

Course Specifications

Valid as from the academic year 2019-2020

Predictive and Prescriptive Analytics (F000801)

Course size (nominal values; actual values may depend on programme)

Credits 6.0 Study time 180 h Contact hrs 45.0 h

Course offerings and teaching methods in academic year 2020-2021

A (semester 2)	English	lecture	10.0 h
		seminar: practical PC room classes	25.0 h
		seminar: coached	5.0 h
		exercises	
		group work	5.0 h

Lecturers in academic year 2020-2021

Van den Poel, Dirk EB23 lecturer-in-charge

Offered in the following programmes in 2020-2021

	crdts	offering
Master of Science in Business Engineering (main subject Data Analytics)	6	A
Master of Science in Business Engineering (main subject Operations Management)	6	A
Exchange programme in Economics and Business Administration	6	A

Teaching languages

English

Keywords

Advanced predictive models, advanced optimization and simulation methods, data mining, machine learning, artificial neural networks, reinforcement learning, ensembles, computational statistics, R

Position of the course

This course builds on the modeling techniques seen in "Analytical Customer Relationship Management" and "Pricing and Revenue Management" by means of more advanced analysis techniques as well as by using the open-source scientific programming language (R). Flipped-classroom teaching will be used in this course. Students will need to watch video's or other material to prepare for class.

Contents

Following an introduction to predictive and prescriptive analytics, the various substantive areas are explored along with the available supporting software: • Refining the model-building process for predictive analytics (with specific emphasis on the performance evaluation of models) using the CRISP-DM methodology o Business understanding o Data understanding o Data preparation (including techniques for interpolation (splines)) o Modeling o Evaluation & deployment • Applying advanced statistical as well as machine learning techniques (including techniques from computational statistics such as Bayesian analysis and decision trees) to solve different business problems: forecasting, pricing, credit scoring, marketing mix modeling, fraud detection, acquisition, retention,... • Customer valuation (life-time value calculations) • Using advanced operations research techniques (optimization and simulation) to translate the predictive outcomes into actionable management information • Using the R programming language and environment for: o Data preparation (database connectivity, statistics) o Model development o Data visualization: representing the analysis results in 2D or 3D

More specifically, this course applies advanced statistical methods and machine learning algorithms (including both supervised and unsupervised learning techniques) to solve business and marketing related problems. The focus lies in the understanding

and implementation of several supervised and unsupervised learning techniques. For every modeling technique, we will focus on the trade-off between model interpretability and accuracy and their impact on the bias-variance problem. The following supervised learning techniques are discussed:

- (Non-)Linear regression
- Logistic regression
- Decision Trees
- Neural Networks
- Bagging of trees
- Adaboosting
- Random forest
- Hybrid ensembles
- Forecasting models such as arima, exponential smoothing, ...

Next to supervised learning, several unsupervised learning algorithms will be taught:

- Hierarchical clustering
- K-means clustering
- Spectral clustering
- Principal component analysis

The evaluation of all these techniques is crucial. Thereby we will rely on several performance measures such AUC, PCC for classification and R^2 , RMSEP for regression.

The second part of the course 'prescriptive analytics' focuses on a mix of mathematical techniques that given a set of complex objectives and constraints can be used to increase business performance. The main focus will be on decision making under uncertainty using auxiliary data sources. We cover several state-of-the-art exact and meta-heuristic optimization methods. The exact optimization methods include:

- Linear programming
- Integer programming

The meta-heuristics exists of a mix of local search and population-based search methods:

- Local search algorithms:
 - Gradient descent
 - Simulated Annealing
 - Tabu search
- Population-based search algorithms:
 - Genetic algorithm
 - Particle swarm optimization

Reinforcement Learning will be used to introduce AI into prescriptive analytics. The exploration-exploitation trade-off will be explained. Automatic learning of optimal policies will be discussed. Reinforcement learning refers to the problem of an agent that aims to learn optimal behavior through trial-and-error interactions (exploration) with a dynamic environment. All algorithms for reinforcement learning share the property that the feedback of the agent is restricted to a reward signal that indicates how well the agent is behaving. In contrast to supervised machine learning methods, any instruction concerning how to improve its behavior is absent. Thus, the goal and challenge of reinforcement learning is to improve the behavior of an agent given only this limited type of feedback.

These optimization algorithms will be taught using several real-life case applications e. g. demand-supply matching, advertising reach maximization, traveling salesmen problem and the location-allocation problem. Special attention will be given to the multi-armed bandit problem. Students will be incentivized to do sensitivity analysis about their solution. Special focus will be put on the linkage between predictive and prescriptive analytics. Students will get practical examples of how the results of their predictive analysis can be integrated and evaluated in their optimization models.

Initial competences

This course builds on the final competences of the courses "Analytical Customer Relationship Management", "Pricing and Revenue Management" and "Marketing Information Systems/Database Marketing" (SQL), which in turn implies knowledge about intermediate statistics and econometrics.

Final competences

Mastering advanced statistical, optimization, and data mining techniques using the R programming language. Being able to apply them effectively in real-life projects.

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

Group work, lecture, seminar: coached exercises, seminar: practical PC room classes

Extra information on the teaching methods

Ex cathedra sessions as well as active class discussions of the different techniques and models with interactive exercises in the PC room.

Learning materials and price

Handbook (free online book):

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. *An Introduction to Statistical Learning: with Applications in R*, 1st ed. 2013. Corr. 4th printing 2014 edition. ed. Springer, New York.

References

Sutton R.S. & Barto A.G. (2018), Reinforcement Learning: An Introduction, 2nd Edition, The MIT Press.

Course content-related study coaching

Numerous exercises are being solved during sessions. In addition, assignments (to be solved in teams) are handed out. Students will receive coaching in the process of solving the assignments and feedback afterwards (collectively, by team and individually). After tests about the programming language R as well as their insights into using advanced aCRM and advanced statistics, students will receive collective as well as individual feedback & coaching.

Evaluation methods

continuous assessment

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

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

Examination methods in case of permanent evaluation

Written examination, oral examination, assignment

Possibilities of retake in case of permanent evaluation

examination during the second examination period is possible in modified form

Calculation of the examination mark

The final score consists of:

50% Written exam to determine to what extent the student mastered: • the principles of predictive and prescriptive analytics, • the higher programming language R, • the principles of advanced analysis techniques, • the use of R to solve non-trivial business problems by means of predictive / prescriptive models.

50% Paper with oral defense as part of a group work where the students will solve a real business problem using analytical techniques (potentially corrected by peer assessment).

Teams will be randomly selected during class to present their solutions to assignments. A maximum of two bonus points (of 20) can be earned in this way.