

Course Specifications

Valid in the academic year 2018-2019

Course size (nominal values; actual values may depend on programme)
Credits 5.0 Study time 150 h Contact hrs 60.0 h

Course offerings and teaching methods in academic year 2018-2019

A (semester 1)	English	self-reliant study activities	5.0 h
		guided self-study	10.0 h
		seminar: practical PC room classes	15.0 h
		lecture	30.0 h

Lecturers in academic year 2018-2019

Benoit, Dries EB23 lecturer-in-charge

Offered in the following programmes in 2018-2019

	crdts	offering
Bachelor of Science in Business Engineering	5	A
Exchange programme in Economics and Business Administration	5	A
Preparatory Course Master of Science in Business Engineering	5	A

Teaching languages

English

Keywords

Univariate and multivariate regression analysis, classification, resampling methods, model selection and regularization, tree-based methods, unsupervised learning, R

Position of the course

Business processes have been digitalized at high pace in recent decades. This lead to a tremendous increase in information stored in databases. With datamining, analysts try to find relevant patterns in this huge source of information that help management in optimizing their decisions. Students will learn how to extract relevant information from databases. Students will learn how to apply a broad set of methods in a correct way to solve real-life business problems. They will also learn the skills of how to interpret the results from such analyses in a meaningful way. All exercises will be solved using the statistical programming language R.
This course builds on the concepts introduced in the course "Statistics I".

Contents

- * Relationship between bias and variance and model complexity
- * Regression analysis
 - simple linear regression
 - multiple regression
 - qualitative predictors
- * Classification
 - logistic regression analysis
 - linear discriminant analysis
 - k-nearest neighbors
- * Resampling methods
 - crossvalidation
 - bootstrap
- * Model selection and regularization
 - variable selection
 - shrinkage methods
 - dimension reduction
- * Tree-based methods

- decision trees
- bagging and boosting
- random forests
- * Support vector machines
- kernels
- * Neural networks
- representation
- parameter estimation
- * Unsupervised learning
- principal component analysis
- clustering

Initial competences

Students need to have a thorough understanding of the basic statistical principles and concepts as treated in the course "Statistics I", <https://studiegids.ugent.be/2018/EN/studiefiches/F000846.pdf>

Final competences

- 1 Apply datamining methods in a correct way, using the programming language R
- 2 Interpret the results in a correct way and communicate these to a non-technical audience
- 3 Choose a well-suited method to tackle a specific business problem, knowing what the advantages and disadvantages are of the chosen approach.

Conditions for credit contract

Access to this course unit via a credit contract is determined after successful competences assessment

Conditions for exam contract

This course unit cannot be taken via an exam contract

Teaching methods

Guided self-study, lecture, self-reliant study activities, seminar: practical PC room classes

Learning materials and price

Course materials will be made available on Minerva.
Course book: see References

References

James, G., Witten, D., Hastie, T. & Tibshirani, R. (2015) An Introduction to Statistical Learning with Applications in R, Springer.

Course content-related study coaching

Students can rely on the teaching assistants and teacher of the course. Communications and interaction through the Minerva platform (group discussions, exercises, ...). Example exam questions will be made available.

Evaluation methods

end-of-term evaluation

Examination methods in case of periodic evaluation during the first examination period

Written examination

Examination methods in case of periodic evaluation during the second examination period

Written examination

Examination methods in case of permanent evaluation

Possibilities of retake in case of permanent evaluation

not applicable

Calculation of the examination mark