Name: The Dictators

Members: Dawid Lominski, William Read, Peter Kirby, Adrian Putra, Joshua Pitcher and Reece Robinson

Tutorial Session: 4

Component 1:

**Machine Learning**

Type of artificial intelligence that provides computers with the ability to learn without being explicitly programmed. Development on computer programs that can change when exposed to new data. Similarly to data mining, machine learning goes through the data provided and looks for patterns, but instead of extracting the data for human comprehension, it uses the data to find patterns and adapt program actions accordingly, without necessarily programming everything. (whatis)

Supervised machine learning

Majority of machine learning uses supervised learning. The goal is to approximate the mapping function so well that when you have new input data (x) that you can predict the output variables (y) for that data. (machinelearningmastery)

Called supervised learning because the process of an algorithm learning from the training dataset can be though of as a teach supervising the learning process. We know the correct answers, the algorithm iteratively makes predictions on the training data and is corrected by the teacher. Learning stops when the algorithm achieves an acceptable level of performance.

**Problems:**

* Classification – a classification problem is when the output is a category, such as “red” or “blue” or “disease” and “no disease”.
* Regression – A regression problem is when the output variable is a real value, such as “dollars” or “weight”

**Examples**:

Linear regression for regression problems  
Random forest for classification and regression problems,  
Support vector machines for classification problems.

Unsupervised machine learning

When you only have input data but no corresponding output variables. The aim for unsupervised learning is to model the underlying structure or distribution in the data in order to learn more about the data.

Called unsupervised learning because unlike supervised learning there is no correct answers and there is no teacher. Algorithms are left to their own devises to discover and present the interesting structure in the data.

**Problems:**

* Clustering – where you want to discover the inherent groupings in the data, such as grouping customers by purchasing behaviour
* Association – where you want to discover rules that describe large portions
* of your data, such as people that buy X also tend to buy Y.

**Examples:**

k-means for clustering problems,  
Apriori algorithm for association rule learning problems

Semi-supervised machine learning

Problems where you have a large amount of input data(x) and only some of the data is labelled (y) are called semi-supervise learning problems.

**Supervised –** All data is labelled and the algorithms learn to predict the output from the input data.

**Unsupervised -**  All data is unlabelled and the algorithms learn to inherent structure from the input data

**Semi-supervised** – Some data is labelled but most of it is unlabelled and a mixture of supervise and unsupervised techniques can be used.

**References:**

<http://whatis.techtarget.com/definition/machine-learning>

<http://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/>

**UML Diagram**



**USE CASE Diagram**

