ACADGILD

Data Analytics with R, Excel and Tableau

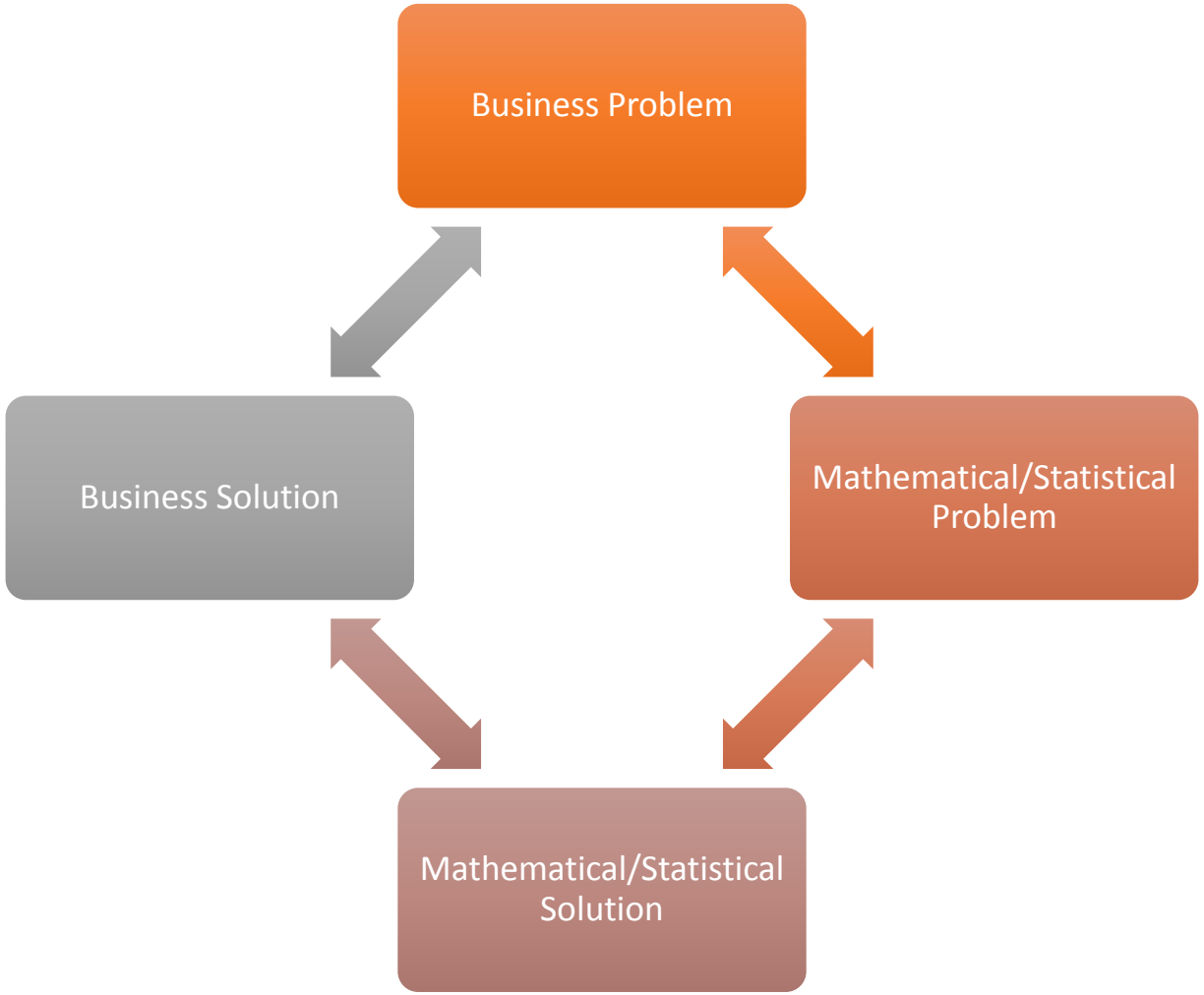
Session 1 - Business Analytics with R

AGENDA

What is Business Analytics?

Copyrights© 2017, AcadGild. All rights reserved

What is Business Analytics?



Copyrights© 2017, AcadGild. All rights reserved

What is Business Analytics?

- Use of data and information (both structured, semi-structured and unstructured) along with human intelligence
- In combination with information technology, software's, frameworks and tools
- Triggering informed business decision making, so that business value can be optimized
- A very structured thinking process
- Involves a lot brainstorming sessions with all the stake holders
- Data can be in any form, that has no defined source

Copyrights© 2017, AcadGild. All rights reserved

Applications/Use Cases

- Digital marketing - How to decide to whom should a digital marketing manager sent an email?
- Promotions - How do a company decide which product to put on sale and how much discount should be given to whom?
- Healthcare - What is the probability that a person will get readmitted to the hospital after 30 days from the date of procedure?
- Insurance - Who is going to buy my insurance? What should be the premium amount?
- Customer relationship - How to identify the pain areas of the customers?

Copyrights© 2017, AcadGild. All rights reserved

Applications/Use Cases

- Stock market - Forecasting stock prices given the historical stock price data
- Sports - How to decide the auction price of all the players in IPL event?
- Entertainment - Who should be the brand ambassador for a product? How the decision is being taken?
- Pricing decisions - How the price of the product determined?

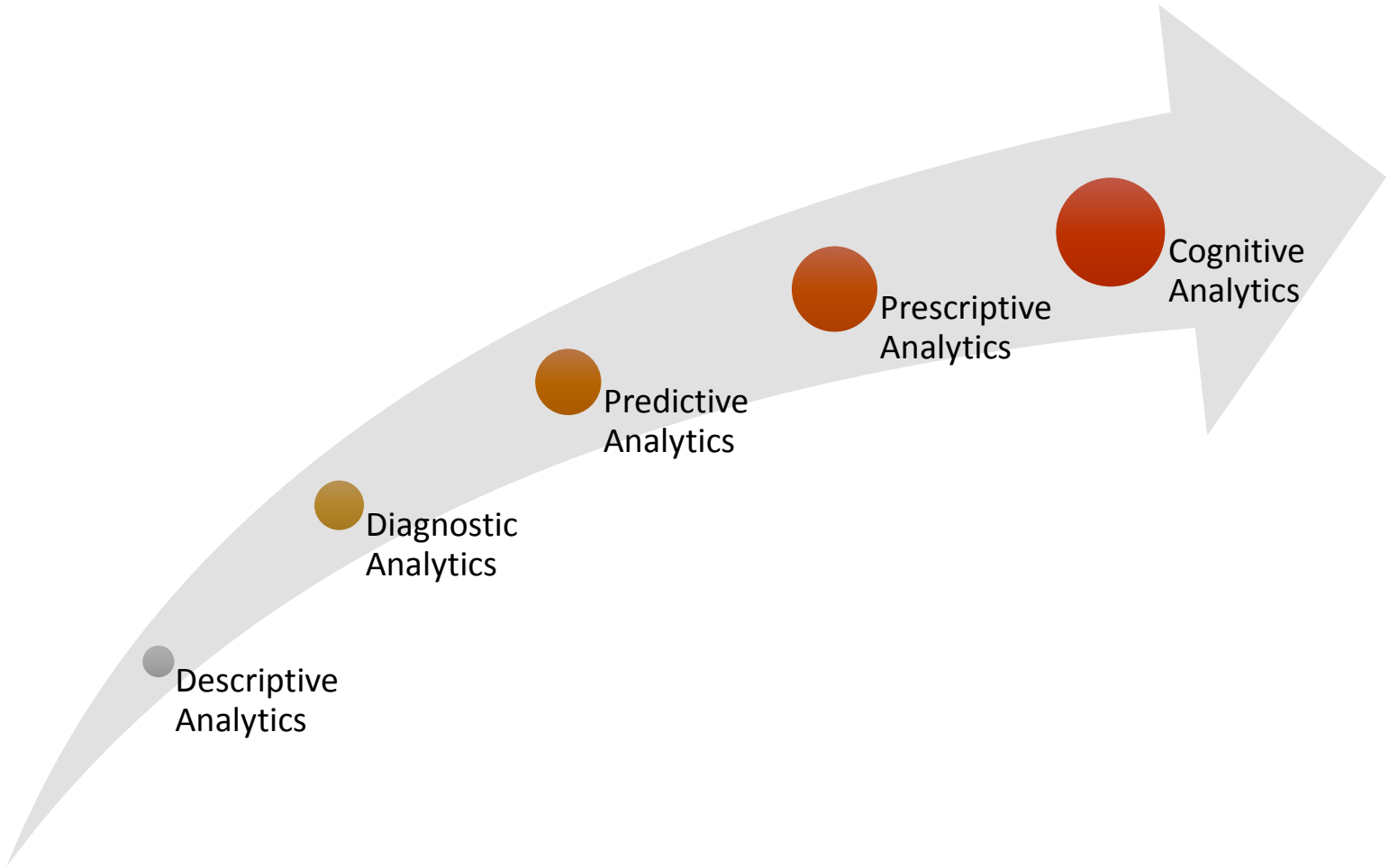
Copyrights© 2017, AcadGild. All rights reserved

Relevance of BA

- Helps in formulation of right strategies in right time
- Helps in driving smart decision making
- Helps in achieving the business target in a time bound manner
- Helps in driving operational efficiency
- Helps in driving profitability
- Helps in getting a clear picture about the data through better data visualization

Copyrights© 2017, AcadGild. All rights reserved

Stages of BA



Copyrights© 2017, AcadGild. All rights reserved

Descriptive Analytics

- A method of finding the historical trends
- Helps in reflecting the as is scenario
- Shows what has happened in past
- Shows how the business has performed in past
- Can be delivered using business intelligence tools
- Shows basic statistics about any business metrics
- Easily interprets Business metrics
- At this stage what happened will be analyzed

Copyrights© 2017, AcadGild. All rights reserved

Diagnostic Analytics

- Key factors will be analyzed in this stage
- Analyses root cause of the problems
- Identifies different stages of the problem
- Analyzes why the problem happened
- Details all the axis of the problem
- Involves a lot of reporting
- Various different reports indicate just another perspective
- Involves some degree of correlation and association

Copyrights© 2017, AcadGild. All rights reserved

Predictive Analytics

- This step shows what is likely to happen?
- Can we predict the future events, given the historical data?
- How well we can predict given the features?
- Depends on the previous stage, where we are creating the factors that affect the problem statement
- Involves a lot of statistical model building
- Statistical model also requires a sense of probability
- Probability distribution and future prediction is related

Copyrights© 2017, AcadGild. All rights reserved

Prescriptive Analytics

- Stage prescribes recommendations
- If a certain business metric predicted to go down, what recommended action is required.
- Helps a business in forming strategies
- Helps in preparing the plan B if the things are not going in right direction
- A reactive analysis, as something happens what action required and if not what alternate action required.
- A set actions always triggered based on predictions

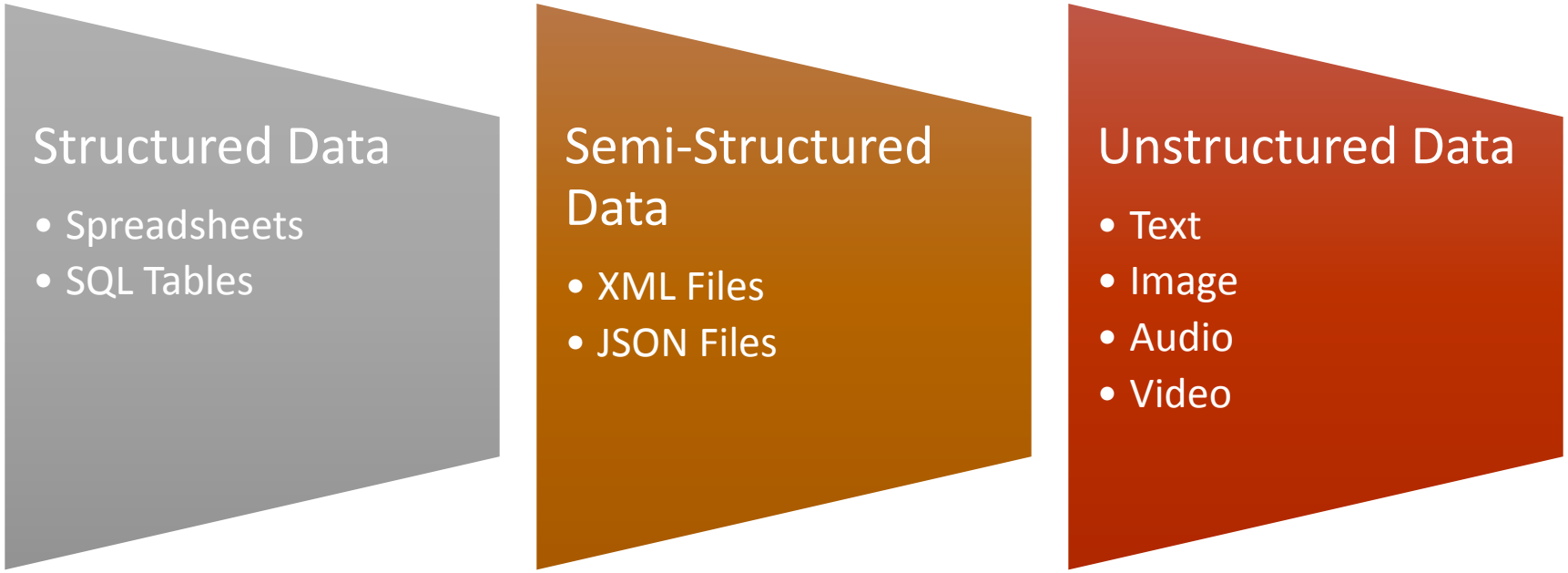
Copyrights© 2017, AcadGild. All rights reserved

Cognitive Analytics

- This is the artificial intelligence and machine learning layer
- If we know the prediction and recommended action to be taken:
  - Can any machine take decisions?
  - Can we do any what-if analysis?
  - Can machines decide what recommendation to be given?
- The cognitive function, based on reasoning and logic
- Involves machine learning models
- Involves a set of rules

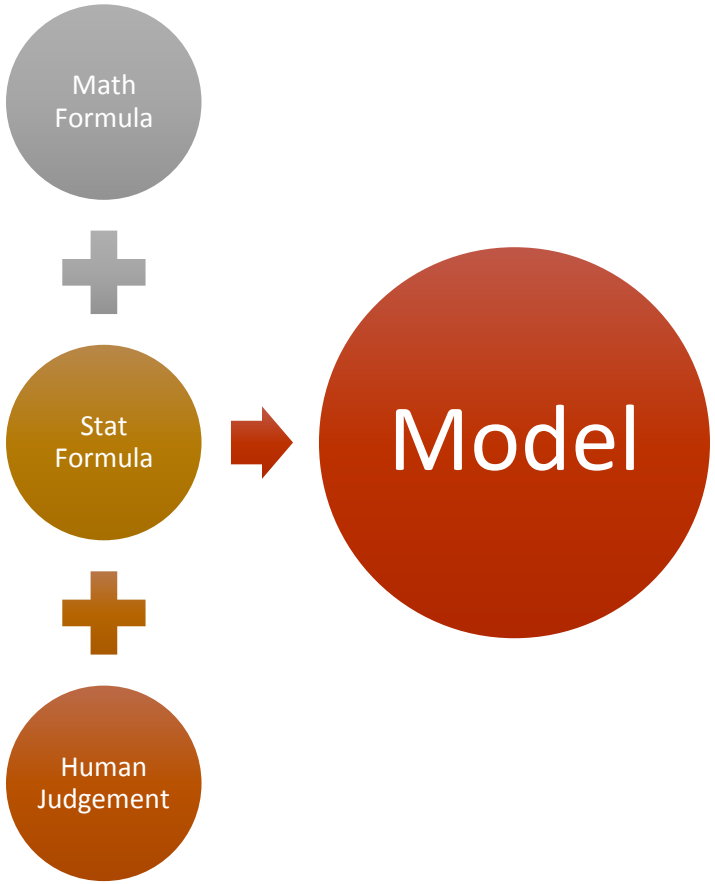
Copyrights© 2017, AcadGild. All rights reserved

Sources of Data



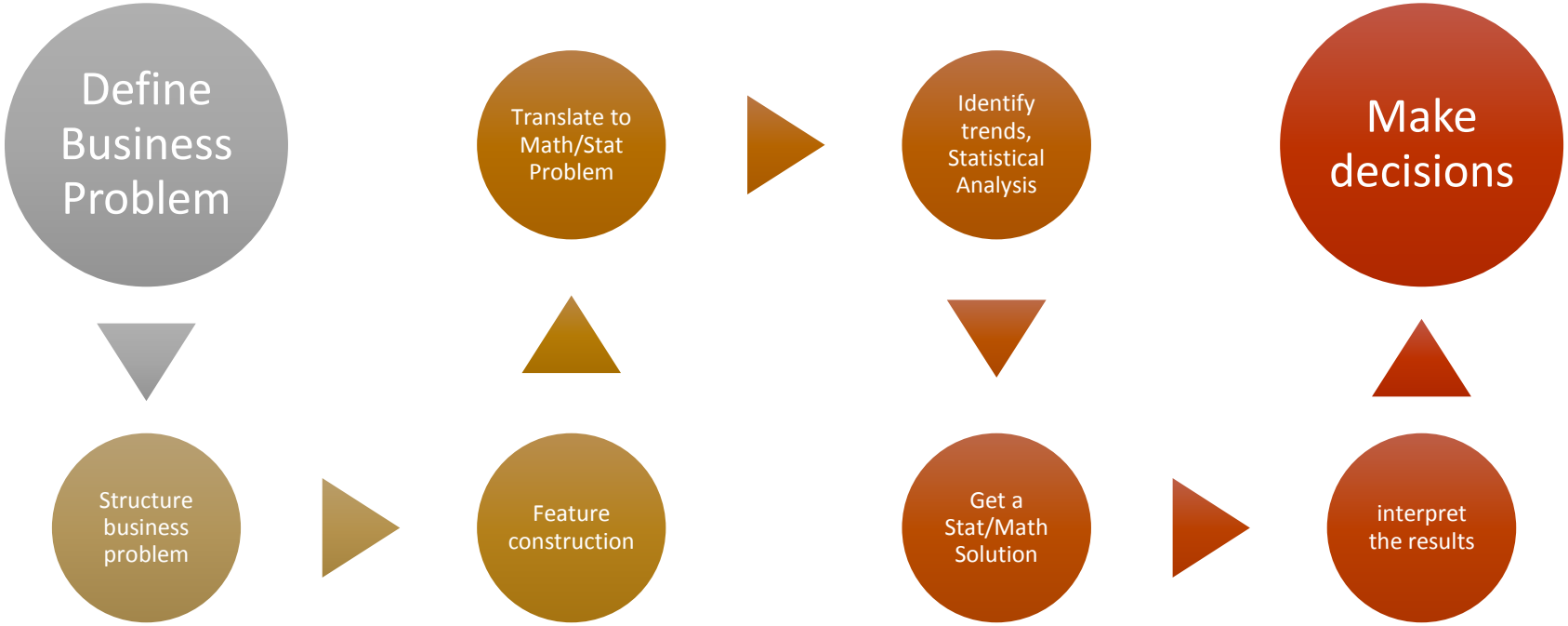
Copyrights© 2017, AcadGild. All rights reserved

What is a Model?



Copyrights© 2017, AcadGild. All rights reserved

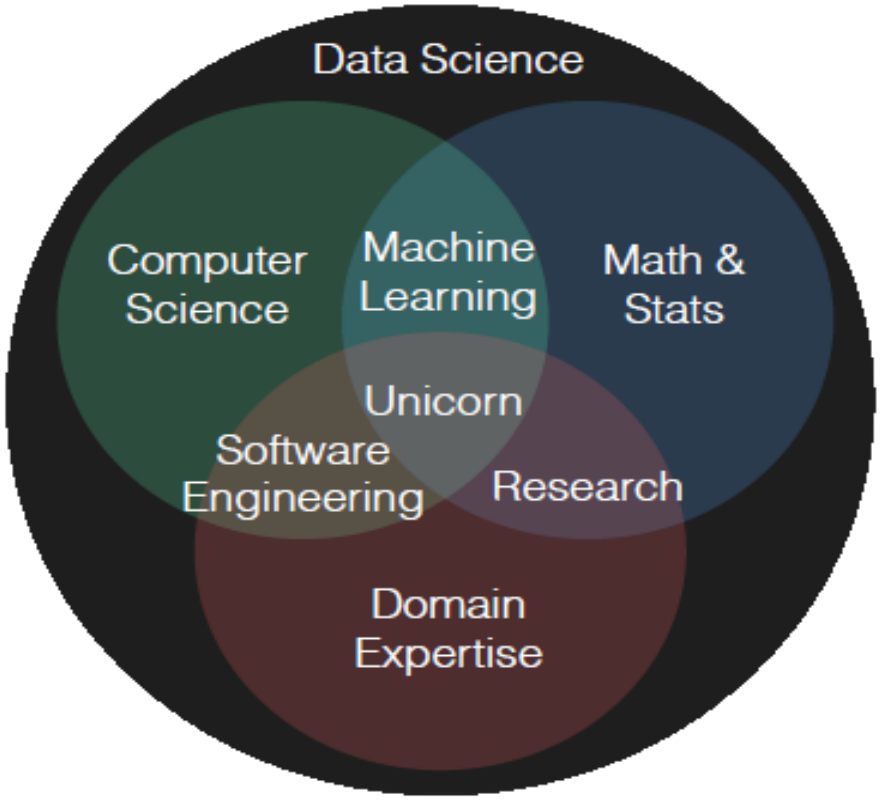
Steps in Problem Solving



Copyrights© 2017, AcadGild. All rights reserved

What is Data Science?

- Data science is a multidisciplinary blend of data inference, algorithm development, and technology in order to solve analytically complex problems.



Copyrights© 2017, AcadGild. All rights reserved

What is Data Science – Domain Expertise

- Analytical Problem-Solving - Know how to approach high-level challenges with a clear eye on what is important; Employ the right approach and methodology to make the maximum use of time and human resources.
- Effective Communication - Detail your techniques and discoveries to technical and non-technical audiences in a language they can understand.
- Intellectual Curiosity - Explore new territories and find new creative and unusual ways to solve challenges.
- Industry Knowledge - Understand how the industry you work in functions and how data is collected, analyzed and utilized to make decisions.

Copyrights© 2017, AcadGild. All rights reserved

What is Data Science – Math and Stats

- Mathematical and statistical skills also equally important.
- Evolution of low cost devices like smarter computing machines provides more flexibility to the mathematicians and statisticians to write complex algorithms that can process data very quickly.
- As a data scientist a person should be good at matrix algebra, linear algebra and basic mathematics and statistics.

Copyrights© 2017, AcadGild. All rights reserved

What is Data Science – Computer Science

- The computer science contributes in making newer programming languages, newer databases, smarter implementation of various newer tools, and integration of applications for smooth data operations.
- As a Data Scientist, a person should have idea about computer science fundamentals and basic programming knowledge in any one language.
- If someone is very good at Mathematics and Statistics, understands computer vision, then computers can be trained to perform certain task automatically by learning new patterns from data.
- Machine Learning is that branch of data science, where the system gets better as it learns new patterns over a period of time, specific to a particular task.
- As a Data Scientist a person should be thinking towards large scale machine learning.

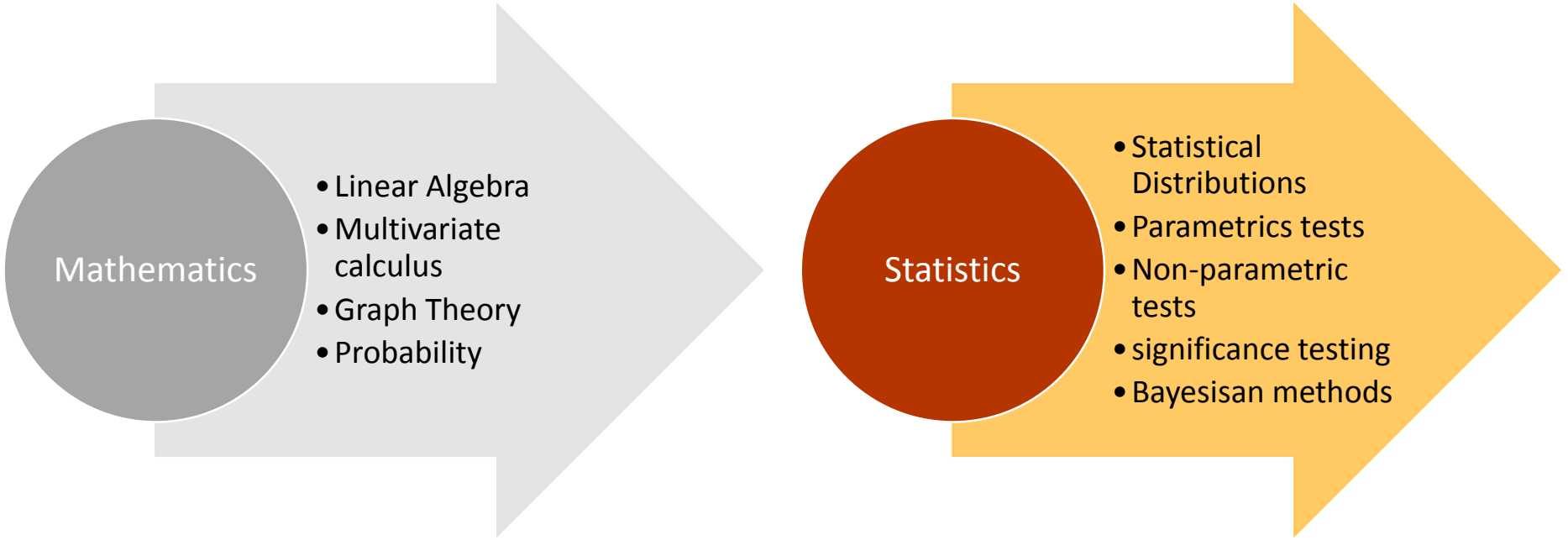
Copyrights© 2017, AcadGild. All rights reserved

A day as a Data Scientist

- A Data Scientist tries to understand the business problem, availability of data relating to a business problem and additional data that they may need to build the solution
- A Data Scientist formulates hypothesis from the data relating to the business problem
- A Data Scientist runs various experiments with data by applying mathematical and statistical techniques for data discovery and pattern recognition
- Finally, a Data Scientist conveys relevant business stories to the stake holders about their business problem and possible recommendations.

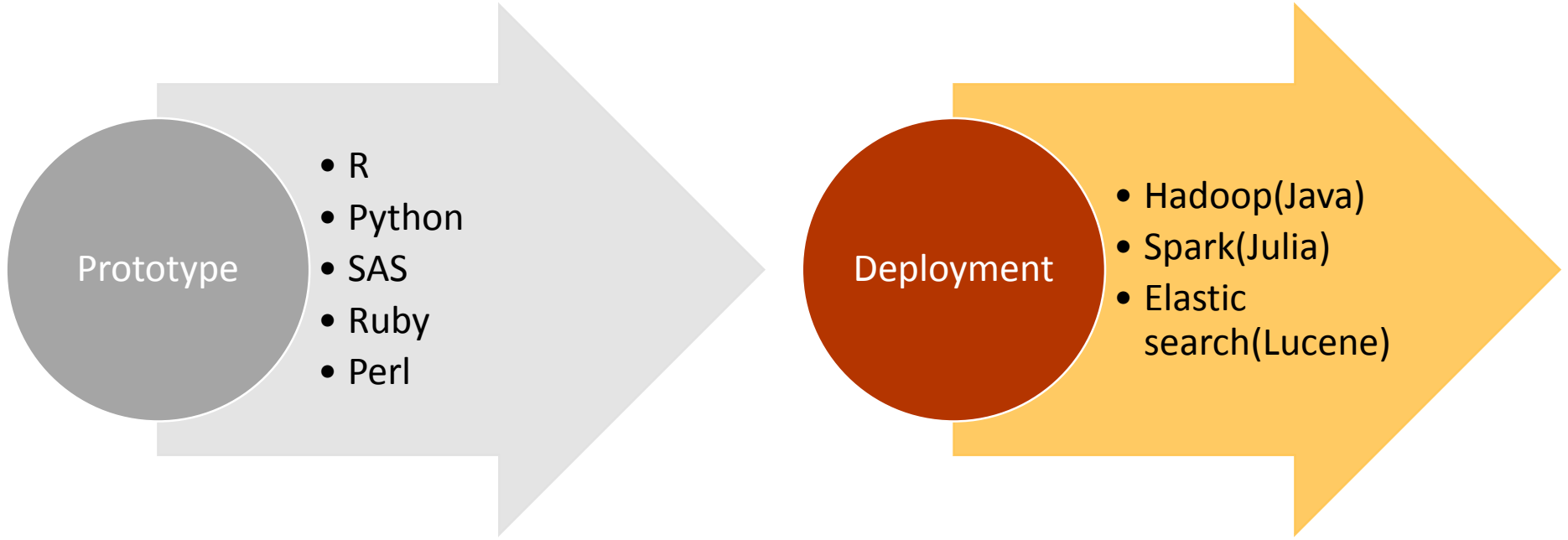
Copyrights© 2017, AcadGild. All rights reserved

Skill Sets of a Data Scientist



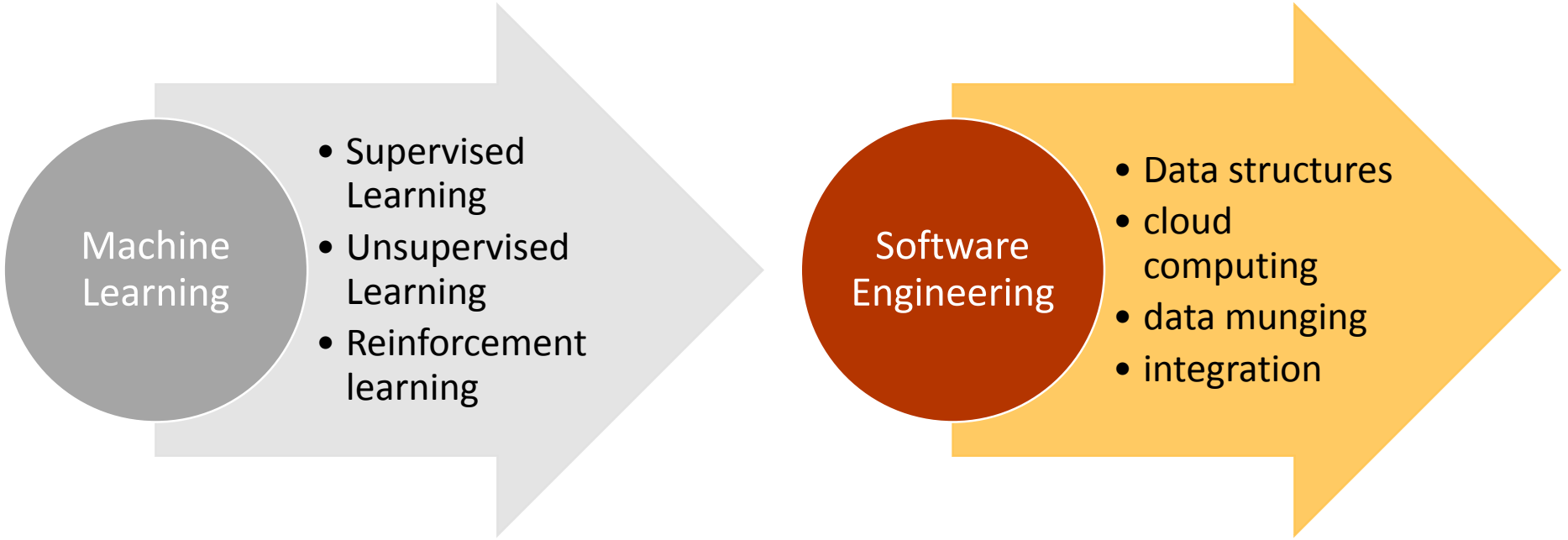
Copyrights© 2017, AcadGild. All rights reserved

Skill Sets of a Data Scientist



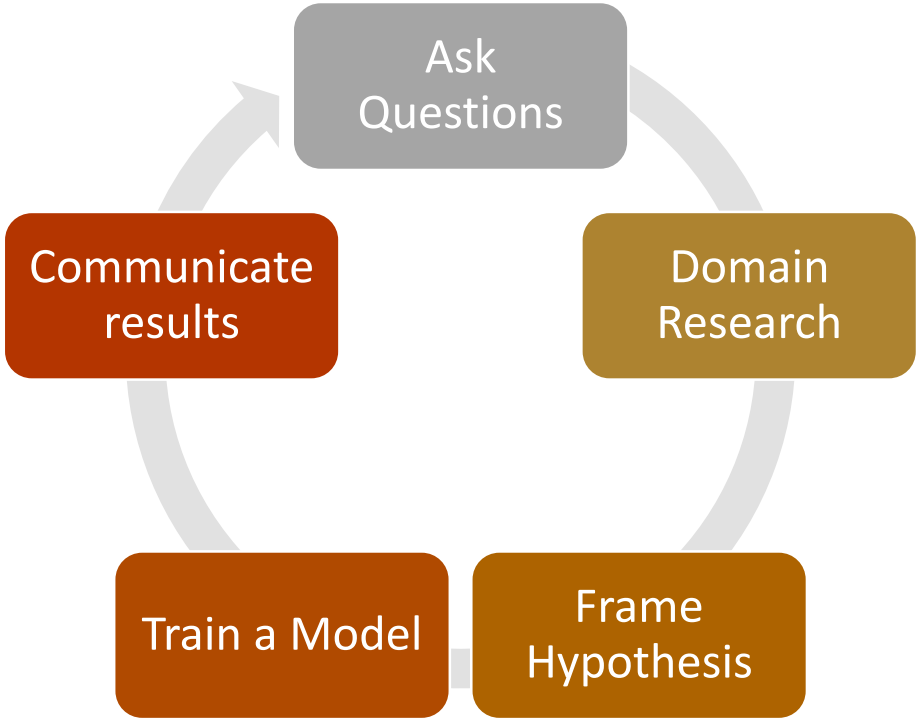
Copyrights© 2017, AcadGild. All rights reserved

Skill Sets of a Data Scientist



Copyrights© 2017, AcadGild. All rights reserved

Data Science Pipeline



Copyrights© 2017, AcadGild. All rights reserved

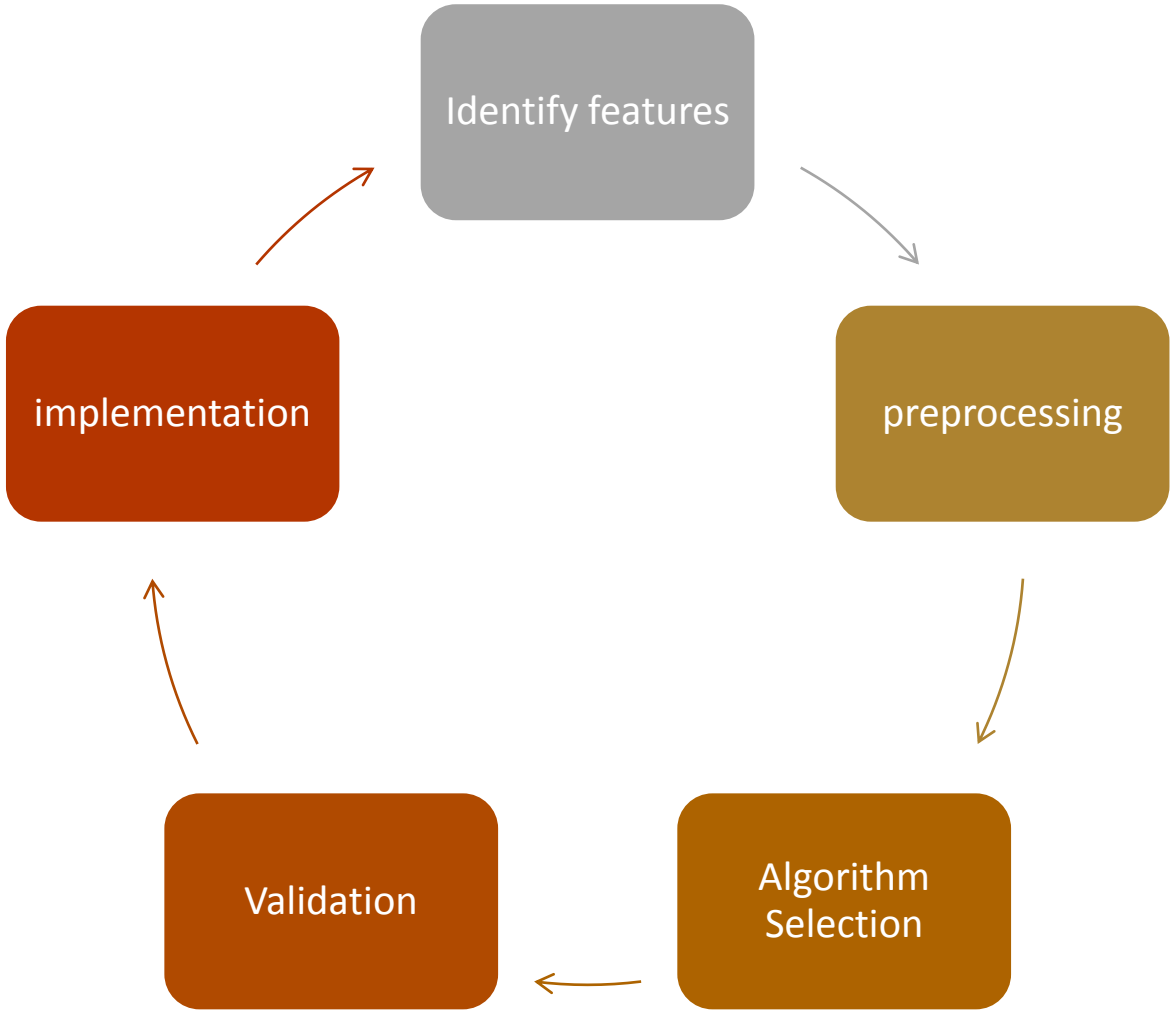
What is Machine Learning?

*A computer program is said to learn from experience **E** with respect to some class of tasks **T** and performance measure P, if its performance at tasks in **T**, as measured by **P**, improves with experience **E**.*

-Tom Mitchell

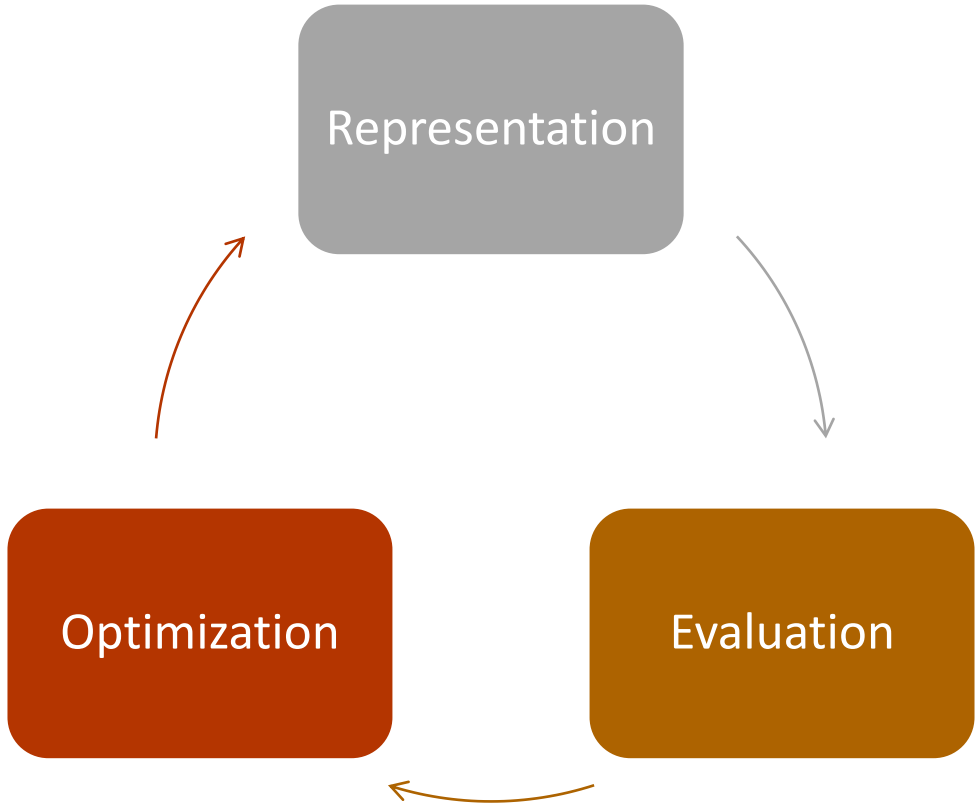
Copyrights© 2017, AcadGild. All rights reserved

Machine Learning Pipeline



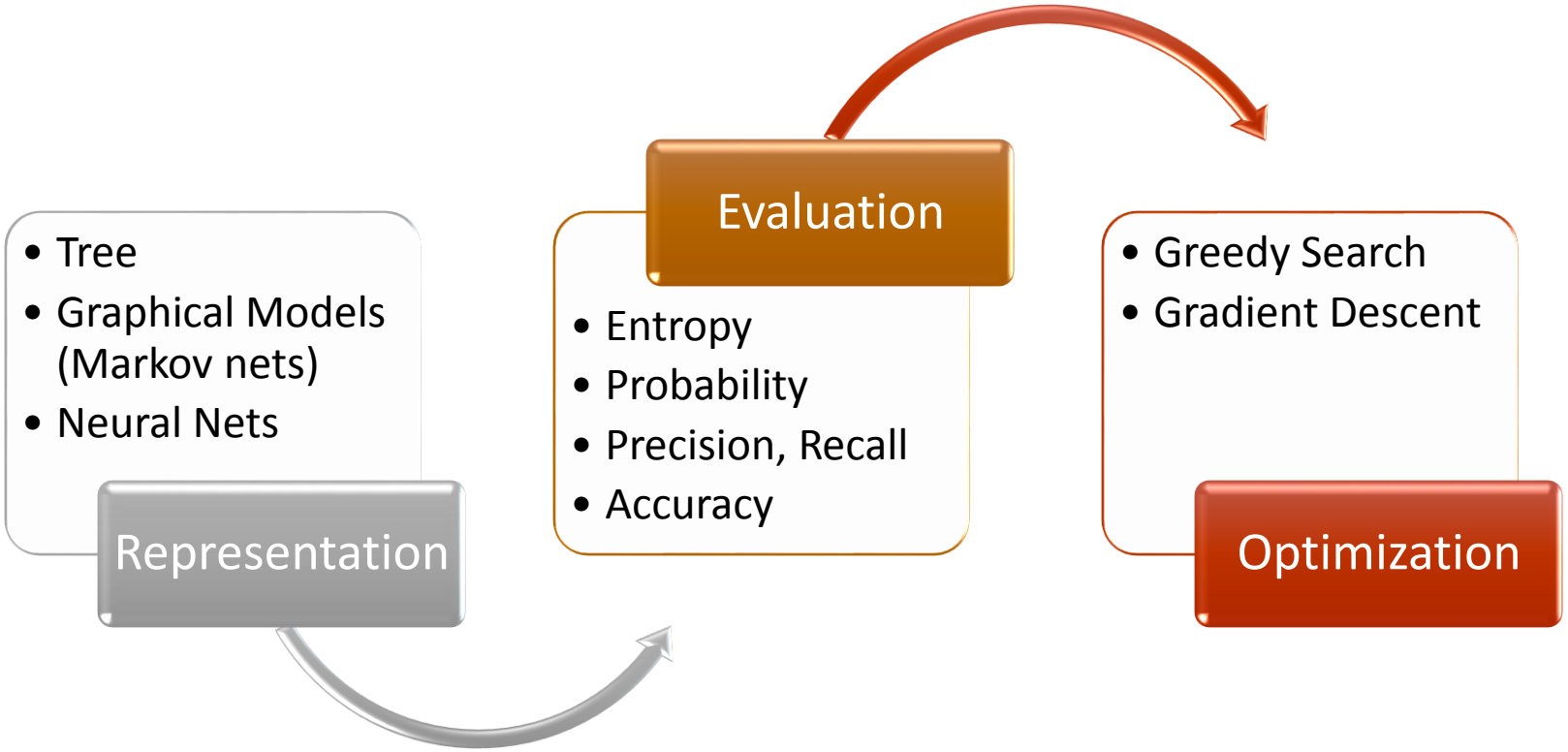
Copyrights© 2017, AcadGild. All rights reserved

What is an Algorithm?



Copyrights© 2017, AcadGild. All rights reserved

Component of Learning



Copyrights© 2017, AcadGild. All rights reserved

Why Data Science?

- Smarter computers to process complex algorithms in near real time (ex. GPUs)
- Minimal dependency on data analysts to fine tune the ML model
- Leverage a cluster of models than a single model (ensembles)
- Auto selection of ML models based on validation result

Copyrights© 2017, AcadGild. All rights reserved

How to become a Data Scientist?

Prerequisites to become a Data Scientist

- Basic Pre-requisites - Mathematics, Algorithms & Databases: Math -Linear Algebra, Analysis of Algorithms, Introduction to Databases
- Statistics - Probability and Statistics for Programmers, Statistical Formulas For Programmers, Data Analysis, Statistics One
- Programming - Google Developers R Programming Lectures, Introduction to R, Scientific Python Lectures, How to Think Like a Computer Scientist
- Distributed Computing and Databases
- Data Munging
- Filter and Mining data

Copyrights© 2017, AcadGild. All rights reserved

What is Next?

- Installing R and R Studio On Desktop
- R Studio Screen, Setting Your Working Directory
- Installing And Using Packages, Installing Packages And Libraries In R Studio
- Data Mining GUI In R, Graphics GUI In R
- Data Structures In R, Data Types - Vectors
- R Data Structures – Matrices
- R Data Structures - Arrays
- R Data Structures - Lists
- R Data Structures - Data Frames
- R Data Structures - Factors

Copyrights© 2017, AcadGild. All rights reserved

Email us – support@acadgild.com

Copyrights© 2017, AcadGild. All rights reserved