

Workshop

Data science for Business and Finance using Python

Workshop Objectives

The aim of the workshop is to provide a broad understanding of the principles and techniques of data analysis and machine learning in business applications through the use of Python coding in Jupiter notebooks with intuitively visualized output. These objectives will be achieved by solving tasks/case studies and discussing theory. In this course, the real datasets from retail, marketing, social media, finance and other interesting sectors will be introduced for business decision support. For example, profiling of customers and their transactions will be used to identify the advertising strategies and create the new market segments for the business. The food reviews and customer feedbacks from social media will be analyzed by using the text mining for generating the ideas to improve its services.

On completion of the participants will be expected to:

- Understand what machine learning is and why it is essential to the design of intelligent machines.
- Understand how big data technologies can improve and extend the opportunity for its business.
- Know how to fit the suitable models for types of business decisions.
- Understand numerical computation, statistics and optimization in the context of learning.
- Have a good understanding of the business problems that arise when dealing with very small and very big data sets, and how to solve them.
- Understand the basic mathematics necessary for constructing novel data analysis and machine learning solutions.
- Be able to design and implement various data analysis and machine learning algorithms in a wide range of real-world applications.

Understanding the problems covered in this workshop can be important for all students seeking a career in the expanding technology-intensive field of financial engineering/Business analytics or grappling with the challenges related to the business data.

1. Reference materials

1.1. Main textbook

Foster Provost and Tom Fawcett, Data Science for Business: Fundamental principles of data mining and data analytic thinking, O'Reilly 2013. (P&F)



Ewen Gallic, Python for economists. (E)

Luis Pedro Coelho and Willi Richert, Building Machine Learning Systems with Python, 2nd Edition, Packt publishing 2013. (C&R)

2. Workshop plan

Part 1: Understand the Basic Python----“Hello World”

Session	Topic	Lecturer/ Reading/Lab
Session 1-2	Introduction to the Course Introduction to Data Analytics Doing Data Science Tools for Business Analytics – Conceptual and Practical Introduction to Python <i>Learn the basics of Python, one of the world's most popular and powerful programming languages, and see how it can be utilized in Economics and Finance.</i>	Ch.1,9 (E) Ch. 1-2 (P&F) Computer Lab: Basic Python: NumPy Case: managing churn in wireless telecom.
	Introduction to NumPy <i>Learn the basics of NumPy, a Python package for handling large, multi-dimensional arrays and</i>	

Session	Topic	Lecturer/ Reading/Lab
	<i>matrices, to quicken your economic and financial analyses.</i>	
Session 3-4	Introduction to Pandas <i>Learn the basics of Pandas, a Python package for data manipulation and analysis, and discover how to access economic and financial data from APIs.</i>	Ch.10 (E) Basic Python: Pandas Case Study: analyzing asset prices -Bitcoin returns
Session 5-6	Visualizing Economic and Finance Data <i>Learn the basics of Matplotlib, a Python package for data visualization, and discover how to optimize a portfolio of stocks.</i>	Ch.11 (E) Case Study: analyzing asset prices -Stock returns
Session 7-8	Statistics for Data Analysis <i>Learn how to calculate and interpret several descriptive statistics using the Python library NumPy.</i>	Ch.9 (E) Case Study: COVID-19 and Resale Price of Luxury Masks
Session 9-10	Hypothesis Testing with SciPy <i>Learn SciPy, a Python module for comparative statistics, in order to perform many different statistical tests in code.</i>	Lecture Note Case Study: Firm Level Innovation and CEO Compensation

Part 2: Born to be Young Data Scientists

Session 1	Intro to Predictive Modeling Supervised Segmentation Conceptual Tools for Creating a Predictive Analytics Solution	Ch. 3-4 (P&F)
	Classifying with Real-world Business Examples <i>uses real data to learn about Decision Tree classification, whereby we train a computer to be able to distinguish different classes of data.</i>	Ch.2 (R&C) Case Study: Predict Loan Funding with Loan Profile
Session 2-3	Classifying with Real-world Business Examples (Cont.) <i>uses real data to learn about Decision Tree classification, whereby we train a computer to be able to distinguish different classes of data.</i>	Case study: Classification Analysis of Suicide Rate
Session 4	Regression <i>explains how to use the classical topic, regression, in handling data, which is still relevant today. We also learn about advanced regression techniques such as the Lasso and ElasticNets</i>	Ch.7 (R&C)
Session 5-6	Regression (Cont.) <i>explains how to use the classical topic, regression, in handling data, which is still relevant today. We also learn about advanced regression techniques such as the Lasso and ElasticNets</i>	Case Study: Predicting the Local House Prices
Session 7	Classification II <i>Sentiment Analysis, explains how Naïve Bayes works, and how to use it to classify tweets to see whether they are positive or negative.</i>	Ch. 9 & 10 (P&F) Lecture Note

		<i>Case Study: Modeling consumer behavior for targeted marketing (banking and/or online advertising)</i>
Session 8-9	<i>Classification II (Cont.) Sentiment Analysis, explains how Naïve Bayes works, and how to use it to classify tweets to see whether they are positive or negative.</i>	<i>Case Study: Sentiment Classification on movie reviews (IMDb)</i>
Session 10	Wrap Up <i>Data Analytics, Strategy & Tactics</i>	Ch. 12 & 13 (P&F)