

Introduction to Machine Learning in Engineering Science

National Cheng Kung University

Department of Engineering Science

Instructor: Chi-Hua Yu

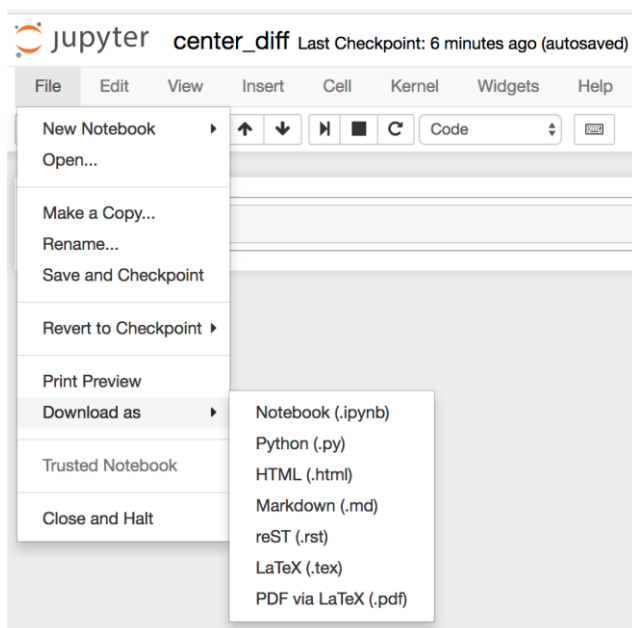
Lab 2

Programming, Due 21:00, Thursday, October 14th, 2021

Late submission before post of solution: score*0.8 (the solution will usually be posted within a week); no late submission after the post of solution

Lab Submission Procedure (請仔細閱讀)

1. You should submit your Jupyter notebook and Python script (*.py, in Jupyter, click File, Download as, Python (*.py)).



2. Name a folder using your student id and lab number (e.g., n96081494_lab1), put all the python scripts into the folder and zip the folder (e.g., n96081494_lab 1.zip).
3. Submit your lab directly through the course website.

Total 120%

1. (20%) Name your Jupyter notebook `gcd.ipynb` and Python script `my_gcd.py`. Write a Python program that finding a greatest common divisor by using Euclid's algorithm. Euclid's Algorithm states that the greatest common divisor of two integers m and n is n if n divides m evenly. However, if n does not divide m evenly, then the answer is the greatest common divisor of n and the remainder of m divided by n .

The interface of `gcd` function is:

Introduction to Machine Learning in Engineering Science

National Cheng Kung University

Department of Engineering Science

Instructor: Chi-Hua Yu

```
Def gcd(m: int,n: int)->int:  
    """ function of finding a greatest common divisor """
```

Below is the running example

```
[1]: from my_gcd import gcd
```

```
[2]: print(gcd(20,10))
```

```
10
```

```
[3]: ?gcd
```

```
Signature: gcd(m: int, n: int) -> int
```

```
Docstring: function of finding a greatest common divisor
```

```
File:      ~/Dropbox/NCKU/Lecture/MLES/2021/Code/Lab 2/my_gcd.py
```

```
Type:      function
```

2. (50%) Name your Jupyter notebook `fraction` and Python script `fraction.py`. A Fraction is user defined datatype to present fraction numbers. The operations for the Fraction type will allow a Fraction data object to behave like any other numeric value. We need to be able to add, subtract, multiply, and divide fractions. We also want to be able to show fractions using the standard “slash” form, for example 3/5. In addition, all fraction methods should return results in their lowest terms so that no matter what computation is performed, we always end up with the most common form.

The interface for Fraction should look like:

```
from my_gcd import gcd  
class Fraction(object):  
    """  
        User-defined object to represent numeric fractions  
        The top value, known as the numerator, can be any integer.  
        The bottom value, called the denominator, can be any integer  
        greater than 0  
    """
```

Below is the running example:

Introduction to Machine Learning in Engineering Science

National Cheng Kung University

Department of Engineering Science

Instructor: Chi-Hua Yu

```
[1]: from fraction import Fraction
```

```
[2]: x = Fraction(1,2)
     y = Fraction(2,3)
     print(x+y)
     print(x == y)

7/6
False
```

```
[3]: print(x!=y)

True
```

```
[4]: print(x-y)

-1/6
```

```
[5]: print(x*y)

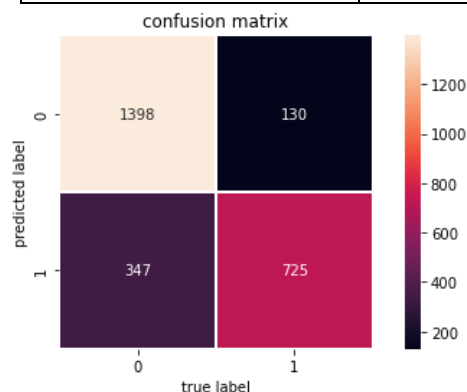
1/3
```

3. (50%) Bank of LAiMM wants to predict who are possible defaulters for the consumer loans product. They have data about historic customer behavior based on what they have observed. Hence when they acquire new customers, they want to predict who is riskier and who is not. You are required to use the training dataset to identify patterns that predict “potential” defaulters. Use Support Vector Machine to identify possible defaulters. There are some hyperparameters that can be adjusted in Support Vector Machine. These hyperparameters affect the results of training.

(a) Use GridSearchCV to search for best hyperparameters and select the parameters and set range of values to be searched. Please report training history, parameters and range of values, confusion matrix and accuracy. Please note that the number of possible combinations of parameters affects the search time, the more combinations will consume more time, especially when using GridSearchCV.

Example:

Parameters	Range
C	[0.2 0.4 ... 1.0]
Break_ties	[True, False]
⋮	⋮



accuracy score: 0.8165

Introduction to Machine Learning in Engineering Science

National Cheng Kung University

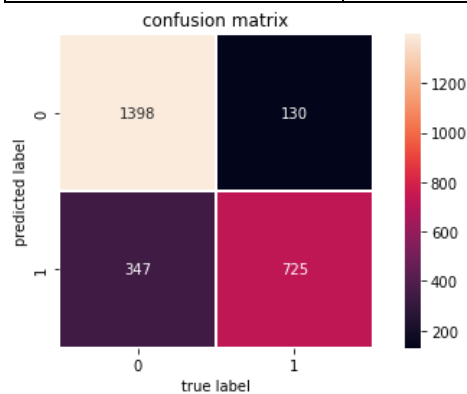
Department of Engineering Science

Instructor: Chi-Hua Yu

(b) RandomizedSearchCV to search for best hyperparameters and select the parameters and set range of values to be searched. Please report training history, parameters and range of values, confusion matrix and accuracy.

Example:

Parameters	Range
C	[0.2, 1.0]
Break_ties	[True, False]
⋮	⋮



accuracy score: 0.8165