COURSE OUTLINE

1. GENERAL

SCHOOL	ECONOMIC SCIENCES		
DEPARTMENT	ECONOMICS		
LEVEL OF STUDY	*Postgraduate*		
COURSE UNIT CODE		**SEMESTER**	2nd
COURSE TITLE	MACHINE LEARNING AND AI FOR BUSINESS		

COURSEWORK BREAKDOWN	**TEACHING WEEKLY HOURS**	**ECTS Credits**
Lectures	2	5

COURSE UNIT TYPE	Compulsory
PREREQUISITES	NO
LANGUAGE OF INSTRUCTION/EXAMS:	English
COURSE DELIVERED TO ERASMUS STUDENTS	NO
MODULE WEB PAGE (URL)	

2. LEARNING OUTCOMES

Learning Outcomes

Recent years have witnessed an unprecedented availability of information on social, economic, and health-related phenomena. Researchers, practitioners, and policymakers have nowadays access to huge datasets (the so-called “Big Data”) on people, companies and institutions, web and mobile devices, satellites, etc., at increasing speed and detail.

Machine learning is a relatively new approach to data analytics, which places itself in the intersection between statistics, computer science, and artificial intelligence. Its primary objective is that of *turning information into knowledge and value* by “letting the data speak”. To this purpose, machine learning limits prior assumptions on data structure, and relies on a *model-free* philosophy supporting algorithm development, computational procedures, and graphical inspection more than tight assumptions, algebraic development, and analytical solutions. Computationally unfeasible few years ago, machine learning is a product of the computer’s era, of today machines’ computing power and ability to learn, of hardware development, and continuous software upgrading.

This course is a primer to machine learning techniques using Stata, Python, and R. These software own today various packages to perform machine learning which are sometimes unknown to many users. This course fills this gap by making participants familiar with (and knowledgeable of) these ML software potential to draw knowledge and value form row, large, and possibly noisy data. The teaching approach will be mainly based on the graphical language and intuition more than on algebra. The training will make use of instructional as well as real-world examples, and will balance evenly theory and practical sessions.

General Skills

After the course, participants are expected to have an improved understanding of Stata potential to perform marching learning, thus becoming able to master research tasks including, among others: (i) factor-importance detection, (ii) signal-from-noise extraction, (iii) correct model specification, (iv) model-free classification, both from a data-mining and a causal perspective.

3. COURSE CONTENTS

PROGRAM

1. The basics of Machine Learning

Machine Learning: definition, rational, usefulness

Supervised vs. unsupervised learning
Regression vs. classification problems
Inference vs. prediction
Sampling vs. specification error

Coping with the fundamental non-identifiability of E(y|x)

Parametric vs. non-parametric models
The trade-off between prediction accuracy and model interpretability

Goodness-of-fit measures

Measuring the quality of fit: in-sample vs. out-of-sample prediction power
The bias-variance trade-off and the Mean Square Error (MSE) minimization
Training vs. test mean square error
The information criteria approach

Machine Learning and Artificial Intelligence

The Stata/Python integration: an overview

2. Resampling and validation methods

Estimating training and test error

Validation

The validation set approach
Training and test mean square error

Cross-Validation

K-fold cross-validation
Leave-one-out cross-validation

Bootstrap

The bootstrap algorithm
Bootstrap vs. cross-validation for validation purposes

3. Model Selection and regularization

Model selection as a correct specification procedure

The information criteria approach

Subset Selection

Best subset selection
Backward stepwise selection
Forward stepwise Selection

Shrinkage Methods

Lasso and Ridge, and Elastic regression
Adaptive Lasso
Information criteria and cross validation for Lasso

Software implementation

4. Discriminant analysis and nearest-neighbor classification
The classification setting
Bayes optimal classifier and decision boundary
Misclassification error rate
Discriminant analysis
    Linear and quadratic discriminant analysis
    Naive Bayes classifier
The K-nearest neighbors classifier
Software implementation

5. Nonparametric regression
Beyond parametric models: an overview
Local, semi-global, and global approaches
Local methods
    Kernel-based regression
    Nearest-neighbor regression
Semi-global methods
    Constant step-function
    Piecewise polynomials
    Spline regression
Global methods
    Polynomial and series estimators
    Partially linear models
    Generalized additive models
Software implementation

6. Tree-based regression
Regression and classification trees
    Growing a tree via recursive binary splitting
    Optimal tree pruning via cross-validation
Tree-based ensemble methods
    Bagging, Random Forests, and Boosting
Software implementation

7. Neural networks
The neural network model
    neurons, hidden layers, and multi-outcomes
Training a neural networks
    Back-propagation via gradient descent
    Fitting with high dimensional data
    Fitting remarks
Cross-validating neural network hyperparameters
Software implementation

7. Neural networks
The neural network model
    Neurons, hidden layers, and multi-outcomes

8. ROC Curve
Introduction to Binary Classification and Performance Metrics
The Receiver Operating Characteristic (ROC) curve
    Comparing Classifiers with the ROC Curve and AUC
Software implementation

4. TEACHING METHODS - ASSESSMENT

<table>
<tr><td><b>MODE OF DELIVERY</b></td><td>online</td></tr>
<tr><td><b>USE OF INFORMATION AND COMMUNICATION TECHNOLOGY</b></td><td>Dynamic powerpoint transparencies
e-class support
Communication via e-mail and course discussion group</td></tr>
<tr><td><b>TEACHING METHODS</b></td><td>
<table>
<tr><th><i>Method description</i></th><th><i>Semester Workload</i></th></tr>
<tr><td>lectures</td><td>26</td></tr>
<tr><td>Individual Assignments</td><td>34</td></tr>
<tr><td>Self study</td><td>65</td></tr>
<tr><td><b><i>Course total
(24 hours of work load per credit)</i></b></td><td><b><i>125</i></b></td></tr>
</table>
</td></tr>
<tr><td><b>ASSESSMENT METHODS</b></td><td>I. Final examination (50%)
I. Individual Assignments (50%)</td></tr>
</table>

5. RESOURCES

Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani (2013), *An Introduction to Statistical Learning with Applications in R*, Springer, New York, 2013.

Cerulli, G. (2023), Fundamentals of Supervised Machine Learning: With Applications in Python, R, and Stata, Springer.

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