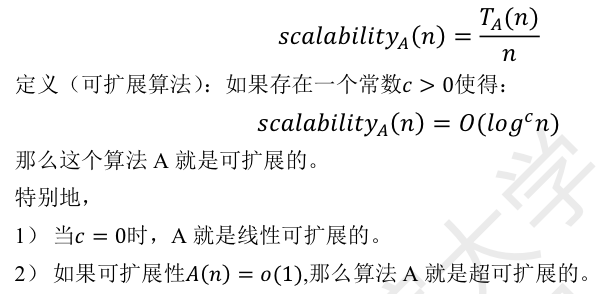
三．

1.



2.

同步更新的原理如下： 批量同步更新设置同步点（通常是每一轮迭代至少设置一个同步点） i） 在同步点等候所有节点完成任务，然后统一更新参数

ii） 更新后进入下一轮迭代 在最大程度上模拟了序列计算

1. 与序列计算结果基本一致
2. 与序列计算具有同样的理论保障(正确性、收敛性等)

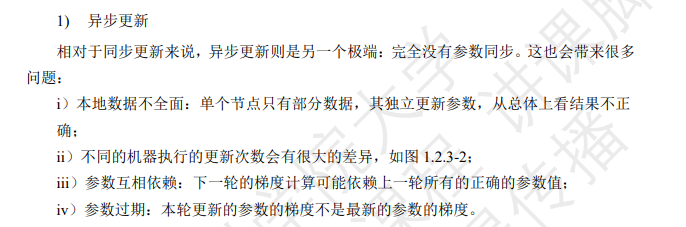
同步更新主要的缺点就是：同步点的设置会导致 BSP 低效：

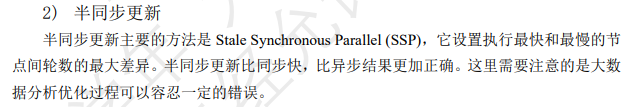
1） 网络速度慢，延时大；

2） 机器计算速度不一致（导致等待）

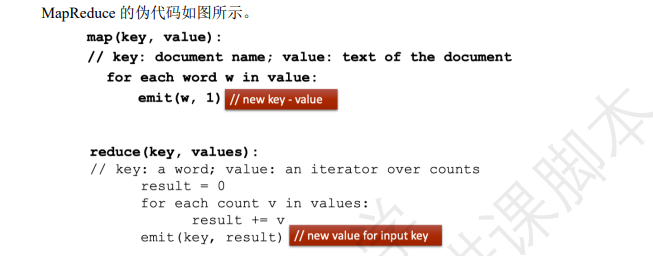
3） 与机器即时状态有关

4） 与需要处理的数据有关

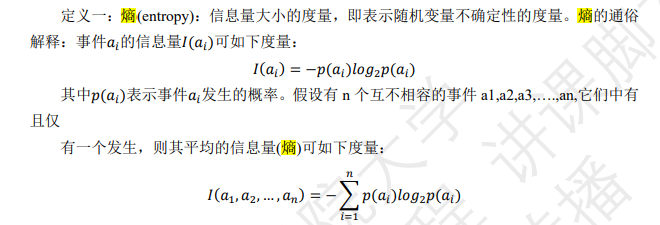


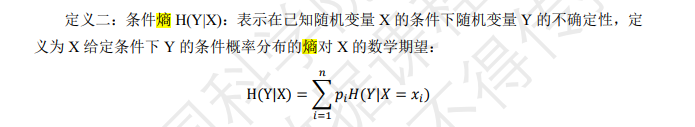


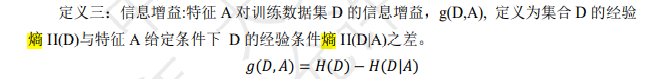
3.



4.

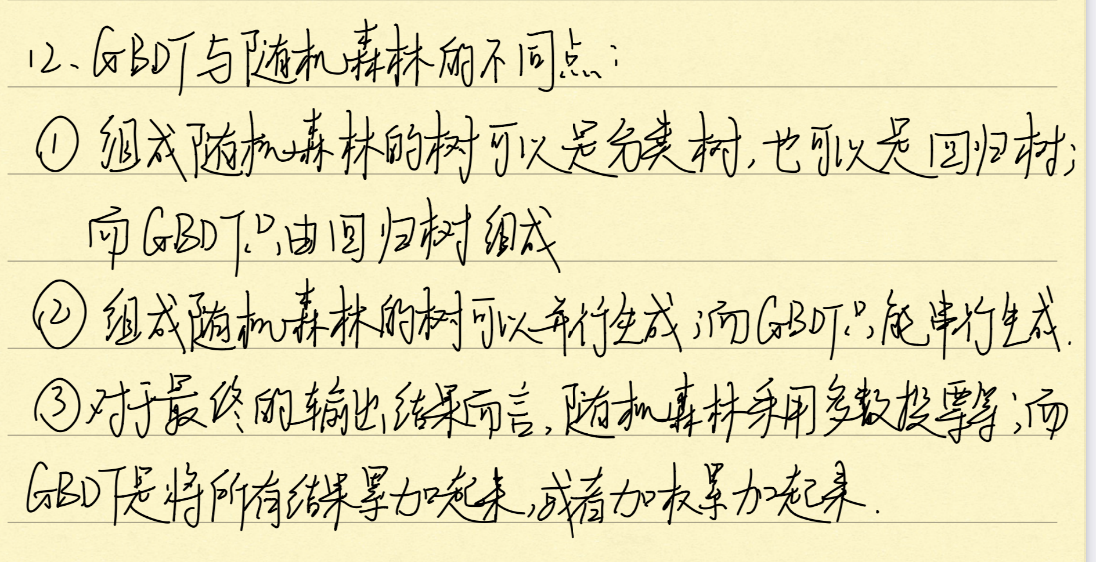




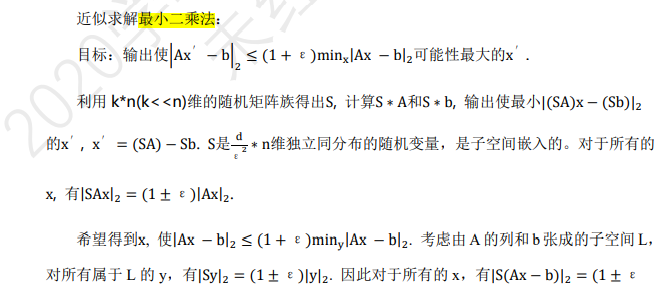
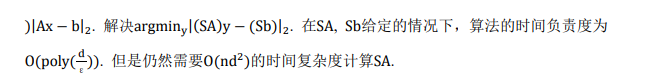


5.

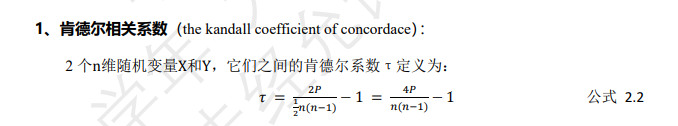
都是由多棵树组成，都属于集成学习算法，两种算法最终的结果都是由多棵树一起决定，RF算法与GBDT算法在训练过程中会生成多棵决策树，并整合这些决策树的输出提升整体模型的分类或回归的效果。

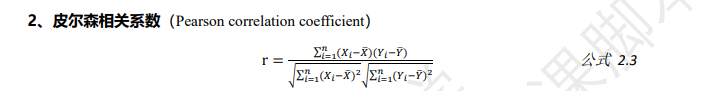


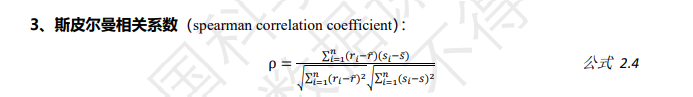
6.

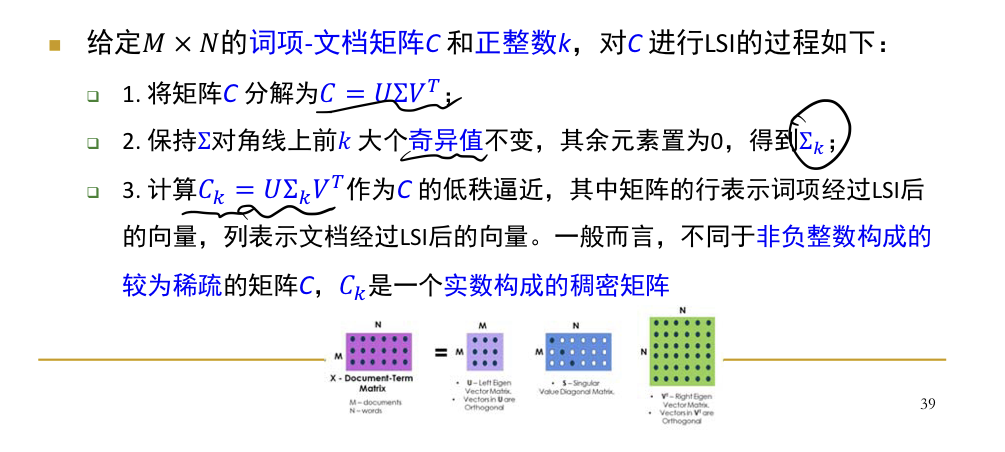
7.

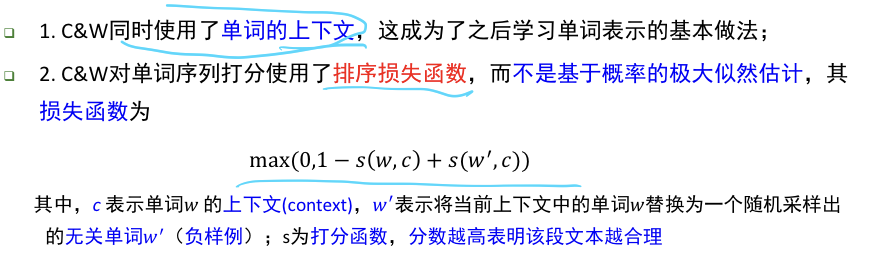




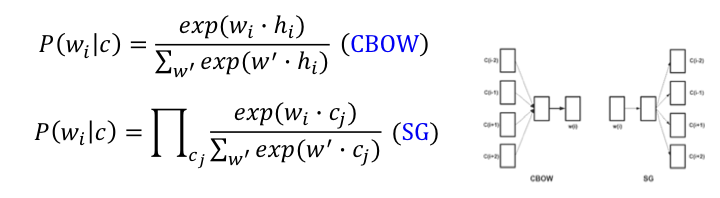


8.

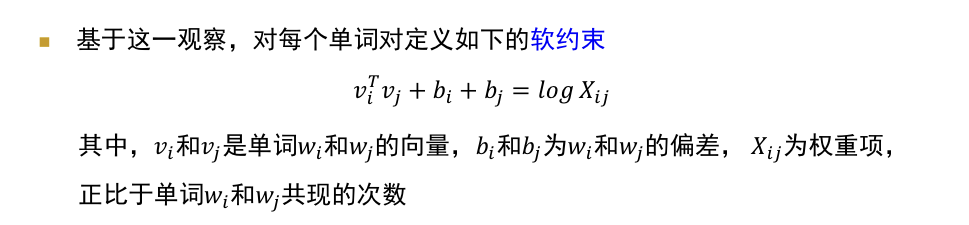


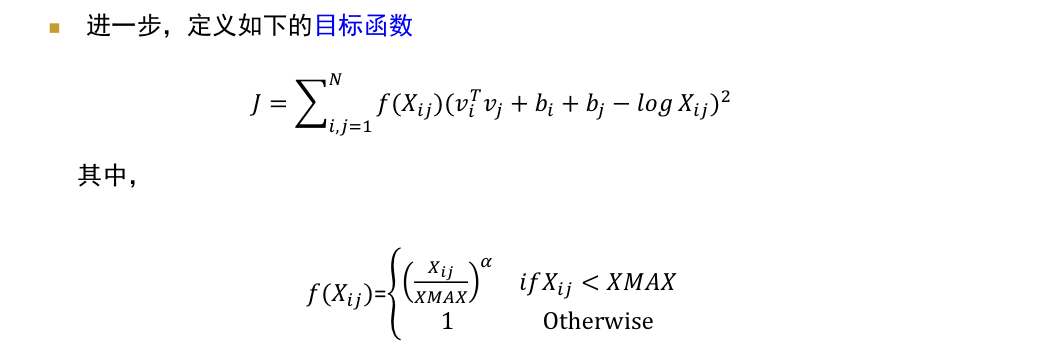
9. 

10.

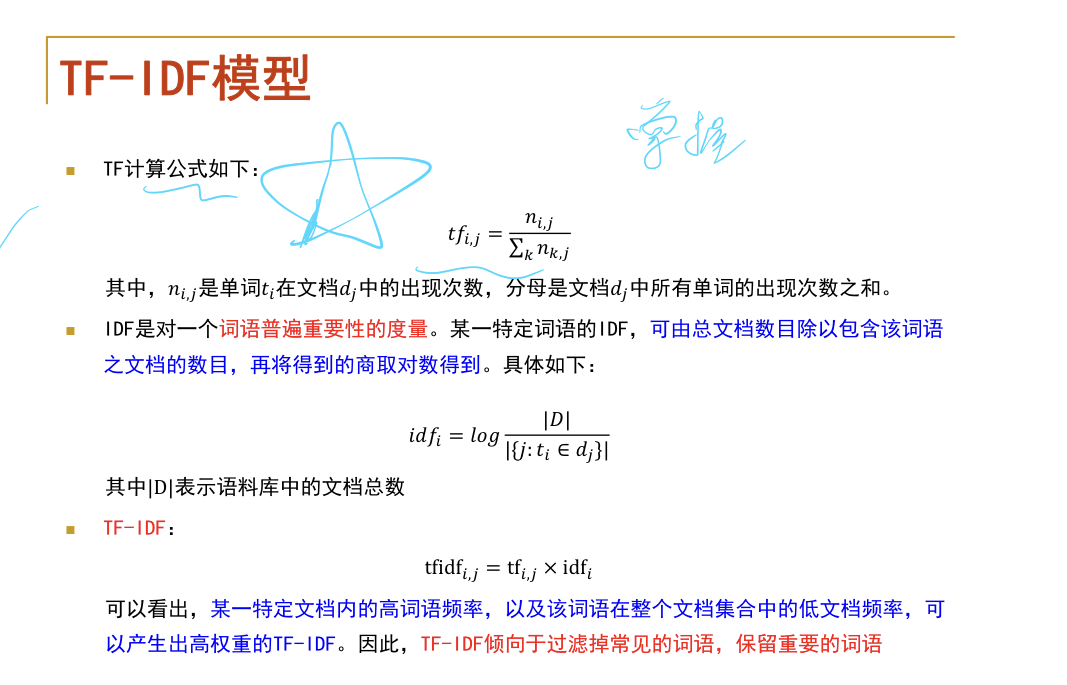


11.

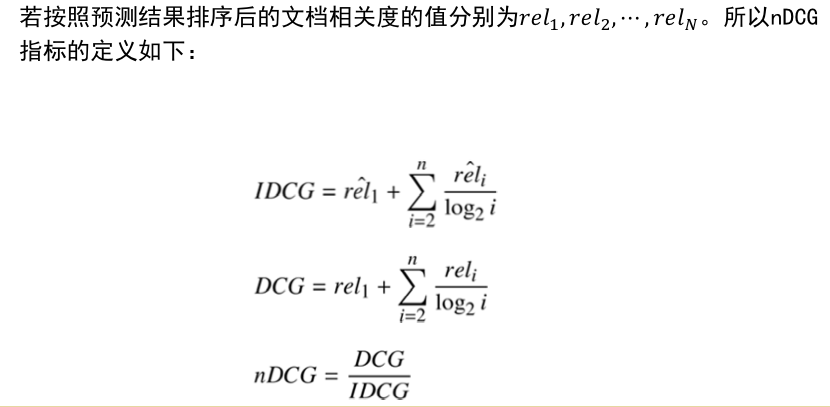




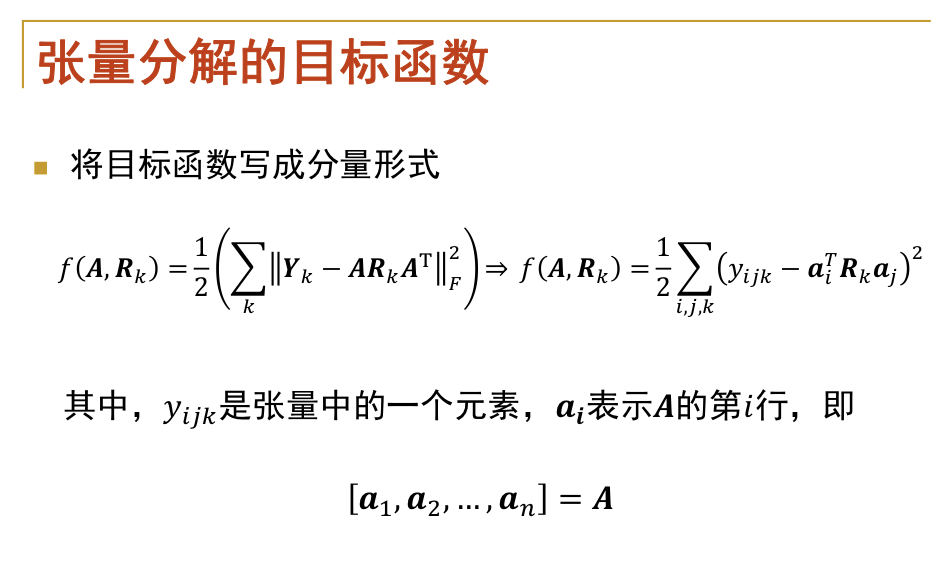
12.



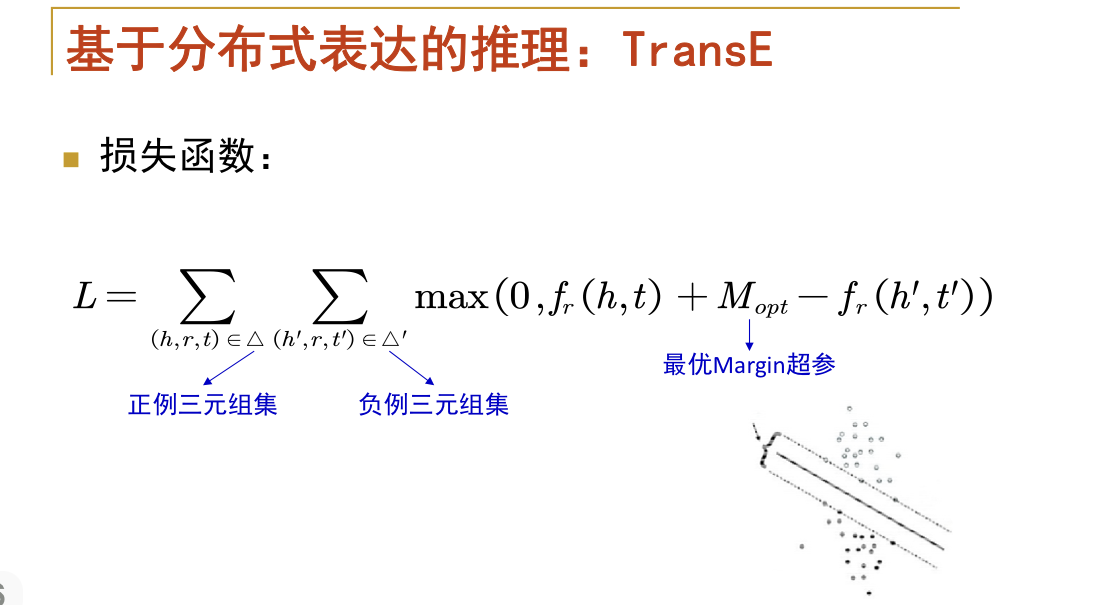
13.



14.

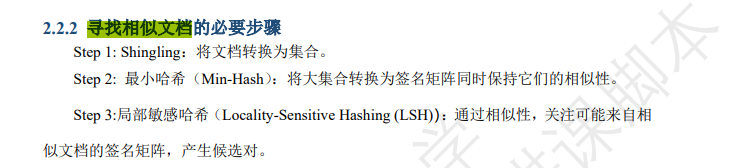


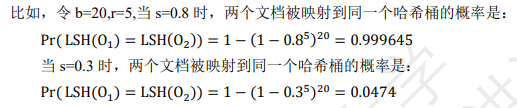
15.



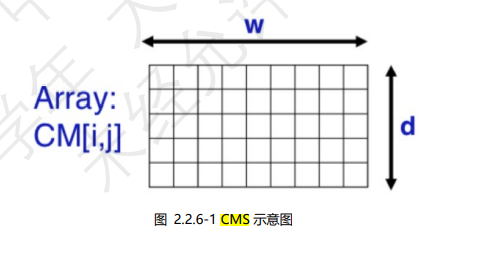
四．

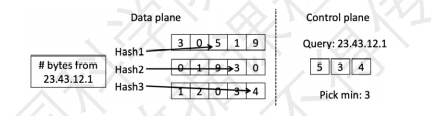
1.



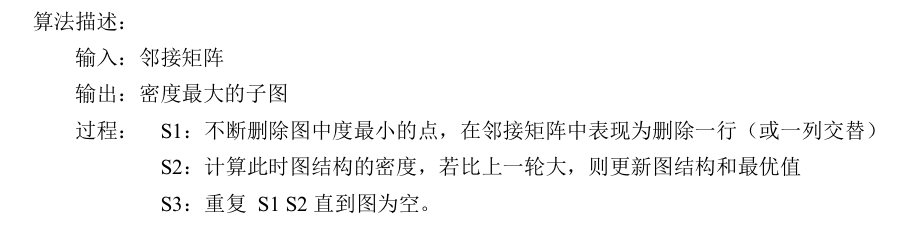


3.



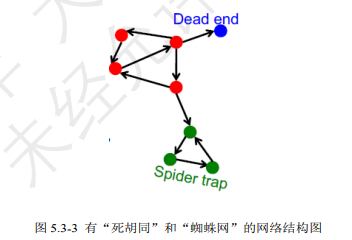


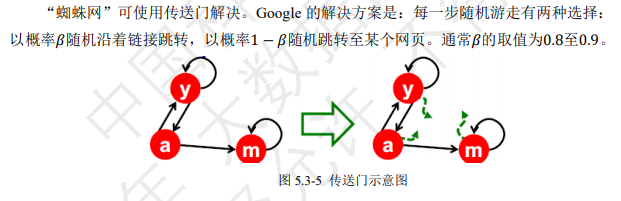
4.



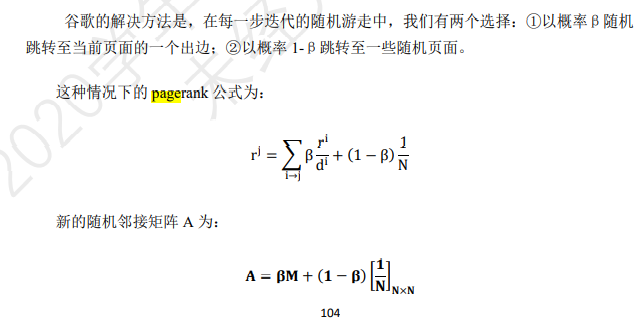
证明见常博笔记16点。

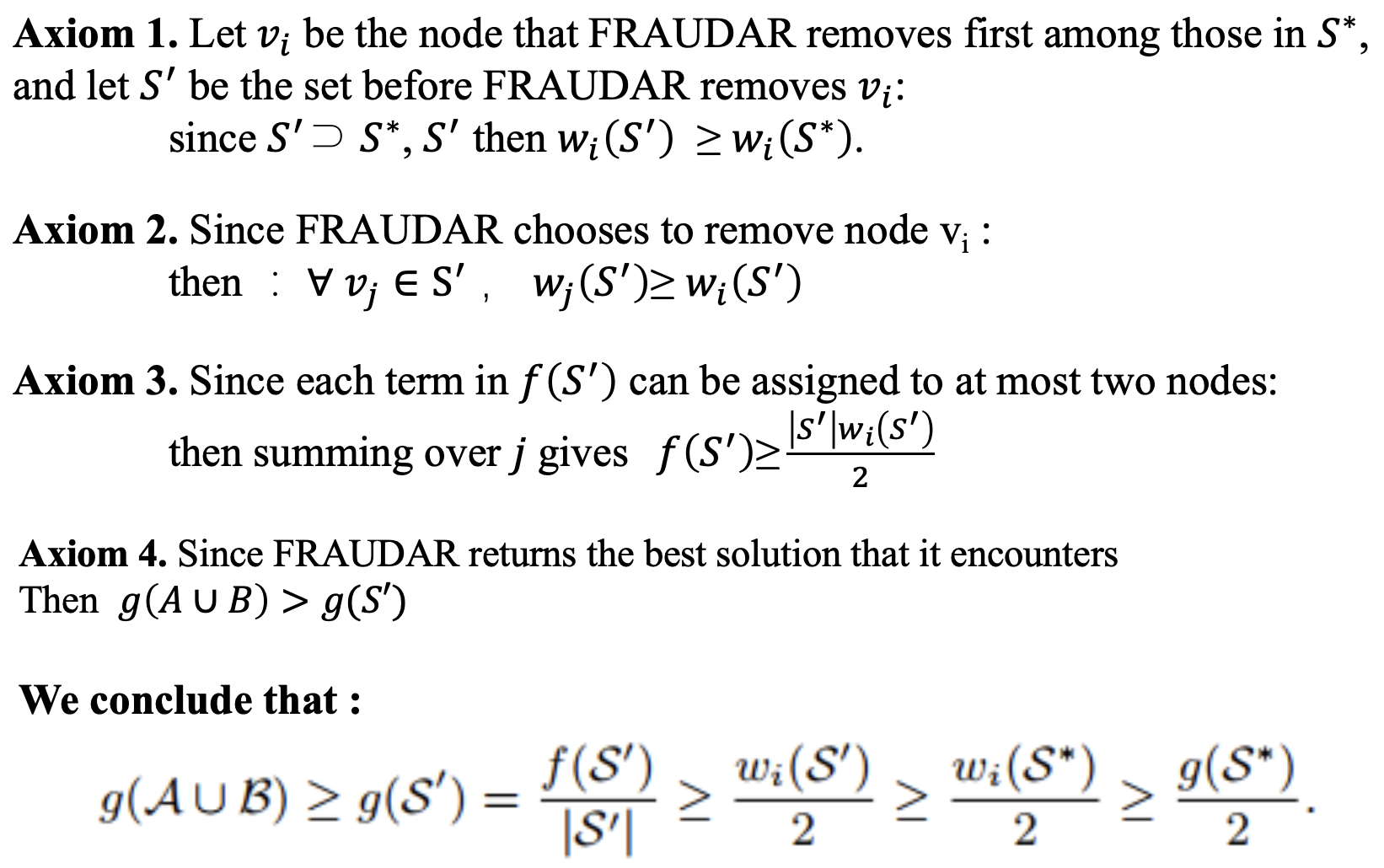
5.





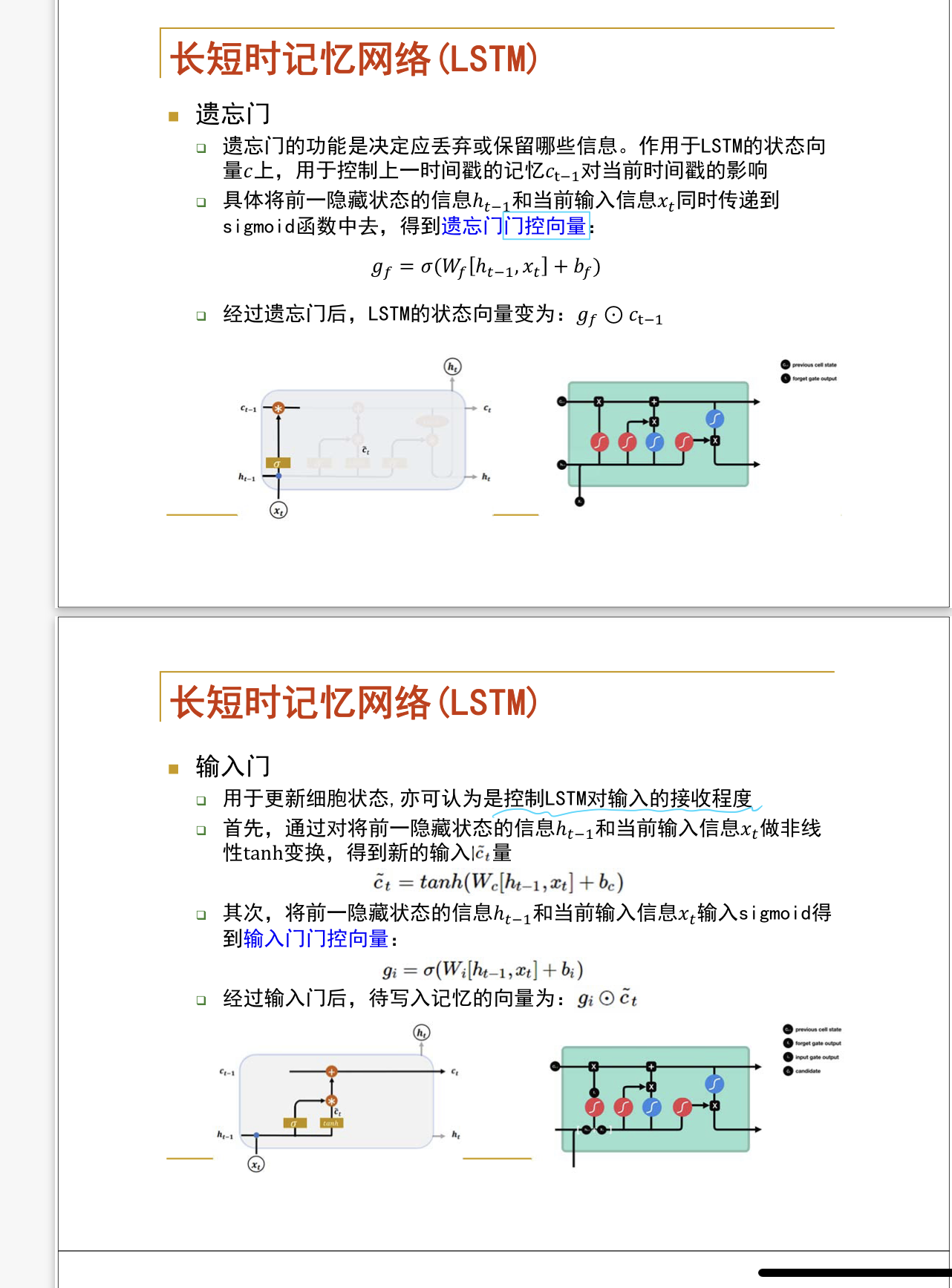




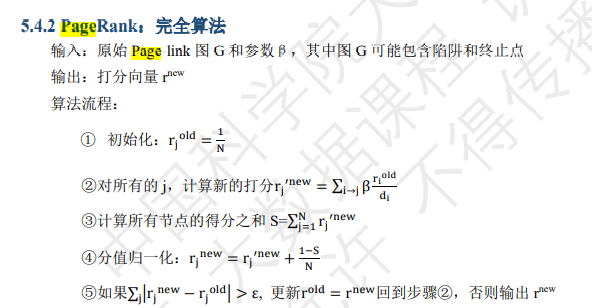


6.

、



7.



SVD是对数据进行有效特征整理的过程。

首先，对于一个m×n矩阵A，我们可以理解为其有m个数据，n个特征，（想象成一个n个特征组成的坐标系中的m个点），然而一般情况下，这n个特征并不是正交的，也就是说这n个特征并不能归纳这个数据集的特征。

SVD的作用就相当于是一个坐标系变换的过程，从一个不标准的n维坐标系，转换为一个标准的k维坐标系，并且使这个数据集中的点，到这个新坐标系的欧式距离为最小值（也就是这些点在这个新坐标系中的投影方差最大化）