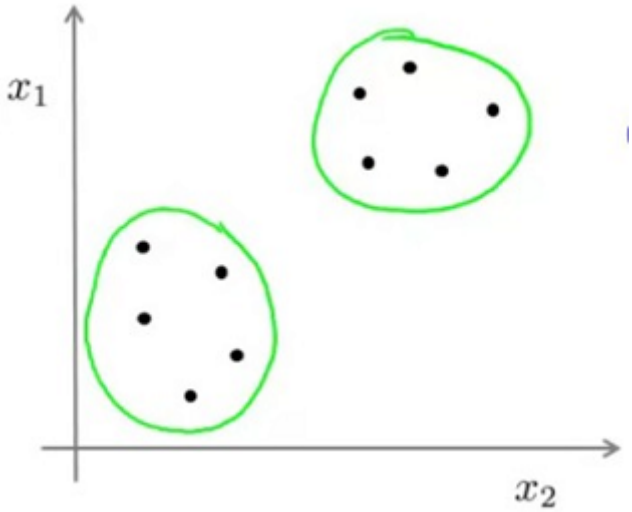
~ Unsupervised Learning - Clustering ~

Unsupervised learning is a machine learning technique, where we do not need to supervise the model. Unsupervised machine learning helps us to finds all kind of unknown patterns.

In a typical supervised learning problem where we're given a labeled training set, the goal is to find the decision boundary that separates the positive label examples and the negative label examples. In contrast, in the unsupervised learning problem we're given data that **does not have any labels associated with it**.

- **Supervised learning:**
  - Given a set of labels, fit a hypothesis to it
- **Unsupervised learning:**
  - Try and determining structure in the data
  - Clustering algorithm groups data together based on data features

Concretely, in the unsupervised learning we're given data that does not have any labels associated with it. The plot below is a set of points and no labels:



Training set: $\{x^{(1)}, x^{(2)}, x^{(3)}, \ldots, x^{(m)}\}$

What is clustering good for:

- **Market segmentation:** Group customers into different market segments
- **Social network analysis:** Facebook "smartlists"
- **Organizing computer clusters:** and data centers for network layout and location
- **Astronomical data analysis:** Understanding galaxy formation

Video Question: Which of the following statements are true? Check all that apply.

> In unsupervised learning, the training set is of the form $\{x^{(1)}, x^{(2)}, \ldots, x^{(m)}\}$.

> Clustering is an example of unsupervised learning.

> In unsupervised learning, you are given an unlabeled dataset and are asked to find "structure" in the data.

- Clustering is the only unsupervised learning algorithm.