2.2769536396899803

Use whole brand

pred = logistic\_reg(dtrain\_X,dtrain\_y,dtest\_X, C=0.019)

log validation -> 2.2820224548209187

log test -> 2.27349 2.27119

Use top10 brand

pred = logistic\_reg(dtrain\_X,dtrain\_y,dtest\_X, C=0.019)

log Validation -> 2.2829285726670925

log Test -> 2.27453 2.27208

Xgb validation -> 2.3462112432259534

Xgb test

Add Time

log validation -> 2.2805166549913034

log test -> 2.27261 2.27065

Xtrain = hstack((Xtr\_brand, Xtr\_model, Xtr\_app, Xtr\_label,Xtr\_tod), format='csr')

validatioin 2.275592462992458

test 2.26625 2.26426

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Features | CV\_los | Kaggle\_test | Kaggle\_true |
| Logistic (origin parm) | Xte\_brand, Xte\_model, Xte\_app, Xte\_label | 2.4808 |  |  |
| Logistic (best parm) | Xte\_brand, Xte\_model, Xte\_app, Xte\_label | 2.2735 |  |  |
| Xgboost (origin parm) | Xte\_brand, Xte\_model, Xte\_app, Xte\_label | 2.3221 |  |  |
| Xgboost (best parm) | Xte\_brand, Xte\_model, Xte\_app, Xte\_label | 2.3059 |  |  |
| RandomForest (origin) |  | 3.2657 |  |  |
| RandomForest (best parm) |  | 2.3769 |  |  |
| Logistic (best parm) | Xte\_brand, Xte\_model, Xte\_app, Xte\_label,Xte\_tod | 2.2728 |  |  |
| Xgboost (best parm) | Xte\_brand, Xte\_model, Xte\_app, Xte\_label,Xte\_tod | 2.3068 |  |  |
| Logistic (best parm) | Xte\_brand\_top10, Xte\_model, Xte\_app, Xte\_label,Xte\_tod | 2.2728 |  |  |
| Xgboost (best parm) | Xte\_brand\_top10, Xte\_model, Xte\_app, Xte\_label,Xte\_tod | 2.3075 |  |  |
| Logistic (best parm) | Xte\_brand, Xte\_model, Xte\_app, Xte\_label,Xte\_tod,Xte\_brand\_top10, Xte\_label\_topbot10 | 2.273 | 2.26614 | 2.26411 |
| Xgboost (best parm) |  | 2.3078 |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

 