SPAI

Introduction to Jupyter Notebooks & Google Colab

AUTHOR: OH TIEN CHENG, PRESENTER: WONG ZHAO WU

Before We Start

RESOURCES

A link to all resources mentioned (e.g. Google Colab) is in the description.

SKIPPING TO THE TUTORIAL

Check the video chapters to skip to the tutorial.

WHAT IS A JUPYTER NOTEBOOK?

1 | WHAT ARE NOTEBOOKS?

2 | WHY NOTEBOOKS?

SETTING UP A JUPYTER NOTEBOOK

3 | LOCAL INSTALLATION

4 | GOOGLE COLAB

AGENDA

What is a Jupyter Notebook?

- **INTERACTIVE DOCUMENT**

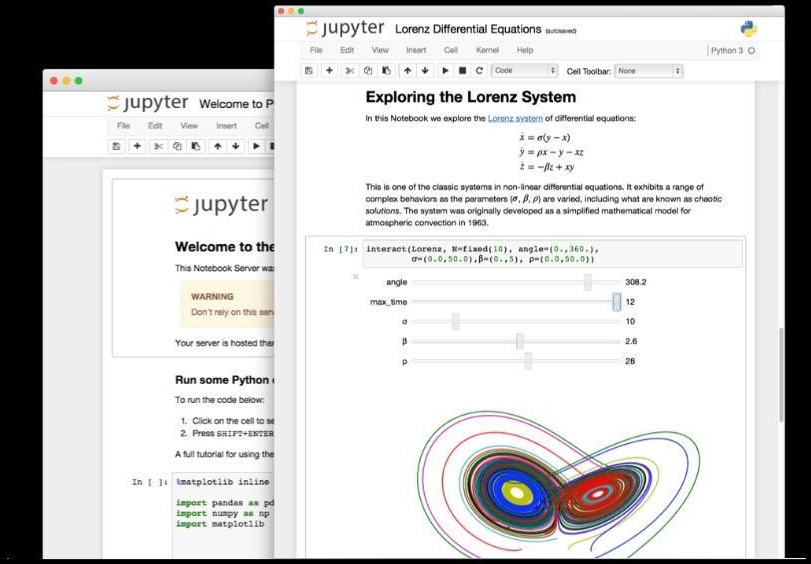
  Contains code, graphics (e.g. Images and Graphs), and text documentation.

- **LANGUAGES**

  Jupyter supports over 40 programming languages, but is typically used in Python and R

- **SHARABLE**

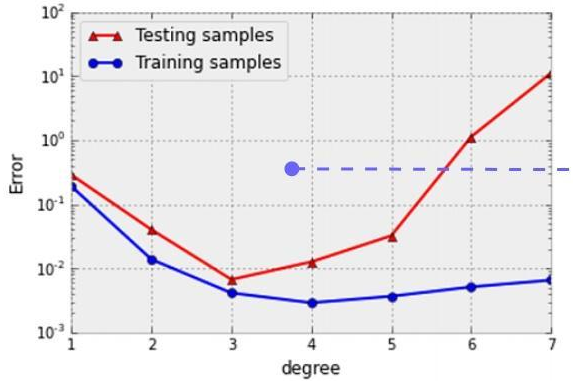
  Notebooks can be shared with others, and even exported to PDF and HTML files.



Anatomy of a Notebook

CODE CELL

```
In [40]: degrees = range(1, 8)
errors = np.array([regressor3(d) for d in degrees])
plt.plot(degrees, errors[:, 0], marker='^', c='r', label='Testing samples
plt.plot(degrees, errors[:, 1], marker='o', c='b', label='Training sample
plt.yscale('log')
plt.xlabel("degree"); plt.ylabel("Error")
  = plt.legend(loc='best')
```



CODE CELL OUTPUT

TEXT CELL

By sweeping the degree we discover two regions of model performance:

- **Underfitting** (degree < 3): Characterized by the fact that the testing error will get lower if we increase the model capacity.
- **Overfitting** (degree > 3): Characterized by the fact the testing will get higher if we increase the model capacity. Note, that the training error is getting lower or just staying the same!.

Why Notebooks?

#1: REPRODUCIBILITY

- Able to run the same programs again from scratch, and get exactly the same result.
- Lets others replicate your results.

#2: TWEAKABLE

- Lets you easily change parameters and algorithms, and see the output.
- A notebook is never finished until the entire project is done

#3: DOCUMENTABLE

- Notebooks permit documentation through text and visuals.
- Lets data scientists communicate what they are doing and why.

Setting Up Your Notebook Environment

Local Installation

1 **INSTALL ANACONDA**

Anaconda is a distribution of Python that comes with many essential libraries preinstalled.

2 **LAUNCH A JUPYTER NOTEBOOK**

Run the Jupyter Notebook shortcut on the start menu OR open the Anaconda prompt and execute the “jupyter notebook” command

SPAI

Google Colab

- **LIKE GOOGLE DOCS, BUT FOR NOTEBOOKS**
- **CLOUD NOTEBOOK**

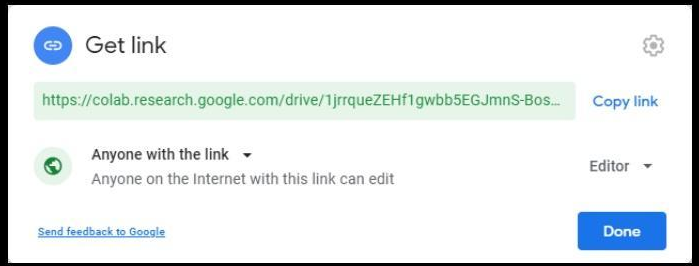
  Notebook runs on Google servers, equipped with powerful hardware

- **ZERO SETUP**

  Commonly used libraries pre-installed

SPAI

Google Colab



COLLABORATIVE FEATURES

Others can edit your notebook, comment on it, or make a copy of it

Tutorial

Any Questions?

- @spai.sp
- sp_ai@ichat.sp.edu.sg
- https://discord.gg/XCPuMWrqyE