

COS40007 Artificial Intelligence for Engineering

Week 4 Studio Activities

ILO	Understand Evaluation Measure of Machine Learning model.
Aim	<ul style="list-style-type: none"> Learn how to find different evaluation measure of ML models Learn how to compare ML models Learn how to build AI from ML
Resources	<p>Books:</p> <ol style="list-style-type: none"> Proise, Jeff. Applied machine learning and AI for engineers. " O'Reilly Media, Inc.", 2022. Raschka, Sebastian, Yuxi Hayden Liu, and Vahid Mirjalili. Machine Learning with PyTorch and Scikit-Learn: Develop machine learning and deep learning models with Python. Packt Publishing Ltd, 2022. <p>Web Resources:</p> <ol style="list-style-type: none"> https://scikit-learn.org/stable/modules/model_evaluation.html
Requirements for to be marked as complete	Demonstrate the table of outcome to your tutor

Disclaimer: The dataset used in this studio was originally collected for a funded research project by Australian Meat processor Corporation. The dataset here is used solely for educational purposes and can be only used for completing activities of this studio. By any mean this dataset is not shareable to others or any public domain.

Dataset: We will be using the same dataset that we used in Studio 3. If you have not completed Studio 3, please go back and complete Studio 3. Otherwise, download the dataset from Studio 3.

Evaluation metrics

Check this link: https://scikit-learn.org/stable/modules/model_evaluation.html

There are several evaluation measures that are used to compare different ML models. The evaluation normally depends on target problem. A common evaluation measure when you solve classification problem is accuracy, precision, recall and F1 score and confusion matrix the visualise the evaluation.

Studio Activity 1: Group formation for Design project

The overview and the list of topics for design project is now available on Canvas. If you have not reviewed the design project specification document already, please go read the document. This is available under Module> design project.

Now, you need to form a group of 5 to complete the project. If you already have a group in mind to form then this is great. Otherwise, if you need help for group formation, please take assistance from you tutor. Once you finalise your group, please submit the google form. The link of the form can be found in the design project specification document.

Studio Activity 2: Classification Report

Check this function https://scikit-learn.org/stable/modules/model_evaluation.html#classification-report

Now go back to your Studio 3 outcome. In Activity 2 you measured the performance of ML model for train-test split and cross validation using only accuracy. And in Activity 7 we measured the accuracy of different ML models using the original dataset. Now let's expand our evaluation outcome. We will conduct this evaluation exercise only using original data and only for train-test split (that is using `y_test` and `y_pred`).

Now, compute the evaluation outcome of 4 models (SVM, SGD, RandomForest and MLP) using classification report and answer the following questions

- 1) Which model has the best F1 score in terms of weighted average. Is this different than the accuracy that you obtained in Studio 3?
- 2) Which model has the best precision value
- 3) Which model has the best recall value
- 4) Which model has the best accuracy value
- 5) Which class has poor performance across all models? How you determined this?

Studio Activity 3: Confusion Matrix

A good way to understand your evaluation outcome is to visualise them in [confusion matrix](#). You will also get an idea of the class distribution of your test set. This also help you to understand true positives, false positives, true negatives and false negative classes.



Now plot the confusion matrix of 4 models. From your observation where are the most false positives coming from? How you can fix it?

Studio Activity 4: Class balancing: Undersampling and oversampling

This dataset has a class imbalance issue so model may overfit with the majority class. You can fix this issue using [undersampling and oversampling](#) . If you see the distribution of classes you will see we have majority of samples from class 2 and few samples of class 1. So, let's fix the issue in our dataset.

Now generate a resamples version of the dataset by

- 1) Oversampling the minority classes using SMOTE
- And
- 2) Undersampling Majority class using Tomeklinks

Now train and test your best classifier with the resampled version of the dataset. Computer the classification report and visualise confusion matrix. What you observed?

Repeat the above process but now adjusting class weights (e.g. 0.2 for class 2, 0.5 for class 1 and 0.3 for class 0)

Which approach resulted better evaluation performance?

Studio Activity 5: Saving ML model

This is very similar to Studio 3 activity. You need to create SVM model using data only from 3 workers (w1, w2 and w3) with the following criteria. And keep w4 data aside.

- 1) Merge w1,w2,w3 data
- 2) Use undersampling and oversampling to fix the class distribution
- 3) Use train-test split (70/30)
- 4) Generate classification report
- 5) Visualise confusing matrix

Now save the model that you genrated

Sample code

```
import pickle
clf = svm.SVC()
clf.fit(_X_train, y_train)
filename = 'mymodel1.pkl'
```



```
pickle.dump(model, open(filename, 'wb'))
```

Do the similar thing as above but now generate and save a RandomForest model (as mymodel2.pkl)

Studio Activity 6: AI from ML

In Activity 4 you got ML models which is trained using data from 3 workers. Now you want to build AI using the trained model for a new worker. So, the model has never seen data from the worker before because we have not used the data from train and validation. Say, we want to use the model in real life deployment settings. That is, assume you model is running in some machine and worker's activity data is coming in real-time every second. You want to use your model to get the data and find the activity that the worker is doing. This then becomes AI. Now we will use w4 data to test our first AI deployment.

First you need to load the model that you saved
You can do this using pickle

```
model = pickle.load(open(filename, 'rb'))
```

Now load the w4.csv data in a dataframe and separate feature (X) and class (Y)

Now write a for loop that

Initialise a counter starting from 0

- 1) Iteratively read 1 line from X and Y (say x_i, y_i for i-th iteration)
- 2) Apply model to predict class $y_{pred_i} = \text{model.predict}(x_i)$
- 3) If y_{pred_i} matches with y_i increment the counter.

So, at the end of the loop, the value in the counter will be the number of correct classifications you obtained for w4 data

If you divide the counter with the length of Y then you will basically find the accuracy for w4 data.

Now repeat the above the model2 (random forest) and measure the accuracy value again for model2.

Which model has given you better performance. Have you observed different result that you obtained in the model evaluation (using classification report) or different?

Studio Activity 7: Understand ML model construction

Sometimes you want to understand how your ML model is constructed, how your model reach to a decision point. This provides an explainability on your outcome prediction. A decision tree is a simple classifier to visualise such measure.

- 1) Using the data generated in Activity 3 after sampling get the data with 10 features using [mutual info classif — scikit-learn 1.5.1 documentation](#) .
- 2) Now build a DecisionTree classifier using 10 selected features and measure the accuracy (use train-test split)
- 3) Save your model
- 4) Load the model
- 5) Plot the tree using [plot tree](#)
- 6) Plot tree using [export text](#)

Recommended readings

Check the following links to visualise other ML models that you used before.

- 1) [Support Vector Machine](#)
- 2) [Stochastic Gradient Decent](#)
- 3) [Radom Forest](#)

Next Steps:

The assessment Task for Week 4 now can be attempted and submitted via Canvas.