



Hyrje në shkencën e të dhënave

Pjesa 5 – Python

Prof. Asoc. Dr. Ermir Rogova

What is Python?



A scripting language available on every platform



Easy to learn, easy to use, quite extensible and robust

Where is Python?

- Many UNIX systems have it pre-installed
- Free to download and install for all platforms
- Command-line:
 - \$ python
- IDE (integrated development environment)
 - Eclipse
 - Jupyter notebook
 - Anaconda
 - Spyder

Basics of Python

- One-line examples
- Arithmetic operators
- Logical operators
- Data types
- Control structures
 - Condition checking with ‘if’ and ‘else’
 - While loop
 - For loop

Statistics with Python

- Storing a set of numbers
- Numpy library
- Doing descriptive analysis
- Visualization with bar graph

Statistical analysis

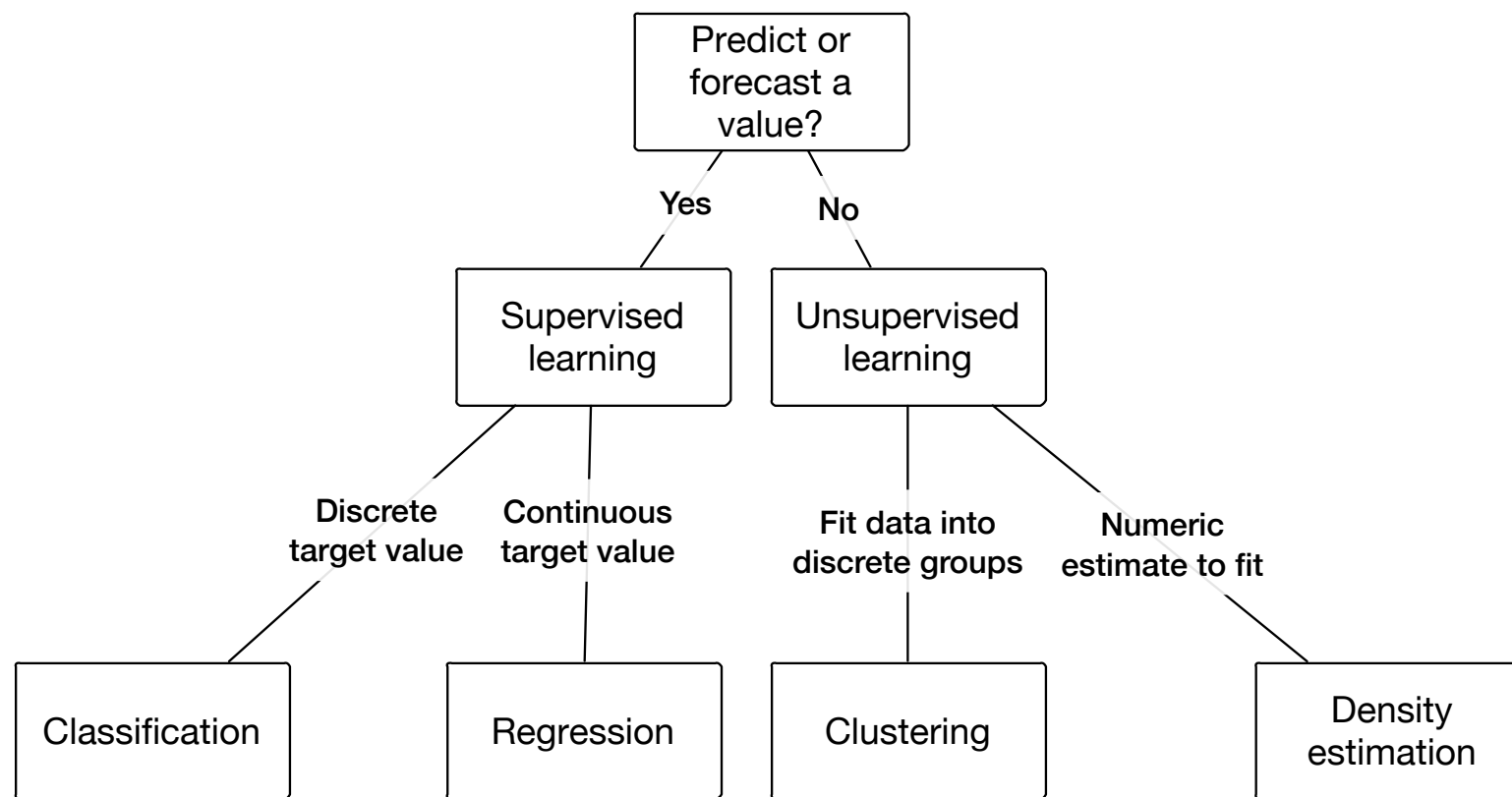
- Loading external data and using pandas library
- Correlation
- Linear regression
- Multiple linear regression

Basics of machine learning with Python

- What is machine learning?
- Classification (supervised learning)
- Clustering (unsupervised learning)
- Density estimation (unsupervised learning)

Outline of core ML problems

Machine Learning Outline



Summary

- Python: the most used language for doing data science
- Simple and easy to learn, yet versatile
- Most common tools for practicing Python: Jupyter, Anaconda
- Most useful libraries for data science: numpy, pandas, matplotlib, sklearn

Temat e javës së gjashtë

- R
- Getting Access to R
- Getting Started with R
- Graphics and Data Visualization
- Statistics and Machine Learning



Pyetje???