总结报告4

（2019.9.18——2019.9.19）

**一、学习内容：机器学习WEEK1**

**1、机器学习（Machine Learning）的概念**

①机器学习就是不用对计算机明确编程，却能使计算机拥有学习能力的研究领域。（the field of study that gives computers the ability to learn without being explicitly programmed）

②TEP定义：A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E. 一个计算机程序可以从经验E中学习某类型任务T和性能评价手段P。它在任务T中的性能(用P度量)随着经验E的提高而提高。

T：Task任务

E：Experience经验

P：Performance measure性能度量 或称 对学习好坏的检验手段

**2、本门课要学习的内容**

①不同类型的学习算法（Different types of learning algorithms）：

√ Supervised Learning（监督学习）：teach

√ Unsuperivised learning（无监督学习）：learn by itself

②强化学习Reinforcement learning

③推荐系统Recommender systems

④应用学习算法的实用建议Practical advice for applying learning algorithm

**3、监督学习（Supervised Learning）**

（1）判别(拥有"right answer")

In supervised learning, we are given a data set and already know what our correct output should look like, having the idea that there is a relationship between the input and the output.

**KEY WORDS：**历史数据集（data set）、价值判断（correct output）、输入输出关系（relationship between the input and the output）

（2）两类问题

(a) 回归问题（**Regression**）：问题连续的

常见问题类型：预测类问题

(b) 分类问题（**Classification**）：问题是离散的（discrete valued output）

常见问题类型：是否患病、结果是两个或多个离散的答案

**4、无监督学习（Unsupervised learning）**

（1）判别(没有"right answer")

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 this structure by clustering the data based on relationships among the variables in the data.

**KEY WORDS：**无需知道结果应该如何（little or no idea what results should look like）

（2）两类问题

(a)聚类问题（Clustering）：找到一种方法自动分组，这些分组在某种意义上有一定的相似或相关性（somehow similar or related by different variables）。

(b)非聚类问题（Non-clustering）：鸡尾酒聚会算法（"Cocktail Party Algorithm"）：可以在混乱的环境中提取结构（to find structure in a chaotic environment）,例如在嘈杂声中分辨出人声和音乐(identifying individual voices and music from a mesh of sounds at a Cocktail Party).

（3）no right answer

在无监督学习中，没有基于预测结果的反馈（no feedback based on the prediction results）

**5、监督学习和无监督学习的区别方法**

①teach **OR** learn by itself

②with right answer **OR** without right answer

（我的理解：是否有明确的评价标准。比如：监督学习能明确地告诉计算机结果哪个是好的，哪个是不好的；而非监督学习，因为人可能也不知道什么是好，什么是坏，所以，计算机仅仅能将其聚类，即分成两类可能不知道代表什么含义）

**二、问题**

1、区别是监督学习，还是无监督学习还有其他的方法吗？

2、无监督学习也可以有历史数据，但是只是这些历史数据人们并没有评价标准，可以这样理解吗？

3、判别是否是无监督学习中，有一个短语是得到结构（derive this structure），这个结构（structure）应该怎么理解呢？仅指聚类吗？

4、课程提到这门课编程最好用Octave，要学习这门语言的编程吗？或是用Matlab？而且，现在我应该着手学习Python吗？