# 机器学习第一次实验作业报告

### 班级：计算机1603 姓名：汪傲雪 学号：1611640322

1. **题目**

使用 Python 实现对数几率回归模型

1. **问题描述**

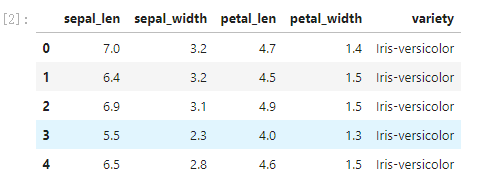
通过编程实现对数几率回归模型，并对 Iris 数据集进行分类以验证模型的效能

1. **数据集描述**

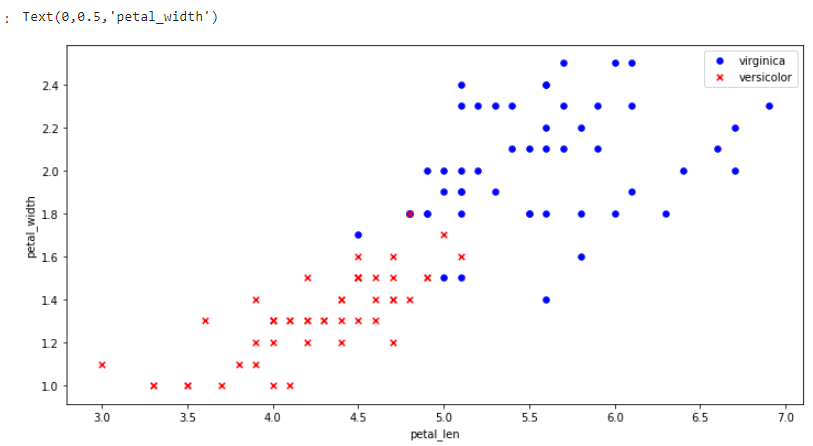
鸢尾花数据集分为三类：setosa、versicolor和virginica。并包含四种特性：花瓣的长度、宽度和花萼的长度和宽度。我们对鸢尾花的这些特性进行分类并预测它是属于花的哪个种类。

1. **实验结果图（包括随着训练次数的增加，模型损失函数 的变化曲线图）**

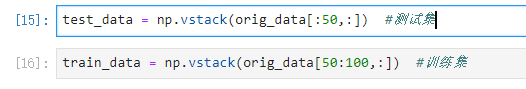
1.实验数据

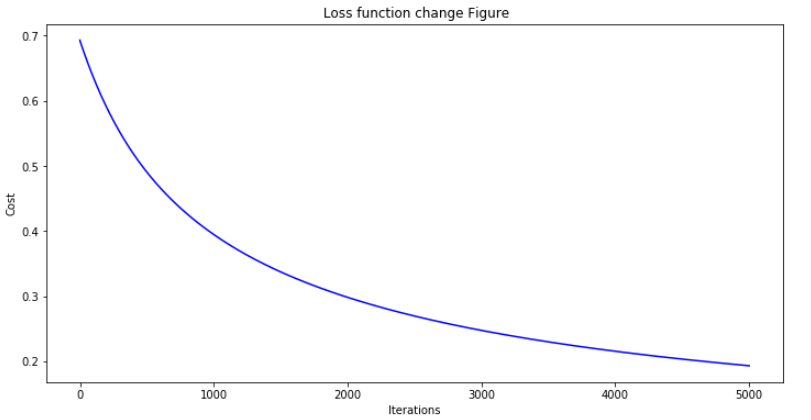


共100组数据（此处截取前5组），我选取了后两种花的种类，即versicolor和virginica。选取他们的花瓣长度和宽度进行研究，并做出了如下散点图



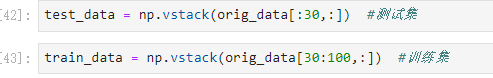
2.将数据集的 50%作为训练集，50%作为测试集，检验模型在测试集上的分类正确率

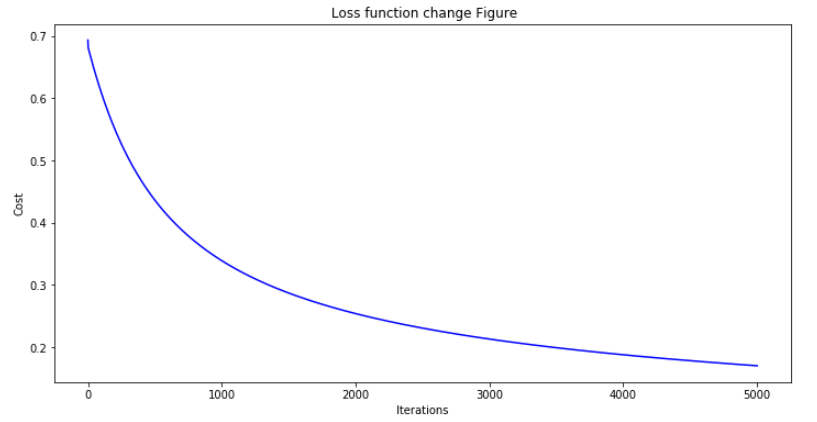






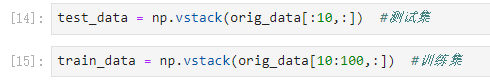
3.将数据集的 70%作为训练集，30%作为测试集，检验模型在测试集上的分类正确率

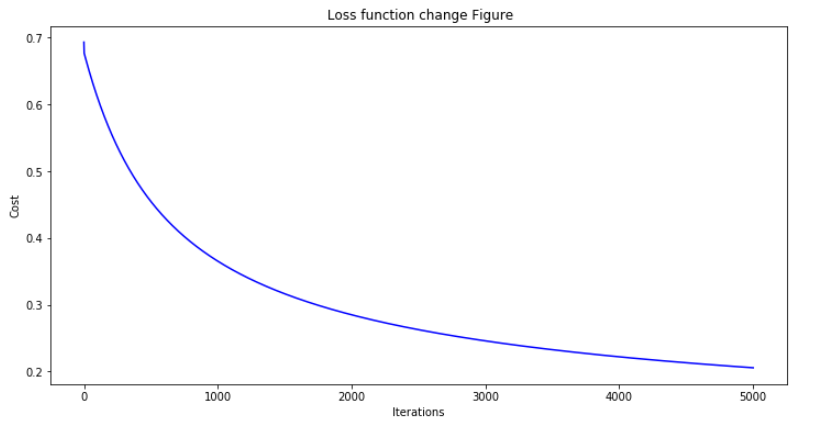






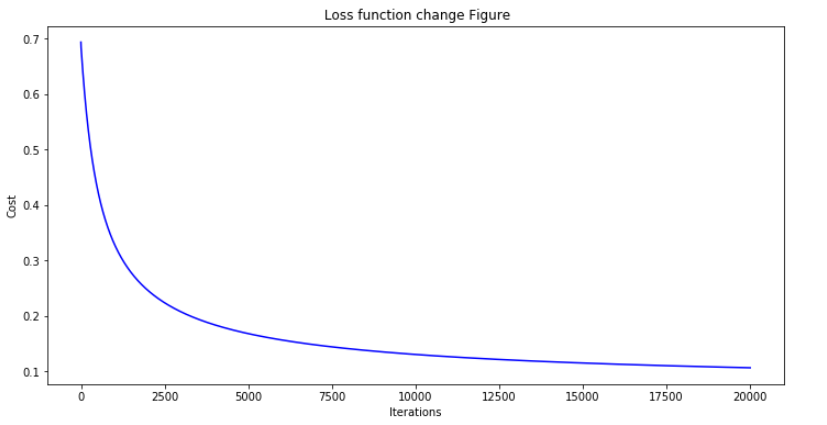
4将数据集的 90%作为训练集，10%作为测试集，检验模型在测试集上的分类正确率







5.以上三幅损失函数变化图都是迭代5000次显示出来的，下图是迭代20000次所显示出来的变化图



可见随着迭代次数的叠加，cost越来越小，趋于平稳。

1. **实验结果分析**
2. 这次实验是通过逻辑回归对数据模型进行分类求解的过程，也可以说是求解梯度下降的过程。首先要得到一个损失函数，通过梯度下降法进行多次迭代，迭代很多次，让损失函数接近最小值，这样就达到了训练模型的目的。
3. 由得到的损失函数变化图可以看到，损失函数逐渐接近一个最小值，当迭代次数越多，越接近这个点，到达平稳的状态，模型训练成功。
4. 我按照3种训练模式：50%训练、50%测试，70%训练、30%测试和90%训练、10%测试 所得到的结果 ，都是处于90%以上，甚至会有100%，说明这个结果是准确无误的，预测正确。