



MLHEP Computing Resources

2018-08-06, Oxford

Andrey Ustyuzhanin

NRU HSE

YSDA

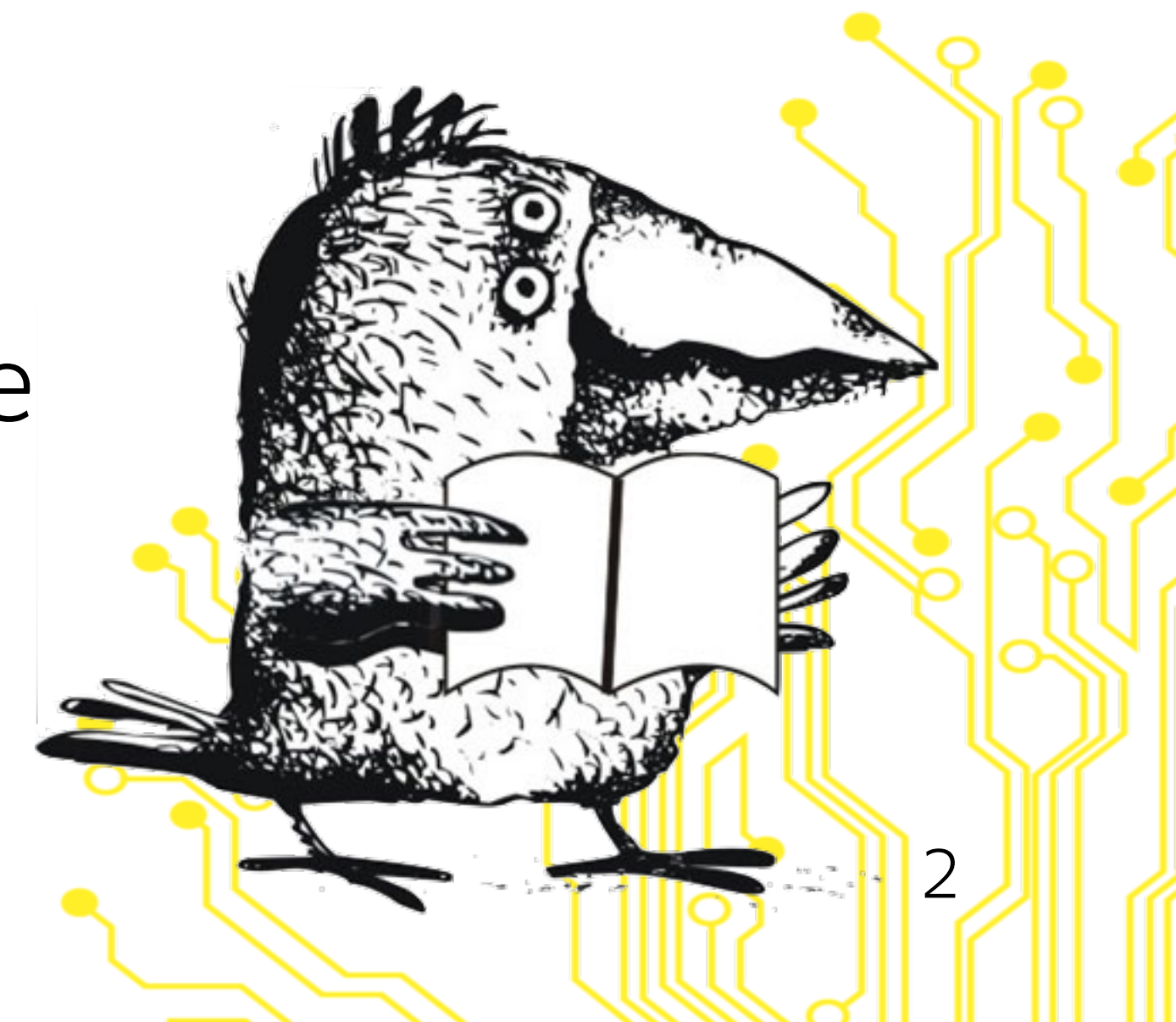
ICL

School dedicated resources



Microsoft + HSE: <URL from registration> + credentials

- › Tesla K80 GPU x 4 on each node!
- › 224GB RAM
- › 24 CPU cores
- › Python3
- › Up to 6 users per node
- › Env `CUDA_VISIBLE_DEVICES` – specifies your GPU core

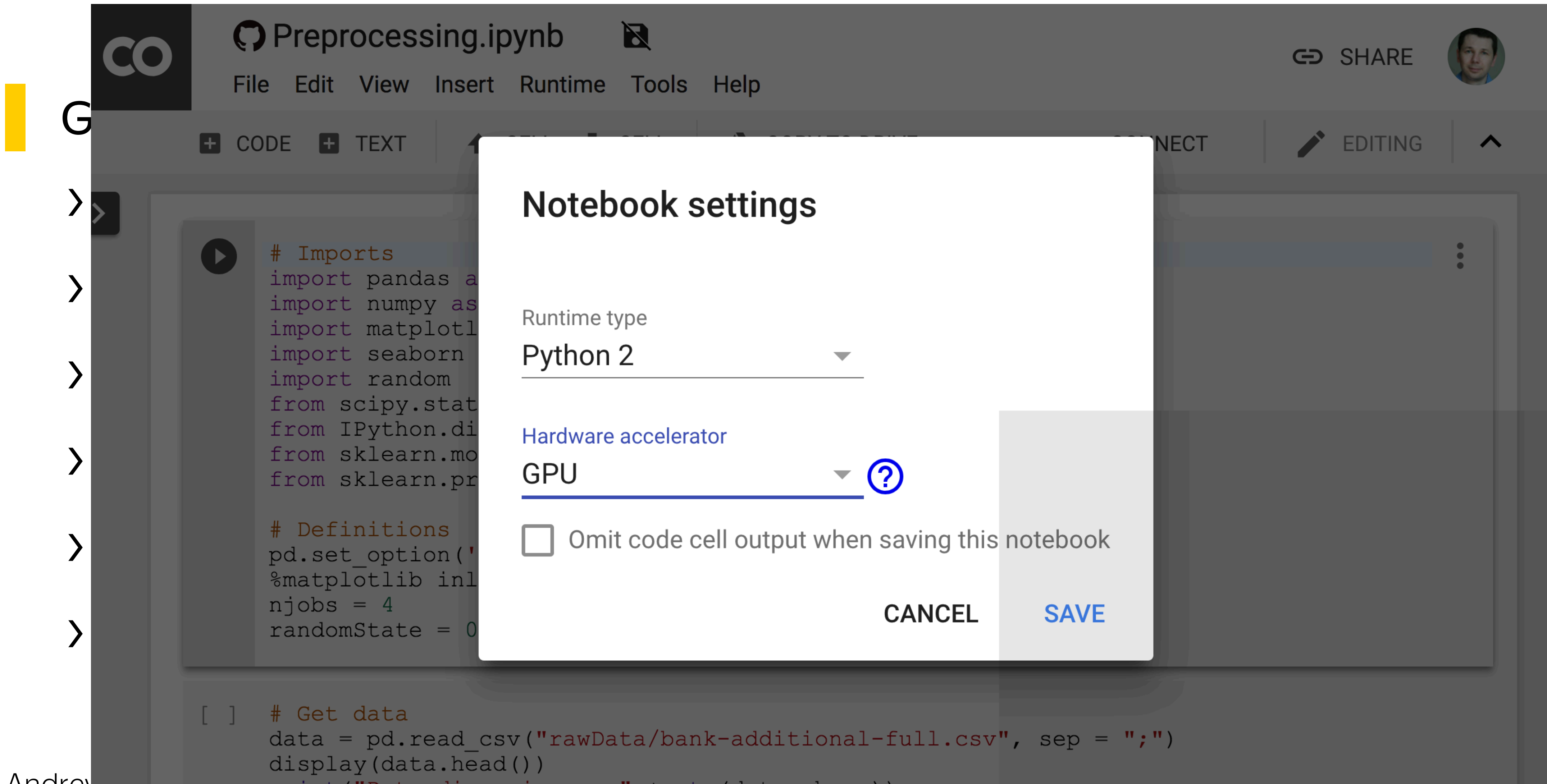


Free Resources

Google: <https://colab.research.google.com>

- › Tesla K80 GPU
- › 12GB RAM
- › 2 CPU
- › Python2/3
- › Integrates with Google Drive
- › May be crowded

Free Resources



The screenshot displays a Jupyter Notebook interface with a 'Notebook settings' dialog box open. The dialog box contains the following settings:

- Runtime type:** Python 2
- Hardware accelerator:** GPU
- ☐ Omit code cell output when saving this notebook

The background shows a Jupyter Notebook with the following code cells:

```
# Imports
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import random
from scipy.stats import norm
from IPython.display import clear_output
from sklearn.metrics import accuracy_score
from sklearn.preprocessing import StandardScaler

# Definitions
pd.set_option('display.max_colwidth', 50)
%matplotlib inline
njobs = 4
randomState = 0

[ ] # Get data
data = pd.read_csv("rawData/bank-additional-full.csv", sep = ";")
display(data.head())
```

Cheatsheet

Jupyter

- › <https://www.cheatography.com/weidadeyue/cheatsheets/jupyter-notebook/> - jupyter keyboard
- › !<command> - run command in shell

Packages

- › !pip install --user <package>
- › !git clone <REPO>
- › !git clone <https://github.com/yandexdataschool/mlhep2018.git>

Data

- › !wget -O <filename> <URL>

GPU

- › !nvidia-smi

SSH access

Good for batch process running
Authorized by public key

- › Either you have sent it to us already or
- › `!echo "YOURKEY" > ~/.ssh/authorized_keys`

