GTB

March 16, 2019

```
In [6]: #Gradient tree bosting
        import pickle
        covertype_dataset = pickle.load(open('covertype_dataset.pickle','rb'))
        covertype_X = covertype_dataset.data[:15000,:]
        covertype_Y = covertype_dataset.target[:15000] -1
        covertype_val_X = covertype_dataset.data[15000:20000,:]
        covertype_val_Y = covertype_dataset.target[15000:20000] -1
        covertype_test_X = covertype_dataset.data[20000:25000,:]
        covertype_test_Y = covertype_dataset.target[20000:25000] -1
```

```
In [7]: from sklearn.model_selection import cross_val_score,StratifiedKFold
        from sklearn.ensemble import GradientBoostingClassifier
        hypothesis = GradientBoostingClassifier(max_depth=5, n_estimators=50, random_state=101)
        hypothesis.fit(covertype_X, covertype_Y)
```

```
Out[7]: GradientBoostingClassifier(criterion='friedman_mse', init=None,
                      learning_rate=0.1, loss='deviance', max_depth=5,
                      max_features=None, max_leaf_nodes=None,
                      min_impurity_decrease=0.0, min_impurity_split=None,
                      min_samples_leaf=1, min_samples_split=2,
                      min_weight_fraction_leaf=0.0, n_estimators=50,
                      n_iter_no_change=None, presort='auto', random_state=101,
                      subsample=1.0, tol=0.0001, validation_fraction=0.1,
                      verbose=0, warm_start=False)
```

```
In [9]: from sklearn.metrics import accuracy_score
        print("GradientBoostingClassifier -> accuracy on test set:", accuracy_score(covertype_t
```

```
GradientBoostingClassifier -> accuracy on test set: 0.782
```

```
In [ ]:
```