UCLouvain

ldats2310

2023

Data science for insurance and finance

3.00 credits	15.0 h	Q1

Teacher(s)	Hainaut Donatien ;
Language :	English
Place of the course	Louvain-la-Neuve
Prerequisites	Concepts and tools equivalent to those taught in teaching units LSTAT2020 Logiciels et programmation statistique de base LSTAT2120 Linear models LSTAT2100 Modèles linéaires généralisés et données discrêtes A first course in probability and statistics is required e.g. : LBIR1203 Probabilités et statistiques I and LBIR1304 Probabilités et statistiques II (or equivalent modules).
Main themes	This module aims to introduce recent developments in the field of statistical learning, applied to the insurance and financial sectors. Statistical methods are used in the insurance industry to assess the risk profile of an insured. This profile presents two sides: one is the frequency of claims and the other is the size of the claim caused by the insured. Both aspects are studied carefully by insurers so as to propose the best price for an insurance coverage. In the financial industry, advanced statistical methods are needed to evaluate the credit risk of a lender. As for an insurance contract, this risk has two sides. The first one is the probability that the lender will not repay is debt (the default risk). The second aspect is the size of the loss when the lender do not redeem is loan. This module present the common tools to study these risks: generalized linear models, additive models, Regression/classification trees. Some new aspects will also be developed among them we quote shrinkage methods (Lasso, Ridge) and random forests that reveals to be powerful tools to explore massive data.
Learning outcomes	**At the end of this learning unit, the student is able to :** 1 At the end of this course, students will be able: • To explain and motivate the choice of a statistical method to analyze insurance or financial data • To use Generalized Linear and Additive models to propose a grid of insurance premium or to propose a model to evaluate the default risk of a counterparty • To use Regression Tree and random forest on insurance or credit datasets. • To adapt the previous cited methods to include constraints of sparsity in the calibration (Lasso Ridge) • To understand the interests of bootstrapping methods and to implement them.
Evaluation methods	The Evaluation is based on an individual report in which methods seen during lectures are apply to a real dataset. Notice that the professor keeps the right to orally question students on the content of their work.
Teaching methods	• Lectures based on readings • Programs in Python (keras & tensorflow) • Case studies
Content	This course focuses on applications of deep learning to insurance and finance. The first part introduces neural networks as tools of regression. We review the various calibration algorithms and loss functions for actuarial or financial applications. We also present the techniques for managing high dimensional datasets (penalization, embedding, randomization) and methods of bias regularization. We next see the tools for interpreting the output of neural networks: global (PDP, ICE) and local (LIME, SHAP) methods. The second part of this lecture is devoted to unsupervised learning, and in particular to neural auto-encoders for data analytics. The last section introduces time-series forecasting with recurrent, long short-term and convolutional neural networks.
Inline resources	Moodle website
Bibliography	Slides available on moodle are based on the following reference: Denuit M., Trufin J. , Hainaut D. 2019. Effective statistical learning III : neural networks and extensions. Springer actuarial lectures notes.
Faculty or entity in charge	LSBA

Programmes containing this learning unit (UE)

Program title	Acronym	Credits	Prerequisite	Learning outcomes
Master [120] in Data Science : Statistic	DATS2M	3		
Master [120] in Actuarial Science	ACTU2M	3		
Master [120] in Statistics: General	STAT2M	3		
Certificat d'université : Statistique et science des données (15/30 crédits)	STAT2FC	3		