International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 05 | May 2022 www.irjet.net p-ISSN: 2395-0072

# Data Science and Machine Learning Approach to Improve Online Grocery Store Sales Performance

**Abhiram Nallamekala(11801535)1, Subbareddy Vanukuri(11801547)2, Dr. Om Prakash Yadav3**

*[1,2]Student, School of Computer Science Engineer*
*[3] Assistant Professor (UID: 26121), School of Computer Science Engineer.*
*Lovely Professional University, Jalandhar - Delhi, Grand Trunk Rd, Phagwara, Punjab 144001.* country

---

**Abstract -** *Data science, artificial intelligence, machine learning, and advanced computing techniques became*

*the catalyst to expand the business and provided a venue to do electronic commerce to reach and interact with customers. There is a need to improve and enhance the sales performance for grocery Store The data science and machine learning approach with handling big data with the transparency of data records using computing techniques to enhance e-commerce is one of the demanding areas of research these days. This paper discusses the data science and machine learning approaches to improve grocery sales performance. This paper can give an idea how can we achieve it. The method can also help in saving the time for customer who visits the store and helps him in tracking the groceries present in the store. It contains the data which is analyzed using data science where it has been designed in a way the data can be collected in faster way. operation can be done quickly.*

***Key Words***: Grocery Retailing Industry, Online Transaction, Data Science, Machine Learning

## 1. INTRODUCTION

The grocery delivery is a huge opportunity for businesses to cash on customer comfort of ordering groceries from home. There's huge scope of machine learning in this space. Alone in India there is Flipkart, Amazon, Jiomart, Swiggy, big basket and others who are betting on grocery service space. Online shopping is becoming increasingly popular for a variety of reasons. There are certainly external factors such as rising fuel prices, difficulty accessing traditional stores and the problems that are often associated with supermarkets and other general stores to contribute to the increase in online shopping interest. Consumers can find out more about the product and its updates from existing users. When someone wants to buy a product, they no longer just ask friends and family because there are so many products reviewed on the web that give feedback to existing users of the product.

Online shopping sites contain a wide variety of high-quality and soft quality products that keep in mind people. The project is a web application which is used to order the daily grocery's we use in our real life. This helps the customers to buy the groceries in fingertips with a easily understandable user interface, any person very new to the website can understand the navigation easily and correctly. The website is made using PHP, HTML5 Stack and using wamp server to simulate the MySQL server and establish the connection between the frontend and backend. The user-friendly interface is very easy to understand by any user who visits the website.

### 1.1 DATA SCIENCE, MACHINE LEARNING

The Data Science (DS), Machine Learning (ML) approaches are indeed required to scientifically develop methods, processes, algorithms with software applications to extract purposeful and timely information from user-provided structured and unstructured data collected from e-commerce websites. The expertise of data mining and big data is also used to analyze business and market trends on a temporal scale. It also unifies data engineering, statistical methods, machine learning algorithms and programming procedures. The strong knowledge and expertise of mathematics, statistics, information sciences, and computer science with artificial intelligence is required and much useful. "Machine learning" is the sub-field of "Computer Science" i.e. it comes under the area of "Artificial Intelligence", which is the scientific study of algorithms developed under stochastic theory which is effective to perform tasks without having explicit program instructions, based on inference and relying on patterns. The mathematical model of sample data i.e. "Training Data" is built via machine learning methods to make decisions on data and predictions are drawn.:

### 1.2 LOCALIZATION & ZIP-CODE LEVEL CUSTOMIZATION

The big shift in food retail is how the local food movement has started to change the assortment landscape. Increasingly more grocery shoppers are returning to their roots, seeking local food, supporting local farmers to reduce the environmental impact of transporting food from long distances. Assortment solutions need to enable local store managers to source, validate and test local products – essentially supporting a curated offer of all that is local – and integrate this with a more efficient central assortment which

takes into account the number of customers who will order online and therefore do not need to choose and purchase the product instore. In short, it requires a blend of the data-driven, machine-assisted recommendations and human approach to truly enable grocery retailers to deliver at a local level. For these reasons, the need for zip-code level tracking of price and assortment is paramount for a highly customized assortment. Retailers can no longer have a central assortment strategy for the entire market.

**Figure 1**: Localization of Store

Data science helps the retailers to find the best places for implementing new stores for selling their products. It uses the decisions of the customer concerning area, for this analysis, there is a large amount of data is required. Such as the customer data available online, market trends in that area, the location of the other nearby shops, etc. Using data science, geospatial analysis, and machine learning techniques, this project aims to provide a solution for this problem and recommending the best neighbourhood for opening the new store.

A quick summary of problems which this grocery store addresses using data science and Machine Learning solutions:

1. maximizing the customer conversion rate through advertisement and targeting customers to right grocery items or stores would generate revenue for owner.

2. Boosting the traffic to a specific store for visibility also increases Instacart revenue.

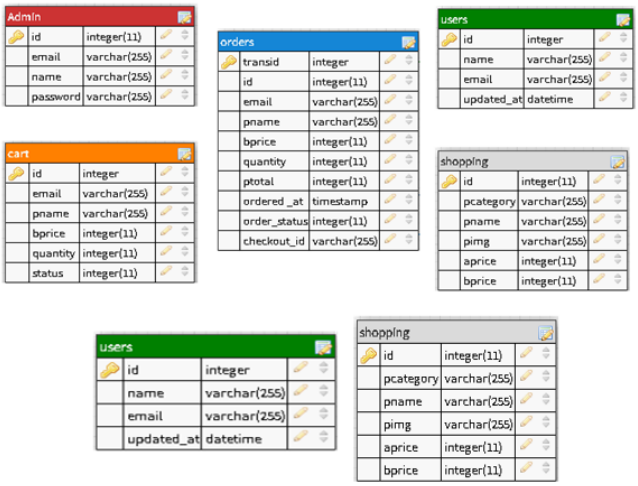
3. It would also need to know how many customers repeat customers are which purchase on regular basis.

4. There is high variance in predicting the time it takes to deliver an order based on various factors like weather, road condition, rush at store, how long will it take for store to pack the order and available delivery agents etc.

5. It is also building its own model for travel time for delivery agent to deliver order to specific customer, as most of orders are repeat orders and there is high possibility.

6. Recommending items dynamically to the user (Which he/she might be interested) would also increase conversion.

## 2. DATABASE TABLE SCHEMA




**Schema -1:** Database Schema

## 3. ALGORITHM APPROACH

Some apparent challenges as shown in below figure. That the online grocery store is facing are

1. Payment security issues in online grocery store
2. Handling logistics and distributions.
3. Lower sales conversation on site.
4. Also, the Warehouse Management.


**Chart -1:** Algorithm Flow chart

- As, above figure no Payment security in online grocery store has always been under the spotlight. Although the payment gateways and processors have outgrown in terms of their subscriber base and average transaction

volume, there is still a deficiency of secure and stable security systems especially in developing markets.

- One of the biggest problems that e-commerce businesses are facing is the efficient handling of logistics and the distributions network. Most logistics companies are not furnished to provide services to modern e-commerce businesses. Therefore, many new e-commerce stores fail to meet the requirements of high efficiency resulting in poor on-time delivery.

- The conversion rate is metric which is calculated as the percentage of the number of users who process the complete purchase on e-commerce website among the total number of visitors of that site." For example, www.daraz.pk, a popular e-commerce website visited by 1000 customers during a particular month that is March to buy a perfume, but only 60 people processed the complete purchase, hence the site conversion rate became 60/1000 equal to 6%.

- Another view of a successful "Sales Conversion" is one in which the user follows the "Call- to-Action" (CTA) and completes the steps that were required to be performed. For example, to get more email newsletter subscribers for a brand, a visitor on a website must have to enter his email address and name. Each successful subscription would be counted as a "conversion" in this case.

- This is one of the most challenging areas for every e-commerce website, especially the new ones in the market, that although receive decent traffic, some thousands of monthly visitors to the website, but still only a minority of those complete the purchase and show a conversion.

- when it comes to online advertising such as Facebook Ads or Google Advertising, specifically for the e-commerce businesses, sales conversion is one of the most important metric i.e., "Key Performance Indicator" (KPI) to analyze the user behavior, engagements, and purchases. All these metrics assist in determining the final return on investment (ROI) from online advertising for the company.

One simply cannot ignore the ROI factor as it is the judging criteria to decide whether a marketing channel should continue to be used for brand awareness or not. Let's say a small amount of $80 for a day was spent on ads that resulted in total of 16 sales on the e-commerce website with an average cart size of $150.

This could be considered as a profitable merchandising or advertising campaign that gives the return on investment and the company would continue to leverage this marketing channel with higher ads spent. Alternatively, however, the same e-commerce company would take if they spent $1000 for seven days on the same ad campaign but fail to get their return on investment. This is where the concept of improving conversions rise.

Usually, warehouse management staff do not possess 360° visibility to their inventory stock. They usually face difficulties such as running out of the required stock at the very last minute at the time of an immediate demand from customers. This fallacy directly impacts in letting the company's cash flow fall.

## 4. METHODS TO IMPROVE CONVERSIONS AND SALES

Improving conversion rate, a metric generally called "Conversion Rate Optimization (CRO)". These are data-driven approaches along with a series of procedures through which any web portal, WebApp or an Advertisement could be improved and optimized to achieve increased sales and maximum revenue with better usability and high on-site user engagement. "Sales" metric is one of the most important metrics highly needed to be improved to achieve higher conversions.

Numerous data science and analytical techniques are implemented on e- commerce websites that are producing data, that later produce considerable results with optimized metrics for decision making. Some of the techniques are still under experiments.

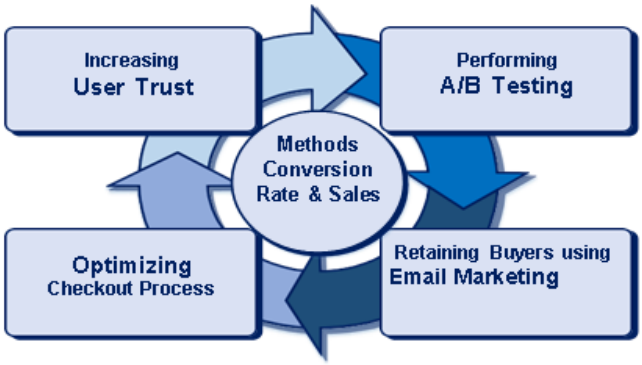



**Chart -2:** Methodology Flow chart

**PERFORMANCE (A/B) TESTING:**

I .A/B testing is an eccentric way for framing the best online promotional and marketing strategies for any business". It is a method to select the design, content, and functionality of a webpage that seems more successful for conversion factor by website visitors. This test is also used to test website copy, sales emails, search Ads, and so on etc. The variation in components or elements of a page that may affect the actions or behavior of site visitor could be tested easily.

II. The iterative A/B testing is more effective to find an improved version. There are two types of A/B testing: (1) Client-side and (2) Server-side, with client-side being the more common. Client-side testing involves showing the same version of a page to every website visitor and then using

JavaScript to make changes within the visitor's browser before the visitor looks to the resulting page. On the other hand, Server-side testing is applied when the webserver shows visitors with different page variations, altering them on the server before they are sent to the visitor's browser. No modifications to the page are permissible at the browser. A/B testing can be applied in the cases where an e-commerce owner wants to compare: (1) Various page-layouts, (2) Navigation organization, (3) Headline effectiveness and content, (4) Website photography and product images, (5) New visual styling for a page, and (6) New pricing strategy or different promotions and offers.

III. Some of the research tools used are included as (1) Analytics, (2) Heatmaps, (3) User Tests, and (4) Surveys. The detailed description of research tools is elaborated in Table below.

| Sr. No. | Tools | Description |
|---|---|---|
| 1. | Analytics | Identify the most popular pages, CTR, traffic sources, pages with a high bounce rate, and paths through the site. |
| 2. | Heatmaps | Discover exactly where and how people navigate the website. See where they click, where they pause, where they get confused, etc. |
| 3. | User Tests | Personally watch users navigate the website to discover where points of friction occur.<br>Proper user testing will also include the testing subjects speaking out loud about what they are experiencing as they complete tasks on the website. |
| 4. | Surveys | Directly ask the customers about their shopping experiences. Did they encounter any problems? What would they like to see improved? |

**Table -1 :** Tools and Description

**Retaining Buyers using Email Marketing**

I. According to the study on returning customer after a successful sale, a customer has a 27% chance of buying from the same e-commerce website again. If they make a second and third purchase, they have a 54% chance of making another i.e. fourth purchase, hence adding value to the business. Today, to get the most out of "customer lifetime value (CLV)", an e-commerce website needs to retain, upsell, and cross-sell its customers. To do that in a world of increasing population and market competition, they must stay top-of-mind. Email marketing is still the top way in which most people are easily reachable. The significance to retain customers in a way that encourage them to return and refer to the product or shopping web again is to make them feel like they are part of a brand's family, where customarily and socially, the families should always stay in touch.

II. The timely emailing to the customer with the release of a brand or any promotional sale with relevant content on regular basis builds trust and importance of that demand in their running life. Customer will surely recommend that product or brand to other members of family or friend first psychologically if they get themselves satisfied with such dealing of care. It costs 5 to 10 times more to acquire a new customer, than the efforts required to retain an existing customer as a returning customer. Every time a customer returns, it is more likely for him to return and buy again since returning customer build businesses. The retention rate could be maintained and increased by email campaigns, push-sales etc. that increase the ROI.

**OPTIMIZING CHECKOUT PROCESS**

I. Optimizing the entire checkout process is a core technical work that requires a lot of testing at the developer's end. For example, an e-commerce website has a separate payment page where 84.71% of the visitors' traffic land to buy a product. If we increase this percentage to 90%, that will result in generating 461 more orders and an additional USD 87,175 per month. If we take a close look, that's a surprising rise in revenue of 23.94%. Although it seems little count, considered a gain in revenue.

II. Similarly, another good technique to apply is to make the shopping cart and products added to it along with the "Checkout" or "Order Now" button to remain visible at all the time until the user clicks somewhere else. Also, changing the "Checkout" button color to something that stands out – mostly yellow or green – is also known to work and could be another advantage.

**Visibility**

Throughout the buyer's journey, the shopping cart feature should display all the products that are added to the cart and show the total cost of the basket. The reason behind this is to avoid card abandonments on the checkout page.

**Control**

The customer must have the control to easily make any changes to the product quantity or even remove them.

Here, another amazingly profitable strategy that many e-commerce business owners today still don't know about is Cart Abandonment. It is very common to see that most visitors on a website who click on a product to add it to the shopping cart, usually don't complete the checkout process and leave the website without making any purchase. This heavily results in a great loss of sales and revenue. As a solution, if we do not let the cart contents expire and

leverage the technology of cart abandonment email reminders and re-targeting ads to bring those potential customers back to the same checkout page can dramatically increase conversions and recover lost sales.

The last phase of the checkout optimization is tweaking the payment fields. If the customer has entered the shipping information first, only then they must be able to get into the billing part. This multi-step form process is used to push the customer towards commitment and consistency. The logic behind this is very simple – when a human brain starts doing something, it focuses the person to eventually finish it. This little psychological hack could significantly add up sales.

**INCREASING THE USER TRUST**

- Security and online privacy are some of the most important concerns for many of us who are doing business activities or even buying activities online. Customers, especially in countries like Pakistan, still are afraid to use credit or debit cards online to avoid any mishaps or potential theft. Therefore, an e-commerce website of any kind needs to ensure the use of end-to-end encryption from the origin server to the client-side via SSL and TLS certificate. If an e- commerce website is already using all the required security protocols, it is still important to tell visitors about it by displaying site security seals and PCI compliant labels prominently on the payment or checkout page.

- A few proven techniques to let customers know that their details are secure are (1) Using different background color for payment form – preferably light green color, (2) Displaying SSL certificate logo or a green padlock icon, and (3) An additional written statement with the SSL logo such as: "Payment secured by 256-bit SSL Encryption".

- In general, most users are not very technically aware and therefore they probably are not aware of terms like SSL or HTTPS secured, so telling them about their data security in plain terms is also a good idea.

## RESULTS & DISCUSSION

For every task we do and for defect we fix, for every feature we add we a need the result. The result can give I a outcome or an idea to fix, this thing helped very well in this project and lead to a perfect outcome as expected. The results play a main to boost up your ideas and help to reach the goal. The managed plan has been played main role to achieve the good results.

Discussion has been made weekly, the mentor has helped us in giving suggestions and approach to solve the issue and use different techniques to achieve the results. The proper communication can help the discussion more interesting in the technical terms, the sync ups are the main things for the issue discussion and solving them using better approach finalized after a discussion.

## CONCLUSION

Concluding the results, we see that people are not happy and consider it a good thing sometimes. People suffer because of its bad effects and are very hostile. They choose according to tradition methods over this new technology. As the population grows companies and markets it grows day by day. So people place threats of cheating, fraud, transactions, etc. In fact, people do not really think that they can be trusted purpose. So, we can clearly conclude that traditional shopping is better than online to buy. This paper helps in expanding business. Saves a lot of time for client's customers who ever visits the store. Merchant convenience and saving merchant's time for other customers. Knowledge gained by the students by building a project and interacting with team members and mentor can help in improving the communication skills and completion of project helps in improving domain knowledge.

Machine learning methods are highly useful where general computing methods are less effective. But it is worth mentioning that machine learning can be used as one method for data analysis. This is because, in essence, machine learning can be thought of as a data analysis method that builds the analytical model automatically. It can help identify and uncover patterns found within the data. Further, these intelligent and smart approaches strengthen the enterprise to lead the market more effectively with timely product marketing, timely introducing product promotions and provide a better quality of services to potential customers.

## FUTURE WORK

In future we can add more enterprises and classes to increase the security of the project time you can use some third party APS to trigger the emails and notifications which can be sent to the user and also about the delivering it can also integrated other payment things which are enabling the net banking internet upi and credit and debit card transactions also interested in having some specific wallets we can also introduce the wallets in the future to the project. Apart from the payments which we can also had a delivery tracking with online GPS when the product is out for delivery so that customer can easily able to track like some of the delivery systems zomato and swiggy.

Also, we can add some authentication things while login in and and also integrated some Google API for authenticating using the Google authenticator and also we can introduce so many things to improvise the security which can be updated in the future in latest. The way US designing can be improvised more in the coming days it requires a good team

with good set of knowledge on the designing where they can help in impressing the designs and make look and feel very well . The authors can acknowledge any person/authorities in this section. This is not mandatory.

## REFERENCES

[1] The Virtual Online Supermarket: An Open-Source Research Platform-Experimental Consumer Research, April 21.

[2] H. Saleem, K. B. Muhammad, A. H. Nizamani, S. Saleem and J. Butt, "Data Science and Machine Learning Approach to Improve E-Commerce Sales Performance on Social Web," International Journal of Advanced Research in Engineering & Technology (IJARET), ISSN: 0976-6499, Scopus Indexed, HEC Recognized "Y" Category Journal, vol. 12, no. 4, pp. 401-424, 2021.

[3] Machine Learning: Algorithms, Real-World Applications and Research Directions,[2021], Iqbal H. Sarker ,SN Computer Science volume 2, Article number: 160 (2021)

[4] 'Data Science and Machine Learning Approach to Improve E-Commerce Sales Performance on Social Web, April 2021 | 12(4):401-424 DOI:10.34218/IJARET.12.4.2021.040

[5] Supermarkets of the future: deploying AI in your grocery store, By Nadejda Alkhaldi, Innovation Analyst, Published on May 7, 2021

[6] Data Science Use Cases in Retail Industry | Mayur Badole — May, 2021

[7] 'How Zomato Uses Machine Learning | Zomato saw 3x improvement in its 'feature store' service compared to last year | By Amit Raja Naik | Published on August 20, 2021 in Opinions.

[8] Are AI Technologies Paving a New Path for Food Retail? | 2020 | India

[9] Grocery Prediction using Machine Learning, ISSN 2321 3361 © 2020 IJESC, Neha Kumari, Department of CSE,VIT University Chennai Campus, India. 2020

[10] How ML/AI is revolutionizing grocery delivery service |Ravi Chandra| Aug 2020, India

[11] 'How Data Science and AI are changing Supermarket shopping', Richard Farnworth [2020], Adelaide, Australia.

[12] Anatomy of Understanding Consumer Behavior When Grocery Shopping in the New Next 2020.

[13] J. Leek, "The key word in "Data Science" is not Data, it is Science," SimplyStats, 2013. [Online]. Available: https://simplystatistics.org/2013/12/12/the-key-word-in-data-science-is-not-data-it-is-science/. [Accessed 2019].

[14] 'How Grocers Are Reimagining the Future With AI? ' | By Amanda Baltazar on Jan., 2019.

[15] Online Grocery Shopping and Consumer Perception, Feb 2018

[16] A Study on Consumers acuity towards online grocery shopping, December 2017

[17] Online grocery shopping: the key factors influencing consumers' purchase intention, May 2017

[18] Web Content Analysis of Online Grocery Shopping Web Sites in India, March 2016

[19] Park, A. C. W., Iyer, E. S., Smith, D. C., Iyer, E. S.,& Smith, D. C., 2016. The Effects of Situational Factors on In-Store Grocery Shopping Behavior: The Role of Store Environment and Time Available for Shopping. Journal of Consumer Research. 15(4),422-433.

[20] Retail, G. T., & Survey, C., 2015. Total Retail 2015 (February). Statistica. (2016). ECommerce. Retrieved June 26, 2016, from https://www.statista.com/markets/413/ecommerce/U.S. Census Bureau News. (2016). Retrieved from http://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf

[21] Tribune., 2015. By 2020: Pakistan's ecommerce market to surpass $1 billion. Retrieved May 23, 2017, from https://tribune.com.pk/story/1016706/by-2020-pakistans-e-commerce-market-to-surpass-1-billion/

[22] Wharton., 2015. Online Groceries in India: Will Consumers Bite? Retrieved July 9, 2016, from http://knowledge.wharton.upenn.edu/article/online-groceries-in-india-will-consumersbite/

[23] Online Grocery Shopping: An exploratory study of consumer decision making processes, Joana, March 2013.

[24] Pozzi, A. (2013). The effect of Internet distribution on brick-and-mortar sales. The RAND Journal of Economics, 44(3), 569–583.

[25] Pozzi, A. (2012). Shopping cost and brand exploration in online grocery. American Economic Journal: Microeconomics, 4(3), 96–120.

https://www.businesstoday.in/latest/trends/reliance-jiomart-opens-online-grocery-service-in-200-cities-key-things-to-now/story/404916.html