

COS40007- Artificial Intelligence for Engineering

Portfolio Assessment – 2

Name : Vivek Saini

Student ID: 103828056

Studio class: 1-3

In this weeks portfolio I have completed the Studio task 3 and have built the outlined SVM models, here is the summary of the accuracy result for activity 6 in Studio3 assignment:

SVM Model	Train-test split	Cross validation
Original features	89.77%	89.18%
With hyper parameter tuning	83.63%	83.58
With feature selection and hype parameter tuning	83.92%	83.78%
With PCA and hyper parameter tuning	83.69%	83.60%

Furthermore, I have also performed the train-test split and cross-validation on the original et of data as outlined in the assignment with different forms of classifiers. Here is the summary of the observations as required per the activity 7 in Studio-3 task:

Model	Train-Test Split	Cross Validation
SVM	89.56%	89.17%
SGD	80.05%	87.26%
RandomForest	92.61%	92.59%
MLP	89.40%	83.76%

In the week 3 portfolio assignments, I have gone step by step to follow all the tasks that have been given us and implementing them correctly gives us the following output. The following table will consist of the step 4 in the portfolio assignment.

Model	Train-Test Split	Cross Validation
Original Features	86.98%	92.00%
With Hyperparameter Tuning	76.59%	75.22%
With Feature Selection and Hyperparameter Tuning	76.59%	75.22%
With PCA and Hyperparameter Tuning	76.59%	75.22%

Finally, In the last step we make different ML models similar to the Studio task consisting of the comparisons between SVM, SGD, RandomForest and MLP classifier. Below are the accuracy results.

Model	Train-Test Split	Cross Validation
SVM	77.56%	75.19%
SGD	86.06%	85.53%
RandomForest	72.71%	73.94%
MLP	77.61%	75.44%

The SVM model with Grid Search appears to be the best choice because of its high accuracy, it provided high results amongst different configurations including when combined with feature selection and dimensionality reduction techniques. This model's ability to adapt to complex decision boundaries in the data, along with the tuned hyperparameters, makes it the most robust option for the classification task.

The **Support Vector Machine (SVM) with the RBF kernel and optimized hyperparameters** is the best model for this problem. It consistently achieved the highest accuracy across both the train-test split and cross-validation, making it the most reliable and effective choice for your dataset.

Appendix

Link to my Github - https://github.com/Vikksaini/AI_for_engineering_portfolio