

Name: Sahithi Dodda
Person Number: 50441731
UB Email: sahithid@buffalo.edu

Assignment 7

EAS 504: Applications of Data Science - Industrial Overview - Spring 2023

Lecture by Anurag Bhardwaj – Data Science in Retail

Q1): Describe the market sector or sub-space covered in this lecture.

Ans: This lecture discusses the market sector of Data Science in Retail Industry. Data science is important in the retail industry as it enables retailers to make data-driven choices and enhance their operations, customer experience, and revenues. Increasing demand for increasing amount of data, particularly in Retail technology industry results in huge need for data science for enabling profits through data driven decisions. There are many different types of discounts for customers to profit from, such as location-based bargains. For example, deals in Western New York differ from those in San Francisco and Manhattan. Data science is used to predict future trends and customer behavior, which enables merchants to make educated decisions about which items to carry and when to conduct discounts. Also, In the retail industry, data science is utilized to detect and prevent fraud. Retailers may utilize data to spot unusual patterns of activity, such as fraud transactions , and take preventive and proactive steps to safeguard their consumers and their business. Coca-Cola analyzes data on lead times, shipping costs, and supplier performance to manage its supply chain. Coca-Cola might utilize supplier performance data to detect possible obstacles to its supply chain and takes necessary steps to minimize them.

Q2): What data science related skills and technologies are commonly used in this sector?

Ans: As the Data is continuously growing every day in a very rapid rate, need for developing new models and new machine learning techniques are very much essential and also the need for data science is well observed for increasing profits by retail technology .Understanding consumer behavior, recognizing patterns, and making data-driven decisions all need data analysis and visualization abilities . Machine learning is an effective technology that allows merchants to create prediction models and make data-driven decisions. Tools like python, R, Tableau, Power BI are essential. Natural language processing (NLP) is a crucial skill for chatbots, voice assistants, and other applications that need natural language interpretation and processing. Machine learning is an important technology in retail for recommendation systems, fraud detection, and other applications. Clustering, Recommender system. Robotic Manufacturing with a lot of sensors is getting very popular in manufacturing of cars and various goods. Data science tools such as computer vision and machine learning may be used to evaluate product photos and detect flaws. This can assist businesses in improving product quality and decreasing the quantity of damaged items supplied to customers. Data and computing-related methods can help retailers manage their inventory more efficiently by predicting demand and identifying the right product

mix. For example, Sephora, a cosmetics business, employs data science to tailor client experiences. Sephora employs machine learning algorithms to generate customized product suggestions based on consumer shopping history and interests. Sephora also employs augmented reality technology to allow customers to digitally test goods, increasing the in-store purchasing experience. Cloud computing like AWS, GCP and Microsoft Azure is becoming increasingly significant in the retail industry because it allows merchants to scale their operations, cut costs, and enhance agility. Predictive analytics can also be used to identify potential issues before they occur, such as supply chain disruptions or inventory shortages.

Q3): How are data and computing related methods used in typical workflows in this sector? Illustrate with an example.

Ans: Data and computing related methods are used extensively in typical workflows in the retail technology sector. Retailers use data and computing methods to manage their inventory. This includes analyzing sales data to identify trends and patterns in customer behavior, forecasting demand, and optimizing inventory levels to reduce waste and ensure that products are in stock when customers want to purchase them. For example, Amazon's recommendation engine employs machine learning algorithms to customize product suggestions based on user different locations, purchase history(how often they buy and what they buy) and browsing activity. This enables Amazon to give customers with a more customized purchasing experience, which may lead to higher sales and customer loyalty. Amazon also may experiment with various algorithms or suggestion tactics to see which ones result in the best conversion rates. Machine learning techniques, such as collaborative filtering and content-based filtering, are used by recommendation systems to assess consumer data and produce product suggestions. These algorithms are trained using past customer data to detect patterns and trends in consumer behavior. Warehouse planning, Warehouse operations and inventory management falls under warehousing. Data and computing-related methods can help retailers manage their inventory more efficiently by predicting demand and identifying the right product mix. As the Data is continuously growing every day in a very rapid rate , need for developing new models and new machine learning techniques are very much essential and also the need for data science is well observed for increasing profits by retail technology .

Q4): What are the data science related challenges one might encounter in this domain?

Ans: The retail industry encounters various challenges related to data science, including demand prediction, personalization, recommendation systems, and inventory management. Accurately predicting product demand is a significant challenge for retailers, who require reliable predictive analytics models that consider factors such as trends, seasonality, and consumer behavior. For instance, Walmart had to adjust its predictive models during the COVID-19 pandemic to predict the demand for essential items like toilet paper and cleaning supplies accurately. Retailers must also personalize their shopping experiences to remain competitive, which involves understanding customer preferences and recommending products accordingly. Amazon has achieved this by using its recommendation engine to suggest products to customers based on their purchase history, leading to increased sales and customer satisfaction. Additionally, retailers need to manage their inventory efficiently to avoid overstocking or understocking products. Overstocking

results in additional inventory carrying costs, while understocking leads to lost sales and dissatisfied customers. For example, Target faced a significant challenge in managing its inventory during the 2013 holiday season. Due to a data breach, Target's reputation suffered, and customers were hesitant to shop in stores. As a result, Target had to rely heavily on its online channel, but the surge in demand caused inventory shortages and missed sales opportunities. Other data science challenges faced by retailers include working with enormous datasets, ensuring data quality and integration, optimizing the supply chain, dealing with millions of customers' preferences, and detecting fraud. To satisfy consumer expectations, retailers must develop advanced machine learning algorithms capable of analyzing client data and producing customized suggestions and experiences. These algorithms must be frequently modified to remain effective in responding to changes in customer behavior and preferences, or risk losing their market, it's that simple.

Q5): What do you find interesting about the nature of data science opportunities in this domain?

Ans: I find the potential of data science in the retail industry domain quite intriguing. One of the most interesting aspects of data science opportunities in the retail industry is the wide range of areas that it covers, from demand forecasting to inventory management, personalization and recommendation systems, supply chain optimization, fraud detection, and many more. With the increasing amount of data generated by retailers and consumers, data science provides a valuable tool for retailers to analyze and interpret data, gain insights into customer behavior, and optimize their operations to increase profits. Another aspect that makes data science in the retail industry fascinating is the scope of the industry itself, with a massive number of companies and subcategories that are constantly growing and evolving. This provides a wide range of opportunities for data scientists to specialize in different areas of the industry, such as warehouse planning and inventory management, pricing strategies, or customer acquisition. Machine learning algorithms may detect individual tastes and purchase habits by evaluating client/customer data, which can then be utilized to give customised discounts and promotions. This can assist merchants in developing closer ties with their customers and increasing customer loyalty. These purchase habits, trends and tastes will be constantly changing as days pass on. So, there is always more demand for Data Scientists in the future, presenting exciting and rewarding career prospects.

(i) Describe some of the data science problems relevant to Manufacturing and Warehousing in the Retail Product Lifecycle. (10 pts of the 80 C+R points in the rubric))

Ans: There are several data science problems relevant to manufacturing and warehousing in the retail product lifecycle. In manufacturing, one challenge is ensuring compliance with specifications and quality assurance. Machine learning models can be used to detect faulty chips, to find possibilities for process improvement and optimize the manufacturing process by evaluating data from sensors, machine logs, and other sources, but the trade-off between precision and false positives and negatives must be carefully balanced. When producing computer chips, If a model indicates that every chip is defective, a human inspector should check

it manually. If the model is overly aggressive, it may classify more chips as faulty than necessary, leading to a decrease in yield due to false positives. False negatives, on the other hand, occur when the model fails to identify faulty chips. It is impossible to optimize all aspects of a machine learning problem, so the focus should be in this case, on minimizing false negatives and then checking on reducing false positives because we cannot have false negatives for sure because where a faulty chip is not detected, it can be particularly damaging and should be prioritized for correction. For example, KLA-Tencor uses computer vision and image processing to inspect chips, but there may be limitations with techniques such as Hough transform for complex images. In warehousing, the primary challenge is determining where to build warehouses and what products to stock in them. For instance, in summer in New York, fishing equipment is in high demand, so inventory management must be done ahead of time to route these items to warehouses. This requires sophisticated algorithms for forecasting and supply chain optimization. Data science can be used to analyze demand in different regions and predict which products will sell in which locations. Additionally, space constraints can be a major problem, and data science can be used to identify the most efficient use of available space and to optimize the layout of the warehouse for maximum efficiency.

(ii) Describe some of the data science problems and techniques that would be useful in Inventory Management and Pricing Optimization. (10 pts of the 80 C+R points in the rubric)

Ans: Data science is crucial in the retail industry for inventory management and pricing optimization. To understand commodity demand trends, historical data on comparable products is gathered. Machine learning algorithms are used to evaluate patterns and remove seasonality from data, allowing for a more accurate assessment of true demand. To forecast demand for certain commodities, LSTM is used to normalized data and detrended data. When selecting the best pricing approach for small enterprises, consider the costs of manufacturing and shipping. Retailers confront additional problems when it comes to generating money and decreasing inventory levels through omni-channel pricing. Inventory management is especially difficult in the fashion and garment industries due to high churn rates and continuously shifting trends. Some techniques that can be useful for inventory management and pricing optimization include demand forecasting, data analysis, and machine learning algorithms. Effective demand forecasting necessitates the use of historical sales data as well as market trends, while data analysis can aid in the identification of patterns and correlations that can inform pricing and inventory decisions. LSTM and other machine learning algorithms can be used to examine massive volumes of data and forecast future demand. Furthermore, when choosing effective pricing strategies, businesses must consider a variety of elements such as production and transportation costs, market trends, and competition. By applying machine learning algorithms and data analysis tools, firms can uncover possibilities to cut expenses, enhance revenue, and optimize pricing tactics. Effective inventory management is also critical, as it helps to ensure that products are available when customers want to buy them, while also reducing waste and minimizing costs.

(iii) Also, answer the following multiple-choice questions: You can list the question number and the letter corresponding to the correct choice as Answer in your report, (2x5 = 10 pts of the 80 C+R points in the rubric)

Ans:

Q1) B

Q2) C

Q3) D

Q4) B

Q5) C