机器学习(Machine Learning)是人工智能（Artificial Intelligence）研究与应用的一个分支。

监督学习（Supervised Learning）：侧重对事物未知表现的预测，一般包括分类问题（Classification）和回归问题（Regression）。

分类问题：类别是离散的，也是预先知道数量的。例如根据人的体重，身高，三维来预测其性别。

回归问题：预测的目标是连续变量，例如根据房屋面积，地理位置等预测其价格。

无监督学习（Unsupervised Learning）：侧重对事物本身特性的分析，常用的技术有降维（Dimensionality Reduction）和聚类（Clustering）

数据降维：对事物的特性进行压缩和筛选，之前做的PCA主成分分析就是一种数据降维方法。

聚类：根据数据的相似性，将相似的数据划分为一个簇。不同于分类问题，大多数情况下我们并不预先知道簇的数量和具体含义。例如，大型点数对用户的信息和购买习惯进行聚类分析，从而针对特定人群投放特定广告。

特征：反映数据内在规律的信息。

特征向量：描述一个数据样本的规律信息。