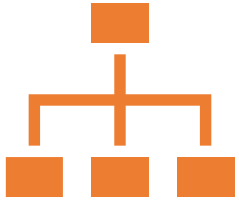


# Introduction to DS Course

# Course Objectives



Classification

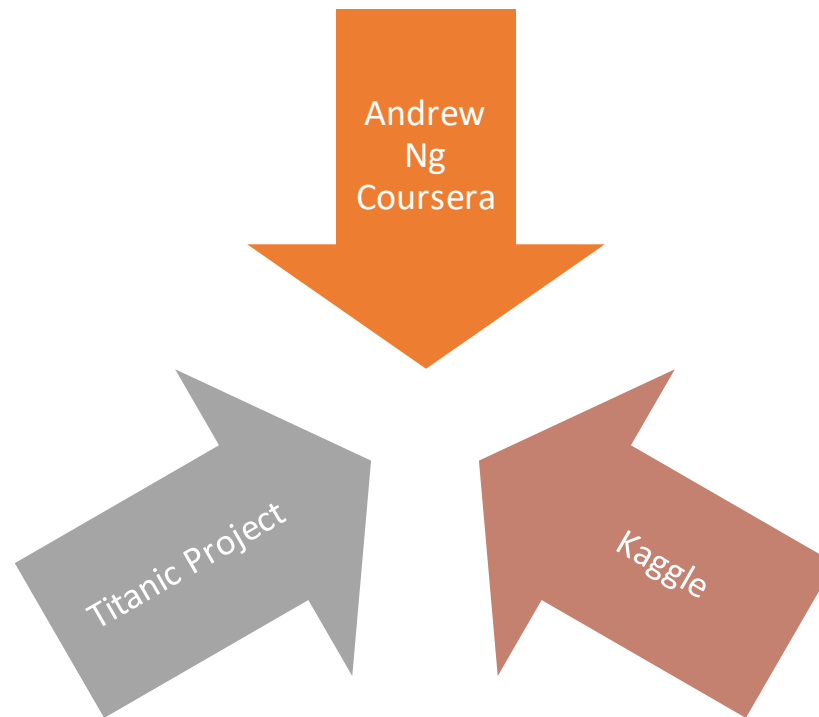


Cluster



Anomaly Detection

# Supporting materials



# Table of content

---

## introduction supervised & unsupervised

- introduction
- WorkFlow of Data Science Project
- Classification - Logistic Regression

## Titanic Project

- Problem definition and data pre-processing
- Data Science Project - Work Flow
  - Data Pre-processing
  - Feature Engineering
  - Feature Selection
  - First Run of Model (logistic regression)

# Table of content

---

## Random Forest

- Theory
- Implementation of Random Forest on Titanic Project
- Machine Learning System Design
- Model Evaluation

# Table of content

---

## Linear Regression

- Linear Regression with One Variable
- Linear Regression with Multiple Variables

## K-FOLD & Cross Validation

## Advice for Applying Machine Learning

## Dimensionality Reduction (PCA)

## \*Statistical Methods For Finding Outliers

# Table of content

---

Unsupervised Learning (Clustering)

Anomaly Detection

Support Vector Machines

Final Exercise - Home Credit