**Introduction:**

Machine Learning (ML) by looking at the definition, we can say that it is a field of computer science. The telecommunications sector has become one of the main industries in developed countries.

**Problem Definition:**

Customer retention means that the company is having the ability to turn the customers into buying the products from their company without switching to any other company. If the customer is one of the daily customers of the company, the companies tend to provide them with discounts and special offers.

**Exploratory Data Analysis (EDA):**

EDA is a phenomenon under which data analysis is used for gaining a better understanding of data aspects like a.) Main features of data. b.) Variable and relationships that hold between them. c.) Identifying which variables are important for our problem.

There are 269 rows and 71 columns in this dataset. . We will plot the box plots, distribution plots, count plots etc. The count plot analysis are: From the above observations, we can say that female are 181 and male are 88. The applicants who shop online from Delhi are 58, Greater Noida are 43, Noida are 40, Bangalore are 37, Karnal are 27, Solan are 18, Ghaziabad are 18, Gurgaon are 12, Merrut are 9, Moradabad are 5, Bulandshahr are 2. The applicants who use Mobile Internet while shopping are 142, who use Wi-Fi while shopping are 76, who use Dial-up are 4. The applicants who use smartphone are 141, laptop are 86, desktop are 30 and tablet are 12 are the devices the applicants use to access the online shopping. The applicants who use mobile device with 5 inch size are 7, 4.7 inch size are 29, 5.5 inch size are 99 and other sizes are 134. The applicants who use Windows OS are 122, Android are 85 and IOS/Mac are 62. The applicants who run Google Chrome browser are 216, Safari are 40, Opera are 8 and Mozilla Firefox are 5. The applicants who followed Search Engine channel to arrive at their favorite online store are 230, Content Marketing channel to arrive at their favorite online store are 20 and Display adverts channel to arrive at their favorite online store are 19. The applicants who took more than 15 minutes to make their first purchase are 123, 6-10 minutes are 71, 11-15 minutes are 46, less than 1minute are 15, 1-5 minutes are 14. The applicants who preferred payment option through debit/credit cards are 148, Cash on Delivery are 76, E-Wallets ( Paytm, Free charge etc.) are 45. Some of the applicants who abandon the bag or shopping cart because they got the best alternative offer and the applicants who choose this are 133, promo code not available are 54, because of the change in price some of the applicants dropped out are 37 and lack of trust are 31.

After plotting the distribution plots for all the columns, we can say that data is not distributed normally in all the columns. After plotting the box plots for all the columns, we can say that the outliers are present in pin code of where you shop online from, what is the screen size of your mobile device, what browser do you run on your device to access the website, which channel did you follow to arrive at your favorite online store for the first time, User friendly Interface of the website, Empathy (readiness to assist with queries) towards the customers, Being able to guarantee the privacy of the customer, Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.), Net Benefit derived from shopping online can lead to users satisfaction. These are the columns where in there were outliers present in it.

The columns where in the outliers are not present are Gender of respondent, How old are you, Which city do you shop online from, Since How Long You are Shopping Online, How many times you have made an online purchase in the past 1 year, How do you access the internet while shopping on-line, Which device do you use to access the online shopping, What is the operating system (OS) of your device, After first visit, how do you reach the online retail store, How much time do you explore the e- retail store before making a purchase decision, What is your preferred payment Option, Why did you abandon the “Bag”, “Shopping Cart”, The content on the website must be easy to read and understand, Information on similar product to the one highlighted is important for product comparison, Complete information on listed seller and product being offered is important for purchase decision, All relevant information on listed products must be stated clearly, Ease of navigation in website, Loading and processing speed, Convenient Payment methods, Trust that the online retail store will fulfill its part of the transaction at the stipulated time, Online shopping gives monetary benefit and discounts, Enjoyment is derived from shopping online, Shopping online is convenient and flexible, Return and replacement policy of the e-tailer is important for purchase decision, Gaining access to loyalty programs is a benefit of shopping online, Displaying quality Information on the website improves satisfaction of customers, User derive satisfaction while shopping on a good quality website or application, User satisfaction cannot exist without trust, Offering a wide variety of listed product in several category, Provision of complete and relevant product information, Monetary savings, The Convenience of patronizing the online retailer, Shopping on the website gives you the sense of adventure, Shopping on your preferred e-tailer enhances your social status, You feel gratification shopping on your favorite e-tailer, Shopping on the website helps you fulfill certain roles, Getting value for money spent, From the following tick any (or all) of the online retailers you have shopped from, Easy to use website or application, Visual appealing web-page layout, Wild variety of product on offer, Complete relevant description information of products, Fast loading website speed of website and application, Reliability of the website or application, Quickness to complete purchase, Availability of several payment options, Speedy order delivery, Privacy of customers information, Security of customer financial information, Perceived Trustworthiness, Presence of online assistance through multi-channel , Longer time to get logged in (promotion, sales period), Longer time in displaying graphics and photos (promotion, sales period), Late declaration of price (promotion, sales period), Longer page loading time (promotion, sales period), Limited mode of payment on most products (promotion, sales period), Longer delivery period, Change in website/Application design, Frequent disruption when moving from one page to another, Website is as efficient as before, Which of the Indian online retailer would you recommend to a friend. So, these are the columns where in the outliers are not present.

**Conclusion:**

Customer retention is the most important factor for the companies to hold on to their customers so that they do not switch to another company to buy the products