

# UNIT IV

## Scoring Your Customers Optimizing the CRM Process

# Introduction

- After a model has been created based on the historical data, it can then be applied to new data in order to make predictions about unseen behavior.
- The process of using a model to make predictions about behaviors that has yet to happen is called scoring.
- The output of the model prediction is called as score.
- Scores can take any form (numbers, strings, etc.).

# Process- Flow

- Eg: Database Marketing

1. The process begins with a database (Data warehouse, Datamart) containing information about customers or prospects.
2. Segmentation of the users in customer database is done based on a criteria.
3. The selected group of customer is scored using a predictive model. The score is a number between 0 and 1 represents the probability that the customer will purchase your product. The scores are placed in the DB with the customer ID.
4. The customers are sorted based on the score value. The top 25% of the scoring customers are stored in the new DB.
5. The customer details of the top 25% scoring customers are pulled out of the DW and they are used for sending the brochure

The marketing department decides on when and where to make campaigns.

Focus the campaign on the targeted customers.

Scoring model against a DB can be time consuming and Error prone activity.

To minimize the errors,

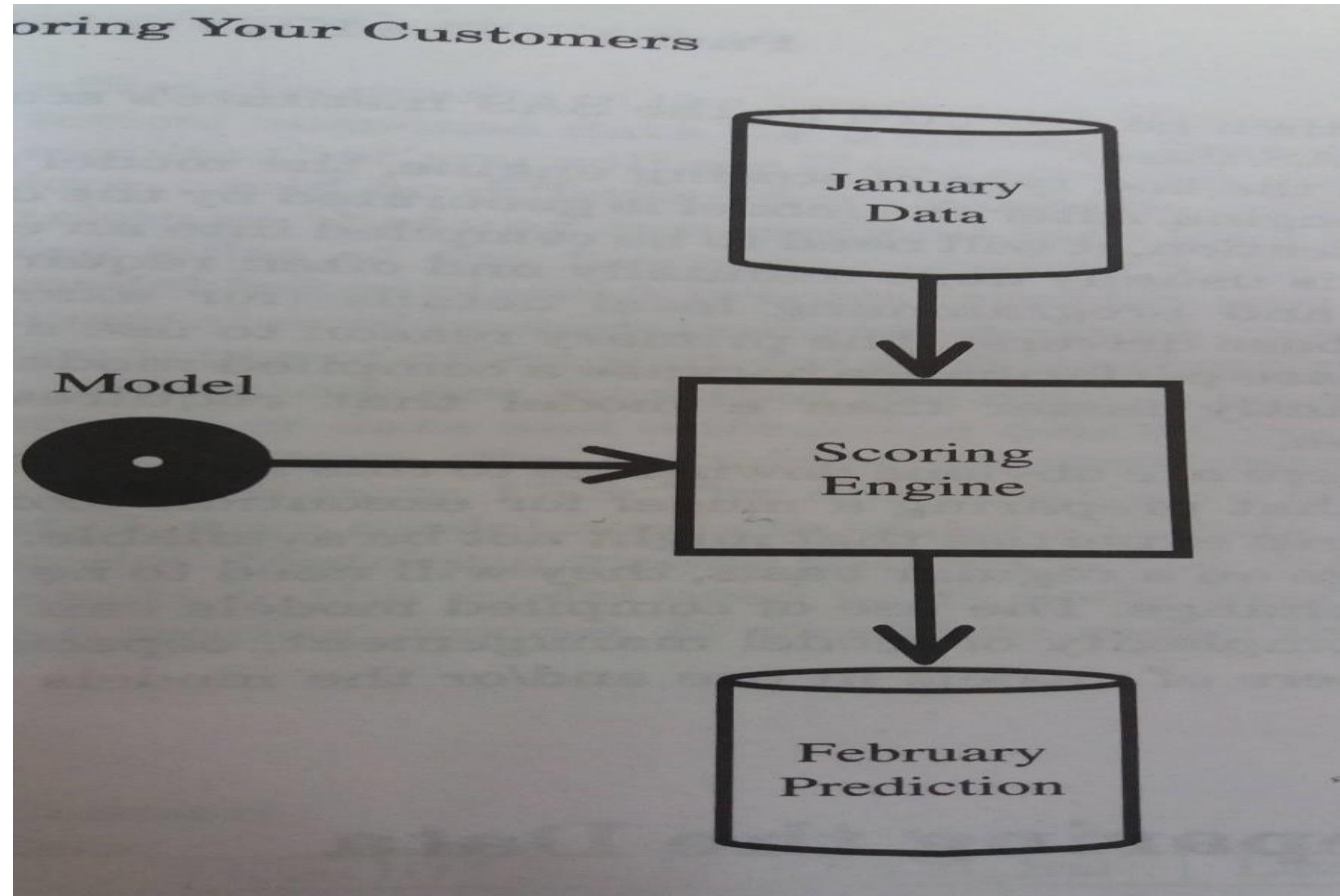
- To make sure that the definition of the segmentation is correct and select the right customer for scoring. The error can be due to incorrect translation.
- Make sure the correct customers are scored. There is a chance of using a wrong table.
- Make sure the correct model is used for scoring. There may be a large number of models available and the correct one is to be chosen.
- Make sure the scores are placed in the right place as there is confusion in the table containing the scores.
- Make sure that you understand how the scores are ordered. (high values are good or bad) understand which score values represent the good customers.

- When the frequency of marketing campaigns increased so that they occur on a daily or weekly basis, there are two significant impacts on the campaign. They are
  - **Decreased time between mailings**  
There is much less room for errors when carrying out the individual steps in the process. If a mistake is found, it needs less time to correct it compared to the less frequent campaigns.
  - The **total number of scoring events will increase** due to the increased frequency of campaigns and increase in the number of segments that need to be scored.

If the marketing campaigns that rely on the scores are run on a continuous basis, this means a lot of phone calls between marketing and IT, as well as marketing and the modelers.

Best approach to solve this problem is to use the campaign management software that is integrated with the scoring engine. If the integrated software is not available care will need to be taken so that difficulties are minimized.

# Scoring Architecture and configurations



# Scoring Architecture and configurations

- Scoring needs a piece of software called the **scoring engine (Prediction code)**.
- This takes the model and the dataset as the input and produces a set of scores as output.
- Common Approaches to scoring Engine
  - A scoring engine is different from the model building application.
  - A scoring engine is a part of the model building application.
  - A scoring engine is produced by compiling the model code. In this case the model itself is the scoring application.
- The type of model generated will depend upon the data mining system that is used. ( Multiple types of models or single type).

- In the first two cases, the scoring engine is a software application that needs to be run by the user. (GUI, Command line program, in which user gives the input when the program is run)
- Three inputs to the scoring engine
  - the model that is to be run
  - the data that is to be scored &
  - the location to store the output.
- In the third type of scoring engine, the model acts as its own scoring engine. After the model is generated by the data mining software application, it will need to be compiled into an executable form. It increases the performance because a compiled model will run faster.
- Cons: 1.Preparing a model for execution (Compiling, Linking, etc.) requires expertise.  
2. If the models change on a regular basis, recompilation is needed.  
3. It increases the complexity of model management.



# Preparing the data

- Before scoring, we need to prepare the data that needs to be scored.
- The **data used for scoring must be consistent** with the customer data that was used to build the model.

(Eg. Model was built using customers aged between 30 to 40, it should not be used for the age 40 to 50)

- (Interaction that will take place with the customer). The **customer interaction must also be consistent** with the original data or else the results will not be correct. (data for mailing in the case of marketing)
- After making sure the data is consistent with the historical customer data and the interaction details, **map individual columns in data set with the inputs of the model.**

# Mapping

**Figure 17-3**  
Mapping the data  
to the model

Customer  
Data:

Cust ID	Income	Cur_Bal	Last_Bal	Acct_Type	Age
---------	--------	---------	----------	-----------	-----

Inputs:

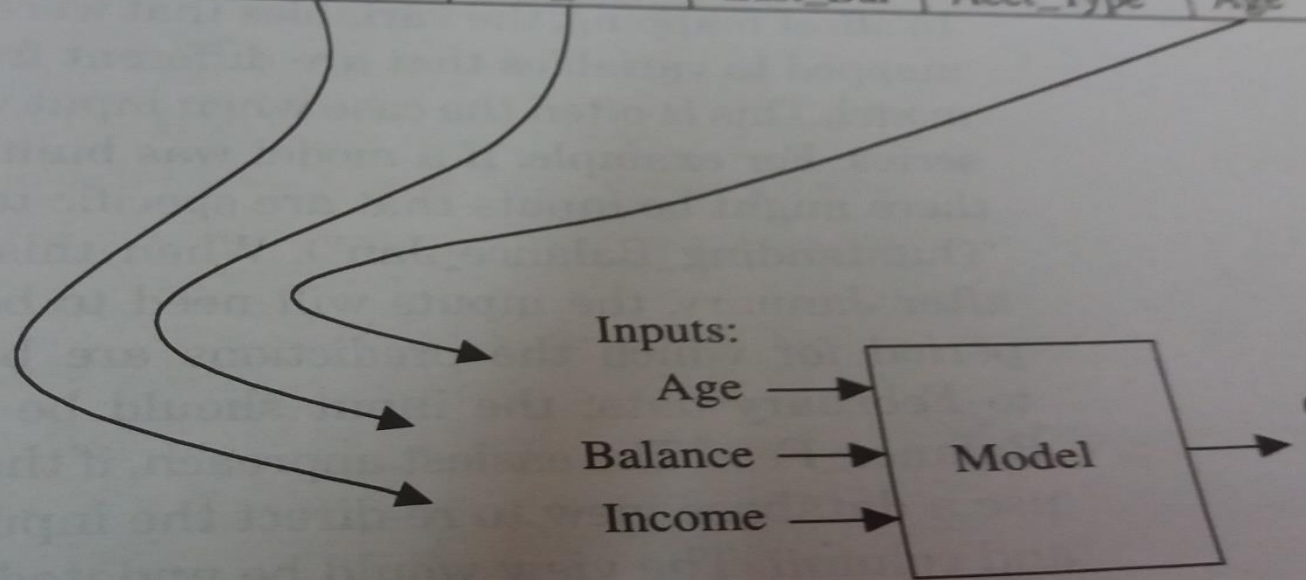
Age

Balance

Income

Model

Output:  
Score



# Types of mapping

While mapping the data in the database to the inputs of the model, there are two types of mapping, they are,

- 1. Direct Mapping :** the variable that was used to build the predictive model is included as the input and is mapped to the same variable.
  - Eg: Account type. Can be mapped to same variable.
  - This approach is used for variables that are not a part of the time series.
- 2. Indirect Mapping :** The variable used as inputs to the model are mapped to variables different from those used to build the model.
  - Eg: “outstanding\_balance\_Jan” is input specific to month.
  - It needs to be changed as “outstanding\_balance\_Feb ” for the month of Feb.

- In real world, scoring combines both direct as well as offset mapping.
- The last step in preparing the data is to transform the I/P to confirm to the specific requirements of the model.
- Eg: Account type in the database is represented as a string (savings or current) might be transformed into numbers (integers) before it can be fed into the model.
- The transformation is done by the person designing the model. Although this functionality should be incorporated into the model itself by the data mining system, some applications require the user to do the transformations manually.

# Integrating Scoring with other Applications.

- Integration of the DM applications with the other software can be done.
- Integrated with ERP, OLAP, campaign management and data visualization.
- Eg: DM system integration with marketing campaign management system.
  - We can segment customers to **yes or no** category and all customers in yes category will receive marketing offer.
  - Closer the DM system to campaign management system, it produces better results.

- The Integration is crucial in two areas.
  - The campaign management s/w must share definition of defined campaign segment with the DM application to avoid the modeling of the entire database.
  - The selected scores from the resulting predictive model must flow seamlessly into campaign segment in order to form targets with highest profit potential. Manual process for movement of the scores must be minimized due to errors and delay.

# Creating a model

- Predictive models are used in DM application.
- Modeling is completely separate from creating a campaign.
- Interaction with the campaign management begins when the model of sufficient quality is found. Then the model is exported.
- The exporting of model tells that the model exist and can be used in future.

## Dynamically scoring the data

- Dynamic scoring is a type of S/W integration that allows the scoring to be used by some other application for some other purpose. (Done based on the need dynamically)
- Dynamic – Scoring is done on the required data in the segment.
- Only relevant items are scored and assures fresh and up to date results.
- In Dynamic scoring, the scores are stored in temporary score table for processing.



# Optimizing the CRM Process

# Improved customer profitability through optimization

- **Optimization** is the science of optimally determining what can be done to make a **customer as profitable** as possible for as long as possible.
- Optimization can be done in order to **achieve the best results and take the right decisions**.
- Optimization is the single most important technique in the CRM system for increasing the customer profitability. It is typically applied to problems where there is a well defined measure of success.
- It is applied only in places where the results can be easily measured.
- Eg: airline scheduling
  - Optimizing the airline on correct schedule minimizes the waiting time in air and also the fuel costs and serve customers better.
- In marketing, only well defined problems are optimized.

- Optimizations can be done with
  - Targeted marketing,
  - Fraud detection,
  - Attrition prediction,
  - Cross selling,
  - Acquisition of new customers.

### To optimize something, you must have control over it

- In optimizing something, we must not only be able to measure the changes, but also control the change.
- In classical marketing, it is difficult to measure, also finding the effects is also difficult.
- But today, every type of marketing is controlled and also the effects are found and measured.

# Why now?

- Optimization is now possible because
  1. New technology and its maturation
  2. Data and its collection and storage.
  3. Changes in business process
- 1. The time is right also because of the maturation of new technologies like
  - Data Mining
  - Data Warehousing,
  - OLAP and
  - World Wide Web.
- 2. due to the quantity and the types of data that is collected and stored.
- 3. The time has come for optimization in the marketing function because of changes in business process like
  - Data based marketing,
  - ERP
  - E-Commerce and
  - CRM.

# Optimized CRM

- There is a lot of ways the marketing is performed. Sales looks like a system that can be optimized.
- Sales involves
  - Collection of data
  - There is something to be changed.
  - There is a way to measure the impact of the change.
- These factors make the sales to be optimized. We can start by recognizing what has worked in the past and trying necessary variations on it.
- Each try makes it closer to creating a more profitable customer.
- For optimization , select possible catalyst for improvement and optimally use the catalyst based on the past behaviour.

# The complete loop

- The three important steps in optimization process are,
  - **Measure : see what happened.**
  - **Predict : figure out something else to try based on what happened.**
  - **Act : Try it out.**
- First measure the outcome of the marketing intervention performed.
- Based on the measurements, make predictions about
  - What will happen next and
  - What changes will give useful results that can be measured.
- With all these information we must use it and act accordingly.

## Optimal CRM process: Measure, Predict, Act.

- They represent the steps (M, P, A) in a cycle optimizing come in cycle of CRM that is continuously improving.
  - This methodology is applied for customer management functions such as the following :
    - Cross selling – Selling a new product to an existing customer
    - Acquisition – Acquiring new valuable customer and
    - Retention of customers – Retaining existing valuable customer
  - To achieve the goals of the CRM, certain process like
    - Targeted marketing,
    - Lifetime value prediction,
    - Channel management (matching channel to the customer in most profitable way)
- must be done over and again.

- Marketing consist of certain interventions.
- Eg: Creative offers created based on the customer feedback.
- These help the customers to motivate and do more business.
- If the interventions are ready, we can make the remaining questions simple. The questions include.
  - What do I do to whom and when?
  - If there is a regular schedule in the marketing process, then the question is “What do I do to whom?”
  - The solution to the question is to create a model of what would be the outcome of the interventions.(predict).
- Whatever is predicted must be measured and recorded to know the predicted future behaviour.



# Optimization techniques (Neural Network, GA, Simulated Annealing)

- The first technique is the Neural Networks where there is a network with the links. There is value for each link and the input nodes, hidden nodes and the output nodes are present.
- The trick in training is to find out the correct numbers for the links.
- The Back propagation is the only way to find out the values for the links.
- The genetic algorithms can also be used to find out the link weights.
- Over time, the best solution gets better and better.

- GA simulate natural evolution on computer. They take the dna in genetics to the list of numbers that represent the link weights in the Neural networks.
- GA gives more number of right link weights and create a population of the link weights.
- The survival of the fittest is used to remove poorly performing networks.
- GA consumes more time and also faces the problem of getting stuck in a sub optimal solution.

# Simulated annealing and Neural networks

- This method helps to constantly improve to a better solution and is similar to the hill climbing.
- For neural networks, modifying the link weights will have a big impact on the performance of the network to be improved.
- Simulated annealing is the art of cooling a metal or glass at the right speed and minimize the number of defects in the structure of the material.

- This is applied to the neural network
  - **Make large changes to the weights in the early stages of training and then slowly decrease the amount of change made to the network so that it can zero into a best solution.**
- Large changes in weights at the beginning and changes become smaller as time goes on and hence the errors are removed at a finer level.