Vak: Statistiek 4 — credits: 3

Vakcode	BFVH4STA4
Naam	Statistiek 4
Studiejaar	2023-2024
ECTS credits	3
Taal	Nederlands
Coördinator	T.A. Wassenaar
Werkvormen	Hoorcollege Werkcollege
Toetsen	Statistiek 4 - Computer, eigen organisatie

Leeruitkomsten

De student kan na afloop van deze module:

- matrix/vector berekeningen uitvoeren in R
- gemiddelden, varianties, covarianties en correlaties berekenen door middel van operaties op vectoren
- uitleggen wat de begrippen variantie, covariantie en correlatie betekenen
- gegeven een dataset modellen opstellen voor multipele lineaire regressie en de parameters voor die modellen schatten met R
- gegeven een dataset de kwaliteit van de fit tegen een (multi)lineair model evalueren
- de resultaten van een (multi)lineaire fit grafisch presenteren
- principale componentenanalyse (PCA) uitvoeren
- uitleggen wat principale componenten zijn in abstracte zin en concreet in relatie tot een gegeven dataset
- dimensiereductie toepassen door middel van principale componentenanalyse
- hoogdimensionele datasets visualiseren en presenteren door middel van projectie op eigenvectoren
- afstandsmatrices bepalen voor een gegeven dataset met behulp van de R functie dist met een gefundeerde keuze voor het soort afstand
- hierarchische clustering uitvoeren op een afstandsmatrix met een gefundeerde keuze voor de methode voor clustering
- de resultaten van hierarchische clustering grafisch presenteren
- de resultaten van hierarchische clustering evalueren
- k-means clustering uitvoeren op een gegeven dataset
- de resultaten van k-means clustering grafisch presenteren
- bepalen van het optimaal aantal clusters voor k-means aan de hand van de total within sum-of squares

Inhoud

Statistiek 4 geeft een introductie in het analyseren van onderlinge relaties en structuren in data
De student begrijpt de begrippen covariantie en correlatie en kan lineaire regressie uitvoeren en samenvatten.
De student kan clusteranalyse uitvoeren in R met verschillende methoden en verschillende afstandsmaten.
De student kan principale componenten analyse uitvoeren in R.

Literatuur:
An introduction to R, 2nd ed. (2009), W.N. Venables, D.M. Smith and the R Developement Core Team Network Theory Ltd. ISBN: 0-9546120-8-6
Tutorial R for Beginners
C 1-11 Introduction...Statistical Models in R
Materiaal beschikbaar gesteld door de docent en computeropdrachten beschikbaar op Blackboard Blackboard course thema 10

Opgenomen in opleiding(en)

Bio-Informatica

School(s)

Instituut voor Life Science & Technology

share your talent. move the world.

De ECTS onderwijscatalogus van de Hanzehogeschool Groningen wordt met de grootst mogelijke zorg samengesteld. Het is echter mogelijk dat de inhoud van de catalogus -en de daarin vervatte informatie- verouderd, incompleet of onjuist is. Aan de inhoud van de catalogus kunnen dan ook geen rechten worden ontleend.