed. Note that E is not offered to customers, 3E is making losses, and (b) Profit and Loss (MYR) graph for E, 2E, and 3E using route X-Y when ten (10) percent of customers using alternative transportation are “migrated”. Note that E is not offered to customers, and 3E makes a profit.

weakness will also impact the development of the training data, which should consist of all household appliances (as well as their brands) and, upon completion, replace SA<sup>2</sup>VING Table 3.

The Transportation Pricing System could be improved by taking (strategic) advantages of Dynamic Pricing and Machine Learning features. Dynamic Pricing (surge, demand, or time-based pricing) is a revenue management strategy

that allows businesses to set flexible prices for products or services based on current market demand [14]. It applies to firms with competing “networks” and a saturation process. Each firm wants its customers to use its “network” rather than its competitors. Due to the positive network internality, the greater the share of a firm’s consumers, the easier it is to attract additional consumers. Each firm can influence the rate at which it attracts consumers by altering its price and level of advertising. However, as the market saturates, a firm with a small market share will find it difficult to attract additional consumers, even if it lowers its price substantially [15]. Dynamic airline pricing accounting for both forces (*i.e.*, intertemporal price discrimination and dynamic adjustment to stochastic demand) with new flight-level data. By disentangling critical interactions between the arrival pattern of consumer types and remaining capacity under stochastic demand, the forces complement airline markets and lead to significantly higher revenues and increased consumer surplus, compared to a more restrictive pricing regime [16]. Dynamic pricing is a powerful alternative to the segmented pricing and A/B testing approach many developers currently use. It optimizes every user’s price in real time, without manually defining or testing complex pricing rules. Instead, the Dynamic pricing tool uses Machine Learning algorithms to make pricing decisions. The Machine Learning model is a regression model that estimates the impact on revenue for each possible price configuration. These estimates are significant because they tell the Dynamic Pricing algorithm which price configuration to display [17]. While Dynamic Pricing offers significant advantages concerning revenue optimization and customer engagement, it also presents challenges. These include potential favoritism toward certain buyers, concerns about privacy invasions and the risk of data fraud and misuse, particularly when businesses utilize Big Data and Internet of Things (IoT) technologies [18].

### Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

### References

- [1] Simonini, T. and Sanseviero, O. (2022) An Introduction to Deep Reinforcement Learning. <https://huggingface.co/blog/deep-rl-intro>
- [2] Lateef, Z. (2019) Introduction to Machine Learning: All You Need To Know About Machine Learning. <https://www.edureka.co/blog/introduction-to-machine-learning/>
- [3] Brown, S. (2021) Machine Learning Explained. <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>
- [4] Marr, B. (2016) What Is the Difference between Artificial Intelligence and Machine Learning? <https://bernardmarr.com/what-is-the-difference-between-artificial-intelligence-and-machine-learning/>

- 
- [5] Malay Mail (2023) Singapore Household Electricity Tariffs to Increase 4.1 pc in 1Q 2024, gas by 3.3 pc. <https://www.malaymail.com/news/singapore/2023/12/29/singapore-household-electricity-tariffs-to-increase-41pc-in-1q-2024-gas-by-33pc/109736>
- [6] Tenaga Nasional Berhad. <https://www.mytnb.com.my/energy-efficiency/home-energy-savings-tips/home-energy-calculator>
- [7] Forrester, J.W. (2009) Some Basic Concepts in System Dynamics. Sloan School of Management, Massachusetts Institute of Technology.
- [8] Yusoff, M.I.M. (2018) The Behavior of the Procurement Process as Described by System Dynamics. *Advances in Operations Research Journal*, **2018**, Article ID: 6721912. <https://doi.org/10.1155/2018/6721912>
- [9] Norford, L.K., and Leeb, S.B. (1996) Non-Intrusive Electrical Load Monitoring in Commercial Buildings Based on Steady-State and Transient Load-Detection Algorithms. *Energy and Buildings*, **24**, 51-64. [https://doi.org/10.1016/0378-7788\(95\)00958-2](https://doi.org/10.1016/0378-7788(95)00958-2)
- [10] Ghahramani, Z. and Jordan, M.I. (1997) Factorial Hidden Markov Models. *Machine Learning*, **27**, 245-273. <https://doi.org/10.1023/A:1007425814087>
- [11] Rabiner, L.R. (1989) A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, **77**, 257-286. <https://doi.org/10.1109/5.18626>
- [12] Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977) Maximum Likelihood from Incomplete Data via EM Algorithm. *Journal Royal Statistics Society*, **39**, 1-38. <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x>
- [13] Antonio (2019) Why Wattics Removed Non-Intrusive Load Monitoring from Energy Analytics. <https://54.194.200.155/why-wattics-removed-non-intrusive-load-monitoring-from-energy-analytics/>
- [14] Surowiecki, J. (2014) In Praise of Efficient Price Gouging. MIT Technology Review. <https://www.technologyreview.com/2014/08/19/74207/in-praise-of-efficient-price-gouging/>
- [15] Pindyck, R.S. (2012) Lecture Notes on Pricing. Sloan School of Management, Massachusetts Institute of Technology.
- [16] Williams, K.R. (2017) Dynamic Airline Pricing and Seat Availability. Cowles Foundation for Research in Economics, Yale University.
- [17] Boshell, B. (2017) How Dynamic Pricing Uses Machine Learning to Increase Revenue. <https://www.linkedin.com/pulse/how-dynamic-pricing-uses-machine-learning-increase-revenue-boshell#:~:text=Dynamic%20Pricing%20and%20Machine%20Learning&text=It%20automatically%20optimizes%20prices%20for,to%20drive%20its%20pricing%20decisions>
- [18] Taylor, C.R. (2004) Consumer Privacy and the Market for Customer Information. *The RAND Journal of Economics*, **35**, 631-650. <https://www.jstor.org/stable/1593765>