

机器学习的主要术语

- 什么是机器学习：是通过学习如何组合输入信息来对未见过的数据进行预测。就是让机器从数据中学习，进而得到一个更加符合现实规律的模型，通过对模型的使用使得机器有更好的表现。
- 标签：标签是我们要预测的事物，可以使任何事物，如：未来房子的价格，图片显示中的是什么等等。
- 特征：特征是指输入变量，简单的机器学习可以只有一个特征，而复杂的机器学习可以多达数百万多的特征。如：垃圾邮箱中的所包含的发件人地址，敏感字段等等。
- 样本：样本是指数据的特定实例。
 - 有标签样本：包含特征和标签。
 - 无标签样本：只包含特征。

模型：模型定义了特征和标签间的关系。

训练表示创建和学习模型。

推断表示将训练后的样本运用于无标签样本。

回归与分类：

- 回归模型：可以预测连续值。如北京的一套房价格为多少？

等等。

- 分类模型：可以预测离散值。如这张照片是男人还是女人？

等等。