6b_final_project_checkin_template

November 22, 2019

Table of Contents

0.1 Final Project Check-in

0.2 Richter

0.3 Student Names

1. Nithish Bolleddula
2. Alan Flint
3. Rushil Sheth
4.

0.4 Load Data

[2]:
```
import pandas as pd
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import precision_score


df = pd.read_csv('../input/train_values.csv')
train_target = pd.read_csv('../input/train_labels.csv')

cat_cols = df.columns[df.dtypes == 'object']
X_train = pd.get_dummies(df, columns=cat_cols)
y_train = train_target.damage_grade
```

0.5 Fit scikit-learn model

[3]:
```
estimator = RandomForestClassifier(n_estimators=30,min_samples_leaf=100)
estimator.fit(X_train, y_train)
```

[3]:
```
RandomForestClassifier(bootstrap=True, class_weight=None, criterion='gini',
                       max_depth=None, max_features='auto', max_leaf_nodes=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=100, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, n_estimators=30,
                       n_jobs=None, oob_score=False, random_state=None,
                       verbose=0, warm_start=False)
```

Evaluation Metric

[4]:
```
preds = estimator.predict(X_train)
score = precision_score(y_train, preds, average='micro')

print(f"micro averaged precision score is = {score}")
```

```
micro averaged precision score is = 0.6755461414192578
```