Executive training

EU Supervisory Digital Finance Academy (EU-SDFA)
2nd Advanced Training Week

Artificial Intelligence and Machine Learning for SupTech
13-16 March 2023

w: eusdfa.eui.eu • m: sdfa@eui.eu

European University Institute – Villa Schifanoia
Via Boccaccio 121, 50133, Firenze

Organisers:
Florence School of Banking & Finance
in cooperation with the European Supervisory Authorities (ESAs)
and the European Commission (DG REFORM and DG FISMA)

Programme

Monday – 13 March 2023

13.30 – 14.00 **Welcome remarks and EU-SDFA presentation, followed by a tour de table**

Part one

Iman Van Lelyveld (*Vrije Universiteit Amsterdam and De Nederlandsche Bank*),
Michiel Nijhuis (*De Nederlandsche Bank*)

14.00 – 14.45 **Lecture 1. Course overview**

- Why is this course relevant?
- What can you expect?
- What we will cover

14.45 – 15.45 **Tutorial 1: How to read data and use sklearn?**

- Getting started with Python and data manipulation.
- How is this different from Excel?
- Read the data and get to know it.
- Introduction to sklearn: where to find the buttons

15.45 – 16.15 *Coffee break*

16.15 – 17.00 **Lecture 2: Introduction to Machine Learning (ML)**

- What is ML? What is ML applied to?
- Linear regression from the ML lens.
- The outlines of the ML approach
  - Supervised vs. unsupervised learning
  - Hyperparameters and how to select them
  - Gradient descent and grid search

17.00 – 18.00 **Tutorial 2: Regressions versus Classifiers**

- Logit as a statistical model vs ML model
- How to find the optimal (hyper)parameters
- A different classifier: Support vector machines
    - Different types of kernels
    - First glimpse: Dangers of overfitting
    - Evaluating performance

18.00 – 19.00 **Cocktail**

19.00 *Shuttle bus from Villa Schifanoia to Hotel San Gallo Palace*

Tuesday – 14 March 2023

8.45 *Shuttle bus from Hotel San Gallo Palace to Villa Schifanoia*

9.15 – 10.00 **Lecture 3: Machine Learning – the basics**

- Importance of pre-processing your data
- Building up to the workhorse classifier: the logit model
- When is a classifier doing a "good" job?
- Confusion matrix, Receiver Operator Characteristic (ROC)
- What are overfitting, bias and variance?

10.00 – 11.00 **Tutorial 3: Data pre-processing and assessing model performance**

- How to pre-process: standardize your data
- Pros and cons of standardization
- Working with the confusion matrix
    - What if costs are not symmetric?
    - The trade-off between precision and recall

11.00 – 11.30 *Coffee break*

11.30 – 12.15 **Lecture 4: Fighting the curse of dimensionality**

- How to reduce dimensionality?
    - K-Nearest Neighbours (KNN)
    - Principal Components Analysis (PCA)
- Feature selection and regularization
    - How to select the most important features?
    - Examples: RIDGE, LASSO, Elastic net
- Is a "good" model always good? What is external validity?
- Cross-validation and holdouts

12.15 – 13.15 **Tutorial 4: Cross-validation applied to LASSO variable selection**

- Looking closer at cross-validation (CV) and holdouts
- K-fold, Leave-one-out, stratified CV
- Splitting your data into training and testing samples
- How to use CV to tune a LASSO model

w: eusdfa.eui.eu • m: sdfa@eui.eu

13.15 – 14.25 *Lunch break*

14.25 – 14.30 *Group picture*

14.30 – 15.15 **Lecture 5: Improving weak learning**

- How to grow a decision tree? How to split?
    - Purity measures
- Can Ensemble Classifiers improve weak learners?
    - Bagging and boosting
    - Examples: AdaBoost, XGBoost

15.15 – 16.15 **Tutorial 5: Decision trees and random forests**

- Growing your own decision tree
- How deep? How many splits? How big are the leaves?
- From trees to random forests
- Comparing performance with the confusion matrix

16.15 – 16.45 *Coffee break*

16.45 – 17.30 **Lecture 6: Unsupervised learning and clustering**

- Supervised versus unsupervised learning
- What can we do with unsupervised learners?
- K-means, t-SNE, DBSCAN, Gaussian mixtures

17.30 – 18.30 **Tutorial 6: Finding clusters and neighbours**

- Implementing K-means and DBSCAN
- Hierarchical clustering: Bottom-up or Top-down?
- Visual inspection of results

18.30 *Shuttle bus from Villa Schifanoia to Hotel San Gallo Palace*

19.00 – 20.30 *Guided tour on the history of banking and finance in Florence (Meeting point at Hotel San Gallo)*

Wednesday – 15 March 2023

8.45 *Shuttle bus from Hotel San Gallo to Villa Schifanoia*

9.15 – 10.00 **Lecture 7: Natural Language Processing (NLP)**

- What are the main approaches in textual analysis?
- Going beyond simple word counts
- How to extract market sentiment?

10.00 – 11.00 **Tutorial 7: NLTK and sentiment analysis**

- Constructing a bag of words
- Classifying sentiments (positive/negative)
- Example with financial news data

11.00 – 11.30 *Coffee break*

w: eusdfa.eui.eu • m: sdfa@eui.eu

w: eusdfa.eui.eu • m: sdfa@eui.eu

11.30 – 12.15 **Lecture 8: Explainability**

- How can we open the black box
- Explain it like I'm a 5 year old (ELI5) and Shapley values

12.15 – 13.15 **From practice to policy: AI ACT supervisor implication**

- Discuss some things that can go wrong
    - Survivorship bias, input errors and deceit
    - Fairness and discrimination
- What is the reaction of authorities?

13.15 – 14.30 *Lunch break*

Part Two

14.30 – 16.00 **Session 1: NCAs use case of Suptech applications**

Michiel Nijhuis *(De Nederlandsche Bank),* **Hans Gmasz** *(ECB),* **Oliver Giudice** *(Banca d'Italia)*

16.00 – 16-30 *Coffee break*

16.30 – 18.00 **Session 2: Future trends, challenges, and collaborations in Suptech: ESAs prospective**

Miguel Caballero *(EIOPA),* **Vaidotas Tamulenas** *(EBA),* **Giulio Bagattini** *(ESMA)*

18.30 – 19.00 *Cocktail*

19.00 – 19.45 *Keynote Speech by Jermy Prenio (BIS)*

19.45 – 21.00 *Dinner*

21.00 *Shuttle bus from Villa Schifanoia to Hotel San Gallo Palace*

Thursday – 16 March 2023

8.45 *Shuttle bus from Hotel San Gallo Palace to Badia Fiesolana*

9.15 – 10.45 **Session 3: Technology application design and prototyping**

Robert Binder (Regnology)

- Agile prototyping and applications development
- Lifecycle management application
- Lean procurement

10.45 – 11.15 *Coffee break*

11.15 – 12.45 **Session 4: Industry dialogue: future market development**
Mathias Strand (Anch.AI), Robert Binder (Regnology), Lubos Pernis (FNA)

12.45 – 13.00 **Closing Remarks**

13.00 *Light lunch*