1、在梯度下降过程中，学习率控制着算法每一轮迭代中的更新步长，如果学习率设置的太大容易振荡，设置太小则\_\_\_\_收敛速度太慢\_\_\_\_。

2、均方误差有非常好的几何意义，它对应了常用的欧氏距离。基于均方误差最小化来进行模型求解的方法称为\_\_\_\_最小二乘法\_\_\_\_。

3、如果使用数据集的全部特征，学习模型在训练集上达到100%的准确率，但在测试集上仅能达到70%左右，这说明存在\_\_\_\_过拟合\_\_\_\_问题。

4、训练对数几率回归分类模型，如果在模型中引入正则项，正则化参数会对模型的性能有很大的影响，如果设置的过大则不能缓解过拟合问题，如果设置的过小\_\_\_\_易导致过拟合问题\_\_\_\_。

5、在预测任务中，给定样本集，其中是样本的真实标记。要评估学习器的性能，就要把学习器预测结果与真实标记进行比较。回归任务最常用的性能度量是均方误差，对应的公式为\_\_\_\_\_\_\_\_\_\_\_\_ *\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_*\_\_\_\_\_\_\_\_\_\_\_\_。

6、线性回归模型是用线性模型的预测值逼近样本的真实标记，对数几率回归模型是用线性回归模型的预测结果去逼近真实标记的\_\_\_\_\_对数几率\_\_\_\_\_\_。

7、Logistic回归（Logistic Regression，LR）是一种常用的处理二分类问题的线

性模型。Softmax 回归（Softmax Regression），也称为多项（Multinomial）或多类（Multi-Class）的Logistic回归，是Logistic回归在\_\_多类分类\_\_问题上的推广。

8、Logistic 回归采用\_\_\_\_对数损失函数\_\_作为损失函数，并使用\_\_梯度下降法\_\_\_\_来对参数进行优化。

9、写出对数几率回归模型中样本属于类别的概率公式。

10、基于一些基本策略，可以利用二分类学习器解决多分类问题。多分类学习的基本思路是“拆解法”，将多分类任务拆为若干个二分类任务求解。最经典的拆分策略有一对一、一对多和多对多。如果给定数据集中包含个样本，对应有个类别，请分析一对一和一对多策略的特点。

一对一策略：

给定数据集D，其中有N个分类，那么一对一的策略将这N个类别两两配对，从而产生N(N-1)/2个二分类任务。例如OvO将为区分类别Ci和Cj训练一个分类器，该分类器把D中的Ci类样例作为正例，Cj类样例作为反例。在测试阶段，新样本将同时提交给所有分类器，于是我们会得到N(N-1)/2个结果，最终结果可通过投票产生，即把被预测的最多的类别作为最终的分类结果。

一对多策略：

每次见一个类的样例作为正例，所有其他类的样例作为反例来训练N个分类器，在测试时若仅有一个分类器预测为正类，则对应的类别标记便是最终的分类结果，如果有多个分类器预测为正类，则同常考虑个分类器的预测置信度，选择置信度最大的类别标记作为分类结果。

11、给出线性模型的基本形式。（向量形式）