

9th Webinar on Data Science and Analytics

Topic:

From 360 degree view to Customer Genome- Hyper Personalization using Data Analytics

Srilatha Kappagantula

Head, Risk Data Repository , State Street Corp

23rd October, 2021, Saturday

03.00 PM to 05.00 PM (India Time)



What is common in all pictures



Winners have methodically aligned data to their advantage



Technology will be the key lever to differentiate the winners against losers in all industries including banks



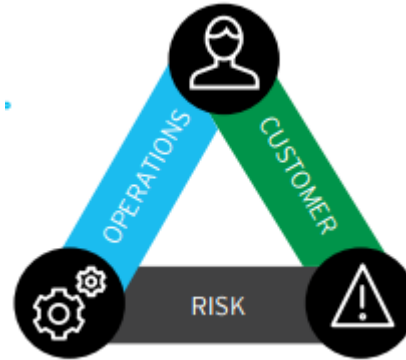
*“Get closer than
ever to your
CUSTOMERS.
So close that you tell
them what they
need well
before they realize it
themselves.”*



The banking and financial services industries have been adopting personalization across functional areas to optimize costs, income and reduce risks

REDUCE COSTS:

- Call Center Analytics
- Campaign Spend
- Spend and Usage
- Behavior Analytics



GROW BUSINESS:

- Market Sizing & Segmentation
- Cross-sell and Retention
- Campaign Effectiveness
- Sales Effectiveness
- Propensity measurement
- Personalized experience
- Customer Life time Value

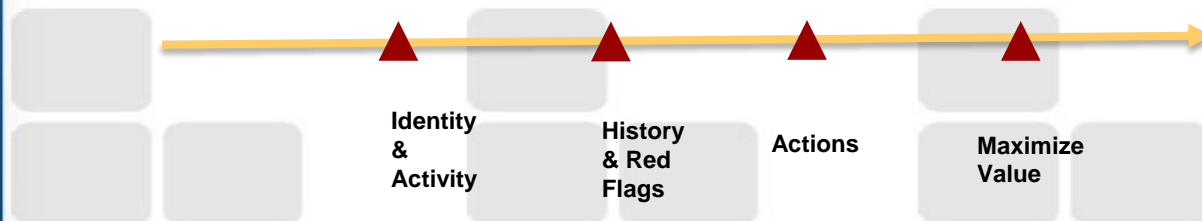
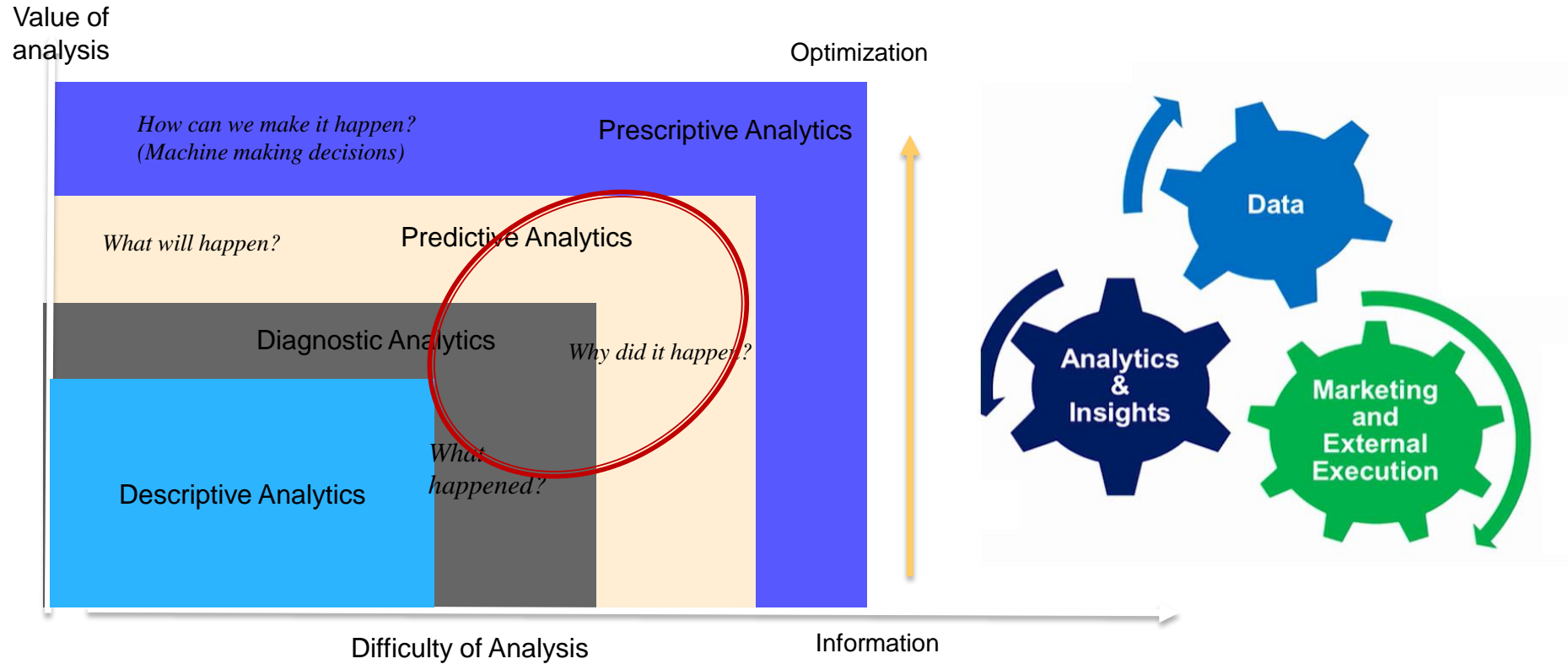
MANAGE RISKS:

- Churn Management
- Customer Profitability
- Customer Satisfaction/ Dissatisfaction
- Red Flags
- Early warning systems
- Behavior pattern analysis

Challenges in achieving personalization

| Data Silos | Data Volumes |
|----------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none">• Multiple Data Silos• Overlapping and Conflicting Info• Multiple Business Units | <ul style="list-style-type: none">• Data growing rapidly• Data dying rapidly |
| New Data Sources | Data Processing Costs |
| <ul style="list-style-type: none">• Semi/ Unstructured Data sources• Streaming real time call center data | <ul style="list-style-type: none">• Prohibitive Cost• GIGO |

Maturity Path of Personalization using Analytics



Illustrative Commercial Loan Portfolio Analytics Screens

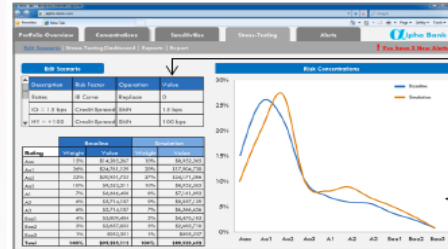
Portfolio Composition Dashboard



Industry Exposure

Risk Allocation

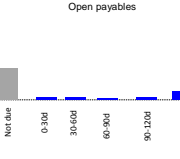
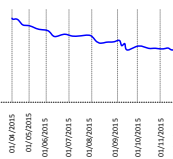
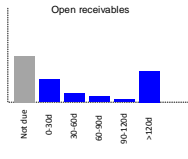
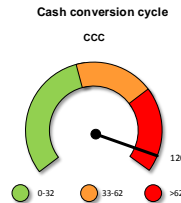
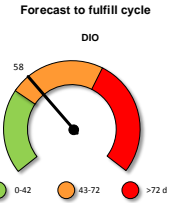
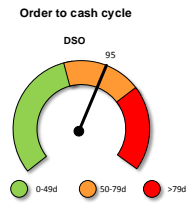
Scenario Simulator



Stress-Testing Scenario Parameters

Changes in Portfolio Risk Decomposition

Early Warning System



| Customer | Overdue (\$) |
|----------------|--------------|
| Cori Engineers | 323469 |
| AA Steel | 491678 |
| Netbau Inc. | 181200 |

| ABC | Inventory Value | Value over target |
|-----|-----------------|-------------------|
| A | 700933 | -73.13 |
| B | 399401 | 65.82 |
| C | 32717 | 346.25 |

| Vendor | Not Due |
|-------------|---------|
| hackneys | 204837 |
| Travis & Co | 82327 |
| ToolStation | 55899 |

Working Capital
\$5,642,720.78



Two explosions killed three people and injured many at the April 15 Boston Marathon. Within 24 hours the FBI had correctly identified the suspects. To do this, they analyzed about 10 TB of data from cell phone tower call logs, text messages, social media, photographs and video from surveillance cameras and cell phones.

How did they identify the correct suspects from the enormous amount of data in less than 24 hours?

Poll Question



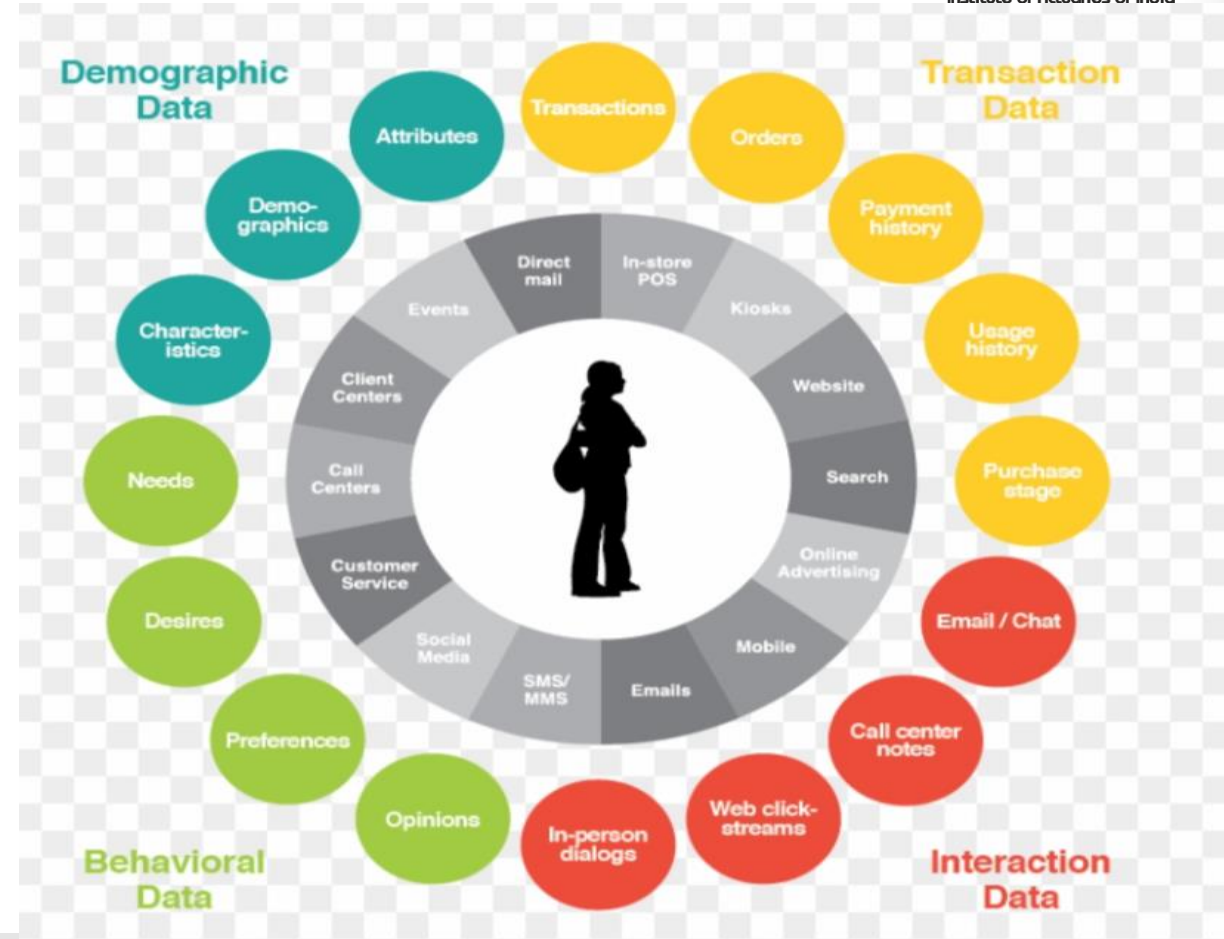
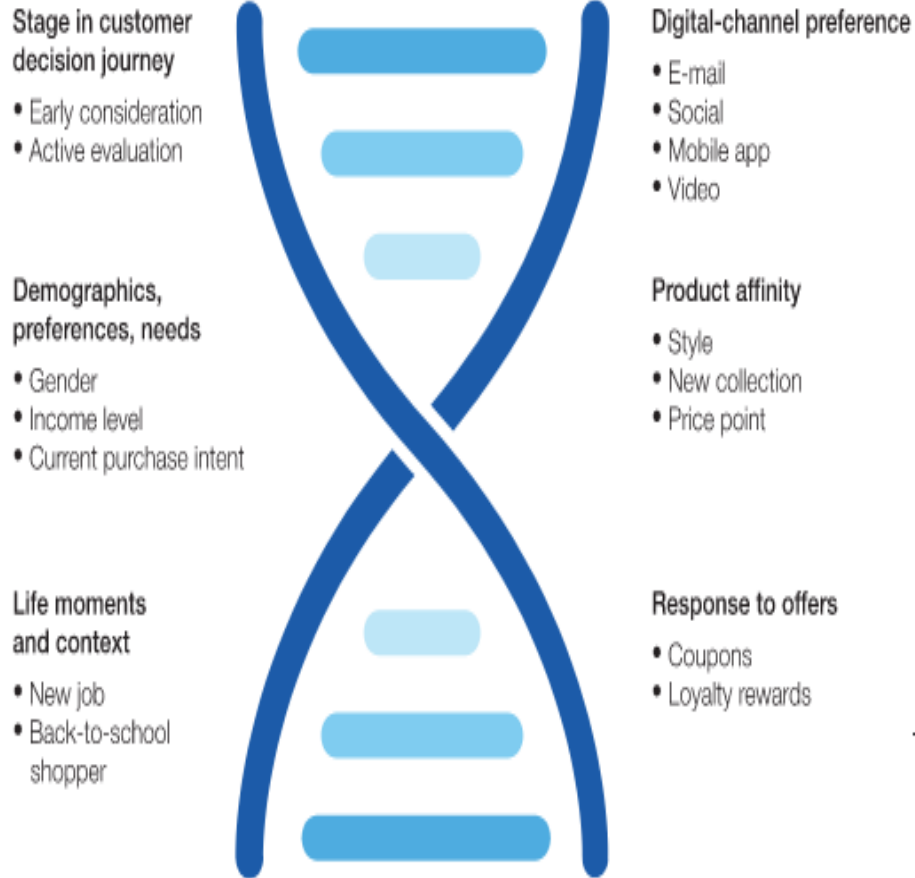
Customer Genome- Personalization to Hyper Personalization

- How to unlock the value of enterprise data
- How to go beyond customer expectations and provide Value



Every Customer interaction / Choice enriches the genome

A genome brings together multiple data sources across channels



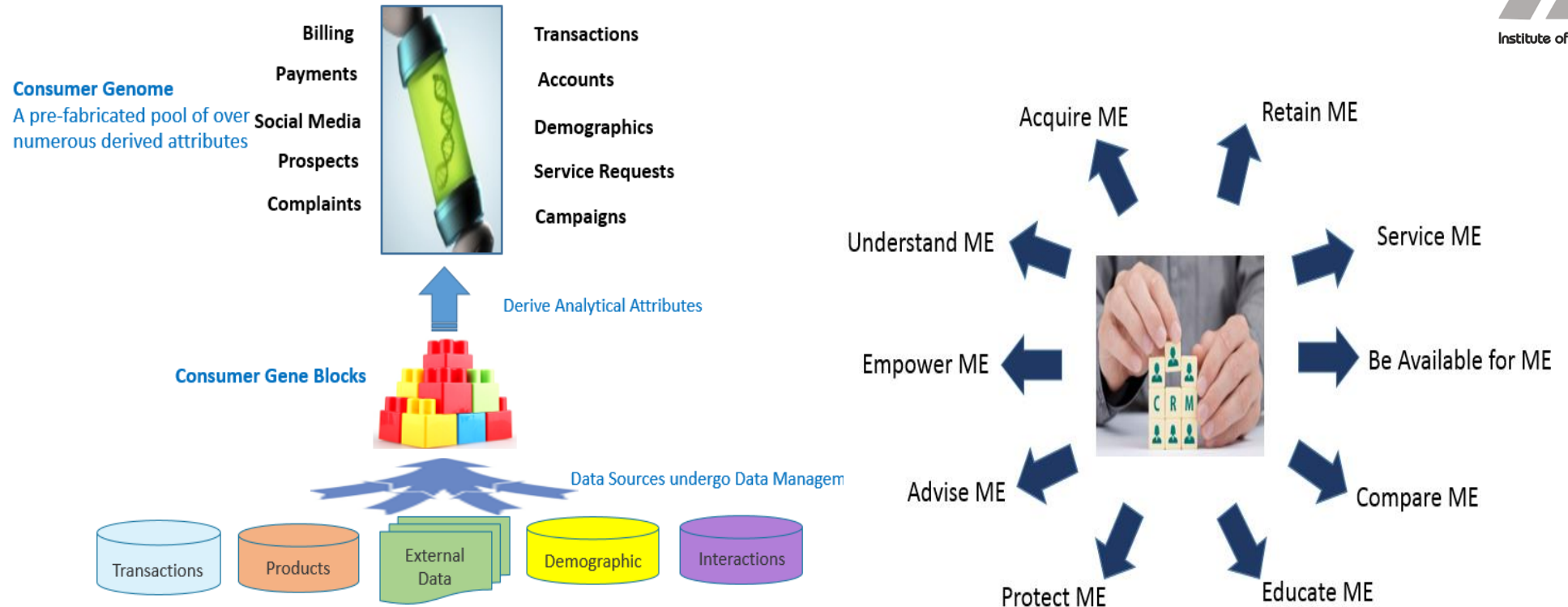
• *Understand Customer decision journey*

• *Understand Digital-channel preference*

• *Response to offers*

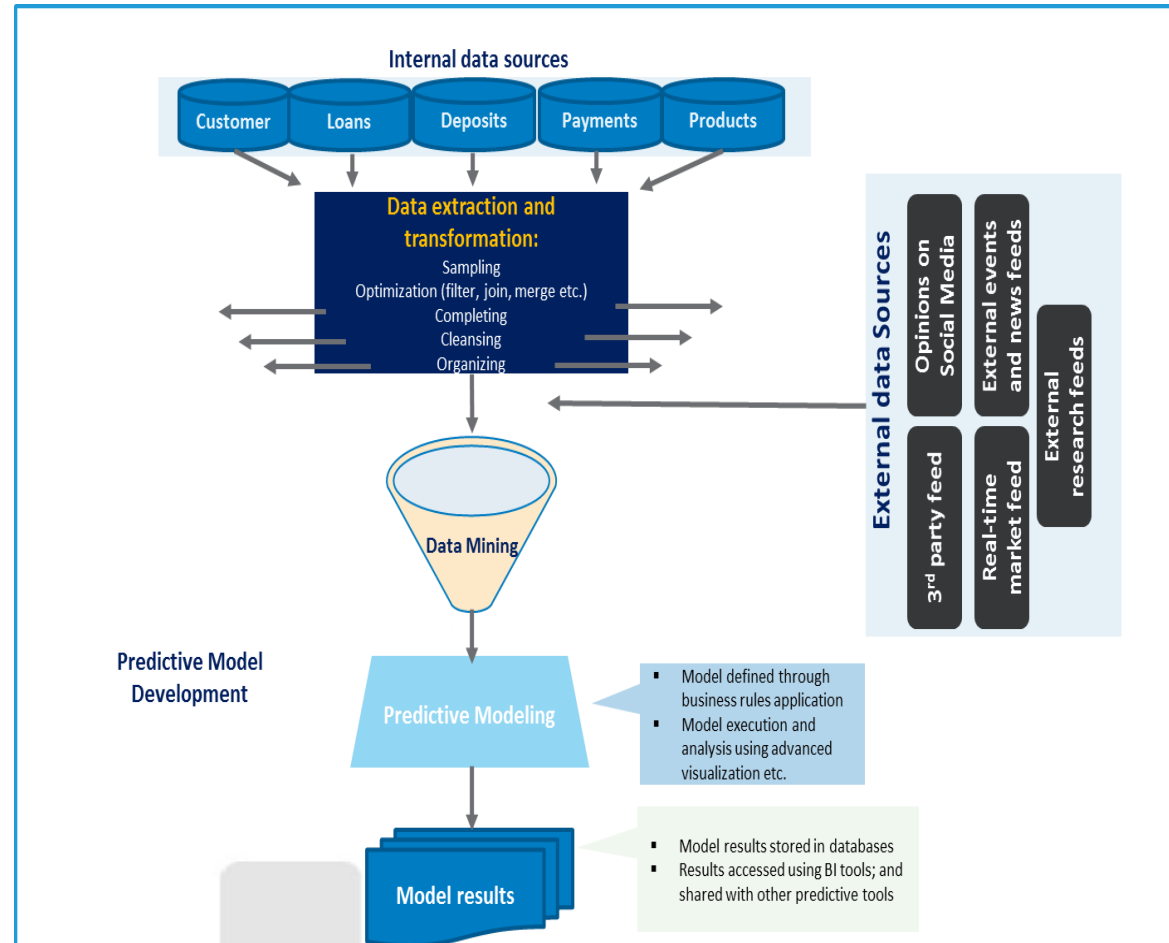
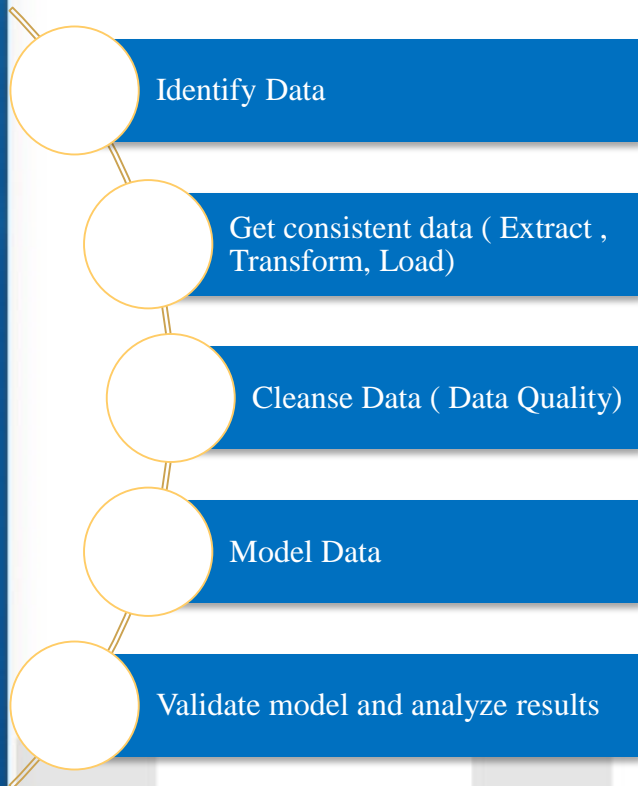
• *Life moments and context*

Customer Genome : Breaking into customer's genes to build a Virtual "Me" (Customer)



Different dimension to customer interaction.....

Implementation: Key imperatives



Major Techniques - Clustering

Customer Hyper Segmentation using multiple gene blocks

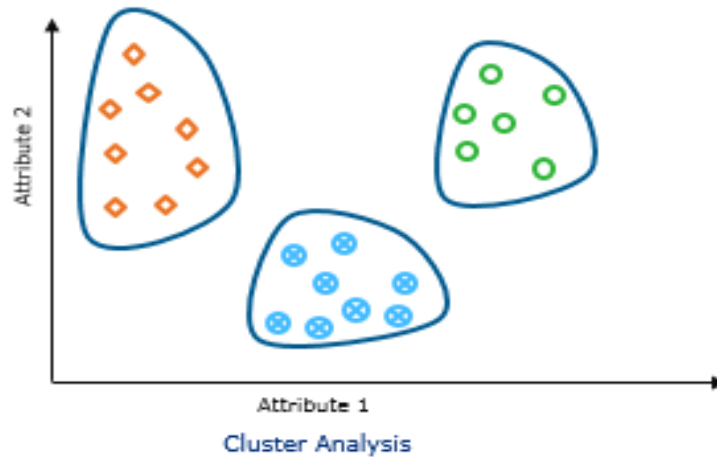


Cluster Analysis

Example: Customer Hyper segmentation

Business Objectives: Identify various groups of customers

Criteria: Similarity/distance among the observations in the available data and parameters



- Customer Segmenting
- Goal: To find groups of Customers that are similar to each other based on the important attributes such as income , spend, age , location , profession
 - Approach: To identify customer groups that are similar to each other. Form a similarity measure based on the attributes. Use it to cluster.
 - Gain: Segments can be used in various ways
 - Create clusters based on spend patters to identify suspicious spend patters
 - Offer products based on the cluster to which a customer belongs
 - Identify risk patterns based on the cluster to which a customer belongs

Intracluster distances
are minimized

Intercluster distances
are maximized

Major Techniques - Classification

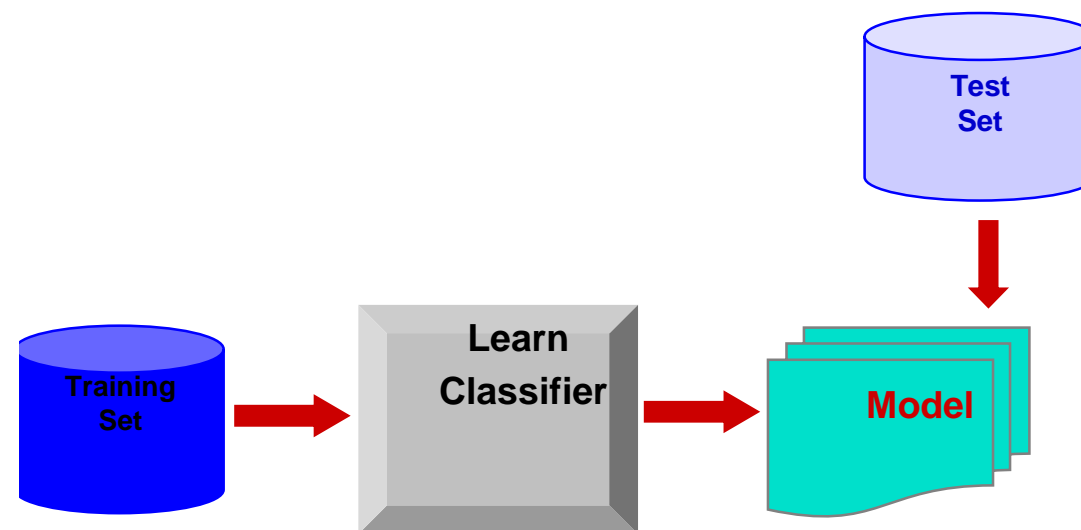
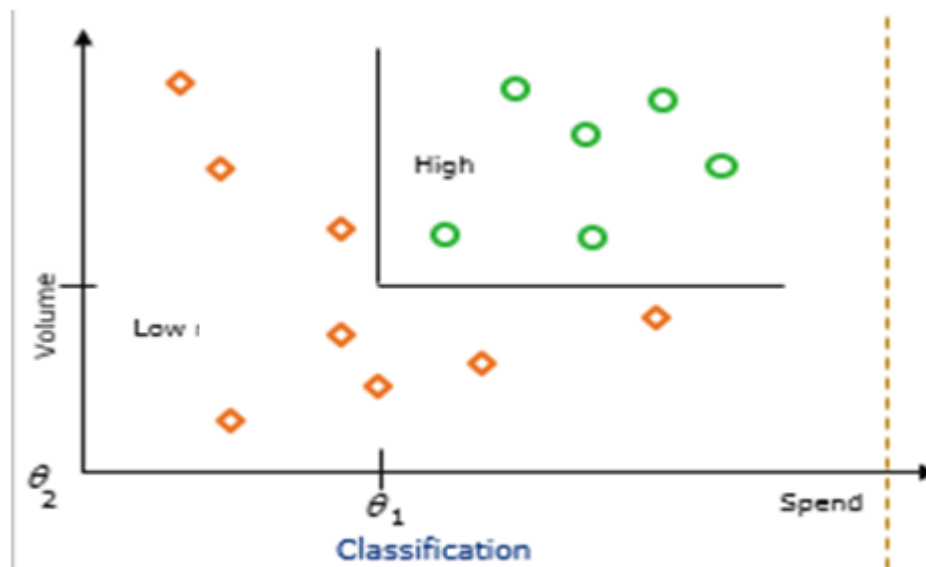


Propensity to accept an offer in a campaign

Example: Propensity to accept an offer in campaign

Business Objectives: Differentiating between low-propensity and high-propensity customers

Criteria: IF $Spend < \mu_1$ AND $Volume < \mu_2$
THEN low-risk ELSE high-risk



Frequency between transactions
Location of two transactions
When does a customer buy,
What does a customer buy, and how much does he / she spend - Spend patterns

ited, India

Major Techniques - Regression

Life Time Value of Customer / Profitability from a customer



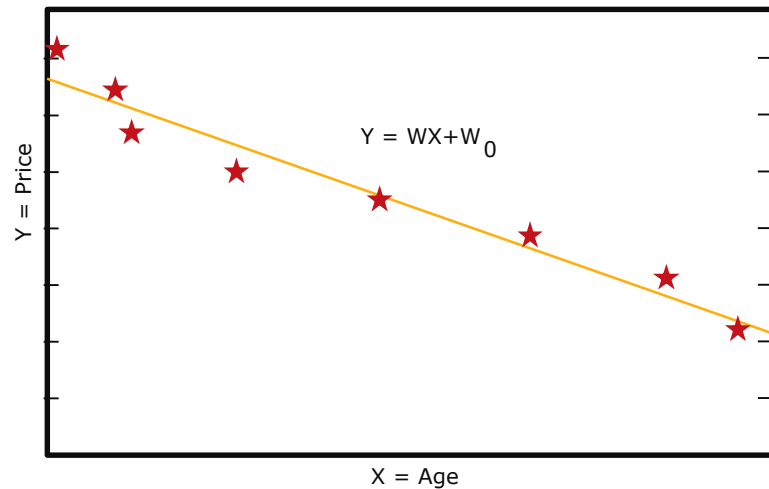
Regression Analysis

Example: Price Prediction

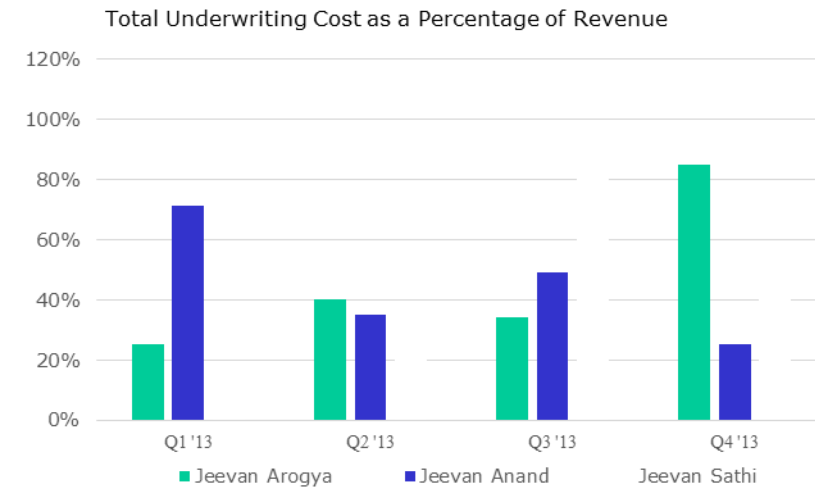
Business Objectives:

Predicting the Car price based on the age of the Car

Model: Car price is function of Car age. Age is a determinant factor of Car price



Insurer can use predictive analytics to filter out the applicants who do not meet a predetermined model score. This type of screening can gently increase insurer efficiency by reducing the employee hours it may have spent in researching and analyzing the applicants who ultimately is not a desired insured.

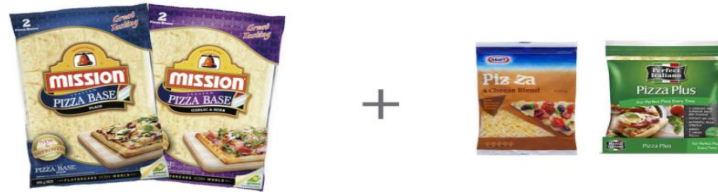


Product & Service Bundling

Market Basket Analysis

Helping online shoppers make smart decisions

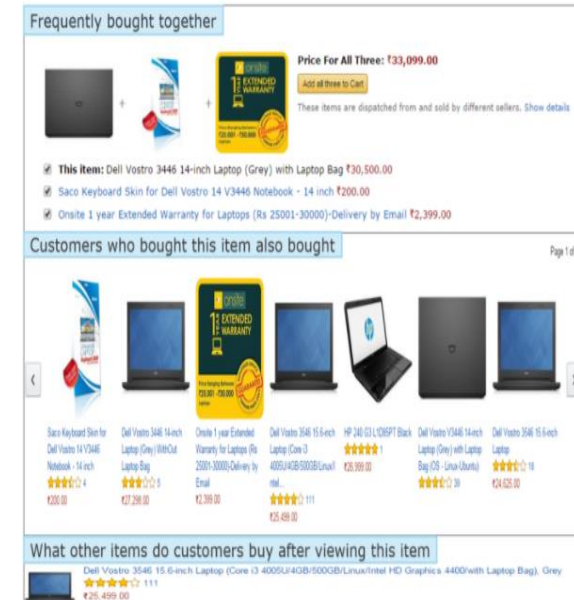
How likely is it that people who purchase Pizza Base are also likely to purchase Cheese?



The Sad Truth About Diapers and Beer



Don't be surprised if you find six-packs stacked next to diapers!



Frequently bought together

Price For All Three: ₹33,099.00

✓ This item: Dell Vostro 3446 14-inch Laptop (Grey) with Laptop Bag ₹30,500.00

✓ Saco Keyboard Skin for Dell Vostro 14 V3446 Notebook - 14 inch ₹200.00

✓ Onsite 1 year Extended Warranty for Laptops (Rs 25001-30000)-Delivery by Email ₹2,399.00

Customers who bought this item also bought

| Product | Price | Rating |
|--------------------------------------------------------------------------------|------------|--------|
| Saco Keyboard Skin for Dell Vostro 14 V3446 Notebook - 14 inch | ₹200.00 | 4.5 |
| Dell Vostro 3446 14-inch Laptop (Grey) with Laptop Bag | ₹30,500.00 | 4.5 |
| Onsite 1 year Extended Warranty for Laptops (Rs 25001-30000)-Delivery by Email | ₹2,399.00 | 4.5 |
| Dell Vostro 3546 15.6-inch Laptop (Core i3) | ₹25,499.00 | 4.5 |
| HP 246 G3 (13.3) i5 (8GB) Black | ₹25,999.00 | 4.5 |
| Dell Vostro 3446 14-inch Laptop (Grey) with Laptop Bag OS - Linux (Ubuntu) | ₹24,525.00 | 4.5 |
| Dell Vostro 3446 15.6-inch Laptop | ₹24,525.00 | 4.5 |

What other items do customers buy after viewing this item

Dell Vostro 3546 15.6-inch Laptop (Core i3 4005U/4GB/500GB/Linux/Intel HD Graphics 4400/with Laptop Bag), Grey

₹25,499.00

- Product Bundling
- Goal: To find groups of products & Services that can be offered together
- Produce dependency rules which will predict occurrence of an item based on occurrences of other items.
 - Example : Travel Card , Travel Insurance , FX Conversion services, Line of Credit (Over Draft)
 - Identify Risk patterns that exist together

Good practices

- Context is important – We don't need all information . We need relevant information
- Agile Systems : System Architecture has to be agile to take in new datasets
- High Quality Data : Should use consistent and high quality data .
- Focus on Use cases: Understand Use cases and use appropriate gene blocks. Bring Incremental data instead of 100%
- Work on data before implementing Analytics
- Data lake to be capable of using structured and unstructured data
- Synchronous data exchange between all parties .
- Customer Genome is not a “Cure all “ and not everything goes by predictions. Have reasonable expectations.



Few Real Life examples

Unprecedented Tapping into Social Media Data Sources



Problem Statement : A large bank in Australia is looking for an early warning system , for commercial lending portfolio . The bank is wanting to consider attributes that have not been considered so far .

High volume day-to-day data and feeds which can be viewed holistically as a predictor of riskiness, especially for Small and Medium Enterprises.



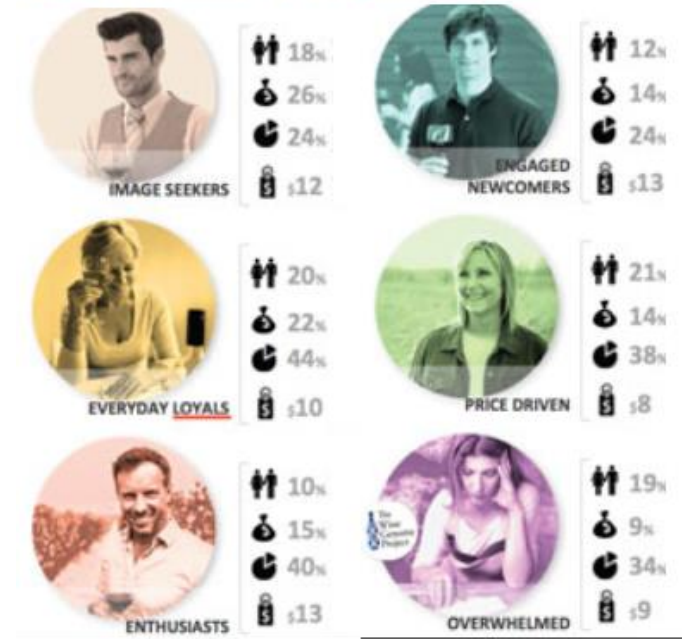
| S. No. | Area | Post | Weightage | Reasoning |
|--------|-------------------------------------------------------------|-----------------------------------------------------------------|-------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | Performance of Existing Loan | Good comments from Customers | Positive | Customer delight and its quantification can be defined, i.e. 10 positive customer statements can result in an increase in the SME's scorecard |
| 2 | Performance of Existing Loan | Marketing new products | Neutral / Positive | If only the marketing / information angle then there will be no impact, but if it is classified into the number of likes and reaches a pre-defined target then can add on to the positivity of the existing loan |
| 3 | Performance of Existing Loan | Conducting customer interaction events | Positive | Ensures Customer connect and along-with the target connects and likes for the connects, scorecard can be increased |
| 4 | Performance of Existing Loan | Announcing positive news by the SME | Positive | Innovative marketing or some news which increases the worth of the SME can result as a re-enforcement of existing loan |
| 5 | Performance of Existing Loan | Utilization of technology/ media | Positive | The usage of new technology / platforms by the SME demonstrates their alertness and knowledge in the area |
| 6 | Performance of Existing Loan / Monitor Potential of Default | Negative comments from Customers | Negative | Will have the opposite effect of the positive comments, decreasing the worth of the SME with a defined number of comments |
| 7 | Performance of Existing Loan / Monitor Potential of Default | Number of followers | Neutral / Negative / Positive | The number of followers can monitor the value of the business, e.g. a low number can be a negative signal, a normal number neutral and a very high following denoting positivity |
| 8 | Monitor Potential of Default | Related business / groups | - | The analysis can track related businesses or groups and their performance can indicate the SMEs performance |
| 9 | Monitor Potential of Default | Increase in unlikes | Negative | If Customers are unliking the SME, product, comments can flag dissatisfaction |
| 10 | Monitor Potential of Default | Enquiry of withdrawal in a possible area | Negative | Indicates the SME has been unsuccessful (poor management, market study, etc.) |
| 11 | Monitor Potential of Default | Enquiry of withdrawal in a possible product | Negative | Denotes failure in an area |
| 12 | Market New Product / Lines | Enquiry of interest in a possible area | Positive | Location expansion possibilities and can be utilized for proactive marketing of the set-up of the new store |
| 13 | Market New Product / Lines | Enquiry of interest in a possible product | Positive | Product expansion possibilities and can signal need for funding of additional infrastructure |
| 14 | Market New Product / Lines | Accumulation of [x] number of positive customer feedback | Positive | Business doing well and more investment can be made to the SME |
| 15 | Fraud Indication | Information of fraud from customers, shareholders and creditors | Negative | Due to heavy use of social media, stakeholders are directly using social media to express any fraud or negative interaction with the company |

Genre Identification by Netflix

Problem Statement : Micro Tag Customers in order to recommend genres, by his video watching patters

- Micro segments all the movies into genres based on the customer watching patterns
- Hundreds of attributes are considered apart from the content such as hues / color shades used, which part of the day the movie is watched , the labels , designs on the cover etc
- Based on such micro segmentation , Netflix has created >80000 genres / segments that could be recommended to the customer
- Each genre is tagged based on the common attributes

GENOME SEGMENTS



Thank You Questions?



CPD Question