技巧：

1、多看、多学、多实践

2、多看kernel, discussion

3、抱团取暖更好，切忌单打独斗

怎么打比赛？

1. 背景知识：数据分布
2. 查阅资料
3. 集思广义
4. 团队协作
5. 操作性极强：需要自己复现代码。

推荐资料：

《统计学习方法》

《利用Python进行数据分析》

sklearn、pandas、xgboost、lightgbm、numpy、ml-en官网

数据明确 -> 解决问题的思路

1. 数据一明确之后就要看到这是一个什么问题？（回归？分类？）
2. 这时候就要想起有哪些算法可以做回归？
3. 那么线性回归要什么样的数据才可以用上呢？
4. 数据中是否含有字符串或者缺失值？如何把它们变成数值型的？
5. 数据特征工程的思路，数据的EDA（pandas\_profling），特征选择，再就是特征组合或者特征分割等等。
6. 然后再就是算法的选择。