Portfolio Week 3

COS40007 - Artificial Intelligence for Engineering
Studio 1-1 (12:30-2:30)

Hanh Nguyen Nguyen
103532674

Studio 3

Activity 6

Table 1. Summary table containing accuracy value of developed SVM models

SVM	Train-test split	Cross-validation
Original Features	89.28%	88.65%
With hyperparameter* tuning	83.89%	83.47%
With feature selection and hyperparameter tuning	84.95%	84.49%
With PCA and hyperparameter tuning	83.92%	83.52%

**obtained by using GridSearch: C = 10, gamma = 0.0001, kernel = 'rbf'.*

```
Accuracy of Splitting Train:  0.8928059615935798
Accuracy of 10 fold cross validation:  [0.91659501 0.84866724 0.92863285 0.92347377 0.92863285 0.88822012
 0.81513328 0.8383491  0.86070507 0.91652324]
```

Figure 1. Before tuning

```
Accuracy of Splitting Train:  0.8389223273144167
Accuracy of 10 fold cross validation:  [0.83404987 0.83404987 0.83490972 0.83319003 0.83319003 0.83319003
 0.83319003 0.83576956 0.83662941 0.83907057]
Accuracy of train test split validation with best SVM features:  0.8495270851246776
Accuracy of 10 fold cross validation with best SVM features:  [0.84522786 0.84522786 0.85210662 0.84264832 0.84436801 0.8460877
 0.83404987 0.83576956 0.84092863 0.86230637]
Accuracy of train test split validation with principle components:  0.8392089423903697
Accuracy of 10 fold cross validation with principal components:  [0.83404987 0.83490972 0.83576956 0.83319003 0.83319003 0.83404987
 0.83319003 0.83662941 0.83748925 0.83993115]
```

Figure 2. Accuracy values after hyperparameter tuning with different models

Activity 7

Table 2. Summary table containing accuracy value of different developed models

Model	Train-test split	Cross-validation
SVM	88.42%	89.20%
SGD	87.90%	88.31%
RandomForest	92.06%	92.55%
MLP	86.36%	84.83%

Cross-validation values = mean 10 fold cross validation values calculated using Excel

[Portfolio 3 + Studio 3 Drive]

```
SVM train test split accuracy:  0.8842075093149899
SVM 10 fold cross accuracy:  [0.89423904 0.89251935 0.88822012 0.88993981 0.89337919 0.8916595
 0.89423904 0.88134136 0.89251935 0.90189329]
SGD train test split accuracy:  0.879048437947836
SGD 10 fold cross accuracy:  [0.87618229 0.88993981 0.88650043 0.88650043 0.87790198 0.87962167
 0.87618229 0.88564058 0.87102322 0.9010327 ]
RandomForest train test split accuracy:  0.9206076239610204
RandomForest 10 fold cross accuracy:  [0.92691316 0.92519347 0.92261393 0.92175408 0.92433362 0.92347377
 0.92863285 0.91401548 0.92949269 0.93459552]
MLP train test split accuracy:  0.8635712238463743
MLP train test split accuracy:  [0.89939811 0.78675838 0.84264832 0.85726569 0.88993981 0.89509888
 0.86586414 0.83662941 0.89423904 0.7151463 ]
```

Figure 3. Screenshot different models' accuracy in Python terminal

Additional resource

- Source Code
- Data

Portfolio 3

Step 1: Data collection

- Source Code
- Data

Step 2: Create Composite columns

- Source Code
- Data

Step 3: Data pre-processing

- Source Code
- Data

Step 4: Training

Table 3. Summary table of accuracy values of developed SVM models for slicing and boning meat

SVM	Train-test split	Cross-validation
Original Features	77.56%	75.60%
With hyperparameter tuning**	77.56%	75.19%
With feature selection and hyperparameter tuning	83.38%	80.76%
With PCA and hyperparameter tuning	77.56%	75.19%

***obtained by using GridSearch: C = 0.1, gamma = 1, kernel = 'rbf'.*

```
Accuracy of Splitting Train:  0.775623268698061
Accuracy of 10 fold cross validation:  [0.75206612 0.76666667 0.775      0.76666667 0.73333333 0.725
 0.75833333 0.76666667 0.76666667 0.75      ]
```

Figure 4. Before tuning SVM accuracy values

```
Accuracy of Splitting Train:  0.775623268698061
Accuracy of 10 fold cross validation:  [0.75206612 0.75833333 0.75833333 0.75       0.75       0.75
 0.75       0.75       0.75       0.75       ]
Accuracy of train test split validation with best SVM features:  0.8337950138504155
Accuracy of 10 fold cross validation:  [0.84297521 0.83333333 0.85833333 0.85       0.7        0.74166667
 0.76666667 0.81666667 0.88333333 0.78333333]
Accuracy of train test split validation with principle components:  0.775623268698061
Accuracy of 10 fold cross validation with principal components:  [0.75206612 0.75833333 0.75833333 0.75       0.75       0.75
 0.75       0.75       0.75       0.75       ]
PS C:\Users\PC\OneDrive - Swinburne University\Desktop\COS40007>
```

Figure 5. Accuracy values after hyperparameter tuning with different models

Portfolio 3 + Studio 3 Drive

Table 4. Summary table of accuracy values of different developed models for slicing and boning meat

Model	Train-test split	Cross-validation
SVM	74.62%	49.80%
SGD	75.43%	65.66%
RandomForest	88.99%	51.81%
MLP	76.94%	73.59%

Cross-validation values = mean 10 fold cross validation values calculated using Excel

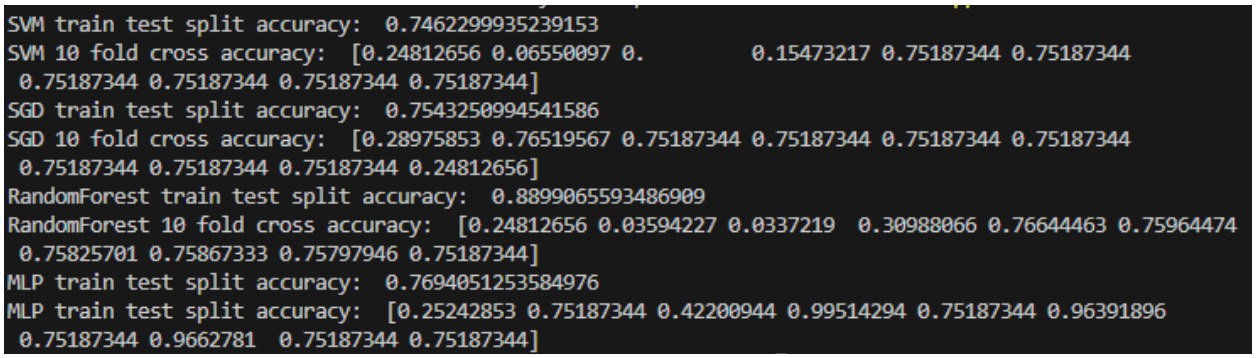

Figure 6. Screenshot different models' accuracy in Python terminal

Additional resource

- Source Code (SVM and Different models)
- Data (SVM and Different models)

Step 5: Model Selection

1) The SVM model best for my dataset of Right and Left hand motions would be the **With feature selection and hyperparameter tuning** model.
   This is because the model provides high accuracy and shows consistent performance in different settings of Train-test split and Cross-validation. With feature selection, it also removes less relevant and not as important features, making the model more reliable and providing more accurate results.

2) The ML model best for my dataset would be the **MLP** model.
   This is because even though MLP does not have the highest accuracy in the Train-test split, the accuracy this model provides is consistent and moderately high in both settings. Hence, Cross-validation is a more advanced tool in assessing machine learning models compared to the Train-test split. For a model to have decent accuracy in that test, it is worthy to utilise that model.