Summary of Lecture of Retail & Marketing Analytics on 26/02/2021

Qian Zhang

During the last terms, we learnt various technique of machine learning, including linear regression, KNN models, random forest, and neural forest, etc. In this course, the focus will be applying those machine learning techniques to the retail and marketing problems in the reality. By exploring the dataset collected from real business situation, we want to examine the relationship between all related variables and making good predictions and forecast the demand for better retail and marketing strategy. To better practice marketing analytics and create more value for the brand manager, we need to train not only the ability to examine the data using mathematical models with the help of modern computational technology, but to draw insights from the results and practice actions.

The focus of the first lecture is demand forecasting and various elementary approaches to forecast the demand. For example, simple moving average, which calculates the average of last k observations as a forecast for the next period, is a naïve forecast. However, simple moving average fails to consider many potential factors that might impact the forecasting of demand. For instance, the importance of the observations goes down over the course of the data, period. Therefore, simply calculating the average of the demand is not an effective method to forecast the demand. We need to implement some modification to this algorithm, like giving less weight to the distant observations in the dataset.

However, even we take the weight of observations in different period into consideration, they are still many other factors that will significantly impact the accuracy of demand forecast. For example, the demand of a particular commodity might follow seasonality pattern, such as electric fans, whose demand is largest during summer but lower during winter. Therefore, more sophisticated methods for forecasting are required to deal with real world situation.

Multiple linear regression is a powerful method that could be used to forecast demand. To formulate a linear regression model that could be applied in the real world, there are some basic factors that should be considered. First, we should take trend, seasonality, and cycles into consideration. Moreover, competition and holiday effect will also significantly impact the demand. Also, many marketing factors, such as advertising, distribution, and promotions, are also important. Finally, since the real world is too complex to be predict preciously, the model should contain the error parameter, which is used to evaluate all those factors that could be included in the model.

Since the relationship between the sales response and all the marketing variables is highly complex, and it is not possible to assume that the errors are independent and normally distributed, the multiple linear regression model might not be very effective to find the relationships between all marketing factors and the forecasted demand. Among the other approaches that could be used to do retail sales forecast in real world, neural network is one of the commonly used machines learning modelling tool, since there is no need to impose any statistical assumptions in a neural network modelling approach. Moreover, we could verify the accuracy of the neural network model by calculating the difference between target value and the forecast demand.