

ML-OPS  
4- Assignment  
Readme File

Name: Rinki Kumari  
Roll no: M21AE248  
Email: [kumari.23@iitj.ac.in](mailto:kumari.23@iitj.ac.in)

**Questions 7:** write the performance metrics from step 3 in the readme.md. in following format  
run svm decision tree.

**Answer:**

1. Import `from sklearn import tree`  
`clf = tree.DecisionTreeClassifier()`
2. Split train, test and validation data set in 5 different and set hyper-parameter turning for both the classifiers

```
def h_param_tuning(h_param_comb, clf, x_train, y_train, x_dev, y_dev, metric, verbose=False):  
    best_metric = -1.0  
    best_model = None  
    best_h_params = None  
    for cur_h_params in h_param_comb:  
        hyper_params = cur_h_params  
        clf.set_params(**hyper_params)  
        clf.fit(x_train, y_train)  
        predicted_dev = clf.predict(x_dev)  
        cur_metric = metric(y_pred=predicted_dev, y_true=y_dev)  
        if cur_metric > best_metric:  
            best_metric = cur_metric  
            best_model = clf  
            best_h_params = cur_h_params  
        if verbose:  
            print("Found new best metric with :" + str(cur_h_params))  
            print("New best val metric:" + str(cur_metric))  
    return best_model, best_metric, best_h_params
```

3. mean and standard deviations of both the classifier's performances  
Using pandas' data frame, I computed mean and standard deviation.

**SVM:**

**Mean: 0.991051**

**Standard Deviation: 0.003058**

**Decision Tree:**

**Mean: 0.8413**

**Standard Deviation: 0.018779**

4. **Model Comparison:**

```
best metric on dev was 0.8729281707933801  
{ 'svm': [(0.9888268156424581), (0.994413407821229), (0.9776536312849162), (0.9832402234636871), (0.994413407821229)], 'decision_tree': [(0.8435754189944135), (0.88268156424581), (0.8156424581005587), (0.888268156424581), (0.8547486033519553)] }
```

5. Identify if there is more to the classifier comparison than just the numbers? -- Can you somehow compare the predicted labels

```
Best hyperparameters were: {'gamma': 0.0007, 'C': 2}
Best Metric on Dev was: 1.0
Best hyperparameters were: {'max_depth': 20}
Best Metric on Dev was: 0.9005524861878453
Best hyperparameters were: {'gamma': 0.0007, 'C': 0.3}
Best Metric on Dev was: 0.994475138121547
Best hyperparameters were: {'max_depth': 10}
Best Metric on Dev was: 0.8729281767955801
Best hyperparameters were: {'gamma': 0.0007, 'C': 0.3}
Best Metric on Dev was: 0.994475138121547
Best hyperparameters were: {'max_depth': 20}
Best Metric on Dev was: 0.850828729281768
Best hyperparameters were: {'gamma': 0.003, 'C': 0.7}
Best Metric on Dev was: 1.0
Best hyperparameters were: {'max_depth': 20}
Best Metric on Dev was: 0.8729281767955801
Best hyperparameters were: {'gamma': 0.0007, 'C': 2}
Best Metric on Dev was: 0.994475138121547
Best hyperparameters were: {'max_depth': 20}
Best Metric on Dev was: 0.8729281767955801
{'svm': [{0.9888268156424581}, {0.994413407821229}, {0.9776536312849162}, {0.9832402234636871}, {0.994413407821229}], 'decision_tree': [{0.8435754189944135}, {0.88268156424581}, {0.8156424581005587}, {0.888268156424581}, {0.8547486033519553}]}
```