1.什么是判别模型与生成模型？两者的优缺点？神经网络是判别模型还是生成模型？

（1）

判别模型不会考虑样本数据产生的模型是什么样的，而对于给定输入直接给出预测输出；生成模型关注的是样本数据是怎么生成的。

从公式角度看：

判别模型直接学习得到 $P(Y|X)$，利用最大后验概率得到对应的类别标签；生成模型学习的是先得到联合概率分布 $P(X,Y)$，然后再得到 $P(Y|X)$，预测时应用最大后验概率得到预测类别标签。

（2）

两者的优缺点：

判别方法利用了训练数据的类别标识信息，直接学习的是条件概率 $P(Y|X)$，直接面对预测，往往学习的准确率更高；

由于直接学习条件概率 $P(Y|X)$，可以对数据进行各种程度上的抽象、定义特征并使用特征，因此可以简化学习问题但是不能反映训练数据本身的特性。

生成方法能够反映同类数据本身的相似度；

学习收敛速度更快。当样本容量增加的时候，学到的模型可以更快地收敛于真实模型；

当存在隐变量时，仍可以用生成方法学习，此时判别方法不能用。（这一点不是特别明白，请帮忙解释一下吧）

（3）

神经网络属于判别模型。

2.多层感知器是什么，它的基本模型是什么，解决什么问题？工作流程？

（1）

多层感知器由简单的相互连接的神经元或节点组成，其模型结果示意图如图 1 所示：

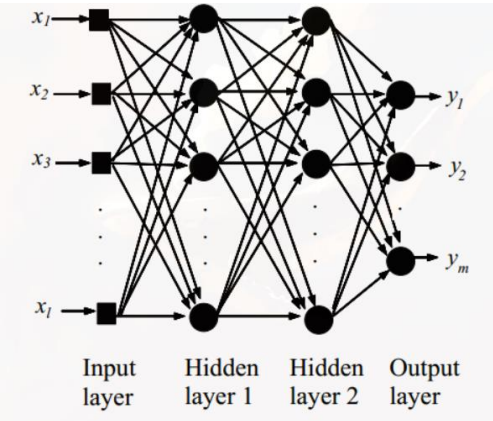

图 1 多层感知机模型示意图

（2）
从图 1 可以看出多层感知机包含输入层、隐层和输出层，不同层之间是全连接的。
（3）
单层感知机的局限性在于无法实现分离非线性空间，而通过叠加多个感知机可以近似非线性函数。
（4）
以图 1 所示的多层感知机可以看出，这个感知机将 $x_1 \cdots x_l$ 作为神经元的输入，将其和各自的权重相乘之后传送至下一个神经元，在下一个神经元中，计算这些加权信号的总和。

和神经网络的区别在于感知机的权重是由人工选定的，而神经网路可以自动地从数据中学到合适的权重参数。

3.训练集、验证集、测试集如何划分，划分的意义是什么？
(1)
常使用均匀随机抽样的方式，将数据集划分为训练集、验证集、测试集，这三个集合不能有交集，常见的比例是 8:1:1。
(2)
训练集：用来训练模型的，更新权重参数；
验证集：来验证训练的模型的好坏。进行超参数的更新，比如降低学习速率，增加迭代次数等。
测试集：用来测试模型的好坏的。

4.提出交叉验证的初衷是什么？有哪几种方式？
(1)
因为在实际情况中数据不充足，为了更好地模型选择提出交叉验证。把给定的数据进行切分，将切分的数据集组合为训练集和测试集，然后反复地进行模型的训练和测试。
(2)
交叉验证的方式：
①简单交叉验证
随机地将数据划分为训练集和测试集，然后多次调整参数后得到不同的模型，并用测试集测试这些模型的性能，选出测试误差最小的模型；
②S 折交叉验证
因为 S 折交叉验证将原始数据切分为 S 个互不相交的大小相同的子集，将每个子集都做一次验证集，其余的 S-1 个子集作为训练集。
③留一交叉验证
留一交叉验证属于 S 折交叉验证的特殊情况，当 S 为样本个数时即为留一交叉验证。一般是样本个数比较少时选择留一交叉验证。