SNS COLLEGE OF TECHNOLOGY
(An Autonomous Institution)
COIMBATORE-35

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

19ITE305 BIG Data Analytics

Agenda

- Classification of analytics
- Top challenges facing big data
- Why is big data analytics important?
- What kind of technologies are we looking towards to help meet the challenges posed by big data?
- Data science

Classifiaction

- Two ways
  1. Classify analytics into
     * basic
     * operationalized
     * advanced
     * monetized
  2. Classify analytics into
     * analytics 1.0
     * analytics 2.0
     * analytics 3.0

First School of Thought

- Basic analytics

1. The primarily slicing and dicing of data to help with basic business insights.

2. This is about reporting on historical data, basic visualization etc.

- Operationalized analytics

  It is about enterprise's business process

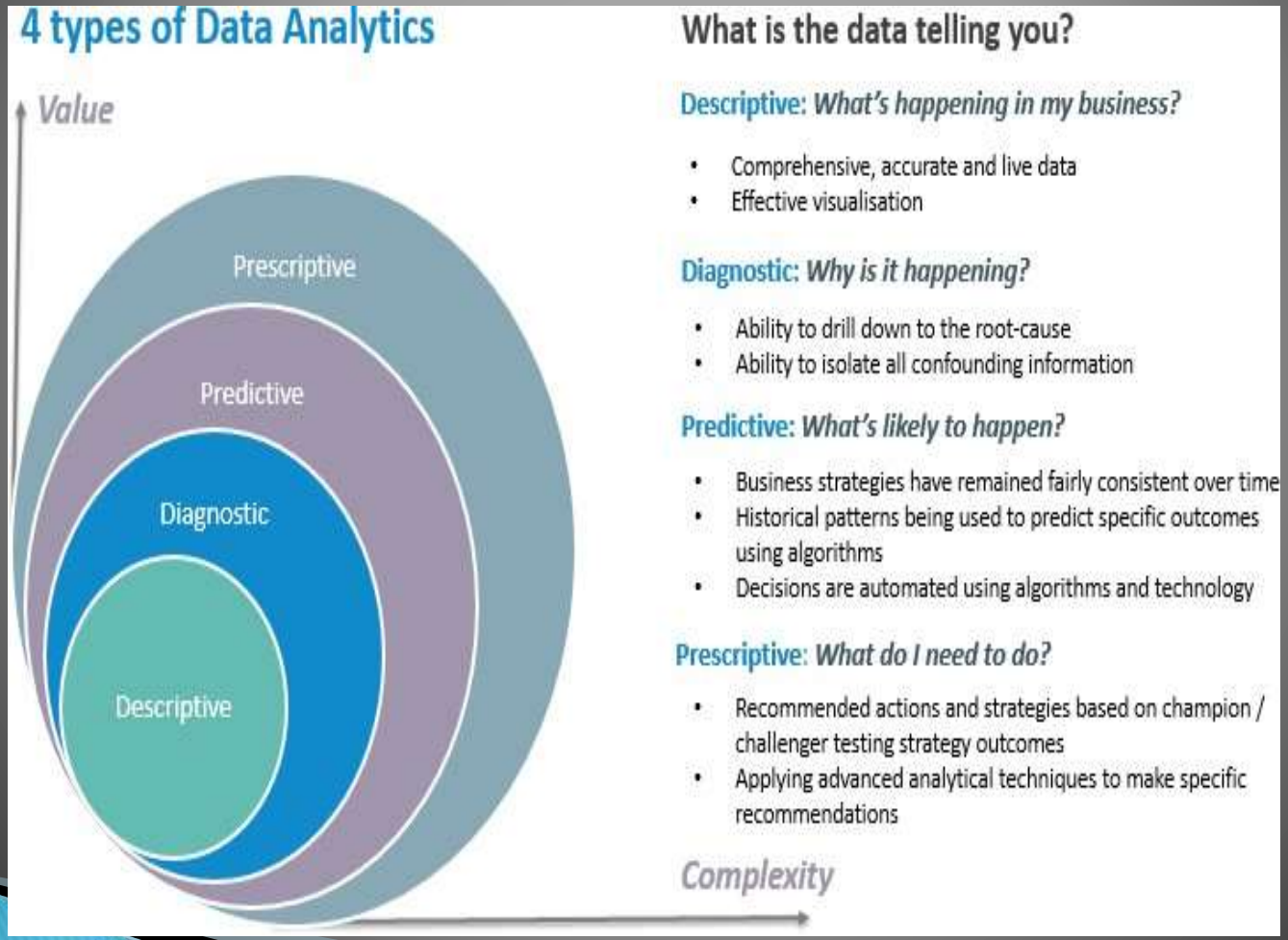
- Advanced analytics

  It is about forecasting for the future by the way of predictive and perspective modelling

- Monetized analytics

  in use to derive direct business revenue

TYPES OF DATA ANALYTICS



What is the data telling you?

Descriptive: ***What's happening in my business?***

- Comprehensive, accurate and live data
- Effective visualisation

Diagnostic: ***Why is it happening?***

- Ability to drill down to the root-cause
- Ability to isolate all confounding information

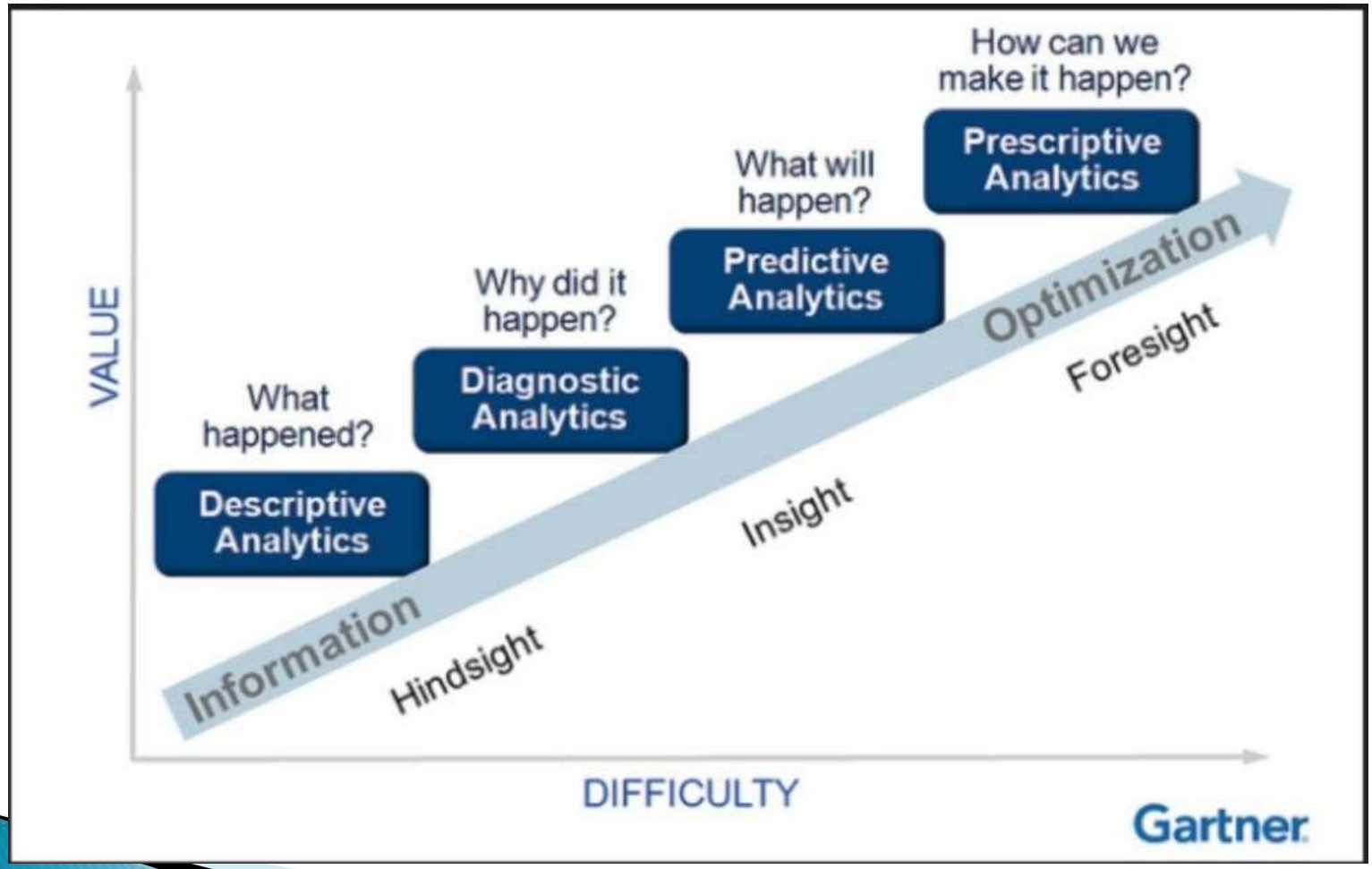
Predictive: ***What's likely to happen?***

- Business strategies have remained fairly consistent over time
- Historical patterns being used to predict specific outcomes using algorithms
- Decisions are automated using algorithms and technology

Prescriptive: ***What do I need to do?***

- Recommended actions and strategies based on champion / challenger testing strategy outcomes
- Applying advanced analytical techniques to make specific recommendations

TYPES OF DATAANALYTICS



- **Descriptive analytics:** What has happened and what is happening right now? Descriptive analytics uses historical and current data from multiple sources to describe the present state by identifying trends and patterns.

- Descriptive analytics can help to identify theare as of strength and weakness in an organization

- **Diagnostic analytics:** Why is it happening? Diagnostic analytics uses data (often generated via descriptive analytics) to discover the factors or reasons for past performance.

Predictive analytics

- What is likely to happen in the future?
- Predictive analytics applies techniques such as statistical modeling, forecasting, and machine learning to the output of descriptive and diagnostic analytics to make predictions about future outcomes.
- Predictive analytics is often considered a type of "advanced analytics," and frequently depends on machine learning and/or deep learning.
- Predictive models help make weather forecasts,develop video games, translate voice-to-textmessages, customer service decisions, and developinvestment portfolios.

Prescriptive analytics:

- What do we need to do? Prescriptive analytics is a type of advanced analytics that involves the application of testing and other techniques to recommend specific solutions that will deliver desired outcomes. In business, predictive analytics uses machine learning, business rules, and algorithms.

Second school of thought

- Analytics 1.0
  * Era: 1950 to 2009
  * Descriptive analytics
  * Data from legacy systems, ERP,CRM and 3rd party apllications
  * small and structured data
  * data stored in data warehouses or data marts
  * Relational database

- Analytics 2.0
  * Era: 2005 to 2012
  * Descriptive statistics +predictive
  * Big Data
  * Unstructured data
  * data stored in massive parallel server
  * Hadoop cluster

- Analytics 2.0
  - * Era: 2012 to present
  - * Descriptive statistics +predictive +persepective
  - * Big Data +data from ERP,CRM and 3rd party
  - * Unstructured data
  - * data stored in massive parallel server
  - * machine learning techniques

Challenges in Big Data

- Scales
- Security
- Schema
- Continous availablity
- Consistency
- Data quality
- Partion tolerant

What kind of technologies are we looking towards to help meet the challenges posed by big data?

- Cheap and abundant storage
- Need a fast processors to help with quicker processing of big data
- Affordable open source,distributed big data platforms such as hadoop
- Cloud computing and other flexible resorce allocation arrangement.
- Parallel processing, clustering, virtualization, large grid enviornment,high connectivity