



Artificial Intelligence I: Introduction to Data Science and Machine Learning

Assoc. Prof. Dr. Taner Arsan

H. Fuat Alsan, PhD(c)

Sena Kılınç, PhD(c)

Instructors

- **Assoc. Prof. Dr. Taner Arsan**
 - Head of Computer Engineering Department
 - E-mail: arsan@khas.edu.tr
- **H. Fuat Alsan**
 - PhD Candidate in Computer Engineering
 - E-Mail: huseyinfuat.alsan@stu.khas.edu.tr
- **Sena Kılınç**
 - PhD Candidate in Computer Engineering
 - E-Mail: s.kilinc@stu.khas.edu.tr

Outline

- Python Basics
 - variables, conditionals, loops, data structures, slicing, file I/O, OOP
- Data Science
 - **Numpy:** ndarrays, vectors, matrices, basic linear algebra, data generation, example math functions, array stacking
 - **Pandas:** Series, Dataframes, reading & transforming data, handling missing data
 - **Matplotlib:** data visualization

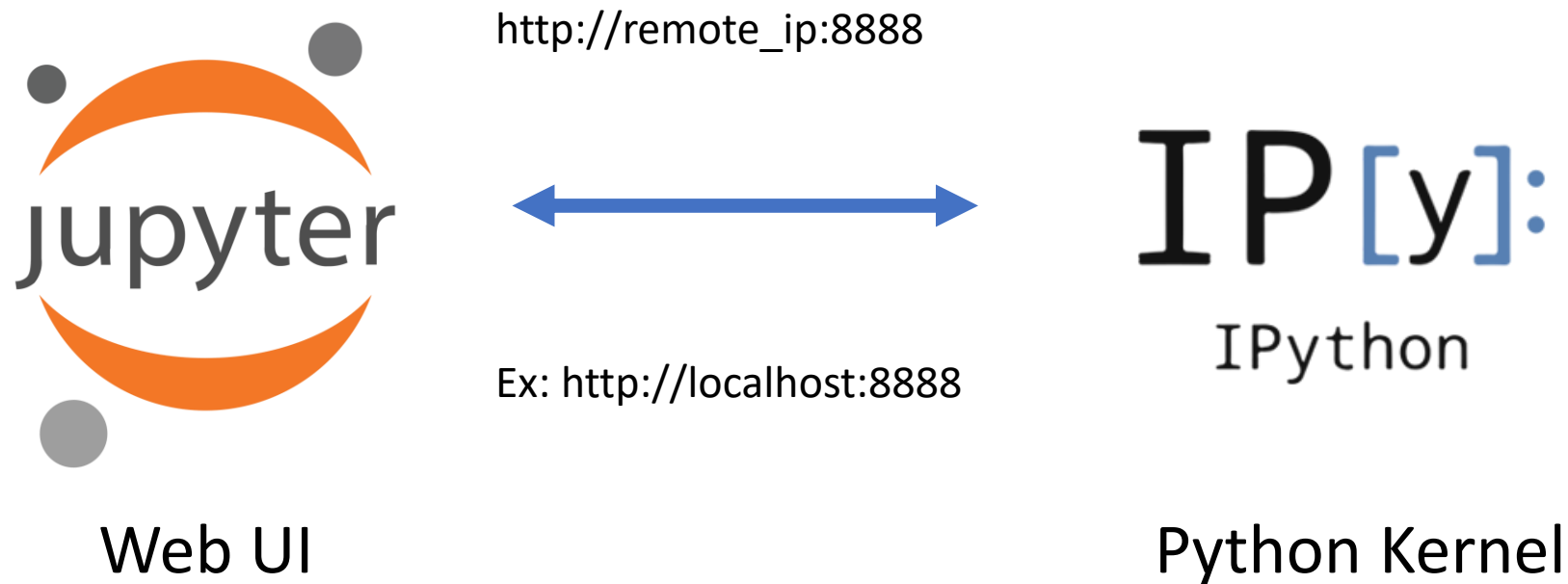
Outline (16th March)

- Machine learning introduction
- Regression with gradient descent
- Machine learning with **Sklearn**
- Feature selection & importance
- Classification (logistic regression)
- Clustering (K-means)
- Train/test split
- Model comparison & selection
 - Naïve bayes, decision tree, random forest, SVM, grid search
- Classification metrics & confusion matrix
 - TP, TN, FP, FN, F1 score, ROC curve
- Class imbalance

Jupyter Lab (Notebook)

- Client-Server based application
 - Client: web UI (browser)
 - Server: Python runtime (kernel)
- Web based interactive environment for working with data
 - Web page has executable cells (code, markdown and raw)
 - Code cells are sent to Python kernel
 - Results from Python kernel are shown in browser
- Jupyter Lab: Newer, with better UI
- Jupyter Notebook: Classic notebook

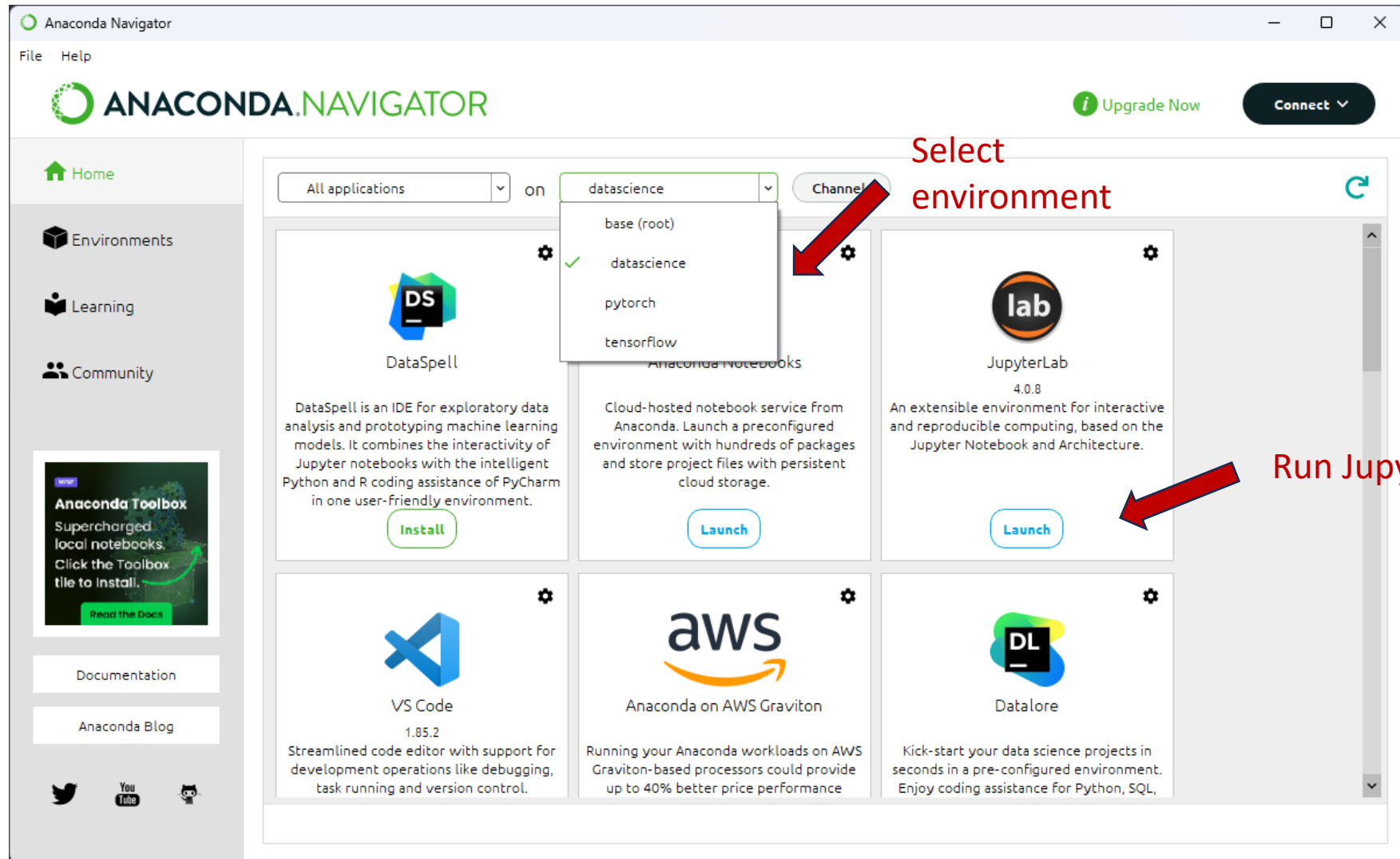
Jupyter Lab (Notebook) Architecture



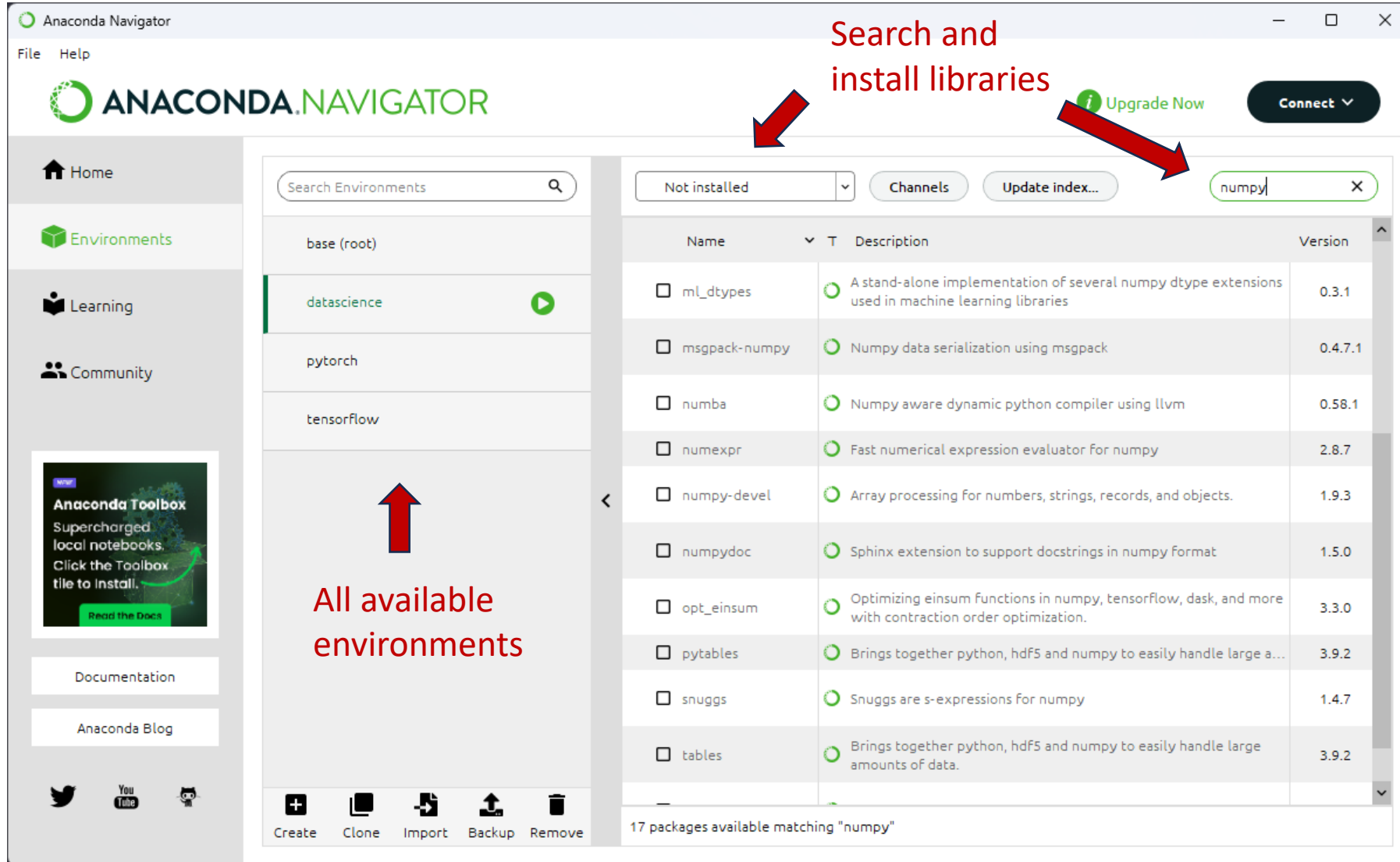
How to Launch Jupyter Lab/Notebook

- With terminal
 - \$ cd <working_dir>
 - \$ jupyter lab
 - \$ jupyter notebook
 - visit <http://localhost:8888/> on your browser
- Without terminal
 - Install Anaconda Navigator
 - Launch Jupyter Lab/Notebook from main page
 - (Browser should run automatically)

Anaconda Navigator (Run Jupyter)



Anaconda Navigator (Environments)



Anaconda Navigator

File Help

ANACONDA.NAVIGATOR

Home

Environments

Learning

Community

Anaconda Toolbox
Supercharged local notebooks.
Click the Toolbox tile to Install.
[Read the Docs](#)

Documentation

Anaconda Blog

Twitter YouTube GitHub

Create Clone Import Backup Remove

Search Environments

base (root)

datascience

pytorch

tensorflow

Search and install libraries

Upgrade Now

Connect

Not installed

Channels

Update index...

numpy

Name	Description	Version
<input type="checkbox"/> ml_dtypes	A stand-alone implementation of several numpy dtype extensions used in machine learning libraries	0.3.1
<input type="checkbox"/> msgpack-numpy	Numpy data serialization using msgpack	0.4.7.1
<input type="checkbox"/> numba	Numpy aware dynamic python compiler using llvm	0.58.1
<input type="checkbox"/> numexpr	Fast numerical expression evaluator for numpy	2.8.7
<input type="checkbox"/> numpy-devel	Array processing for numbers, strings, records, and objects.	1.9.3
<input type="checkbox"/> numpydoc	Sphinx extension to support docstrings in numpy format	1.5.0
<input type="checkbox"/> opt_einsum	Optimizing einsum functions in numpy, tensorflow, dask, and more with contraction order optimization.	3.3.0
<input type="checkbox"/> pytables	Brings together python, hdf5 and numpy to easily handle large a...	3.9.2
<input type="checkbox"/> snuggs	Snuggs are s-expressions for numpy	1.4.7
<input type="checkbox"/> tables	Brings together python, hdf5 and numpy to easily handle large amounts of data.	3.9.2

17 packages available matching "numpy"

All available environments

Python Environments

Python == 1.12.1
...

base

Python == 1.11.7
NumPy == 1.24.1
...

datascience

Python == 1.11.7
PyTorch == 2.1.1
...

dl_pytorch

Python == 1.9.8
PyTorch == 1.13.1
...

dl_pytorch_old