Business Intelligence

Dr. Dave E. Marcial

B. Common Functions of Business Intelligence Technologies

Common Functions of Business Intelligence Technologies

Reporting

data mining

predictive analytics

online analytical processing

text mining

Analytics

Benchmarking

business performance management

process mining

Data Mining

- Is the process of identifying valid, novel, potentially useful and ultimately comprehensible information from databases that is used to make crucial business decisions
- Predicts future trends and behaviors, allowing businesses to make proactive, knowledge-driven decisions
- the analysis step of the Knowledge Discovery in Databases process, or KDD
- is the process of extracting patterns from large data sets by combining methods from statistics and artificial intelligence with database management.

Questions are answered…

- What goods should be promoted to this customer?
- What is the probability that a certain customer will respond to a planned promotion?
- Can one predict the most profitable securities to buy/sell during the next trading session
- What medical diagnosis should be assigned to this patient?

Task Solved by DM

- **Predicting.** A task of learning a pattern from examples and using the developed model to predict future values of the target variable
- **Classification.** A task finding a function that maps an example into one of several discrete classes
- **Detection of relations.** A task of searching for the most influential independent variables for a selected target variable
- **Explicit modeling.** A task of finding explicit formulae describing dependencies between various variables
- **Clustering.** A task of identifying a finite set of categories or clusters that describe data
- **Deviation detection.** A task of determining the most significant changes in some key measures of data from previous or expected values

DM business opportunities

- **Automated prediction of trends and behaviors**
  - Targeted marketing, forecasting bankruptcy, identifying segments of a population likely to respond similarly to given events
- **Automated discovery of previously unknown patterns**
  - Analysis of retail sales data to identify seemingly unrelated products that are often purchased together
  - Detecting fraudulent credit card transaction and identifying anomalous data that could represent data entry **keying errors**
- **Databases can be larger in both depth and breadth**

Technologies Used in DM

- Neural networks
  - Nonlinear predictive models that learn through training and resemble biological neural networks in structure
- Rules induction
  - The extraction of useful if-then rules from data based on statistical significance
- Evolutionary programming
  - Automatically formulates hypothesis about the dependence of the target variable on other variables, in the form of programs expressed in an internal programming language
- Case based reasoning
  - To forecast a future situation, or to make a correct decision, such systems find the closest past analogs of the present situation and choose the same solution, which was the right one in those past situations.. That is why this method is called the nearest neighbor method

Technologies Used in DM

- Decision trees
  - Tree-shaped structures that represent sets of decisions
- Genetic algorithms
  - Optimization techniques that use processes such as genetic combination, mutation, and natural selection in a design based on the concepts of evolution
- Nonlinear regression models
  - Based on searching for a dependency of the target variable on other variables. Most applied in financial markets or medical diagnostics

Other DM algorithm-based technologies

- Classification
  - To predict one or more single numeric variables based on other variables in the data
  - Ex: decision tree
- Regression
  - To predict what a continuous stream of numbers-based, use of patterns
  - Ex: whether sales are lower in the summer moths or peak during the holidays
- Segmentation
  - To divide data into units or groups that have similar attributes (clustering algorithm)
- Association
  - To hunt correlations between particular items of data
  - EX: beer and diaper are often purchased at the same time, at single transaction, at a grocery store (market basket analysis)
- Sequence analysis
  - To identify common sequences in a series of data
  - EX: analyze your website, discover people often click its link in a certain order. (sequence clustering)

Business Cases for the data-mining algorithms

- Market Basket Analysis
  - To identify which items are generally purchased in the same check-out or shopping basket
- Churn Analysis
  - To identify the patterns behind customer churn (turnover)
- Market Analysis
  - Assist in grouping similar customers into different segments in order to better understand customer demographic
- Forecasting
  - Allows to input past data in order to predict future values such as inventory levels or sales information

Business Cases for the data-mining algorithms

- Data Exploration
  - Permits to explore the various components of data, analyzes profit margin of a particular product across demographic segments
- Unsupervised learning
  - Identifies relationships between components of your business that you ma not have known existed
- Web site Analysis
  - To fully understand how customers and potential customers use your website.

Business Cases for the data-mining algorithms

- Campaign analysis
  - Targets a marketing campaign and attempts to quantify the results (e.g. analyze how a particular product or demographic responds to a particular promo offer
- Information quality
  - Helps clean and organize data coming into a system
- Text analysis
  - To analyze feedback coming in from customers or clients

Common Functions of Business Intelligence Technologies

Reporting

data mining

predictive analytics

online analytical processing

text mining

Analytics

Benchmarking

business performance management

process mining

Online Analytical Processing (OLAP)

- A type of application that allows a user to interactively analyze data
- Online transaction processing that focuses on processing transactions such s orders, invoices or general ledger transactions
- Referred before as DSS

OLAP Applications

- Sales and marketing analysis
- Financial reporting and consolidation
- Budgeting and planning
- Product profitability and pricing analysis
- Activity based costing, Manpower planning
- Quality analysis

A dozen of rules in OLAP

1. Multidimensional conceptual view
   - Supports EIS “slice-and-dice” operations and is usually required in financial modeling
2. Transparency
   - OLAP should be part of an open system that supports heterogeneous data sources
3. Accessibility
   - Should be present the user with a single logical schema of the data

A dozen of rules in OLAP

4. Consistent reporting performance
   - Performance should ne degrade as the number of dimensions in the model increases
5. Client/server architecture
   - Requirement for open, modular systems
6. Generic dimensionality
   - Not limited to 3D and not biased towards any particular dimension

A dozen of rules in OLAP

7. Dynamic sparse-matrix handling
   - Related both to the idea of nulls in relational databases and to the notion of compressing large files, a sparse matrix is one in which not every cell contains
8. Multiuser support
   - Need to support multiple concurrent users, including their individual views or slices of a common database
9. Unrestricted cross –dimensional operations
   - All dimensions are created equal, and operations across data dimensions do not restrict relationships between cells

A dozen of rules in OLAP

10. Intuitive data manipulation
    - Ideally users should not have to use menus or perform complex multiple –step operations when an intuitive drag-and-drop action will do
11. Flexible reporting
12. Unlimited dimensional and aggregation levels

OLAP Key features

- Multidimensional views
- Calculation-intensive capabilities
- Time intelligence

Common Functions of Business Intelligence Technologies

Reporting

data mining

predictive analytics

online analytical processing

text mining

Analytics

Benchmarking

business performance management

process mining

Data Analytics (DA)

- is the science of examining raw data with the purpose of drawing conclusions about that information.
- is used in many industries to allow companies and organization to make better business decisions and in the sciences to verify or disprove existing models or theories.
- Data analytics is distinguished from data mining by the scope, purpose and focus of the analysis.
- Data miners sort through huge data sets using sophisticated software to identify undiscovered patterns and establish hidden relationships.
- Data analytics focuses on inference, the process of deriving a conclusion based solely on what is already known by the researcher.

Data Analytics (DA)

- is the application of computer technology, operational research, and statistics to solve problems in business and industry.
- is carried out within an information system: while, in the past, statistics and mathematics could be studied without computers and software, analytics has evolved from the application of computers to the analysis of data and this takes place within an information system or software environment.

Common Functions of Business Intelligence Technologies

Reporting

data mining

predictive analytics

online analytical processing

text mining

Analytics

Benchmarking

business performance management

process mining

Process mining

- is a process management technique that allows for the analysis of business processes based on event logs.
- The basic idea is to extract knowledge from event logs recorded by an information system.
- aims at improving this by providing techniques and tools for discovering process, control, data, organizational, and social structures from event logs.

Process mining

- techniques are often used when no formal description of the process can be obtained by other approaches, or when the quality of an existing documentation is questionable.
- For example, the audit trails of a workflow management system, the transaction logs of an enterprise resource planning system, and the electronic patient records in a hospital can be used to discover models describing processes, organizations, and products.
- Moreover, such event logs can also be used to compare event logs with some *a priori* model to see whether the observed reality conforms to some prescriptive or descriptive model.

Common Functions of Business Intelligence Technologies

Reporting

data mining

predictive analytics

online analytical processing

text mining

Analytics

Benchmarking

business performance management

process mining

Business performance management

- is a set of management and analytic processes that enable the management of an organization's performance to achieve one or more pre-selected goals.
- Synonyms for "business performance management" include **"corporate performance management"** and **"enterprise performance management"**

Common Functions of Business Intelligence Technologies

Reporting

data mining

predictive analytics

online analytical processing

text mining

Analytics

Benchmarking

business performance management

process mining

Benchmarking

- is the process of comparing one's business processes and performance metrics to industry bests and/or best practices from other industries.
- Dimensions typically measured are quality, time and cost.
- In the process of benchmarking, management identifies the best firms in their industry, or in another industry where similar processes exist, and compare the results and processes of those studied (the "targets") to one's own results and processes.
- In this way, they learn how well the targets perform and, more importantly, the business processes that explain why these firms are successful.

Benchmarking Types

- Process benchmarking
- Financial benchmarking
- Benchmarking from an investor perspective
- Performance benchmarking
- Product benchmarking
- Strategic benchmarking
- Functional benchmarking
- Best-in-class benchmarking
- Operational benchmarking
- Energy benchmarking

Common Functions of Business Intelligence Technologies

Reporting

data mining

predictive analytics

online analytical processing

text mining

Analytics

Benchmarking

business performance management

process mining

Text mining

- sometimes alternately referred to as *text data mining*, roughly equivalent to *text analytics*
- refers to the process of deriving high-quality information from text. High-quality information is typically derived through the devising of patterns and trends through means such as statistical pattern learning.
- usually involves the process of structuring the input text (usually parsing, along with the addition of some derived linguistic features and the removal of others, and subsequent insertion into a database), deriving patterns within the structured data, and finally evaluation and interpretation of the output.
- 'High quality' in text mining usually refers to some combination of relevance, novelty, and interestingness.

Text mining Tasks

- text categorization,
- text clustering,
- concept/entity extraction,
- production of granular taxonomies,
- sentiment analysis,
- document summarization, and
- entity relation modeling (*i.e.*, learning relations between named entities).

Common Functions of Business Intelligence Technologies

Reporting

data mining

predictive analytics

online analytical processing

text mining

Analytics

Benchmarking

business performance management

process mining

Predictive analytics

- is an area of statistical analysis that deals with extracting information from data and using it to predict future trends and behavior patterns.
- relies on capturing relationships between explanatory variables and the predicted variables from past occurrences, and exploiting it to predict future outcomes.
- the accuracy and usability of results will depend greatly on the level of data analysis and the quality of assumptions.

Predictive analytics

- **Predictive models**

- analyze past performance to assess how likely a customer is to exhibit a specific behavior in the future in order to improve marketing effectiveness.
- encompasses models that seek out subtle data patterns to answer questions about customer performance, such as fraud detection models.
- perform calculations during live transactions, for example, to evaluate the risk or opportunity of a given customer or transaction, in order to guide a decision.
- With advancement in computing speed, individual agent modeling systems can simulate human behavior or reaction to given stimuli or scenarios.
- The new term for animating data specifically linked to an individual in a simulated environment is avatar analytics.

Predictive analytics

- **Descriptive models**

- quantify relationships in data in a way that is often used to classify customers or prospects into groups.
- Unlike predictive models that focus on predicting a single customer behavior (such as credit risk), descriptive models identify many different relationships between customers or products.
- do not rank-order customers by their likelihood of taking a particular action the way predictive models do.
- can be used, for example, to categorize customers by their product preferences and life stage.
- can be utilized to develop further models that can simulate large number of individualized agents and make predictions.

Predictive analytics

- **Decision models**

- describe the relationship between all the elements of a decision — the known data (including results of predictive models), the decision and the forecast results of the decision — in order to predict the results of decisions involving many variables.
- These models can be used in optimization, maximizing certain outcomes while minimizing others.
- Decision models are generally used to develop decision logic or a set of business rules that will produce the desired action for every customer or circumstance.

Common Functions of Business Intelligence Technologies

Reporting

data mining

predictive analytics

online analytical processing

text mining

Analytics

Benchmarking

business performance management

process mining