Untitled

May 18, 2018

1 CLASSIFIERS WITH NEW FEATURES

```
In [1]: import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        import warnings
        warnings.simplefilter('ignore')

In [2]: df = pd.read_csv('new_train.csv')

In [3]: target = ['booking_bool', 'click_bool']

In [4]: data = ['year', 'month', 'prop_score','prop_location_score1','prop_location_score2','dif
                        'usd_diff', 'star_diff', 'srch_query_affinity_score', 'orig_destination_
                        'prop_review_score','position','price_rank','stars_rank','score_rank']

In [7]: from sklearn.cross_validation import cross_val_score

C:\Users\Ryan\Anaconda2\lib\site-packages\sklearn\cross_validation.py:41: DeprecationWarning: Th
  "This module will be removed in 0.20.", DeprecationWarning)


In [5]: df.fillna(value='0', inplace=True)

In [6]: df_sample = df.sample(frac=0.1)
        X = df_sample[data]
        y = df_sample[target]
```

1.1 K-Neighbors Classifier

```
In [8]: from sklearn.neighbors import KNeighborsClassifier
        knn = KNeighborsClassifier(n_neighbors=5)
        print cross_val_score(knn, X, y, cv=10, scoring='accuracy').mean()

0.9531838825629324
```

1.2 Random Forest

```
In [11]: from sklearn.ensemble import RandomForestClassifier
         clf = RandomForestClassifier(max_depth=2, random_state=0)
         clf.fit(X, y)
         print cross_val_score(clf, X, y, cv=10, scoring='accuracy').mean()
```

```
0.95462071597051
```

1.3 Neural network

```
In [12]: from sklearn.neural_network import MLPClassifier
         clf = MLPClassifier()
         clf.fit(X, y)
         print cross_val_score(clf, X, y, cv=10, scoring='accuracy').mean()
```

```
0.9545590767483214
```

1.4 Decision Tree

```
In [14]: from sklearn import tree
         clf = tree.DecisionTreeClassifier()
         clf = clf.fit(X, y)
         print cross_val_score(clf, X, y, cv=10, scoring='accuracy').mean()
```

```
0.9073922551850027
```

2 ENSEMBLE CLASSIFIERS

```
In [12]: from sklearn.ensemble import BaggingClassifier
         from sklearn.neural_network import MLPClassifier
         from sklearn.ensemble import RandomForestClassifier
```