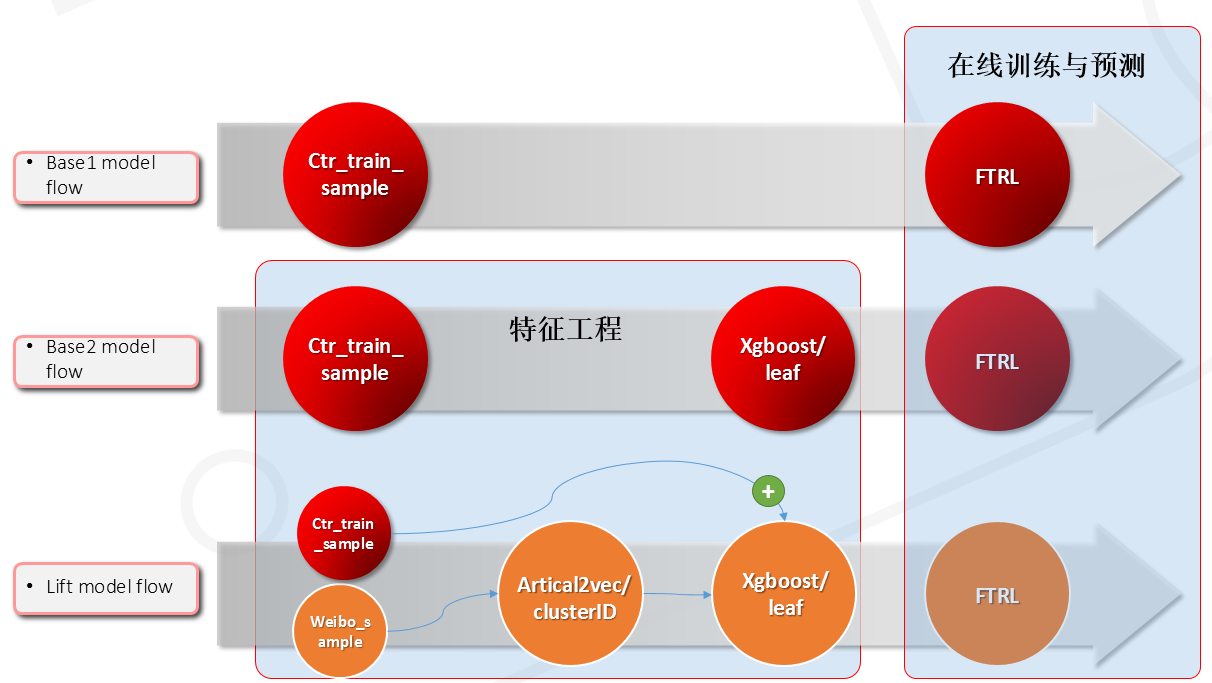
### 实验设计流程：



#### 实验说明：

* Base1 model flow
  + 使用数据ctr\_train\_sample，不做任何特征工程，直接用ftrl进行训练和测试，计算误差：**error1**
  + FTRL原理参考：<http://www.cnblogs.com/EE-NovRain/p/3810737.html>
* Base2 model flow
  + 使用数据ctr\_train\_sample，采用xgboost+leaf做特征工程，然后用ftrl进行训练和测试，计算误差：**error2**
  + xgboost+leaf原理参考：<http://www.cbdio.com/BigData/2015-08/27/content_3750170.htm>
* Lift model flow
  + 增加weibo文本语料（weibo\_sample），进行特征转换（article2vec+clusterID），然后使用xgboost+leaf进行特征转换，生成leaf文件，添加到ctr\_train\_sample尾部，然后采用ftrl进行训练和测试，计算误差：***error3***
  + article2vec+clusterID原理参考：<http://www.jianshu.com/p/d34d61188ab5>
* 比较三个误差的大小，理论上应该是Lift model 最小