Customer Acquisition and Retention in E-commerce Using AI & Machine Learning Techniques

Mr. Rasik Patel[1], Dr. Mahammad Idris I. Sandhi[2]
[1]Assistant Professor,
1Sankalchand Patel College of Engineering(MCA) ,
Sankalchand Patel University, Visnagar, Gujart , India.
[2]Associate Professor & Head,
2Sankalchand Patel College of Engineering(MCA) ,
Sankalchand Patel University, Visnagar, Gujart , India.

Abstract-Since 2012, Big Data has been attached more and more importance by companies in this data exploding era. Companies utilize Big Data Analytics to make data driven decisions in many aspects of the business to gaincompetitive advantages. This paper mainly focuses on analyzing how e-commerce companies use big data analytics toboosts their customer acquisition and retention and propose a feasible framework to apply big data analytics based ona case study of Walmart. It also brings the limitation and concerns of using big data in the last section to remind companies to utilize it in the right way. It's worth mentioning that although this paper mainly studies an e-commerce company, Walmart, due to its in-nature advantages of easy access to a ton of in-house customer data, the conclusion, framework and methodology can be generalized to other types of companies as well. Besides, it's been observed that there are still companies struggling to embed big data analytics into their company ecosystem during the pandemic in which companies move their business online to survive. The research is done in this paper hopefully can bring some inspiration to those companies and help them adopt big data technologies.

Keywords: Big Data ecosystem, Customer Acquisition, Customer Retention, Loyalty Program, Customer Experience

1. Introduction

Nowadays, Big Data is becoming a key factor for business to succeed in customer relationship management because it allows the business to gain deeper insights into customer behaviors from the massive amount of data and customize its service and offering to customers. Research shows that companies applying big data analytics outperformed their peers by 5% in productivity and 6% in profitability and 92% of company executives from across 19 countries were satisfied with the results produced by big data analytics and they expected big data to have bigger impacts in their organizations specifically on customer relationship management [1].

However, there are still findings that 70% of the customer data is never used for making improvements and only 30% of organizations use customer experience to help succeed in today's market [2]. There are concerns over big data analytics that customer behavior is changing so fast that the historical data and model cannot be trusted, and the patterns are hard to discern. This can be remedied by tapping high quality data such

as company's own in-house data or third-party data and employing and iterating a better model in an agile setting [3]. During pandemic, a lot of businesses and customers transitioned online, and some marketers are doubling down on precision marketing, since they know customers will not abandon their online shopping habits even after pandemic. Gathering online customer data to perform big data analytics to guide its decision making and business strategy is crucial for a company to gain a leading position in the market during and post pandemic.

This paper studies the theory that how big data is used to boost the customer acquisition and retention, investigates how the theory is implemented by Walmart and analyzes the outcome. Since e-commerce firms are the most adopters of big data, this paper also tries to formulate some patterns from their successful

experience, produce a viable framework and increase other business' awareness of using big data analytics in customer relationship management.

2. Big Data Analytics In Customer Acquisition And Retention

The e-commerce company has easy access to massive data daily. Customers are required to create online accounts with their personal information such as name, age, gender, date of birth, locality, etc. to use the payment and shipment feature. Data like customer shopping and browsing history is collected and captured by the platform all the time from the very beginning when customers research products online, to customer shops and purchases goods till the very end when a customer shares their opinions about the product or brand in social media, comments, rates the goods, and returns for refund. All the customer shopping behavior is recorded and quantified for later analysis easily by partnering with SAS institutes. This data is of paramount value to the company because it represents the customer and is the source of truth to create a more detailed segmentation of customers. Companies then proactively tailor or personalize the service or product for different types of customers to optimize customer shopping experience, increase the customer loyalty and enhance the brand image. Company connects well with the customer because it speaks the same language with customers and understands what customers truly need and how their product can meet those needs. If customers feel they are specifically serviced and appreciated, customer sticks to the product or service. This applies to not only existing customers but also potential customers. For example, marketers analyzing customer information can offer different types of coupons or discounts to different customers. Customers new to the product or service might get bigger discounts for better acquisition. Through personalization, company can understand individual customers' needsand serve them better to increase sales and customer loyalty and as a result, customer acquisition and retention are promised.

Apart from custom offering and personalization, another prominent capability of big data analytics ispredictive analysis. Santana, a car manufacture company, applies data mining and machine learning to generate features from its historical data and utilize those features to predict customer perceptions of the products. Santana can predict the problem even before engaging the customer and ends up with a better customer relationship management [4].

The reason we utilize big data analytics is that the amount of data is too huge for traditional database to store and process and it is ideal when all or most of the data should be analyzed in our case to understand each customer's shopping behavior and the sampling of data is not as effective as a larger dataset [5].

3. Online Walmart Study

Walmart is widely thought to be the largest retailer company in the world. It was founded back in 1945 and has been operating for more than half a century. Itsbusiness model focuses on one-stop shopping convenience by stocking a variety types of products. The last decades have seen the firm starting to embrace e-commerce by acquiring a bunch of online firms including Jet.com, Moosejaw, Shoebuy, etc. to reply to other newer entrants such as Amazon. Considering many customers have demonstrated a shift to digital, Walmart lab was also founded to empower big data analytics to fuel the business online growth. Walmart paid loyalty program is first analyzed to understand howthe firm utilizes big data analytics to acquire and retain the customer. Other applications of predictive analysis and personalization revealed by Walmart are subsequent.

3.1 *Loyalty program*

Loyalty program has been changing over the last decade to adapt to the digital channel. Consumers spend more time on their phones, tablets and other digital channels. And it is more straightforward for merchants to approach consumers to market their products in digital format. The barrier to establish a digital brand is no longer high and there are so many brands that customers are easily attracted by other brands and tendto be less loyal [6]. Brands are striving to

develop a strong relationship with the shopper. In this big data era, digital channel produces a ton of data. Big data analytics is widely utilized to figure out how the brand is routed to the customer for acquisition and design the loyalty program with more useful patterns for retention. For example, customer behavior patterns can help identify customers who are more likely to establish the relationship with the brand. Merchants can precisely market their brands to those customers for better acquisition.

McKinsey indicates that predictive analysis helps better understand different types of customers to calibrate and hone customer loyalty program for different customer segmentation. It allows sellers to design custom products and services in the loyalty programs to drive a particular profitable behavior that customers may portray when shopping and capitalize from that. For example, through BDA, it is found that customer who buys product X also buys Y. Then company can design incentive into the program to drive that specific behavior. The important thing to keep in mind is that loyalty program can become costly so easily that it doesn't always bring value for companies. Company should use BDA to find the gap between a customer's perceived value of the product and the actual cost of providing it to see whether the program is viable[6].

Walmart implemented the loyalty program called Walmart+ with intention to drive high customer acquisition and retention. Customer is charged 98$ annually to enroll in this program and the benefit includes unlimited free delivery on more than 160k items, scan and go shopping feature and fuel discounts. All these benefits aim to make customers' life easier by seamless shopping experience and great one stop shopping convenience. Walmart uses big data analytics to analyze the customer needs and provide either discounts or other custom features to provide an enhanced shopping experience. The firm's cash back fuel reward card builds the customer loyalty as well as retention [7]. This loyalty program has proved to work well and poses a threat to existing business rivals such as Amazon's loyalty program Prime [8].

3.2 Predictive Analysis in offlinebusiness

Walmart employs predictive analysis in its smart forecasting system to help the managers to estimate the demand for over 500 million products across just US stores. This system is utilized to analyze around 100TB of data generated daily over 11000+ stores and 12 online websites [9] globally in order to ensure customers can always buy the desired product. Predictive analysis is also applied in its pharmacy business, Walmart simulates the customer pharmacy usage to identify the number of prescriptions filled at a certain time and to infer the busiest time of a day or days in a month. This information helps its staff scheduling to cover the busiest time to reduce customer waiting time for the prescription to be fulfilled. Additionally, it notes the most selling drugs and stocks them to avoid shortages. When customers are assured, they can get their specific drugs at Walmart without delay and short supply, they hardly shift to other stores. Furthermore, Walmart also uses predictive analysis to anticipate demand at certain hours in its bricks and mortar stores to determine how many staff are needed at the counter. Customer checkout experience is thus greatly improved [10].

3.3 *Predictive Analysis in onlinebusiness*

Walmart utilizes big data to personalize the online shopping experience for its customers based on customer's shopping data. For example, A user searching for a desktop computer online might also needa monitor, Walmart can anticipate this need and recommends the complementary goods to the customer appropriately. This greatly eases the customer shopping journey and increases customer loyalty. Besides, for the bricks and mortar business, Walmart utilizes the same preference data to decide the way to display the merchandise on the store shelves to the customers, what brands to discontinue and what new brands to carry. Therefore, Walmart develops a consistent and delightful shopping experience for both online and offline customers.

Walmart's customer acquisition and retention strategy have helped it improve e-commerce sales in the last few years since it embarked on the online sales journey. In three years, the firm has tripled its online sales. In 2022 , it made $21.9 billion, up from $7.9 billion in 2016 in the US market. Although the 2022 growth rate of the e-

commerce sales is 27%, which was a drop from 40% in 2018, the company expects to experience huge gains in grocery pickup and delivery programs and expects to surpass $28B in 2023 [11].

The new Walmart Grocery App was developed to assist customers in paying for online groceries. This app recorded a rise in 2023 , following Covid-19, surpassing Amazon by 20% [12]. The app downloads proved that Walmart is performing well in terms of customer acquisition and gaining more opportunities for conversion. People have been using the app to make online purchases, particularly during the lockdowns, which contributes a lot to online sales as mentioned above. Walmart would later integrate the app with Walmart App to make it a single app inspired by Amazon adding online groceries to their main app to unify the customer experience based on analysis of customer behavior that customer keeps pulling those two apps together. This also indicates that Walmart makes calculated data driven decision to retain the customer. Additionally, the firm has been hiring more employees to cater to the increased e-commerce orders and avoid unnecessary delays or shortage. With much data collected daily, the retailer can tailor services based on consumer preferences and provide new features, such as purchasing online and picking up in-store, to attract new customers.

With such success in its online business, Walmart wants to extend its best customer experience to its in- store customers as well via the mobile apps. It has been noted that more than 240 million clients go to brick and mortar stores every week. It's just as significant as online shopping. The Walmart app is used by more than 22 million active users who rank the company as one of the best worldwide [13]. The app enhances a good in store shopping experience with features as checking and picking an online order from the nearest Walmart stores,searching for goods in the store and the scan and go feature. All these efforts are to improve customer experience guided by the data behind the scenes.

4. Proposed Architecture

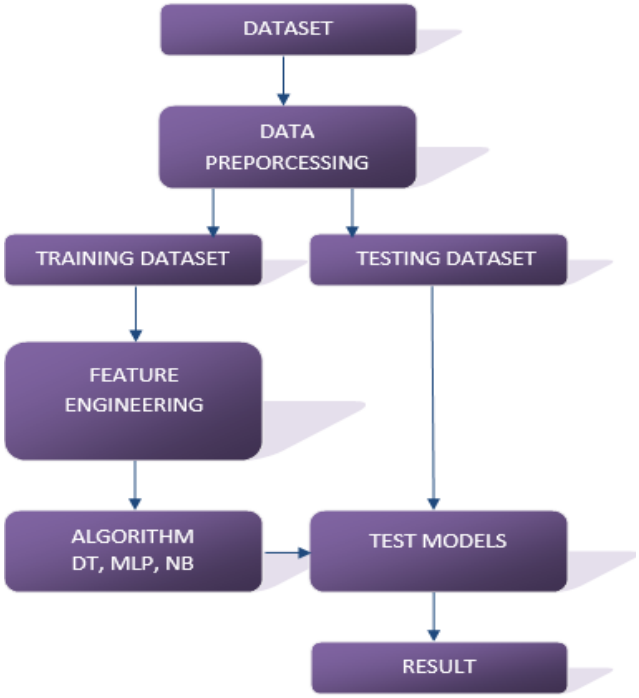


Fig-1 Proposed Architecture

5. Results

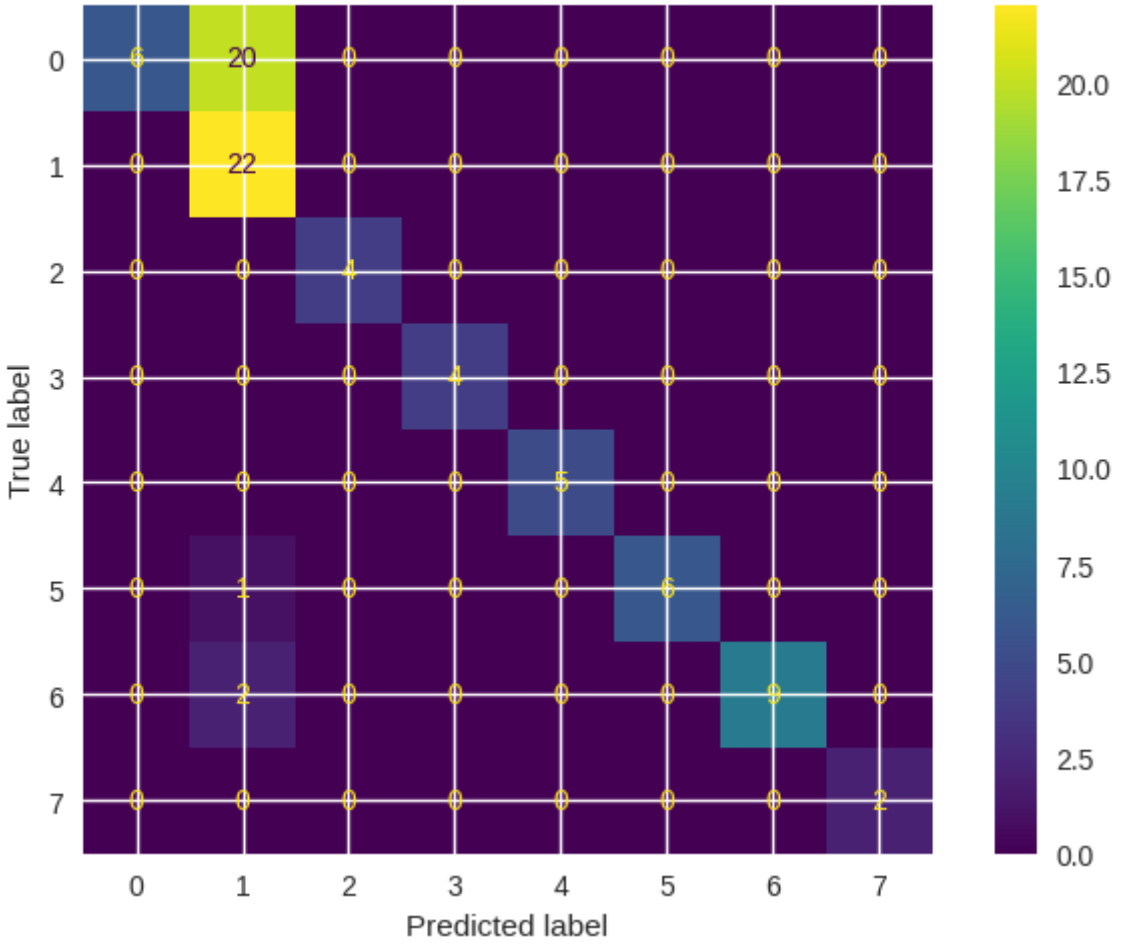
5.1 Decision Tree Implementation Code :

```
from sklearn.tree import DecisionTreeClassifier as DT
DTmodel = DT(max_depth=5,random_state=0)
DTmodel.fit(xtrain,ytrain)
DTmodel.score(xtest,ytest)
Y_predict = DTmodel.predict(xtest)
DTaccuracy = accuracy_score(ytest,Y_predict)
print("Decision Tree Accuracy is: %.2f%%"% (round(DTaccuracy,2) * 100.0))
```

```
Decision Tree Accuracy is: 72.00%
```

5.2 Decision Tree Confusion Matrix :

```
confusion matrix for Decision Tree
<sklearn.metrics._plot.confusion_matrix.ConfusionMatrixDisplay at 0x7b3317
```



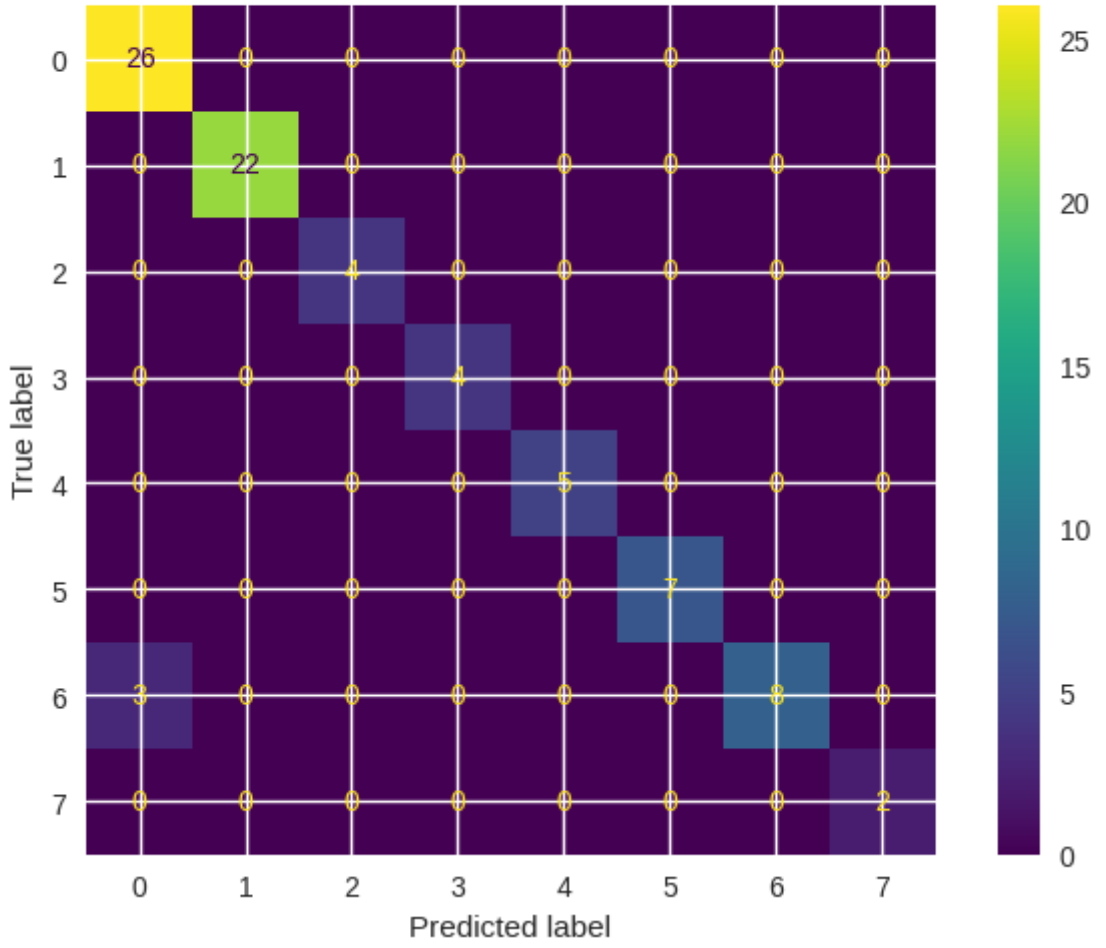
5.3 Naïve Bayes Implementation Code:

```
from sklearn.naive_bayes import GaussianNB as NB
NBmodel = NB()
NBmodel.fit(xtrain,ytrain)
NBmodel.score(xtest,ytest)
Y_predict = NBmodel.predict(xtest)
NBaccuracy = accuracy_score(ytest,Y_predict)
print("Naive Bayes Accuracy is: %.2f%%" % (NBaccuracy * 100.0))
```

```
Naive Bayes Accuracy is: 96.30%
```

5.4 Naïve Bayes Confusion Matrix:

```
confusion matrix for Naive Bayes
<sklearn.metrics._plot.confusion_matrix.ConfusionMatrixDisplay at 0x7b33
```



5.6 Multilayer Perceptron Implementation Code :

```
from sklearn.neural_network import MLPClassifier
MLP_Classifier = MLPClassifier(random_state=1, max_iter=16).fit(xtrain, ytrain)
MLP_Classifier.score(xtest, ytest)
```

```
0.8148148148148148
```

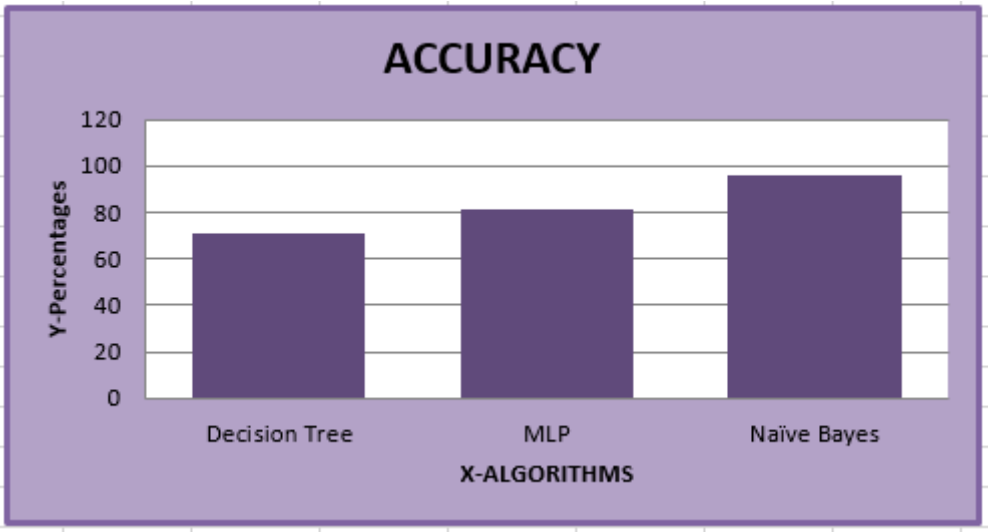


Fig-2 Accuracy Comparison Chart

6 Conclusion

Big data analytics is vital for an e-commerce company to survive in this highly competitive market. To increase customer acquisition and retention, company uses predictive analysis and custom offering to enhance a good customer experience. It also helps the business to gain valuable insights into their customer shopping data and design custom loyalty programs to improve customer loyalty and brand image. Other online companies such as Amazon, Capital One, American Express, and T-Mobile, among others, have discovered the need for big data and are using it to boost customer service and win more customers. Indeed, big data analytics is acquiring and retaining customers in this big data era and companies need more aggressively adopt the big data technology to benefit from their own data to gain competitive edges over competitors.

However, there is still the limitation of big data analytics which needs careful attention when applying in practice. First, data is prone to infringements and breaches. Walmart cleanses and anonymizes the raw data, such as names and emails, before storing and analyzing to protect customer privacy. Second, Big data analytics is not perfect and sometimes requires human judgment to interpret the non-intuitive result because of outdated data or anomaly models. I got dataset from kaggle.com and I have implemented 3 ML algorithms Decision Tree , Naïve bayes and MLP. So I found good accuracy in Naïve Bayes like 95% accuracy to predict the customer retention and acquisition.

References

[1] Ghandour, A. (2021). Big Data Driven E-Commerce Architecture. International Journal of Economics, Commerce & Management (IJECM) [online]. 3. [Viewed 5 July 2021]. Available from: https://www.researchgate.net/publication/277701850 6_Big_Data_Driven_E-Commerce_Architecture

[2] Satish, L. and Yusof, N. (2021). A Review: Big Data Analytics for enhanced Customer Experiences with Crowd Sourcing. Procedia Computer Science [online]. 116, 274-283.

[Viewed 6 July 2021].
Available from: doi: 10.1016/j.procs.2021 .10.058

[3] Bibby, C., Gordon, J., Schuler, G. and Stein, E. (2021). The big reset: Data-driven marketing in the next normal [online]. Mckinsey & Company.[Viewed 8July2021].

[4] Available from: https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/the-big-reset-data-driven-marketing-in-the-next-normal

[5] Santana, J. (2022). How companies will use big data for customer retention [online]. ALTA. [Viewed 8July 2021]. Available from: https://altadigital.com/blog/big-data-customer- retention Kubina, M., Varmus, M. and Kubinova, I. (2021). Use of Big Data for Competitive Advantage of Company. Procedia Economics and Finance[online]. 26, 561-565. [Viewed 10 July 2021]. Available from: doi: 10.1016/S2212- 5671(15)00955-7

[6] Mckinsey & Company. (2023). Customer loyalty:The new generation [online].

[7] Gallagher, B. (2023). Walmart's rise of e-commerce & brand loyalty [online]. ANNEX CLOUD.[Viewed 12 July 2021].

[8] Available from: https://www.annexcloud.com/blog/walmarts-rise- to-ecommerce-brand-loyalty

[9] Springer, J. (2023). Walmart Reveals Paid Loyalty Program [online]. Insight Grocery Business. [Viewed 12 July 2021]. Available from: https://www.winsightgrocerybusiness.com/retailers/walmart-reveals-paid- loyalty-program

[10] Malur, R. "Pillars of Walmart's DemandForecasting." Medium, 1 Aug, 2022 , https://medium.com/walmartglobaltech/pillars-of- walmarts-demand-forecasting f6722de86e1aAccessed 31 Jul, 2021.

[11] Walmart Staff. (2021). 5 Ways Walmart Uses Big Data to Help Customers [online]. Walmart. [Viewed 12 July 2021]. Available from: https://corporate.walmart.com/newsroom/innovatio n/2021 0807/5-ways-walmart-uses-big-data-to-help-customers

[12] Kaziukenas, J. (2023). Walmart's Online Sales to Reach $28 Billion in 2023 [online]. Marketplace Pulse. [Viewed 12 July 2021]. Available from: https://www.marketplacepulse.com/articles/walmar ts-online-sales-to-reach-28-billion-in-2023

[13] Perez, S. (2023). Walmart Grocery app sees record downloads amid COVID-19, surpasses Amazon by 20% [online]. TechCrunch. [Viewed 13 July 2021].Available from:https://techcrunch.com/2023/04/09/walmart-grocery-app-sees-record-downloads amid-covid- 19-surpasses-amazon-by-20

[14] SAS. (N.d.) How Walmart makes data work for its customers [online]. SAS. [Viewed 14 July 2021]. Available from: https://www.sas.com/en_us/insights/articles/analyti cs/how-walmart-makes-data-work-for-its- customers.html

Mckinsey & Company. [Viewed 12 July 2021]. Available from:https://www