Advanced School in Artificial Intelligence

Introduction to Python

Ing. Zese Riccardo

riccardo.zese@unife.it

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Regione Emilia-Romagna

FERRARIAE UNIVERSITAS · 13 91 · EX LABORE FRUCTUS

Università degli Studi di Ferrara

Outline

- Introduction to Python
- Introduction to Neural Networks
- Convolutional NN
- Recurrent NN
- Autoencoders and self supervised learning

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Outline

- Introduction to Python
- Introduction to Neural Networks
- Convolutional NN
- Recurrent NN
- Autoencoders and self supervised learning

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Modules

- Modules allow to define a group of functions and variables that are related.
- The name of the module is the name of the file without **.py**
- To use a module, one needs to **import** it.
  - NOTE: Importing a module causes python to run each line of code in the module, i.e., if there are operations outside functions they are executed.
- To call a function in a module: **module_name.function_name()**

```
import my_module
my_module.function_name()
```

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Modules

- Each module has an implicit variable `__name__`.
- If we import a module called `module_m`, then

```
module_m.__name__ == "module_m"
```

- But if we run a module, then

```
__name__ == "__main__"
```

- Recall that if we are running a module, we don't need the module name as a prefix.
- By checking if the module is the running module we can define a sort of function main per module that will be performed only if the module is the running one.

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Modules

a.py

```
import b

def f(x):
    return x * 2

if __name__ == "__main__":
    print(1)
    print(b.f(10))
    print(f(10))
```

b.py

```
import a

def f(x):
    return x / 2

if __name__ == "__main__":
    print(1)
    print(a.f(10))
    print(f(10))
```

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Modules

```
from module_name import fn_name1, fn_name2
```

- Import only functions **`fn_name1`** and **`fn_name2`**
- The imported functions can be called using their name
- Can also use `*` as a wildcard to import all the functions.

```
from module_name import *
```

- If two modules have a function with the same name, the most recent one stays.

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Modules

a.py

```
import b


def f(x):
    return x * 2


if __name__ == "__main__":
    print(1)
    print(b.f(10))
    print(f(10))
```

b.py

```
from a import f


def f(x):
    return x / 2


if __name__ == "__main__":
    print(1)
    print(a.f(10))
    print(f(10))
```

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

More on modules

- Importing a module means creating a variable having the name of the module and containing the reference to the module itself.
- The import from can be simulated also by assigning the reference of the function to a variable.

```
func=a.f
```

```
func(10)
```

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

More on modules

- Importing a module means creating a variable having the name of the module and containing the reference to the module itself.
- We can use `as` to rename this variable.

```
import a as module_a
module_a.f(10)
```

```
from a import f as func
func(10)
```

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Functions

- The definition of a function is

```
def function_name(arguments):
    # statements
```

- def is a python keyword
- Arguments are 0 or more variables.
- The body of the function (statements) must be indented.
- If the block contains the keyword **return**, it returns a value; otherwise it returns the special value **None**.

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

NumPy e other useful modules

- sys
- NumPy
- Scikit-learn
- Managing CSVs (csv and Pandas)

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Module sys

- It is a built-in module giving functions and parameters useful to interact with the operating system.
- It is usually used to get command arguments from CLI

```
import sys
print("Script name: " + sys.argv[0])
for i in range(1,len(sys.argv)):
  print("Arg " + str(i) +":" + sys.argv[i])
```

argv[0] is the name of the file

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Module sys

- It can be used to redirect standard IO

```
import sys
save_stdout = sys.stdout # keep previous streams
save_stderr = sys.stderr
file1 = open("output.log","w")
file2 = open("error.log","w")
sys.stdout = file1  # redirection
sys.stderr = file2
print("This is the output message\n") # print on file1
sys.stderr.write("This is the error message\n") # print on file2
sys.stdout = save_stdout # redirection to original streams
sys.stderr = save_stderr
file1.close()
file2.close()
```

`sys.stderr.write()` and `sys.stdout.write()` to write on standard output and error

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Module sys

- Interesting parameters:
  - **`sys.executable`** path of Pyhton interpreter
  - **`sys.maxint max`** value for integers
  - **`sys.modules`** dictionary containing all modules loaded by the interpreter
  - **`sys.path`** list of folders where Python looks for modules
  - **`sys.platform`** name of the SO
  - …

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Module NumPy

- NumPy is the fundamental package for scientific computing with Python. It contains among other things:
  - a powerful N-dimensional array object
  - sophisticated (broadcasting) functions
  - tools for integrating C/C++ and Fortran code
  - useful linear algebra, Fourier transform, and random number capabilities
  - Part of the SciPy framework (https://www.scipy.org/)
  - Do you remember arrays? We will use NumPy especially for that.
- See https://www.numpy.org/ for documentation and tutorials. We will explain what we need step by step.

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Module Scikit-learn

- Machine Learning in Python
  - Simple and efficient tools for data mining and data analysis
  - Accessible to everybody, and reusable in various contexts
  - Built on **NumPy**, SciPy, and matplotlib
  - Open source, commercially usable - BSD license
- Allows performing classification, regression, clustering, …
- See https://scikit-learn.org/stable/ for documentation and examples.

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Module for working with csv

- CSV files are often used to store data in machine learning.
- To read and write CSV we can use the module `pandas`, a module from the SciPy framework.

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Reading a CSV with `pandas`

- Pandas can read a CSV and save it into its **`DataFrame`**

```
import pandas
df = pandas.read_csv('my.csv')
print(df)
```

- **`DataFrame`**s have many attributes (for example containing the types of the columns) and methods.
- As NumPy and SciPy, Pandas uses its own types that are automatically translated into built-in Python types but that bring more power.

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Reading a CSV with `pandas`

- Pandas automatically uses the first line as header of the table and represents it as a matrix where the column index is the name of the column while the row index is an integer from 0.
  - If we need to use a different indexing for rows, using the values contained in a certain column (for example an ID) we can set this during reading

```
import pandas
df = pandas.read_csv('my.csv', index_col='ID')
print(df)
```

Name of the column

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Reading a CSV with `pandas`

- Pandas is usually able to automatically convert special types. In case it isn't (for example with dates), we can force the conversion.

```
import pandas
df = pandas.read_csv('my.csv',
                     parse_dates=['Date Column'])
print(df)
```

- Pandas will use the **Timestamp** datatype

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Reading a CSV with **pandas**

- We can also read CSV managing the header row and passing the name of the columns as argument

```
import pandas
df = pandas.read_csv('my.csv',
                     header=0,
                names=['My', 'Column','Names'])
print(df)
```

- **header** argument tells the row where the headers are. Used in combination with **names** we can replace headers of the CSV. If the CSV doesn't contain headers we can use **names** alone to define the headers (or assigning **None** to **header**).

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Writing a CSV with `pandas`

- Writing with pandas can be done easily by instantiating a DataFrame and write it into a file

```
import pandas
d = {'col1': [1, 2], 'col2': [3, 4]}
df = pd.DataFrame(data=d)
df.to_csv('my.csv')
```

	col1	col2
0	1	3
1	2	4

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Exercise 1

- Define the following variables: n1 = 1, n2 = 2, s1 = 'hi', s2 = 'hello'
- Use the if statement to test whether the type assigned to n1 and n2 is the same. If this is true, print the values of the variables and their relation for example in this way: «1 is greater than 2»
- Use the if statement to test whether s1 and s2 are of the same type. If this is true, check if s1 is a substring of s2 (if true print the result), else if s2 is contained in s1. Otherwise print that s1 and s2 are different.

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Exercise 2

1. Create a list of integers containing numbers from 2 to 7
2. Loop on the list to sum the values contained, print this value (don't use function sum())
3. Print the content of the list
4. Loop on the list to change its values by summing every element with the sum computed before
5. Print the content of the list

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Exercise 3

1. As exercise 2 but use arrays.
2. Create an array of integers containing numbers from 2 to 7
3. Loop on the array to sum the values contained, pint this value (don't use function sum())
4. Print the content of the array
5. Change the elements of the array's values by summing every element with the sum computed before
6. Print the content of the array
7. Try with the different definitions of array

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Exercise 4

- Sometimes we want to figure out what the key corresponding to a given value is.
  - This is impossible to do naively.
- We need to build a dictionary were values are used as keys.
- Problems?

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Exercise 4

- Sometimes we want to figure out what the key corresponding to a given value is.
  - This is impossible to do naively.
- We need to build a dictionary were values are used as keys.
- Problems?
  - While the keys in a dictionary must be unique, the values don't have this restriction.
  - So multiple keys can have the same value.
- So?

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Exercise 4

- Dictionary can take everything as value (mutable and immutable objects – remember keys are immutable)
- We can use for example a list.
- Try it by creating a dictionary containing the following pairs

```
'Boss Nass':'Star Wars'
'Tom Bombadil':'The Lord of the Rings'
'Hari Seldon':'Foundation series'
'Polliver':'Game of Thrones'
'Jules Winnfield':'Pulp Fiction'
'The Mule':'Foundation series'
'Flynn Rider':'Rapunzel'
'Yoda':'Star Wars'
'Vince Vega':'Pulp Fiction'
```

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Exercise 4

- Print the dictionary created.
- Invert the dictionary.
- Print the new dictionary.

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021

Exercise 5

1. Create two Python modules called respectively **`module_one`** and **`module_two`**.
2. **`module_one`** defines the functions:
   1. **`f`** takes one or two arguments x and y, with y having the default value 10 and returns their sum.
   2. **`g`** takes one argument x and computes x^3 (use the exponentiation.
   3. Define a main that calls **`f(2)`** and **`f(g(2))`** and prints their results.
3. **`module_two`** imports only function **`f`** from **`module_one`** and calls **`f(10,5)`** printing its result.

Progetto di alta formazione in ambito tecnologico economico e culturale per una regione della conoscenza europea e attrattiva approvato e cofinanziato dalla Regione Emilia-Romagna con deliberazione di Giunta regionale n. 1625/2021