Andrew Ng ML学习总结

**概述**

断断续续，一个月的时间，把吴老师的机器学习视频教程看完，收获很多，从一无所知到概念的理解、公式的推导、算法的探究等等，可以说基本上算是快要入门了，接下来将是继续学习。当然，在学习的过程中需要不断的总结、实践和提升，所以，接下来我将对所学到的知识进行简单总结，以加强自我学习。

**机器学习**

机器学习是人工智能的核心，它在数据挖掘、计算机视觉、自然语言处理等领域有着广泛的应用。机器学习可以分为**监督学习**和**无监督学习**。

**监督学习**

监督学习是指：利用一组已知类别的样本调整分类器的参数，使其达到所要求性能的过程。例如，线性回归、逻辑回归、神经网络等算法都属于监督学习。

**线性回归**

线性回归根据其特征数，可分为一元线性回归和多元线性回归，这里以通用模型进行讲解。线性回归的**假设函数**模型如下：

x为**特征值**，θ为我们的参数也叫**权重**，其中x\_0=1。经过推导和演变，可以得到我们的代价函数（方差）：

我们的**梯度下降**公式如下：

(for j=0 and j=m)

其中𝛼为**学习率**，决定算法收敛的快慢，但取值过大可能造成无法收敛。通过迭代次数与J(θ)函数的可视化，可以查看梯度下降是否正确收敛。

根据多元微分法，我们可得到：

这样就可求得θ的值，进而得到我们的算法模型。

在梯度下降中，我们可以进行**特征缩放**，将我们的特征值缩放到一个相近的范围，如[-1,1]，这样我们的梯度下降会更快的收敛（特征缩放不能用于x\_0）。**均值归一化**是一个不错的特征缩放方式。其公式为： x\_i = (x\_i – 平均值) / (最大值 – 最小值)

除了梯度下降法求解θ的值，我们还可以用**正规方程法**直接求解θ。

**逻辑回归**

**正则化**

**神经网络**

**支持向量机**

**无监督学习**

**K-means**