Jupyter & Tools

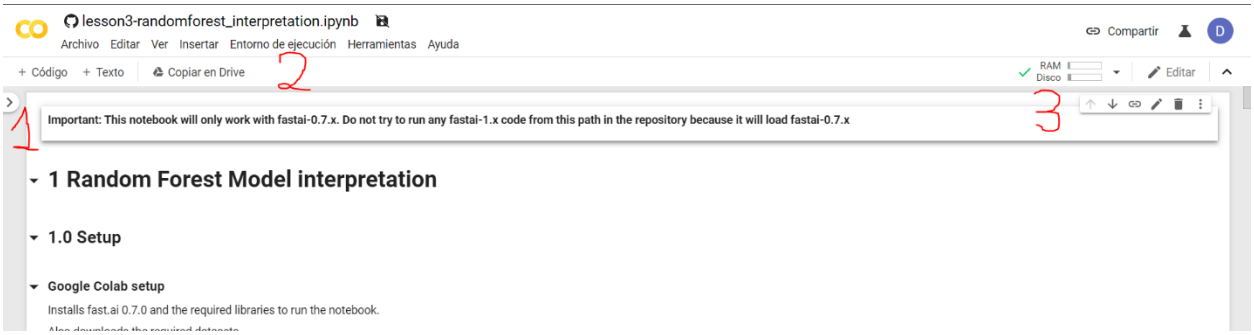
Tutorials for Jupyter

https://www.datacamp.com/community/tutorials/tutorial-jupyter-notebook

Google Colab

It is accessed from the following link :
https://colab.research.google.com/notebooks/welcome.ipynb#recent=true

It is a free platform that allows you to run python notebooks on a GPU/TPU machine to speed up model training.



In the image you have indicated the tab (1) where you can upload data files that we will use to train, the environment section (2) where the settings are to verify that the environment has a GPU and the environment information (3) that you it tells if it is connected and what resources the machine has. I remember that although you can access documents from your google drive, what you upload to the environment is not in google drive but in the machine that google has left you for this. We also remind you that every time you access colab, google will give you a new instance, so the documents you have uploaded will not be there and the special libraries you have installed will not be either. Finally, if you want to access the operating system on the machine that they leave you, you can put "!" at the beginning of the line of code in the cell you are in and you can write bash commands (pwd, ls, smear, mkdir, ...). Do not worry if you are not familiar with this, we will help you along the way. Extra notes:

- ML: Each time you enter you will have to execute the following line : !pip install fastai .

Local installation

Instructions are generally taken from this link: https://forums.fast.ai/t/fastai-v0-7-install-issues-thread/24652
This installation I only recommend it for anyone who is going to do the ML track as it is a cpu-ready installation without having the graphics card drivers. We move on to the prerequisites:

- Anaconda (python3.7)
- Git

Once these two are installed, we can proceed with the installation:

- **Windows**: Abrid anaconda prompt, you can find it in the search bar

1. git clone https://github.com/fastai/fastai.git
2. cd fastai
3. conda env create -f environment-cpu.yml
4. With these commands executed, you can now open anaconda navigator, choose the environment fastai (in navigator above a little on the left) and you've got everything ready

- **Linux/macOs:** Open a terminal

1. git clone https://github.com/fastai/fastai.git
2. cd fastai
3. conda env create -f environment-cpu.yml
4. With these commands executed, you can now open anaconda navigator, choose the environment fastai (in navigator above a little on the left) and you've got everything ready

Anaconda installation

https://www.aprendemachinelearning.com/instalar-ambiente-de-desarrollo-python-anaconda-para-aprendizaje-automatico/
Any mistakes or questions you have to throw it around Slack ;)

Introduction

Welcome to AI Saturdays Madrid! We hope these weeks are really enriching and let's learn together.

You will see that the learning methodology that we follow differs greatly from conventional models where many concepts are studied, and the teacher is the center of attention.

In AI Saturdays the center of attention is you. You yourself are the protagonist of learning and the more you want to be the more we all win.

Why do we follow a model of centralizing resources and mentoring? After following several authors and psychology and neuroscience techniques, we realized that learning should not be embodied in a role. Your learning is demonstrated by the understanding and skill you gain when you put yourself to the test.

This itinerary is organized from less to more complexity. At first you will not understand everything. It is about putting different skills into practice.

Introduction to Machine Learning

https://www.youtube.com/watch?v=WSbgixdC9g8

1ª Week - Exploratory Data Analysis & Classification

This is probably the hardest session and where you should spend more time.
During this first week we will see and practice the following concepts. For your first contact with the ML it is important to learn how to deal with the data and analyze it. We will see later how data processing occupies most of the process of developing a project.
Goals:

- Numerical & Visual EDA
- kNN & Decision Trees
- Measure Model Performance
- Data Preprocessing
- Basics ML models

Being the first week we recommend you follow the resources as we have ordered in this platform. Then start with reading and practice. Soak up everything you can :)

2ª week - Regression & Model Validation

Goals:

- Overfitting/Underfitting
- Bias/Variance
- Cross-validation
- Regularization
- Model evaluation

3ª week - Random Forest Deep Dive

Goals:

- Hyperparameter tunning
- Precision-recall curve
- Classification vs Regression Evaluation
- Random Forest
- The vine
- Feature Engineering
- Out of Bag
- Boosting

4ª week - Unsupervised Learning

- Clustering (PCA, k-means)
- Dimension Reduction
- Interpretable features, NMF

5ª week - Neural Networks, Gradient Descent

- Introduction to NN
- Gradient Descent
- Backpropagation
- Normalization, Loss Function, Learning rate, Transfer Learning
- Evaluation, metrics

6ª week - NLP

- Basic concepts
- NLP Workflow
- Regular expressions & word tokenization
- DOWN
- Feature Engineering

Week 1 Guide

1. If you need an introduction, in the last module of " *Introduction* ", we have left a couple of readings and a video that were very useful.
2. After the introduction, go to the EDA module where you have an essential video and reading. (The video you *must follow along with the notebooks that are attached in the module)*
3. Then in the Classification module we have left two very important videos too. (The *2 videos you must follow along with the notebooks that are attached in the module*)
4. We leave you an exercise to practice this week. See you on Saturday! 🚀

Practice

If during the week you want to practice before the Saturday session. **Important**: *do not look at the solution until you have finished*.

RESOURCES

topic1_practice_pandas_titanic.ipynb

topic1_practice_pandas_titanic_solution.ipynb

titanic_train.csv

Github Link:

https://github.com/SaturdaysAI/Itinerario_MachineLearning/tree/master/module_1_introduction