F:

其中：

所以：

拆分成子问题，对每个子问题分别建模：

File1: user\_id,song\_id,gmt\_create,action\_type,Ds

File2: song\_id,artist\_id,publish\_time,song\_init\_plays,Language,Gender

Result: artist\_id,plays,ds

特征维度：

User 307425

Song 10842

Artist 50

切入点：

每个song的播放记录、每个user的播放记录

先做基本的回归

分析每首歌每天的user的重复率和新的user的规模。并可以考虑给user赋权。