Data Analytics and Machines Learning in FInance

Frankfurt School of Finance & Management
Spring 2018

Prof. Dr. Gregory Wheeler g.wheeler@fs.de
TA: Mr. Ruishen Zhang r.zhang@fs.de
TA: Mr. Xiang Zheng x.zheng@fs.de
Course Server: https://damlf-course.de

Course Description

Advanced data analytics employs techniques from machine learning and artificial intelligence to sift through large amounts of data, sometimes even unstructured data, to reveal patterns and identify trends to yield more accurate judgments and better-informed decisions. The aim of machine learning is to make a computer learn from data without explicitly programming it how to do so, and the fruits of machine learning are all around us: email spam filters classify your messages; online businesses recommend products to customers; speech-to-text transcribers now match the accuracy of human transcribers, opening the possibility of real-time automated translation; and pharmaceutical companies analyze social media feeds to identify symptoms of early onset cognitive disorders. Financial institutions increasingly apply these very same techniques to an expanding range of problems, leveraging the growing volume of data from daily operations and third-party sources to manage portfolio risk, perform trades, detect fraud, segment markets, comply with regulations, and much, much more.

This course is a hands-on introduction to contemporary machine learning, with a focus on supervised learning algorithms (used to make accurate predictions about the future from current data) and unsupervised learning (used to discover unknown structure in your current data), and deep learning.

Because applications in this field are fast moving, the focus of this course is to give students a working understanding of core ML techniques backed by a solid theoretical understanding of each algorithm.

Requirements

You will be asked to complete a series of Python programming assignments using the Jupyter Notebook computational environment to modify and run IPython notebooks. All assignments will be released and submitted exclusively through the course website: https://damlf-course.de.

Assignment	Topic	m / d due dates	points
ps0	Introduction to Python	31.mar.2018	16
ps1	Linear Regression & Gradient Descent	m 15.apr.2018 d 24.apr.2018	16
ps2	Logistic Regression & BFGS Algorithm	m 22.apr.2018 d 30.apr.2018	16
ps3	Deep Learning with Keras & TensorFlow	m 06.may.2018 d 10.may.2018	16
ps4	Unsupervised Learning & Dimension Reduction	m 22.may.2018 d 23.may.2018	16
	Final Exam	26.may.2018	40

Note the separate deadlines for the M01 and M02 sections (m) and the D01 and D02 sections (d) of the course. Each student's deadline is determined by her section enrollment as of 01.apr2018.

Late Policy

Each student has a **48 hour budget for LATE ASSIGNMENTS** for the entire semester to account for unforeseeable minor problems. After a student's 48 hours have been used, each submission arriving between 00:00 and 23:59 after deadline will be worth a maximum of 50% of the total available points and any submission arrive 24 hours after deadline will be worth zero. A major problem is one that has documentation, such as a doctor's note. There are no exceptions to this policy.

Note that you are free to submit each assignment as many times as you like but that the system only saves your last submission.

Collaboration

You may collaborate with other students currently enrolled in the course, and I encourage you to do so. However, you must:

1. **Write your own code.** You may share pseudocode with your collaborators, but you should perform all final implementations on your own.
2. **Compose your written answers independently.** You may discuss questions with your fellow classmates, but you must compose your final responses on your own.
3. **List the people you worked with for each notebook.** Failure to report your collaborators may result in a zero for the entire problem set.

In general, you should not give or receive aid that makes the assigned tasks significantly easier, and should document any significant help that you do receive.

Course Materials

The course can be successfully completed by attending class, reviewing the lecture slides, and carefully reading the instructions in the programming assignments. In addition, the following textbooks are recommended.

Optional Texts: Ian Goodfellow, Yoshua Bengio, and Aaron Courville (2016). *Deep Learning*, MIT Press.

Trevor Hastie, Robert Tibshirani, Jerome Friedman (2008). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer.

Foundations: Jim Hefferon (2017). *Linear Algebra*. http://joshua.smcvt.edu/linearalgebra/book.pdf.

Kevin P. Murphy (2012). *Machine Learning: A Probabilistic Perspective*, MIT Press.

Course Plan

Each section of the course consists of 12 ninety-minute lectures and 5 three-hour labs. Each lab will be devoted to a programming assignment, so you are encouraged to bring a laptop with a working browser and wifi. A portion of each lab may be used as spill-over for lecture material or software demonstrations necessary to complete that week's assignment, but the bulk of the time will be reserved for you to work on your assignments and ask us questions.

Lecture	Lecture Topic	Lab	Lab Topic
1	Welcome to the Course!		
2	Univariate Linear Regression, Gradient Descent for One Feature	1	Implementation of Gradient Descent
3	Multivariate Linear Regression, Gradient Descent for Multiple Features	2	Matrices, arrays, and vectorization
4	Classification, Logistic Regression, BFGS		
5	Deep Learning	3	Bias-variance trade-off, Dropout Regularization
6			Introduction to Keras and TensorFlow
7	Unsupervised Learning: Clustering		
8	K-means Algorithm, optimization objective, distortion cost function	4	In-class implementation of K-means
9	Dimension Reduction: SVG Autoencoders	5	Model Comparison: two-layer encoders/decoders versus several-layer encoders/decoders
10	Deep Autoencoders		
11	Industry Application: Guest Speaker		
12	Socially Responsible Artificial Intelligence		

APRIL

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
1	2	3	4	5	6	7
8	9	10	11	12	13	14
	Lecture 1&2 M	Lab 1M				Lecture 1&2 D
15	16	17	18	19	20	21
	Lecture 3&4 M	Lab 2M		Lab 1D	Lecture 3&4 D	
22	23	24	25	26	27	28
	Lecture 5&6 M	Lab 3M		Lab 2D	Lecture 5&6 D	Lab 3D
29	30					

MAY

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
		1	2	3	4	5
6	7	8	9	10	11	12
	Lecture 7&8 M	Lab 4M			Lecture 7&8 D	Lab 4D
13	14	15	16	17	18	19
	Lecture 9&10 M	Lab 5M		Lecture 9&10 D	Lab 5D	
20	21	22	23	24	25	26
			Lecture 11&12 D	Lecture 11&12 M		
27	28	29	30	31		