Hyper parameter tuning and cross validation for Support vector machine:

Learning the parameters of a prediction function and testing it on the same data is a methodological mistake: a model that would just repeat the labels of the samples that it has just seen would have a perfect score but would fail to predict anything useful on yet-unseen data. To avoid over-fitting, it is common practice to use cross validation when performing a (supervised) machine learning experiment to hold out part of the available data as a test set.

It is possible and recommended to search the hyper-parameter space for the best cross validation score. Any parameter provided when constructing an estimator may be optimized in this manner.
A search consists of:

- an estimator
- a parameter space;
- a method for searching or sampling candidates;
- a cross-validation scheme; and
- a score function

we used GridSearchCV which exhaustively considers all parameter combinations

We determined The best parameters that can be determined by grid search techniques to search the hyper-parameter space for the best cross validation score. the below tuned parameters is used in our GridSearchCV:

```
tuned_parameters = [{'kernel': ['rbf'], 'gamma': [1e-3, 1e-4],
                     'C': [1, 10, 100, 1000]},
                    {'kernel': ['linear'], 'C': [1, 10, 100, 1000]}]
```

the below data show the grid search for each on of the SVM for the 13 features we have:

grid search for First SVM:

Best parameters set found on development set:

accuracy:25.6%

{'C': 100, 'kernel': 'linear'}

Grid scores on development set:

0.128 (+/-0.145) for {'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}

0.205 (+/-0.192) for {'C': 1, 'gamma': 0.0001, 'kernel': 'rbf'}

0.128 (+/-0.192) for {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}

0.154 (+/-0.126) for {'C': 10, 'gamma': 0.0001, 'kernel': 'rbf'}

0.128 (+/-0.192) for {'C': 100, 'gamma': 0.001, 'kernel': 'rbf'}

0.128 (+/-0.073) for {'C': 100, 'gamma': 0.0001, 'kernel': 'rbf'}

0.103 (+/-0.145) for {'C': 1000, 'gamma': 0.001, 'kernel': 'rbf'}

0.103 (+/-0.145) for {'C': 1000, 'gamma': 0.0001, 'kernel': 'rbf'}

0.179 (+/-0.192) for {'C': 1, 'kernel': 'linear'}

0.205 (+/-0.192) for {'C': 10, 'kernel': 'linear'}

0.256 (+/-0.192) for {'C': 100, 'kernel': 'linear'}

0.179 (+/-0.073) for {'C': 1000, 'kernel': 'linear'}

grid search for the second SVM:

Best parameters set found on development set:

{'C': 10, 'kernel': 'linear'}
accuracy:23.1%

Grid scores on development set:

0.077 (+/-0.218) for {'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}

0.077 (+/-0.218) for {'C': 1, 'gamma': 0.0001, 'kernel': 'rbf'}

0.103 (+/-0.192) for {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}

0.077 (+/-0.218) for {'C': 10, 'gamma': 0.0001, 'kernel': 'rbf'}

0.103 (+/-0.192) for {'C': 100, 'gamma': 0.001, 'kernel': 'rbf'}

0.103 (+/-0.192) for {'C': 100, 'gamma': 0.0001, 'kernel': 'rbf'}

0.103 (+/-0.192) for {'C': 1000, 'gamma': 0.001, 'kernel': 'rbf'}

0.103 (+/-0.192) for {'C': 1000, 'gamma': 0.0001, 'kernel': 'rbf'}

0.103 (+/-0.073) for {'C': 1, 'kernel': 'linear'}

0.231 (+/-0.218) for {'C': 10, 'kernel': 'linear'}

0.179 (+/-0.073) for {'C': 100, 'kernel': 'linear'}

0.205 (+/-0.073) for {'C': 1000, 'kernel': 'linear'}

grid search for third SVM:
Best parameters set found on development set:

{'C': 100, 'kernel': 'linear'}
accuracy:48.7%

Grid scores on development set:

0.308 (+/-0.126) for {'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}

0.282 (+/-0.192) for {'C': 1, 'gamma': 0.0001, 'kernel': 'rbf'}

0.333 (+/-0.192) for {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}

0.333 (+/-0.145) for {'C': 10, 'gamma': 0.0001, 'kernel': 'rbf'}

0.333 (+/-0.192) for {'C': 100, 'gamma': 0.001, 'kernel': 'rbf'}

0.333 (+/-0.145) for {'C': 100, 'gamma': 0.0001, 'kernel': 'rbf'}

0.333 (+/-0.192) for {'C': 1000, 'gamma': 0.001, 'kernel': 'rbf'}

0.333 (+/-0.145) for {'C': 1000, 'gamma': 0.0001, 'kernel': 'rbf'}

0.282 (+/-0.192) for {'C': 1, 'kernel': 'linear'}

0.436 (+/-0.316) for {'C': 10, 'kernel': 'linear'}

0.487 (+/-0.363) for {'C': 100, 'kernel': 'linear'}

0.462 (+/-0.332) for {'C': 1000, 'kernel': 'linear'}

grid search for fourth SVM:
Best parameters set found on development set:

{'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}
accuracy=20.5%

Grid scores on development set:

0.205 (+/-0.073) for {'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}

0.205 (+/-0.073) for {'C': 1, 'gamma': 0.0001, 'kernel': 'rbf'}

0.179 (+/-0.145) for {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}

0.205 (+/-0.073) for {'C': 10, 'gamma': 0.0001, 'kernel': 'rbf'}

0.179 (+/-0.145) for {'C': 100, 'gamma': 0.001, 'kernel': 'rbf'}

0.205 (+/-0.073) for {'C': 100, 'gamma': 0.0001, 'kernel': 'rbf'}

0.179 (+/-0.145) for {'C': 1000, 'gamma': 0.001, 'kernel': 'rbf'}

0.179 (+/-0.145) for {'C': 1000, 'gamma': 0.0001, 'kernel': 'rbf'}

0.128 (+/-0.073) for {'C': 1, 'kernel': 'linear'}

0.154 (+/-0.126) for {'C': 10, 'kernel': 'linear'}

0.103 (+/-0.073) for {'C': 100, 'kernel': 'linear'}

0.103 (+/-0.073) for {'C': 1000, 'kernel': 'linear'}

grid search for fifth SVM:

Best parameters set found on development set:

{'C': 1000, 'kernel': 'linear'}
Accuracy:28.2%

Grid scores on development set:

0.231 (+/-0.000) for {'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}

0.179 (+/-0.073) for {'C': 1, 'gamma': 0.0001, 'kernel': 'rbf'}

0.231 (+/-0.000) for {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}

0.179 (+/-0.073) for {'C': 10, 'gamma': 0.0001, 'kernel': 'rbf'}

0.256 (+/-0.073) for {'C': 100, 'gamma': 0.001, 'kernel': 'rbf'}

0.231 (+/-0.000) for {'C': 100, 'gamma': 0.0001, 'kernel': 'rbf'}

0.256 (+/-0.073) for {'C': 1000, 'gamma': 0.001, 'kernel': 'rbf'}

0.231 (+/-0.000) for {'C': 1000, 'gamma': 0.0001, 'kernel': 'rbf'}

0.256 (+/-0.145) for {'C': 1, 'kernel': 'linear'}

0.256 (+/-0.073) for {'C': 10, 'kernel': 'linear'}

0.256 (+/-0.073) for {'C': 100, 'kernel': 'linear'}

0.282 (+/-0.073) for {'C': 1000, 'kernel': 'linear'}

grid search for sixth SVM:
Best parameters set found on development set:

{'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}

Grid scores on development set:

0.282 (+/-0.261) for {'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}

0.282 (+/-0.145) for {'C': 1, 'gamma': 0.0001, 'kernel': 'rbf'}

0.282 (+/-0.261) for {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}

0.282 (+/-0.145) for {'C': 10, 'gamma': 0.0001, 'kernel': 'rbf'}

0.282 (+/-0.261) for {'C': 100, 'gamma': 0.001, 'kernel': 'rbf'}

0.282 (+/-0.145) for {'C': 100, 'gamma': 0.0001, 'kernel': 'rbf'}

0.282 (+/-0.261) for {'C': 1000, 'gamma': 0.001, 'kernel': 'rbf'}

0.282 (+/-0.145) for {'C': 1000, 'gamma': 0.0001, 'kernel': 'rbf'}

0.282 (+/-0.073) for {'C': 1, 'kernel': 'linear'}

0.231 (+/-0.126) for {'C': 10, 'kernel': 'linear'}

0.231 (+/-0.126) for {'C': 100, 'kernel': 'linear'}

0.231 (+/-0.126) for {'C': 1000, 'kernel': 'linear'}

Detailed classification report:

The model is trained on the full development set.
The scores are computed on the full evaluation set.

grid search for seventh SVM:

Best parameters set found on development set:

{'C': 10, 'kernel': 'linear'}
accuracy:33.3%

Grid scores on development set:

0.205 (+/-0.073) for {'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}

0.205 (+/-0.073) for {'C': 1, 'gamma': 0.0001, 'kernel': 'rbf'}

0.205 (+/-0.073) for {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}

0.205 (+/-0.073) for {'C': 10, 'gamma': 0.0001, 'kernel': 'rbf'}

0.205 (+/-0.073) for {'C': 100, 'gamma': 0.001, 'kernel': 'rbf'}

0.205 (+/-0.073) for {'C': 100, 'gamma': 0.0001, 'kernel': 'rbf'}

0.205 (+/-0.073) for {'C': 1000, 'gamma': 0.001, 'kernel': 'rbf'}

0.205 (+/-0.073) for {'C': 1000, 'gamma': 0.0001, 'kernel': 'rbf'}

0.308 (+/-0.218) for {'C': 1, 'kernel': 'linear'}

0.333 (+/-0.261) for {'C': 10, 'kernel': 'linear'}

0.333 (+/-0.316) for {'C': 100, 'kernel': 'linear'}

0.333 (+/-0.316) for {'C': 1000, 'kernel': 'linear'}

grid search for eighth SVM:

```
Best parameters set found on development set:

{'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}
accuracy:20.5%

Grid scores on development set:

0.179 (+/-0.073) for {'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}
0.179 (+/-0.073) for {'C': 1, 'gamma': 0.0001, 'kernel': 'rbf'}
0.205 (+/-0.073) for {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}
0.205 (+/-0.073) for {'C': 10, 'gamma': 0.0001, 'kernel': 'rbf'}
0.205 (+/-0.073) for {'C': 100, 'gamma': 0.001, 'kernel': 'rbf'}
0.205 (+/-0.073) for {'C': 100, 'gamma': 0.0001, 'kernel': 'rbf'}
0.205 (+/-0.073) for {'C': 1000, 'gamma': 0.001, 'kernel': 'rbf'}
0.205 (+/-0.073) for {'C': 1000, 'gamma': 0.0001, 'kernel': 'rbf'}
0.103 (+/-0.192) for {'C': 1, 'kernel': 'linear'}
0.103 (+/-0.073) for {'C': 10, 'kernel': 'linear'}
0.154 (+/-0.126) for {'C': 100, 'kernel': 'linear'}
0.205 (+/-0.073) for {'C': 1000, 'kernel': 'linear'}


grid search for ninth SVM:
Best parameters set found on development set:

{'C': 100, 'kernel': 'linear'}
accuracy:28.2%

Grid scores on development set:

0.205 (+/-0.073) for {'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}
0.154 (+/-0.126) for {'C': 1, 'gamma': 0.0001, 'kernel': 'rbf'}
0.205 (+/-0.073) for {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}
0.205 (+/-0.073) for {'C': 10, 'gamma': 0.0001, 'kernel': 'rbf'}
0.205 (+/-0.073) for {'C': 100, 'gamma': 0.001, 'kernel': 'rbf'}
0.179 (+/-0.073) for {'C': 100, 'gamma': 0.0001, 'kernel': 'rbf'}
0.205 (+/-0.073) for {'C': 1000, 'gamma': 0.001, 'kernel': 'rbf'}
0.179 (+/-0.073) for {'C': 1000, 'gamma': 0.0001, 'kernel': 'rbf'}
0.231 (+/-0.126) for {'C': 1, 'kernel': 'linear'}
0.231 (+/-0.126) for {'C': 10, 'kernel': 'linear'}
0.282 (+/-0.145) for {'C': 100, 'kernel': 'linear'}
0.256 (+/-0.192) for {'C': 1000, 'kernel': 'linear'}


grid search for tenth SVM:
Best parameters set found on development set:

{'C': 100, 'kernel': 'linear'}
accuracy:41%

Grid scores on development set:

0.282 (+/-0.192) for {'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}
0.282 (+/-0.192) for {'C': 1, 'gamma': 0.0001, 'kernel': 'rbf'}
0.282 (+/-0.192) for {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}
0.282 (+/-0.290) for {'C': 10, 'gamma': 0.0001, 'kernel': 'rbf'}
```

```
0.282 (+/-0.192) for {'C': 100, 'gamma': 0.001, 'kernel': 'rbf'}
0.308 (+/-0.218) for {'C': 100, 'gamma': 0.0001, 'kernel': 'rbf'}
0.282 (+/-0.192) for {'C': 1000, 'gamma': 0.001, 'kernel': 'rbf'}
0.308 (+/-0.218) for {'C': 1000, 'gamma': 0.0001, 'kernel': 'rbf'}
0.333 (+/-0.261) for {'C': 1, 'kernel': 'linear'}
0.385 (+/-0.332) for {'C': 10, 'kernel': 'linear'}
0.410 (+/-0.261) for {'C': 100, 'kernel': 'linear'}
0.410 (+/-0.261) for {'C': 1000, 'kernel': 'linear'}


Best parameters set found on development set:

{'C': 100, 'kernel': 'linear'}
accuracy:35.9%

Grid scores on development set:

0.282 (+/-0.145) for {'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}
0.282 (+/-0.192) for {'C': 1, 'gamma': 0.0001, 'kernel': 'rbf'}
0.282 (+/-0.145) for {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}
0.308 (+/-0.126) for {'C': 10, 'gamma': 0.0001, 'kernel': 'rbf'}
0.282 (+/-0.145) for {'C': 100, 'gamma': 0.001, 'kernel': 'rbf'}
0.282 (+/-0.073) for {'C': 100, 'gamma': 0.0001, 'kernel': 'rbf'}
0.282 (+/-0.145) for {'C': 1000, 'gamma': 0.001, 'kernel': 'rbf'}
0.282 (+/-0.073) for {'C': 1000, 'gamma': 0.0001, 'kernel': 'rbf'}
0.308 (+/-0.251) for {'C': 1, 'kernel': 'linear'}
0.333 (+/-0.261) for {'C': 10, 'kernel': 'linear'}
0.359 (+/-0.261) for {'C': 100, 'kernel': 'linear'}
0.308 (+/-0.126) for {'C': 1000, 'kernel': 'linear'}

Detailed classification report:

grid search for 11th SVM:

Best parameters set found on development set:

{'C': 1, 'gamma': 0.0001, 'kernel': 'rbf'}
accuracy:30.8%

Grid scores on development set:

0.179 (+/-0.261) for {'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}
0.308 (+/-0.218) for {'C': 1, 'gamma': 0.0001, 'kernel': 'rbf'}
0.231 (+/-0.251) for {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}
0.256 (+/-0.261) for {'C': 10, 'gamma': 0.0001, 'kernel': 'rbf'}
0.231 (+/-0.251) for {'C': 100, 'gamma': 0.001, 'kernel': 'rbf'}
0.256 (+/-0.261) for {'C': 100, 'gamma': 0.0001, 'kernel': 'rbf'}
0.231 (+/-0.251) for {'C': 1000, 'gamma': 0.001, 'kernel': 'rbf'}
0.256 (+/-0.261) for {'C': 1000, 'gamma': 0.0001, 'kernel': 'rbf'}
0.231 (+/-0.126) for {'C': 1, 'kernel': 'linear'}
0.282 (+/-0.073) for {'C': 10, 'kernel': 'linear'}
0.231 (+/-0.218) for {'C': 100, 'kernel': 'linear'}
0.231 (+/-0.332) for {'C': 1000, 'kernel': 'linear'}


grid search for 11th SVM:
Best parameters set found on development set:
```

{'C': 100, 'kernel': 'linear'}
accuracy:23.1%
Grid scores on development set:

0.154 (+/-0.218) for {'C': 1, 'gamma': 0.001, 'kernel': 'rbf'}
0.179 (+/-0.192) for {'C': 1, 'gamma': 0.0001, 'kernel': 'rbf'}
0.128 (+/-0.145) for {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}
0.154 (+/-0.218) for {'C': 10, 'gamma': 0.0001, 'kernel': 'rbf'}
0.128 (+/-0.145) for {'C': 100, 'gamma': 0.001, 'kernel': 'rbf'}
0.128 (+/-0.145) for {'C': 100, 'gamma': 0.0001, 'kernel': 'rbf'}
0.128 (+/-0.145) for {'C': 1000, 'gamma': 0.001, 'kernel': 'rbf'}
0.128 (+/-0.145) for {'C': 1000, 'gamma': 0.0001, 'kernel': 'rbf'}
0.179 (+/-0.192) for {'C': 1, 'kernel': 'linear'}
0.205 (+/-0.192) for {'C': 10, 'kernel': 'linear'}
0.231 (+/-0.218) for {'C': 100, 'kernel': 'linear'}
0.154 (+/-0.126) for {'C': 1000, 'kernel': 'linear'}

After the training part, we predict the classes for each 13 features of a human face, then get the majority voting and choose the 2 common class predicted by our SVM classifier.