MODULE 3

- ***Business Intelligence Concepts and Application***

- *By : ANIL MAHESHWARI*
- *1st EDITION – McGraw Hill*

TOPICS

- Business Intelligence Concepts and Application
- Data Warehousing.
- Data Mining
- Data Visualization.

3.1 : Business Intelligence Concepts and Application

- Introduction
- BI for better Decision.
- Decision type
- BI Tools
- BI Skills
- BI Application

Introduction

- BI used to analyze an organization's data and communicate the information to relevant users.

- The nature of life and business is to grow. Information is a life blood of business.

- Decision are made from fact and feelings. Data-based decision are more effective than those based on feelings alone.

- Actions based on accurate data, information, knowledge, experimentation, and testing using fresh insight and lead to sustained growth.

Introduction

- In hyperconnected world where everything is potentially connected to everything else; with potentially infinite correlations, data represents the impulse of nature in the form of certain events and attributes.

- A skilled business and motivated person is to use this cache of data to harness nature, and to find new niches of unserved oppertunities that could become profitable ventures.

- Khan Academy – BI in Education – Bill Gates Promoted.

BI FOR BETTER DECISIONS

- People uses crystal balls, astrology,palmistry, ground hogs, and also mathematics and numbers to migigate risk risk in decesion making.

- The goal is to make effective decesion, while reducing risk.

- Relaible knowledge about the future can help managers make the right decision with lower level of risks.

- Business calculate risks and make decesion based on a broad set of facts and insights.

BI FOR BETTER DECISIONS

- The speed of action has risen exponentially with the growth of the Internet, the speed of the decesion and the action can be key advantage.

- The internet and mobile technology allow decesion to be made anytime, anywhere.

- A sentiments expressed on social media should also be utilized as a potential sales and promotion oppertunity, while it lasts.

DECISION TYPES

- There are two kinds of decesion

   – Strategic Decesions

   – Operational Decesions.

- When people and organization act, new fact data are generated, where BI can be used.

- An unending process of generating fresh new insigths in real time can help make better decesions, and thus can be a significant competative advantage.

Strategic Decesion

- SD are those that impact the direction of the company.

- Decesion to reach out to a new customer set would be startegic decesion.

- In SD, the goal itself is may be maynot be clear and same is true for path to reach the goal.

- The consequence of decesion would be apparent some time later.

- Thus, one is constantly scanning for new possibilities and new path to achieve the goals.

- BI can help with what-if analysis of many possible scenarios.

Operational Strategy

- OS are more routine and tactical decesion, focused on developing greater efficiency. Updating an old website with new features will be an OS.

- OS can be made more effective using an analysis of old data.

- BI can help automated operation level decision making and improve effeciency by making millions of microlevel operational decesion in a model driven way.

- A decesion based model can be provided consistently .

- Developing such decesion tree models is one of the main applications of data mining tech.

BI TOOLS

- Information can be provided about the current state of affairs with the capability to drill down into details, and also about emerging pattern which leads to projections into the future.

- BI tools include data warehousing, online analytical processing, social media analytics, reporting, dashboards, querying and data mining.

- BI tools can be simple tools that could be considered end-user tools, to very sophisticated tools that can offer very broad and complex set of functionality.

BI Tools

- A Spreadsheet tool, such as Microsoft Excel can act as an easy but effective BI tool by itself.

- Data can be download and stored in the spreadsheet, then analyse to produce insights and presented in the form of grap and tables.

- This system offer limited automation using macros and other feature.

- The analytical feature include statistical and financial function.

- Pivot tables help do sophisticated what-if analysis.

BI Tools

- A dashboard system, such as IBM Cognos and Tableau, can offer a sophisticated set of tools for gathering, analyzing and presenting data.

- At the user end, modular dash boar can be designed and redesigned easily with a GUI.

- The back end data analytical capability include many statistical functions. The dashboard is linked to the datawarehouse at the backend to ensure that the tables and graph and other elements of the dashboard are updated in real time.

BI APPLICATION

- Every manager today needs access to BI tools to have up-to-date metrics about business performance. The following are some of the areas of BI and Data Mining.

Customer Relationship Management (CRM)

- A Business exists to serve a customer. A happy customer becomes a repeat customer.

- A Business should undersatnd the needs and sentiments of the customers.

- Sell more of its offering to the existing customers, and also expand the pool of customers it serves.

- BI application can impact many aspects of marketing.

Maximize the Return on Marketing Campaigns

- Understand the customers pain point from data ~~base analysis can ensure that marketing messages are fine~~tuned to better resonate with customers.

Improve Customers Retention

- It is more difficult and expensive to win new customers than it is to retain existing customers.

- Discount, free services to retain profitable customers in a cost effective manner.

Identify and Delight Highly Valued Customers.

- By segmenting the customers, the best customer can be identified.
- They can be proactively contacted, and delighted, with greater attention and better service.
- Loyalty program can be managed more effectively.

Fund-Raising from Alumini and Other Donors

- School can develop predictive models of the alumini that are most likely to pledge financial support to the schools.

- School can create a profile for alumini more likely to pledge donations to the schools.

- This can lead to a reduction in the cost of mailing and other forms of outreach to alumini.

Financial Services

- Stock brokers are an intensive users of BI systems.
- Fortunes can be made or lost based on access to accurate and timely information.

- INSURANCE
- This industry is a prolific user of prediction models in pricing insurance proposals and managing losses from claims against insured assets.

Banking

- Banking make loan and offer credit card to millions of customers.

- They are most interested in the improving the quality of loans.

- They also want to retain more good customers and sell more services to them.

Improve store layout and Sales promotions

- A market basket analysis can develop predictive models of the products often sold together.

- This knowledge of infinities between products can help retailers co-locate those products.

- Alternatively those infinity products could be located further apart to make the customers walk the length and breadth of the store, and thus be exposed to other products.

- Promotional discounted products bundle can be created to push a non selling items along with a set of products that sell together.

Manage Brand Image

- A business can create a listening post to listen to social media chatter about itself.

- It can do some sentiment analysis of the text to understand the nature of comments, and respond appropriately to the prospects and customers.

Education

- As higher education becomes more expensive and competative, it becomes a greater user of data-based decision-making.

- There is a strong needs for effeciency, increasing revenue, and improving the quality of students at all levels of education.

Healthcare and wellness

- Helathcare is one of the biggest sector in advanced economies.
- Evidence based medicine is the nevest trend in data based management.
- BI application can help apply the most effective diagnoses and prescription for various alignment.
- They can also help manage public health issues and reduce waste and fraud.

Other Areas of Applications

- Manufacturing
- Improve Product quality
- Telecom
- Marketing and Product Creation
- Network Failure Management
- Fraud Management
- Public Sectors
- Law Enforcement
- Scientific Research
- Prevent Machinery Failure.
- Maximize Customers Value.

Data Mining

- Data mining is the art and science of discovering knowledge pattern in the data.

- It is the act of extracting useful patterns from an organized collection of data.

- Pattern must be valid, potentially useful, and understandable.

- Past data can reveal pattern of activity that can be projected into the future.

ADVANTAGES

- The utilize the knowledge of data quality and data organizing from the database.

- It uses modeling and analytical technique from statistics and computer sceince areas (AI).

- Decesion making from the field of business management.

- Pattern recognization in defence "freind or foe" on a battle field.

Example

- Customers who buy cheese and milk also buy bread 90% of the time. Would be useful pattern for grocery store.

- People with blood pressure greater than 160 and age greater than 65 were at high risk of dying from heart attack, diagonestic value for doctor.

- Data scientist could at best predict correctly 59% of the time.

Gathering and Selecting Data

- The total amount of data in the world is doubling every 18 months. There is an ever growing avalanche of data coming with high velocity, volume and variety.

- Smart data mining requires chossing where to play, one has to make the decesion about what together and what to ignore, based on the purpose of the data mining exercise.

- To learn from data, quality data needs to be effectively gather, cleaned and organized and then effeciently mined.

- One requires the skills and technology for integrating of data elements from many resources.

Gathering and Selecting Data

- Most of the organization uses Enterprise Data Model (EDM) to organize their data
- EDM
  - Unified , high level model of all the data store in a organizations databse.
  - It is an inclusive of the data generated from all internal system.
  - Basic menu to create data in dataware house system for decesion making.
  - Can access data in Easy and usable manner.

Gathering and Selecting Data

- Gathering and curating unstructured and semi-structured data take time. Unstructured data can come in many form like DB, blogs, images, videos, audios and chat.

- There are stream of unstructured data like social media, blogs and chat

- There are streams of machine generated data like RFID tags, IoT etc.

- Choosing right data at right place is essential.

Data Cleansing and Preparation(GIGO)

- Data certainly needs to be cleansed and transformed before it can be used for data mining. There many ways in which data can be cleansed.

  – Filling missing values.

  – Transforming fields

  – Binning continous variables.

  – But it 60-80% time needed for the data mining project.

Data Cleansing and Preparation(GIGO)

- Duplicate Data need to be removed.

- Missing values needs to be filled in

- Data Elements should be comparable.

- Continuous value may need to be binned.

- Outlier data elements needs to be Removed.

- Data May need to be selected to increase information density.

OUTPUT OF DATA MINING

- Decesion Tree :

    – To make model based decesion.

    – Probabilities can be assigned to each branch.

    – If-then statement state the causality and can be mapped to the business rules.

    – Prediction based output.

- Regression Equation or Mathematical Fucntion

    – Curve to represent the data.

    – Classification execrcises.

    – Forcasting fromula.

OUTPUT OF DATA MINING

- Centroid : statistical measure for describing central tendencies of a collection of data points.

DATA MINING BEST PRACTICE

- Effective and successful use of data mining activity requires both business and technology skills.

- The IT components help fetch the data from many sources, clean up the data, assemble it to meet the needs of business problem, and then run the data mining technique on the platform.

- It divide and conquer the problem with smaller amount of data, and get closer to heart of the solution in iterative sequence of steps.

- The data mining industry has proposed a Cross-industry standard process for Data Mining (CRISP-DM). It has six essential steps

CRISP-DM

1. Business Understanding : *find the right business questions, while aswering it would lead to large payoff's for the organization, finacially or otherwise. Thinking outside the box is important, both in terms of a proposed model as well in the datasets available and required.*

2. *Data Understanding : undersatnd the data available for mining, Without related data, the hypothesis cannot be tested.*

3. *Data Preparation : The data should be relevant, clean and of high quality. It is important to assemble a team that has a mix of technical and business skill, who undersatnd the domain and data, this phase will take 60-7- percent of the time in data mining project.*

DIAGRAM

CROss-Industry Standard Process for Data Mining
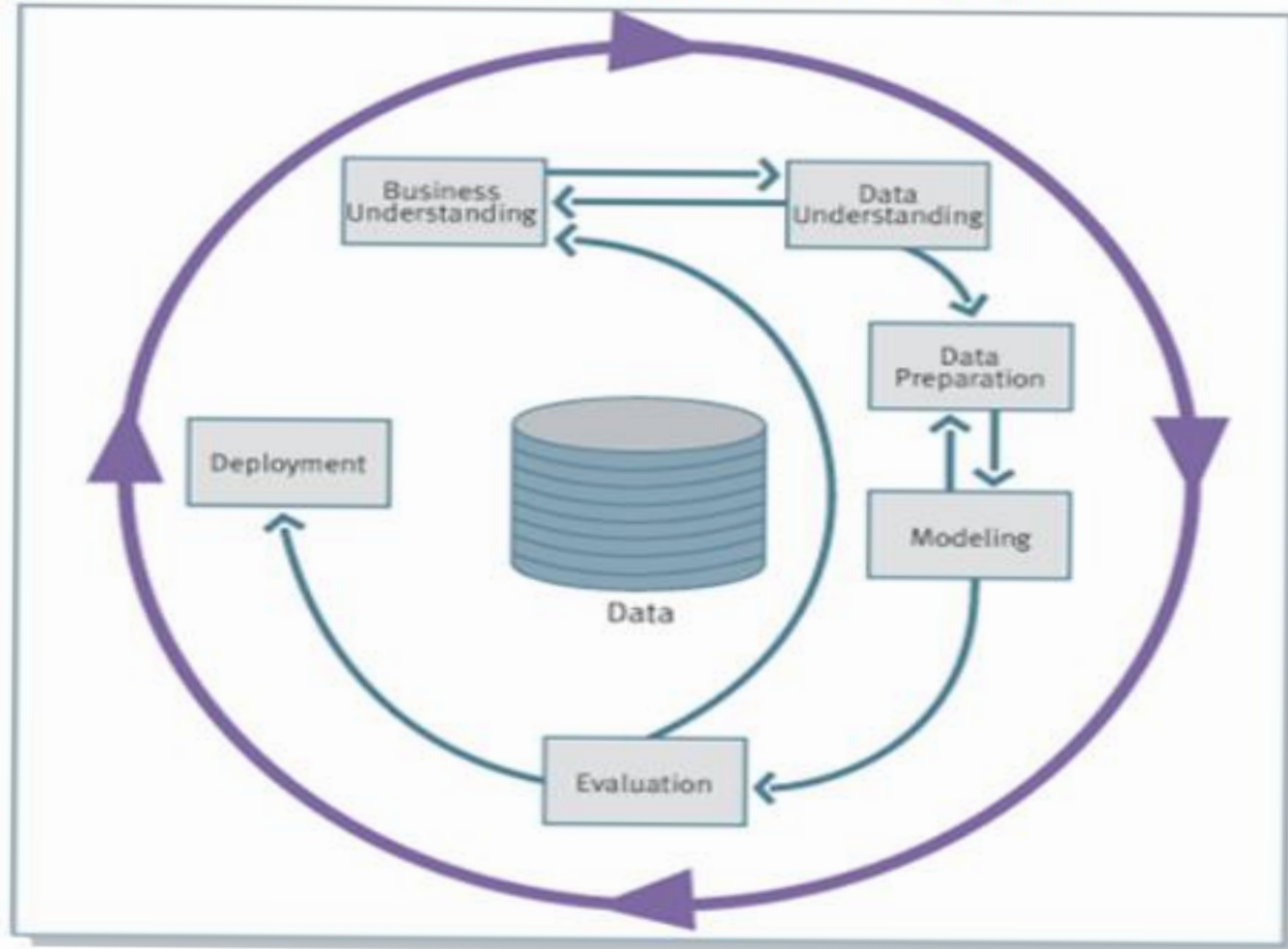
CRISP-DM

- 4. Modeling : task of running many algorithms using the available data to discover if the hypothesis are supported, patience is required in continuous engaging with the data until the data yields some good result.

- 5. Model Evaluation: one should not accept what the data says at first, it is better to apply multiple data mining techniques, when the accuracy has reached some satisfactory level, then the model should be deployed.

- 6. Deployment : represent the data mining solution with ket stakeholders and is deployed in the organization. Otherwise the project will be waste of time and setback for establishing and supporting a data-based decesion process culture in organization.

Data Visualization - Defination

- *Data Visualization is the art and science of making data easy to understand and consume for the end user.*

- *Ideal Visualization shows the right amount of data, in the right order, in the right visual form, to convey the high priority information.*

- *The visualization is the last step in the data life cycle. This is where the data is processed for presentation in an easy-to-consume manner to the right audience for the right purpose.*

EXCELLENCE IN VISUALIZATION

☐ Data can be represented in the form of rectangular tables or colorful graph of various types.

☐ Small, non-comparative, high labelled data-sets usually belongs to tables.

☐ When the amount of data grow, graph are preferable because it gives shape to data.

☐ The objectives of graphical excellence are

☐ **Show, and even Reveal, the data** : **the data should tell a story, specially the story hidden in large mass of data.**

☐ **Induce the viewer to Think of the Substance of the data**: **The format of the graph should be so natural to the data, that it hide itself and lets the data shine.**

Continue

- Make large data set coherent : Data will tell a comprehensive story when you bring large set of data together.

- Encourage the Eyes to Compare Different Pieces of Data : concentrate on different parts of the data.

- Reveal the data at several levels of details: Graphs leads to insights, which raise further curiosity, thus presentation help get to the root cause.

- Serve a Reasonably Clear Purpose : Information or decesion-making.