Feature List

Subject

* **duration**: 加入课程的时间/记录log的时间跨度
* **duration\_less2day**: duration是否小于2天（<=2）
* **duration\_less1week**: duration是否小于1周（<=7）
* **duration\_2week**: duration是否小于2周（<=14）
* **duration\_1month**: duration是否小于20天
* **request\_count**: log记录的总个数（排除page\_close）
* **request\_count\_lst2week**: 最近2周的log记录的总个数
* **request\_count\_lst1week**：最近1周的log记录的总个数
* **event\_count**: log记录的总个数（Note：不排除page\_close）
* event\_count\_lst2week：最近2周的log记录的总个数
* event\_count\_lst1week：最近1周的log记录的总个数
* active\_days：总共有多少天用户时有log记录的
* active\_days\_freq: #active\_days / #duration
* active\_days\_per\_week: 平均每周有多少天有log
* active\_days\_lst2week：最近2周的activedays
* active\_days\_lst1week：最近1周的activedays
* server\_{access, nagivate, wiki, discussion,problem}\_count: source为server的不同动作的个数
* server\_{access, nagivate, wiki, discussion,problem}\_count\_lst2week
* server\_{access, nagivate, wiki, discussion,problem}\_count\_lst1week
* browser\_{access,video, problem, page\_close}\_count: source为browser的不同动作的个数
* browser\_{access,video, problem, page\_close}\_count\_lst2week
* browser\_{access,video, problem, page\_close}\_count\_lst1week
* problem\_request\_count: 请求problem的次数（多次访问同一个problem的情况存在）
* problem\_request\_count\_lst2week
* problem\_count: 回答问题的个数
* problem\_count\_lst2week
* problem\_per\_day： 平均每天（active days）回答问题的个
* problem\_per\_day\_lst2week
* video\_request\_count：请求video的次数
* video\_request\_count\_lst2week
* video\_count：video的个数
* video\_count\_lst2week
* video\_per\_day：每天平均看过的video的个数
* video\_per\_day\_lst2week
* request\_module\_count：请求课程资源的次数
* request\_module\_count\_lst2week
* request\_module\_count\_lst1week
* module\_count：总共看过课程module的个数
* module\_count\_lst2week
* module\_count\_lst1week
* module\_per\_day：平均每天看过多少module
* module\_per\_day\_lst2week
* module\_per\_day\_lst1week
* session\_count：session的个数（最小间隔为30分钟）
* session\_count\_lst2week
* session\_count\_lst1week
* session\_per\_day：平均每天session的个数
* request\_hour\_count：e.g., 如果用户在7am和9am这两个小时里面有动作，那么这个feature为2
* request\_hour\_mean
* request\_hour\_std
* request\_lag\_min: 最小间隔天数
* request\_lag\_max
* request\_lag\_mean
* request\_lag\_std
* total\_lag：所有间隔的总和
* lag>2d\_times
* lag>3d\_times
* lag>5d\_times
* lag>7d\_times
* sumlag<1week
* sumlag1<>2week
* sumlag>2week
* staytime\_min：每天停留在course上的最少时间
* staytime\_max
* staytime\_mean
* staytime\_std
* staytime\_median
* staytime\_25p
* staytime\_75p
* request\_weekend\_count：周末请求的个数
* request\_weekend\_percentage：周末请求的个数所占总请求的比例
* more\_weekend：用户是否大部分在周末上课
* day\_time：发生在白天的动作个数 > 发生在晚上的动作个数
* night\_time：发生在白天的动作个数 < 发生在晚上的动作个数
* daytime\_ratio: 发生在白天的动作个数 / 总个数
* exam\_time\_fstday：第一次课程记录是否发生在考试月（6月，7月，12月，1月）
* holiday\_fstday：第一次课程记录是否发生在假期内 （7月，8月，2月）
* exam\_time\_lstday：最后一次课程记录时发呕发生在考试月内
* holiday\_lstday
* in\_exam：所有课程记录是否发生在考试月内
* in\_holiday：
* module\_and\_problem\_count：看过课程module同时看过problem的个数
* module\_problem\_ratio
* module\_not\_problem\_count
* module\_not\_problem\_ratio
* problem\_over3minutes\_count：看过problem大于3分钟的次数
* video\_over10minutes\_count
* server\_count：来源是server的动作个数
* browser\_count
* browser\_server\_ratio
* year\_2013：记录是否发生在2013年
* year\_2014