“Use of Analytics to forecast product demand in the retail industry: *A summary report.*”

Naga Chaitanya Paladugu

AA-5000-26: Foundation of Analytics

Saint Louis University

Introduction

Demand forecasting is the process of predicting the future shopping list for a store. It helps retail companies guess what customers will want to buy next. They predict this by feeding customers' historical sales data, market trends, customer behavior on online websites, seasons, and more data variables into machine learning algorithms to make these predictions more accurate. By doing this, stores can minimize stockouts and overstock situations, resulting in better inventory management and ensuring that products are available for customers when they want them. This results in a win-win situation for both the retailers and customers.

I am always curious about this part of retail analytics whenever I shop. Particularly, I always wonder how certain retail chains, such as 'D-Mart' in India, and Walmart and Costco in the USA effectively manage, and replenish such a large inventory with a wide and targeted range of products with fewer instances of "Out of stock" situations and in case of overstock, they manage to move those items into customer's basket with better deals and eye-catching promotions.

To fulfill this curiosity, I have selected "Usage of analytics to forecast product demand in the retail industry" as my final project topic.

For information collection, I primarily used SLUth-search Plus, Google Scholar, SpringerOpen, Whitepapers, and blog posts by credible companies in the data science field.

Although I initially struggled to prepare the plot and identify the resources, I managed to identify appropriate resources with time and effort effectively. Why do I feel so?

Here is the reason- I divided my research into two parts, one addressing the academic research aspect of demand forecasting and the other the industry aspect of how it is used in the corporates by taking a case study by a big retail company: **Amazon**.

This summary report can be useful for anyone in a leadership position in a retail organization, who values decision-making driven by data rather than intuition *(McAfee & Brynjolfsson, 2014)*. It can be from a Chief Operations Officer (COO) to an inventory operations manager, anyone working in areas related to inventory management, operations, and strategic planning in the retail sector.

This report gives them an outline of the importance of demand forecasting, how businesses can significantly perform cost-saving and improve profitability with the adoption of demand forecasting, various methods of demand forecasting, going further by discussing the benefits of advanced ML-based demand forecasting for accurate prediction, challenges in the adoption of demand forecasting, and finally practical working of the flow of outline mentioned above by taking a use case of a real-world retail company. Thus, the reader of the report can know where their company stands in adopting demand forecasting, addressing challenges, and improvising to drive their business growth.

Summaries

Source 1:

***Demand forecasting for the modern supply chain | SAP*. (n.d.). SAP.**

https://www.sap.com/products/scm/integrated-business-planning/what-is-supply-chain-planning/demand-forecasting.html

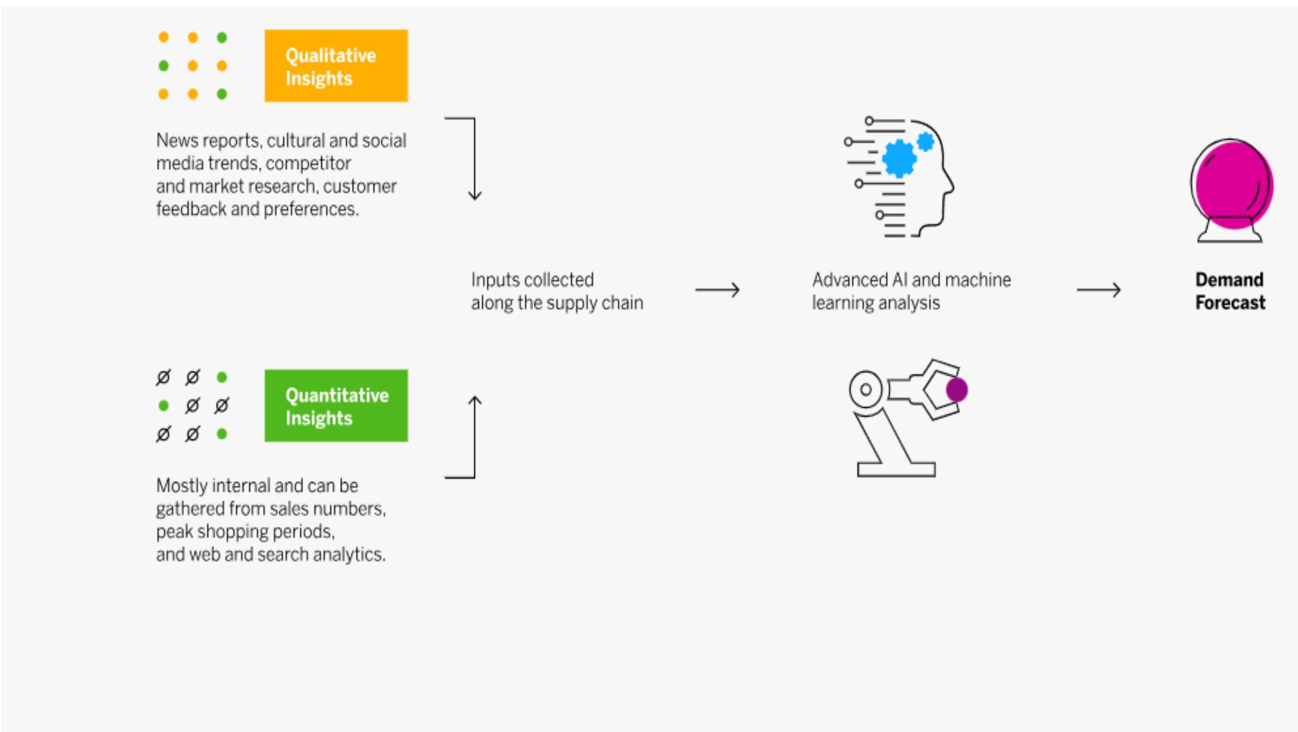
In short, the post on "demand forecasting for the modern supply chain" by a leading software company, SAP, gives a significant overview of the topic in depth. It explains how demand forecasting is used in retail analytics, especially for the core operational processes of modern supply chains. I know, the term "modern supply chain" sounds interesting. By this, the SAP means that it is an evolution of traditional supply chain management with the integration of technologies like data analytics especially predictive technologies, real-time connectivity, and cloud-based solutions to cater the modern business needs. Next, it discusses how the post-pandemic business climate has changed with altered consumer behavior, and how demand forecasting is driving companies in this evolving, and face-paced market. Then, it discusses the various methods of data collection for forecasting models, and types of demand forecasting used, and, lastly, it concludes by discussing various factors influencing demand forecasting in the supply chain.

The two key ideas from the post that I felt more interesting to discuss here are,

Qualitative and Quantitative Forecasting-

We generally tend to have thought of collecting quantitative data which is mostly internal like the data gathered from sales numbers, peak shopping periods, web analytics, etc. for analysis. But with the evolving trends and globalization, global market impacts are felt on businesses like never before (the COVID-19 pandemic is the best and most recent example). So, qualitative data can be collected from external sources like news reports, social media trends, competitor research, and internal sources, like customer feedback, and with the help of modern AI and machine learning technologies, a more accurate forecasting picture can be laid out.

Here is a sample picture from the post that outlines the path flow of accurate demand forecasting,



Demand forecasts are achieved through advanced analysis of qualitative and quantitative supply chain insights.

<u>Short-term and Long-term Forecasting-</u>

Forecasting predictions become more accurate and reliable for future demand when they are segregated by short-term and long-term based on business operations. With short-term forecasting, we can optimize demand for day-to-day, month-to-month, and quarter-to-quarter operations and thus guide the sales and marketing teams for potential demand spikes. A good example of this is a simple Halloween shop or a winter clothing business which are often seasonal rather than a whole year business. Long-term forecasting typically involves looking for a period of usually more than a year, in making an informed decision like business expansions, investments, acquisitions, etc. The best example is the rise of many Ed-tech companies after the COVID-19 outbreak and the rise of investment in data centers in India by many prominent companies from late 2015 with the vast availability of the Internet at cheaper rates resulting in huge amounts of data generation added with a huge user base in India because of its population.

This post by the SAP company is very significant for my research. Not all companies have implemented demand forecasting, and even if they did, then again, the leadership who is going to review this report must get a basic overview of demand forecasting before going into depth. Thus, this post serves as a foundation that introduces them to the world of demand forecasting, especially in the context of modern supply chains.

It starts by stressing the importance of adopting unique business strategies like demand forecasting in this changing trend of consumer behavior and evolving global markets. Alongside this, it also discusses the use of various data collection methods for forecasting demand, both quantitative and qualitative, which is extremely important to make accurate predictions. It says that the type of demand forecasting method that we choose is very important. This is true that in the real world, businesses often tend to be either seasonal or full-time. So, choosing an appropriate type of either short-term or long-term forecasting method for a business is extremely important in making valid and accurate predictions. Finally, it closes off with a discussion on the various factors influencing demand forecasting which is another prior knowledge to be known.

Thus, even an inventory operations manager in a company that has not yet adopted demand forecasting would find this paper extremely important as a foundation.

Source 2:

Mahya Seyedan, & Fereshteh Mafakheri. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. Journal of Big Data, 7(1), 1–22. https://doi.org/10.1186/s40537-020-00329-2

This article introduces us to the various demand forecasting models used in the supply chain. It explores the applications of predictive big data analytics (BDA) in supply chain demand forecasting. For this, the author has studied many previous articles by doing literature reviews and found out that with time, the research on the forecast demand in the supply chain has been on the rise, and the most frequently appeared forecasting models are: Neural networks, Regression, Time series forecasting (ARIMA), Support vector machine, Decision tree.
Also, most importantly, it is noted that the author has explicitly stated that there is a shift from traditional statistical forecasting approaches to intelligent, advanced forecasts driven by BDA techniques like the machine learning models like those mentioned above which can perform the analysis by taking many complex variables into consideration of the modern global supply chains.

Other key findings include a taxonomy of data sources that are to be considered for accurate demand forecasting, unlike a traditional approach where very few predictors were considered which limits the overall prediction. And coming to the aspect of models, the author says that there is no one-fit model to solve all business problems. For example, time series models are appropriate for prediction over equal time intervals, clustering algorithms are appropriate by tailoring predictions to specific customer groups, K-nearest neighbor (KNN) is appropriate especially for pattern recognition and versatile in irregular demand predictions, thus used for forecasting demand for automotive spare parts, and Walmart's supply chain planning, Support vector machines (SVM) are appropriate for complex and multidimensional data and thus adopted by household and personal care supply chain demand forecasting.

This article is significantly important for my research. After giving an overview of demand forecasting, this article has explained the significance of machine learning models which are the main core functionalities that forecast demand. It then slowly emphasizes the need for advanced machine learning models for multi-dimensional data that can more accurately predict demand than traditional models.

It then signifies the importance of both quantitative and qualitative data to accurately forecast demand, which I believe is essential in this current dynamic world. Another important point that I learned from this paper is that there is no one-size-fits-all model for all business cases by going through real-world examples like K-nearest neighbor (KNN) for forecasting demand for automotive spare parts and Walmart's supply chain planning.

Source 3:

Kharfan, M., Chan, V. W. K., & Firdolas Efendigil, T. (2021). A data-driven forecasting approach for newly launched seasonal products by leveraging machine-learning approaches. Annals of Operations Research, 303(1–2), 159. https://doi-org.ezp.slu.edu/10.1007/s10479-020-03666-w

As the multitude of forecasting models were introduced in the previous article, this article is an extension of a use case of those ML models on a company's product forecast.

The overview of the article with the key findings is as follows. The fashion industry in general is dynamic and has a fast-paced nature of fashion trends. The article, therefore, discusses the challenges of demand forecasting in the fashion industry that requires agile supply chain management. Even the author of this article also says that modern ML algorithms more accurately forecast the demand by considering many predictors than traditional forecasting methods. In the case study, a demand for a new fashion product that is going to be launched soon is to be predicted. As this type of case does not involve past historical data, it was not possible to apply a time series model, rather the researcher developed a methodology into a three-step model: clustering, classification, and prediction. Thus, the study evaluated the model's performance, provided some practical results, and suggested strategies for improved decision-making processes in the fashion sector of the retail industry for such a use case.

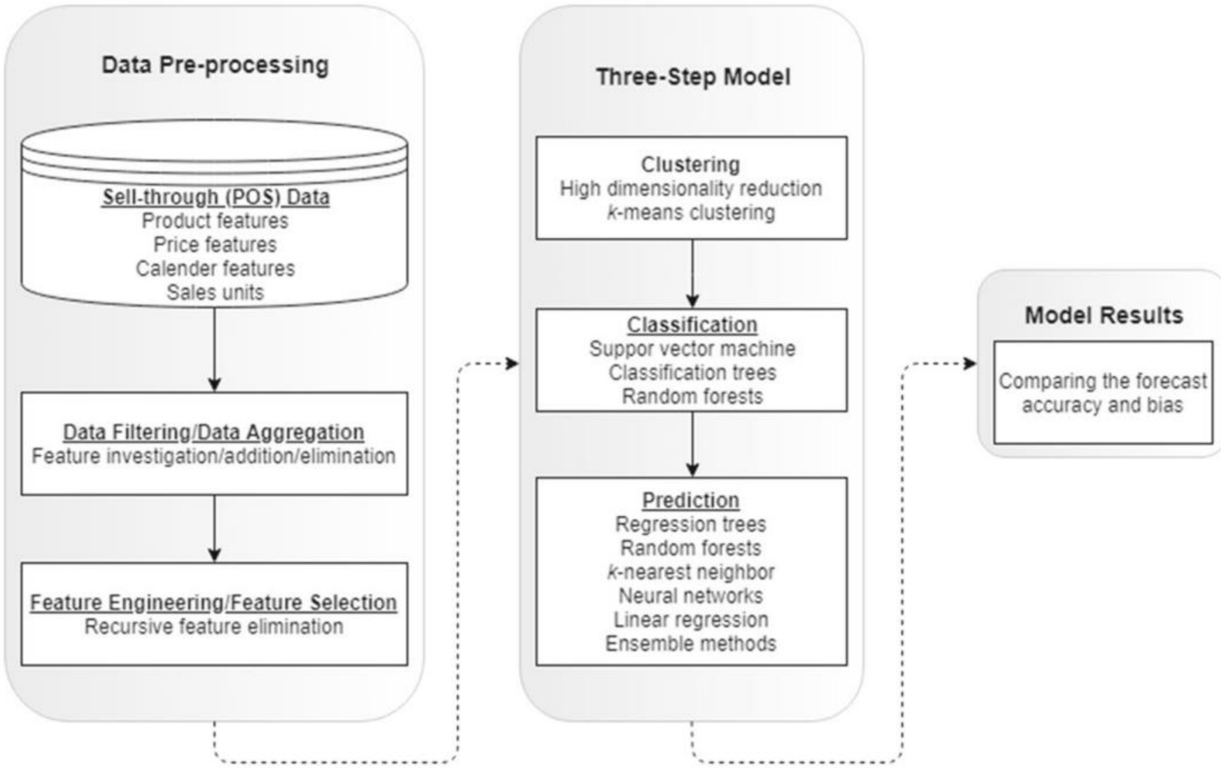


Fig. 1 The proposed methodology

Other key findings are the benefits of cluster analysis for segmentation strategies which groups similar products together by considering factors such as lifecycle length, sales volume, AUR, MSRP, and store count as factors influencing clusters (refer to the article for a complete understanding of the various factors used in detail). This cluster analysis helps us understand the unique qualities of each group. Thus, companies can better predict how a certain group of products are performing and make smart decisions about what to sell and where to sell them. Lastly, performance measurement metrics tools called Weighted Mean Absolute Percentage Error (WMAPE) and Weighted Mean Percentage Error (WMPE) are used to figure out how far the predictors used were accurate enough at prediction or if there is any bias in the forecasting which then means that there might be an overestimating or underestimating happening in there.

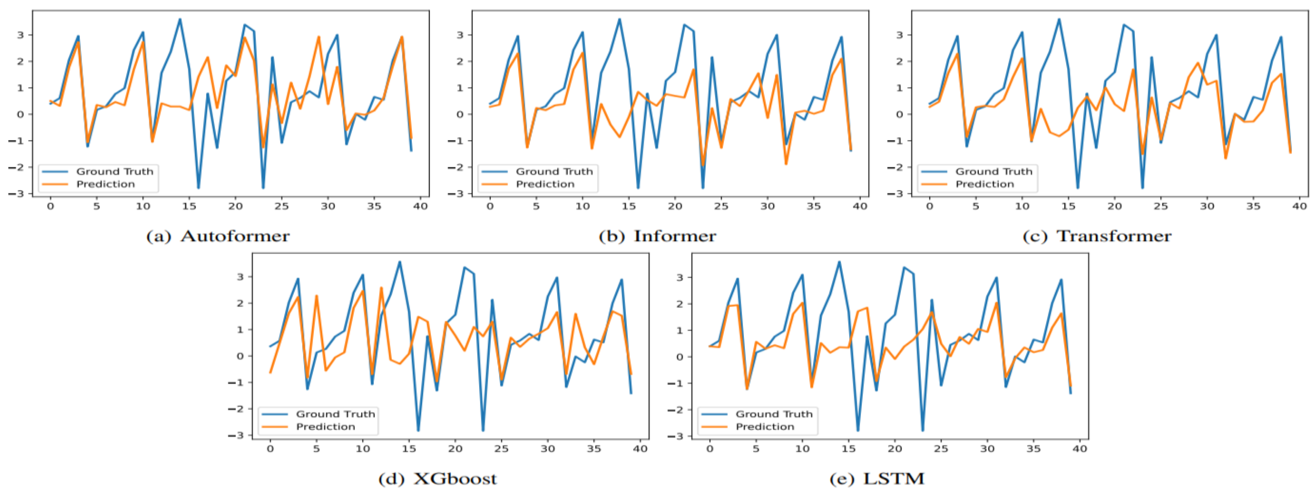
The article holds significant importance for the report as it discusses machine learning models beyond theories to practical applications in the fashion industry. The specific focus on the fashion industry, which is a dynamic and fast-paced environment with changing trends resonates with the challenges faced in diverse business contexts in the real world.

Particularly speaking, trying to forecast demand in the absence of historical data for a newly launched product is a real-world use case that even the inventory operations manager may encounter in the future. I observed the benefits of cluster analysis for segmentation strategies as a practical exploration that improves the accuracy of the demand forecasting models. This time, I learned how the reliability of these models is checked using performance measurements which is essential for us to determine the better model out of all.

Source 4:

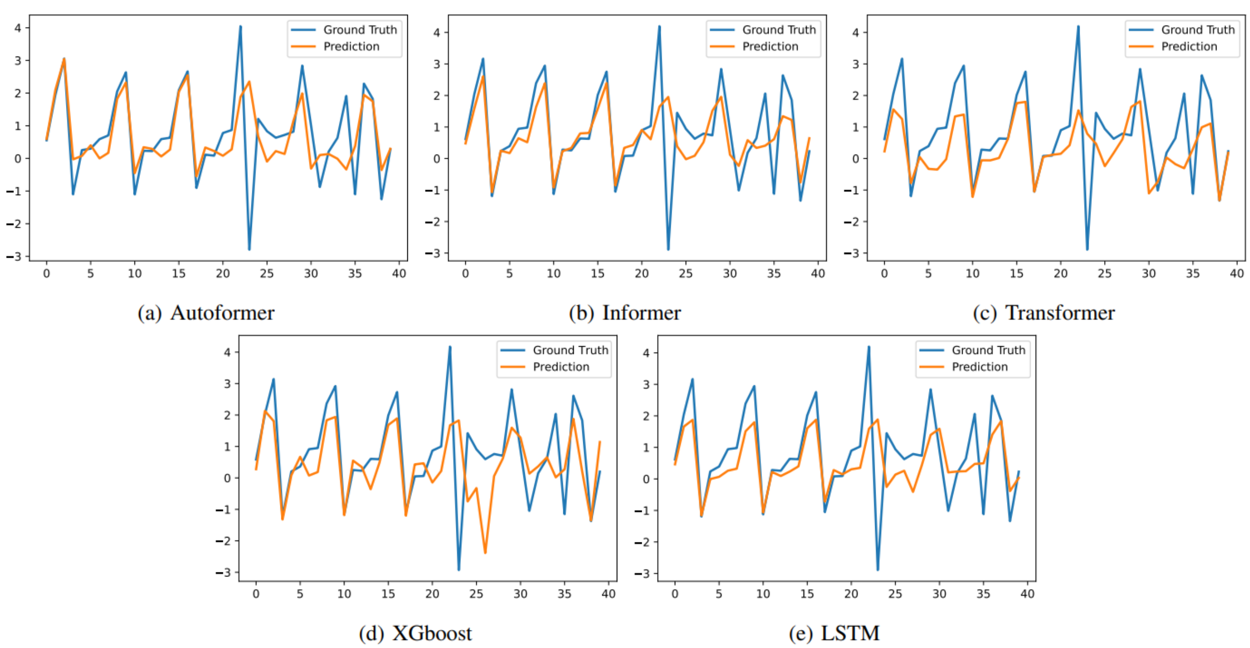
Agbemadon, K. B., Couturier, R., & Laiymani, D. (2023). Overstock Prediction Using Machine Learning in Retail Industry. 2023 3rd International Conference on Computer, Control and Robotics (ICCCR), Computer, Control and Robotics (ICCCR), 2023 3rd International Conference On, 439–444. https://doi-org.ezp.slu.edu/10.1109/ICCCR56747.2023.10194060

This article takes us from regular ML models to advanced ML models including deep learning ones. Overstock prediction in the retail industry is quite a challenge to address and this article emphasizes the significance of accurate demand forecasting by analyzing it for one of the Belgian's largest supermarket companies with transactional data from Colruyt sales focusing on dairy and fish product families. For this prediction, various advanced ML models like XGBoost, LSTM, Autoformer; an attention-based model, Informer, and Transformer are applied, and the model's significance is tested using metrics like Root Mean Square Deviation (RMSE) and Mean Absolute Error (MAE). Of all the above-mentioned models, Autoformer stood out with a superior performance in predicting overstock situations. A suggestion was even made to include more external data sources like weather to increase the accuracy of the forecasting.



(a) Autoformer (b) Informer (c) Transformer (d) XGboost (e) LSTM

Fig. 2: Product category A dataset



(a) Autoformer (b) Informer (c) Transformer (d) XGboost (e) LSTM

Fig. 3: Product category B dataset

From the above picture, we can see that for both the categories of products, Category A: Fish, and Category B: Dairy, the deep learning models like Autoformer and Informer outperformed the other ML models in forecasting the overstock situation. But at the same time, the author has acknowledged the fact that no "existing" work provides a clear answer as to the effectiveness of ML models for the overstock prediction problem which means that this topic is still an open question with a scope for further research in coming to a fixed conclusion.

This article is significantly important for my report. This is a kind of enlightening article for me. It's because until now I was thinking in a bubble that demand forecasting is used to predict the positive consequences, for example, forecasting product demand for this month to serve the customers well. But this article has reminded me that demand forecasting, at the same time, is also used to predict the negative consequences, for example predicting overstock situations that might impact the business profit margins.

Also, I got to know that in real-world scenarios, many models are tested and compared over their effectiveness results before finalizing the best-suited model. Thus, the article explains well the potential of machine learning and especially deep learning models to address overstock issues in demand forecasting of retail products.

Source 5:

***Demand Forecasting - Demand Forecasting*. (n.d.).**

https://docs.aws.amazon.com/whitepapers/latest/demand-forecasting/demand-forecasting.html

This is a whitepaper on various demand forecasting challenges faced by organizations in general and the possible solutions for client companies using Amazon's Web Services (AWS) solutions to address those challenges. Then, it discusses Amazon's in-house developed product namely **"Amazon Forecast"** integrated with the other AWS services for demand forecasting.

As conveyed and in line with the above summary sources (1-4), the paper acknowledges the importance of both quantitative and qualitative demand forecasting in the current trends. Also, it explicitly says that traditional demand forecasting methods may have limitations in accurately predicting the demand as they can naturally factor only limited data into their models. So, it emphasizes the usage of ML models to handle such large volumes of data and the increasing number of demand signals for accurate demand forecasting.

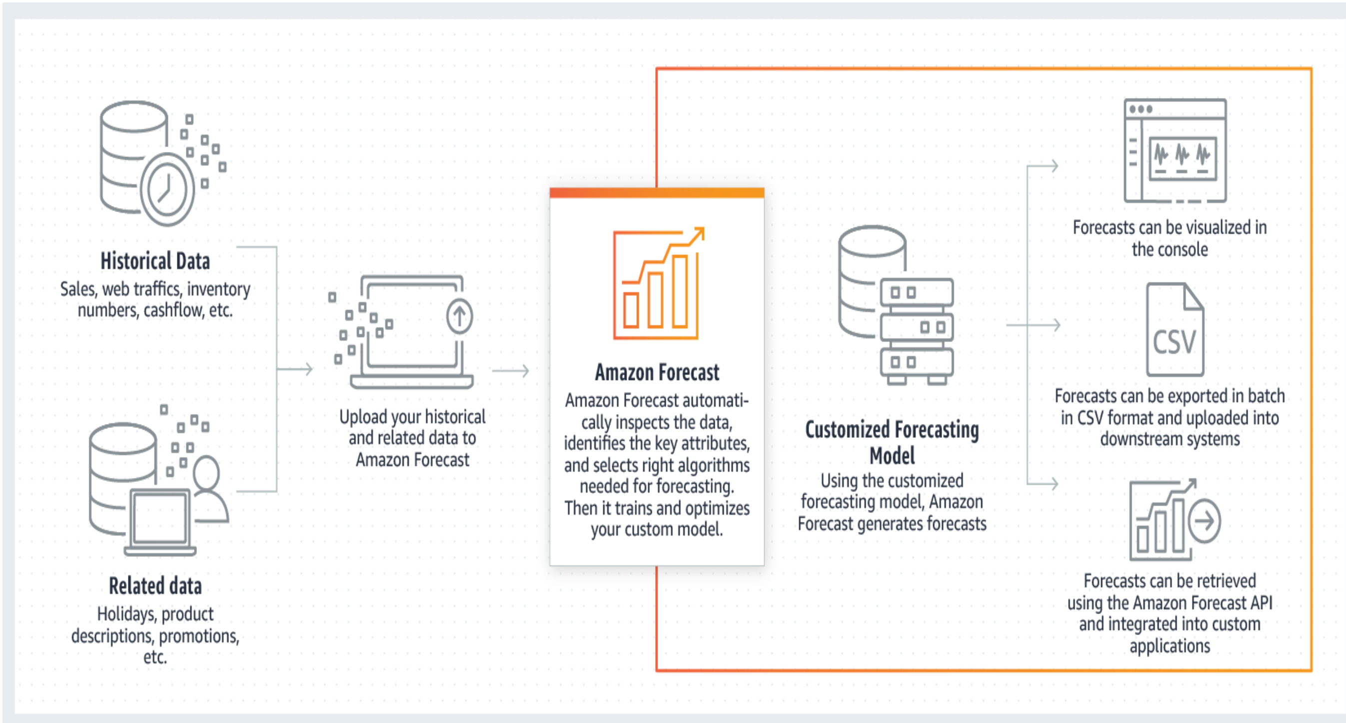
The paper acknowledges that even though companies recognize the value of using data science, they are facing challenges in adopting AI/ML-based demand forecasting. It is because,

- Although they have some expertise, they either they have not enough data to start with or adequate data to accurately forecast the demand.

- Lack of dedicated data science expertise in their companies or a lack of business focus on data science teams by the existing leadership in the company.
- Unable to build effective and competitive models as they are highly resource and time-intensive which the businesses might not be interested in taking a risk for. Even if they did, keeping the model up to date and maintaining it to be reliable and accurate with time is a significant effort that requires a constant monitoring of the model lifecycle and a good MLOps team that again incurs high costs on the company.
- Lastly, interpreting and incorporating the model results into business decision-making is still a challenge.

Then the paper gives hope that these challenges can be effectively addressed using the solutions from AWS, which is kind of irrelevant for this report. So, let's now talk about Amazon's demand forecasting product: **Amazon Forecast**.

The company says that the product uses high-end ML models that can help industries like retail, Consumer Packaged Goods (CPG), utilities, manufacturing, etc., and many more to accurately forecast demand. It employs the following algorithms: Autoregressive Integrated Moving Average (ARIMA), CNN-QR, DeepAR+, exponential smoothing (ETS), Non-parametric time series (NPTS), and Prophet. It is no surprise that Amazon developed such a sophisticated demand forecasting product using advanced ML models with deep learning techniques and neural network architectures as it is a pioneer company with tremendous data sources, zeal for technology adoption, and good financial health. The company claims that with these models the accuracy in forecasting demand has increased by 50% compared to those of traditional methods.



The next most important question that strikes the reader of this report (let's say an inventory strategy manager) is this: **Enough of these plain words! Any specific numbers to back the claims of your summary report that proclaim demand forecasting can help a business?**

Here are the 'numbers' of a couple of clients of Amazon as reported in the paper:

1). **Foxconn**, a widely recognized electronics manufacturer, had to manage staffing by making a better prediction of demand and production needs. This is because overstaffing results in unused worker hours, whereas understaffing means paying overtime pay for workers. So, their two questions to be addressed through demand forecasting are,

1). How many workers are needed to meet the short-term production needs; 1 week ahead?

2). How many workers do they need to meet the long-term production needs; 13 weeks ahead?

With the usage of Amazon Forecast to predict the worker's demand, an initial solution itself has resulted in an 8% improvement in the accuracy of demand forecasting and saved the company $553K for that year.

Here are, specifically, the numbers of a retail company,

2). **More Retail Ltd. (MLR)**, one of India's top grocery retailers had a challenge of predicting how much fresh food they needed to stock their shelves. Both over-stocking and under-stocking were a problem as they resulted in food wastage or poor customer experience. With the usage of Amazon Forecast, their forecasting accuracy has increased from 24% to 76% leading to a fresh-food wastage reduction of up to 30%, improving the in-stock rates by 10%, and thus increasing the gross profit by 25%.

Finally, according to a recent study by **McKinsey** *(Chui et.al., 2019),* companies that have used ML-based demand forecasting techniques saw a 10% improvement in their accuracy, a 5% reduction in inventory costs, and a 2-3% increase in revenue.

(This paper is such interesting that it has summed up the gist of almost all the past articles of the AA- 5000-26 class in the aspect of current trends. This whitepaper posted in 2022 by Amazon seems a kind of counterargument to 'my' perception that I used to give myself whenever I read those posted articles, being very old and so probably outdated. But, after reading this whitepaper I understood that various challenges in the field of data science persist even today mainly in the aspect of big data, Machine Learning, and AI.)

This is an extremely significant paper for the report. “Action speaks louder than words”- this is true because although until now various benefits of demand forecasting have been discussed in the previous articles, this specific paper by Amazon that was published recently is apparent concrete evidence that all these theatrical concepts show positive results in solving real-world business problems even in current times. This paper gives a big boost to the confidence of the operations manager in adopting demand forecasting.

I also have seen advanced machine learning models in action producing impactful and more accurate results. Along with the benefits, I was reminded even through this article that the challenges in the adoption of ML-based demand forecasting persist and need to be addressed to produce productive results.

Synthesis

As I have discussed how the demand is forecasted in the retail industry along with references from various sources, it is time to see the key takeaways to see an overview of the report.

The main key point that was observed is the crucial role of advanced machine learning (ML) models, like those employed by Amazon that take the prediction of demand forecasting to another level. In the current fast-paced world with ever-changing trends along with quantitative data, it is essential to include qualitative data as well to improve the effectiveness of the prediction. Advanced machine learning and deep learning models indeed generate more accurate predictions than the traditional methods. However, not all business cases require the same prediction technique. Rather, it depends on each case. Various model results must be compared using certain performance metrics to identify better models of all.

Alongside the benefits, it is equally important to address the challenges in adopting AI/ML for demand forecasting including the need for adequate relevant data, expertise, model maintenance, and lowering the gap between model results and their incorporation into the business insights. Finally, the potential benefits ranging from improved accuracy in predictions to significant cost reductions are a few of the many benefits of adopting demand forecasting techniques into businesses.

At this point, I confidently believe that the inventory strategy manager would seriously consider demand forecasting as an effective strategy for improving the business. Now, thinking through the mind of that strategic manager, I presume some potential questions that strike the mind are,

1). First things first, where do we stand in the adoption of demand forecasting?

2). As a whole lot of competitors have already started adopting demand forecasting, did we as a company completely adopt "advanced" machine learning models to more accurately forecast the demand to get a competitive advantage over them? If not, how do we adopt them?

3). Among the challenges in the adoption of AL/ML demand forecasting mentioned in the report, which of them is our company particularly facing? What can be done to address them?

4). Are the results of demand forecasting models from the data science teams being effectively considered to alter business decision-making? How many forecasted results from those models that had been implemented for the business have brought a positive change? If the change is significant, then how can we bridge the gap between these AI/ML model results and their incorporation into the business-decision making? If the change is not significant, then are we using the appropriate models to forecast the demand that suits the end goal that we are trying to address? Can better models that suit more aptly for the business case be used by bringing in a few applied AI/ML scientists into our data science teams?

There were quite a handful of possible questions that could strike the manager's mind after reviewing the report. Now I take a couple of these questions (Q2, Q3) and try to give an outline of how they can be addressed.

Firstly, addressing the adoption of advanced ML models for demand forecasting, an assessment of the current usage of demand forecasting models must be done and thus first identify the adoption level. Then a competitor's analysis of their success with advanced ML models must be assessed to identify the gaps in our company's adoption of such models. Then, immediately a roadmap for the incorporation of such models, necessary resources, and any essential training for the data science teams must be given for a positive change to come.

This question is extremely relevant to the inventory operations manager. Why? It is indeed true that the adoption of demand forecasting helps companies save costs and improve profitability. But when even the competitors have adopted it as well, then it is essential to make sure how the demand forecasting can be made better to serve the customers better than their rivals to get a competitive advantage. Thus, when all the companies are using traditional machine learning-based demand forecasting, then how well we take care of customer satisfaction by making sure that the product is available when required and at the same time avoiding any potential overstock situations can be achieved by this advanced ML-based demand forecasting.

Secondly, by addressing the challenges in the adoption of AI/ML demand forecasting, inadequate data problems can be solved by identifying all possible data sources and types of data both quantitative and qualitative including social media, twitter feed, global news, etc. Expertise can be handled by training and upskilling the data science teams or by hiring more knowledgeable employees, and model maintenance can be handled with continuous effort and robust MLOps practices. I believe that effective communication between business and ML teams is essential and an established defined process for seamless ML result integration is recommended.

This specific question is also extremely important for the manager. This is because identifying and addressing challenges in the adoption of ML-based demand forecasting is directly related to operational efficiency. For the inventory operations manager, understanding the challenges being faced by the company in the adoption of demand forecasting is crucial in streamlining their operations and improving forecasting accuracy. For example, a problem identified with the decreased accuracy of the model can be addressed by improving the MLOps team operations.

References

McAfee, A. (2014, October 8). *Big Data: The Management Revolution*. Harvard Business Review. https://hbr.org/2012/10/big-data-the-management-revolution

McKinsey. (2022, May 2). Educational technology in the age of COVID-19 | Israel. Retrieved December 5, 2023, from https://www.mckinsey.com/il/podcast/educational-technology-in-the-age-of-covid

Data center industry in India. (n.d.). Wikipedia. Retrieved December 5, 2023, from https://en.wikipedia.org/wiki/Data_centre_industry_in_India

Chui, M., Henke, N., & Miremadi, M. (2019, March 7). Most of AI's business uses will be in two areas. McKinsey & Company. https://www.mckinsey.com/capabilities/quantumblack/our-insights/most-of-ais-business-uses-will-be-in-two-areas