QN A

The Pre-modelling stage of the data mining process involves four steps.

Step 1: Identification of business problem.

The business problem could be to increase customers’ overall satisfaction with Sam’s based on various contributing factors. This ensures that the pizza restaurant remains competitive despite of the aggressive marketing by restaurants in the newly opened shopping mall.

Step 2: Translation of business problem into data mining application.

The objective of the data mining application is to model the relationship between customers’ satisfaction score for each factors and their overall satisfaction with Sam’s so as to determine the areas of improvement for the pizza restaurant.

Step 3: Assessment of data needed for the data mining application.

The data required can be generated internally by conducting a customer satisfaction survey. The survey result can be consolidated into a datasheet for use in the analysis. Additionally, customers’ transaction records or demographic information can be leveraged on to enrich the model.

Step 4: Preparation of data for data mining.

The preparation of data requires the data collected to be complete and accurate. Any discrepancies or noise in the data would impact the model’s accuracy. If the data come from various sources, efforts would be needed to ensure that they are stored in a standard format.

QN B

A potential business problem for Sam’s would be to increase customers’ overall satisfaction based on various contributing factors. In view of the aggressive marketing by restaurants in the newly opened shopping mall, it is natural that Sam’s existing customers may flock to the competitors. To prevent natural attrition, the customer satisfaction survey conducted by Sam’s can be used to better understand the pizza restaurant’s standing from customers’ perspective. In general, customers who are more satisfied are more likely to maintain their patronage with the restaurant. By analyzing the satisfaction score for each contributing factors, the customers’ overall satisfaction can be increased and thereby, aiding Sam’s in customer retention.

A data mining solution to this problem would be to construct a model with the objective of analyzing the relationship between customers’ satisfaction score for each factors and their overall satisfaction with Sam’s. This allows the pizza restaurant to better understand their customers’ point of view, what is good about the restaurant and what can be further improved on. The model result can be leveraged on to develop marketing strategy or increase operation efficiency of Sam’s.

QN I

Organization background and context

The organization discussed in this proposal is a financial institution with global footprints; it will be referred as SBM in this report due to privacy constraints. It has a large customer base of over a million clients and operates branches in various countries. This in turn generates a huge amount of data from the transactions and processes taking place in the day to day operations. The primary business of SBM is to provide financial services such as loans, securities, and futures and options to its clients. While the front office proactively engages the clients and brings in business for the firm, the back office provides a strong and comprehensive support to the firm.

The extensiveness of SBM’s business has become a challenge for the Credit Risk department to effectively mange the clientele portfolios. Situated in the back office, Credit Risk is responsible for setting credit limits, stress testing and reporting of clients’ risk profile and SBM’s credit health. To carry out its functions, Credit Risk need to transform the large amount of data associated with SBM’s clients and transaction into useful information where the team can perform analysis on. This is essential as the analysis will be depended on during decision making. To manage the data explosion, a data mining tool can be employed for the data analysis that the team needs to do. A data mining tool has strong analytical capabilities and is able to handle large amount of data. IBM SPSS Modeler, SAS Enterprise Miner and Statisca are examples of commercial data mining software available in the market.

Software 1 (description / functionalities / pro con / context fitting) = IBM SPSS Modeler

IBM SPSS Modeler is a powerful and versatile analytic workbench that allows user to build predictive models spontaneously without the use of programming language. It comes with a graphical interface for ease of visualization. The data mining process is represented as a workflow in the software; this promotes users’ interaction with the information where user can easily zoom into the details at any point of the model instead of spending time on programming task. It adopts the Cross Industry Standard Process for Data Mining (CRISP – DM), a structured and well proven methodology for solving business problems with data mining.

Software 2 (description / functionalities / pro con / context fitting) = SAS Enterprise Miner

SAS Enterprise Miner is an advanced data mining tool

Providing the risk associated with , it is experiencing difficulty in transforming the big data into useful information in digestible pieces for the management

Powerful analytical tool

Transform huge amount of data into digestible information for decision making

Dependent on availability and quality of data

Not perfectly accurate

Require domain knowledge + tool knowledge + IT skills

Substantial investment of resources,

Software 1 (description / functionalities / pro con / context fitting) = IBM SPSS Modeler

Software 2 (description / functionalities / pro con / context fitting) = SAS Enterprise Miner

Software 3 (description / functionalities / pro con / context fitting) = Statistica

Recommendation = SAS Enterprise Miner