



Voorspellende en voorschrijvende analytische methoden (F000801)

Wegens Covid19 kan mogelijk afgeweken worden van de onderwijs- en evaluatievormen. Dergelijke afwijkingen zullen via Ufora worden gecommuniceerd.

Cursusomvang (nominale waarden; effectieve waarden kunnen verschillen per opleiding)

Studiepunten 6.0 Studietijd 180 u Contacturen 45.0 u

Aanbodsessies en werkvormen in academiejaar 2020-2021

A (semester 2)	Engels	Gent		
			hoorcollege	10.0 u
			werkcollege: PC- klasoefeningen	25.0 u
			werkcollege: geleide oefeningen	5.0 u
			groepswork	5.0 u
			online hoorcollege	0.0 u
			online werkcollege: geleide oefeningen	0.0 u
			online werkcollege: PC-klasoefeningen	0.0 u

Lesgevers in academiejaar 2020-2021

Bogaert, Matthias EB23 Verantwoordelijk lesgever

Aangeboden in onderstaande opleidingen in 2020-2021

	stptn	aanbodsessie
Master of Science in Business Engineering (afstudeerrichting Data Analytics)	6	A
Master of Science in Business Engineering (afstudeerrichting Operations Management)	6	A
Uitwisselingsprogramma Economie en Bedrijfskunde	6	A

Onderwijstalen

Engels

Trefwoorden

Geavanceerde marketingmodellen, geavanceerde optimalisatie- en simulatiemethoden, data mining, machine learning, artificiële neural netwerken, reinforcement learning, ensembles, computationele statistiek, R

Situering

Deze cursus omvat een verdere uitdieping van de cursussen 'analytische klantenbeheerssystemen' en 'prijszetting en omzetbeheer', m.a.w. van meer geavanceerde analysetechnieken alsook het aanleren van een wetenschappelijke programmeertaal R. In deze cursus wordt de "flipped classroom" onderwijsmethode gebruikt. Studenten dienen dus video's of ander materiaal door te nemen vooraleer ze naar de les komen.

Inhoud

Na een inleiding inzake voorspellende en voorschrijvende modellen, worden de diverse substantieve deelgebieden belicht, en wordt aangegeven welke modellen er kunnen gehanteerd worden. Aldus worden de volgende deeldomeinen belicht : • Verder verfijnen van het modelleerproces voor voorspellende analytische methoden (hierbij gaat vooral aandacht naar de evaluatie van de performantie van modellen) gebruik makende van de CRISP-DM methodologie. o Inzicht in het bedrijfsprobleem o Inzicht in de data o Voorbereiding van de data (inclusief gebruik van interpolatie (splines)) o Modelleren o Evaluatie en uitrol • Toepassen van meer geavanceerde statistische en machine learning technieken (waaronder technieken vanuit de computationele statistiek, zoals Bayesiaanse analyse en beslissingsbomen) om uiteenlopende bedrijfsgerelateerde problemen op te lossen zoals prijszetting, marketing mix

modellering, fraude detectie, acquisitie, retentie, forecasting, ... • Waardering van klanten ('life-time value' berekeningen) • Vertalen van de resultaten van een voorspellend model naar bruikbare management informatie door middel van technieken uit het operationele onderzoek (optimalisatie en simulatie). • Het inzetten van de open-source R-programmeertaal voor: o Data voorbereiding (connecteren met een database, statistiek) o Model ontwikkeling o Data visualisatie: het weergeven van de analyseresultaten (2D en 3D)

Meer in het bijzonder wordt in dit vak statistische en "machine learning" technieken (zowel voor "supervised" en "unsupervised" leermethoden) aangewend om bedrijfsrelevante problemen op te lossen. Volgende "supervised learning" technieken worden besproken:

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

Volgende "unsupervised learning" methoden worden besproken:

- Hierarchical clustering
- K-means clustering
- Spectral clustering
- Principale componenten analyse

Volgende performantiematen worden behandeld: AUC, PCC voor classificatie en R^2 , RMSEP voor regressie.

Het tweede deel van de cursus 'voorschrijvende analytische methoden' focust op het aanleren van wiskundige technieken die gegeven een complexe doelfunctie en een aantal beperkingen gebruikt kunnen worden om bedrijfsresultaten te verbeteren. Hierbij focussen we vooral op beslissingen nemen onder onzekerheid met extra databronnen. Verschillende exacte optimalisatie methodes en meta-heuristieken worden gebruikt. De exacte optimalisatie methodes omvatten:

- Linear programming
- Integer programming

De meta-heuristieken bestaan uit lokale zoekalgoritmes en zoekalgoritmes gebaseerd op een populatie:

- Lokale zoekalgoritmes
 - Gradient descent
 - Simulated Annealing
 - Tabu search
- Zoekalgoritmes gebaseerd op een populatie:
 - Genetic algorithm
 - Particle swarm optimization

"Reinforcement learning" wordt gebruikt om tot betere beslissingen te komen en introduceert AI (artificiële intelligentie) in voorschrijvende analytics. De exploratie-exploitatie trade-off wordt uitgelegd. RL probeert automatisch "policies" af te leiden, die tot verbeterde beslissingen moeten leiden. Reinforcement learning verwijst naar het probleem waarbij een "agent" optimaal gedrag tracht te leren door "trial-and-error" interacties (exploratie) met een dynamische (d.w.z. veranderende) omgeving. Alle algorithmen voor reinforcement learning hebben met elkaar gemeen dat de feedback naar de agent toe beperkt is tot een zogenaamd "reward signal", dat een indicatie geeft hoe goed de agent zich gedraagt. In tegenstelling tot "supervised machine learning" methoden wordt niet aangegeven hoe de agent zijn/haar gedrag dient aan te passen. Het doel en de uitdaging voor reinforcement learning is het gedrag van de agent te verbeteren met slechts een beperkte vorm van feedback.

Deze optimalisatiealgoritmes zullen worden onderricht aan de hand van praktische voorbeelden zoals het afstellen van vraag en aanbod, het maximaliseren van het bereik van advertenties, het handelsreizigersprobleem en het locatie-allocatie probleem. Speciale aandacht zal worden besteed aan het "multi-armed bandit" probleem. Studenten zullen worden aangemoedigd om sensitiviteitsanalyse te doen van hun oplossing. Extra aandacht zal eveneens worden gevestigd aan het linken van voorspellende en voorschrijvende modellen. De studenten zullen in de praktische voorbeelden zien hoe voorspellingen kunnen geïntegreerd en geëvalueerd worden in hun optimalisatiemodellen.

Begincompetenties

Sluiten aan op de eindtermen van de cursus analytische klantenbeheerssystemen alsook van marketinginformatiesystemen/database marketing (SQL) en "Prijszetting en Omzetbeheer", die op hun beurt kennis veronderstellen van gevorderde statistiek en econometrie.

Eindcompetenties

Kunnen doorgronden van geavanceerde statistische -, optimalisatie- of dataminingmethoden voor het oplossen van uitdagende bedrijfsproblemen. Een uitgebreide kennis van de R-programmeertaal en deze ook effectief kunnen inzetten in reële projecten.

Creditcontractvoorwaarde

Toelating tot dit opleidingsonderdeel via creditcontract is mogelijk mits gunstige beoordeling van de competenties

Examencontractvoorwaarde

Dit opleidingsonderdeel kan niet via examencontract gevolgd worden

Didactische werkvormen

Groepswerk, hoorcollege, werkcollege: geleide oefeningen, werkcollege: PC-klasoefeningen, online hoorcollege, online werkcollege: geleide oefeningen, online werkcollege: PC-klasoefeningen

Toelichtingen bij de didactische werkvormen

Ex cathedra uiteenzetting alsook actieve bespreking in groepsverband van de diverse technieken en modellen met interactieve oefeningen in het PC-lokaal.

Leermateriaal

Handboek (gratis online boek):

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.

Referenties

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

Vakinhoudelijke studiebegeleiding

Talrijke oefeningen worden tijdens de sessies behandeld. Bovendien worden taken opgegeven, die in groepsverband dienen te worden opgelost. Hierbij krijgen de studenten uitgebreide (studie)begeleiding en na afloop feedback (met de volledige groep, in team en individueel). Na de periodieke tests over de programmeertaal R wordt collectieve alsook individuele feedback gegeven.

Evaluatiemomenten

periodegebonden en niet-periodegebonden evaluatie

Evaluatievormen bij periodegebonden evaluatie in de eerste examenperiode

Schriftelijk examen

Evaluatievormen bij periodegebonden evaluatie in de tweede examenperiode

Schriftelijk examen

Evaluatievormen bij niet-periodegebonden evaluatie

Mondeling examen, werkstuk, peer-evaluatie

Tweede examenkans in geval van niet-periodegebonden evaluatie

Examen in de tweede examenperiode is enkel mogelijk in gewijzigde vorm

Eindscoreberekening

De totaalscore bestaat uit: 50% Schriftelijk, waarbij gepeild wordt naar de mate waarin de student zich: 1 de principes van gevorderde predictieve en prescriptieve technieken eigen heeft gemaakt, 2 in staat is om een analytisch project uit te werken, 3 inzicht heeft in de praktische applicaties van voorschrijvende en voorspellende technieken. 50% presentaties en paper met mondelinge verdediging als deel van een groepswerk waar de studenten een realistisch bedrijfsprobleem oplossen aan de hand van analytische methoden (mogelijk gecorrigeerd door "peer assessment"). Teams zullen op lukrake wijze worden geselecteerd om hun resultaten voor groepswerk te presenteren. Tijdens de verdediging zal ook gepeild worden naar hun kennis van de programmeertaal R.