



Big Data Science (E018210)

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 60.0 u

Aanbodsessies en werkvormen in academiejaar 2020-2021

| | | | | |
|----------------|--------|------|-------------|--------|
| A (semester 2) | Engels | Gent | hoorcollege | 30.0 u |
| | | | practicum | 30.0 u |

Lesgevers in academiejaar 2020-2021

| | | |
|------------------|------|---------------------------|
| Mannens, Erik | TW06 | Verantwoordelijk lesgever |
| Dimou, Anastasia | TW06 | Medelesgever |
| Verborgh, Ruben | TW06 | Medelesgever |

Aangeboden in onderstaande opleidingen in 2020-2021

| | stptn | aanbodsessie |
|--|-------|--------------|
| Brugprogramma Master of Science in Bioinformatics (afstudeerrichting Engineering) | 6 | A |
| Brugprogramma Master of Science in de ingenieurswetenschappen: bedrijfskundige systeemtechnieken en operationeel onderzoek | 6 | A |
| Brugprogramma Master of Science in Industrial Engineering and Operations Research | 6 | A |
| Master of Science in Business Engineering (afstudeerrichting Data Analytics) | 6 | A |
| Master of Science in Bioinformatics (afstudeerrichting Engineering) | 6 | A |
| Master of Science in Business Engineering (afstudeerrichting Operations Management) | 6 | A |
| Master of Science in de ingenieurswetenschappen: bedrijfskundige systeemtechnieken en operationeel onderzoek | 6 | A |
| Master of Science in de ingenieurswetenschappen: computerwetenschappen | 6 | A |
| Master of Science in Computer Science Engineering | 6 | A |
| Master of Science in Industrial Engineering and Operations Research | 6 | A |
| Uitwisselingsprogramma informatica (niveau master) | 6 | A |
| Postgraduate in Internet of Things | 6 | A |

Onderwijstalen

Engels

Trefwoorden

Big Data platformen, interactieve datavisualisaties, semantische webtechnologieën, grootschalige datamining, grootschalige machine learning, streambeheer, streaming algoritmen, aanbevelingssystemen, technologie-ethiek en privacy

Situering

Het belangrijkste doel van deze cursus is om de studenten praktische ervaring te laten opdoen met de belangrijkste concepten van End-to-End Big Data Science. Ze leren hoe ze Big Data kunnen beheren, analyseren en visualiseren en nadenken over de maatschappelijke impact ervan.

Inhoud

- AI Ethiek
 - Ethiek
 - Privacy en GDPR
 - Maatschappelijke uitdagingen

Big Data Systemen

- Big Data beheersystemen
- Batch versus stream AI
- Cloud vs. Edge AI

Big Data Analyse

- Algoritmes op het web (PageRank, Adwords, ...)
- Machine Learning (aanbevelingssystemen, classificatie, regressie, ...)
- Schaalbare datamining (Dimensionality Reduction, Clustering, ...)

Kennisgrafen

- Grafentheorie
- FAIR, Open en LOD gegevens
- Gedecentraliseerde en gefedereerde query's

Streamverwerking

- Stream Management Systemen
- Heterogene stroom-voorverwerking
- Stream datamining

Interactieve datavisualisaties

- Menselijke perceptie
- Ontwerpprincipes en interactie
- Dashboard Raamwerken

Gastcolleges van Belgische Big Data bedrijven

Begincompetenties

- Basis programmeervaardigheden
 - Vooral ervaring met Python, Java en JavaScript is voordelig (maar, met wat extra inspanning, geen noodzaak)
- Elementair begrip van basisgegevensformaten
 - CSV, TSV, etc.
- Lineaire algebra, inleidende cursus statistiek

Eindcompetenties

- 1 De mogelijkheden en beperkingen van Big Data technologieën begrijpen
- 2 De verschillende componenten van Big Data systemen begrijpen
- 3 De industriële toepassingen van Big Data begrijpen
- 4 Big Data componenten combineren in een systeemarchitectuur om aan specifieke productbehoefte te voldoen
- 5 De levenscyclus van Big Data begrijpen
- 6 Big Data opschonen voor productiegebruik
- 7 Visueel en niet-visueel Big Data verkennen
- 8 Interactieve dashboards maken over Big Data
- 9 Schaalbare datamining opzetten en uitvoeren
- 10 Machine learning algoritmen toepassen op Big Data
- 11 Beheer van datasets met meerdere uitdagende dimensies (grootte, formaat, kwaliteit, ...)
- 12 Omgaan met snel-veranderende gegevens via real-time messaging en streamverwerking
- 13 Het verbinden van heterogene gegevens door middel van semantische technologieën
- 14 Het beargumenteren van de ethische en privacy-aspecten van grootschalige gegevensverwerking

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

Hoorcollege, practicum

Toelichtingen bij de didactische werkvormen

Hoorcolleges, practica, project

Leermateriaal

geannoteerde presentaties, artikelen (uitreksels) en hoofdstukken uit tekstboeken (vrij online beschikbaar)

Referenties

- Mining of Massive Datasets, Jure Leskovec, Anand Rajaraman, Jeffrey Ullman, ISBN: 978-1-107-07723-2
- LAMBA Book

Vakinhoudelijke studiebegeleiding

Evaluatiemomenten

periodegebonden en niet-periodegebonden evaluatie

Evaluatievormen bij periodegebonden evaluatie in de eerste examenperiode

Openboekexamen, mondeling examen

Evaluatievormen bij periodegebonden evaluatie in de tweede examenperiode

Openboekexamen, mondeling examen

Evaluatievormen bij niet-periodegebonden evaluatie

Werkstuk, vaardigheidstest, verslag

Tweede examenkans in geval van niet-periodegebonden evaluatie

Examen in de tweede examenperiode is enkel mogelijk in gewijzigde vorm

Toelichtingen bij de evaluatievormen

- Periodieke evaluatie
- mondeling examen met schriftelijke voorbereiding, waarbij het internet en alle cursusmaterialen toegankelijk zijn
- Niet-periodegebonden evaluatie
- beoordeelde labsessierapporten in groepen
- beoordeelde projectrapporten

Eindscoreberekening

De eindscore is de gemiddelde score voor de twee onderdelen (examen en practica). Indien de score voor één onderdeel meer dan 7/20 maar minder dan 10/20 bedraagt, wordt de eindscore begrensd op 9/20.

Indien de score voor één onderdeel 7/20 of minder bedraagt, wordt de eindscore begrensd op 7/20.