DISCLAIMER NOTICE

Due to the condition of the original material, there are unavoidable flaws in this reproduction. We have made every effort possible to provide you with the best copy available.

Thank you.

The images contained in this document are of the best quality available.
Improving E-Commerce sales using Machine Learning

by

Shaheen Thobani

Submitted to the System Design and Management Program
on January 15, 2018, in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Engineering and Management

ABSTRACT

Trends show promising growth of the online e-Commerce industry. While the e-Commerce companies are aggressively moving towards digital sales and marketing, the customers are being bombarded with frequent and often irrelevant marketing communication from myriad sources. The thesis proposes understanding the digital purchase journeys of the customers from the lenses of both sellers and customers to make online sales and marketing efforts relevant and intelligent. The thesis applies the improved customer journey framework to identify the needs of the customers and goals of the seller at various stages of customer purchase journey. It discusses the need to take an integrated view of the purchase journey to improve the customer experience at the journey level. It illustrates with an example how to design end-to-end journeys – a starting point for consciously shaping the purchase journeys.

Larger companies are using Machine Learning to improve marketing technologies and processes to create a competitive advantage and capture market share through digital presence. The thesis aims to understand and illustrate the applications of Machine Learning to digital sales and marketing ecosystem for the e-Commerce industry. It first understands the e-Commerce touchpoints using which customers interact with the brands and delves deeper into the underlying technologies powering these touchpoints. Then it illustrates and analyzes the application of Machine Learning to the e-Commerce website which includes search, recommendation system, and Product Detail Page with an aim to improve conversion, and to the advertising ecosystem which includes Data Management Platform and Demand Side Platform in order to enable prospecting and customer targeting. The thesis also illustrates and proposes the use of a framework called ‘Machine Learning Canvas’ to systematically apply Machine Learning to any system while keeping value proposition for the business in the center.

Thesis Supervisor: Michael A. Cusumano
Title: Sloan Management Review Distinguished Professor of Management
MIT Sloan School of Management
ACKNOWLEDGEMENTS

I am grateful to all the fantastic teachers I have met during my education at MIT. They have contributed tremendously to my growth and learning. I am also grateful to all my schoolmates and friends for the time I have spent with them as it helped me evolve intellectually and culturally.

First, I would like to thank my thesis advisor Professor Michael Cusumano for his advice, dependable support, and patience during my thesis journey. In his classes and presence, I have learned more than what can be taught.

I am grateful to Joan Rubin and William Foley for their constant support and guidance throughout the program, and making the journey at MIT pleasant.

Finally, I am thankful to my family, my pillar of strength and support, for always being with me in all my decisions.
ABSTRACT

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

1. Introduction
   1.1 Motivation
   1.2 Approach and Objective
   1.3 Scope
   1.4 Structure

2. Customer Purchase Decision Journey
   2.1 Introduction
   2.2 Evolution of the Customer Purchase Journey
   2.3 Understanding the customers' purchase decision journey (CDJ) through the lenses of sellers and customers

3. Customer Touchpoints
   3.1 What are touchpoints?
   3.2 Mapping touchpoints
   3.3 Designing the journey
   3.4 Considerations for designing the customer journey
   3.5 Touchpoints - an integrated view
   3.6 Important touchpoints for Ecommerce
      3.6.1 Online Ads
      3.6.2 E-commerce website

4. Technologies Underlying Touchpoints
   4.1 Recommendation Systems
   4.2 Applications of recommendation systems in e-Commerce website
   4.3 Approaches to recommendation systems
4.3.1 Collaborative filtering
4.3.2 Content-based filtering
4.3.3 Hybrid recommender system

4.4 E-Commerce Product Search

4.5 Advertising
4.5.1 Programmatic Advertising
4.5.2 Key players and components of the advertising ecosystem
4.5.3 Technologies enabling advertising
4.5.4 Programmatic advertising – a closer look

5. Applying Machine Learning
5.1 Introduction to Machine Learning
5.1.1 Description
5.1.2 Classification of ML
5.1.3 Choosing a model

5.2 Machine Learning Canvas

5.3 Application of ML to Recommendation Systems

5.4 Application of ML to e-Commerce Search

5.5 Application of ML to advertising ecosystem (DSPs) for prospecting

5.6 Application of ML to Propensity Modelling for customer targeting

5.7 Application of ML to Product Detail Page Optimization

6. Conclusion
6.1 Thesis summary
6.2 Challenges and opportunities for future work
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Correlation of Retail e-commerce with smartphone penetration</td>
</tr>
<tr>
<td>2</td>
<td>Classic funnel metaphor of a linear customer journey Vs evolved non-linear journey</td>
</tr>
<tr>
<td>3</td>
<td>Evolved loop metaphor customer purchase journey</td>
</tr>
<tr>
<td>4</td>
<td>Classic Journey Vs Accelerated Loyalty Journey</td>
</tr>
<tr>
<td>5</td>
<td>Customer Journey Framework</td>
</tr>
<tr>
<td>6</td>
<td>Detailed TV purchase decision journey analysis from customer lens</td>
</tr>
<tr>
<td>7</td>
<td>Brand Touchpoint Wheel for all company touchpoints</td>
</tr>
<tr>
<td>8</td>
<td>Brand Touchpoint Wheel for customer touchpoints</td>
</tr>
<tr>
<td>9</td>
<td>Customer journey maps through touchpoints</td>
</tr>
<tr>
<td>10</td>
<td>Purchase Decision Journey of Persona 1</td>
</tr>
<tr>
<td>11</td>
<td>Purchase Decision Journey of Persona 2</td>
</tr>
<tr>
<td>12</td>
<td>Correlation of customer satisfaction and Willingness to recommend with touchpoints Vs Journeys</td>
</tr>
<tr>
<td>13</td>
<td>Method for inbound marketing</td>
</tr>
<tr>
<td>14</td>
<td>Recommendations on the landing page of the e-Commerce website</td>
</tr>
<tr>
<td>15</td>
<td>Recommendations on the Product Detail Page</td>
</tr>
<tr>
<td>16</td>
<td>Feedback profile, a form of recommendation system on eBay</td>
</tr>
<tr>
<td>17</td>
<td>Steps for memory-based approach</td>
</tr>
<tr>
<td>18</td>
<td>Architecture of content-based system</td>
</tr>
<tr>
<td>19</td>
<td>Representation of Programmatic Advertising Ecosystem</td>
</tr>
<tr>
<td>20</td>
<td>Representation of Adobe’s Data Management Platform</td>
</tr>
<tr>
<td>21</td>
<td>Programmatic transaction types</td>
</tr>
<tr>
<td>22</td>
<td>Representation of advertisement inventory transaction process</td>
</tr>
<tr>
<td>23</td>
<td>Real-Time Bidding process</td>
</tr>
<tr>
<td>24</td>
<td>ML Algorithm Cheat Sheet</td>
</tr>
<tr>
<td>25</td>
<td>Machine Learning Canvas</td>
</tr>
<tr>
<td>26</td>
<td>ML canvas applied to ML for recommendation system</td>
</tr>
<tr>
<td>27</td>
<td>User Probability Matrix</td>
</tr>
<tr>
<td>28</td>
<td>ML algorithms for recommendation systems</td>
</tr>
<tr>
<td>29</td>
<td>ML algorithms for e-Commerce search</td>
</tr>
<tr>
<td>30</td>
<td>ML algorithms for customer prospecting</td>
</tr>
<tr>
<td>31</td>
<td>Representation of Propensity Model</td>
</tr>
<tr>
<td>32</td>
<td>Product Detail Page</td>
</tr>
</tbody>
</table>
1. Introduction

1.1 Motivation

Research forecasts that e-commerce will grow from approximately $2.8 trillion market – currently, 10 percent of the pie that represents $28 trillion retail markets worldwide to a $4 trillion market globally by the year 2020. A study by Nielson shows that the retail e-commerce growth is strongly correlated with smartphone penetration in most of the markets. Though the smartphone penetration does not completely explain the growth of retail e-commerce, the trend does imply growth in online retail shopping [1].

Figure 1, Correlation of Retail e-commerce with smartphone penetration [1]

A report published by Forrester observed and predicted that though the e-commerce market shows promising future, the gap in the market is widening. [2] Some of the top players such as Amazon will continue to aggressively gain market share and out-do many small e-commerce retailers. Among the losers are mostly the smaller e-commerce retailers partly because they do not have more sophisticated ways of reaching out to their potential customers. Larger firms have recently started adopting the advanced data analytics, Machine Learning, and Artificial Intelligence in order to make their customer reach accurate.

On one hand, the e-commerce retailers are increasing the possibilities of reaching out to potential customers, on the other end customers are getting bombarded with advertisements and recommendations through their smart devices. Though the sellers now have so much more data about the customers, not knowing how to use that data is resulting in innumerable and unsophisticated sales and marketing efforts. This is making consumers insensitive and ignorant due to their irrelevance of content, time and place, to the extent of having them look out for tools such as Ad-blockers. Using the predictive algorithms of machine learning can let companies identify their potential customers and target them, improving their return on investment on their marketing and sales efforts.
In order to improve their sales, online retailers do not only need to target right customer, they also need to target them at the right time and during the right stage in their purchase decision journey and on the right place (device or location) where the probability of catching their eyeballs is highest. In addition to improving the customer reach, the companies also need to leverage analytics methods such as Machine Learning to make their processes efficient.

1.2 Approach and Objective

The objective of this thesis is to systematically study the application of machine learning in the online sales process with a goal to improve it. This thesis also attempts to understand the progress that has been made in the application of Machine Learning in the space of retail e-commerce. The thesis approaches the e-commerce sales using the framework of Customer Purchase Journey and then further diving deeper by looking at the level of touchpoints used by customers to navigate through this journey. Then it attempts to understand the technologies used to run these touchpoints and applications of Machine Learning to these technologies.

1.3 Scope

The thesis limits the scope of its application and relevance to commodity products in retail e-commerce including apparels, FMCG, and electronic products. It is important to note that the customer purchase journeys for products such as automobiles, real estate properties, music subscriptions and all services is different than for the commodities. Further, the touchpoints considered in this thesis are online touchpoints, whereas the touchpoints used during the purchase journeys of above-mentioned products and most services are mix of online touchpoints (used for research and awareness) and offline touchpoints such as customer service phone calls and retail stores (for other activities of purchase decision journey including purchase).

1.4 Structure

This thesis has four chapters. The first chapter starts by describing the evolution of customer’s purchase decision journey from the metaphor of ‘funnel’ which implies more of a linear approach to understanding customers’ actions and steps from the motivation of purchase to the moment of purchase and later to metaphor of ‘loops’ that takes into consideration the non-linear nature of the journeys today as a result of connection between various stages of the journey and the stage of purchase. Further, the chapter presents a framework of purchase journey from the seller and customer point of view, and delves deeper into their objectives at each stage, and ends with an example of the application of the similar framework to explore the journey of a customer who wants to purchase a TV.
The second chapter defines ‘touchpoint’ – the front end that the customers interact with and briefly describes touchpoint mapping. Next, it recommends the process and considerations for designing customers journeys with an understanding of touchpoints. The last section of the chapter describes most major online touchpoints for e-Commerce customers which are broadly various types of online advertisements and the e-Commerce website.

The third chapter delves into the backend technology of the touchpoints covered in the second chapter. Before applying Machine Learning, it is important to understand the process and system that manages the data, so the chapter starts with understanding the complex data management platform end-to-end starting from the sources and ways of data collection to the data delivery to the publishers of the online advertisements. Further, it presents the study of advertisement technology and finally ends with the brief study of the technology behind systems of major touchpoints on the e-Commerce website.

The fourth and final chapter presents and analyzes the current Machine Learning application in e-Commerce website Search and Product Description Page, the Recommender System and the Ad targeting.
2. Customer Purchase Decision Journey

2.1 Introduction

This chapter discussed the evolution of the Customer Purchase Decision Journey (CDJ) and describes the framework that can be used to understand the journey from the lenses of the customers as well as the sellers.

CDJ is the model that describes how customers make purchase decisions that start with the day to day influences of brands and go up to the point when customers purchase and reconsider or recommend the product they bought. [3] The survey by Salesforce shows that 86% of senior level marketers believe that it is very important to have a cohesive customer journey. This journey needs to be consistent and make the process as painless as possible for the customer. [4] Google provides their analytics solutions at not only touchpoint levels but journey level. They assert ‘It's important to understand the entire customer journey so you can measure all of the elements that contribute to your campaigns, attribute the right value to them, and adjust your marketing budgets where appropriate.’ [5]

2.2 Evolution of the Customer Purchase Journey

Traditionally, the consumer decision journey has been understood through the metaphor of a “funnel” (Figure 2 Left) - the consumers start with a large number of options at the top of the funnel for the products they want to purchase and systematically narrow down the choices to purchase one. This notion was true until nearly a decade back when consumers were introduced to options through very few channels or through their visit to the brick and mortar stores after which they evaluated and narrowed the number of options. Today, the consumers are continually exposed to various products through multiple channels that leave the impressions of products and brands in their minds. The multiplication of products, fragmentation of the marketing channels and evolution of media have led to the reduction of the number of brands consumers remember and start their journey with. [6] When the consumers begin to consider buying an item, these accumulated impressions inform the potential options they may start with. The number of options might increase rather than decrease in the next step of the purchase journey when consumers reach out to digital channels to evaluate their decision, or when consumers go to on the product page of an e-Commerce website and see recommendations and options for a product they set out to purchase.

The analogy of funnel was appropriate until a decade back because then the purchase journeys were relatively linear. However today, the increase in the types of digital channels that potential consumers use and are exposed to has resulted in non-linear purchase decision journeys. The customers might interact with new brands at any stage of their purchase journeys. In some cases, brands that were initially not being considered by the consumer may enter into customer’s decision journey in the later phases, of the purchase
journey, sometimes resulting in customers dropping the brand they were most considering earlier. This non-linear nature of the journey provides opportunities for brands to enter into the customer’s decision consideration when it is most efficient and makes the most impact on the customer [6]. The customers’ decision journey evolved and became more complex than a linear funnel as sellers engage with their multi-touchpoint customers with a multichannel approach [7] (Figure 2 Right).

As the conversation between sellers and the consumers started becoming two-way from one-way, the metaphor of the consumers’ purchase journey evolved from a funnel to a circular loop (Figure 3) [6]. The loop implies that many sellers can successfully convert buyers to become long-term loyal customers giving rise to a loyalty cycle between the ‘Trigger’ stage, a time when consumers discover their intent to purchase an item and ‘Purchase’ stage, time when they purchase the same brand. For example, Sephora, a French brand of high-end cosmetic products has a very strong loyalty program and it creatively stays in touch with its customers post their purchases through Sephora mobile app by sending notifications and reminders to reorder a particular Sephora product their customers precisely when they are about to run out of it. The communication from customers back to the sellers (and other customers) also consists of feedback about the performance of products. The Sephora app helps get feedback from customers about a product’s performance after days and weeks of the purchase, and it also helps Sephora get information about whether a particular customer is using right products for her skin type and tone – a factor critical for the success of a cosmetic product.
As sellers have the ability to reach customers by several digital channels and at all time, they can actively shape the customers’ decision journey to deliver value to both the customer and their brand. (Figure 4). Some of the important capabilities that companies need to deliver such value are ability to proactively personalize the customer interaction and messaging, contextually and timely reaching out to the customers at intended stages of their decision journey to have best influence on their purchase decisions, and identifying and delivering new sources of value, such as new services, for both the customer and seller. [8].

Figure 3, Evolved loop metaphor customer purchase journey [6]
The underlying technical capability that companies require is the ability to leverage customer’s personal data and learn to predict their behavior from a continuous stream of data about their activities and presence on various digital channels. The companies need to leverage machine learning to intelligently automate the intensive tasks of predicting customer behavior and personalize selling at the level of individual customers. They can use orchestrate their communication with customers through various stages of the journey, to improve their sales and marketing spend effectiveness.

2.3 Understanding the customers’ purchase decision journey (CDJ) through the lenses of sellers and customers

The purchase decision journey of the customers largely shapes their experience and perception of a brand, and so they have the large potential be a source of competitive advantage for improving the sales. This Thesis adopts from and modifies the traditional CDJ framework. The modified framework (figure 5) proposes to perceive the decision journey from the standpoint of the sellers and the customers in order to understand their objectives during various stages of the journey. This understanding will help sellers strategically align their sales and marketing efforts with the needs and actions of the customers. Following are various stages of the customer journey that identify various stages of the journey by the goals or objectives of customers and sellers.
1. **Customers want to be inspired for a purchase;**
   Sellers want to trigger the desire for purchase

The journey begins typically with a customers’ need or inspiration arising from an encounter with a brand, a product or its advertisement. This stage includes making the customers aware of the brand; the set of options considered at the outset could be thrice more likely to be purchased finally than the brands outside that set [6]. Today, the customers start their journey with less number of brands in mind due to the exposure to a plethora of advertisements compared to a decade back. The sellers’ objective at this stage is to trigger that desire for purchase. They want to either make customers feel that their current solution is not sufficient or enhance the urgency and make them aware of a solution to overcome their problem and solve their needs. And so, whether or not the customers’ consider the sellers’ products in their decision journey depends largely on seller’s capability to outreach their potential customers.

2. **Customers want to explore options**
   Sellers want to inform more about their brand/products

At this stage, the customers want to explore options while minimizing their search efforts. The sellers’ objective should be to help them explore options easily by partnering in the research with the customer. A research by HubSpot about how much part of overall shopping online research is found that half of the shoppers spend 75% of time conducting online research. The number of options initially considered might increase as customers proactively seek reviews, ratings, and recommendations, learn about new options during comparison or update their criteria as a result of learning new information. The recommendations and reviews have become significantly more important part of the product exploration stage. According to the Nielson study on the influence of reference on purchasing decisions, it was found that 82% customers proactively seek referrals before making a purchase. [9]
The marketing is increasingly shifting from the time when sellers “pushed” marketed and reached out all customers with the same content without considering much about the timing and the stage of their decision journey to the time when customers “pull” information needed to make decisions and inform themselves. A study shows that the customers engage with only one-third of the touchpoints that are activated by the company in the evaluation stage of the decision journey compared to two-thirds the number of touchpoints that are customer driven such as reviews and ratings, social recommendations from peers and influencers and also include the experiences of past interactions with the brand. [6] In this stage, the sellers focus on finding on which touchpoints can they can meet their customers and deliver the relevant content through the relevant touch points. In addition, they should help the customers to move to the next stage at the decision journey or direct them to the relevant touchpoints for more information.

3. **Customers want to evaluate the solution**
   **Sellers want to convince that their solution is superior**

While sometimes customers explore and evaluate simultaneously, other times active evaluation of a solution follows the exploration stage. At this, the customers are clearly intent on purchasing, and their main objective is to find the best value. Companies’ ability to relevantly adapt to different consumer needs and behavior is based upon their knowledge of customers’ profile and this knowledge can significantly improve customers’ experience with seller’s brand during their journey. If the customer is new to the category, the sellers need to inform or educate them about the solutions, type of products, features, brands, prices and where to buy and where to get the best deal, if the consumer is in the replacement stage, sellers can update them about their latest products and their features, special value packs, most recommended products by other consumers or celebrities, competitor brands and prices and where to buy to get the best deal.

4. **Customers want to purchase**
   **Sellers want to encourage the purchase of their product**

This stage mainly includes the moment of purchase by the customers. Customers need easy and frictionless purchase experience. For more expensive or technical products they want the assistance of the sellers during their purchase process. They also want to be sure that they are getting the best value.

The sellers’ objective is to encourage the consumer to close the deal. They also encourage purchase upgrades, cross-sell, and prompt adding accessories to the purchase to drive basket size up. The purchase is also called conversion, generally in the context of e-Commerce sales. To improve the sales, the sellers often drive promotional conversion using tactics including:
- Price discounts: Sellers can drive customers’ consideration and conversion by offering price discounts. However, lowering down the base price in the digital channel can hurt their brand perception, bring value loss over time and create channel conflict if the discounted price is not offered in-store or other online channels.

- Coupons or cashbacks: These promotional tactics do not impact base price, and are not picked up by price comparison engines online. Sellers could use them to drive traffic while reduce channel conflict and working close to the moment of purchase.

- Bundles: In e-Commerce, not only physical bundles but also virtual bundles can be setup at retailers without having a cost related to special packaging.

5. Customers want to receive value
Sellers want to delight the customer

After purchasing a product, the customers judge its value throughout its lifetime. Though sellers’ objective post-purchase is to improve the consumer loyalty, it comes as a by-product of reassuring the customers about the value of the product by helping them to reach ‘point-of-value’ and delighting them by serving a best-in-class experience. Sellers could do so by delivering an enhanced post-purchase experience in the product usage. For example, for the electronic devices they can have a digital channel that provides how to use guidance, how to maintain/clean/charge instructions, and how to obtain the best results or product life tips. They need to support the users in case of issues with the products, or the usage, maintenance, repair, and all the generic customer care related issues. In addition, they should register customers on their database and track customer activities on social channels in order to growth efficiency their database and carry out efficient Customer Relationship Management activities to initiate dedicated loyalty programs.

6. Customers want to share their experiences and reconsider the brand
Sellers want to motivate to positive recommendations and ensure loyalty

Sellers do inspire repurchases from their customers by creating regular incentives for loyal consumers and interacting with them post-purchase to communicate the innovative propositions. But because of an array of options and other sellers trying hard to acquire new customers, sellers need to turn their customers form ‘passive’ loyalists – who tend to stick with the brand to ‘active’ loyalists – who go beyond and recommend it. Sellers should spend on the relevant touchpoints with a focus on expanding their base of active loyalists.

It is evident from numerous product reviews on the internet that customers often express their opinions about the products though ratings and review not only on the channels of purchase but also socially. So,
after the purchase, the sellers should aim to seek advocacy by stimulating some form of sharing and promoting using relevant channels. Facilitating rating and review generation will help sellers convince other consumers in their active evaluation phase by providing them ‘social proof’ of their products, will contribute to triggering new consumers and also raising the visibility in search engines. Many sellers have started enabling, stimulating and incentivizing user-generated content about the propositions, so as to fuel the content on the touch points such as forums, YouTube, e-Commerce platforms, etc. used by other consumers in their active evaluation phase. It improves the visibility on search-engines and contributes to triggering new potential customers.

In addition to understanding the goals of the customers at each stage of the purchase journey, it adds higher value to consider what’s happening at these stages more granularly from the lens of the customer. The following framework (adopted from an article in Harvard Business Review) dives deeper into the process at the customer end by asking questions about each stage of the journey about four factors –

Actions: What are the actions of the customers at this stage and what actions do they take to move to the next stage? [10]

Motivations: Where are the customers coming from? What are their needs? What motivates the customers to advance to the next stage? What are their main criteria? [10]

Questions: Which questions customers might seek answers to? What do customers surely care for of they want to move advance to the next stage? Could there be any issues in understanding the value proposition or technical and operational aspects of the product? [10]

Barriers: Are there potential issues such as a tedious payment process, missing element in the services with the product, unavailability of inventory, or cost that could hinder consideration of product purchase or movement of the customer to the next stage. [10]

Following is an example of the application of the above framework for customer TV purchase decision journey. (figure 6) The seller needs to consider what are all the potential Actions, Motivations, Questions and Barriers that can come up at customer end during all the stages of that journey. [10]
<table>
<thead>
<tr>
<th>Activations</th>
<th>Research</th>
<th>Purchase</th>
<th>Out-of-box experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Hears about it from friends or sees it in friend’s home</td>
<td>• Asks friends for advice</td>
<td>• Open the package</td>
<td>• Open the package</td>
</tr>
<tr>
<td>• Sees it on TV, web or magazine</td>
<td>• Visit stores and asks sales person questions</td>
<td>• Look for directions</td>
<td>• Get it done fast</td>
</tr>
<tr>
<td>• Reads websites, blogs, magazines for reviews and to gain understanding</td>
<td>• Leans jargons, brands</td>
<td>• Program and learn the setting</td>
<td>• Check out new features</td>
</tr>
<tr>
<td>• Looks for sales</td>
<td>• Consider ad-Ons and installation services</td>
<td>• Call customer service</td>
<td>• Relish new purchase</td>
</tr>
<tr>
<td>• Learn about it from friends or sees it in friend’s home</td>
<td></td>
<td>• Share reviews and concerns online</td>
<td>• Invite friends and trump about the purchase</td>
</tr>
<tr>
<td>Motivations</td>
<td>▸ House remodeling</td>
<td>▸ Make the best choice</td>
<td>▸ Get it done fast</td>
</tr>
<tr>
<td>▸ Broken TV</td>
<td>▸ Make the best choice</td>
<td>▸ Promotion in a store</td>
<td>▸ Check out new features</td>
</tr>
<tr>
<td>▸ Super Bowl</td>
<td>▸ Satisfy needs</td>
<td>▸ Release of a new model</td>
<td>▸ Relish new purchase</td>
</tr>
<tr>
<td>▸ Wedding gift</td>
<td>▸ Get the best deal</td>
<td>▸ Influencer recommended</td>
<td>▸ Invite friends and trump about the purchase</td>
</tr>
<tr>
<td>▸ Children want a TV</td>
<td>▸ Trump neighbors and friends</td>
<td></td>
<td></td>
</tr>
<tr>
<td>▸ Upgrade TV to make it compatible with new gadgets to pair with it</td>
<td>▸ Know enough to not get duped</td>
<td></td>
<td></td>
</tr>
<tr>
<td>▸ Technology change – maybe need a 3D TV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questions</td>
<td>▸ Which is best option</td>
<td>▸ Do I need it</td>
<td>▸ Is it damaged?</td>
</tr>
<tr>
<td>▸ How much? (Can I afford it?)</td>
<td>▸ What can I get for my budget?</td>
<td>▸ How can I get it home?</td>
<td>▸ Do I need help and directions to set up?</td>
</tr>
<tr>
<td>▸ Is it worth it?</td>
<td>▸ Is it future proof?</td>
<td>▸ Is it on sale or can I get discount</td>
<td>▸ What do I do with the old TV?</td>
</tr>
<tr>
<td>▸ Is it an improved product?</td>
<td>▸ Will it work with my other devices?</td>
<td>▸ What free accessories accompany the TV?</td>
<td>▸ How do I use it and change the settings?</td>
</tr>
<tr>
<td>▸ What will my friends think?</td>
<td>▸ Will it be durable?</td>
<td>▸ Is TV compatible with my older accessories?</td>
<td>▸ Is it meeting my expectations?</td>
</tr>
<tr>
<td>▸ Status quo – satisfied with TV?</td>
<td>▸ What criteria is important to me?</td>
<td>▸ Is it in stock</td>
<td></td>
</tr>
<tr>
<td>• Not tech aware or interested</td>
<td>▸ Is it user friendly?</td>
<td>▸ Does retailer deliver it?</td>
<td></td>
</tr>
<tr>
<td>▸ Honest, unbiased reviews</td>
<td>▸ Where can I get best offer?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>▸ Too much to learn</td>
<td>▸ Will it go on sale soon?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>▸ No time to do necessary research</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>▸ Too many seemingly good or bad choices</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barriers</td>
<td>▸ Don’t have desired model in stock</td>
<td>▸ Is it meeting my expectations?</td>
<td></td>
</tr>
<tr>
<td>▸ What to do with old peripherals and furniture?</td>
<td>▸ Unhelpful/unconvincing sales person</td>
<td></td>
<td></td>
</tr>
<tr>
<td>▸ Not compatible or too complex for pairing up with other devices</td>
<td>▸ Discovery of bad reviews</td>
<td></td>
<td></td>
</tr>
<tr>
<td>▸ Involved set up time consuming and no help from customer care</td>
<td>▸ Offer on competing brand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>▸ Status quo – satisfied with TV?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 6, Detailed TV purchase decision journey analysis from customer lens [10]

To summarize, companies should start with structurally looking into the customer purchase decision journey and analyzing the objectives from the lenses of both sellers and customers. Next chapter discusses the touchpoints that customers use to move through their decision journey and process for designing efficient customer journeys by identifying the relevant touchpoints.
3. Customer touchpoints

3.1 What are touchpoints?

Touchpoint Metrics (2003) defines touchpoint as “every point of contact—online and off; each communication, human resource, branding, marketing and sales process initiative creates touchpoints. The quality of touchpoint experiences drives perceptions, actions, and relationships.” [11]

This definition of touchpoint provides a brand level definition of touchpoints. Accordingly, a touchpoint is every point of interaction between the company and all its stakeholders outside including the customers. One can conceptualize the brand level touchpoints using Denise Lee Yohn’s framework called ‘Brand Touchpoints Wheel’ (figure 7, presented in her book ‘What Great Brands Do’) for various teams in a company to systematically identify and assemble their brand’s touchpoints. [12]

![Brand Touchpoint Wheel for all company touchpoints](image)

**Figure 7**, Brand Touchpoint Wheel for all company touchpoints

Teradata (2001) defines touchpoint as “customer interaction channels such as call centers, web sites, automated teller machines and web kiosks.” [11]

For the purpose of understanding the touchpoints in the customer purchase decision journey, we will go with the Teradata’s definition of touchpoints. So, customer touchpoints in our discussion will be considered as any point of interaction between the company selling products or services and the customers during their...
purchase journey. In ‘Brand Touchpoints Wheel’ framework that was originally conceptualized by Davis and Dunn (Davis and Dunn, 2002, p 6-7), customer touchpoints are categorized and explained in three stages as follows: “the pre-purchase experience touchpoints, the purchase or usage experience touchpoints and the post-purchase experience touchpoints”. [13]

![Brand Touchpoint Wheel](image)

**Figure 8, Brand Touchpoint Wheel for customer touchpoints**

### 3.2 Mapping touchpoints

This thesis explains the Customer Purchase Decision Journey in six stages. By adopting the idea of mapping the customer touchpoints to various phases of their purchase (Brand Touchpoint Wheel, Davis and Dunn), map all the touchpoints to various stages of purchase decision journey. This derived framework considers 2X2 matrix to map touchpoints across the stages of the journey, because it makes the possible mapping of multiple touchpoints with multiples stages of the journey. In reality, the customer could be interacting with same touchpoints in different stages of their purchase decision journey. In addition to mapping the touchpoints using this framework, the companies can map the existing and potential customer journeys through those touchpoints to ensure they move through the stages to purchase and share as intended by the companies. Following table maps potential journeys of a furniture buyer:

<table>
<thead>
<tr>
<th>Customer’s Objective</th>
<th>Be Inspired</th>
<th>Explore</th>
<th>Evaluate the solution</th>
<th>Purchase</th>
<th>Receive Value</th>
<th>Share and Reconsider</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seller’s Objective</td>
<td>Trigger</td>
<td>Inform</td>
<td>Convince</td>
<td>Encourage Purchase</td>
<td>Delight</td>
<td>Motivate to Share and encourage Loyalty</td>
</tr>
</tbody>
</table>


Mapping would help sellers evaluate the journey by understanding the following: [14]

- The points where customers face obstacles
- The message which is difficult to understand or the information that is complicated for customers
- The points in the journey where consumers lose engagement or interest in the brand or the product and would prefer to not engage with this brand
- The points that provide consumers most satisfaction to motivate them to move to the next steps

3.3 Designing the journey

Customers’ experiences about a brand are created when they interact with its touchpoints, and that experience plays an important role in whether they would consider that brand for their ultimate purchase decision. The goal of designing the customer journey is to design the customer experience a brand intends to provide to its customers. Today, it is not enough to only be reactive to consumers during their purchase decision journey but be proactive to influence their decisions by designing it. [8] Good journeys have potential to provide a competitive advantage to the sellers. Companies who optimize their journeys can not only acquire and retain more customers but also provide value for themselves and the customers.
Bellos and Kavadias (2011) state that “the optimal design decisions depend non-monotonically on two important parameters: i) the variability of the experiential outcome from touchpoint interactions and ii) the underlying interdependencies between tasks, which may give rise to correlated experiential outcomes across touchpoints.” Customer experience is partly a function of variability caused by multiple ways a customer can move across the touchpoints and different quality of experience each touchpoint provides to a customer. It is desirable for the sellers to present a consistent customer experience by optimally designing the journey across touchpoints in order to positively determine the journey outcome. A well-designed journey will consistently, effectively and clearly communicate the brand across all the touchpoints throughout the journey.

3.4 Considerations for designing the customer journey

1) Identify key touchpoints

For a particular product, a seller needs to identify what are the primary touchpoints that its existing and potential customers use during a particular stage of their purchase decision journey. Sellers can choose for several touchpoints these days. Some of the following criteria will help select the touchpoints over others. The effectiveness of the touchpoint, which depends on the reach of the touchpoints among the customer segments that seller intends to target, and the size of the segments interacting with those touchpoints. The touchpoints also need to be assessed by their usage on various devices if the seller wants to interface with the customer on particular type of device – the customers are more likely to use and get influenced by different touchpoints on mobile, desktop and tablets. For example, a certain type of content and touchpoints are more suitable for mobiles such as the Ads and in-App notifications triggered based on geolocation.

In most stages, there are multiple touchpoints that the customers use to achieve their goals. Also, many of the same touchpoints could be used in more than one stages of the journey. How can a seller prioritize the touchpoints to be used? Attribution of sales to various touchpoints using data analytics can provide insights into the contribution of each touchpoint. Tools such as Google Analytics Solutions use data-driven machine learning approach to determine the impact of each touchpoint in the customer journey. However, the caveat is that the insights from attribution should be derived not completely on the face value of the attribution percentage that the tool provides, because the impact of many non-digital touchpoints responsible for influencing customers’ purchase decisions such as billboards cannot be measured. Also, the impact of some of the digital touchpoints such as Display Advertisements and social media Advertisements, which might create repeated visual impact but is not clicked by the customer might not be attributable but does have a role in the purchase. This touchpoint might indirectly create impact; however, some other
touchpoint might get higher attribution despite its lesser role in creating the influence on customer’s purchase decision.

2) Identify the role of the touchpoint in each stage of the Customer Purchase Journey

The questions to ask about the role of touchpoints with respect to the customer journey stages are:

- What specific things are sellers doing at each touchpoint?
- Is a particular touchpoint answering customers’ questions and addressing concerns? Are they targeted to desired customer segments?
- Are the touchpoints going beyond addressing the needs and delighting the customer's unstated needs that the competitors are not? [16] The role of social touchpoint such as Facebook Advertisements is to trigger the need for purchase and spread brand awareness in the exploration phase, direct the online customers looking for discounts and deals to the sellers’ website to motivate them in their exploration phase and provide social proof of the likeability of the brand and recommendation (measured by the Likes and Shares) in the evaluation phase and post-purchase loyalty phase. Identifying why customers are using particular touchpoints during a particular stage of the purchase journey helps sellers prioritize the touchpoints based on how efficiently they meet the desired objectives of the sellers and goals of the customers. It also helps identify the content needs of the touchpoints ad adapt the message/creative to those touchpoints.

3) Identify what should be the next action of the customer

By predicting the next move of the customer using data and the stage of the purchase decision journey the customer is in, the seller can guide the customer to the next relevant touchpoints. The seller can also identify if there are specific hindrances in the journey from one stage to the next that led to disengagement of the customers’ or dissatisfaction that calls for use of more expensive steps to retain or attract the customer back. [16]

For example, if data about a particular customer shows that she is evaluating options for purchasing a coffee machine, the seller could guide her to the review video of the coffee machine of its brand, or to a comparative chart of various coffee machines by variants of its brand or all comparable brands. These touchpoints could in-turn guiding them further to the online or offline channels for purchase, creating a seamless journey unlike the broken journeys we see online today.

Example: YouTube (people seeing a music video) → banner ad (with clear call to action) → clicking leads them to Retailers YouTube channel (where seller describes in detail the product’s value proposition)
For motivating the customer’s potential next steps through the touchpoints, the seller needs to determine:

- Which stage of the journey customers are in, what is the relevant next touchpoint that the seller wants them to take based on sellers pre-designed customer journey for that product, and what is the ‘call to action’ that a maximum number of the target customers would respond to.

So, by deciding upfront how various potential journeys might look like for their customers, sellers can lead the target customers to the specific content and touchpoints that motivate the customers to progress into the next stages of their journeys.

Following are some of the pitfalls that the journey designers need to consider: Though the prioritization of touchpoints is important, planning such that the customers who do not land on those top predicted touchpoints do not make into the end of the journey would be a rigid journey design. So, having a minimum viable ecosystem of touchpoints in place is important for pulling into the journey maximum set of target customers. Also, in order to shorten the customer journeys to bring the customers to purchase phase faster, the designers could make the mistake of cutting short the journeys at wrong stages, losing the customers at those stages. So, the call to actions on all the touchpoints should be considered as bridges in the customer journeys and designers should check for any shortcuts in their designs.

4) **Incorporate the understanding of various customer personas**

“*Personas are the starting point,*” says Michael Hinshaw, customer experience strategist and president of MCorpCX. “*Because a journey map is the story of a customer*” experience. [17] Customer personas help develop a user-centric design of journeys and help think about the possible variations of customer journeys comprehensively. As real research and customer interviews are conducted to identify personas, they provide an additional basis to identify various possible touchpoints particular customer persona might be more inclined to use while interacting with the company over the other touchpoints which company might originally expect of the customers.

To understand how the touchpoints and the customer journey differs for different personas, refer to the following example in which Persona 1 purchases phone for personal use as the desire for new technology triggers her compared to Persona 2 who purchases phones for her corporate employees for their (primarily) office and personal use. Though example uses a high level understanding of the two personas, sellers should look at the potential customer journeys of all major personas more granularly.

High-level **Persona 1** description (figure 10) – purchases for personal use, digital savvy, frequently purchase new phones, performs research and comparison on online channels as well as prefers to have the touch-and-feel of the phone in-store (Omni-channel research)
High-level **Persona 2** description (figure 11) – purchases for corporate employees, does moderate online research but depends mainly on the agents and dealers for a corporate deal, uses the corporate channel of purchase, bulk purchases new phones, may require to replace and repurchase phones in future.

**3. 5 Touchpoints - an integrated view**

A media company was facing the problem of retaining its customers and acquiring new ones. The company executives decided to improve the customer experience to improve the situation. But when they measured the customer experience, they found that score for each of their touchpoint was strong 90 percent. They concluded that their service was great. However, a focus group in their company contrarily reported that the customers left their company because of poor customer service. Upon deeper analysis, the executives understood that the overall satisfaction that was a result of cumulative experiences of customers across touchpoints throughout their journey with that company was low, although each touchpoint individually performed well. If a customer interacted with four touchpoints through their journey, each with a 90 percent chance of going well, the overall customer experience in their journey fell by average 40 percent. So though
touchpoints performed well, it did not ensure great end to end performance. A supporting evidence was found by McKinsey & Co. in one of their surveys – “the gap on customer satisfaction between the top- and bottom-quartile companies on journey performance was 50 percent wider than the gap between the top- and bottom-quartile companies on touchpoint performance.” [18] It is evident that more than the performance of touchpoints, the performance of journeys is much more effective at driving business outcomes such as churn and repeat purchases, and thus revenue. It is because journey performance is strongly correlated with the total customer satisfaction. The following exhibit shows results of a survey that explain the correlation between customer satisfaction and their willingness to recommend the brand with touchpoints Vs journeys. (figure 12) [18]

Figure 12, Correlation of customer satisfaction and Willingness to recommend with touchpoints Vs journeys [18]

3.6 Important touchpoints for Ecommerce

3.6.1 Online Advertisements

- **Search (organic)**

Most consumers use search engines to find what they need. Search engine (e.g. Google, Baidu, Yahoo) results are based on content available in other touchpoints after a consumer made a search query. The results
of this search are based on the rankings of search engine algorithm. Organic search is the critical touchpoint for every Customer Decision Journey (CDJ).

Presence in the Search Engine Result Page (SERP) of the results is a powerful way to trigger customers and to create brand awareness and initial consideration. Sellers can influence these results by doing search behavior analysis. It needs the right content in the right places of the CDJ to work efficiency. So, accordingly produce relevant content and ensure that they deployed it in the key touchpoint that has been identified. The keywords and phrases that customers use while searching particular information about a product category or brand need to be identified from sources such as Google trends, etc. because search engine displays the results of the search query based on keyword matching.

- **Search (Paid)**

  When users put search query in the search engine such as Google, the initial few search results associated with that search are sponsored and indicated as ads. These web pages pay to appear as top results of the search query. [19] Paid search works very well in the trigger phase. If sellers do not have the content in the right places, the invested money in paid search will be wasted (not efficient) because it will drive people to a touchpoint that does not have relevant content and customers will not convert.

- **Display**

  Display advertisements are the form of advertisements that appear on websites or search engine results pages such as publishers (example New York Times) or social media (example Facebook). [20] Display advertisements are also called Banner Ads. Various forms of Banner Ads are Traditional Banner Units, Mobile Banner Units and In-Banner Video Units. These ads are scalable and easy to deploy, and results can be measured to a high degree. Banner Ad campaigns are measured using the following metrics: Clicks, impressions, conversions, spend and video views.

- **Video**

  Video advertisements are run either before, during or after the online video content. It enables advertisers to connect their message to an engaged audience. Also, as the draw of viewers towards internet TV is increasing, and the video ads are spanning across four screens – computers, smartphones, tablets and connected TVs, the potential of these type of ads to outperform other types of online ads for branding is increasing. Watching the video online has gone from a niche activity to mainstream and Video click rates are far higher than image format ads. Users click on video ads about five times as often as they do image ads. [21] Major performance metrics for video ads are 100/75/50/25 % Completed Rate, Impressions, Clicks, and Click-to-rate (CTR).
Mobile advertising involves displaying text, graphic images and animated ads on data-enabled mobile devices. These have transitioned from being a small format of the web display ads to being a separate category of ads. Mobile advertising is becoming a more crucial part of the marketing mix for brands of all sizes, as companies develop more sophisticated methods to engage consumers, influence their buying behavior, and even complete sales via a mobile device.

The major types of mobile advertisement as are: Click-to-download ads - the user is directed to the Appstore or Google Play, Click-to-call ads - the user calls to a phone number after clicking the button, image text and banner ads - a click opens users’ browser and re-directs them to a page, and push notification – messages that pop up looking like updates in the notification area of the devices. They are generally pushed from the Apps that are installed on the devices. [22]

The major types of mobiles ad formats for Mobile Web are Optimized - in which the advertisement is responsive to the device size. In this, the format of the Ads reformats itself to suit the screen dimensions. Non-Optimized - some websites still have their site layout independent of the screen size of the mobile and the advertisements are called. Another common type is In-App mobile ads that could include the banners and expandable ads. The metrics used for measuring the performance of these types of ads are same as those for display ads.

Some of the newer formats of mobile ads are as follows:

- Interstitials ads that appear intermittently during the app usage and are larger than other ads to allow for more engagement of the mobile users. [23]
- Native mobile ads that coherence in formatting and content with the hosting app or site.
- Deep Linking that are actually the URLs of the mobile apps. These ads have an embedded link called Uniform Resource Identifier (URI) URI which an address for an app or a location within the app, and is used from bringing the potential customer to a particular location in an app such as a product page of a particular item or a description page of a deal. [23]
- Beacons are external devices that connect to the app using device’s Bluetooth technology. They detect nearby smart devices and send them media such as ads, coupons, location-based deals, or additional product information.

Native

Generally speaking, native advertising is advertising that fits ‘form and function’ into the context of the webpages on which it appears. For example, the advertisements that appear as the sponsored feed or news
on Facebook along with the content posted by the regular user and Facebook friends are the Facebook native ads. In one of its reports, the Interactive Advertising Bureau (IAB), the primary organization responsible for developing ad industry standards detailed six categories of native ads: [24]

- In-Feed Ad Units: These units appear as normal content feeds of the website, as if the content is written to match the surrounding stories of the website.

- Search Ads: Paid search ads are also considered native ads because of their nature that they appear as the part of the main content of search engine, but are generally above the organic search results.

- Recommendation Widgets: These form of native ads are also part of the content like In-Feed Ads, however, they are differentiated using external references or indicators such as “Elsewhere from around the web”, “You may have missed”, etc. like suggestions.

- Promoted Listings: These are similar to the paid ads in that they appear like the main content of the website on which they are hosted but are sponsored. For example, on the food delivery ordering website such as Yelp, some of the restaurants appear on the top of the search results along with other search results below them.

- In-Ad (IAB Standard): This ad fits IAB standard size container found outside the feed, containing "...contextually relevant content within the ad, links to an offsite page, has been sold with a guaranteed placement, and is measured on brand metrics such as interaction and brand lift."

- **Social Media**

Advertisements on the social media intent to display content that the social media consumers would interact with (share, comment, like, retweet, play, click and take subsequent action) on their social networks. The channel partners gather the data about all these interactions to evaluate high-value customers for a particular type of ad and target them and other similar prospects. The key players in this category of ad publishers are Facebook, Instagram, Snapchat, Twitter, Pinterest and LinkedIn.

- **Blogs and News (Inbound Marketing)**

Sellers reach the potential customers who might have high interest in a particular category of product through the blogs, articles, news or reports about these categories or topics related to these categories. For example, if a customers need a new smartwatch, they generally want to compare the brands or read reviews on blogs by technology product experts. These articles could include recommendations, features and finally links to websites of smartwatches that sellers want to attract customers too. Sellers often partner with blog writers and reviews who have high number of followers to get them to write sponsored articles. E-commerce
websites such as Amazon often drive customer by sponsoring articles such that might be about ‘5 coolest products on Amazon you don’t know about’.

HubSpot, one of the leading companies that provide inbound marketing tool and training to the sellers, has defined a four-step methodology for inbound marketing. First, attract right people, ones who are more probable of becoming customers. This could be done by writing educational blogs that answer the potential questions of the customers, placing this type of content on the places where people are more likely to search it, and placing high quality and valuable content on social media to get it shared. After people visit the website through the links, the second step is to convert them into leads by personalizing the engagement from there on by understanding their requirements through forms, chatting or exchanging messages with them, providing them easy and quick ways to reach out to business for any queries and systematically documenting the interaction with them using the Customer Resource Management tools. Once the business has leads, the third step is to convert them into customers. In this step, identifying the best leads and nurturing them is important. Not all leads are qualified for the investment required to convert them, so predictively scoring the more qualified lead is recommended. Further, understanding the stage of purchase they are in and communicating with them by providing focused and relevant information. Finally, in the fourth step, the framework recommends staying connected with the customers after converting them by sending out the content smartly based on their product lifecycle stage and customizing the conversations by keeping track of comments and requests. [25]

![Figure 13, Method for inbound marketing](image)

3.6.2 E-commerce website

E-Commerce sites encompass a full CDJ, hence they have a role in all the phases of the purchase decision journey. Many people start their journeys from the search page of the e-commerce website or the catalogue, and end it with the purchase on that website and become repeat purchasers and sometimes loyal members, which is the case with Amazon Prime members. Various touchpoints on the website correspond to stages of the CDJ such as: Search page -> Exploration stage, Product Recommendations -> Trigger, Product Page -> Inform

The sites themselves are massive touchpoint that the customers visit, therefore e-Commerce sites have wide reach and can serve as media platforms to buy media and do advertising. E-commerce sites span across
different ecosystems such as mobile, desktop, r-reader (for example, Amazon.com, Kindle, Amazon Mobile app), and so they attract huge consumer base and are usually in the top 10 most visited sites of the respective country. Hence the media deployed in these entities get a broader reach. In addition to the reach, as customers are in the mindset for shopping, the CDJ is shortened and probability for call to action is stronger to move the shopper towards the purchase and next stages. Finally, e-commerce sites profile their consumers and their shopper behavior, that enables targeted and customized communication possibilities (e.g. Amazon Media Group deploys the media banners to different shopper personas based on lifestyle such as vegan, young professional etc.).

Some of the strategies that e-Commerce websites employ following strategies to increase their conversion:

a) For bannering:

- **Seamless execution of banners across platforms** exposing customers to multiple and consistent exposures. For example, a customer who has seen Nike ad on amazon.com sees the same ad on Amazon mobile website or kindle

- **Rich banners on static ads**: For example, ratings and reviews are embedded in banners, directly add to cart or wish list button are embedded in banners, use of expandable ads that extend the banner when hovered over that shows complete product detail or category pages are enriched with how-to videos and other relevant content

- **Focus on KPIs and objectives** are measured that are linked to business strategy (for example, cost per action)

b) Next to the bannering there are other ways to trigger consumers within e-Commerce sites:

- Internal search: product-related search is usually higher in volume than the amount of search in search engines.

- Deal of the day, Top sellers

- People who bought also bought, products recently viewed

- Bundle suggestions. Example, accessories such as sleeve, headphones, wireless charger, etc. with tablets

- Buying/gift guides

- **Display network**

**Search engine display network**: Next to paid and organic search, search engines also own display networks. For example, the basket of placement in websites or high traffic portals that the ads are deployed to through rich or static display banners. Display networks of search engines work almost exactly as display networks of media agencies. One key advantage of display networks that are run by search engines is their
power of retargeting. Search engines, based on the queries of their visitors, can show the display banners on the websites that are relevant to the shopper at the right place at the right time.

**E-Commerce Site Display networks:** In addition to displaying ads on their owned web entities, e-Commerce sites also own display networks. An example is Amazon media group display network.

To summarize, customers interact with several touchpoints in the purchase decision journeys. The experience of the customers and the success of sellers depend on how well the journeys are designed. Further, it is important for companies to know various touchpoints on which they can meet their customers and should map the potential journeys based on which companies should plan their sales and marketing efforts. Next, the thesis discussed the technology and working of the above category of touchpoints in the next chapter.
4. Technologies Underlying Touchpoints

4.1 Recommendation Systems

Recommendation Systems use algorithms that predict the customer preference for any item on the website and recommend those items to the customers. [26] They are meant to be profitable for the customers as they reduce search cost and time, minimize risk while maximizing profits. [27]

Recommendation systems help improve the sales on an e-Commerce website in several ways throughout various steps of the customer journey. Most importantly, it has the potential to convert a visitor into a customer. When a customer is in the product discovery phase and lands on the e-Commerce website, the landing page itself, when personalized to display and recommend relevant items, drastically improves customer experience and reduces effort of searching the desired product. During the evaluation phase of the journey, when customer understands the product features in more detail (generally on the Product Detail Page), the recommender system can increase sales by cross selling and up selling by bundling the product with accessories or peripherals which are based on the information about past purchases of customers who bought multiple related items in a single purchase. It also improves customer loyalty due to the value addition made by websites that they can do by learning about the customers and providing personalized recommendations and communications consisting of recommended items.

Larger companies are heavily investing to improve the personalization of their products. Recommendation systems are a big part of that effort. For example, every year Netflix makes $150 million investment to improve their content recommendations as it has very small amount of time to convince a user to watch a particular piece. Larger e-Commerce companies make recommendation engines in-house. However, may e-Commerce players prefer plug and play solutions slightly modified for their products, whereas smaller players use recommendation software as a service. [58] Some examples of recommendation products are Adobe Target, Recommendations API offered by Microsoft. Companies, and recommendation engines offered by Barilliance and Strands. [59]

4.2 Applications of recommendation systems in e-Commerce website

Some examples of the applications of recommendation systems on the e-Commerce website such as Amazon.com and eBay are:

Different types of recommendations on Amazon.com (Figure 14 and Figure 15):
Figure 14. Recommendations on the landing page of the e-Commerce website [28]

Inspired by your browsing history

Frequently bought together

Customers who viewed this item also viewed

Inspired by your shopping trends

Figure 15, Recommendations on the Product Detail Page [28]

Figure 16 shows a feedback recommender system on eBay.com. Buyers and sellers provide feedback on this system after doing business with each other. The feedback includes rating and comments. This recommendation can be pulled up by either party to verify the other before engaging.
4.3 Approaches to recommendation systems

Recommendation systems are one of the subclasses of the information filtering system. They are filtering problems because recommending an item to the customer is like discarding the not useful or unimportant items. [30] Typically recommendations have been produced in two ways – Collaborative and Content-based filtering. Often the above-mentioned filtering approaches are combined to form Hybrid recommender systems.

4.3.1 Collaborative filtering: This mechanism of filtering recommends items to customers based on the preferences of similar customers. The algorithms of these systems collect behavior data of the customers, their activities, and their preferences, ratings and/or remarks to predict what similar customers might like. [31] The key idea is to find the notion of similarity between customers of an item.

The main advantage of these systems is that they do not depend on the understanding and analysis of the products or services it considers for recommending. The disadvantages are that these systems depend on the large extent of customer data to make useful recommendations, and so recommendation systems of e-Commerce websites that use collaborative filtering approach might suffer from ‘cold start’ due to the lack of data in the beginning when there are fewer customers. Another disadvantage of these systems is that only a few out of several thousand products on the
e-Commerce websites are liked or rated by the customers, also termed as 'sparsity', so the number of products that the system can recommend to other customers is limited and most of the products remain out of or at the bottom of the recommendation list.

Collaborative recommendation systems are of two main categories: memory based and model-based.

- **Memory-Based**: It is called memory-based approach because it consists of storing all the customers’ information and predicting the rating or preference of the target customer by retrieving the information of customers similar to the target customer. This approach has three steps. In the first step, the algorithm evaluates the similarity between the target customer and other customers with respect to the item under consideration. Next, using the similarity as weight, the algorithm finds the nearest neighbors (similar customers). Finally, it predicts the preference of the target customer using similarity weights. (Figure 17)

The weighted mean is one of the algorithms generally used in the memory based approach.

\[ \hat{v}_{aj} = k \sum_{i=1}^{m} w(a, i)v_{ij} \quad k = 1/ \sum_{i=1}^{m} w(a, i) \]

In the above equation, a is the customer to whom the item is to be recommended any object o

The equation finds a rating to predict how likely customer a is to like object o. It basically combines normalized rating of users. Here the sum is taken over all users. As not all users contribute equally to the average, the weight controls the inference of all customers on the prediction. This weight is related to the similarity between u, and u, the more the contribution user u, will make in predicting the preference of user u,. Here, k is a step for normalization that takes the sum of weights of all users.
To determine the weight function $w$, there are many approaches. One of the widely used approaches is Pearson Correlation Coefficient which measures the linear relationship between two variables. For example, whether two customers both tend to give lower or higher rating to a similar item. [32]

- **Model-based**: Memory-based approach is generally real time and so not very fast or scalable because of the extensive datasets of continuously generated customer data. A model-based approach attempts to overcome these problems. In model-based approach, models are build based on the partial data set instead of the complete data set for making recommendation each time. [33]

There are a number of techniques to build models. Some techniques are:

- **Probability problems**: It is a probability problem when a task of recommending an item is carried out by predicting the probability of a preference rating taking a particular value. Clustering models such as K-means and Bayesian network are some examples of probability techniques for the model-based filtering.

- **Enhanced memory-based technique**: This is the enhancement of the concept discussed above in the memory based approach which predicts the customer’s preference rating for an object by using the similarity between customers as weights. The idea is to build the model based on a particularly limited number of data points enough to provide the desired accuracy and store the similarity measure to predict customer rating.

4.3.2 **Content-based filtering**: This approach is based on the item similarity rather than customer similarity. In this system, the products or items are identified by keywords. Also, the system makes customer profiles by considering the products that customers like and then recommends products that are similar to the ones customers like. The basic process followed in this approach is of recommending the best-matched products with the ones rated in the past.

The architecture of the content based filtering system would look as shown in Figure 18. The item's information is the input to the system (read Doc Source) which is classified by the binary classifier. The binary classifier has data about the interests of the customers which the system keeps track of. There is a utility function that works as a decision module. It evaluates which items will pass through the binary classifier based on its utility based decision-making criteria. This utility function sets a threshold value that the items need to measure over to be recommended to the customers. It is not based on the ranking of items but measures utility absolutely. The initiator module takes the input such as customer’s keywords, etc. which are input along with the attributes of the initiator.
Content-based recommendation approach would do reasonably well in practical applications and could be a starting point because it is dependent on a limited number of customer ratings. It can recommend items based on its similarity to other items even though it has not been rated, so there is less dependency on all the items on customer ratings.

One of the most used techniques for content-based filtering is similar to TF-IDF (Term Frequency – Inverse Document Frequency). The algorithm abstracts the text-based attributes of the products in the database of the e-Commerce website. The product is weighed based on the frequency in which the customer’s searched keyword for product search matches the product attributes in the database. [35] However, nowadays sophisticated Machine Learning techniques including decision trees, clustering, Bayes classification, etc. are taking over to predict the likelihood of customer liking a product. [26]

4.3.3 Hybrid recommender system: Hybrid approach generally combines collaborative filtering approach with content-based filtering. It might also unify demographic and knowledge-based approaches in the hybrid system’s model. The demographic approach includes demographic attributes of the customers, and this information could be very useful in predicting customers’ product preferences when they vary more by locations. The knowledge-based approach is a more advanced one as it infers the customers’ need and might use the functional knowledge of the products that is abstractable by the algorithm to recommend better products that meet those inferred customer needs. [26]

Several approaches are used for hybridizing the techniques such as:
- The weighted approach in which the preference score of a recommended product is a weighted sum of the preference scores of various products. The fundamental logic behind this approach is to optimize the weights for set of products using the past data [31]
- Switching approach in which the recommendation system switch between results of different techniques and chooses one to recommend. If same product recommendation is resulted by using two different techniques, say both content-based and collaborative filtering, then those products are ranked higher

- Mixed recommendation approach in which the top-ranked results generated from various approaches are presented together

- Meta-level approach in which one of the techniques first generates a model that becomes an input for the next technique

The studies have empirically shown that the unification of approaches in the hybrid system produces more precise product recommendations compared to individual approaches, and also overcomes problems in the individual approaches discussed before such as ‘cold start’ and ‘sparsity’. The problems with the hybrid system are that it relies on the quality of integration of various approaches such that the limitations of combined approaches are not inherited. Another issue is to update the recommendation system by taking updated inputs of customer rating for the products. The key to the problem of making a model adaptable is to use Machine Learning algorithms to incrementally train the model using new inputs. [31]

4.4 E-Commerce Product Search

Numerous purchase journeys start with the product search when the customer inputs a query in the search bar of the e-Commerce website. Thus, optimizing the search results that are the result of product ranking algorithm saves customers time and effort of discovering the desired product from several hundred products on the e-Commerce website. Ranking algorithm scores the products in the database and presented to the customers in the order most relevant to the user. Ranking algorithm factors many features for ranking the products. The types of features are: search query features – the product description keywords in the search text entered by the customer, user behavior or interaction features such as revenue from a particular customer, clickthroughs, number of add-to-carts, etc., product features such as price, brand, variations, ratings, and reviews.

Product search is an information retrieval process in which the objective function is to retrieve the product in an order that maximizes conversion (product purchase). For a particular search query, the objective function combines the various features such as product, query and user behavior features with proper values of weights assigned to them to construct an optimized ranking of the products. That is, the probability that a particular product is relevant to the customer is a function of the above-mentioned features. Assigning correct weights to various features in the function is a challenging task where Machine Learning algorithms can contribute.
Following features (not an exhaustive list) might be used as inputs for the search algorithm for e-Commerce websites [37]:

**Product features:**

Price: Algorithm makes use of the price to score a product

Title: Product titles are matched with the keywords used by the customers during the search

Image: Having a high definition image of the product that could be zoomed counts in the algorithm. However, animated images that show 360-degree view or how to use videos may still not be the part of the algorithm for search results

Variations: The variants of the same products such as sizes or colors have more importance in the algorithm because multiple customers might look out for variations of the product which they searched

**Customer Behavior features:**

Type of membership: For example, on Amazon.com amazon prime members (subscribers of their loyalty program) might have a different result compared to the nonprime members. The prime members might be displayed items that are eligible for prime delivery at a higher rank

Purchase history: Various parameters of past purchase might form key inputs about customers behavior

**Key business metrics:**

CTR (Click through rate): This feature could be more important as a feedback feature. If some products are more interacted with by clicking and proceeding through stages of the customer journey, they would be considered more important by the algorithm.

Sales Rank: If a product has sold more, it has higher probability of selling more

Learning to Rank (LETOR) method has been used as a widely used method not only for web search but also for e-Commerce search.

**Other features:**

Stock-outs: If products are stocked out, they will be ranked below in the search results

Seller rating: Products sold by poorly rated sellers will be ranked low

The above features are the input to the algorithms of Machine Learning model that is discussed in the next chapter.
4.5 Advertising

4.5.1 Programmatic Advertising

The advertising ecosystem has changed from traditional manual transactions of advertisements between advertisers and publishers to programmatic. Traditionally, buying and selling were made by salespeople who made Request for Proposals (RFPs) and negotiated prices. Now, the programmatic technology has enabled automatic and thus efficient tractions in the space of digital advertising. Before getting into the process, it is important to understand the main players, components and key technologies in the ecosystem. [38]

![Programmatic Advertising Ecosystem](image)

Figure 19, Representation of Programmatic Advertising Ecosystem [38]

4.5.2 Key players and components of the advertising ecosystem

The advertising ecosystem (Figure 19) has two sides – buy and sell side. The buy side mainly consists of advertisers and/or ad agencies on behalf of advertisers. The sell-side consists of publishers (entities such as New York Times) who publish ads on their websites and apps, Ad networks who aggregate advertisers space inventory and sell-side platform.

**Advertisers:** Advertisers want to spend on marketing in order to grow their revenue. Some of the examples of advertisers are Bank of America and Samsung. Nowadays they have customer data like never before and they are trying to make digital marketing more efficient by knowing their customers better. They are enriching their existing customer data with third party data from partners to increase their advertising accuracy and reach. Companies are working on attribution – measuring which channel and marketing effort has contributed how much to the sales. Overall, their attention to improving advertising performance has drastically increased. The recent trends show that they are using rich media such as videos in their marketing because of the reach of smartphones and internet. Also, availability of data and smart advertising
is creating a level playing field. Some large players are shifting all the operations of advertising from partnerships model to propriety or in-house model so as to have marketing as strategic competency.

**Agency:** Agencies specializes in helping the advertisers with their creatives, managing the bidding process in the auctions, and budget allocation. Some examples are Omnicom and Publicis. Agencies are facing market pressure from the larger clients as they are shifting the advertising operations in-house for getting more control and reducing costs.

**DSPs (Demand Side Platforms):** These are the platforms used by advertisers and media buying agencies helping advertisers to perform real-time and automated media buying using real-time bidding. Examples are Google Bid Manager is Adobe Media Optimizer. DSPs are central to advertising campaign management because they facilitate campaign design, implementation, and optimization by allowing marketers to set limits for bids and parameters for campaigns and target customers. As many DSPs interface directly with the advertisers, the need for people handling media transactions and agencies is decreasing.

**Publishers:** Publishers are generally business who have an online presence through website and apps. They have space (also called inventory) where they can serve ads to their visitors. Most publishers have a revenue model of earning through ads served such as Facebook, Yahoo, etc. They try to generate traffic on their websites and serve relevant ads to the target customer segments. The cost of the inventory sold by publishers is based on the amount of customer traffic they get. They have their premium inventory/space (the more visible space on the website that catches majority eyeballs) and remaining space which is below the fold (website needs to be scrolled down to see it) or at the bottom of the page. Traditionally all the inventory used to be sold by the publishers through the ad agencies who in turn would negotiate deals with the marketing teams of the advertisers. The programmatic advertising technology has changed the way in which the remnant inventory is sold. Though premium inventory is being negotiated traditionally today, publishers are increasingly moving towards the programmatic technology.

One of the major advantages of programmatic technology is the data it offers about the post-advertising/post campaign performance which lets the publishers make an informed decision about how to monetize the space more efficiently and helps advertisers use customers’ behavioral data to target them better with relevant content and customization.

**Supply Side Platform (SSP):** It is a platform used by publishers to sell the ad inventory on their websites using programmatic technology and enable routing to demand side including ad exchanges and DSPs. Some examples are Pubmatic and Google Admeld. The algorithm in SSP is designed to manage ad yield of publisher ad inventory. SSP’s function is also to ensure the best price for supplier inventory during
transaction or bidding. Using SSPs, publishers can control the transaction to a great extent such as setting a minimum price for their ad space, and buyers whom they want to sell to. They can also dynamically decide to switch between channels through which they want to sell their inventory – for example, from direct selling in the market to Real Time Bidding. The process that publishers follow in the SSP to make a transaction deal is to enter details of their desired criteria in the SSP interface such as the type of space, media format, customer segments, etc. A deal ID is created and send to the DSPs or an agency/advertiser. Some SSPs are enabling a private marketplace for some selected agencies/advertisers to provide them priority access and first chance at the premium inventory. [39]

**Industry bodies:** One of the primary industry body is a global trade organization called IAB (Interactive Advertising Bureau) consisting of approximately 650 members. It includes companies of stakeholders in the digital advertising ecosystem. IAB helps develop standards, conducts and publishes research and best practices on related topics. In addition, it offers training and education to its members through classes, seminars, etc. Other such industry bodies include TAG (Trustworthy Accountability Group) and ANA (Association of National Advertisers). They also work on topics including fraud detection, digital security, and privacy. [39]

**Specialised Industry vendors:** These include third-party data providers who collect customer data from multiple sources such as credit card companies, package the data and sell it. This data is used by companies such as Amazon to enrich their propriety data with missing information.

**DMP (Data Management Platform):** This is a complex system that consolidates data from multiple sources both offline and online by plugging into different data sources, and analyses that data. It helps unify all the data for a particular customer to provide a single view. DMP’s have interface that helps send data to multiple channel partners so that they can activate their campaigns. Examples of DMPs are Adobe Audience Manager and Google Doubleclick Audience Center).

Publishers, advertisers, and agencies use DMP in different ways. Publishers use the data and analyze it to accordingly manage the inventory of space they have for offering to advertisers efficiently. When publishers augment their various types of inventory spaces with the data of customer response to ads in those spaces, it can improve the value of publishers inventory. One of the primary ways of analyzing the data is to create customer segments based on their ad interactions and visits. This is done either manually by creating groups of common interests among customers or using look-alike modeling where the application of Machine Learning models can learn to segment and target new customers. Advertisers and agencies use DMP to create and manage datasets for activating the campaigns from clients on different channels. [39]
DMP has three step process:

**Aggregate**: DMP (Figure 20) helps identify most useful data from disparate sources such as ERP, CRM, customer subscription data, and analytics systems. It also provides insight into the audiences (target customers) by helping track data from multiple devices.

**Segment and find target customers**: First, the DMP helps identify important indicators and insights from the gathered data. Second, it helps organize the data to create customer profiles based on a variety of customer-related information including device, purchasing history, contextual, interest, location, etc. Third, it helps capture the results of previous campaigns and optimize new ones. Finally, it helps generate customer segments, turning data into a set of customer who can be targeted.

![Figure 20, Representation of Adobe’s Data Management Platform](image)
**Action:** DMP has links with marketing tools and channels, for example, some companies integrate DMP with ad servers, email tools, and interfaces of publishers such as Facebook, Pinterest, etc. It operations of inward and outward data flow take place in real time resulting in quick insights. Segments are updated in the DMP after the feedback from the campaigns.

The step two above consists of data analysis that helps companies know more and granularly about their customers and find the look-alike customers is the most fertile ground for applying Machine Learning because of its potential to automate the task of analysis and because of the learning required with each new customer and each new information about the existing customer that requires updates to the segmentation. This is discussed further in the next chapter.

**4.5.3 Technologies enabling advertising**

**Cookies:** Also called HTTP cookie, it is a piece of data that the website (example, Amazon.com) sends to the customer’s web browser (example, Google Chrome webpage). After the cookie is sent, the browser can fetch that data from the server and update the cookie. The intent behind designing cookies was to keep track of the state of users browsing on a particular website. With the use of cookie, e-Commerce websites can maintain the state of their customers’ information from previous browsing events which buttons they clicked on, how much was their browsing session time, their time on various pages, products they added to cart, their information such as name, delivery address and passwords for a website, etc. There are two types of cookies – first, a persistent cookie that has a long-term expiry date which could span months or years and is designed to maintain a long-term state of customers’ activities. Second, session cookie that is temporary and maintains a state of browsing only till the browsing session ends. [39]

When the domain attribute of a cookie is same as the server which sets it, it is called the first-party cookie. But websites have some images or Advertisements which could be contents of other domains. This content from other domains when read and retrieved called third-party cookies. Generally, first-party cookies can track activities within the same domain, whereas the third party cookies can track activities across other domains. For example, the advertisement on Amazon.com if has third-party cookie can track customers activities across other websites. Because of the privacy concerns, most web browsers servers are starting to block third-party cookies. [41]

A cookie is useful because it identifies customers, stores their information about various website visits – a compilation of which can be used for can be analyzed for identifying customer interests and also categorize the customer into the segment and serve them relevant content or ads. It is important for the advertising ecosystem that the cookie data is synchronized across various domains.
**Pixels**: A pixel, also called as a stage, is a piece of HTML code or JavaScript in the form of a minuscule image that could be embedded in the webpage or in the marketing email sent to the customers. This pixel tracks the customer’s information such as their activities of websites, the information about customer’s interaction with the marketing emails or ads. Pixels from a particular sender identifies whether a cookie from the sender was set at the customer’s browser. When a customer opens an email or any interface with a pixel, the pixel enables its sender website to place a third party cookie at the customer end. Digital marketing activities require managing these tags for efficient information tracking, and various systems such as Google or Adobe tag manager does this task. [39]

**Ad servers**: Ad servers store and deliver online ads. Other functions of ad servers are to manage campaign traffic, optimize check ad performance, monitor and report, and campaign analysis. There are two categories of servers – publishers’ and advertisers’. Publishers are parties that publish ads on their website such as New York Times. Publisher ad servers do campaign management and prioritization of campaigns from several clients. Remote servers (third-party servers) distribute ads across multiple domains owner by several independent publishers. Another type of servers called local servers are owned and operated by the single publisher to publish ads on its own domain, providing them better control. These servers Advertisers buy advertising space from various websites, apps, etc. Ad servers for advertisers help them manage ad distribution across myriad media and publishers, track the purchase of media, send and monitor ads, and report performance. [42]

**Ad networks**: These are brokers who aggregate the ad space from various publishers and sell this inventory to the advertisers with a markup.

**Ad exchanges**: This is a marketplace where with the help of Real-Time Bidding (RTB), publishers sell the advertising space (on their website/app) or inventory to the advertisers at a maximum price by means and advertisers attempt to purchase the inventory at fair value with some information about the past bid prices. The examples of ad exchanges are Google’s Double Click and Microsoft Ad Exchange. [39]

4.5.4 **Programmatic advertising – a closer look**

There is more than one way in which the transactions take place in the space of digital advertising. Broadly, the transactions can be categorized into fixed price and auction-based.
Broadly, there are four types of transactions as seen in the Figure 21 based on whether the inventory is reserved and whether the price is set. Open auctions are most common of all types of transactions and generally can be accessed by all buyers. Invitation only transactions, on the other hand, are open to select advertisers and publishers. In the automated guaranteed transactions, as opposed to unreserved fixed rate, the inventory of space is set aside for the buyers who want to have a choice on where they want to show their advertisement and want to ensure advertising space for particular campaign delivery. The process of these deals is partially automated. It is enabled by the Open Direct protocol and is maintained by IAB which works to not only set standards and make the communication between stakeholders but also enable growth and innovation in programmatic advertising.

The auction-based transactions are enabled by Open Real-Time Bidding (RTB). The auctions take place within 100 milliseconds. The complex transaction process is explained schematically in the Figure 22.
An Advertiser can buy directly from the Publisher when premium inventory is sold typically on a guaranteed basis by the Publisher’s sales team. An advertiser can buy from Ad networks that are aggregators of the ad inventory from Publishers. This is typically sold on a non-guaranteed basis. Advertisers can also buy from Ad exchangers through DSP which has RTB capabilities and can buy inventory with the advertiser on the Ad exchange. The trend is moving increasingly toward Real-time bidding.

Real-Time Bidding: Publisher make the ad inventory available for advertisers to buy on the Ad exchange and they want to maximize their price. Advertisers work with DSPs who enable them to access Ad exchanges. DSPs bid on the inventory based on the past data analysis, and try to pay a fair price. When the customer visits a website, before the ad is displayed, an ad call is made to the exchange by SSP. Ad exchange hold auction for each impression of the ad. It sends bid request to each DSP to bid on that impression. DSP sends out a bid amount and winning the bid is chosen by the ad exchange and the ad is displayed on Publisher’s website.
To summarize, the major technologies underlying e-Commerce can be classified as e-Commerce website and technologies in advertising ecosystem. Major technologies on which application of Machine Learning will be illustrated are discussed. They include recommendation systems, product search and product detail page. Advertising ecosystem which includes platforms including Ad exchange, Data Management Platform, Supply Side and Demand Side Platforms, Ad server, and third party system, as well as transaction process are discussed. Mainly the complex process of RTB is discussed in detail. With the understanding of these technologies powering e-Commerce ecosystem, we discuss application of Machine Learning to these systems in the next chapter.
5. Applying Machine Learning

5.1 Introduction to Machine Learning [ML]

5.1.1 Description

Algorithms that enable machines to “learn without being explicitly programmed”. ML models learn by identifying patterns from experience (provided data called training data) and apply this learning to new data. [45] ML can be applied to tasks in many areas such as face recognition, document classification, disease diagnosis, speech translation, etc.

ML problems can be described by three parameters: task, training data and performance metric. Following description discusses ML problem with an example [46]:

**Task:** Spam email detection

**Training Data:** Data that consists of a list of emails with label $Y$ indicating whether an email is a spam. $Y$ will be either 1 or 0 depending on whether it is spam or not. The data will also have attributes of email, say $X$ that email can be represented by.

Suppose email $E$ can be represented by attributes $X_1$, $X_2$, $X_3$... $X_l$ could stand for ‘email source unknown’, $X_2$ for ‘email subject line contains $$', and so on

The goal is to come up with a function $Y = f(X_1, X_2, X_3...)$ using this training data. This function results in a decision rule which model learns from training data. When a new email is provided to the model, it can decide based on this decision rule, whether the email is spam or not.

**Decision rule** could be that email is a spam, that is $Y=1$, if $X_1$ and $X_2$ and $X_3$

**Performance metric:** The metric could be to reduce probability of mistake

Fundamentally, model selection in ML for this problem is finding a set of decision rules that can be generalized well to classify a new set of emails into a spam or not spam.

5.1.2 Classification of ML

Machine Learning can be classified on a couple of bases. First, it can be classified into following categories on the basis of how model learns [45]:

Supervised learning: These are models that learn based on examples or feedback. The example of spam email classification was a supervised learning task because it knows from training data which emails could be spam based on their attributes.
Unsupervised learning: These models do not have feedback or information about the outcome to learn from. These tasks are generally like partition or clustering data into similar groups or describing the structure of the dataset.

Reinforcement learning: This model works with a feedback in the form of rewards or punishments which is given dynamically when a machine performs a certain task. For example, a robot is required to successfully hit a ball. When it hits, it gets rewarded and when it fails, it gets punished. The objective of the model is to maximize its rewards.

Second, ML tasks can be categorized as follows based on the type of output (major categories mentioned, not exhaustive):

Classification: Inputs are divided into classes

Regression: Output is continuous rather than discrete

Clustering: Output is groups which the inputs are divided into

Density estimation: output is in the form of distribution of inputs

Dimensionality reduction: When inputs have many dimensions, the output is input mapped to low-dimensional space

5.1.3 Choosing a model

Steps for choosing a model are [47]:

1. Categorize the task by input and output

2. Find applicable and practically implementable algorithms

3. Apply all of them and compare their performance

4. Select one and optimize its parameters

Following (Figure 24) by Microsoft is a rule-of-thumb for selecting models, for people who are learning the basics. It is compiled from contributions, experiences, and feedback of data scientists.
5.2 Machine Learning Canvas

At times, the data science teams invest considerable time in building Machine Learning models, but at the end of the exercise and application of models, the teams do not see Returns on Investment. This framework starts with the idea of value proposition or a business case that managers need to define before beginning to work on Machine Learning tasks. This is a good starting point for a designing system that uses ML.
The framework starts with defining the value proposition – the what (the goal of the system), the why (the system is needed) and the who (is going to use the system and get impacted by the system) of the system. Next step is to think about how the system will predict (left half of the canvas) and how it will learn (right half of the canvas). ML task asks about the background information – what the system is trying to achieve, what input is required and what output analyst is trying to predict, what type of ML problem it is such as classification, regression, etc.? Next step ‘decisions’ is to understand how are predictions being used to make decisions and ‘making predictions’ is to plan details around making predictions such as constraints, time for making predictions, plan to evaluate the system offline prior to deployment, etc. [49]

Next, (to the right of the canvas) the factors to consider are first know how to collect data and what are the data sources, second which features are to be used. Features are the characteristics or attributes of the inputs such as gender, location, age, etc. of the customer. Finally, consider the details about building the ML model. The final step is to evaluate and verify system performance.

5.3 Application of ML to Recommendation Systems
Recommendation systems originated in the 1990s and have evolved since then. Today, these systems have become more and more important for helping customers make decisions about the products or services they want to consume as the consumption is becoming digital (for example songs consumption of Spotify, videos on Netflix, products on eBay, etc). The researchers have been studying the use various Machine Learning (ML) algorithms to different applications of the recommendation system in order to provide better recommendations. There are many ML algorithms and approaches used in the recommendations systems till date, and they are based on the logic of the recommendation system, the data about the customers and the recommended items. Because of the sheer variations in ML algorithms, their combinations and the number of different algorithms used in the past to design recommendation systems, it has become challenging to choose and track applications of ML algorithms in these systems. [50]

Application discussion: Suppose that a Product Manager wants to design a recommendation system for a PDP (Product Detail Page) of an e-Commerce website that recommends ‘what other people bought’ or ‘what you might like’ products on the page. Following is the example of recommendation canvas for recommendation system. So the objective is to predict the product affinity. Let’s say that customer C is interested in purchasing a laptop P. Website knows that she prefers ratings of 4.5 and which all laptops online have that rating. The problem is to find how the customer will rate laptop P. The function can be represented as Affinity (CIP) (Affinity of customer C for product P).

![ML Canvas: Recommenders](image)

Figure 26, ML canvas applied to ML for recommendation system [50]

Here, the collaborative filtering technique (discussed in recommendation systems, Chapter 3) can be applied. The fundamental concept is to use the information about what other customer likes to predict what
the other customer might like. The problem can be solved by Matrix Factorization method. Customer-Product affinity matrix can be represented with all the customers in the database on one axis and all the products on the other axis (Figure 26). When all customers ratings for all the products are filled in the matrix, the observation is that the matrix is very sparse. This means that very few affinities are known explicitly as a few customers rate a few products. The problem is to predict the unknown affinities – the ratings a customer might give to a product she has not rated (used for ranking and displaying as a recommended item). The issues are that very small amount of data is available, and the matrix is very large. The Machine Learning model needs to try to predict the missing values or the ratings for the product with available only around 1% of the values for product ratings.

**Input:** Input will be the features/attributes of the all products P and all customers C, and the rating of customers C for products P

**Output:** Rating of new customer for a product P

**Model:** The abstraction use is that the large matrix is projected by two small matrices X (representing the customer) & Y (representing the product) (Figure 27). This is called as Link Factor Model. First, the model is trained by using the data that is already available and then apply the learning to the new data. The very first step is to run similarity using a function such as a cosine similarity to find the products that are similar to other products for which ratings are known. Cluster algorithms can be used for these. To start with, this helps get the initial approximate for the missing values of product ratings. Getting the customer data about their possible interest in products they haven’t rated (P_c) can also be done by tracking the real-time customer behavior. For example, if a customer spends more time interacting with a product, her interest in the product is inferred to be high, and sample data can be collected and plugged in the model for retaining it with real data to obtain predictions. Next step is to rank the products with the information about the customer based on prediction (X_c). [50]

![Figure 27, User Probability Matrix](image)

Figure 27, User Probability Matrix [50]
Now, Approximate $X_c = X(CIP_c)$

$$Z = X.Y^T$$

$$Z_c = X_c.Y^T$$

$$P_c \sim X_c.Y^T$$

$$P_c(Y^T)^{-1} \sim X_c.Y^T(Y^T)^{-1}$$

$$X_c \sim P_c Y (Y^T)^{-1}$$

Consequently, all the unknown values including $X_c$ can be generated.

Next, parameters need to be included in the function in order to factor the assumption about the number of existing data points to be used ($Y^T$), the number of iterations to train the model for, and parameter to avoid overfitting. Finally, use Root Mean Square (RMS) error to find how well does matrix approximate the data to evaluate the model to find which value of parameters present the best value for RMS. [50]

**Improvements** – For websites such as Amazon.com, the matrices are very large because of the long tail of products. So, collapsing technique can be applied to the matrix – collapsing products by categories or by some features. Second, currently, the matrix is two dimensional with products and customers, however, adding the third dimension of time helps predict better because ranking for a customer changes with time.

**Other ML methods:** A systematic literature review of recommendation systems has been done to study the most used applications of ML algorithms. Some ML algorithms found in the literature review can be classified easily because they are small variations of widely used and established algorithms, but others cannot be classified intuitively. [51] Out of the 35 sources used for the review, 26 were retained based on the quality and clarity of approach and validity of sources. Following ML algorithms have been used in recommendation system (listed by order of frequency of use in the retained sources).
5.4 Application of ML to e-Commerce Search

E-Commerce search has been discussed in Chapter 3. In this section, we discuss the application of Machine Learning to e-Commerce search. Learning to Rank technique has been widely applied to search and ranking widely such as in the web search. We discuss the application of Learning to Rank method (specifically Logistic Regression ML model) to this problem. Continuing the discussion on search – the problem is mainly about combining many different attributes of the query that determine the quality of output (accuracy of products presented as a result of the search query) to generate a ranking of products. The objective is to rank the products by their relevance to the customer and improve accuracy and robustness of ranking. ML model is used to optimize the weights of attributes while combining them by making use of the customer feedback about product relevance in the form of clicks, add to basket, purchase, etc.

**Application discussion:** Consider for a query-product pair \((Q,P)\), there can be many attributes \(X_i\) \((Q,P)\). Inputs \(X_i\) include:

- Content-based attributes such as the keywords mentioned in the query used as a basis for matching product by application of a retrieval function (a function such as Query Light Code and PL2 that retrieves product whose title or description matches with the search query entered by the customer)
looking for a particular product on the e-Commerce website). The features include various product features as discussed in e-Commerce search section in Chapter 3.

- Behavioral features of the customer searching the product, business metrics and other features such as product availability, etc. as discussed in Chapter 3.

**Output:** Probability that the customer will interact with the product presented as a result of the search function is $p(R=1 \mid Q,P)$. This function represents the hypothesis that the probability that a customer finds a product relevant is related to the attributes factored in the inputs “through a particular form of the function that has some parameters. These parameters can control the influence of different features on final relevance.” This hypothesis is validated experimentally. [32] [34]

The next idea is to understand the relationship between the input attributes with the output by understanding these parameters and to understand the which attributes are more influential on the relevance (the output). This is a Machine Learning task. It is based on the past data that is the training data which provides evidence of which attributes are important by providing information about when customers interacted with a particular product. Simply put, the training data will provide clues and evidence on when a customer found a product relevant (a successful output for the search function) and what was the relation of the input features with the output in the successful instances. The assumption here is that the products that are clicked or purchases are better with respect to the searched query than those that are not. [32] [34]

$$p(R=1 \mid Q,P) = S(X_1(Q,P), X_2(Q,P), \ldots, X_n(Q,P), \lambda)$$

$\lambda$ is a set of parameters. It can be learned by optimizing the above function on the training data.

**Model:** A method based on regression will be used to analyze the relationship between input attributes and the output in the training data. Many different regression-based models can be used, but here Logistic Regression model is considered primarily because the output required is the prediction whether the product is relevant to the customer – binary output – yes (1) or no (0). Also, this regression method is very simple to start with and has been widely applied to this category of problems successfully.

In this case, we assume that the input attributes can be linearly combined. Let $\lambda_i$ be the parameter that controls the weight of the input attributes. A higher value of $\lambda$ means greater weight of a respective attribute and higher influence on the output. $P$ is the probability of product’s relevance to the customer and it will take the value between 0 (not relevant) and 1 (relevant). The attributes could be combined simply using Linear Regression, using the expression on the right-hand side of the equation below. However, then it could take any value even greater than 1. So, using the log function on the left-hand side of the equation, we convert it into probability value. [32] [34]
\[
\log \frac{P(R = 1 \mid Q, P)}{1 - P(R = 1 \mid Q, P)} = \lambda_0 + \sum_{i=1}^{n} \lambda_i x_i
\]

Next, the function above is written in the form of the probability function. If the expression on the right-hand side of the above equation is large, then the probability value in the equation below is high. That is the probability of relevance of the product is high. The equation below connects the attributes of the search with the probability of product relevance. [32] [34]

\[
P(R = 1 \mid Q, P) = \frac{1}{1 + \exp(-\lambda_0 - \sum_{i=1}^{n} \lambda_i x_i)}
\]

Next step is to estimate \( \lambda \), the parameter values using the training data in order to apply this function to the new data - Machine Learning. To illustrate this point in simplified manner, suppose the input attributes \( X_1, X_2, X_3 \) have values 0.7, 0.11, 0.65 in one of the instance \( \Lambda \) of the training data in which the customer found the product relevant (\( R=1 \)), and values 0.3, 0.05, 0.4 in another instance in which she found it irrelevant (\( R=0 \)). Using the above equations, the value of parameter \( \lambda \) can be estimated. Several such instances of the training data will help estimate \( \lambda \) accurately (called leaning). With the equation having the \( \lambda \) value defined, the machine can predict the relevance of the new instances of data. In this use case, with any new customer search query, the model can predict the product relevance and rank products based on that. [32] [34]

**Comparison of ML methods for e-Commerce search:** Many regression and classification approaches have been applied to Learning to Rank for search. LambdaMART has shown better performance empirically for the e-Commerce search by a study that compared these algorithms.

![ML algorithms for e-Commerce search](image)

**Figure 29, ML algorithms for e-Commerce search** [52]

**5.5 Application of ML to advertising ecosystem (DSPs) for prospecting** [53]
The DSPs (Demand Side Platform), as discussed in advertising ecosystem in Chapter 3, need to decide whether an ad is to be shown to a customer and set optimal bid amount for each ad impression for that customer. So, it needs to find a set of target customers. Depending on the goal of a marketing campaign, the target customer (or customer prospecting varies).

Consider a problem of targeting customers who have never interacted with a particular brand (as per the past knowledge of the customer from their browsing history that the advertisers have by means of pixels and cookies – ref. Chapter 3). This is a key area for applying Machine Learning because of the large amount of granular customer behavior data, data on customers’ interaction with advertisement or brand, and need for DSPs to make real-time decisions and hundreds of decisions every second on whom to deliver ads at what price while communicating with the RTB on the Ad exchanges.

There are several challenges that have been faced in applying Machine Learning in this area. The dimensionality of the data is very high having millions of attributes (input variables for the model). Not all are important or worth considering. High dimensionality results into a problem of ‘cold start’ for a Machine Learning model, that is, when the training data has so many attributes but the new data on which the model us to be applied does not have data for those attributes, the model predictions for the new marketing campaigns are not accurate. Further, collecting training data that has all values for these millions of attributes is beyond expensive and time-consuming. Another main challenge is a high possibility of bias in the models. The data available for marketing campaigns have relatively very low positives or conversions (instances of customers interacting with the advertising brand as a result of viewing the ads is very low). So, a limited sample of training data set with very low instances of positives does not train the model accurately and bias gets build into the model. Carefully selecting and experimenting with a randomized sample is very costly. So, it is important to have a training data set with a correct population of positives and negatives.

One of the recent researches has tried to address these problems with application of Machine Learning in this area by developing a two-step model training method. In the first step, the model uses a large amount of raw data (though biased) to deal with dimensionality and sparseness. In this step, the data is collected from multiple sources, model learns the relationships and the important attributes. The model in step 2 uses the learned functions from the step 1 as inputs, and combine them to produce calibrated output. This approach is termed as transfer learning. Its aim is to apply its learning to the tasks that are different from the task model is trained on in terms of relationships of attributes with the outcome and/or have a different distribution of data. Despite these differences, the learning can be applied to a real task that is dissimilar as mentioned before. This aspect of the method makes it relevant for application to targeting customers for various marketing campaigns as each marketing campaign is different in terms of performance criteria.
Application discussion: The goal of the model is to predict using the high dimensional customer data whether that customer will make a purchase or convert (in general terms) within seven days of being shown an ad. Based on this prediction, the advertiser will bid for an impression to be shown to a particular customer.

**Step 1:** Model is high dimensional and so complex. Not many positives and so biased for the purpose of learning. Positives are the actions of customer indicates actions of the customers that indicate interaction with the advertiser such as a visit to the advertisers’ website. Other browsing events are negatives.

**Input 1:** Anonymized URLs (converted to binary form) of pages browsed by the customers in the past. The sample consists of both positives, that is, instances when a customer interacted with the brand and negatives, that is general browsing without interaction. The input features $X$ are in binary format.

**Output 1:** $Y = 0$ or 1 based on whether the browsing event was negative or positive respectively. 1 indicates conversion or customer activity that advertiser considers important from point of view of understanding customers interest in the brand, such as website visit, clicks, purchase, etc. These observations are based on real data collected from past campaigns.

**Model 1:** For the Step 1, Logistic Regression is used for training the model. SGD (Stochastic Gradient Descent) is used to train Logistic Regression model

**Step 2:** The model in step 2 consists mainly of the output of model from step 1. The output of step 1 reduces the dimensionality. The benefit is that the model gets rid of bias because the ratio of a number of positives to the entire population in the training sample is near to real universal data set.

**Input 2:** Sample of customers who have seen an ad (without being targeted) fewer positives (customers who interacted with ads). This type of sample is taken because it is like the final target population for which the model is trained - prospects (customers who haven’t seen that ad before). Input feature $X = (I, X_{step1})$ where $X_{step1}$ is set of all functions $f(X)$ trained in step 1. $X_{step1} = [f_1(X), f_2(X), f_3(X)...]$

**Output 2:** $Y = 0$ or 1 based on whether customer only browsed or interacted with the advertisement respectively with n days of advertising

The trained model from step-2 is applied to the real new data which is in the raw form to make predictions in real time. Applying this Machine Learning technique to targeting prospects for marketing campaigns is important. Marketers do a new campaign almost every week, and several campaigns go on simultaneously, and the timespan for these campaigns is often short (a few days). It is prohibitively time consuming and expensive to make Machine Learning models for each new campaign. Hence there is a new to apply transfer learning method.
Other ML methods: The method of transfer learning has been discussed in previous studies and application of following Machine Learning techniques have been demonstrated.

| Hierarchical relationships within the feature set |
| Laplaces smoothing into Poisson regression       |
| Use of alternative outcomes in a classification model |
| Multitask Learning                              |

Figure 30, ML algorithms for customer prospecting [53]

Problem identified with the above techniques is that though they use transfer learning in a way that can be applied to marketing campaigns, the practical challenge is that cross campaign data transfer or use is not permitted due to the privacy of marketers’ data. The technique of transfer learning discussed above in the detail solves this problem as it only learns from data which is relevant to the campaign but collected from different sources. These sources are in the form of distributions and relationships of attributes with the target outcome.

5.6 Propensity Modelling for customer targeting

Any e-Commerce company generates data about the customers from the transactions customer do with them or customer subscriptions on the website. So, in all, they would have data of only the customers who previously interacted with them. This is a small number of customers, and propensity modelling helps them to find look-alike customers (who are not their customers or who did not purchase the product for which marketing campaign needs to run), ones who are similar to their existing customer and might be interested in buying particular products for which marketing campaign is to be run. The propensity models can be in-built in the DMP (discussed in the advertising ecosystem, Chapter 3).
Examine purchasers vs Non-purchasers

---

**Input:** Features of purchase behavior, browse behavior and other remaining information including the customer demographics, quote details, etc.

**Output:** The model should give the probability of customer buying a product or responding to marketing campaign

**Model:** As this is a regression problem, and outcome required is probability, logistic regression is used. The model considers all the input features discussed in input (independent variables) and regresses them to predict the odds of success by determining the probability.

**Application description:** Suppose an e-Commerce company wants to do a marketing campaign for selling a microwave. The first step is to find the microwave buyers in last say 1 year (depending on how much data the company has generated). Then, find the non-buyers in the same period. Do a propensity score matching which helps match buyers with similar non-buyers (similar one to one for most features such as demographics, etc.). So, the data has the same number of buyers and non-buyers who are similar. Also, enrich the data with variables about their recent purchase activities that could indicate that they may buy a microwave (such as their purchase of the refrigerator, or new household items). Run the regression model to understand the key features that most differentiate the buyers from non-buyers. These factors or features could be the causal factors for purchase (for example, socioeconomic status, or moving to a new house, etc.). After identifying these important factors with the training data, improve the model, and then validate...
with new data. The model will provide the probability distribution of customers purchase likelihood, and company can target customers with high likelihood depending on their marketing budget.

5.7 Product Detail Page Optimization

When a customer clicks on the product icon in the search results on the e-Commerce website, the icon redirects to the Product Detail Page (example in Figure 32) that describes the product features, price, delivery date, and shows related pictures and videos. This page brings the customer closest to purchase. Also, if a customer is on the product detail page, it indicates high purchase intent. Therefore, conversion at this stage of the customer journey is critical to purchase. Similarly, product detail page is a place for upselling, cross-selling, and reading customer reviews.

**Application details:** There are several details that e-Commerce product pages can entail, but it needs to satisfy the customer with the needed information and to motivate purchase. The company needs to identify which of the components on the page matter most to the customer for driving purchase. They could be different for different customer segments or even for individual customers. Therefore, companies want to use Machine Learning to optimize and go a step further, personalize the product detail pages for their customers.

**Input:** Various attributes of the product page (as numbered in Figure 32) and data about customers’ past purchase behavior.

**Output:** Output is to predict conversion. The idea of conversion could vary, as it could include customer actions from activities that indicate engagement to customer purchase. For the discussion, we will consider conversion as customer’s engagement with the product page. Measuring product page quality by purchase metric will not provide accurate idea because a customer’s decision to purchase may be a result of product features or some other reasons that are not related to the product page. Product page can motivate a purchase only by providing accurate and desired information in best possible layout or format and make the purchase decision easier and faster. It cannot change the customer’s opinion about a product. Engagement can be measured by behaviors such as the number of clicks on the product page, number of scrolls, and the time spends on the product page.
Model: This is fundamentally a regression problem, and techniques such as Linear Regression can be used. However, it is a practice to explore the application of various Machine Learning models and understand which model provides the best prediction, before deciding which model to use for the training. In many applications, Random Forest has shown most accurate predictions. Random forest is an advanced technique which includes the characteristics of classification and regression method. It can use much more observations than regression while capturing the variances in the dependent variables or the input features. [55]

After comparing the predictive power of various models for a problem, next step is to improve the model. Finding the significant features that have more influence on the outcome/ conversion need to be identified. This could be done by adding and dropping the features and observing the change in the outcome of the model. With the Machine Learning applications, generally, a function is not observed but learned from the data and iterations. [53]

To summarize, the chapter opens with an introduction to ML, and proposes and discusses the use of ML canvas while applying ML to any systems. Then, the chapter analyses and illustrates applications of ML to e-Commerce technologies in detail. The application is structured systematically into sections namely the problem statement, application details, input, output, model, and other possible models that can be/have been applied. This chapter tries to include the logic behind choosing a model where possible. It mentions other models that have been applied to the problems in the past.
6. Conclusion

6.1 Thesis summary

Customer journeys are a series of steps that customers take before and after making the purchase. Purchase journeys are becoming non-linear due to continual access to digital channels and increasing number of touchpoints. Customers can jump between the steps, and with the exposure to competitors’ advertisements, they can quickly change their purchase decision in the favor or again a brand. Competing brands can enter and exit the journey even in the stages right before purchase, and steal the deal. Also, as the customers are purchasing digitally more and more, and as the number of digital touchpoints between customers and purchasers is increasing, the noise of marketing is becoming an irritant to customers.

The thesis uses a framework to accurately predict the customer needs and help companies build loyal relationships by targeting right prospects at the right time on the rightmost relevant touchpoint to assist them in their purchase journeys. Sales and marketing departments need to look at the sales journey of the customer not as discrete events at the touchpoint level, but in a connected way. The thesis illustrates, with an example, an end-to-end view of the customer journey across the touchpoints and designing the journeys of their customers’ in order to consciously shape their decisions.

The thesis analyzes and illustrates the application of Machine Learning techniques to the systems of e-Commerce website – Search, Recommendations and Product Detail Pages, and to the systems in the advertising ecosystem – Data Management Platforms and Demand Side Platforms. The objective of all these applications is to increase the e-Commerce sales and make the process efficient.

6.2 Challenges and opportunities for future work

E-Commerce websites have started doing journey analytics of the customers from the time they land on their website to when they make the purchase or end the session. There is a fertile ground for applying Machine Learning to predict and influence customers’ next move and direct them to relevant touchpoint or channel towards purchase. By using the past data of journeys that led to successful purchases, companies can train models to predict the winning moves and use the information to help shape their customer’s journey. Further, they can optimize the journeys to have least number of roadblocks or exposure of competing brands. “80% of leading marketers today agree that a capability in Machine Learning will be critical in providing personalized experiences along the customer journey.” [56]

However, the e-Commerce companies are yet to apply the end-to-end purchase journey analytics and application of Machine Learning to end-to-end customer journeys. Tarun Thapar, Marketing Analytics Manager at Home Depot, pointed out a practical challenge that e-Commerce players face in the process of
getting a complete picture of their customers’ journeys. They can collect the real-time data about customer actions in the domains of their and partners’ websites, however, that is a limiting factor. During their purchase journeys, customers move across touchpoints in multiple domains, and also across digital and non-digital channels (Omni-channel). Third party cookies which collect data across domains are being used less because of the privacy concerns, and companies cannot place pixels on all possible places where potential customers visit. The challenge is about the broken journeys because of discontinuous journey data. More sophisticated Machine Learning techniques need to be employed to predict the journeys with this limitation in the ecosystem.

Another challenge is about the need for using real-time data for effective marketing. Abhi Yada, Founder of ZyloTech pointed out that “the customer data is highly perishable”. [57] There is a lag between when customers express their interest in a product or brand and when companies learn about that customer and target that customer to drive conversion. The ecosystem has many players and their data processing and data transfer are naturally not connected. For example, an advertiser collects data from multiple sources, makes creatives and ads with the help of agencies, connects with Ad exchanges and SSPs for placing the advertisements, and by the time companies move through the ecosystem from beginning to the end, they are working on stale data, and probably the customer is no more interested. Currently, companies are doing mindless retargeting. For example, if a customer visits an e-Commerce website and clicks on a cell phone, the sellers will keep displaying cell phone ads and recommendations on all touchpoints for next many days. Some players in the ecosystem such as ZyloTech have started to use sophisticated Machine Learning and Artificial Intelligence to tackle these challenges for the companies. However, this is the beginning.
References

[4] Salesforce
[7] https://go.forrester.com/blogs/12-10-04-buyer_behavior_helps_b2b_marketers_guide_the_buyers_journey/
[15] https://www.google.com/analytics/attribution/capabilities/


[22] https://en.wikipedia.org/wiki/Mobile_advertising

[23] https://www.mobileads.com/blog/best-mobile-ad-formats-sizes-display-ad-campaigns/


[25] https://goo.gl/kHT7DW


[27] https://arxiv.org/abs/1511.05263

[28] https://www.amazon.com/

[29] https://goo.gl/CdfyWB


[34] https://www.coursera.org/learn/text-retrieval/lecture/QORNe/lesson-6-5-recommender-systems-content-based-filtering-part-1

[35] https://en.wikipedia.org/wiki/Tf%E2%80%93idf


[38] https://www.youtube.com/watch?v=xnX1nxMM-R0

[39] Book

[40] https://www.adobe.com/experience-cloud/topics/data-management-platform-dmp.html
[41] https://en.wikipedia.org/wiki/HTTP_cookie


[43] Interactive Advertising Bureau, 2013

[44] Forrester Research, E-Consultancy


[46] Machine Learning course slides, MIT EECS


[49] https://www.youtube.com/watch?v=Qpo5XzlUrwE

[50] https://www.youtube.com/watch?v=l2wLFYOUX8o


[55] https://www.analyticsvidhya.com/blog/2014/06/introduction-random-forest-simplified/


[57] https://martechseries.com/analytics/interview-with-abhi-yadav-founder-ceo-zylotech/
