Applied Machine Learning & System Simulators

Abhimanyu Mittal, Shivam Chaurasia, Nilesh Patil

SCAN ME

Product generation for Dream Sports users

- Demand forecasting and policy optimisation in a pricing context to quantify the price elasticity of Dream Sports users.
- Measure long term impact of products on user retention and transactions to enhance user experience and customer lifetime value.
- Optimal product catalogue for improved product-mix and user experience.

Demand Prediction for Revenue Maximization

Determine production quantities in a single step with no immediate reward to maximize total revenue, factoring in defined constraints.

State Space : The state ***s*** represents the current state of the system, including the current inventory levels, historical demand data, and external factors influencing demand.

Action Space : The action ***a*** corresponds to the production quantities of each product. The agent selects ***a*** from the action space ***A*** to maximize total revenue.

Reward Function : The reward ***r*** is the total revenue obtained from selling the products, calculated based on the chosen production quantities and product prices.

$$r(s,a) = \sum_i P_i \times a_i$$

Transition Dynamics : The transition function ***T(s,a,s')*** models the transition from the current state ***s*** to a new state ***s'*** based on the selected actions ***a***. This includes updating inventory levels and incorporating the effect of external factors on demand.

Objective : The goal is to find the optimal policy ***π*** that maps states to actions in a way that maximizes the expected total revenue within the given time window:

$$\pi^* = \text{argmax}_\pi \mathbb{E}\left[\sum_{t=0}^{T} r_t\right]$$

where ***T*** is the specified time window.

Constraints :

- *Irreversibility Constraint* : The agent cannot undo production decisions.
- *Shelf Life and Unsold Product Constraint :* Ensure that the products are sold within the fixed time window to avoid direct revenue loss.
- *Interdependence Constraint* : Model the interdependence of product demands in the state representation and consider it in the decision-making process.
- *Opportunity Loss Constraint* : Minimize opportunity loss by adjusting production quantities to meet or exceed demand.
- *Bulk Production Cost Constraint* : Encourage bulk production by considering the manufacturing cost in the reward function.

Research Objectives

1. Develop a state representation that captures the relevant information for demand prediction and production decision-making.
2. Design a reward function that considers total revenue, opportunity loss, and manufacturing cost.
3. Explore reinforcement learning algorithms, such as Q-learning or Deep Q Networks (DQN), to learn an optimal policy within the specified time window.
4. Train and evaluate the model using historical data and simulate production decisions to validate its effectiveness in maximizing total revenue while adhering to the specified constraints within the given time frame.

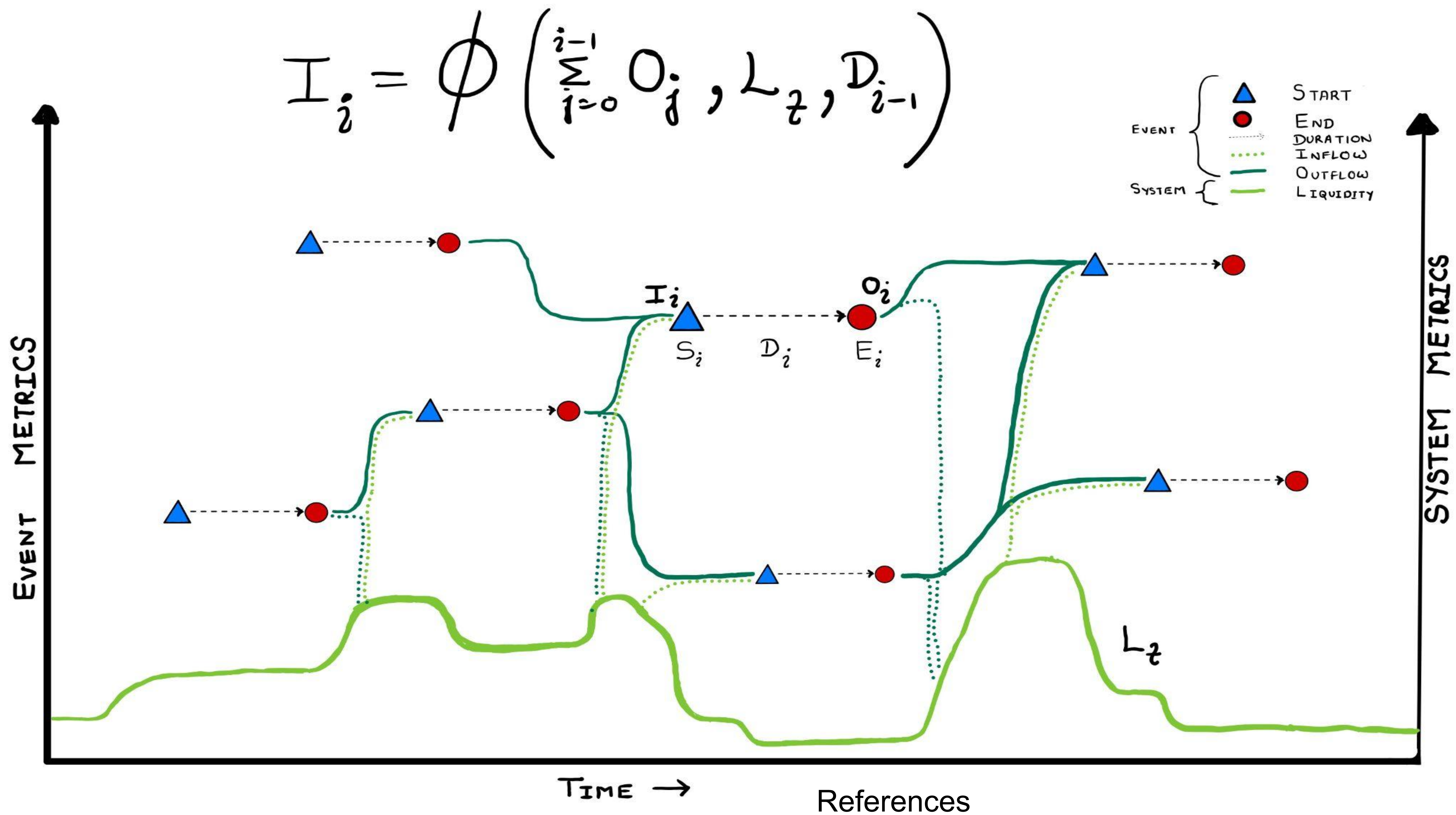
Flow Modeling & System Simulations

The inability to model multi-scale phenomena and their cause-effect mechanisms bottlenecks many promising concepts from ever being prototyped let alone deployed.

We aim to build simulated playgrounds for human experts and AI agents to design and experiment with state-of-art efficiency and fidelity, simulating counterfactuals and understanding downstream effects before deployment.

Potential avenues :

- Mechanisms for automated causal reasoning
- Multi-scale modeling
- Integration of simulators and online data streams
- Continuous learning & action algorithms



References

1. Lavin, Alexander, et al. "Simulation intelligence: Towards a new generation of scientific methods." arXiv preprint arXiv:2112.03235 (2021)

2. Zheng, Stephan, et al. "The AI economist: Optimal economic policy design via two-level deep reinforcement learning." arXiv preprint arXiv:2108.02755 (2021).