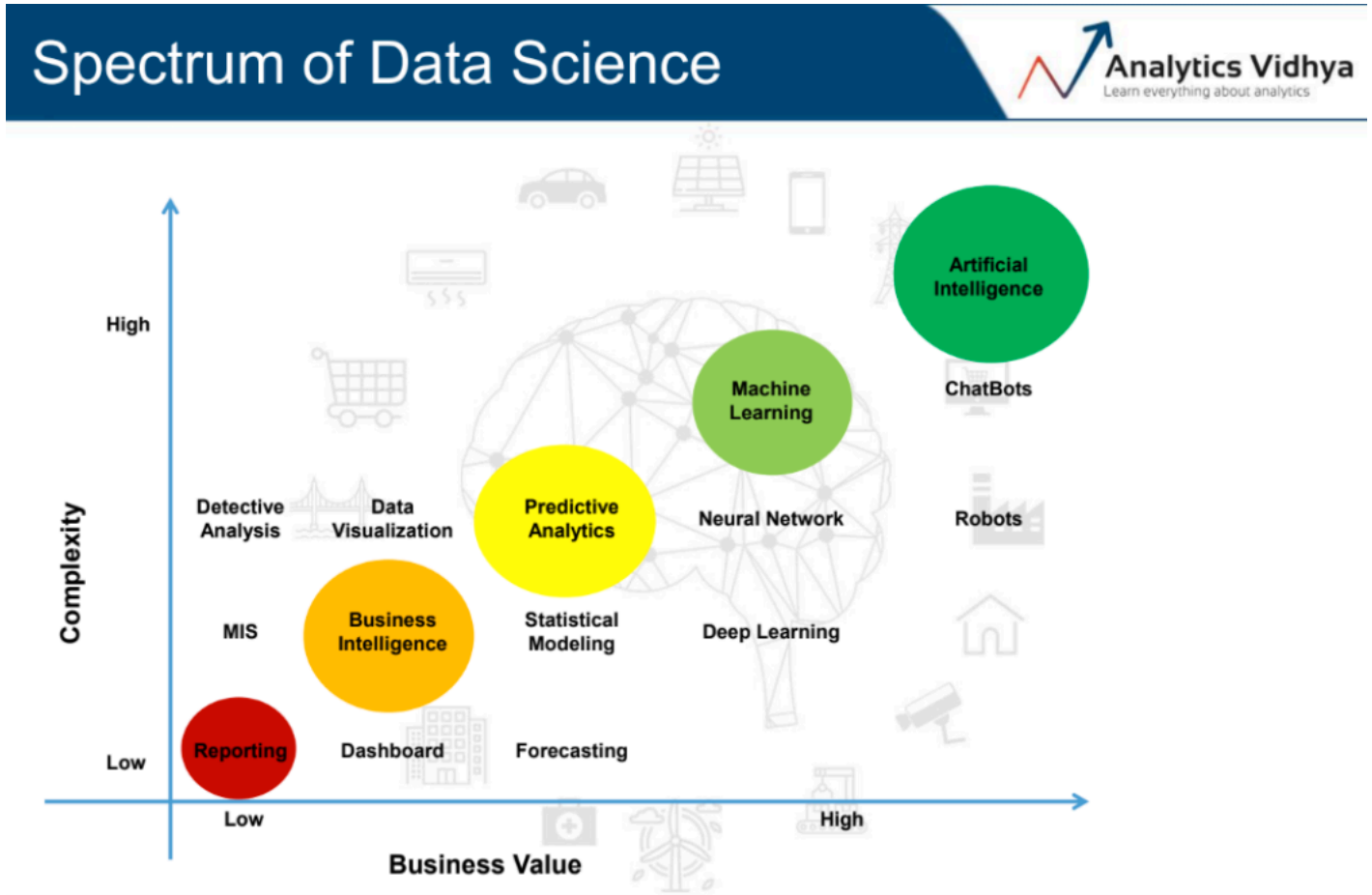
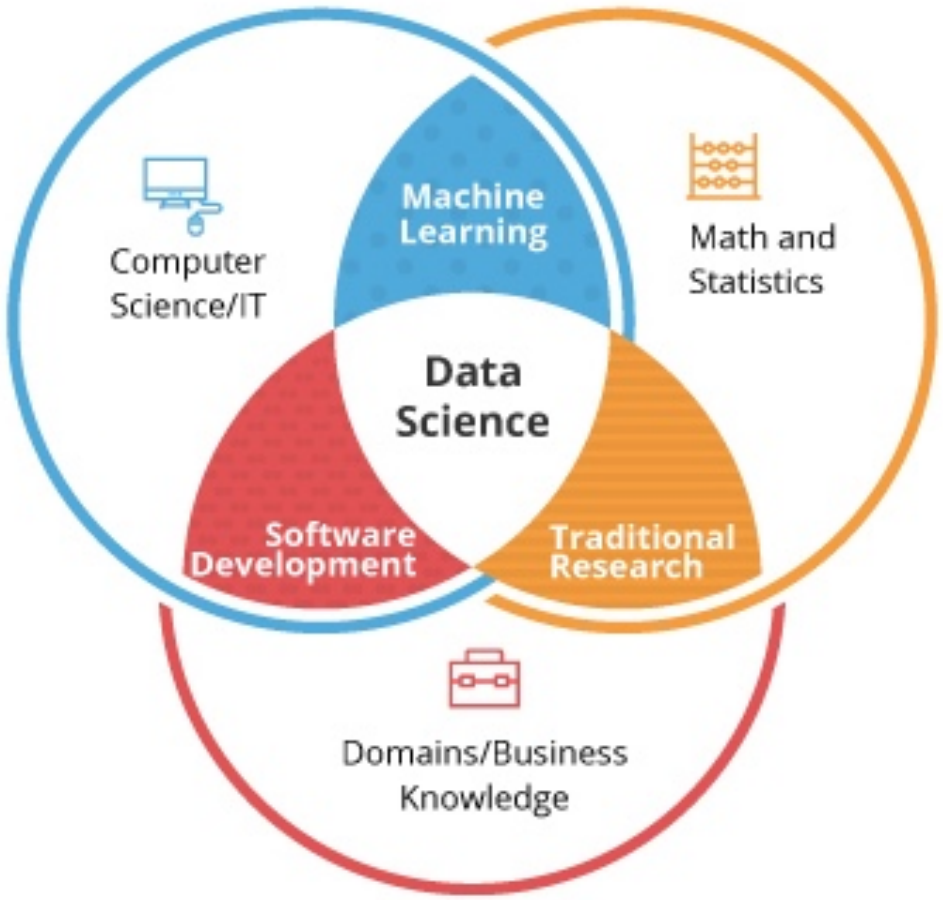
Master of Space Big Data Analytics
Course MSDZ06

Introduction to Common Tools for Data Science

Min Ding 丁忞
Tuesday, Jan. 19, 2021

What is Data Science?





- It is impossible to cover the full spectrum of data science.
- This course will be focused on statistics, inference, ML, and implementation of some widely-used methods.

Intended Learning Outcomes

This course is to introduce commonly used programming languages and data analysis methods. At the end of this course, you will be able to:

1. Code using Matlab, python or other programming language;
2. Implement most of the data analysis methods covered in the class;
3. Choose and apply a proper method to address a real-world problem based on real data.

Grading: 4 x 20% Assignments + 20% Final Project

Tentative Course Schedule

Module	Date	Contents	Assignment/Project	Speaker
1	1/19	Introduction to this Course Statistical Analysis and Hypothesis Tests		MD
2	1/26	Linear Regression Logistic Regression Relaxed Lasso Regression Principal Component Analysis		MD
3	2/2	Introduction to programming languages (Matlab, python, R)	Assignment 1 (20%)	MD
4	3/2	Frequency Domain Analysis		MD
5	3/9	Geostatistics/Kriging		MD
6	3/16	ArcGIS	Assignment 2 (20%)	F. ZHANG
7	3/23	Bayesian, Monte Carlo, Julia		C. GU
8	3/30	Perturbation analysis		L. YE
9	4/6	Inverse Theory/Optimization?	Assignment 3 (20%)	
10	4/13	Deep learning		G. WANG
11	4/20	Reinforced learning Transfer learning		G. WANG
12	4/27	Causal Inference		Y. LUO
13	5/4	Model-Data Fusion	Assignment 4 (20%)	
14	5/11	Student Presentations	Term Project (20%)	

Data Flair

The world is one big data problem.

- Andrew McAfee

Co-director of the MIT Initiative on the Digital Economy