摘要

**本笔记是针对斯坦福大学2014年机器学习课程  
视频做的个人笔记**

黄海广

haiguang2000@qq.com

qq：10822884

最后修改：2017-10-12

**斯坦福大学2014机器学习教程个人笔记（V4.51）**

**斯坦福大学2014机器学习教程中文笔记目录**

第1周 1

一、 引言(Introduction) 1

1.1 欢迎 1

1.2 机器学习是什么？ 4

1.3 监督学习 6

1.4 无监督学习 10

二、单变量线性回归(Linear Regression with One Variable) 15

2.1 模型表示 15

2.2 代价函数 18

2.3 代价函数的直观理解I 20

2.4 代价函数的直观理解II 21

2.5 梯度下降 23

2.6 梯度下降的直观理解 26

2.7 梯度下降的线性回归 29

2.8 接下来的内容 31

三、线性代数回顾(Linear Algebra Review) 32

3.1 矩阵和向量 32

3.2 加法和标量乘法 34

3.3 矩阵向量乘法 35

3.4 矩阵乘法 36

3.5 矩阵乘法的性质 37

3.6 逆、转置 38

第2周 39

四、多变量线性回归(Linear Regression with Multiple Variables) 39

4.1 多维特征 39

4.2 多变量梯度下降 41

4.3 梯度下降法实践1-特征缩放 43

4.4 梯度下降法实践2-学习率 45

4.5 特征和多项式回归 46

4.6 正规方程 48

4.7 正规方程及不可逆性（可选） 51

五、Octave教程(Octave Tutorial) 53

5.1 基本操作 53

5.2 移动数据 60

5.3 计算数据 69

5.4 绘图数据 76

5.5 控制语句：for，while，if语句 82

5.6 向量化 88

5.7 工作和提交的编程练习 93

第3周 96

六、逻辑回归(Logistic Regression) 96

6.1 分类问题 96

6.2 假说表示 98

6.3 判定边界 100

6.4 代价函数 102

6.5 简化的成本函数和梯度下降 107

6.6 高级优化 111

6.7 多类别分类：一对多 115

七、正则化(Regularization) 118

7.1 过拟合的问题 118

7.2 代价函数 120

7.3 正则化线性回归 122

7.4 正则化的逻辑回归模型 124

第4周 126

第八、神经网络：表述(Neural Networks: Representation) 126

8.1 非线性假设 126

8.2 神经元和大脑 128

8.3 模型表示1 132

8.4 模型表示2 136

8.5 特征和直观理解1 138

8.6 样本和直观理解II 140

8.7 多类分类 142

第5周 143

九、神经网络的学习(Neural Networks: Learning) 143

9.1 代价函数 143

9.2 反向传播算法 145

9.3 反向传播算法的直观理解 148

9.4 实现注意：展开参数 151

9.5 梯度检验 152

9.6 随机初始化 154

9.7 综合起来 155

9.8 自主驾驶 156

第6周 159

十、应用机器学习的建议(Advice for Applying Machine Learning) 159

10.1 决定下一步做什么 159

10.2 评估一个假设 162

10.3 模型选择和交叉验证集 164

10.4 诊断偏差和方差 166

10.5 正则化和偏差/方差 168

10.6 学习曲线 170

10.7 决定下一步做什么 172

十一、机器学习系统的设计(Machine Learning System Design) 174

11.1 首先要做什么 174

11.2 误差分析 175

11.3 类偏斜的误差度量 178

11.4 查准率和查全率之间的权衡 179

11.5 机器学习的数据 181

第7周 185

十二、支持向量机(Support Vector Machines) 185

12.1 优化目标 185

12.2 大边界的直观理解 192

12.3 数学背后的大边界分类（可选） 198

12.4 核函数1 205

12.5 核函数2 207

12.6 使用支持向量机 209

第8周 212

十三、聚类(Clustering) 212

13.1 无监督学习：简介 212

13.2 K-均值算法 215

13.3 优化目标 217

13.4 随机初始化 218

13.5 选择聚类数 219

十四、降维(Dimensionality Reduction) 220

14.1 动机一：数据压缩 220

14.2 动机二：数据可视化 223

14.3 主成分分析问题 224

14.4 主成分分析算法 226

14.5 选择主成分的数量 227

14.6 重建的压缩表示 228

14.7 主成分分析法的应用建议 230

第9周 231

十五、异常检测(Anomaly Detection) 231

15.1 问题的动机 231

15.2 高斯分布 233

15.3 算法 234

15.4 开发和评价一个异常检测系统 236

15.5 异常检测与监督学习对比 237

15.6 选择特征 238

15.7 多元高斯分布（可选） 240

15.8 使用多元高斯分布进行异常检测（可选） 243

十六、推荐系统(Recommender Systems) 246

16.1 问题形式化 246

16.2 基于内容的推荐系统 248

16.3 协同过滤 250

16.4 协同过滤算法 251

16.5 向量化：低秩矩阵分解 252

16.6 推行工作上的细节：均值归一化 254

第10周 255

十七、大规模机器学习(Large Scale Machine Learning) 255

17.1 大型数据集的学习 255

17.2 随机梯度下降法 256

17.3 小批量梯度下降 257

17.4 随机梯度下降收敛 258

17.5 在线学习 260

17.6 映射化简和数据并行 262

十八、应用实例：图片文字识别(Application Example: Photo OCR) 263

18.1 问题描述和流程图 263

18.2 滑动窗口 264

18.3 获取大量数据和人工数据 266

18.4 上限分析：哪部分管道的接下去做 267

十九、总结(Conclusion) 268

19.1 总结和致谢 268

附件 270

1 常用的数学公式 270