



廈門大學  
XIAMEN UNIVERSITY

## 信息学院

### 机器学习 ——编程热身

姓 名： 黄静  
年 级： 2020 级  
学 号： 23020201153756  
完成日期： 2021 年 3 月 7 日

## 一、实验目的

根据机器学习的定义，设计一个机器学习的类（Class），并使用 txt 文件进行测试，要求能支持不同类型的模型，比如多项式模型要能支持不同阶的模型。

## 二、实验环境

本次实验采用 Visual Studio Code+numpy1.18.1 版本,使用 Python 作为实现语言，输入的数据为 ex1\_data.txt 文件。

## 三、实验过程

1.模型名为 polynomial\_model，模型包括以下属性最高阶数、损失函数、学习率、训练轮次、训练数据、训练结果（每个属性代表的默认值已在代码中标注）：

```
class polynomial_model:
    '''f = kx + b'''
    model_order = 1          # 默认最高阶数为1，即y = kx+b
    loss_function = "Square_loss" # 默认损失函数为平方损失函数
    learning_rate = 0.05      # 默认学习率为0.05
    epoch = 1000              # 训练轮次默认为1000轮
    x = []                    # 输入数据x
    y = []                    # 输入数据y
    poly_k = []               # 最终训练的结果，即每次幂的系数
    b = 0                     # 最终训练结果的偏置项
```

2.初始化模型方法 (polynomial\_model.init\_model(self, model\_order, loss\_function, learning\_rate, epoch))：

```
# 初始化模型
def init_model(self, model_order, loss_function, learning_rate, epoch):
    self.model_order = model_order
    self.loss_function = loss_function
    self.learning_rate = learning_rate
    self.epoch = epoch
    self.poly_k = np.zeros(shape=(1, model_order))
    print("Successfully init the model!")
```

3.读取数据方法 (polynomial\_model.data\_read(self, filepath))：

该方法是在指定路径上的文件中读取 txt 文件，并将训练数据输入训练模型中（即

polynomial\_model.x 和 polynomial\_model.y 中)。

```
# 读取数据
def data_read(self, filepath):
    f = open(filepath, 'r')
    lines = f.readlines()
    for line in lines:
        data = line.strip('\n').split('\t')
        self.x.append(float(data[0]))
        self.y.append(float(data[1]))
    print('Successfully loading data!')
    print('Length of data is:', len(self.x))
```

4.训练模型方法 (polynomial\_model.train(self)):

本次实验中此方法的实现过程省略，只打印出相关信息。

```
# 训练模型
def train(self):
    '''略'''
    print("Starting Training.....")
    print("Sucessfully Finish Training!")
```

5.类方法的测试，测试代码如下：

```
# 类的使用示例
# 测试模型初始化
a = polynomial_model()
a.init_model(2, 'Square_loss', 0.1, 1500)
print("The order of model is:", a.model_order)
print("The loss function of model is:", a.loss_function)
print("The learning rate of model is:", a.learning_rate)
print("The epochs of model is:", a.epoch)
print("The poly_k of model is:", a.poly_k, '\n')
# 测试模型数据读入
a.data_read('C:/Users/96342/Desktop/ex1_data.txt')
print("The value of x is:", a.x)
print("The value of y is:", a.y, '\n')
# 测试模型训练
a.train()
```

测试结果如下：

```
PS C:\Users\96342> & E:/Users/96342/Anaconda3/envs/gluon/python.exe c:/U
Successfully init the model!
The order of model is: 2
The loss function of model is: Square_loss
The learning rate of model is: 0.1
The epochs of model is: 1500
The poly_k of model is: [[0. 0.]]

Successfully loading data!
Length of data is: 3
The value of x is: [1.0, 3.0, 5.0]
The value of y is: [4.8, 11.3, 17.2]

Starting Training.....
Sucessfully Finish Training!
```

由测试结果可知，初始化模型的相关属性值参数已经成功输入模型。

#### 四、实验总结

通过本次试验，我初步设计了一个多项式机器学习模型类，并结合老师上课所讲，设计了相关类属性和方法，对机器学习模型的结构有了初步认识。