

Classification

Classification

In this lab you will perform **two-class classification** using **logistic regression**. A classifier is a machine learning model that separates the **label** into categories or **classes**. In other words, classification models are **supervised** machine learning models which predict a categorical label.

The German Credit bank customer data is used to determine if a particular person is a good or bad credit risk. Thus, credit risk of the customer is the classes you must predict. In this case, the cost to the bank of issuing a loan to a bad risk customer is five times that of denying a loan to a good customer. This fact will become important when evaluating the performance of the model.

By the completion of this lab, you will:

- Prepare data for classification models using scikit-learn.
- Construct a classification model using scikit-learn.
- Evaluate the performance of the classification model.
- Use techniques such as reweighting the labels and changing the decision threshold to change the trade-off between false positive and false negative error rates.


Lab Steps

1. Make sure that you have completed the setup requirements as described in the Lab Overview section.
2. Now, run jupyter notebook and open the “Classification.ipynb” notebook under Module 4 folder.
3. Examine the notebook and answer the questions along the way.

Question 1

1.0/1.0 point (graded)

Knowing the class imbalances in the data, theoretically, what is the best accuracy you can get without creating any machine learning models?

- ☐ 0% - you cannot really predict the label without creating any machine learning models
- ☐ 50% - this is a binary classification problem, random guessing will translate to 50% accuracy
- ☒ 70% - this can be achieved by guessing all customers will have good credit

- ☐ 100% - because of the imbalances, all customers with bad credit will be easily identified

Submit

You have used 2 of 2 attempts

Question 2

1/1 point (graded)

What is the AUC of the model?

☐ 0.71

☐ 0.75

☐ 0.77

☒ 0.80



Submit

You have used 2 of 2 attempts

✓ Correct (1/1 point)

Question 3

1.0/1.0 point (graded)

What three metrics may change by changing the threshold?

☒ Accuracy

☒ Precision

☒ Recall

☐ AUC



Submit

You have used 2 of 2 attempts

[Learn About Verified Certificates](#)

