Liberal Journal of Management & Social Science
https://liberaljournalofms.com/index.php/Journal
Vol. 5 No. 1 (2026)

AI-DRIVEN DYNAMIC SHELF-SPACE OPTIMIZATION IN RETAIL: A DEEP REINFORCEMENT LEARNING APPROACH USING GROCERY CHAIN DATA

Sajjad Ali
PhD Scholar, Department of Management Sciences, University of Haripur
Dawood Abbas
BS Student, Department of Management Sciences, University of Mansehra

ABSTRACT
We propose a novel approach using Deep Reinforcement Learning (DRL) with data from a major grocery store to adjust shelf spaces on-the-fly. Inspired by prior work on retail analytics and machine learning (Ahumada & Villalobos, 2009; Ferreira et al., 2015), we model shelf allocation as a multi-stage decision problem under uncertainty. Retail shelf-space management: Retail shelf-space assignment is a fundamental problem in operations and supply chain management, which can have substantial influence on sales, customer satisfaction and profit (Chen et al., 2015). Conventional optimization algorithms are insufficiently adapted to the volatile nature of consumer tastes and conflicting trade-offs prevalent in GM, with erratic demand and limited shelf-space. In this work, by making the trade-off among targeted sales volume, stockout rate and inventory turnover, a DQN as well as PPO algorithm are used. We observed that DRL-based mechanisms can potentially improve (up to 18% in case of revenue-optimization and up to 12% in case of stockout-mitigation) over heuristic/lp based policies. These findings highlight the potential of DRL in on-line, adaptive retail decision making (Mnih et al., 2015). The research contributes to the literature by demonstrating the potential use of DRL for making shelf space management a high velocity act, with practical significance for retail planners keen on gaining competitive advantage in fast moving markets.
Keywords: Retail shelf-space, Deep reinforcement learning, Grocery dataset, Inventory optimization, Dynamic allocation, Machine learning, Operations management

INTRODUCTION
Visibility and placement of products have substantial implications on demand, such as the position of product shelves in the three-dimensional setting store layout. However, out- of-sight is typically out-of-mind (Dr è ze et al. 1994; Campo et al. A thousand stock-keeping unit (SKU) supermarket, for example, makes the already difficult problem of not-enough-product-on-the-shelf even worse since all these different items need to be taken into account. Also the perishibility, cyclicality of promotion etc., these are some of factors which are difficult to account for in this kind of decision making mechanism." (Desmet & Renaudin 1998). In these intense competitors, the preferences of shelves are altered re ally very frequent and as well Because customers change their attitude quite rapidly so does the shelf management practice for present era (Singh, Bansal & Gupta 2018). The SSAP has traditionally been solved by deterministic optimization models. Linear and non-linear programming, as well as mixed-integer programming, heuristic algorithms other than these methods have also been used in solving the contradiction between yield and space constraints (Lim et al. 2004; Hwang et al. 2005). Although those approaches are very pronounced out of the solution, on the other hand they mostly rely on a static demand function instead of operating according to such a function that is adjusted for changes in reality (Yang & Chen 2011). Solutions always go on for too long, and when one perfects it, he can hardly retain it. For instance the sleep seasonality may be overlooked and constant throughout in the whole cycle while a new product range or promotional activities by an

LIBERAL JOURNAL OF MANAGEMENT & SOCIAL SCIENCE LJMSS

opponent might suppress demand over the cycle at any time (HPC 2006.
The emergence of big data analytics has opened up new roads for retail decisions making. Point-of sale data, loyalty card records and digital receipts all produce raw material for us to continue our work in honing our models of consumer demand to higher-order effectiveness (Fader & Hardie, 2009). To a new level of refinement, predictive analytics have been harnessed for consumer demand forecasts (Ferreira, Lee, & Simchi-Levi, 2015), price optimization performance (Chen, Mislove, & Wilson, 2016) and personalization in recommendation systems (Jannach et al., 2016).
Nevertheless, the majority of machine learning are indeed descriptive (Fader & Hardie, 2009; Lu et al., 2006) and for prediction in long term demand forecasting rather than prescribing what conforms to good allocation strategies. This leaves a gap between predictive analytics-based insights and real-time decision making support - in operational situations where decisions depend on demand responsiveness as well as inventory control (shelf management, for example).
In this light, RL is an optimal paradigm for dynamic retail decision making. RL is particularly well suited to problems that involve sequential optimization under uncertainty, as it provides a mechanism for agents to learn policies by interacting with the environment via trial and error (Sutton & Barto, 2018). As opposed to traditional search algorithms, RL is personalized per dialogue system in which the system?s policy as well as conversation environment varies dynamically and its goal is the maximization of longterm cumulative reward. These properties render it relevant for retailing environments characterized by demand variability, product perishability and an interdependent allocation decision.
RL has indeed been found to have early-stage insight value for related areas of retailing. It has been examined in scenarios such as dynamic pricing (Choi, Wallace & Wang,, 2018), personalized promotions (Wang, Zhang & Zhao, 2021) and inventory control \cite {Gijsbrechts et al., 2019}. These findings highlight the potential for RL to surpass static or heuristic methods, at least in settings involving large state and action sets. However, despite its potential, the research of RL towards shelf-space allocation problem is still scarce: thus far only few preliminary studies have adopted this approach (Wang et al., 2021).
Deep reinforcement learning (DRL) is a novel combination of deep neural network and RL method, used to approximate value function or policy. It is an opportunity to alleviate the high dimensionality of SSAP DRL in reality. Many of today 's applications take place Anywhere -and ever-increasing Worlds behind yet far beyond all this lunch Select a cottage or two or more. The best-known may be refrigerators. In shelf management, an area that involves managing for thousands of SKUs, DRL has the potential to generate guidelines on assigning the space in a way that is flexible and profitable.
Against this backdrop, the research question for our study reads as follows: For dynamic shelf-space allocation in grocery retailing is deep reinforcement learning superior to its traditional counter part by optimization? For this purpose, the analysis seeks to achieve three goals: 1. Design a DRL model for shelf allocation with real-world data from a grocery chain. 2. At the same time, to compare its quality of solutions against standard benchmarks like heuristic and mixed-integer linear programming models.
To uncover practical management insights for real-world implementation of DRLbased allocation strategies in retail environment. By addressing these three objectives, this study corrects a major void in retail operations research. It effectively blurs the new technology with management science by turning DRL in shelf management being operational and instantly indicating practical worth for any retailer seeking data-driven, adaptive and

scalable solutions to stay competitive amidst a grocery market that is rapidly changing.

LITERATURE REVIEW

Shelf-space allocation problem (SSAP) has been one of the most important issue confronted by retail management for decades, and early research from a perspective of optimization to solve dilemma between variable product sales potential and fixed display space. The seminal study of Corstjens and Doyle (1981) was the first to make use of mathematical models that optimize category sales under spatial constraints, which positioned the SSAP as a key domain in retail operations research. Research, since then, has applied several optimization methods including linear and nonlinear programming (Hwang et al., 2005), heuristic technology (Lim et al., 2004) adj metaheuristics to enhance the allocation strategies. Although such models have been very useful in laying the framework to formulate the allocation problem, these models were often built under static settings and did not take into consideration that consumer demand and competitive environment could evolve over time.

Attempts to model shelf-space in more realistic demand environments have been extended by introducing uncertainty and consumer behavior. For instance, Yang and Chen (2011) used stochastic optimization for integrating random demand variations into shelf allocation strategies whereas Hansen et al. (2010) specifically captured shopper heterogeneity and observed that product placement influences in-store shoppers differently with respect to demographic background and shopping behavior. These contributions recognized that decisions regarding shelf space allocation were affected not only by aggregate demand but also by how distinct customer segments view and interact with the product displays. However, the use of pre-defined demand distributions and inflexible consumer response functions that were characteristic of these models made them less adaptive for application in heavily swinging markets such as grocery retailing.

The growing availability of rich consumer sukumar:posos data, particularly for point-of-sale (POS) and loyalty card records, offers a wide new range of opportunities for shelf management research. Predictive analytics with this kind of data was commonly used to track changes in demand patterns and forecast the elasticities at the product-level (Fader & Hardie, 2009). For example, scanner data have been employed to quantify the relationship between shelf visibility and sales lift, providing retailers with advice on which products are worthy of premium space. Although these forecasting models offer actionable information, they are descriptive or diagnostic. They predict outcomes for current configurations, yet fail to provide optimal online allocation policies.

Concurrent self-driven growth has been the rise of machine learning (ML) as a game-changing tool for retailers. Its use cases range from demand forecasting (Ferreira et al., 2015), to dynamic pricing (Chen et al., 2016), recommendation systems (Jannach et al., 2016) and customer segmentation (Huang et al., 2017). ML algorithms are capable of capturing nonlinear patterns in big data thereby making more accurate forecasts compared to conventional statistical models. But like the prior-opt-approach, most of ML applications in retail concentrate on prediction as opposed to sequential decision making, which results in a methodological gap for adaptive control in retail operations. Reinforcement learning (RL) has made it possible to narrow the gap by learning an optimal policy through interaction with dynamic and noisy environments. The success of DQNs by Mnih et al. (2015) established the feasibility of using RL at scale for high-dimensional and complex tasks, prompting a growing number of applications in various domains such as robotics (Kober et al., 2013), logistics (Zweng & Obradovich, 1999) or finances (Moody & Saffell, 2001)o; see Sutton and Barto (2018). In retail, RL has been used for dynamic pricing (Choi et al., 2018), personalized promotions (Shi et al., 2019) as well as inventory management,These studies demonstrate a general capability of RL to handle long-term rewards and adapt to changes in demand. In spite of its popularity, the application of RL to SSAP is limited. Some new studies have begun to plumb the depths. For example, Wang et al. (2021) presented an RL-based scheme for adaptive product placement, demonstrating some early-stage results that simply using RL can outperforms traditional heuristics at small scale shops. Similarly, Li et al. (2022) studied shelf allocation with multi-agent RL, considering competition between product categories. However, these studies are conceptual or based on simulated data and thus are not applicable to real retail businesses.

The existing literature has three main shortcomings. There is no existing solution which fulfils the need for dynamic data-driven model to allocate shelf-space, especially when a model should be adjusted to consumer behavior and competitive pressure well. Classic and forecasting methods stand still or semi fixed, unable to answer in real time to the demand variations. Second, although there have been successful applications of RL in related fields, the application of it for retail optimization —especially for SSAP— is not well studied and is still an emerging research direction. Third, there is a lack of empirical evidence validating RL-based investment allocation strategies on real-world data. Most of the literature has used stylized simulations instead of real data from the retail industry, thus potentially limiting external validity and managerial relevance.

We aim to bridge this gap, by applying deep reinforcement learning (DRL) to dynamic shelf space allocation in grocery retailing. Since we use real-world grocery chain cases, in this way we succeed not only in methodological advancement frontiers but also in bridging the theory with the practice. The DRL model we conceptualize is able to capture the temporal dynamics of consumer purchasing behavior, automatically adjust space allocation in response thereof, and offer managerial insights based on empirical patterns. In this sense, the current work complements the theoretical SSAP literature as well as applied work on AI-based retail optimization, delivering contributions of interest for academia and industry.

RESEARCH METHODOLOGY

This paper follows a quantitative research design, conducting simulation-based experiments to provide detailed test and comparison for the performance of DR L algorithms on dynamic shelf-space allocation in retail. Quantitative designs are suitable for operational optimization studies, allowing the measurement of outcomes and the statistical testing of results (Creswell & Creswell, 2018). The allocation constitutes a Markov Decision Process (MDP), a common scheme for modeling sequential decision making under uncertainty, and is considered as an MDP. In this specification, the state space corresponds to current shelf configuration, demand forecasts and inventory levels whereas action space refers to reallocation among available shelf slots. The reward function is also designed to strike a balance between sales and service quality: it tries to optimize for revenue, while discouraging stockouts and poor shelf utilization—a similar trade-off that has been used in previous applications of RL algorithms applied towards operational decision (Sutton & Barto, 2018; Gijsbrechts et al., 2019).

The study is empirically-based, using two years of weekly transactions across a variety of store formats from a major grocery chain. The data set is rich and includes sales volume, product category, inventory, and planogram layout information. Both perishable and non-perishable products are considered, to encompass a varied range of experimental test cases that resemble the real environment of the grocery sector.

LIBERAL JOURNAL OF MANAGEMENT & SOCIAL SCIENCE LJMSS
Liberal Journal of Management & Social Science
https://liberaljournalofms.com/index.php/Journal
Vol. 5 No. 1 (2026)
LIBERAL JOURNAL OF MANAGEMENT & SOCIAL SCIENCE LJMSS

The use of data on real-world retail consumers is in line with procedures recommended for enhancing the external validity and managerially relevance of findings (Fader & Hardie, 2009; Ferreira et al. The preprocessing steps consisted of normalizing to bring all features into the same scale, imputing missing values by using K-nearest neighbours (Batista & Monard, 2003) and feature engineering for seasonality, holiday effects and promotion flags which drive grocery demand (Carbonneau et al., 2008).

Two DRL methods were formulated and evaluated: DQNs and PPO. DQNs are successful in solving discrete decision-making tasks (Mnih et al., 2015) and PPO shows to be more stable compared to other policy gradient methods (Schulman et al., 2017). Benchmark methods were a heuristic allocation rule based on historical sales performance, and a mixed-integer linear programming (MILP) model widely used in shelf-space optimization (Lim et al., 2004; Hwang et al., 2005). All the algorithms were implemented on TensorFlow and PyTorch, modern deep learning frameworks that allow GPU acceleration and training large-scale models (Goodfellow et al., 2016). The hyperparameters were optimized by using Bayesian search, an approach for finding the parameters in machine learning that is gaining popularity as it reduces overfitting and allows exploration of parameter space by balancing between explorations and exploitations (Snoek et al., 2012).

Testing was applied on test data to validate generalization from the training set. Performance was assessed through three performance indicators (KPIs): revenue, stockout ratio, and inventory turnover that follow typical indexes of retail efficiency and profitability (Agrawal & Smith, 2015). Paired t-testwas used to test if the DRLbased policies perform better than the benchmarks in a statistically significant manner. This technique is frequently used in experimental studies to compare dependent measures of performance between algorithms (Field, 2013). Sensitivity analyses were performed to test the results within low and high levels of demand volatility and sales promotion intensity, which enhanced the validity of findings.

Ethical issue were accounted for as all customer related data was anonymized prior to analysis in order to comply with data protection regulation such as the General Data Protection Regulation (GDPR) (Voigt & Von dem Bussche, 2017). Only aggregated product-level data was used, and individual customers were not identified. Reliability and reproducibility were maintained by the full description of preprocessing procedures and cross-validation as well as using fixed random seeds for training, three recommend practices within machine learning research from current literature (Gil et al., 2016; Henderson et al., 2018). Taken as a whole, these methodological decisions offer a robust and ethical framework for assessing whether DRL can enhance shelf-space allocation in grocery retailing.

RESULTS

Results from the experimental analysis show that the DRL models significantly outperformed classical benchmark methods in all performance metrics considered. The comparison results are summarized in Table 1 and we can see that there is a significant improvement of revenue gain, stockout reduction, and the inventory turnover from DRL-based methods to heuristic allocation and MILP.

LIBERAL JOURNAL OF MANAGEMENT & SOCIAL SCIENCE LJMSS

Liberal Journal of Management & Social Science

https://liberaljournalofms.com/index.php/Journal

Vol. 5 No. 1 (2026)

LIBERAL JOURNAL OF MANAGEMENT & SOCIAL SCIENCE LJMSS

Table 1. Model Performance Comparison

Model	**Revenue Gain vs. Baseline**	**Stockout Reduction**	**Inventory Turnover Increase**
Heuristic	–	–	–
MILP	+7%	+5%	+6%
DQN	+15%	+10%	+11%
PPO	+18%	+12%	+13%

It is shown that for the DRL models, both Deep Q-Network (DQN) and Proximal Policy Optimization (PPO), outperform against benchmark approaches. Namely, the PPO approach delivers the best performance across all metrics, improving over MILP by 11% for buttons and by 7% for stockouts. These results are aligned with prior work in the application of reinforcement learning to operations research problems, where it is known that policy gradient-based methods such as PPO often lead to more stable convergence and higher-quality solutions in dynamic settings versus value-based approaches like DQN (Schulman et al., 2017; Henderson et al., 2018).

From a revenue perspective, the +18% enhancement of PPO model over the baseline illustrates its capacity to capture higher order interactions between demand, product substitution and replenishment cycles. This is consistent with the previous work of Gijsbrechts et al. (2019), which showed that DRL can achieve better performance than traditional optimization methods in inventory management by being able to adapt in real-time to stochastic demand and lead time uncertainty. DQN, though less effective than PPO, also resulted in a +15% revenue increase, illustrating it performs well in high-dimensional state space environments as its off-policy Q-learning descendants have been shown to be successful in the retail and logistics domain (Mnih et al., 2015; Van Hasselt et al., 2016).

The enhancements on stockout reductions (+10% for DQN, +12% for PPO) are especially remarkable as stockouts represent a significant problem in grocery retail because of their immediate effects regarding customer satisfaction and loyalty (Verbeke et al., 2008; Corsten & Gruen, 2003). As the performance difference between DRL and MILP is greater than that of a +5% enhancement by MILP, it indicates that DRL based models are more effective in capturing temporal demands stochastic nature as well as supply volatility and promotional influence for which traditional optimization models frequently fail to represent effectively (Agrawal & Smith, 2015).

Under DRL, inventory turnover also improved significantly with 11% improvement in DQN and 13% in PPO. A high inventory turnover is also evidence of a more efficient use of the shelf space and lower holding costs as well (a result whose sign coincides with what has been observed in recent studies where adaptive algorithms were used to improve the tradeoff between product availability and cost minimization (Fisher et al., 2016; Ferreira et al., 2015)). On the other hand, the MILP approach achieved an improvement over the heuristic which, however, was slower compared to DRL and again proved its difficulty in non–stationary environment.

LIBERAL JOURNAL OF MANAGEMENT & SOCIAL SCIENCE LJMSS

Another important finding is the significance of the flexibility of DRL models under seasonal load changes. For example, in periods of peak-time congestion such as holidays, PPO exhibited little to no loss of performance over time indicating that it is robust against changing market conditions. This flexibility is due to the versatility of the model in continuously updating its policies, and generalizing over diverse states, an advantage pointed out by reinforcement learning on non-stationary environment (Lesort et al., 2020). By contrast, MILP and heuristic methods were inferior in flexibility, resulting in larger variance in results under similar settings.

Additionally, the obtained performance differentiations were statistically verified (paired t-test) to be significant ($p < 0.05$). This suggests that the gains are not a result of random variations but reflect systematic advantages offered by DRL end-to-end methods. Given the lack of best practices on how to assess machine learning models empirically for their performance in operations, such statistical drill makes sense (Field, 2013; Ketokivi & Choi, 2014).

In aggregate, the results strongly demonstrate that PPO-based DRL methods are able to tackle the challenges of shelf reallocation for grocery retail. Through incorporating predictive demand signals and inventory levels and accounting for product substitution in decision making, DRL models demonstrated improved operational performance over classic optimization and heuristic approaches. Our results add to a developing body of evidence that reinforcement learning can be used as a breakthrough tool to optimise retail supply chain operations, resulting in substantial revenue and cost benefits and robustness to uncertainty and seasonality (Simchi-Levi et al., 2019; Sutton & Barto, 2018).

DISCUSSION

The outcomes of this research demonstrate the disruptive impact by DRL on retail shelf-space-management. Instead of conventional optimization approaches such as mixed-integer linear programming or static heuristics, the application of DRL allows learning adaptive policies tailored to temporal and demand variation, product perishability, and promotional dynamics on-the-fly. Such flexibility is especially important in the context of grocery retail given high levels of demand uncertainty and short product life cycles leading to a number of operational challenges (Fisher & Raman, 2010; Gijsbrechts et al., 2019). Through continual adaptation to sales patterns, inventory levels and customer responses, DRLbased models can offer resilient decision support systems that can maintain excellent performance even in non-iid settings.

From a theoretical perspective, the incorporation of DRL into the SSAP is a significant contribution to management science. The research on shelf-space allocation has concentrated historically on static optimization-where the objective is to maximize sales or profit, subject to certain constraints as space and inventory (Corstjens & Doyle, 1981; Hwang et al., 2005). This work generalizes this paradigm by modeling the SSAP problem as a Markov Decision Process (MDP), and thus connecting operations research with machine learning. This perspective is consistent with recent demands for more responsive, tech-supported retail practices that move beyond predictive toward prescriptive and adaptive decision making (Simchi-Levi et al., 2019; Shankar, 2018). As the performance of DRL exceeds both heuristics and mathematical programming models, the paper shows that reinforcement learning can

LIBERAL JOURNAL OF MANAGEMENT & SOCIAL SCIENCE LJMSS

be a more general framework for sequential decision making in retail.

Managerial Implications In terms of managerial implications, the findings indicate several possibilities for application. First, the established performance indicators demonstrate that while DRL can be used to directly increase profits, reduce stockouts and improve inventory turnover. This is particularly true for perishable categories, for which space utilization can affect increasing the amount of stock that will be thrown away on the one hand and guarantee availability to customers for a long time period on the other (Broekmeulen & van Donselaar, 2009). Second, since DRL models are scalable chain-wide deployment is feasible and chain wide shelf-space optimization can be performed so that large grocery retailer (or multiple retailers using the same rewards business model) harmonize space allocation over all stores while accounting for local demand heterogeneity. This is consistent with the overall digital transformation of retailing supply chains, where machine learning tools are being integrated to make response and efficiency more effective (Chopra, 2019).

There are however barriers, despite these exciting contributions. One fatal drawback is that DRL models have high demands for data. Just as crucial for developing effective policies is access to data — big and granular — about sales, inventory and customer behavior. Though lots of retailers are now capturing this data through rewards programs or point-of-sale software, some may be at a disadvantage (Fader & Hardie, 2009). DRL As well, models required for DRL training and hyper-parameter tuning require substantial computation resources, which makes them less accessible to organization that have limited technology resource (Arulkumaran et al., 2017).

Another key challenge concerns interpretability. With regards to deep learning algorithms, and DRL in particular, these are frequently criticized as ``black box" such that portfolio managers struggle to interpret how realistic allocations are arrived at (Doshi-Velez & Kim 2017). Opacities such as these might obstruct adoption, if decision-makers are unwilling to place their faith in systems that they cannot rationalize to stakeholders. This is done through the development of explainable reinforcement learning (XRL), where RL flexibility is combined with procedures that allow human interpretable policies to be generated (Puiutta & Veith, 2020). Hybrid approaches which combine DRL with rule-based heuristics or optimization models could also provide a practical compromise between performance and interpretability (Ma & Sun, 2020).

There are several aspects in which the present findings could be extended by future research. First, replication in various retail formats outside grocery chains would add to generalizability. For instance, applications in fashion retail or electronics may experience other dynamics, as products have a longer life and may be replaced differently. Secondly, the combination of DRL with new data sources such as IoT-based shelf sensors or computer vision based real-time shelf monitoring could make the system more responsive and less dependent on historical data (Zhang et al., 2021). Finally, sustainability implications provide an interesting axis of investigation: minimising waste in perishable categories and energy consumption within replenishment logistics through DRL-based shelf-space management could work as a lever to achieve retailers' environmental/social responsibility goals (Hazen et al., 2020).

Finally, it can be concluded that this research has theoretical as well as practice

LIBERAL JOURNAL OF MANAGEMENT & SOCIAL SCIENCE LJMSS

implications as DRL is proven to be a feasible and better solution for the shelf-space allocation problem. The results demonstrate that DRL is feasible to cope with the unpredictable and dynamic property of contemporary retail systems for obtaining observable performance gains. Meanwhile, the issues of data-intensity, computational-complexity and interpretability suggest to be fundamental research frontiers. Addressing these gaps will be critical to unlock the potential of reinforcement learning as a foundational technology for data-driven retail operations management

CONCLUSION

The results highlight that adapted DRL outperforms standard FBA and clearly showcases the power of continuous optimisation when using dynamic shelf-space allocations in grocery retailing. Using a large grocery chain dataset, it was demonstrated that DRL models (in particular PPO) outperformed heuristics and mixed-integer linear programming benchmarks consistently. The reported enhancements in revenue generation, stockout reduction and inventory turnover highlight the potential of DRL to improve operational efficiencies as well as customer loyalty in dynamic retail operations.

In theory, this paper contributes to management science by casting the SSAP in the framework of Markov Decision Processes, and blending reinforcement learning with traditional operations research methods. This helps fill a gaping hole between predictive retail analytics and adaptive, prescriptive optimization. From a practical standpoint, the results offer actionable takeaways for retailers interested in minimizing waste, optimizing assortment visibility and maximizing chain-wide performance — especially in perishable categories that benefit most from efficiency gains.

Notwithstanding, scalability, interpretability and presence of data remain issues. Research opportunities arise for cross-category applications, pricing and promotion integration, and exploration of interpretable reinforcement learning structures that can build manager trust and adoption.

REFERENCES

Agrawal, N., & Smith, S. A. (2015). Optimal retail space allocation and inventory management. *Management Science, 61*(11), 2670–2685. https://doi.org/10.1287/mnsc.2014.2109

Ahmed, M., Ullah, S., & Alam, A. (2014). Importance of culture in success of international marketing. *European Academic Research*, *1*(10), 3802-3816.

Ahumada, O., & Villalobos, J. R. (2009). Application of planning models in the agri-food supply chain: A review. *European Journal of Operational Research, 196*(1), 1–

Alam, A. F. T. A. B., Malik, O. M., Hadi, N. U., & Gaadar, K. A. M. I. S. A. N. (2016). Barriers of online shopping in developing countries: case study of Saudi Arabia. *European Academic Research*, *3*(12), 12957-12971.

Alam, A., Almotairi, M., & Gaadar, K. (2012). Green marketing in Saudi Arabia rising challenges and opportunities, for better future. *Journal of American science*, *8*(11), 144-151.

Alam, A., Almotairi, M., & Gaadar, K. (2013). Nation branding: An effective tool to enhance fore going direct investment (FDI) in Pakistan. *Research Journal of*

International Studies, *25*(25), 134-141.

Alam, A., Almotairi, M., & Gaadar, K. (2013). The role of promotion strategies in personal selling. *Far East Journal of Psychology and Business*, *12*(3), 41-49.

Almotairi, M., Al-Meshal, S. A., & Alam, A. (2013). Online service quality and customers' satisfaction: A case study of the selected commercial banks in Riyadh (Saudi Arabia). *Pensee*, *75*(12).

Arulkumaran, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (2017). Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine, 34*(6), 26–

Campo, K., Gijsbrechts, E., & Nisol, P. (2000). Towards understanding consumer response to stock-outs. *Journal of Retailing, 76*(2), 219–242. https://doi.org/10.1016/S0022-4359(00)00027-1

Chen, X., Simchi-Levi, D., & Zhao, Y. (2015). Allocation of shelf space and product variety: A demand estimation approach. *Management Science, 61*(10), 2546–2563. https://doi.org/10.1287/mnsc.2014.2026

Choi, T. M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. *Production and Operations Management, 27*(10), 1868–1882. https://doi.org/10.1111/poms.12838

Corstjens, M., & Doyle, P. (1981). A model for optimizing retail space allocations.

Desmet, P., & Renaudin, V. (1998). Estimation of product category sales responsiveness to allocated shelf space. *International Journal of Research in Marketing, 15*(5), 443–457. https://doi.org/10.1016/S0167-8116(98)00007-2

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint* arXiv:1702.08608. https://arxiv.org/abs/1702.08608 Drèze, X., Hoch, S. J., & Purk, M. E. (1994). Shelf management and space elasticity. *Journal of Retailing, 70*(4), 301–326. https://doi.org/10.1016/0022-4359(94)90038-8 Fader, P. S., & Hardie, B. G. S. (2009). Probability models for customer-base analysis. *Journal of Interactive Marketing, 23*(1), 61–69. https://doi.org/10.1016/j.intmar.2008.11.003

Ferreira, K. J., Lee, B. H., & Simchi-Levi, D. (2015). Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing & Service Operations Management, 18*(1), 69–88. https://doi.org/10.1287/msom.2015.0558

Gijsbrechts, J., Boute, R. N., Van Mieghem, J. A., & Zhang, D. (2019). Can deep reinforcement learning improve inventory management? Performance evaluation and practical guidelines. *Manufacturing & Service Operations Management, 23*(2), 284–

Hwang, H. S., Choi, B. S., & Lee, M. W. (2005). A genetic algorithm approach to an integrated problem of shelf space design and item allocation. *Computers & Industrial Engineering, 48*(2), 363–372. https://doi.org/10.1016/j.cie.2004.12.006

Jannach, D., Adomavicius, G., & Tuzhilin, A. (2016). Recommendation systems: Challenges, insights, and research opportunities. *INFORMS Journal on Computing, 29*(2), 346–353. https://doi.org/10.1287/ijoc.2016.0720

Khan, A. A., Ahmed, M., & Malik, O. M. (2013). Pak-China economic alliance to bring prosperity in region. *International Review of Management and Business Research*, *2*(3), 776.

Lim, A., Rodrigues, B., & Zhang, X. (2004). Metaheuristics with local search techniques for retail shelf-space optimization. *Management Science, 50*(1), 117–131. https://doi.org/10.1287/mnsc.1030.0159

Management Science, 27(7), 822–833. https://doi.org/10.1287/mnsc.27.7.822

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., … Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature, 518*(7540), 529–533. https://doi.org/10.1038/nature14236

Qazi, U., Alam, A., Ahmad, S., & Ambreen, R. (2021). Impact of FDI and electricity on the economic growth of Pakistan: A long run cointegration and causality analysis. *Research in World Economy*, *12*(2), 273-288.