# Lecture\_4a

Hello and welcome to the first lecture in this week. We're going to be discussing applying unsupervised lending. In this lecture video we are going to discuss when to use unsupervised lending, unsupervised learning techniques with a focus on clustering or cluster analysis, selecting a suitable algorithm, to implement unsupervised learning.

In machine learning unsupervised learning works by finding intrinsic structure or patterns in data that can be used to draw inferences for datasets that do not have labelled responses i.e. datasets have inputs only and no outputs. As a rule of thumb, when there is no existing data for how the dataset might be grouped, unsupervised learning is to be used. Unsupervised learning usually takes the form of clustering or cluster analysis. Conventionally, in cluster analysis data is partitioned into clusters of groups according to the measure of similarity or shared attributes between the observations and the data set, Clusters are usually formed in such a way that objects in the same cluster are very similar and objects in different clusters, very dissimilar. Clustering analysis in machine learning usually takes two forms. Hard clustering and soft clustering. In hard clustering each data point or observation in the data set belongs to only one cluster. In other words, data items are grouped all clustered a such a way that each data point is only assigned one cluster. Take, for instance, clustering a data set of tweets to determine whether a tweet is positive or negative.

A very popular hard clustering algorithm. is the k-Means In contrast to hard clustering, in soft clustering, each data point or observation in the dataset can belong to several clusters. In other words soft clustering is about grouping the data items or observations in the data set such that a data point can exist in multiple clusters. For example, clustering a dataset of the customers of a retail store by assigning a probability to each customer to belong to all the clusters of the retail store's customers. A very popular soft clustering algorithm is Fuzzy c-Means. For most unsupervised learning problems the choice of algorithm usually depends on the following cases, finding

possible structures or patterns in the data set. Finding the best number of groups or clusters in the dataset. In this video we've discussed when to use unsupervised learning, unsupervised techniques, with a focus on clustering, or clustering analysis And selecting a suitable algorithm to implement unsupervised learning.