1.假如我们有如下要求：

* 在测试集上运行时间不高于10秒钟。
* 消耗内存小于10MB

应该选择哪个模型？ D

1. 测试集准确率97%，运行时间1秒，消耗内存3MB。
2. 测试集准确率99%，运行时间13秒，消耗内存9MB。
3. 测试集准确率97%，运行时间3秒，消耗内存2MB。
4. 测试集准确率98%，运行时间9秒，消耗内存9MB。

2.以下说法中正确的是：A

1. 准确率是一种优化指标（尽可能提升），运行时间和内存大小是一个满足要求即可的指标。
2. 准确率是一种需要满足要求指标，运行时间和内存大小是一个优化指标。
3. 准确率、运行时间和内存大小都是优化指标。
4. 准确率、运行时间和内存大小都是需要满足要求的指标。

3.以下划分中哪个最合适 A

1. 训练集9,500,000，验证集250,000，测试集250,000。
2. 训练集6,000,000，验证集100,000，测试集300,000。
3. 训练集3,333,334，验证集3,333,333，测试集3,333,333。
4. 训练集6,000,000，验证集300,000，测试集100,000。

4.（多选）假设已经有了100万张街边摄像头拍摄的鸟的照片作为数据集，若新增了100万张手机拍摄的鸟的照片，则这些手机拍摄的照片不能加入数据集当中，因为：AD

1. 街边摄像头和手机摄像头拍摄的鸟的图片中的信息可能会有所不同。
2. 100万张手机拍摄的鸟的照片没有明确的x到Y的映射（和纽约/底特律房价的例子类似）。
3. 数据集太大会使模型训练和测试变慢。
4. 街边摄像头和手机摄像头拍摄的鸟的数据的分布存在差异。

5.一些人工标注数据的错误率如下：

* 专家1标注的数据错误率为0.3%。
* 专家2标注的数据错误率为0.5%。
* 非专家1标注的数据错误率为1.0%。
* 非专家2标注的数据错误率为1.2%。

加入我们的目标为使模型达到“人类水平”，那“人类水平”应该是多少？B

1. 0.0%。
2. 0.3% (专家1的准确率)。
3. 0.4% (专家1和专家2的准确率的均值)。
4. 0.75% (上述数字的均值)。

6.以下说法中正确的是：A

1. 一个学习算法的性能可以比人类水平更好，但它永远不会比贝叶斯误差更好。
2. 一个学习算法的性能永远不会比人类水平更好，但它可以比贝叶斯误差更好。
3. 一个学习算法的性能不可能比人类水平更好，也不可能比贝叶斯误差更好。
4. 一个学习算法的性能可以优于人类水平的性能，也可以优于贝叶斯误差。

7.（多选）一组专家标注数据达到了0.1%的错误率，若将0.1%作为人类水平，得到的训练集和开发集的错误率分别为2.0%和2.1%，以下说法中正确的是：BD

1. 增大训练集减少方差。
2. 添加正则化操作。
3. 减少正则化操作。
4. 训练一个更大的模型，以在训练集上表现更好。

8.假设人类水平的错误率为0.1%，训练集、开发集、测试集上的错误率分别为2.0%、2.1%、7.0%，则下列说法中正确的是：A

1. 对开发集欠拟合。
2. 测试集应该增大。
3. 对开发集过拟合。
4. 开发集应该增大。

9.假设人类水平的错误率为0.1%，训练集、开发集上的错误率均为0.05%，则下列说法中正确的是：CD

1. 数据错误，因为模型对于数据的的识别能力不可能超过人类水平。
2. 由于0.05%的错误率已经非常接近于0，则将模型错误率降低到0也是比较容易的。
3. 如果测试集足够大，则0.05%的误差估计是准确的，这意味着贝叶斯误差不大于0.05%。
4. 将错误率从0.05%降低到的难度会非常大。

10.使用原有数据集（100万条数据）训练的模型已经部署，但是有一种原有数据集未出现过的新的数据被输入到模型当中，假设新数据数量只有1000，客户要求迅速修改原有模型以适应新的数据，则以下任务应该先做哪一个？D

1. 将新数据加入测试集或者开发集，然后在原有模型基础上调参。
2. 将新数据加入训练集重新训练模型。
3. 尝试数据增强/数据合成来获得更多和新数据比较类似的假数据。
4. 将新数据添加到数据集中，并重新分为新的训练/开发/测试集，然后重新训练模型。