At the Chair of Logistics and Supply Chain Management of TUM School of Management, we are looking for an interested and qualified student to conduct his/her Interdisciplinary Project on the topic: 

Deep Reinforcement Learning for the Stochastic Joint Replenishment Problem under Non-Stationary Demand and Partial Observability.

The Joint Replenishment Problem (JRP) is a classical multi-product inventory management problem. In particular, the task involves simultaneously determining the order quantities for all products to minimize costs. The costs include inventory holding costs, penalty costs for shortages, order costs for each product, and common order costs stemming from, e.g., transportation resources. The common order costs make the problem difficult, as they create dependencies between the optimal order quantities and the inventory levels of the other products. The JRP with deterministic demand is known to be NP-hard, and the structure of the optimal policy is unknown.

In the stochastic version, it is commonly assumed that the underlying demand distribution is known and stationary (i.e., constant over time). However, this is not necessarily the case in practice. In general, the demand process may be non-stationary and unknown. For example, during COVID-19, the demand distributions of most companies likely evolved throughout the pandemic phases (e.g., pre-crisis, crisis, post-crisis). A common way to model this behavior is to have a set of demand distributions, one active at a time. Under partial observability, the currently active distribution is unknown and can only be partially observed through the resulting demand realizations.

Deep Reinforcement Learning (DRL) has proved successful in a variety of sequential decision-making tasks, including inventory management. Furthermore, DRL can, in principle, deal with partial information, as has been demonstrated in various domains such as robotics and games (e.g., Atari, StarCraft II, Dota 2). Thus, this project aims to develop a DRL approach for the stochastic JRP with non-stationary demand and partial observability. Furthermore, the approach should be benchmarked against heuristics that work with complete demand information or estimated distributions.

Key project tasks:

- Implementation and testing of the JRP environment
- Implementation and testing of benchmark heuristics
- Development of a suitable DRL method for the problem
- Systematic performance comparison of DRL and heuristics
- Analysis of results and derivation of implications
Requirements:
The project is suitable for master’s students in Informatics who require an IDP for their studies. The ability to work independently, as well as analytical skills, are required. Profound programming skills in Python and good knowledge of (deep) reinforcement learning are required. Knowledge of inventory management is beneficial, but not required. The courses of our chair relevant to this IDP include “Inventory Management” (JRP) and “Computational Logistics” (Python).

Earliest begin:  February 2024
Supervisor:     Patrick Helm
Application:  Email with curriculum vitae and transcript of records to logtheses.log@mgt.tum.de