## 3.特征工程

1.Filter:过滤法

(Filter: 故滤法) (1965年) 未确定处于局性的复数度,然后对价有尾性按照重要程序,从后到价的选择属性。

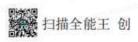
2.Embedding:嵌入式

到Emedia: 版人式 版外 把特性选择的过程作为考了过程的一部放,在考习的过程中进行特征选择。即先使用某些有器分的有法和模型进行训练,将到各个特征的对值系数。根据市值系数从大到小选择等证。这些和值系数,往往代表了等行还对于模型的贡献、对多更广道。(决解)

3.Wrapper:包裹式

型是特征选择与算法训练同时进行的版。可以调用 coef- 式 facture\_importang.
是以来完成特征选择。
但不同的是,我们往往使用一个程序的数据标准需要,并且办法们选项特征。
但是是在初始特征保护上训练评估器,并且通过 coef- 名性 或 facture\_importang。
属业是获得每个转位的重要,然后,从当前的一组特征中的制度不重要的特征。在信息的集合上选出地重复该过程,直到最终到达 所容数是的特征。
区别于过滤波和嵌入 这做一次训练的决价有 问题, 向装达要使用转征 3集进行的次训练, 因此包裹这的成本转的。

者典型666法: 益日辖证消除法(RFG) 定是一种贫心的优化第四,省在找到一旦能最佳的特征3集。它从复分连模型 新车与次选成时保留着生转但对别除最美转位,下一次选代时,它会使用上次 及模中没有被选中的特征来构建下于模型,直到幼有特征新栽建社。 然后,根据包保留或别除特征的顺序来对特征进行利息,



Date:\_	_\Page:_
---------	----------

最终选出一播生3集。 包裹这是所有特征选择3这中最有利于提升模型表现的,它可以使用很多的特征概达到很优秀的效果。但其计算是大,不太数5天大型的数据。