

均值预测

——一次效果欠佳的记录

“Deadline 并不是第一生产力，它只是给了你将一堆 shit 一样的东西 push 上去的勇气” /笑 cry

问题描述：给定 3 月-6 月的用户行为数据和歌曲信息，预测 7-8 月艺人每天的播放量，测评公式如下

$$\sigma_j = \sqrt{\frac{1}{N} \sum_{k=1}^N ((S_{j,k}) / (T_{j,k}))^2}$$
$$\phi_j = \sqrt{\sum_{k=1}^N T_{j,k}}$$
$$F = \sum_j (1 - \sigma_j) * \phi_j$$

where $S_{j,k}$ 表示艺人 j 在第 k 天的预测播放量， $T_{j,k}$ 则表示其实际播放量

数据集： <https://pan.baidu.com/s/1eRR3fAM>

总体思路：

读取数据→预处理数据→可视化观测数据→简单粗暴地使用均值法尝试

最终得分： 5670.30805795