

* Unsupervised Learning

→ Algorithms:

Unsupervised learning

Data only comes with inputs x , but not output labels y .
Algorithm has to find **structure** in the data.

Clustering

Group similar data points together.

Dimensionality reduction

Compress data using fewer numbers.





Anomaly detection

Find unusual data points.

→ Question:

Question

Of the following examples, which would you address using an **unsupervised** learning algorithm?

-  ☐ Given email labeled as spam/not spam, learn a spam filter.
-  ☒ Given a set of news articles found on the web, group them into sets of articles about the same story.
-  ☒ Given a database of customer data, automatically discover market segments and group customers into different market segments.
-  ☐ Given a dataset of patients diagnosed as either having diabetes or not, learn to classify new patients as having diabetes or not