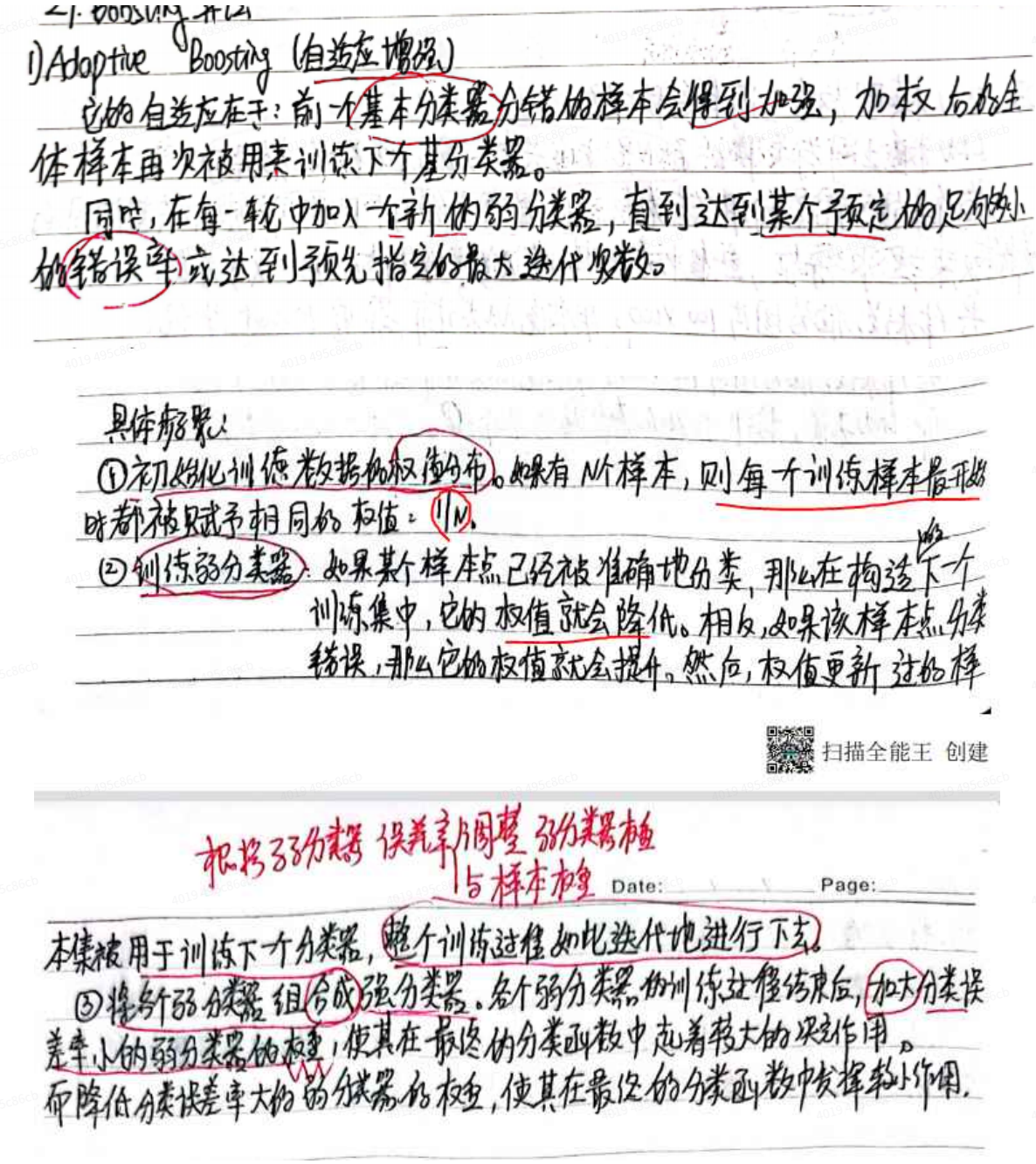
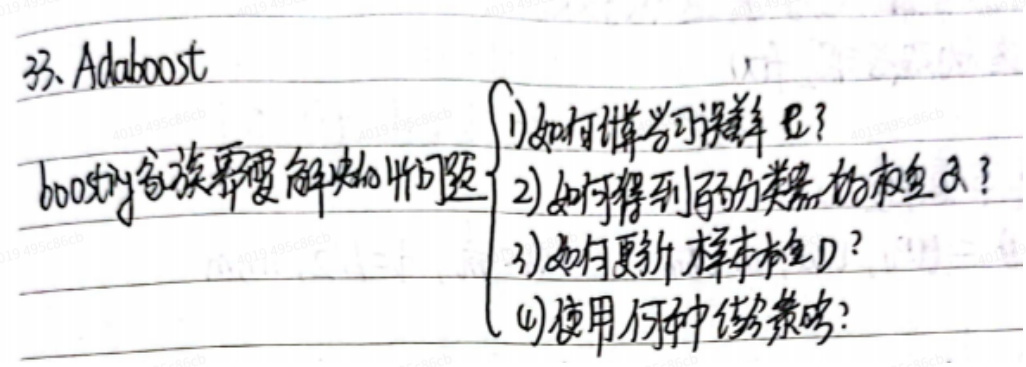
4.Adaboost

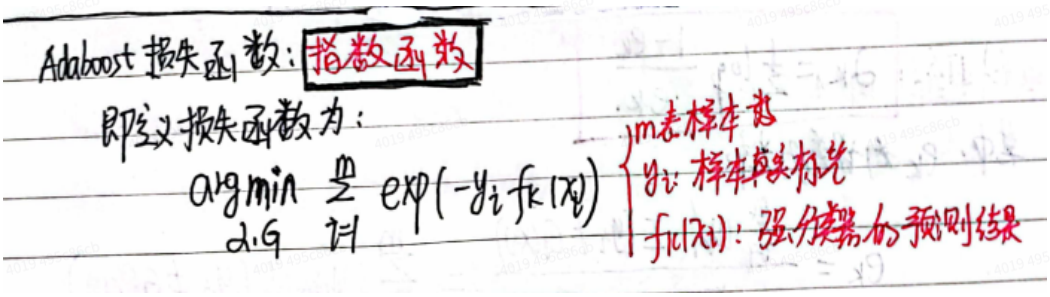
4.1Adaboost实现原理



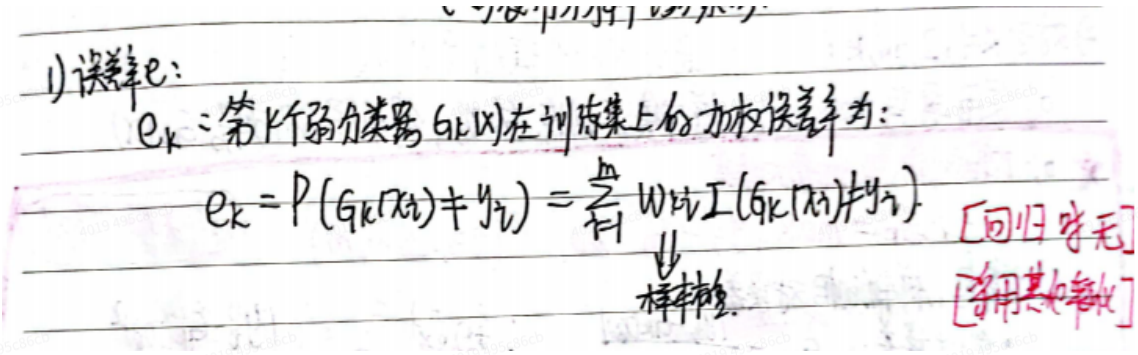
4.2Adaboost做具体实现



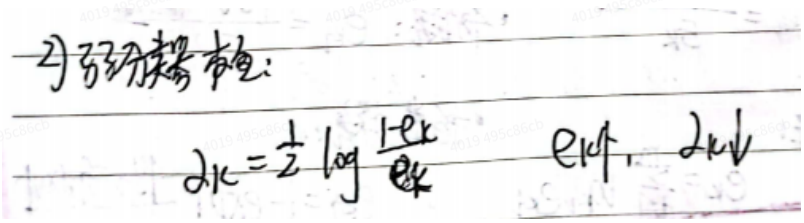
4.2.1分类



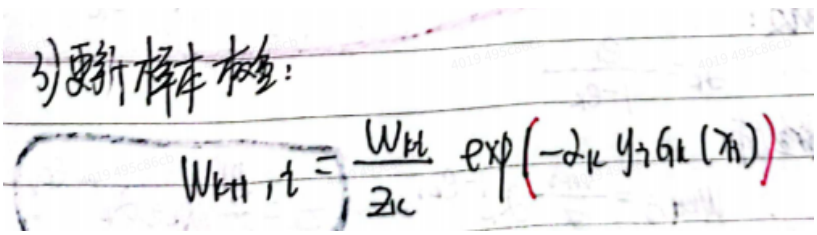
1）误差率e



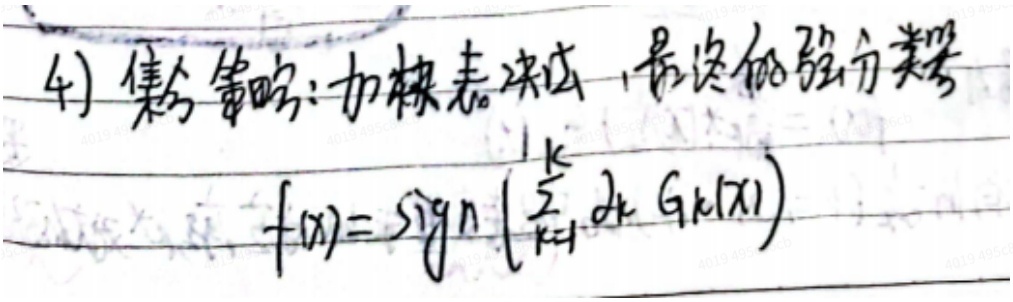
2）弱分类器权重



3） 更新样本权重

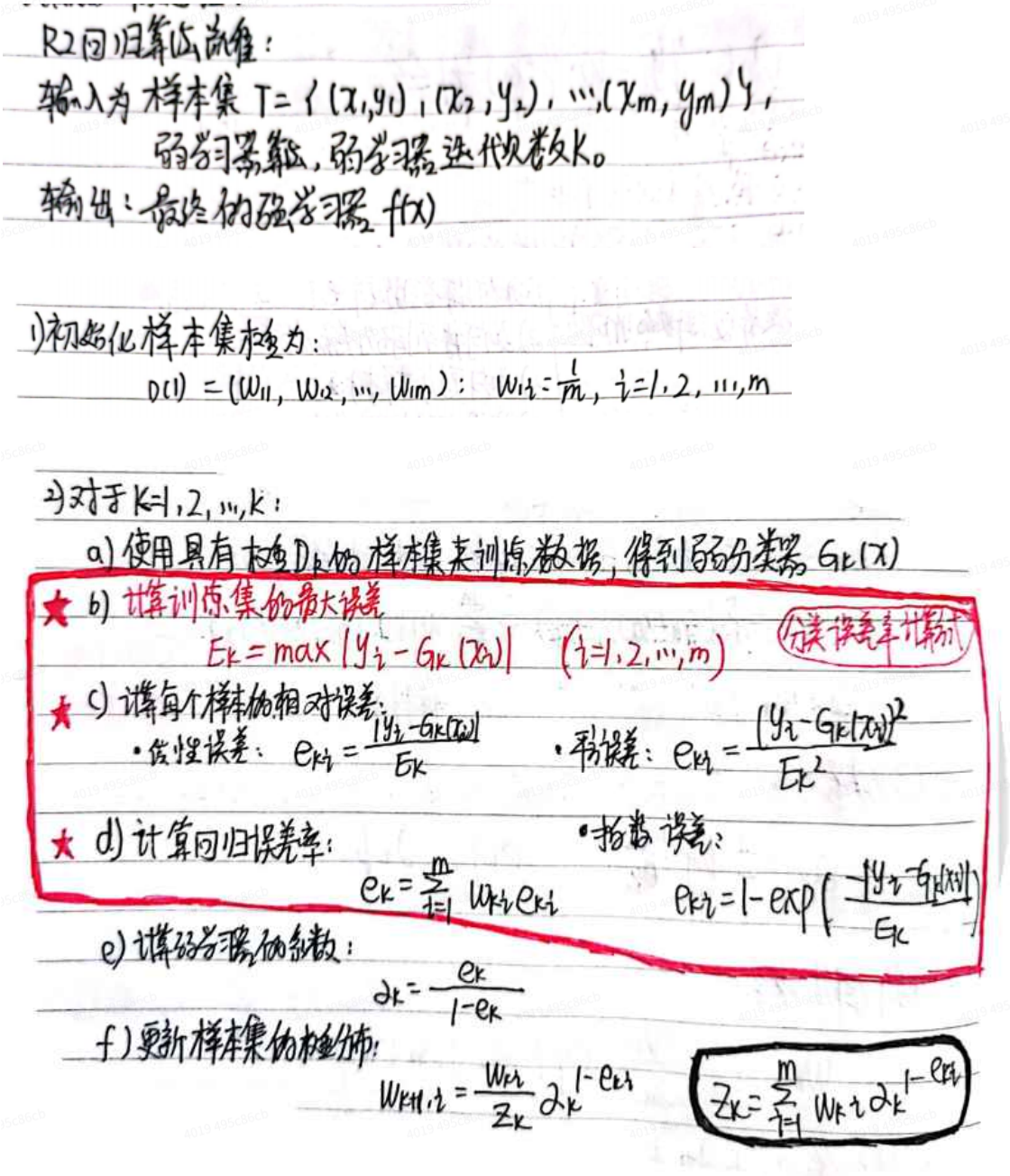


4） 集成策略



4.2.2回归

R2回归流程：



$$f(x)=\sum_{k=1}^{K}\left(\ln\left(\frac{1}{\alpha_k}\right)\right)g(x)$$

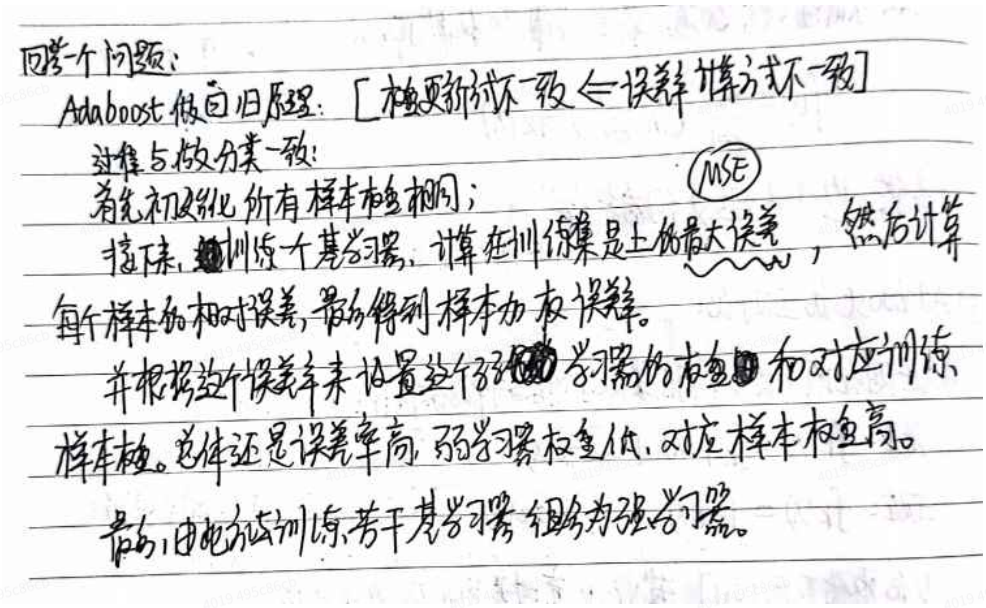
其中 $g(x)$ 是所有 $\alpha_k G_k(x)$，$k=1,2,\dots,K$ 的中位数。

但有点看不懂依据，还是采用加权平均法：

$$f(x)=\sum_{k=1}^{K}\left(\ln\frac{1}{\alpha_k}\right)G_k(x)$$

当然，也可以自定义组合策略。

背诵版



4.3Adaboost优缺点

⑥ Adaboost优缺点

Adaboost的基学习器最常用的是：决策树与神经网络。

对于决策树：Adaboost分类用于CART分类树，

Adaboost回归用了CART回归树

优点：1）作为分类器时，分类精度很高

2）可以完成回归分类与回归

3）不易发生过拟合

Date:\ \ Page:

缺点：

对异常样本敏感，异常样本在迭代中可能会获得较高权重，最终影响强学习器的预测准确性。