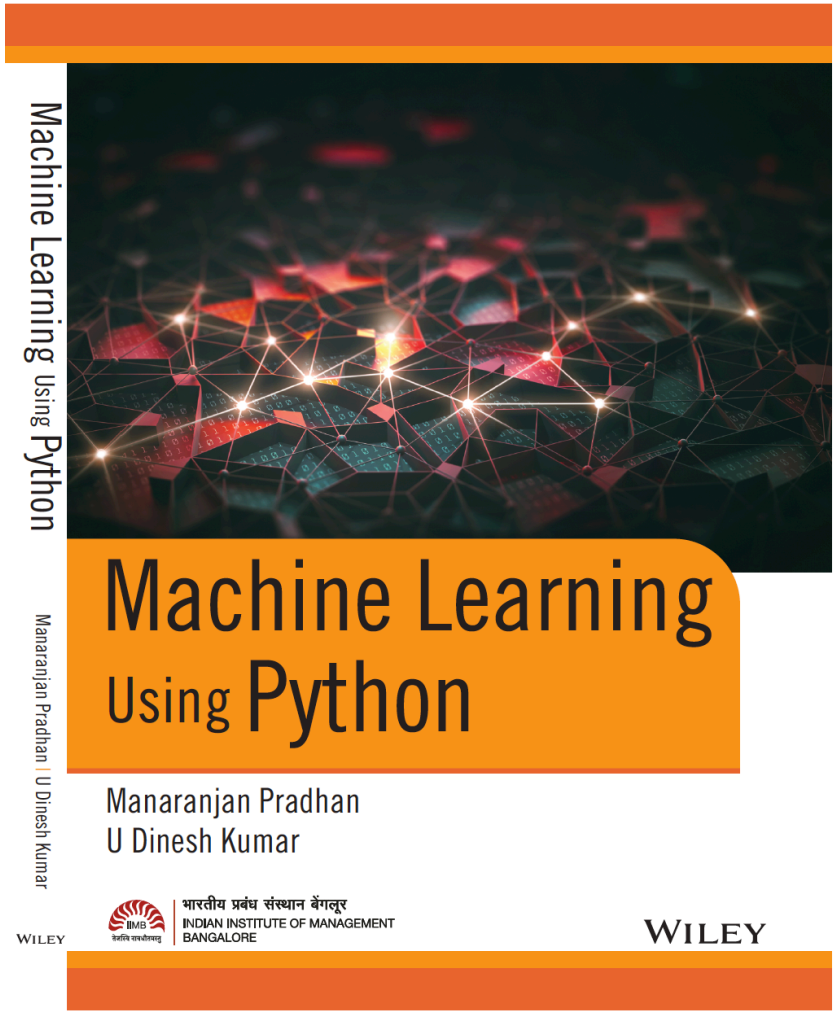

Chapter 01: Introduction to Machine Learning

Prepared By: Purvi Tiwari

Learning Objectives

- Understanding how machine learning is used to solve problems in real world.
- Understanding types of machine learning algorithms and frameworks
- Learning the key libraries in Python, setting up Anaconda platform
- Learning the basic features of Python language to get started with machine learning tasks.

Introduction

- Analytics – a collection of techniques such as artificial intelligence, machine learning and deep learning and tools used for creating value from data.
- Artificial Intelligence (AI) : Algorithms and systems that exhibit human-like intelligence.
- Machine Learning (ML): Subset of AI that can learn to perform a task with extracted data and/or models.
- Deep Learning (DL): Subset of machine learning that imitate the functioning of human brain to solve problems.

Relationship between AI, ML, and DL

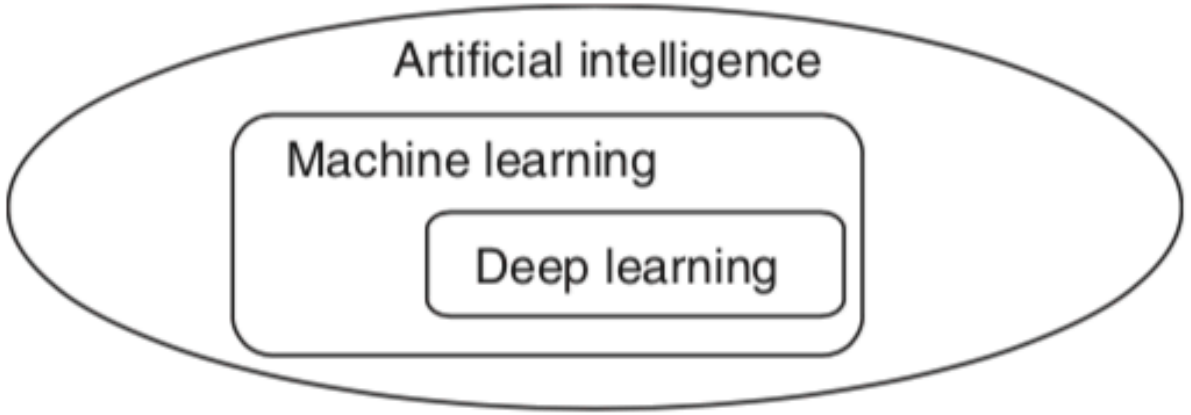

FIGURE 1.1 Relationship between artificial intelligence, machine learning, and deep learning.

Classification of ML algorithms

Supervised learning algorithms

- require the knowledge of both outcome variable (dependent variable) and the features (independent variable)
- Algorithm learns by defining a loss function
- Loss function is usually a function of the difference between the predicted value and actual value of the outcome variable.
- Examples: linear regression, logistic regression, discriminant analysis, etc.

Classification of ML algorithms (Cntd.)

Unsupervised learning algorithms

- Set of algorithms which do not have the knowledge of the outcome variable in the dataset.
- Algorithms must find the possible values of the outcome variable.
- Examples: clustering, principal component analysis, etc.

Classification of ML algorithms (Cntd.)

Reinforcement learning algorithms

- Algorithms that have to take sequential actions (decisions) to maximize a cumulative reward
- There is uncertainty around both input as well as the output variables.
- Examples: Markov chain, Markov decision process, etc.

Evolutionary learning algorithms

- Algorithms that imitate natural evolution to solve a problem.
- Examples: genetic algorithms and ant colony optimization.

Why Machine Learning?

- It helps in understanding the association between key performance indicators (KPIs).
- Identifying the factors that have a significant impact on the KPIs for effective management.
- Knowledge of the relationship between KPIs and factors would provide decision maker with appropriate action items.
- Used for identifying the factors that influence the KPIs, which helps in decision making and value creation.
- Organizations such as Amazon, Google, Capital One, IBM, Facebook are using ML algorithms to create new products and solutions.

Steps for ML algorithm

- Identify the problem or opportunity for value creation.
- Identify sources. Of data and create a data lake.
- Pre-process the data for issues such as missing and incorrect data.
- Generate derived variables and transform the data if necessary.
- Divide the datasets into subsets of training and validation datasets.
- Build ML models to identify the best models(s) using model performance in validation data.
- Implement Solution/Decision/Develop Product

Framework for Developing ML Models

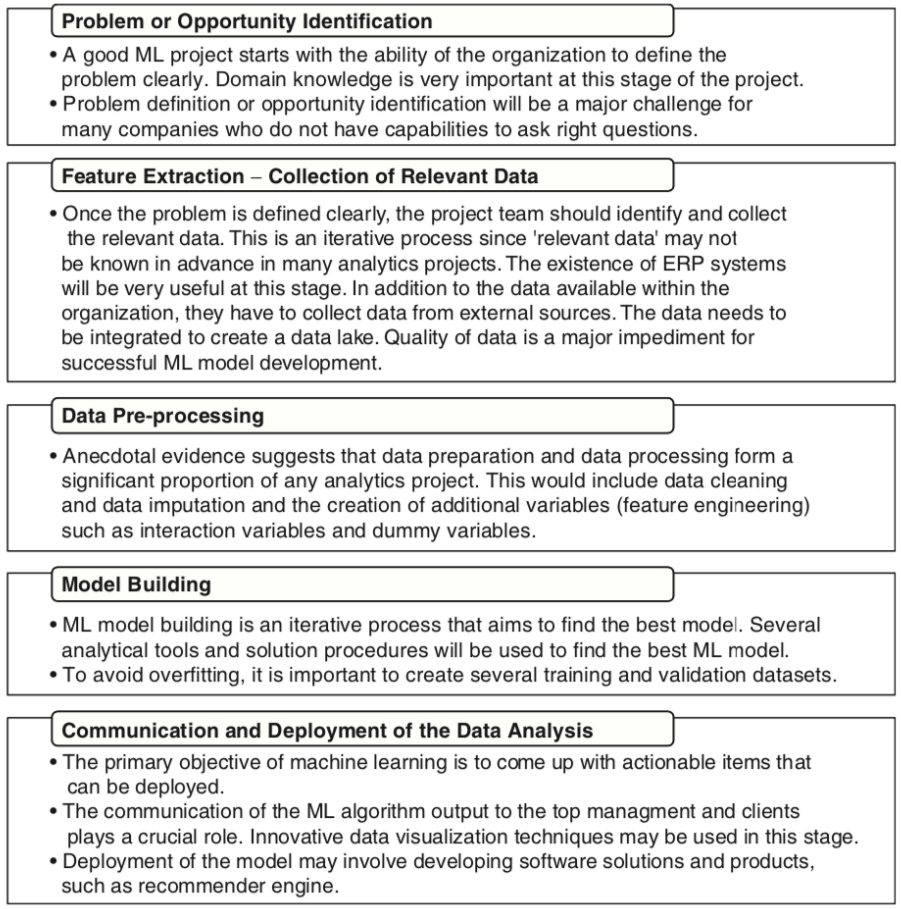

FIGURE 1.2 Framework of ML model development.

Why Python?

- An interpreted, high-level, general-purpose programming language
- Ease of use and high productivity
- Has a comprehensive set of core libraries for data analysis and visualization.
- Can be used to build web applications, enterprise applications
- It has libraries for linear algebra computations, statistical analysis, machine learning, visualization, optimization, stochastic models, etc.

Why Python? (Cntd.)

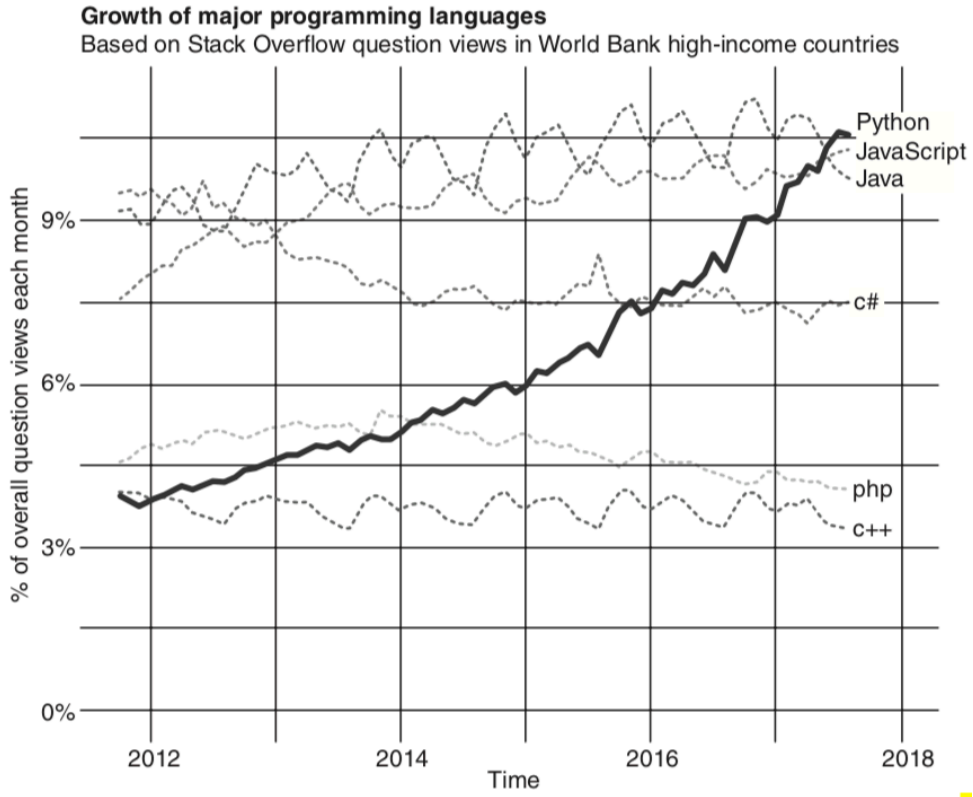

FIGURE 1.3 Overall question views for Python.

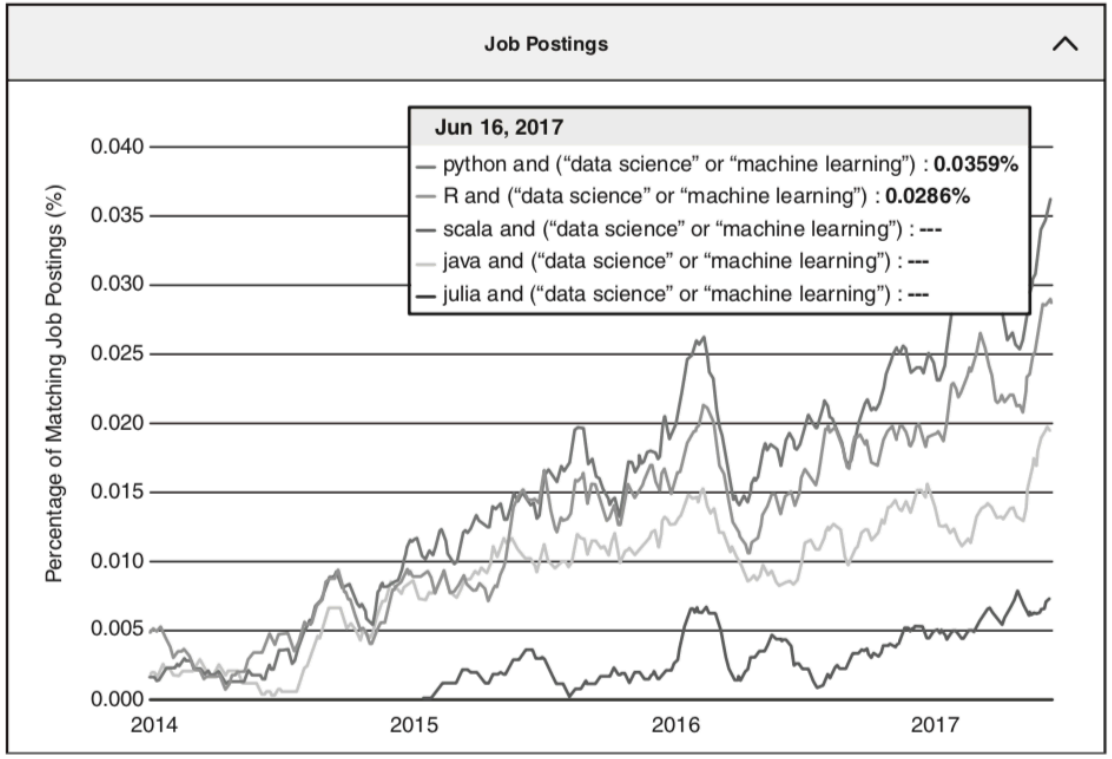

FIGURE 1.4 Trend on job postings.

Python Stack for Data Science

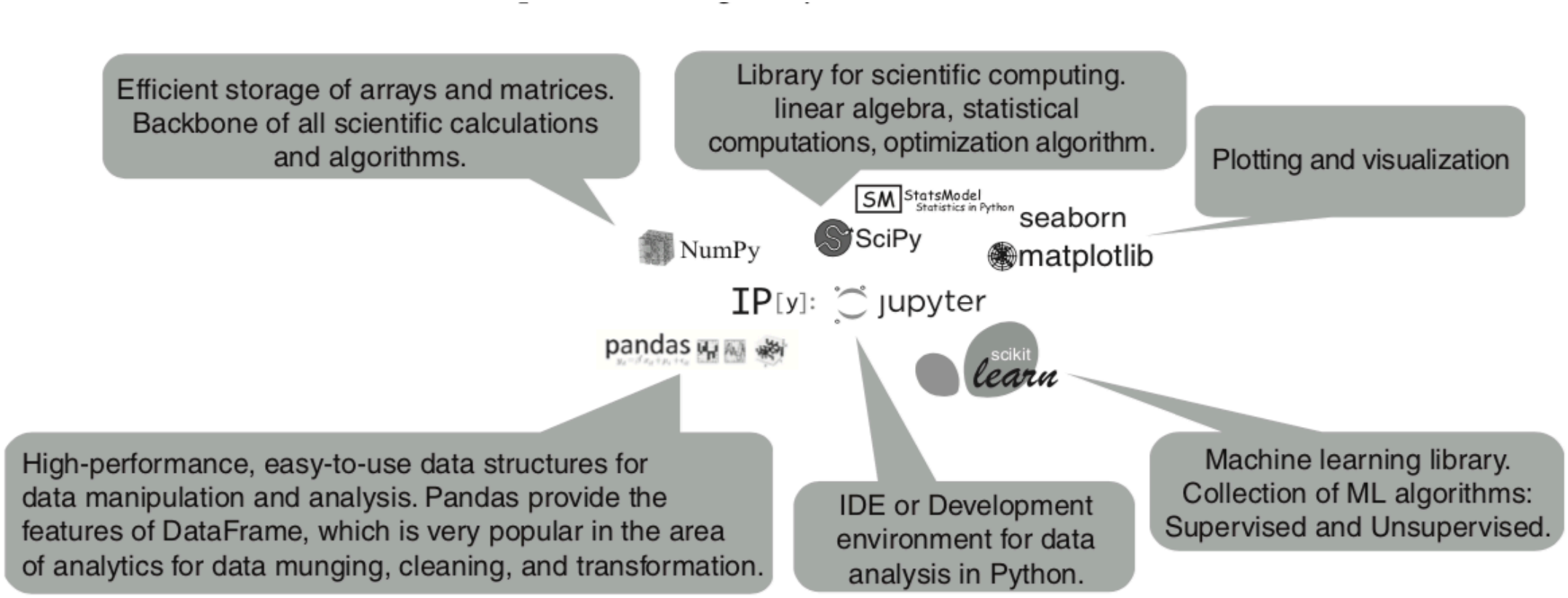

FIGURE 1.5 Python core libraries for data science applications.

Getting Started with Anaconda Platform

STEP 1: Go to Anaconda Site
Go to *https://www.anaconda.com/distribution/* using your browser window.

STEP 2: Download Anaconda Installer for your Environment
Select your OS environment and choose Python 3.7 version to download the installation files as shown in Figure 1.6.

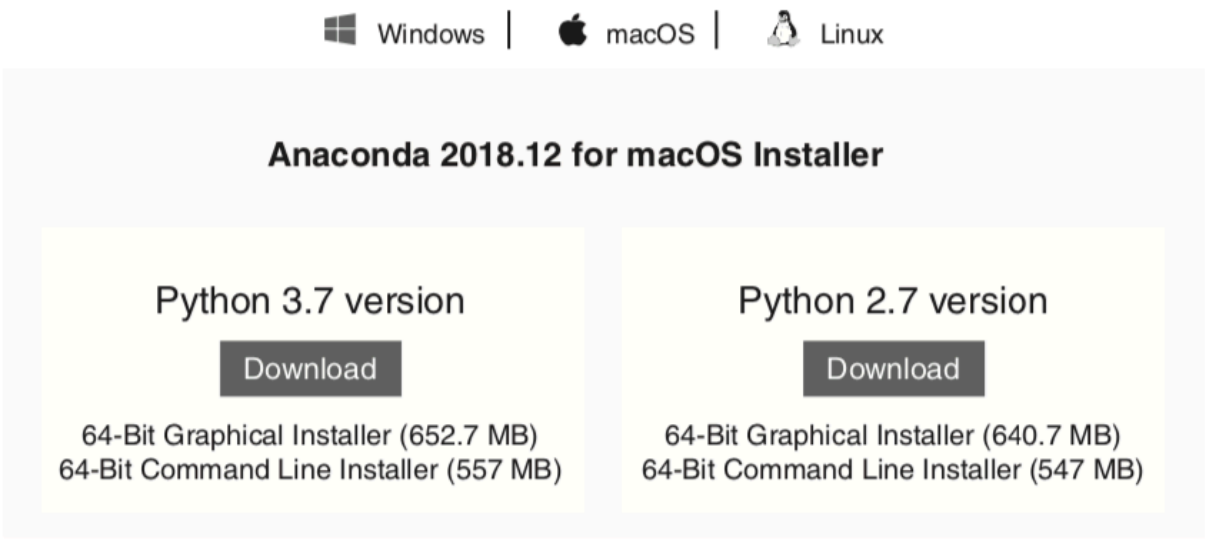

FIGURE 1.6 Anaconda distribution site for downloading the installer.
Source: www.anaconda.com

Getting Started with Anaconda Platform (Cntd.)

Step 3: Install Anaconda

Double click on the downloaded file and follow the on-screen installation instructions, leaving options as set by default. This will take a while and complete the installation process.

Step 4: Start Jupyter Notebook

Open the command terminal window as per your OS environment and type the following command, as shown in Figure 1.7.

*jupyter notebook -- ip=**

```
Manaranjans-MacBook-Pro:~ manaranjan$ jupyter notebook --ip=*
[I 10:20:14.928 NotebookApp] [nb_conda_kernels] enabled, 2 kernels found
[W 10:20:15.636 NotebookApp] WARNING: The notebook server is listening on all IP addresses and not using encryption. This is not recommended.
[I 10:20:15.805 NotebookApp] [nb_anacondacloud] enabled
[I 10:20:15.811 NotebookApp] [nb_conda] enabled
[I 10:20:15.893 NotebookApp] ✓ nbpresent HTML export ENABLED
[W 10:20:15.893 NotebookApp] ✗ nbpresent PDF export DISABLED: No module named 'nbbrowserpdf'
[I 10:20:15.900 NotebookApp] Serving notebooks from local directory: /Users/manaranjan
[I 10:20:15.900 NotebookApp] 0 active kernels
[I 10:20:15.900 NotebookApp] The Jupyter Notebook is running at: http://[all ip addresses on your system]:8888/?token=ae80c575b3993c6dcf5ba8d0b
f5f577da1c019e1380e4eef
[I 10:20:15.900 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[C 10:20:15.901 NotebookApp]

    Copy/paste this URL into your browser when you connect for the first time,
    to login with a token:
        http://localhost:8888/?token=ae80c575b3993c6dcf5ba8d0bf5f577da1c019e1380e4eef
[I 10:20:16.320 NotebookApp] Accepting one-time-token-authenticated connection from ::1
[W 10:20:16.537 NotebookApp] 404 GET /apple-touch-icon-precomposed.png (::1) 7.52ms referer=None
[W 10:20:16.541 NotebookApp] 404 GET /apple-touch-icon.png (::1) 1.01ms referer=None
```

FIGURE 1.7 Screenshot of starting the jupyter notebook.

Getting Started with Anaconda Platform (Cntd.)

This should start the Jupyter notebook and open a browser window in your default browser software as shown in Figure 1.8.

FIGURE 1.8 Screenshot of the file system explorer of Jupyter notebook open in the browser.

Getting Started with Anaconda Platform (Cntd.)

The reader can also start browser window using the URL highlighted below. The URL also contains the password token as shown in Figure 1.9.

```
[I 10:20:15.900 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[C 10:20:15.901 NotebookApp]

    Copy/paste this URL into your browser when you connect for the first time,
    to login with a token:
        http://localhost:8888/?token=ae80c575b3993c6dcf5ba8d0bf5f577da1c019e1380e4eef
[I 10:20:16.320 NotebookApp] Accepting one-time-token-authenticated connection from ::1
[W 10:20:16.537 NotebookApp] 404 GET /apple-touch-icon-precomposed.png (::1) 7.52ms referer=None
[W 10:20:16.541 NotebookApp] 404 GET /apple-touch-icon.png (::1) 1.01ms referer=None
```

FIGURE 1.9 Screenshot of using notebook URL.

Getting Started with Anaconda Platform (Cntd.)

Step 5: Create a New Python Program

On the browser window, select "New" for a menu. Clicking on the "Folder" will create a directory in the current directory.

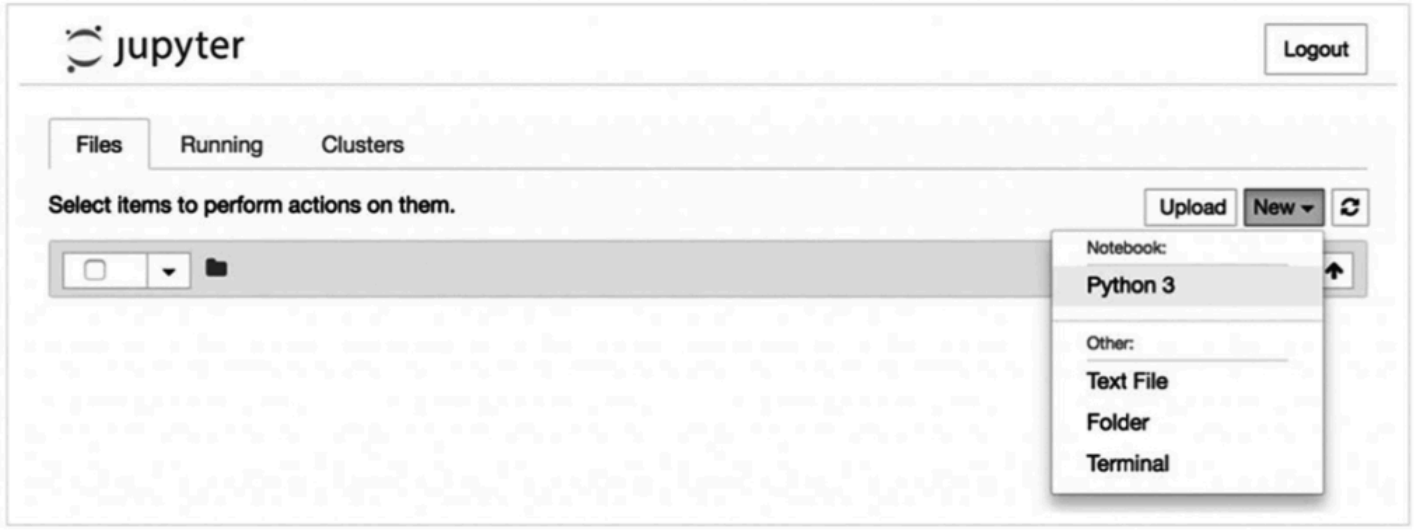

FIGURE 1.10 Screenshot of creating a new Python program.

To create a Python program, click on "Python 3". It will open a new window, which will be the program editor for the new Python program as shown in Figure 1.10.

Getting Started with Anaconda Platform (Cntd.)

Step 6: Rename the Program

By default, the program name will be "Untitled". Click on it to rename the program and name as per your requirement. For example, we have renamed it to "My First Program" as shown in Figure 1.11.

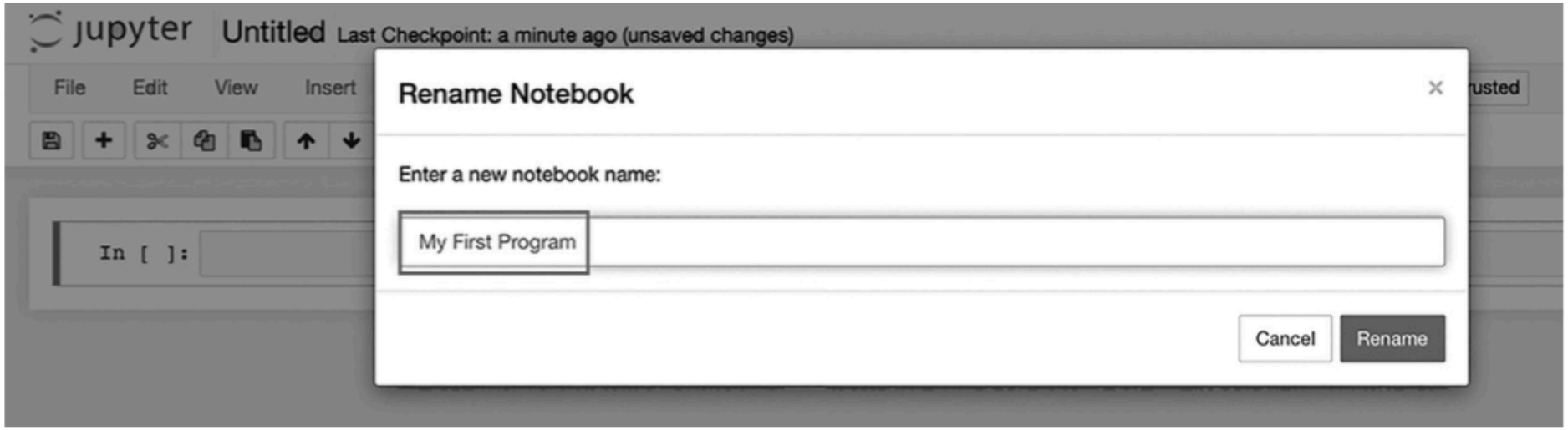

FIGURE 1.11 Screenshot of renaming the python program file.

Getting Started with Anaconda Platform (Cntd.)

Step 7: Write and Execute Code

Write Python code in the cell and then press SHIFT+ENTER to execute the cell as shown in Figure 1.12.

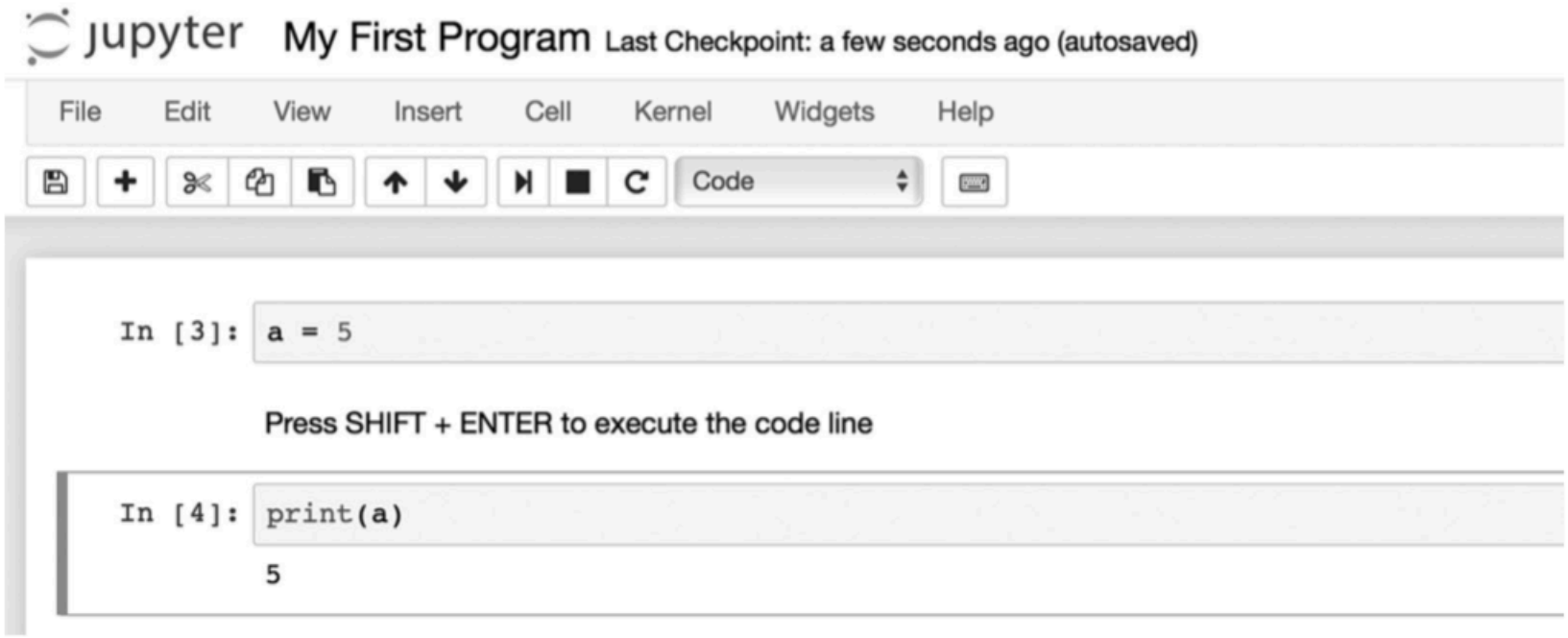

FIGURE 1.12 Screenshots of Python code and execution output.

Getting Started with Anaconda Platform (Cntd.)

Step 8: Basic Commands for Working with Jupyter Notebook
Click on "User Interface Tour" for a quick tour of Jupyter notebook features. Or click on "Keyboard Shortcuts" for basic editor commands as shown in Figure 1.13.

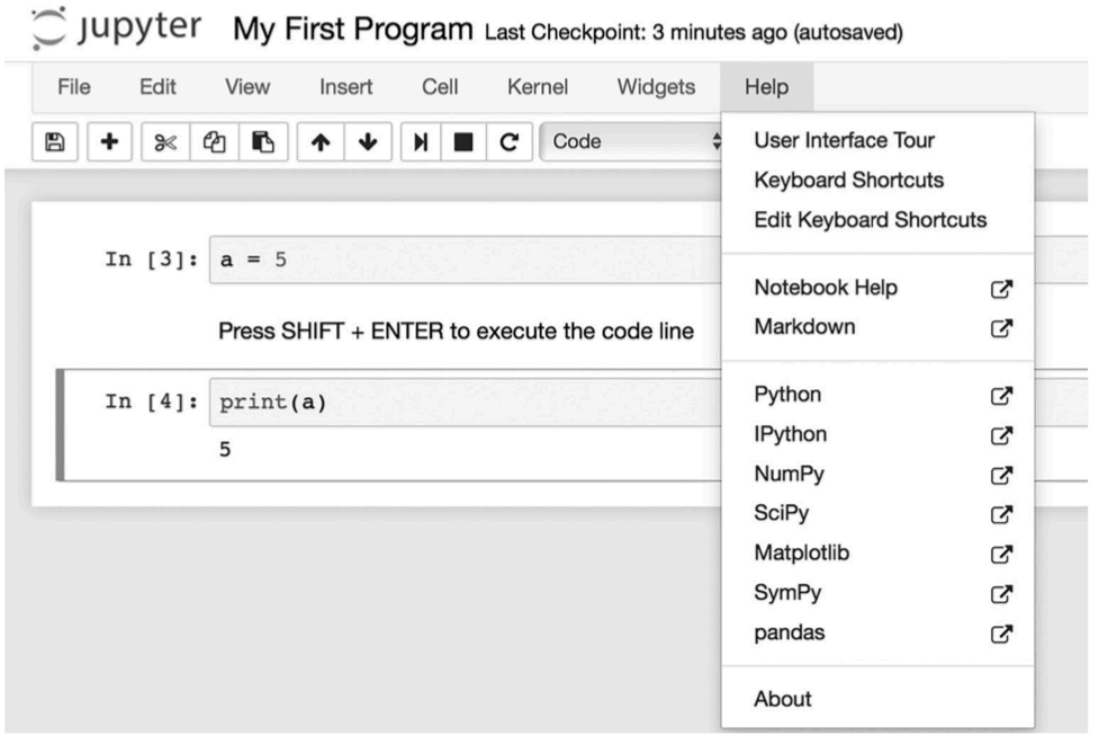

FIGURE 1.13 Screenshot of basic commands in Jupyter notebook.

Introduction to Python

- List of features that will be discussed

1. Declaring variables
2. Conditional Statements
3. Control flow statements
4. Collections
5. Functions
6. Functional Programming
7. Modules and packages

Declaring Variables

- Python automatically infers the variable type from values assigned to it.
- The initialized value of a variable can later be re-assigned as value of a different type.
- Python supports the following variable types:

1. *int* – Integer type.
2. *float* – floating point numbers.
3. *bool* – assigned values using *True* and *False.*
4. *str* – textual data.