IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101276

Digital Infrastructure for Predictive Inventory Management in Retail Using Machine Learning

Raviteja Meda

Lead Incentive Compensation Developer, ORCID ID: 0009-0009-1578-3865

Abstract: This paper is intended to present a digital structure able to forecast the demand of articles in a daily time span and form these forecasts an advice of replenishment orders, in such a way that forecasted incoming sales are always satisfied and stock outs avoided. The quality of this system is evaluated through a simulation process that bases its decisions in data coming from a large retail store and through on field operation in a smaller retail store. In both cases, results show a good performance of the proposed model, with substantial sales increase and costs decrease. All the above, plus the estimation of useful theoretical results, leads to a second part to present a choice support tool for replenishment orders. As this second model makes use of pre-existing ones and adjusted results of these two models allied to possible statistical simulations of the incoming orders' behaviour, it is expected an easy implementation in any retail company at no or low cost. Therefore, despite this later solution doesn't put forward a huge technological solution with great consequences over the existing job structure in the retailer, it can still be considered an improvement in the forecasting of incoming orders.

The digital structure proposed in the first part can significantly increase the accuracy of forecasts with several advantages behind it. However, in spite of the clear advantages of the proposed digital structure, it also represents a huge change concerning the structure and information flow within the retailer, with many risks of instability and huge working effort behind it. The same change brought significant no-forecast problems in the past, with severe consequences. Therefore, it was decided to propose a second part based on the assumption that the proposed model will show desirable results. First targets with this goal are a day to day analysis of the existing filling rate levels, in order to check that the given product supply level is correctly pursued, and of the weeks with bigger tumbles in order to analyze the stock outs and control promotions.

Keywords: Predictive Inventory Management, Machine Learning in Retail, Digital Supply Chain Infrastructure, AI-Driven Inventory Optimization, Retail Demand Forecasting, Smart Inventory Systems, ML-Based Stock Replenishment, Real-Time Inventory Analytics, Retail Data Infrastructure, Cloud-Based Inventory Solutions, Inventory Prediction Algorithms, Retail Forecasting Models, Automated Inventory Control, Big Data Inventory Management, Intelligent Stock Management.

I. INTRODUCTION

In contemporary retail, predictive inventory management necessitates the analysis of external information, particularly demand predictions. This business intelligence is pivotal for inventory optimization and serves as a basis for the design of certain managerial hierarchies in retail. With the proliferation of e-commerce and mobile commerce, demand modeling has developed from conventional time-series models to advanced machine learning models. In wagonload rail transport or unit train operations in which vastly more goods continuously flow than a human operator can keep track of, the correct diagnosis of service devaluing events is complicated by data streaming from a plethora of service related sensors. Here both supervised and unsupervised machine learning methods provide competitive solutions to a previously daunting task. In road transport optimization with a mix of traditional mathematical programming accompanied by new traffic simulation approaches, machine learning helps coordinate cartography tasks reliant on edge agents in a 5 digital geographic regime. Forecasting demand for a multi-branded retailer necessitates both a highly structured in-store demand generation and a stochastic consideration of factory operations. Here the faster than proportional inventory-growth policy is complemented by a simple and interpretable neural network feed-forward arrangement for demand prediction.

Inserting the prediction of demand nonlinearities, the forecast error of demand was reduced without sacrificing interpretability. Encoding the numerically generated demand into a highly polytopic geometry, a convex sample policy supervenes on a Markov decision process model incorporating machine learning forecast input embedding. While challenging, it reigns down the performance disparity amongst the producers of alternative brands of goods and obfuscates decision making in unbranded goods retailing.

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 12, December 2021
DOI: 10.17148/IJARCCE.2021.101276

The convergence of high dimensional state-action space to a finite space via clustering greatly eases learning while sample fitted approximating the value function while meeting the rapid state variation criterion of open maximum mean discrepancies.

1.1. Background and Significance

Advances in new technologies have resulted in increasing the speed of data generation and accessing larger data storage. This is evident from the increase of data generation speed in household products such as smart fridges that are able to scan and keep track of the product inside the fridge. In retail, these systems help create a new data-rich environment in data generation speed and volume since many of the products are RFID tagged and tracked through the whole supply chain tracing the history of each product. Supermarket chains are scanning, counting and investigating the items in their shelves on 24 hour speed and the data is stored in their data lakes.

These environments have not been utilized to pool and analyze the underlying data to gain insights as the same amount of data could exist in a process without new technologies. The emergence of multilayer perceptron neural networks allows for the processing of larger dimensional data. The other major group of new algorithms is the tree-based methods such as the classification and regression trees. Such methods are robust to noise, scalable computationally, and do not require much feature engineering.

Fig 1: Inventory Management in Retail Using Machine Learning.

While ML applications have been researched in logistics and supply chain domains, such works have been limited compared to other business domains. In addition, only a few papers on a generic review of ML applications in supply chain decision-making problems exist. In addition, most of the review papers focus on specific supply chain tasks such as transportation, inventory, or scheduling. The costliest part of the supply chain is the labour cost and wage prediction cost estimation. Better prediction could mean winning tenders and projects or avoiding losses as a subcontractor. This sub-problem has been less studied and very limited attempts with ML exist.

II. LITERATURE REVIEW

Nowaday, the ability to collect and analyze large volumes of structured and unstructured data provides new opportunities for improved decision-making in business. However, to implement big data/information solutions and achieve satisfactory results, companies need to overcome both organizational and technological challenges. Exploring the management of rare data for predictive inventory, this research delivers some preliminary pragmatic insights through an in-depth longitudinal case study. Unique industry characteristics and data ownership/sharing challenges expose the obstinate enigma of predictive inventory management; accordingly, proposed future research directions and potential managerial implications are discussed. The impact of such advanced retail analytics on the future of retailing—especially the survival of shops in this big data age—is speculated. Grasping this “silver lining” as an opportunity rather than a challenge is regarded as the first crucial step towards a new and exciting world of retail analytics.

An exploration of big data practices in the retail sector reveals that, to implement big data solutions, there are prerequisites and challenges that should be overcome. First, there are companies that have actually started using classical data analysis techniques and feel no need for more advanced big data solutions. Second, there are companies that are interested in the opportunities provided by potential big data solutions. In addition, the feasibility of such big data techniques needs to overcome three challenges: responsibility and accountability of people, people's skill set, and the necessary support from their suppliers.

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101276

In case of the big data practices in retail operations observed, the processes need to be integrated, or the resulting improvement on effectiveness in one operation could be hindered by other operations. The processes need to be integrated, or the resulting improvement on effectiveness in one operation could be hindered by other operations.

Machine learning is one of the most prominent technologies for predictive inventory management. Various machine learning approaches have been proposed for inventory management in retail, such as. However, these approaches generally require complicated and cumbersome dataset construction efforts, as they cannot incorporate an efficient simulator for readily generating synthetic dataset and therefore must rely on complicated dataset generation processes. To address this issue, a learning-based approach is developed wherein the simplest (s, S)-policy inventory management is learned. The learning-based approach for inventory management has been studied since the year 2000; this approach treats the demand data as a time series, and determines the order quantity at each time step using reinforcement learning.

2.1. Historical Context of Inventory Management

In every supply chain, one of the main challenges is managing inventories that are not under or overstocked. A well-managed inventory can meet customer demands with minimal holding costs. This is a goal that all businesses have to fulfill according to their conditions. Regardless of their importance, a poorly managed inventory can result in a stopped production line, sales loss due to back-ordered customers, out-of-stock materials, a huge amount of extra costs, and more unmanageable problems. There are many components affecting supply chain inventory management strategies such as forecasting, lead time variability, lumpy demands, and inventory control policy. Traditionally there are two kinds of methods; optimization-based techniques and simulation-based methods. The former has high computational complexity and requires a realistic mathematical model of the concerned system with parameters. The latter tries to replicate the current system better but a lot of effort is needed to obtain changes in numbers. After the 1990s, AI technology, which had not been emphasized in supply chain management systems, has been leading the improvements due to its ability to analyze a huge amount of raw data. On the one hand, a lot of researchers have been attracted to analyze company datasets and proposed models for reducing costs based on these analyses such as demand forecasting and supplier evaluation. These models do improve efficiency, however, without controlling the policies they do not guarantee the minimizing of operational costs. On the other hand, in order to control supply chain systems, researchers started to focus on system-wide optimization e.g. determining production and distribution decisions based on demand forecasts. However, regarding the increasing size and complex structure of real SC networks, even a single production-distribution pair is out of reach for existing optimization methods. The lack of research on controlling large SC is the motivation of this dissertation.

Equ 1: Performance-Based Agent Reward Function.

$$R_{agent} = w_1 \cdot A + w_2 \cdot S + w_3 \cdot T$$

- A: Accuracy
- S: Stability
- T: Timeliness
- Weights w_i optimized via reinforcement learning

2.2. Machine Learning in Retail

The recent improvements in data availability, computational power, and algorithmic approaches such as neural networks are impacting several domains, and Inventory Management is no exception. For inventory-based supply chains, understanding demand for items (skus) can allow firms to value and plan their resources better. Different firms are structured differently, with differing levels of automation for ordering decisions. The main challenge remains on using data on sales and other features to choose the best order placed at each time window, balancing overstocking against stockouts. Order placement rules have thus far relied on heuristic rules, but retailers aspire to become more data-driven. Knowing retail-level predictive demand enables firms to order exactly what they need. They need to know demand, and the Forecasting task comprises the prediction of predicted demand at each time step for each action (sku-store pairs). It includes the construction of highly parallelizable ML models, able to cope with the problem's scale.

Alternatively, an order placement model takes this predicted demand and predicts the order quantity achieved from it. The recent supply chain crisis has raised prices across the market worldwide. Pre-COVID times, s-order policies assured low waste. The increased number of inventories thus needs a data-driven approach to work accurately. Reinforcement Learning (RL) becomes more widely adopted outside academia in various industries, such as finance, HR, and Supply Chains. There are more studies based on RL on Pricing, a close alternative to the Inventory Management task. Other works on Inventory Management generally apply rule-based models.

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 12, December 2021
DOI: 10.17148/IJARCCE.2021.101276

Nonetheless, RL-based methods generally outperform similar heuristics in other demanding tasks such as Pricing. They define the environment, agent, and reward function according to the characteristics of the task. These choice flaws become critical when unrealistically simplifying.

2.3. Current Trends in Predictive Analytics

With the increased complexity of selling processes and the need for timely decisions, predictive analytics is gaining more attention from retailers. The rapid evolution in artificial intelligence and machine learning brings lots of opportunities to improve the precision of frequency estimates and item sales trends and to automate the process of managing decision makers. Most implementations, though, are in the retail domain; other industries, such as fast-moving consumer goods, media and entertainment, or transportation, companies have clung to off-the-shelf software, which primarily focuses on pre-prepared reports in general formats, limiting the value of insights gained.

AI forecasting methods, often on the same level as algorithms, can generate parameter estimates or frequency predictions only, making it tough to design better use of them. In the framework of these algorithms, the stakeholder intervention in the design of better inputs and assessments such as the up-to-date exposition of newly arriving data takes high effort and resources. Furthermore, historical data usage is evaluated, potentially limiting the information input to the forecasting models. Source Remapping provides flexibility in quickly redeploying solution environments, mitigating the specialized method entrance cost.

Finally, the importance of the compressing history is pointed out, aiming to promote practicality in managing enormous amounts of data and computational load. Evaluation based on how the multi-layer mapping structure controls more archetype class regression learning systems cannot validate generative decisions due to their abstraction. Comparison in systematic evaluations of prediction precision versus state-of-the-art methods can better approach extensive changes undertaken in AI.

III. DIGITAL INFRASTRUCTURE OVERVIEW

The retail sector is increasingly recognizing the significance of managing its supply chain digitally to respond to unpredicted demand concerning disruptions. Human-centric, error-prone, and inefficient manual monitoring methods are giving way to automation dyads of machine learning (ML). This paper presents a systematic indication of digital infrastructure, based on five technologies and ten capabilities, designed to augment predictive inventory management in retail using ML. Retailers can self-assess their readiness to conduct predictive inventory management using ML on a five-point ordinal scale by operationalizing each capability quantitatively. Users can identify the desired digital capacities, select the related technologies, provide a path to implement those technologies, and manage their actions to provision all the warranted capabilities once pre-implementation assessment is complete.

The retail sector consistently generates massive volumes of data as a result of digitization initiatives and behavior changes of customers, quickly rendering organizations to collect data without a concrete purpose in mind and even fear of being overwhelmed instead of effectively managing them to generate business value. The data generated by numerous stakeholders across the supply chain is nevertheless, too large, too bready, and/or too fast to be utilized as planned. Nevertheless, prior to these changes, a number of retailers experienced detrimental consequences due to over-in-validated, inefficient, or relationally warehoused early steps of digital data infrastructure development concerns of IT and BI maturity levels. Retailers should thus competently manage transformative big data overhauls of many of existing systems concurrently, while at the same time augmenting the data maturity close to its requirement initially.

Retailing is additionally very sensitive to macroeconomic factors through their great impacts on customer behavior. For instance, the COVID-19 outbreak and its pervasive disruption on daily life previously outmanoeuvred all scenarios and preventive measures in place and subsequently instigated panic buying habits in several countries. Occurred Black Swan events or other exogenous threats should thus be professionally accommodated by taking the defined risk management steps meticulously. Digitalisation of the interior systems and processes intensively, on the other hand, comes with severity, complexity, speed, reliability, security, and ethics requirements. As a consequence, wide-scale advocacy of stochastic and abduction-based resource allocation using AI enabled forecasting in the binomial classification of its aligned decision making, and alert systems to substitute for anticipated unpredictable movements using ML-ML systems are deemed required to be included.

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101276

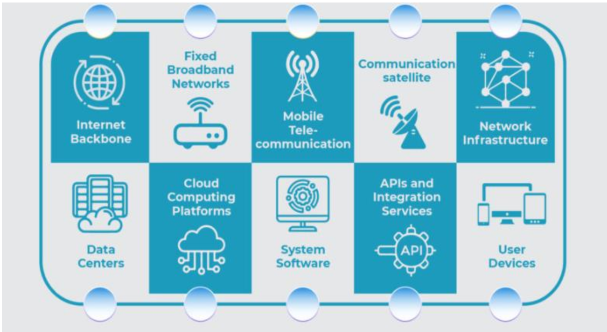

Fig 2: Modernizing Digital Infrastructure.

Convincingly operationalising the definition, the capability of predictive resource allocation, and its IT in the retailing literature is a missed industry opportunity. Whereas on the basis of typically consuming historical data and putting reliance on random forest, machine learning (ML) is instead preferred along with the richness of inputs upon bounding errors and responding in multiple stages.

3.1. Components of Digital Infrastructure

Related to the application of AI-enabled decision-support systems in retail analytics, a digital infrastructure is recommended to help retailers work on AI solutions with sufficient preparation. This digital infrastructure is represented in terms of four components: data architecture, business applications, cloud architecture, and coalition, and is further elaborated upon with subsequent subcomponents. Representing the data architecture component, the data scientist toolkit describes the data architecture required for data engineering and data mining. Data engineering can follow a state-of-the-art batch processing pipeline or real-time streaming processing pipeline befitting data types and use cases. The data mining part concerns how retailers' business and macro policies can shape customer behavior and thus must be incorporated into the development process of AI solutions.

With the increasing variety of business applications ranging from basic applications such as dashboards, forecasting, and clustering, to advanced predictive applications with internal facing and customer facing, it is suggested that retailers can work on business applications robustly if they pursue business applications incrementally. The cloud architecture enables retailers to build a scalable AI model training and serving infrastructure. It also provides customer-facing applications a better availability and performance.

Finally, not necessarily related to technical components but deemed necessary for achieving maximum benefits from AI implementation, the coalition component represents best practices for working on AI solutions using collaborative and cooperative efforts among roles in different domains. With regards to AI solutions, the coalition scheme is further elaborated in terms of overall guidelines, intensity, and forms of interaction, as well as concrete roles and their cooperation. Implementing an AI-enabled decision-support system is a complex task requiring a broad range of expertise across different domains. Although retailers can employ third party vendors to handle the implementation, it may reduce their ownership of the solution and its underlying knowledge, and prevent them from enhancing internal capabilities.

3.2. Cloud Computing in Retail

Data accessibility has all of a sudden improved, allowing capable analytics computations on cloud computing environments at affordable prices. The cloud has the ability to combine these two previously distinct streams of capability in a brand new and exciting way. The contentious nature of cloud computing has been recognized with academic literature addressing its definition, characteristics, and applications. Despite the rapid growth of e-commerce, ecommerce companies still lack the guidance to build the cloud computing architecture. Companies typically ineffective management of information technology performance (IT) strategy implementation after investing extensively in IT systems. There's a lack of literature introducing the retailers cloud computing architecture.

For several decades retailers have been collecting point-of-sales (POS) transactions, inventory level, shipment, and forecast data, culminating in increasingly large and rich data stores. Recent advances in search and retrieval technologies have rendered it practical for retailers to use this data. A few research efforts have focused on the storage and retrieval of this data for customer analysis; optimization; or decision support. These products help companies examine point-of-sale data to monitor service and sell-through, thereby helping corporations react to seasonal variance and gradual changes in consumer behavior. Further, information brokers are using them to monitor inventory levels, prices, and promotions of competitors, as well as for processing levies and trade offers changes.

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101276

The second step is data analysis, where cloud-based data analysis is combined with retail expertise for predetermined queries. This step involves exploring data to formulate behavioral business meaning primarily through optimization and statistical modeling techniques. It leads to a list of tactics, best practices, and policies available to retailers to consider. An example would be to automatically develop symbolic comparisons for list-price competition. Verifying these types of tactics would be more intensive data and compute-wise, leading to a set of prescribed behavioral insights to report. This step reduces the subset of insights from the commercial strategies that will drive business value to drive business value, similarly to other scientific pursuits. This output will determine the physical form of the gallery of displayed and recommended solutions. Retailers will still want control of their proprietary strategies developed over years of experience and arbiter, though still employing the analytic expertise revealed in this step.

3.3. Data Management Systems

In relation to Point 2.2, data regarding price discount, inventory and promotions should be collected from the information management systems within the company accordingly. These data can be replicated on a normalized relational database where each table serves as a dimension in a star-like model. Dimensions include Time (date), Store (characteristic store data session), Item (categories of items for analysis and weekly segment sets), product (characteristics of each product), and Fleet (characteristics of promotional events). Facts include sales, also including price, inventory and promotion information.

A micro database operational prototype can be immediately developed to store item's prices, sales and permutations, and be tested in the company's data. It allows queries similar to SQL; e.g. view last year sales or how average price affects sales; the only weakness being the relationship between tables, thus not enabling complex querying. Further solid, industrialized results can be achieved by developing a relational database in the server for working with larger amounts of data, where the need of external assessment and audits is increasingly important.

427 retail stores in Portugal with hourly prices, promotional events on markdowns on cover prices and footfall (daily) data from 10 mins past 0h until 00h the next day. Data covering 3 years with remaining behaviour and sales of 600 products and their categories on markets are available. 1900 promotional events and 5 day windows (i.e. an average of daily transactions before, to be added). Typically, a 4-8 week lead-time prior to an event is assumed within the retail industry. Predictive modelling has been addressed in the scope of their effect over reference performance metrics to suggest efficient set pricing. Peer group learning models are tested across hundreds of products and thousands of stores. Further, knowledge about the behaviour of a product's peer group on other stores may improve the estimation of sales, hassles and the return on the promotional investment on the other stores requiring such information.

IV. MACHINE LEARNING TECHNIQUES

Consistent with current changes in retail, shoppers today want a wide selection of items. As a consequence, businesses are inundated with requests for enormous amounts of data, events, and signals. Retailers can no longer depend on their historical data alone; they must embrace change and seek outside assistance in order to find service opportunities based on a variety of indicators that can be understood in real-time. Due to the requirement for immediate and ongoing profitability, the retail supply chain's legacy approach is insufficient. In addition to historical trends, it is also anticipated that sudden developments will be taken into account. New types of data sources are proposed, not just based on transactional balance data but also giving a picture that can incorporate other types of sources such governmental, news, weather, etc. Monitoring service indicators will provide retailers with far more information than can be handled, which will list dozens or hundreds of alerts for a single day in order to identify service chances.

Numerous smart systems that can provide a wider perspective on changes related to stock-out chances have been researched and built. They apply supervised learning methods to any incoming pool of service indicators in order to determine service opportunities. Nevertheless, it's still difficult to recognize unanticipated chances based on a continuously updated picture without reliance on additional labels. The majority of services are odd and must be quickly evaluated. Companies rely on their knowledge of the business environment and CSS experts to guarantee utilization in order to avoid the massive investment predicted by unsolicited services that may never be used. A method is presented that allows for the speedy and iterative discovery and assessment of such service opportunities based on a continual stream of indicators and past CSS. A case study is used to illustrate the different aspects of the methodology. User-friendly tools are proposed to simplify its use in actual business settings.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101276

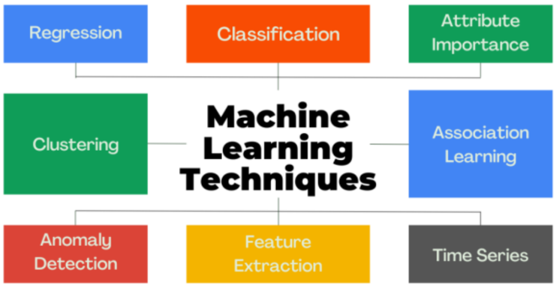

Fig 3: Machine Learning Techniques.

4.1. Supervised Learning Approaches

Supervised learning approaches that have been adopted to properly model the average demand and its uncertainty. Demand forecasting is critical since it can fully enable policy-making processes based on knowledge regarding the future using micro data on customer behaviors and their interrelations. A variety of forecasting models have been explored in MRP systems: from simple 'naïve' models and classical time series smoothing models such as Holt-Winters to the modern 'data mining' techniques such as regression or the data warehouse approach. Recently the advances in the field of artificial intelligence and machine learning (ML) have created new approaches for demand forecasting. The power of numerous machine learning techniques has been explored for demand compliance and their suitability for optimization engines.

Since demand forecasting is the bottleneck of the MRP systems it is worth investigating extensions of the knowledge base with learn-trial-adapt capabilities using supervised machine learning (ML) techniques. Supervised learning requires a certain amount of past data for inferring the predictive models enabling the examination of the potential of machine learning in demand forecasting. Traditional techniques range from the naïve approach to the sophisticated exponential smoothing and regression techniques such as ARIMA, GARCH, Holt-Winters, and various adaptive filtering methods depending on the nature of the time series data distribution with regards to seasonality. In addition to their robustness and minimal requirements, on supply chain design higher accuracy levels have been achieved with sophisticated regularization and linear term based methods at the expense of performance prediction horizons. Following the initial preprocessing, regression trees and ensembles of regression trees have also been tested and outperformed all the other methods.

In advance of the proper demand model, the uncertainty needs to be captured for risk controlling. Short-term forecasts are generally assumed to be uncontrollable and the focus is solely on modeling average demand. In practice, however, their distribution must also be estimated as they play a dominant role in setting safety stocks, that is the minimum stock level that a shop should keep at all times. With regards to uncertainty modeling, most methods estimate the variance or empirical distribution of demand deviations through statistical inference and interpolation. Only a few methods integrate model-free ML-based approaches as in the case of quantile regression to model the conditional quantiles of the uncertainty band during and after level curb regulations. However, this approach assumes that demand distributions remain unchanged and stationary in the long run, whereas non-parametric density change detection methods using kernel density estimation are also available as filtering methods.

Equ 2: Supplier Ranking by Autonomous Agents.

$$\text{Score}_i = \alpha Q_i + \beta R_i + \gamma C_i - \delta D_i$$

- Q_i: Quality score
- R_i: Reliability
- C_i: Cost
- D_i: Delivery delay

4.2. Unsupervised Learning Techniques

Machine learning (ML) models can broadly be divided into supervised, unsupervised, and semi-supervised learning. This section provides an overview of these three general machine learning techniques. Under each subsection, a description of the different algorithms was provided along with a review of use cases and limitations.

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101276

Unsupervised learning models learn from data that has no target value. The central idea is to find a structure in the input data by grouping together similar data points. Unsupervised learning models can be divided into clustering and dimensionality reduction. Clustering models categorize a set of items into groups by measuring their similarities using distance metrics and defining a criterion function. K-means is one of the most popular clustering algorithms and is frequently and appropriately applied to various research fields. However, K-means has numerous limitations that should be addressed while using it.

For example, it cannot detect noise, and it requires defining the number of clusters beforehand. To address some of these drawbacks, many variations and improvements for K-means clustering have been proposed. Helping to categorize the clusters and selecting the best number of clusters was a study conducted by. Another method used to improve search performance was designed and implemented by a real-time adaptive dynamic evolution of K-means clustering with decentralized information processing and data locality in a distributed dynamic environment. Although the efficiency of K-means clustering could be enhanced, data normalization still had to be addressed before executing the clustering stage. To solve this issue, local K-means-i clustering data normalization was developed and successfully applied in numerous use cases.

Clustering models have been applied to categorical data within many intelligent mobile phone applications. One such application is approval of wholesale orders. Several clustering models were utilized to find similar vendors and analyze their purchasing habits to automatically approve wholesale orders. Usually, there are several clustering models available, including K-means, spectral clustering, Agglomerative clustering, CURE, DBScan, etc. Since K-means ties down data at the beginning stage, other algorithms with better exploration capabilities were also considered. On the other hand, Dunn's index was computed and validated in several clustering models.

4.3. Reinforcement Learning Applications

In recent years, reinforcement learning (RL) has found increasing applications in retail supply chains. A RL algorithm is typically implemented in two parts: the controller and the environment. The controller executes a policy which consists of actions, such as adjusting stock levels in response to sales and inventory. The environment responds with a reward typically equal to the increase or decrease in profit over the time parameters of the simulation which may be days, weeks or longer. The controller learns from exploratory input samples, the current policy dictates which action to take, using a model that must comply with the constraints and be executable. This learnt model is usually near optimal for that environment and can suggest actions from a small number of samples. The demand is often assumed to be a stationary process in RL applications by many researchers. With the covid pandemic causing a structural shift in demand for many retail sectors, addressing the challenge of long-term adaptability to changing demand in live or near live RL applications is a frontier of interest. A thorough review of the RL and Supply Chain Literature has been presented to describe various RL architectures for RL applications in Retail Supply Chains. As RL has matured and become more widely adopted as software libraries become more widely available, it is which of several standard RL techniques might be suited to retail. The RL methods that demonstrated success in Retail Applications have been evaluated and their suitability as open-source software projects. To demonstrate flexibility and ease of usage, Library of Standards RL Control Problems comprising classic RL examples with the OpenAI Baselines software library are then implemented.

Reinforcement learning (RL) is a machine learning paradigm which has gained increased prominence in recent years as RL techniques, algorithms and software libraries have matured. RL has applications in stochastic sequential decision making and optimal control in domains where the environment is not fully known and cannot be pinpointed precisely, involving uncertainties and where unintended consequences might arise in control actions. The environment typically behaves like a Markov Decision Process (MDP), a set of states and discrete time steps, with observables drawn from the state. Actions taken may have delayed consequences upon the environment and trigger transitions between states. Each environment has rewards drawn from a probability distribution which quantify the goodness of the resulting state. The goal of RL is to derive the optimal policy of action.

V. DATA COLLECTION AND PREPARATION

The advent of some robust Data Analytical mechanisms has paved a new window for space exploration in significant functional areas. Such paradigms that now come under the umbrella of Analytics are Predictive, Prescriptive & Diagnostic. While Predictive Analytics talks about the past studies over the data, Prescriptive Analytics talks about the solutions or steps to be taken by analyzing the past. On the other hand, Diagnostic Analytics is the mechanism of Predicting and knowing the demand for different SKUs (Stock Keeping Units) so as to manage the inventory better. According to the understanding of the Dynamic Market Structures, the 'Out of Stock' issue prevails generally due to improper demand per store nor per SKU.

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 12, December 2021
DOI: 10.17148/IJARCCE.2021.101276

In order to get rid of it, established MNCs usually have robust Demand Forecasting Mechanisms to forecast the possible demand Error, thereby managing the inventory subsequently. Various methodologies without determining the performance criteria are available in analyzing the past data. Rigorous experimental analysis using considered Models sheds light on Demand Forecasting with best and feasible methodology as Discounted Holt-Winters Exponential Smoothing. In many Multinational Companies Independent Demand of Products and its capacity to be forecasted through Descriptive Models of different Orders of ARIMA, TS Models like HWinters etc have been already tried and tested. Presently Non-Demographic, Reliability Factors as well as Lagged, Weather, Calendar Variables etc. Variables influencing Market behavior are well studied with Generative Forecasting in Time Sliding Window as Extended State Space Models with Auxiliary Particle Filters.

5.1. Sources of Retail Data

Retailers use different sources of data to support decision-making processes. Historical sales record is the main data source for forecasts. Weather observatories and public record offices have historical data about weather conditions. Weather is significantly related to the sales of ice-cream, raincoats, and other products. Historical sales help to measure holiday effects, for instance. Many retailers have planned discounts for their products over holidays. The most common ones are ‘Monday Sales’ and ‘Black Friday’. With historical sales data, such effects can be measured during different holidays. The first and the second halves of the months usually yield different fulfilment rates due to the prime of business. Most of the shelf products are replenished through purchase order forecasting. However, new product estimates cannot be based on sales data because there is no historical data available. Instead, sustainability management has been practiced since the introduction of international standards and voluntary initiatives. Relevant data can be gathered from the customers, suppliers, and authorities. For instance, the carbon footprint of greenhouses, transportation, and customers should be considered. In addition to governments, meteorology departments also share data about real time and forecasted weather data. Historical sales data would enrich the meteorology dataset. The potential sources would include informational kitchens, industrial entities, and agricultural organizations. There have been studies about agriculture, however, because of lacking access to actual data, the results of these modelling would not be totally usable. Major weather elements and indicators for ranging and forecasting must be extracted from these data. The database should be modularized for easy maintenance and personalisation. Processed records would include meteorological values having significant relationships with unknown target data. Other retailers have been requested through appointed sales representatives to collect their sales data. Past and current data of cars have been collected whenever possible. Forensics Unit of Corporate Service Department would make on-site inspections to avoid misuse.

5.2. Data Cleaning and Preprocessing

The data used in this work consists of 7 data frames that contain different features for modeling sales in a retail context. This section presents each of the data sets used and how they are treated to put them in the best conditions possible before feeding them to the machine learning models.

The set is formed by the sales transaction of a big supermarket. It contains in its raw form thousands of transactions on gigabytes of data over 1 year and 10 stores. The construction of a data set to model a forecast is a big challenge, as it requires significant data cleansing and preprocessing steps to expose the data in the best condition for time series forecasting using machine learning techniques. The sales data is made available down to single receipts. In addition to the date, time, and value of the transactions, a bunch of information is included such as branches, department’s code, and name, SKU, and family names. A data set containing prices of the SKU is also supplied, which indicates that sales are very likely to be price dependent and this has a significant impact on the demand for products. The tables and how these are used will be introduced afterwards, but first, the pre-processing steps taken to generate a base sales forecast, what is assumed to be the worst case scenario, will be presented.

The data set needs to be transformed to obtain a sales measure that can be modeled. One of the focuses during the training phase is to assure that the model has generalization properties. The model can memorize the training data, focusing on its specific properties. A phenomenon known as overfitting occurs as, for instance, this will base the forecast on unique particularities of the training set. In this case, any forecast simply constant and equal to the last value of sales is akin to a perfect model. To avoid this phenomenon, the model complexity has to be restrained. In order to tackle this challenge, a cross-validation approach that evaluates the model’s performance with respect to different views of the problem has been adopted.

Sales forecast modeling is tackled as an ensemble of local sales forecasts for semi-stationary time series, i.e., time series in which statistical properties can be expected to change when some exogenous parameters change states. One of the issues here is the data volume, which hinders any extensive model search to be deployed. An approach is proposed that aggregates the hierarchical time series of a multi-branch, multi-product and multi-department into a consolidated national

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101276

sales forecast that is produced quickly and with minimum computational efforts before being disaggregated into branch and product forecasts. The forecast combination strategy considered here is a sweep and prune search implemented in the time domain. A weighted average forecast is the most common data-fusion strategy, but plausible alternative forecasts can be generated by manipulating both the model weights and the forecast weights.

5.3. Feature Engineering for Inventory Management

Effective utilization of large amounts of data generated during the execution of SCM processes (in particular data from entering, informing, and monitoring SCM processes) enables a swift shift from data-driven planning decisions to real-time proactive reaction to SCM events and assists the efficient resolution of modeling and prediction uncertainties that can be better managed by different kinds of ML techniques. A novel sampling method was proposed to utilize the latest available incoming data more effectively. In addition, a multi-tier storage mechanism offers a balance between data availability, data accessibility, and data if there is a better performance within a certain storage tier. Hence, the present data warehouse infrastructure needs to shift toward the following sophisticated features: Multi-tier data storage. The current static, single-tier data storage system lacks a regulatory control that performs proactive diagnosis and monitoring of data and access performance, resulting in ineffectiveness in storing, pre-processing, integrating, and accessing vast amounts of raw data. Effective indexing of pre-processed data blocks via an SPG and integration into an aware storage architecture is drafted. A proactive data delivery approach issues volumes of time-sensitive and heterogeneous raw data to computing environments in advance.

Fig 4: Inventory Management System.

In addition, a data pruning assistant is devised to reduce data-transfer loads and costs. Since massive meteorological and sales datasets would at least be processed in the same time span at different computing environments, the pre-processing operations would be performed on incoming raw datasets collected in new periods as well. The pre-processed data would be stored in external storage networks, which would present a further delay in integrating such computing environments with on-going time-sensitive marketing plans. Hence, prior to the arrival of measured weather-data points, datasets are derived to forecast temperature for inquiring purposes. With the pre-processed data, it is because the length prediction is a way more sensitive technique than the former method to recover probabilistic forecasts, which need faster access such as a Graphic Processor Unit access available locally. The thunderstorms that would occur on Monday evening would be recovered at an earlier time range.

VI. PREDICTIVE MODELING

There is an increasing interest to incorporate machine learning techniques to improve supply chain mechanisms. A complete workflow adopting ML techniques to efficiently model large-scale demand or inventory forecasting is presented showing its application on a real-life Spanish grocery retail dataset. E-commerce supply chains are characterized by uncertainty due to turbulent user demand and complex information and material flows. The proactive allocation of various resources is needed in the entire supply chain to eliminate or control disruptions. develop an Internet of Things (IoT)- and machine learning (ML)-based proactive framework for e-commerce supply chains to eliminate disruptions with the collaborative engagement of supply processes, demand forecasting processes, and inventory management processes. The recommended algorithm can be applied to off-the-shelf order arrival datasets. Its capacity to handle large datasets from different sources enhances data reusability while the designed multi-agent system is able to optimize the supply chain configuration according to the actual need and requests to achieve maximum benefit.

6.1. Model Selection Criteria

The criteria applied in selecting the model and parameters for the prediction task includes the prediction error, i.e., the MSE, the time needed in training, and the number of model parameters. The more complex models would typically achieve a lower/divergent MSE compared to the counterparts.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101276

Models that take longer in training can be expected to better apply the historical data and thus yield a better prediction performance. In that sense, the MSE is selected as the first criteria to select the model structure, while the time needed in training or evaluating filters out the models that consume an infeasible amount of computation resources. There is a linear trade-off between the training time and model depth. Future work should compare alternate model architectures for this task. Comparing different model structures, having more regions almost always improves the performance for the traffic prediction. The majority of the models studied, taking as little as around half a minute to as much as around two hours, eventually reach a low prediction error (the validation loss converging to around-4.75). However, the performance on the other dataset is more delicate or tricky. The MLPs with up to three layers work without too much of a problem on that dataset; models having four layers not only achieve a greater (though not diverging) prediction error, but also fail to converge as the learning rate would not go below a certain threshold. More test runs are needed to fully diagnose the problem and choose the model more wisely with a clearer understanding of the effect of the shared parameters. Also, there are opportunities for using a larger training dataset with data aggregation, refining or regrouping features.

6.2. Training and Testing Models

On the data preparation, data for both demand and external features are cleaned while preserving the information for training purposes. If a data point is still helpful for prediction (normalized to use historical quantity), but has a missing external feature (mean imputation), it would not be removed. Features with wholly missing value (due to old arrival date) would be removed at last. An acceptance ratio of around 95% is observed, meaning the discarded data points would not degrade prediction. Sales of the first two weeks after entering the training set are considered new points and predictions are made as a halt basis, 495 samples at most for each SKU. Five features (June, July, Holiday, Datetime Stamp, Quantity) are discarded for the purpose of model training/testing. This is because a proportion of those variables must be known in advance when fitted at production. Purposes for the acceptance ratio and halt basis choice in prediction are for a fairer comparison among prediction approaches. 495 samples are more than one year's weekly demand on average and sufficient for the purpose.

The predictive machine learning model used in this study is the popular structure of Multi-Layer Perceptrons (MLPs). A hidden layer refers to any layer of neurons after the input nodes, but before the output nodes. The number of neurons in a fully connected layer is the number of input nodes to which this layer connects. The previously mentioned hyper-parameters are tuned one by one according to the structure and the number of layers is selected first. Approximation monotone accuracy is observed by increasing layer number. Nevertheless, networks with many parameters would become challenging to train. It is practical to limit the number of hidden layers to three, which was generally found sufficient for most applications. As for the training iteration process, when increasing the maximum iteration (starting from 200) for MLP with hyper-parameters set from their smaller choices, validation up's and down's seem to flip-flop (e.g., hold one value for epochs).

6.3. Model Evaluation Metrics

In any prediction problem, apart from the core tasks of feature engineering, modelling, hyperparameter tuning, and productionization, the features, models and metrics used to assess prediction accuracy are equally critical. In inventory management, if the metrics to evaluate are not thought through, it may happen that an apparently accurate prediction model does not translate into better actual stocking decisions. Also, having effective metrics itself is not sufficient, because if measurement has a high cost, the model accuracy will look very good for one metric and disappointing in reality. The metrics used to evaluate the fulfilment and utilization index models are "at demand" relative mean absolute error, and the "coverage" metric measures how much of the predicted demand was actually consumed. Therefore, although the prediction models in production had a lower accuracy than the proposed baseline models on these metrics, they highly successfully met the overall goal of ensuring good standards of customer service. For example, in the utilization index model, there are days when hundreds of value SKUs are predicted and just one or two are actually consumed (trained with 14K training data points). An astounding 98% accuracy can be claimed on the "coverage" metric. In reality though, an accurate prediction has occurred of no SKUs in the "predicted" set to then stock given a huge opportunity cost of future revenue foregone.

Accuracy of predictions relative to a time series or set of time series at any point means errors of the target predictions of distributions of these time series will generally be computed at the next point in time or later. Predicting a definite value quantile (say pledge time) is conceptually similar but definitely more complex, as it means estimating numerous, probably millions, of candidate regression hyperplanes (the quantile "estimators") that predict quantile relative to any chance quantiles (the "covariates"). Regarding skilfulness evaluation of probabilistic predictions of time series, Mannshardt provided an ingenious means to assess the predictive validity of pairwise multi-state forecasts of quality percentiles from point estimates.

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101276

VII. CASE STUDIES

This section presents a collection of use cases outlining the value of machine learning and data science as well as practical deployment challenges faced by large teams. The first customer case study outlines demand forecasting for an online fashion retailer in India, and the second use case covers a short-term forecast of order arrival for a cross-border e-commerce scenario.

It is also worth noting that both case studies do not cover the technical aspects in detail, but rather their usage and implications. Customers who successfully adopt certain use cases often engage in business need articulation and implication on working methodology in greater detail than the technical or mathematical aspects. Therefore, both case studies outline the initial motivation and primary reasons (both domain related and methodological) on why the deployment is deemed successful today once a model is fully working. This approach intends to provide a better understanding on what is considered "value" in theoretically advanced applications from a practical perspective in more detail instead of technology.

A Fashion E-Commerce Giant

The landscape of fashion e-commerce is rapidly changing at present. A decade earlier, new e-commerce players emerged without having a formal supply chain setup. Most offline retailers invested in the e-commerce boom, while inventory and supply chain strategies were tested on a platform of confusion and chaos. However, these players have to reshape their whole logistics and supply chain design, shifting their sights from the executive to the operational domain. The first challenge faced by the players was to uniform the size of SKU's, as there were over 5-10 units of length measurement due to differing vendors. They already possessed a formal data analytics team and leveraged data science and machine learning to estimate sizes of variations. The next challenge faced by players was the bulk manufacturing of SKU's, an absolute necessity in a highly price-hunting market. Pricing and inventory were optimally decided using deployed methodology on a short-term basis and the customers acquired a better pricing approach.

A Cross-Border E-Commerce Company

Traditional applications in the cross-border e-commerce supply chain often treat the allocation of resources in a passive manner, leading to longer order processing time and dissatisfied customers. Nowadays, consumer confidence is a key factor affecting supply chain performance in CBEC, and this confidence hinges on the processing time of orders. As customers do not see the product in person, they use indicators like delivery speed to evaluate satisfaction while the migration of sales channels from physical to electronic ones intensifies the competition among platforms. Retailers with the best service quality, in particular the shortest processing times, have a dominant impact on the market. Moreover, it is common for CBEC systems to consist of a number of retailers and platforms, leading to a multi-agent scenario.

7.1. Successful Implementations

The success story of Myntra is one of Supply Chain Optimization and Predictive Inventory Management using a complex multi-echelon Inventory Management Model that incorporates a consumer trend model as well as sales predictions at the beginning of every season for each of its products. Myntra is able to execute both the Procurement Mechanism (for Purchase Orders) close to the start of a season as well as a Reorder Mechanism (for Stock Transfer Orders) through Flower Bouquets (combination of similar SKUs) or through combination orders every week. A multi-tier architecture is utilized for processing historic and current data that includes multiple industry norms and best practices, standards, thresholds and exceptions that are Sahar-ally tuned, and optimized datasets to be used in the model. The relevance of data as it pertains to speed, volume, and veracity to the core system is paramount for implementing the model. The Quality Control Plausible, Actionable, and Measurable (PAM) business strategies become critical success factors for business adoption of this initiative. The implementation of the model took place during the pandemic/lockdown period when manual execution was the only option, however, there were no speed/ramp-up constraints on the model's execution. The methodology includes pre-deployment UAT and post-deployment safeguards for the model that need to be followed. The model engages business Ground Rules (GRs) to simplify the procurement process of price shelves. Transparency is another crucial Design Principle of the model. All the Headcount KPIs are aligned to model outputs. The model uses a plethora of optimization techniques on the inputs to comply with the KPIs, and each of these techniques has its own design principles. The Retail Optimization model is leveraged to determine pricing methods, processes are defined to calculate escalation and de-escalation reasons, and pre-requisites are listed for implementation.

7.2. Lessons Learned from Failures

It would be prudent to examine some failures of predictions and the lessons learnt. There were notable incidents of processes being built that were mathematically sound and devoid of human intervention. This resulted in a situation where sale of certain unusual or rare items was not adequately considered leading to a rapid drop in on-hand inventory

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101276

stock during the end of sales of an iPhone model resulting in delays of multiple weeks in fulfillment. This resulted in not only monetary losses as the items went out of stock but also significant erosion of trust in the fulfillment predictions.

The rollout of a new set of ML-model-based frequency predictions increased the ranges of predictive variable values. This resulted in an initial inability of older processes to process the new scale on which calculations had to happen. Also a significant change in hue of data added noise to the predictions. By the time the effects were being analyzed a data tsunami arose resulting in slower executions of logic especially in splicing systems.

Integration of Traditional Bots with the ML-driven inventory management became an issue as the complexity of the ML bots increased. Older code bases had to be reverse engineered to integrate them with ML-driven predictions. In one instance the integration of new pieces of infrastructure allowed for carrying out another 4-5 months of debugging effort instead of helping the process as it was originally meant to.

New stories added systems of up to 1.5 million item-keys to analyze and route inventory fulfillment. Mistaken predictions meant redirecting inventory worth several hundred millions manually across millions of routes and the set of reporting suggestions had to be applied to re-align items being sold to prediction maturity. A suitable reporting mechanism had to be spawned.

The advent of ML drove the experimentation rate exponentially up. It became difficult to keep track of what promotions were being tested. In another instance there were wrong recording still being reflected in the output affecting analysis. Complexity pipelines absorbed months of debugging to settle subtle issues.
Lastly there were the outreach costs such as redesigning ancient row-based reporting mechanisms requiring learning complex logging systems to extract the required data and understanding the disparate summaries to estimate the effects of new predictions for five thousand items per request.

In conclusion it is important to consider the extensions of the project a wider view of what constituents a success case must be taken. Improvement of the modelling infrastructure to abstract and encapsulate the maths behind robustness of infrastructure choices and/or methods of production rollout might be good areas to pursue together with the wider roll of trust from experiment to experiment in the requests to the inventory input considered as commodity holes in different sorts. Scalability loss on higher layers make it difficult to integrate them.

Equ 3: Lead Time Prediction Using ML.

$$\hat{L}_t = f_{\mathrm{ML}}(O_t, D_t, W_t)$$

- O_t: Order volume
- D_t: Distance
- W_t: Weather/logistics constraints

7.3. Future Trends in Case Studies

Motivated by the rapidly changing environment for retail, many retailers have begun to recognize the necessity of speedy business model adjustments and product mix changes. Furthermore, changing consumer behavior has led retailers to pivot toward a more hybrid approach of online and in-store selling. In response, several digital transformations have been made, as illustrated with Ikea. Several new research endeavors on the above themes could be pursued. Yield manager and inventory scouters could also be further researched as AI-enabled replenishment systems that optimize inbound product allocation for per-product and per-store granularity. The growing presence of social and live commerce could also incite efforts in generating novelty in the fields of media planning and sales prediction. On the other hand, exploratory research is needed on unintended outcomes of data analytics in setting redesigned supply chain processes in tangible product supply chains. The concept of unintended consequences from technology adoption could be applied to the adoption of digital technologies in retail.

Lastly, there is a lively debate on the rise of generative AI. Generative AI has gained notable attention and been widely adopted in content generation and customer service automation. However, how it will impact retail and supply chains remains a question with no definitive answer. It could be hoped that generative AI will lower the barrier for data analytics advancements among retailers, enabling faster, deeper, and wider explorations of insights from big data. Unmitigated exploration is possible, with great data revealing rendered mundane ways of data use, but it is a double-edged sword, as unforeseen ramifications could also emerge on dimensions of data ownership and data reliability.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101276

VIII. ETHICAL CONSIDERATIONS

Prior to deploying machine learning forecasting models in production, businesses must give careful thought to potential ethical issues. Such considerations must be addressed before implementation, and continuous monitoring of the systems is required to mitigate any unfolding consequences. In addition to social aspects, the potential risks of ML technologies must be considered. Model bias and the impact on low-involvement articles in supply chains are some of the questions to be considered before deployment. A mechanism to study the models' performance based on their inputs should also be considered, as well as the potential consequences of such outcomes. Product clustering and actionable item sets that lead to biases towards some articles represent relevant future work in this project. This exploratory work would, however, have to be contrasted with the risks of exposing possible wrongful information.

In terms of potential social implications, privacy is another point to consider, especially with regard to prediction explanations in publicly accessible environments. Misinformation. Moreover, if the methodology is to be evaluated as a service offered wholesale to other retailers, it must be ensured that explaining the model and the reason why a certain prediction was made does not expose critical business information. Proper separation and safeguarding of models are crucial. Since the present model deals with potential stock expansions and adjustments at a store, there is ambiguity about whether dynamically modifying the stock levels of the system enables the delivery of an extensive knowledge structure that optimally explains ML functioning at a reasonable expenditure of time and effort.

Finally, the modelling of different alternative outcomes and behaviours across retailers and items is essential, since randomness in the sales processes may very well lead to biases across branches. Any outlier methodologies would have to reconcile biases with denominators that appropriately parse context, balancing the coverage of explorations and counterfactual hypotheses. Furthermore, the narratives delivered would have to be articulated correctly; else, the wrong conclusions could be drawn. In general, biases introduced or unveiled by the system should be understood so as to avoid manipulating them inadequately.

8.1. Data Privacy Issues

The insights of a company often remain a secret, and subsequently cannot be shared with third parties or be scrutinized directly. This challenge of data privacy is even bigger while dealing on a delicate data such as sales and demand data. The potential risk of prevention of competitors' data usages on common items, 1) Realization of interoperability among federated stores, such as sharing data while preventing competitors' insight acquisition; 2) Realization of federated and decentralized supply chain network management. 3) Improvement of forecast accuracy; 4) Improvement of forecast robustness. Each research goal posed in the previous chapter gives a corresponding data privacy challenge to tackle.

The concept of Federated Learning (FL) is used to maintain data privacy at retailers. Instead of instantiating a cloud to store and process the data, a cyclical procedure is followed. First, some functioning retailers instantiate a model with their local data. Second, the local model parameters are sent to an orchestration server, where global model parameters are computed. Third, global parameter updates are sent to retailers, who compute a new local model based on previous parameters and their local data. This procedure could keep the data from being shared, and only model parameter weights are sent. However, there exist several vulnerabilities. There might be the risk of parameter staleness, and malicious workers could disrupt the learning. While individual item series' data cannot be shared, the reasoning on the basis of all data availability is still possible through a collective mechanism. For example, retailers could pool their data to jointly manage demand for items, or one retailer could pool with its suppliers to maintain the fresh food supply chain. This calls for the opportunity to make third parties collaborative.

In this mutual project, the information having been transferred to third parties is managed in an abstracted profile space. A constructive mechanism consisting of a data encoder is provided to control which information such as item supply operation input and prediction output is aggregated for reducing the risk of preventing competitors by duplicating items. Data poolers are enabled to be employed to have more than two edge nodes participating in the inference. A joint deduplication approach is introduced for a new scope of information sharing among multiple edge nodes.

8.2. Bias in Machine Learning Models

In machine learning (ML), the word bias describes situations in which the training sets are not representative for the test data, thus producing unreliable models. In terms of covariate distribution, data is biased when $Pt(X) \neq Ps(X)$, which is called covariate shift, sampling bias, or sample selection bias. In this case, the class distribution in the training set is different from that in the test set, but the two datasets are still drawn from the same conditional distributions. An example application of this bias would be the identification of fraudulent customers on industrial electricity grids.
Normally, companies apply a data-analysis approach to detect customers with irregular power usage based on their previous usage patterns and thus suspect that they may incur electricity stealing. A set of training customers where this

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101276

information is known is given to model the suspicious behavior. However, the company was interested only in the set of customers where an analysis has been planned, in which customers with irregular usage may be present, and thus it is expected that only a small subset of the whole is suspicious. Therefore, the training set, used to create the predictor, only contains regular usage samples, while the test set contains a mixture of both. As a consequence, the learned model is not sufficient for the detection and predicts almost all customers as possible non-suspicious ones. Also, it can happen that the actual desired output (in a supervised-learning scenario) is actually missing for the whole test set, either because of its policy towards transparency or simply because of its novelty. Still, the need for prediction arises. In this situation, the test data shows the same distribution as the actual desired output r=X, but the output variable is completely missing, i.e. Pr(Y|X)=1 for all X while Ps(Y|X)=0. In this case (Y is causally independent from X), the shifts to the distributions are much more severe, and since the random variable Y is no longer part of the underlying generative process, the same estimates and methodologies cannot be used to create a model for the output.

8.3. Regulatory Compliance

Retailers experience intense competition and a growing number of channels due to the shifting landscape of e-commerce. In the contemporary retail environment, achieving sustainability and high customer satisfaction depends on efficiently managing omnichannel distribution networks. One of the most pressing issues for omnichannel retailers is managing their inventory and efficiently configuring their distribution policy. Therefore, predicting inventory requirements accurately over an extended time horizon is of great significance in solving these problems. Following this framework, a LSTM-FHGP-based model combined with an integrated index-based optimization algorithm is proposed. Modeling the estimation and the prediction approaches separately is employed to reach the requirement of non-stationary behavior of time series. Several forecast horizons are taken into consideration for the time-series prediction step. Several numerical studies based on three months of data are conducted to evaluate the effectiveness of the proposed model. Results demonstrate that the proposed LST-FHGP model consistently outperforms other benchmark models and can be successfully applied for demand modeling. Additionally, a sensitivity analysis is performed to provide some valuable insights into the relationships among key model parameters and apex prediction errors.

The recent advent of many sports-action cameras has considerably increased the role of visual analysis of videos. Even though interested parties traditionally rely on human analysis of camera streams to find admissible plays of a given real-world event, the advent of fast algorithms for detecting defective video-content has opened the possibility of creating automatic video-analysis systems (VAS). These automatic systems can alert the user about un-satisfactory sequences. Such video analysis systems can be coupled with digitized databases for automatic content archival and retrieval based on abstraction models. Some of the main technical problems of building a VAS comprise. Please provide a browse element that renders the 'target' item fully using a suitable editor including text, video, audio, image and other elements as necessary. The rendering element should provide a play-button able to start and stop viewing of the video and/or audio in case these time-based data are presented. Preferably the viewing should be presented inside an IFRAME element able to pop-up. Please provide a scoring function able to assign a score to a given video-unit. Exploitability of the scoring function to render admissible items using an editor with slider(s) over the score's space.

IX. FUTURE DIRECTIONS

In recent years, the retail industry confronts pressure from fast-evolving customer needs. By leveraging Artificial Intelligence (AI), retailers can adapt faster to changing customer behavior. For instance, numerous AI-powered tools have been developed and adopted to enhance retailers' decision-making in managing inventory and supply chains. However, several areas remain unexplored, which provides excellent opportunities for future research.

Retailers face unseen events that drastically alter demand patterns, such as the COVID-19 pandemic. Current demand forecasting tools often assume constant demand signals over the prediction horizon and rely on pre-collected data to generate demands to be fulfilled. When faced with unforeseen supply chain disruptions, such as a sudden explosive change of demand in an asset class due to spillovers from an investment strategy or policy news, demand forecasting would fail to provide accurate information. Future research could explore event-driven forecasting architectures that incorporate both demand and supply signals such that the tool would decide what demand to predict based on the current situation of demand and supply situation state variables.

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 12, December 2021
DOI: 10.17148/IJARCCE.2021.101276

Fig 5: Future Directions of inventory management in retail using machine learning.

Despite the rapid procurement speed and low operating cost, robust supply allocation remains an unsolved problem in dedicating the resources to assets. Most of the optimization models assume that suppliers are deterministic, so substantial effort could be invested in developing optimization approaches such that prediction and prediction capex are also taken into consideration. New industrial events could also trigger new types of vehicles, which differ from bandwidth, speed, supply capacity, and operating cost. Next-understood optimization frameworks would help to redesign the networks and incorporate suppliers larger than the bandwidth of the current supply vessels.

9.1. Emerging Technologies in Retail

Economies worldwide are experiencing multiple waves of digital transformation powered by the relentless attempt to outperform competitors in a hyper-competitive environment. Retailers especially have shown great interest towards reaping the benefits prompted by recent advancements in artificial intelligence (AI), machine learning (ML), big data, and Internet of Things (IoT). Beyond pandemic-induced uncertainty challenges, a professor suggests AI–ML–IoT-based techniques can help retailers better forecast customers' shopping behavior and improve inventory management and restocking efficiency in a sales channel and supply chain context. While usage of AI–ML–IoT has been rapidly and widely discussed in academia and retail practice, it remains unclear how retailers can leverage these functions to enhance operational performance. Digital maturity, or the ability of a firm to utilize and exploit technology (business, managerial, and technical), is similarly important for firms to reap the greatest benefits from using AITs for improved performance.

Digitalization has great potential to improve business resilience and business processes. Implementation and incorporation of new digital technologies such as big data analytics (BDA), artificial intelligence (AI), and machine learning (ML) could promote and facilitate proactive systematic recovery mechanisms across supply chains including risk prevention, mitigation, preparation, response, and recovery. In addition to restoration of disrupted supply chain nodes or processes, proactivity addresses the adoption of preemptive observational measures, detection measures, and contingency measures, which could minimize the effects of or prevent disruptions. Implementation of technologies including ML, multi-criteria decision-making (MCDM) methods, Internet of Things (IoT), simulation, and social media analytics has shown tremendous potential to improve the proactivity, effectiveness, and timeliness of governmental logistic-undertaken risk mitigation measures.

9.2. Scalability of Predictive Models

As one of the most important pillars of the digital transformation in retail, inventory management requires a revamp of the digital infrastructure that allows predictive insights to be generated at scale. The last decade has seen significant advances in all aspects of predictive inventory management. This is mainly caused by the advance of machine learning technologies and the huge digital transformation of retailer operations that generated large amounts of diverse and rich data at every instant of time. Finding patterns and relationships in this data provides better understanding of how the retail business is affected by earlier decisions making and allows better predictions of future events and their consequences. How retailers, with their very powerful data and computer infrastructures, can manage the potholes on the way of building predictive models at scale? This infrastructure must allow several predictive models of any complexity and runtime to be developed around the globe and deployed rapidly. It must provide automated feedback on the performance of each predictive model of any complexity and nature on all future unseen data. Models with insufficient performance must be automatically remediated and re-deployed using automated pipelines allowing for human supervision. With the premise that "learning from failure is the best learning" it seems natural to create instant alerting of and analytics in conjunction with alerts and proper remediation to all very rare and evolving wrong predictions.

A predictive inventory management model monitors the performance of hundreds of existing pool months for each item-store combination by outputting per store the projected sales, the orders and the order limits that must be held to achieve these goals. Every few hours the historic and near future data from the operational systems allows tracking these key performance indicators on every store. Each day and every 16 hours ahead an ML model reads the same data and projects the subsequent number of basket transactions per store over a 16-hour frame for each inventory planning algorithm.

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 12, December 2021
DOI: 10.17148/IJARCCE.2021.101276

For each of hundreds of models, forecast error is watched closely and complemented by thresholded annotation about how hard it was to generate. Every executed prediction is joined with its original context. This data reveals what was the whole network status under which the model generated that prediction, the empirical prediction error together with its predicted central value and uncertainty and subsequently feeds the pivot tables providing critical business insights into how the data changed and what patterns the model fails to understand.

9.3. Long-term Sustainability of Solutions

Firstly, the built solution is flexible and able to sustain under a different product category. It is the common scenario to have a multi-product category for a retailer when the demand can be forecasted separately. For retailers having data for other product categories currently utilizing a simpler method and willing to make a switch, it could be done easily. Current system would be switched automatically at the runtime without needing additional coding. This solution makes it possible for the retail chain to have an automated process that helps in adjusting their inventory against the forecast. The architecture of the solution makes it possible to preserve these functionalities without needing interference. This would save the maximum manpower after implementation of the system.

Secondly, the current solutions are receiving real time updates regarding the stock and sales data. The built API with the user interface helps maximum end users and the overall architecture with storage of model input and computation output allows a better tracking of model performance over time and generating predictions. This would help to assess the model performance over time and enhance the current practice, if better improved models or input feature generation techniques are discovered later on. Moreover, this solution helps in the initial heavy lifting for the other retailers who plan on moving into machine learning algorithms based predictive analytic. It is designed in a manner which allows easy dynamic updating of procedures, another separate analyst can build up a similar one easily with their data just by changing a few flags of the models or changing the location of maintaining the source data or data engineering procedures.

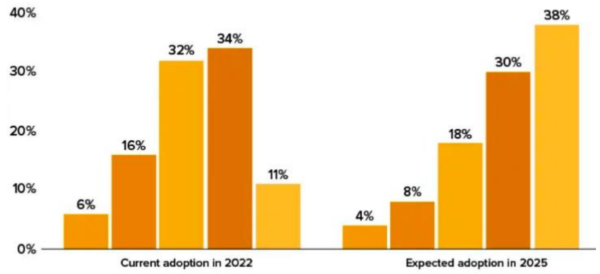

Fig 6: predictive inventory management in retail using machine learning.

Lastly, there are constant updates on relevant data science and machine learning literature; this dynamic changing environment allows better understanding of the models and procedures. Better refinement of current implemented models could be discovered later on with more extensive tuning and testing on evaluation showing improvements. Improvements for the current process architecture handled by other forms of data supporting would be feasible. For such changes, the use with Docker in the cloud computing environment enables a better interchangeability of the codes.

X. CONCLUSION

Based on the results of the projects carried out in this thesis, several conclusions can be drawn. The latest development is more accurate than the previous work of this topic, since it is able to incorporate more input variables, leading to more robust and generalized predictions regarding future sales. With the help of the digital infrastructure and the developed prediction models, future sales of a defined time horizon of new items can be computed such as retail forecasting. Future sales of a defined time horizon can be predicted at an article-store-week level regarding the complete article store set. Then, the lattice search and advising of actions to be performed on the inventory of articles in the upcoming week can be implemented. These methodologies are capable of advising accurate replenishment quantities of articles in order to generally maintain stock levels in a given range. By combining such methodologies with a prediction methodology of another discount type, advising other types of price frontend would be also possible. These digital infrastructures, prediction models and advising methodologies provide the top management of the retailers with useful tools in the strategic and tactical decision stages of the inventory management task. They maintain prediction consistency and robustness over time and quickly compute predictions of new items regarding a complete store article set, outperforming the targets set.

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 12, December 2021
DOI: 10.17148/IJARCCE.2021.101276

It is highly feasible integrating the developed prediction models and digital infrastructure in the prediction side of the systems used in the commercial environment. It can bring substantial advantages as forecasting accuracy improvements or a large set of many operational forecasting tasks with less workload. Good follow-up work regarding the field of stock level advising can be the development of methodologies able to integrate methods of forecasting new base demand models on top of the now-built forecasting methodologies. Besides stock-level advising, advising on conveniently performing actions regarding the other input variables can be approached in follow-up work regarding this topic.

REFERENCES

[1] Paleti, S., Singireddy, J., Dodda, A., Burugulla, J. K. R., & Challa, K. (2021). Innovative Financial Technologies: Strengthening Compliance, Secure Transactions, and Intelligent Advisory Systems Through AI-Driven Automation and Scalable Data Architectures. Secure Transactions, and Intelligent Advisory Systems Through AI-Driven Automation and Scalable Data Architectures (December 27, 2021).
[2] Gadi, A. L., Kannan, S., Nanan, B. P., Komaragiri, V. B., & Singireddy, S. (2021). Advanced Computational Technologies in Vehicle Production, Digital Connectivity, and Sustainable Transportation: Innovations in Intelligent Systems, Eco-Friendly Manufacturing, and Financial Optimization. Universal Journal of Finance and Economics, 1(1), 87-100.
[3] Someshwar Mashetty. (2020). Affordable Housing Through Smart Mortgage Financing: Technology, Analytics, And Innovation. International Journal on Recent and Innovation Trends in Computing and Communication, 8(12), 99–110. Retrieved from https://ijritcc.org/index.php/ijritcc/article/view/11581.
[4] Sriram, H. K., ADUSUPALLI, B., & Malempati, M. (2021). Revolutionizing Risk Assessment and Financial Ecosystems with Smart Automation, Secure Digital Solutions, and Advanced Analytical Frameworks.
[5] Chava, K., Chakilam, C., Suura, S. R., & Recharla, M. (2021). Advancing Healthcare Innovation in 2021: Integrating AI, Digital Health Technologies, and Precision Medicine for Improved Patient Outcomes. Global Journal of Medical Case Reports, 1(1), 29-41.
[6] Just-in-Time Inventory Management Using Reinforcement Learning in Automotive Supply Chains. (2021). International Journal of Engineering and Computer Science, 10(12), 25586-25605. https://doi.org/10.18535/ijecs.v10i12.4666
[7] Koppolu, H. K. R. (2021). Leveraging 5G Services for Next-Generation Telecom and Media Innovation. International Journal of Scientific Research and Modern Technology, 89–106. https://doi.org/10.38124/ijsrmt.v1i12.472
[8] Adusupalli, B., Singireddy, S., Sriram, H. K., Kaulwar, P. K., & Malempati, M. (2021). Revolutionizing Risk Assessment and Financial Ecosystems with Smart Automation, Secure Digital Solutions, and Advanced Analytical Frameworks. Universal Journal of Finance and Economics, 1(1), 101-122.
[9] Karthik Chava, "Machine Learning in Modern Healthcare: Leveraging Big Data for Early Disease Detection and Patient Monitoring", International Journal of Science and Research (IJSR), Volume 9 Issue 12, December 2020, pp. 1899-1910,https://www.ijsr.net/getabstract.php?paperid=SR201212164722,DOI: https://www.doi.org/10.21275/SR201212164722
[10] AI-Based Financial Advisory Systems: Revolutionizing Personalized Investment Strategies. (2021). International Journal of Engineering and Computer Science, 10(12). https://doi.org/10.18535/ijecs.v10i12.4655
[11] Cloud Native Architecture for Scalable Fintech Applications with Real Time Payments. (2021). International Journal of Engineering and Computer Science, 10(12), 25501-25515. https://doi.org/10.18535/ijecs.v10i12.4654
[12] Innovations in Spinal Muscular Atrophy: From Gene Therapy to Disease-Modifying Treatments. (2021). International Journal of Engineering and Computer Science, 10(12), 25531-25551. https://doi.org/10.18535/ijecs.v10i12.4659
[13] Pallav Kumar Kaulwar. (2021). From Code to Counsel: Deep Learning and Data Engineering Synergy for Intelligent Tax Strategy Generation. Journal of International Crisis and Risk Communication Research, 1–20. Retrieved from https://jicrcr.com/index.php/jicrcr/article/view/2967
[14] Raviteja Meda. (2021). Machine Learning-Based Color Recommendation Engines for Enhanced Customer Personalization. Journal of International Crisis and Risk Communication Research, 124–140. Retrieved from https://jicrcr.com/index.php/jicrcr/article/view/3018
[15] Nuka, S. T., Annapareddy, V. N., Koppolu, H. K. R., & Kannan, S. (2021). Advancements in Smart Medical and Industrial Devices: Enhancing Efficiency and Connectivity with High-Speed Telecom Networks. Open Journal of Medical Sciences, 1(1), 55-72.
[16] Chava, K., Chakilam, C., Suura, S. R., & Recharla, M. (2021). Advancing Healthcare Innovation in 2021: Integrating AI, Digital Health Technologies, and Precision Medicine for Improved Patient Outcomes. Global Journal of Medical Case Reports, 1(1), 29-41.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101276

[17] Kannan, S., Gadi, A. L., Preethish Nanan, B., & Kommaragiri, V. B. (2021). Advanced Computational Technologies in Vehicle Production, Digital Connectivity, and Sustainable Transportation: Innovations in Intelligent Systems, Eco-Friendly Manufacturing, and Financial Optimization.

[18] Implementing Infrastructure-as-Code for Telecom Networks: Challenges and Best Practices for Scalable Service Orchestration. (2021). International Journal of Engineering and Computer Science, 10(12), 25631-25650. https://doi.org/10.18535/ijecs.v10i12.4671

[19] Srinivasa Rao Challa. (2021). From Data to Decisions: Leveraging Machine Learning and Cloud Computing in Modern Wealth Management. Journal of International Crisis and Risk Communication Research, 102–123. Retrieved from https://jicrcr.com/index.php/jicrcr/article/view/3017

[20] Paleti, S. (2021). Cognitive Core Banking: A Data-Engineered, AI-Infused Architecture for Proactive Risk Compliance Management. AI-Infused Architecture for Proactive Risk Compliance Management (December 21, 2021).

[21] Vamsee Pamisetty. (2020). Optimizing Tax Compliance and Fraud Prevention through Intelligent Systems: The Role of Technology in Public Finance Innovation. International Journal on Recent and Innovation Trends in Computing and Communication, 8(12), 111–127. Retrieved from https://ijritcc.org/index.php/ijritcc/article/view/11582

[22] Venkata Bhardwaj Komaragiri. (2021). Machine Learning Models for Predictive Maintenance and Performance Optimization in Telecom Infrastructure. Journal of International Crisis and Risk Communication Research, 141–167. Retrieved from https://jicrcr.com/index.php/jicrcr/article/view/3019

[23] Transforming Renewable Energy and Educational Technologies Through AI, Machine Learning, Big Data Analytics, and Cloud-Based IT Integrations. (2021). International Journal of Engineering and Computer Science, 10(12), 25572-25585. https://doi.org/10.18535/ijecs.v10i12.4665

[24] Kommaragiri, V. B. (2021). Enhancing Telecom Security Through Big Data Analytics and Cloud-Based Threat Intelligence. Available at SSRN 5240140.

[25] Rao Suura, S. (2021). Personalized Health Care Decisions Powered By Big Data And Generative Artificial Intelligence In Genomic Diagnostics. Journal of Survey in Fisheries Sciences. https://doi.org/10.53555/sfs.v7i3.3558

[26] Data Engineering Architectures for Real-Time Quality Monitoring in Paint Production Lines. (2020). International Journal of Engineering and Computer Science, 9(12), 25289-25303. https://doi.org/10.18535/ijecs.v9i12.4587

[27] Mandala, V. (2018). From Reactive to Proactive: Employing AI and ML in Automotive Brakes and Parking Systems to Enhance Road Safety. International Journal of Science and Research (IJSR), 7(11), 1992-1996.

