**DIGITAL SOLUTIONS TO RETAIL INDUSTRY USING MACHINE LEARNING**

**INTRODUCTION :**

Data-driven decisions have been defining the success of retailers long before machine learning and AI were readily available and applicable. More and more retailers track customer shopping habits through data sources such as CRM databases, loyalty programs, social media activity, purchase history, consumer demand, and market trends. But, the ability to apply complex mathematical calculations to big data automatically — iteratively and quickly — is now attainable with machine learning.

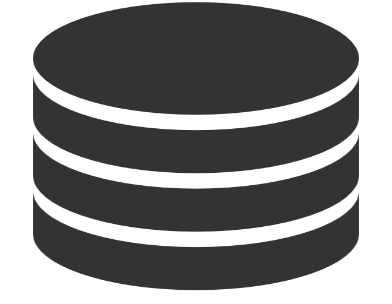
**OBJECTIVE:**

To provide better solution for various retail activities, exploring the usage of both structured and unstructured data.

**ARCHITECTURE DIAGRAM:**

Inventory Management

Pricing



DATA

Recommendation

Customer Segmentation

Target Marketing

**APPROACHES USED -**

**DESCRIPTIVE ANALYTICS**

Descriptive analytics is the interpretation of historical data to better understand changes that have happened in a business. Descriptive analytics simply describes the past using a range of data to draw comparisons.

**PRESCRIPTIVE ANALYTICS**

Prescriptive Analytics is an advanced analytics technology that can provide recommendations to decision makers and help them achieve business goals by solving complicated optimization problems.

**PREDICTIVE ANALYTICS**

Predictive analysisis a commonly used statistical technique to predict future behaviour.

**LINEAR REGRESSION**

Linear regression is used for finding linear relationship between target and one or more predictors.

**LOGISTIC REGRESSION**

Logistic Regression is a classification algorithm. It is used to predict a binary outcome (1 / 0, Yes / No, True / False) given a set of independent variables.

**RFM SEGEMENTATION**

RFM (Recency,Frequency,Monetary) segmentation allows marketers to target specific clusters of customers with communications that are much more relevant for their particular behaviour

**ASSIOCATION RULE MINING**

Association rule learning is a [rule-based machine learning](https://en.wikipedia.org/wiki/Rule-based_machine_learning) method for discovering interesting relations between variables in large databases.

**COSINE SIMILARITY**

Cosine similarity is a measure of similarity between two non-zero vectors of an inner product space that measures cosine of the angle between them.

**CUSTOMER SEGMENTATION**

Customer Segmentation is the practice of dividing a customer base into groups of individuals that are similar in specific ways relevant to their purchases. This is implemented using RFM and Logistic Regression.

**PRICING**

Logistic Regression is a classification algorithm. It is used to predict a binary outcome (1 / 0, Yes / No, True / False) given a set of independent variables. The product features like space, state, quantity, sub category and age were considered for predicting the discount. (Predictive Analytics)

**TARGETED MARKETING**

Targeted Marketing is the process of dividing a large market into smaller pieces based on one or more meaningful shared characteristics and they will respond similarly to marketing action with RFM score. Based on the products what they are buying more and making more profits to the company for every customer their interested sub-categories are decided. The offers are given on their interested products to people in some clusters in products interested by that group by sending promotions through different channels.

**PRODUCT RECOMMENDATION**

Recommendations are based on user behaviour. These are items that have been frequently viewed, considered, or purchased with the one the customer is currently considering. From the result obtained from classification of Products Recommendation system generates score for all products taking one product at a time into consideration.

**CUSTOMER BASED RECOMMEDATION**

Customer Based Recommendation is a term frequently used in marketing. It is a measure of how products and services supplied by a company meet or surpass customer expectation. The best customers and best products were considered .Using Cosine Similarity and Pearson Correlation a similarity score is generated .If the score > threshold, compare the purchase pattern and recommended products.

**INVENTORY MANAGEMENT**

Extracted the pattern of inventory for all products for a period of 48 months (Descriptive Analytics). Based on profits and quantity sold of products they are classified as best, moderate, worst and High, Medium, Low respectively. Based on product & given date the predicted sales for next 3 months are calculated by using Linear Regression (Predictive Analytics)**.**

**FUTURE ENHANCEMENT**

In Customer Based Recommendation, if product reviews are available an NLP approach can be used to enhance the recommendation. Information about the age of the customers can be useful in making age-based clusters like Youth, Old Age, kids etc. In Pricing, instead of telling the retailer whether to give discount or not we can give the discount value to the Retailer.