HMM，CRF模型怎么训练的，分词标注等具体怎么进行的

由softmax怎么推导得到逻辑回归

Svm的hinge函数

Word2vec中的hierarchy softmax怎么求梯度的

随机梯度，牛顿法和拟牛顿法

RNN的层次上的梯度爆炸和时间域上的梯度爆炸和消失怎么解决的

CNN中的pooling有哪两种padding方法

RNN还有除了SDG还有哪些方法