**Machine Learning Revision Week 1**

# What is Predicative Data Analytics?

Predicative Data Analytics is the combination of business and data processes with computational models to enable businesses to make data driven decisions. The three steps, in brief, are from data to insights to finally decisions. Examples of applications using this model are document classification, risk assessment, price predication etc.

# What is Machine Learning?

There are two types of machine learning

* Supervised
  + This type learns from the model relationship by the descriptive features and a target feature already implemented from historical examples.
  + Training Set 🡺 Machine Learning Algorithm 🡺 predication model
  + Query Instance 🡺 predication model 🡺 predication
  + Some examples that encompasses this type of learning are predicative model, consistent predicative models, feature design and feature selection
  + ML uses descriptive features to obtain target feature(s)
* Unsupervised
  + Will return to this after the notes have been completed

# How does machine learning work?

This works by looking through the predicative models for the best description of the relationship between the descriptive features and the target feature. One criteria necessary is the consistency between the data and the model. This can become an ill-posed problem as a result of training dataset cannot compute between different potential models. So, we try to reach the goal where the model can generalise beyond the dataset without the influence of noise using inductive bias. This has two forms, known as restrictive bias and preference bias to allow the machine learning to go beyond the capabilities of the dataset.

# What can go wrong with Machine Learning?

If we use the wrong inductive bias, we might find ourselves in the situation of underfitting or overfitting the model. We need to strike a balance between the model and dataset in machine learning.

# The predicative data analytics project lifecycle: Crisp-DM

