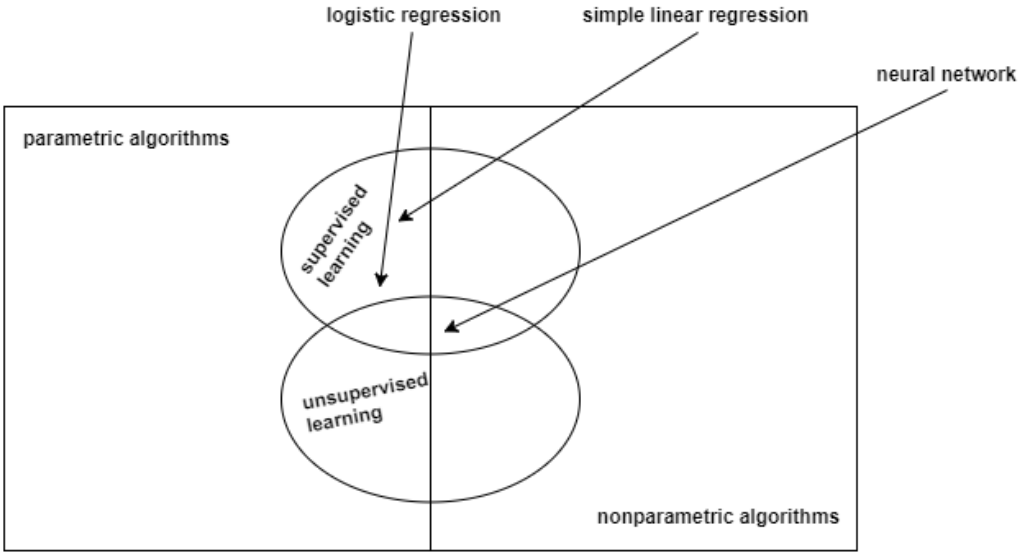
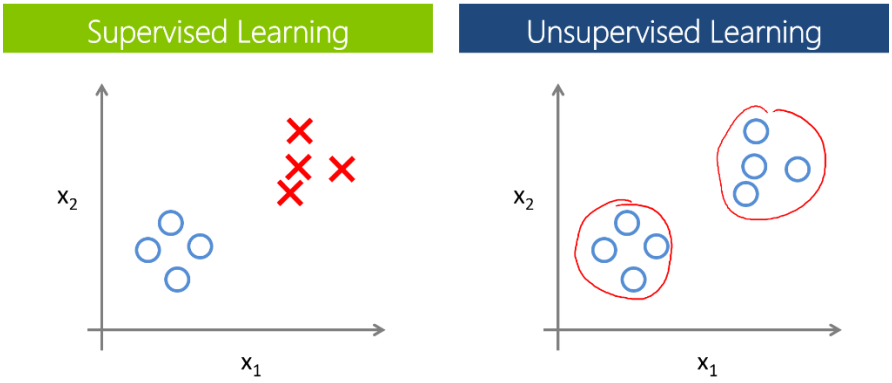
Machine Learning

- **What is machine learning?**
  - A field of study that gives computers the ability to learn without being explicitly being programmed
  - A computer is said to learn from experience E with respect to some tasks T and some performance measure P, if its performance on T, as measured by P, improves with experience E.
- **Machine learning algorithms:**
  - Parametric machine learning algorithms:
    - Algorithms that simplify the function to a known form are called parametric machine learning algorithms.
    - The algorithms involve two steps:
      1. Select a form for the function.
      2. Learn the coefficients for the function from the training data.
    - Examples:
      - Logistic Regression
      - Linear Discriminant Analysis
      - Perceptron
  - Nonparametric machine learning algorithms:
    - Algorithms that do not make strong assumptions about the form of the mapping function.
    - Examples:
      - Decision Trees like CART and C4.5
      - Naive Bayes
      - Support Vector Machines
      - Neural Networks
  - Supervised learning
  - Unsupervised learning
  - Others: reinforcement learning, recommender systems

Machine Learning



- **Example of machine learning:**
  - Playing checkers
    - E = the experience of playing many games of checkers
    - T = the task of playing checkers
    - P = the probability that the program will win the next game



- **Supervised learning**
  - Example: housing price prediction
  - In supervised learning, we are given a data set (training data) and already know what our correct output should look like, having the idea that there is a relationship between the input and output.
  - Supervised learning problems are categorized into:
    - Regression problem:
      - Predict results within a continuous output (maps input variables to some continuous function)
      - Example: housing price prediction

        - Classification problem:
            - Predict result in a discrete output (maps input variables into discrete categories)
            - Example: determine whether a tumor is benign our malignant

- **Unsupervised learning**
    - Idea: given a data set, can we find some structure within it?
    - An unsupervised learning algorithm might decide that the data lives in two different clusters
    - Clustering problem & unsupervised learning algorithm are used in many problems
    - Unsupervised learning is used o organize large computer clusters
    - Example:
        - Organize computer cluster
        - Social network analysis
        - Market segmentation
        - Astronomical data analysis
    - Unsupervised learning allows us to approach problems with little or no idea what our results should look like
    - We can derive structure from data where we don't necessarily know the effect of the variables
    - We can derive the structure by clustering the data based on relationships among the variables in the data
    - With unsupervised learning, there's no feedback based on the prediction results.