Reinforcement Learning in Retail Operation

Reinforcement Learning has ability to train systems to respond to unforeseen environments, is being increasingly adopted in product management to improve forecast accuracy, solve supply chain optimization challenges, and train systems to respond to unforeseen circumstances.

more about RL...

----> *Reinforcement learning is emerging as the dominant technology. The key reason for it's prominence is its ability to autonomously handle massive amounts of data, countless variables, and incredible levels of complexity without requiring much human labour. The technology keeps humans in the loop, but only to guide the technology to follow the strategies of the company and to evolve the technology itself further.*

Possible use cases:

- Cognitive Demand Forecasting
- Product End-of-Life
- Forecasting
- Demand Integrated Product Flow

PROBLEM STATEMENT

- A retail firm has many products in their inventory, and very few of them tend to sell and many of the products only have a single sale in the course of a year.
- So a common question is probable that how much inventory should I carry? In one hand, inventory means working capital costs, operational costs and a complex operation. On the other hand lack of inventory leads to lost sales, unhappy customers and a damaged brand.

So, The sales and growth team of the retail firm wants to determine which products from their inventory should they retain to sell and the ones to discard.

So this leads to two major issue :-

what are they?

1. Product management : Retail operations face the challenge of managing their Product effectively to ensure they have the right products in stock, without overstocking.
2. Price management : Retail operations face the challenge of determining an effective price strategy that balances profitability and customer satisfaction.

How Reinforcement learning helps

Early work in these areas looked at classical algorithms to improve on challenges such as network flow and graphs. But the recent disruptions have made it critical for supply chains to have the resiliency to handle unexpected events. The biggest challenge lies in matching supply with demand given the uncertain environment.

Earlier attempts at building AI models for forecasting involved using time series modeling by using a combination of statistical and ML models to forecast sales based on the sales history and trends. But there were inherent limitations wherein such models could forecast numeric data but would be unable to determine policy dynamics.

This is where our Reinforcement Learning comes in.

RL algorithms has :-

- Ability to optimize strategies and handle unexpected scenarios
- Ability to cater to a diverse set of use cases
- Ability to work with both structured and unstructured data

Solution

So, If we see above two problems. (i.e. Product management & Price management)

- product analysis :- current demand high/low for particular product , particular product stocks , any fault in product broken.
- product management strategy :- real time monitoring of product's stock so that we can re-order accordingly for ensuring no product outage and their smooth availability.
- order stock :- fast and reliable system for reordering products , supplier relation should be good enough for negotiation.
- cost analysis :- cost of product and its overhead expenses to determine the minimum price that can be offered to customer , price range that is appropriate for the product
- Discounts and offer :- We can also add up some coupan or discount to sell the overhead expenses.

How we are going to implement above solution?

We may implement Q learning algorithm which describes as :

Q-Learning is a basic form of Reinforcement Learning which uses ***Q-values*** to iteratively improve the behavior of the learning agent.

Q-values are defined for states and actions. **Q(S,A)**

$$Q(S,A) \leftarrow Q(S,A) + \alpha\left(R + \gamma Q(S',A') - Q(S,A)\right)$$

(i) we train an agent to make decision that when to order products, in

what quantity to order, and which products should be given priority.

(ii) we train the agent on 'sales history data' about products in real time that will give us the idea about current demand.

(iii) we train the agent on 'customer feedback data' , which eventually will tell us which items to priortize. we also get the information of customer 'preferences' and 'expectations' which we can use for recommendation.

(iv) Market data we can see the market trends to set our product price little low but still profiltable to attract customer.

Relevant articles :

i. https://arxiv.org/pdf/2104.14398.pdf

ii. https://towardsdatascience.com/reinforcement-learning-with-openai-d445c2c687d2

iii. https://www.daisyintelligence.com/blog/reinforcement-learning-right-technology-retail

iv. https://towardsdatascience.com/data-science-case-study-optimizing-product-placement-in-retail-part-1-2e8b27e16e8d

v. https://medium.com/towards-data-science/data-science-case-study-optimizing-product-placement-in-retail-part-2-5e4dc855a535