Proceedings of the 8th World Congress on Electrical Engineering and Computer Systems and Sciences (EECSS'22)
Prague, Czech Republic - July 28- 30, 2022
Paper No. CIST 128
DOI: 10.11159/cist22.128

Optimizing Production Decisions with Lead Time Dependent Demand Using Reinforcement Learning

Chi-Yang Tsai[1], Michelle Melsha Sugiarto[1]
[1]Department of Industrial Engineering and Management, Yuan Ze University
135 Yuan-Tung Road, Chung-Li, Taoyuan 32003, Taiwan
iecytsai@saturn.yzu.edu.tw; s1065453@mail.yzu.edu.tw

Extended Abstract

There are more and more sales channels available for consumers to choose. Delivery lead time has become an important factor when consumers choose what and where to purchase. Abundant amount of supply may help to maintain satisfactory delivery lead time, but it increases the risk of overproduction and overstocking. On the other hand, longer lead time due to insufficient amount of supply discourages customers and future demand may decline as a consequence. Facing customer demands that are sensitive to delivery lead time is a great challenge for companies to production decisions that generate maximal profit with limited capacity.

This research explores the optimal production planning problem with lead time dependent demand. Demand from customers exhibits uncertainty. Unsold units at the end of a period are carried over to the next period. Unmet demand is fully backordered. The manufacturing company receives demand at the beginning of a period and decides the production strategy of that period. Subsequently, the amount of supply in this period is determined. How the supply meets the demand further determines the delivery lead time in this period and consequently affects next period's demand. The production decision in a period not only affects the profit of this period but also influences future demand. To respond to the changes of demand in each period, it is necessary for the company to monitor the consequent profit and demand of the next period after the decision on the strategy to take in this period is made. From what is observed, the company makes adjustment on the strategy for the next period in order to pursue higher long-term profits. For solving this dynamic sequential decision-making problem, this research applies reinforcement learning techniques to identify the optimal production planning and pricing strategy. Application of machine learning on complex problems has been receiving more and more attention in recent years. Reinforcement learning is one of machine learning approaches. It is capable of learning from past experiment and solving complex sequential decision-making problems.

Two types of value-based reinforcement learning techniques, SARSA (State-Action-Reward-State-Action) and Q Learning, are applied in this research. Simulation experiments are conducted to evaluate the best sets of hyperparameter values, such as learning rate, discount factor and greedy factor. The performance of the two techniques are next compared. Generally, SARSA and Q Learning achieve similar average profits. SARSA tends to converge more quickly. To further access the ability of these two techniques to make high quality production decisions, two base-stock-type inventory control policies are applied and evaluated. The control mechanism of a base-stock policy is established based on average demand and is commonly used in practice. It is clearly shown in the conducted experiments that the reinforcement learning techniques outperform the two base-stock inventory control policies, especially when backorder cost is high, production cost is low, and inventory carrying cost is high. It is also observed that reinforcement learning performs particularly well when average demand is far lower than production capacity and when the degree of demand variability is high.