

# Rainforest Post Process - Private LB 0.974!

In this notebook we demonstrate a post process for rainforest comp. We will use the output from the following notebook: <https://www.kaggle.com/meaninglesslives/rfcx-minimal> We will increase it's private LB score from 0.964 to 0.974!

To learn more about this post process, read the discussion [here](#)

## Post Process Parameters

To use the following post process, load your `submission.csv` file and first try `MODE=1` . Next try `MODE=2` . Next try `MODE=3` . If those three don't increase your LB, then try `MODE=1` with different `FUDGE` values. Try values 0.5, 1, and 3.

```
In [1]: # USE MODE 1, 2, or 3
MODE = 1

# LOAD SUBMISSION
import pandas as pd, numpy as np
FUDGE = 2.0
FILE = '../input/rfcx-minimal/submission.csv'
df = pd.read_csv(FILE)
for k in range(24):
    df.iloc[:,1+k] -= df.iloc[:,1+k].min()
    df.iloc[:,1+k] /= df.iloc[:,1+k].max()

# CONVERT PROBS TO ODDS, APPLY MULTIPLIER, CONVERT BACK TO PROBS
def scale(probs, factor):
    probs = probs.copy()
    idx = np.where(probs!=1)[0]
    odds = factor * probs[idx] / (1-probs[idx])
    probs[idx] = odds/(1+odds)
    return probs

# TRAIN AND TEST MEANS
d1 = df.iloc[:,1:].mean().values
d2 = np.array([113,204,44,923,53,41,3,213,44,23,26,149,255,14,123,222,46,6

for k in range(24):
    if MODE==1: d = FUDGE
    if MODE==2: d = d1[k]/(1-d1[k])
    if MODE==3: s = d2[k] / d1[k]
    else: s = (d2[k]/(1-d2[k]))/d
    df.iloc[:,k+1] = scale(df.iloc[:,k+1].values,s)

df.to_csv('submission_with_pp.csv',index=False)
```