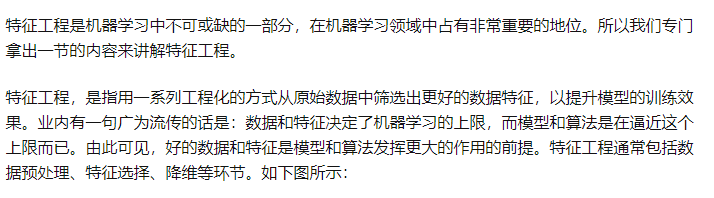
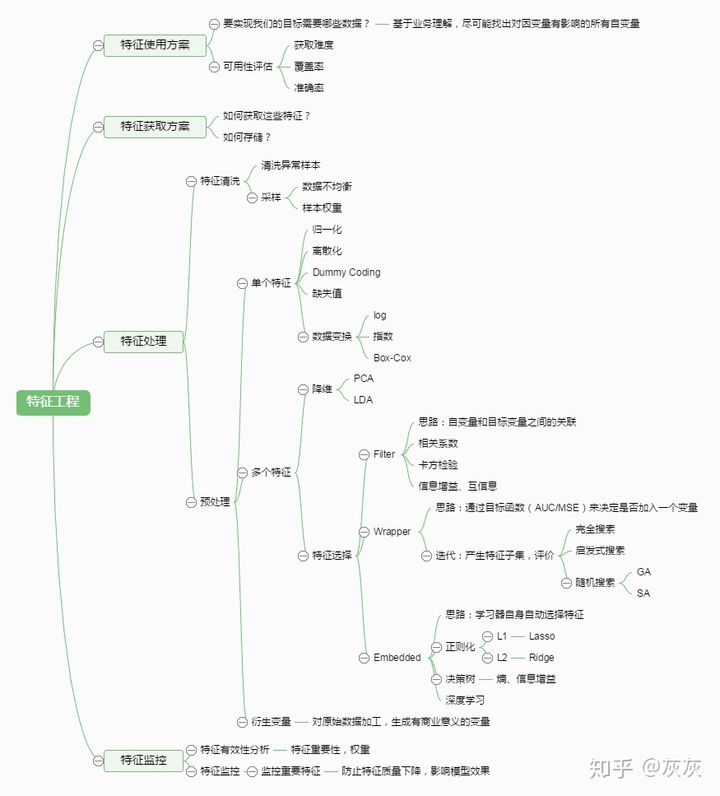
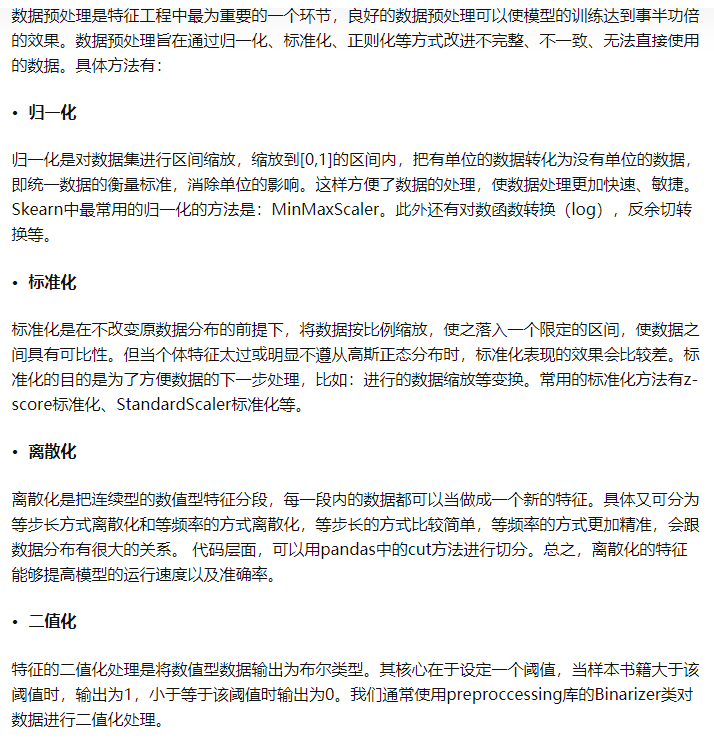
# 原理阶段

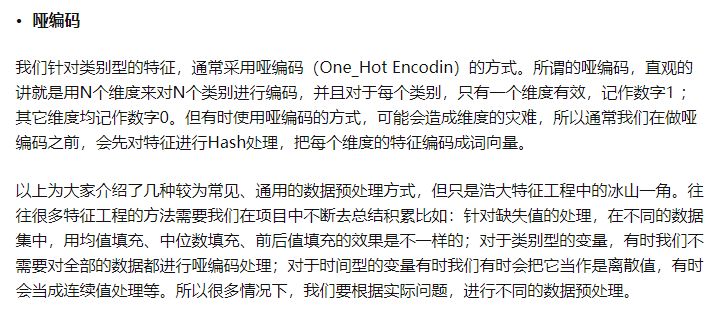
1. **特征工程的作用？**



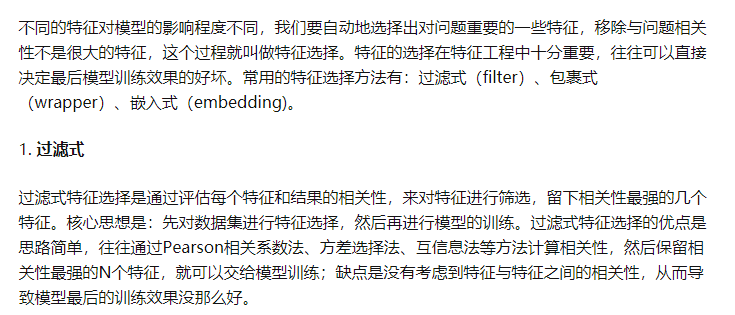


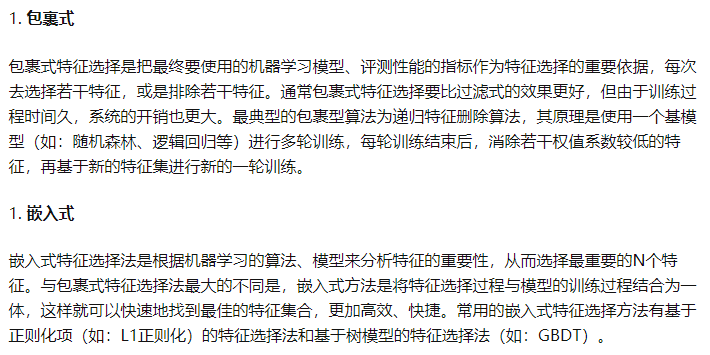
1. **数据预处理**



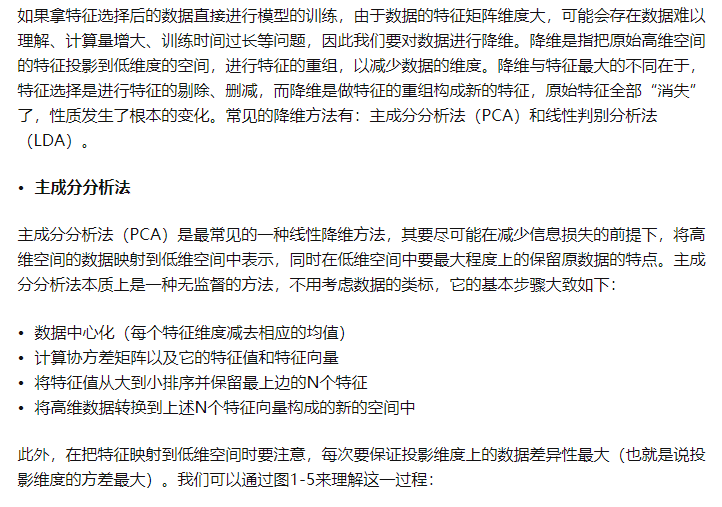


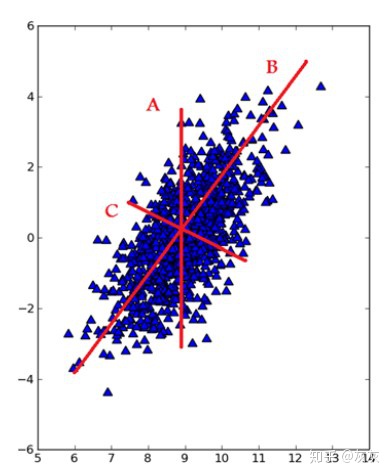
1. **特征选择**

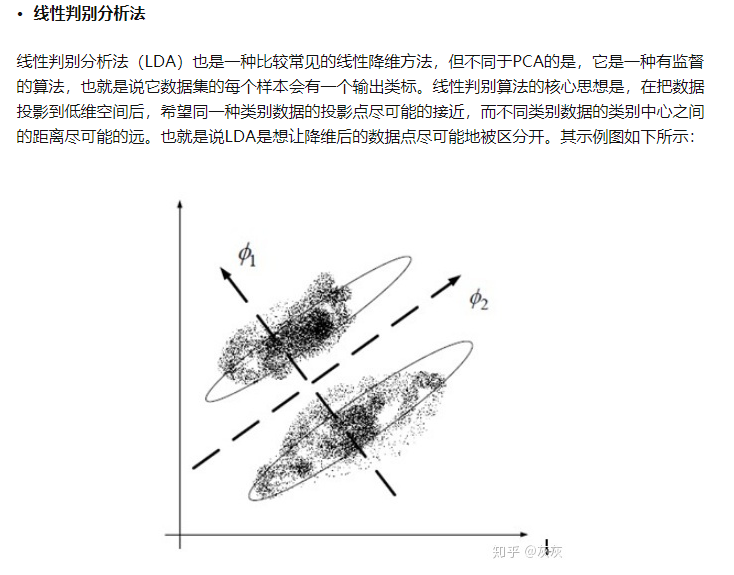




1. **降维**



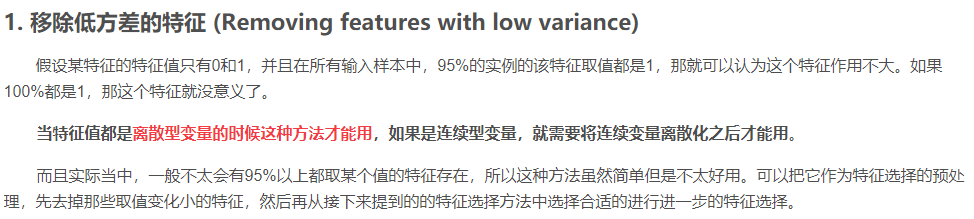




1. **再说特征选择**



1. Filter方法

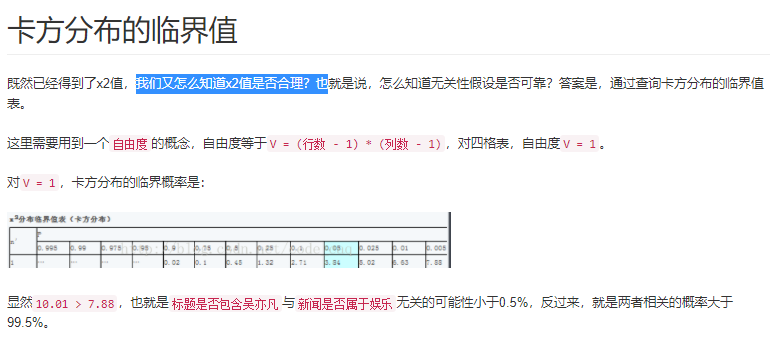




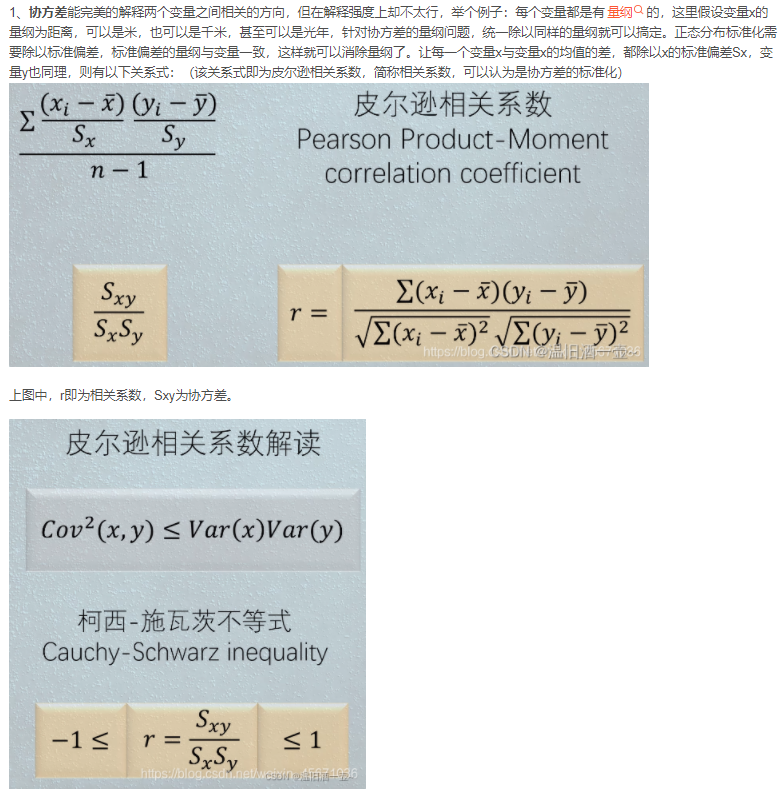
1. 卡方检验 (chi2)

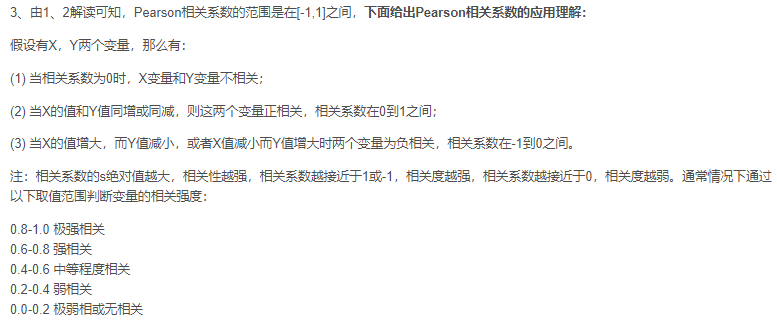


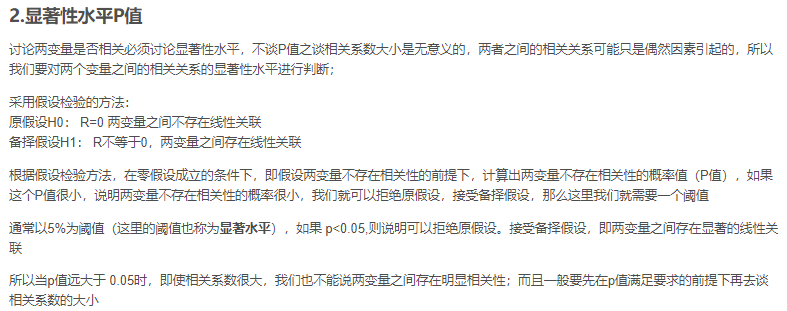




1. Pearson相关系数与P值

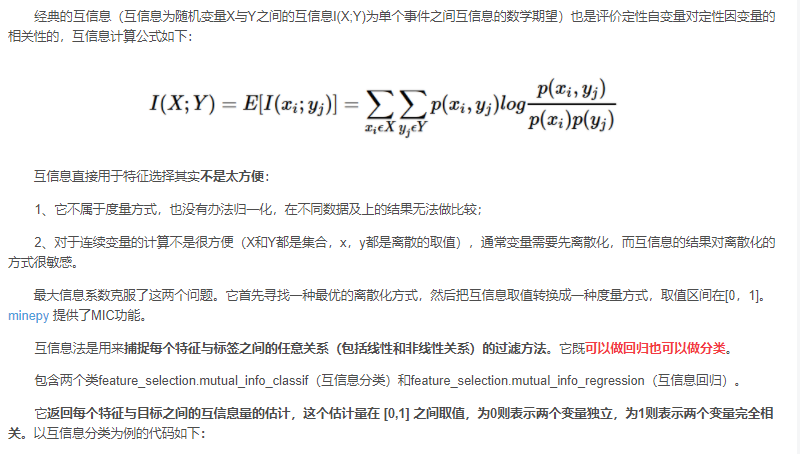




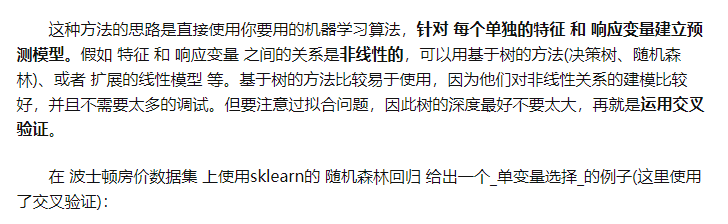


<https://zhuanlan.zhihu.com/p/115265700>(显著性水平)

1. F检验
2. 互信息和最大信息系数(MIC)

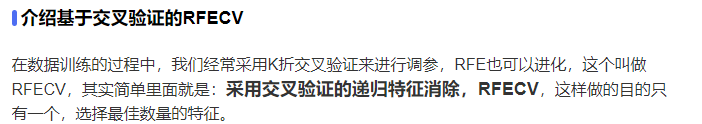


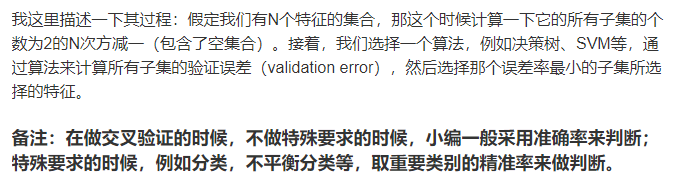
1. 基于模型的特征排序(Model based ranking)



1. Wapper-递归消除特征







1. Embedded-使用SelectFromModel方法特征选择