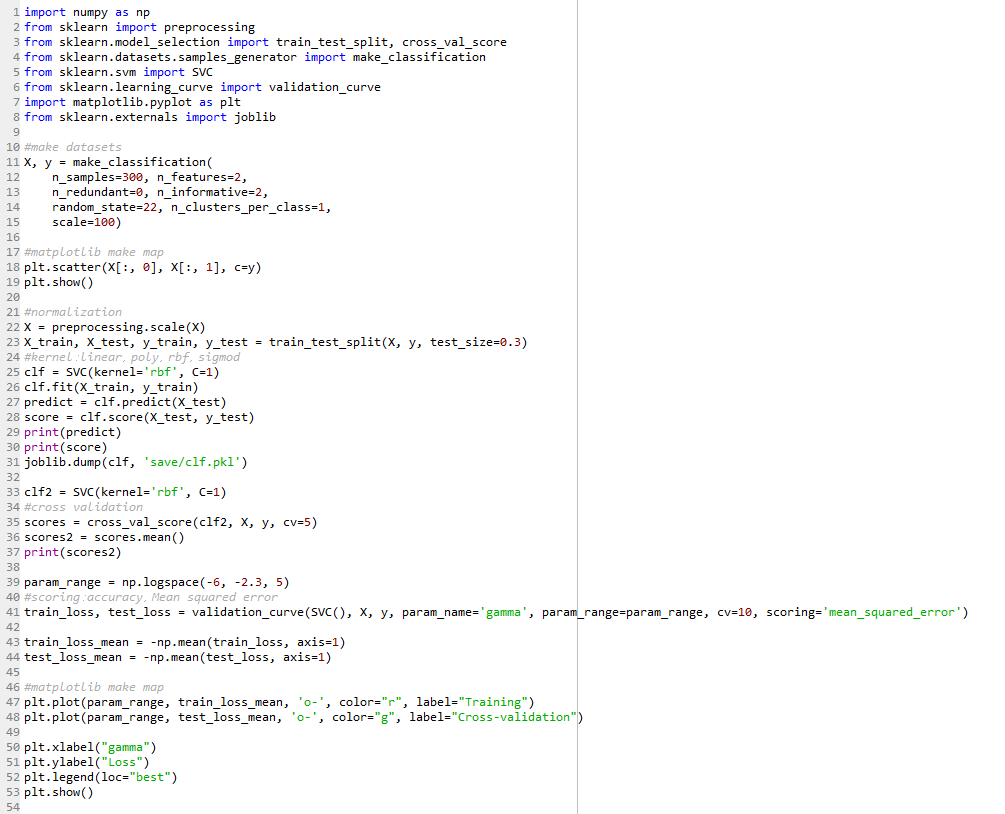
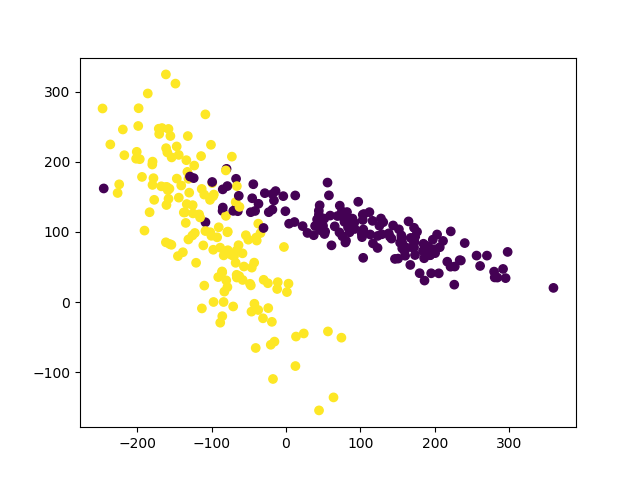
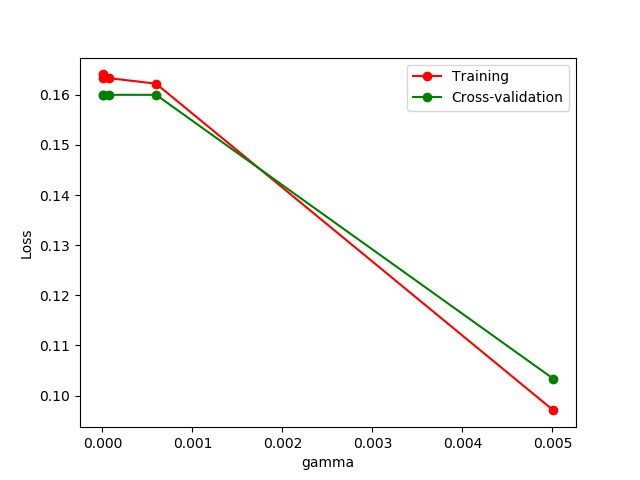
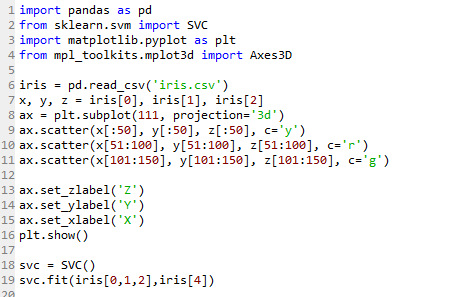
通过下图所示代码，我们希望能通过SVM即支撑向量机解决二分问题，这是一个极其经典的找到最大支撑的超平面从而预测分类问题的机器学习或数据挖掘算法。



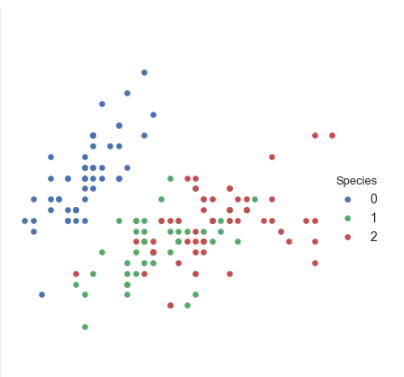
1. 我们这里引用sklearn中datasets.samples\_generator.make\_calssification包创造了一个数据集；
2. 利用matplotlib绘制数据集分布情况如下图所示；
3. 然后通过train\_test\_split对数据集进行训练集和测试集的分片；
4. 通过sklearn的SVM（支撑向量机）中的SVC（二分类模型）进行建模；
5. 对模型fit（喂）入数据，并打印predict和score，从而的到预测新的数据的分类情况和测试模型的精度；
6. 对于训练好的模型，我们选择了joblib.dump来保存知识库；
7. 下面我们选择了新的加入kernel function（核函数）的SVM模型，为了解决高维线性不可分的分类问题；
8. 然后matplotlib画出基于交叉验证方式的新模型损失函数的优势。

此为我们通过sklearn包中带有的数据集分析，然后我们又找到了两组经典的数据集，然后尝试着解决新环境问题。

首先我们找到了iris sets，此为经典的鸢尾花问题；其次我们又找到了Titanic生存问题的数据集。同理上述模型，然后我们将新的两组数据进行录入，就得到了我们对这两个新的实际问题的分析从而得到的预测模型。



下图为3中iris花的分类情况：





下图为Titanic中存活与死亡比：

