

Data Science, de basis

Synthèsecursus van 2 dagen - 14u

Ref : DTX - Prijs 2024 : € 1 950 excl. BTW

Datawetenschap is een belangrijke strategische kwestie voor organisaties en maakt gebruik van wiskundige hulpmiddelen om het gedrag van gegevens bloot te leggen en de gebeurtenissen die ze beschrijven te analyseren. Deze cursus behandelt de grondbeginselen van data science en leert hoe u gegevens analyseert.

PEDAGOGISCHE DOELSTELLINGEN

Na afloop van de opleiding kan de cursist:

De basisprincipes van datawetenschap begrijpen en begrijpen hoe u het proces organiseert

Inzicht in de toepassing van datawetenschap om problemen op te lossen en de beperkingen ervan

Uw vermogen ontwikkelen om cijfers te analyseren en te interpreteren met behulp van grafische weergave

Begrijpen hoe u data science tools gebruikt en modellen ontwikkelt voor professionele doeleinden

De uitdaging aangaan om gegevens te gebruiken in een context van concurrentie en voortdurende verbetering

Inzicht in de organisatie en infrastructuur voor data science-diensten en -projecten

HET PROGRAMMA

laatste update: 04/2024

1) Wat is datawetenschap?

- De basis: big data, data lake, datamining, kunstmatige intelligentie, machine- en deep learning, tekst mining.
- De nieuwe uitdagingen: de opkomst en vermenigvuldiging van nieuwe gegevensbronnen.
- Er moet rekening worden gehouden met heterogeniteit van gegevens, real-time stromen en explosie van gegevensvolumes.
- Het technologische ecosysteem van Big Data.
- De wereld van data science ontsluiten: beschrijvende, voorspellende en prescriptieve analyse.
- De baan, tools en methoden van de datawetenschapper.
- Inleiding tot machinaal leren, analyse onder toezicht en analyse zonder toezicht.
- Noties van over- en onderleren.

Use cases voor data science in een zakelijke waardeketen (klantgedrag, productaanbod, enz.).

2) Methoden en modellen voor gegevenswetenschap

- Verzamelen, voorbereiden en onderzoeken van gegevens.
- Het belang van de aanpak van gegevenskwaliteit (opschonen, transformeren, verrijken).
- Definitie van metriek.

DEELNEMERS

IS-managers, projectmanagers voor gegevensanalyse, managers van statistisch onderzoek.

VOORAFGAANDE VEREISTEN

Geen speciale kennis vereist.

VAARDIGHEDEN VAN DE CURSUSLEIDER

De deskundigen die de cursus leiden zijn specialisten op het betreffende vakgebied. Zij werden geselecteerd door onze pedagogische teams zowel om hun vak kennis als hun pedagogische vaardigheden voor elke cursus die zij geven. Zij hebben minstens vijf tot tien jaar ervaring in hun vakgebied en oefenen of oefenden verantwoordelijke bedrijfsfuncties uit.

BEOORDELINGSMODALITEITEN

De cursusleider beoordeelt de pedagogische vooruitgang van de deelnemer gedurende de gehele cursus aan de hand van meerkeuzevragen, praktijksituaties, praktische opdrachten, ... De deelnemer legt ook van tevoren en naderhand een test af ter bevestiging van de verworven kennis.

PEDAGOGISCHE EN TECHNISCHE MIDDELEN

- De gebruikte pedagogische middelen en cursusmethoden zijn voornamelijk: audiovisuele hulpmiddelen, documentatie en cursusmateriaal, praktische oefeningen en correcties van de oefeningen voor praktijkstages, casestudies of reële voorbeelden voor de seminars.
- Na afloop van de stages of seminars verstrekt ORSYS de deelnemers een evaluatievragenlijst over de cursus die vervolgens door onze pedagogische teams wordt geanalyseerd.
- Na afloop van de cursus wordt een presentielijst per halve dag verstrekt, evenals een verklaring van de afronding van de cursus indien de stagiair alle sessies heeft bijgewoond.

TOEGANGSMODALITEITEN EN -TERMIJNEN

De inschrijving dient 24 uur voor aanvang van de cursus plaatsgevonden te hebben.

TOEGANKELIJKHEID VOOR MINDERVALIDEN

Is voor u speciale toegankelijkheid vereist? Neem contact op met mevr. FOSSE, contactpersoon voor mindervaliden, via het adres psh-accueil@ORSYS.fr om uw verzoek en de haalbaarheid daarvan zo goed mogelijk te bestuderen.

- Statistische basismethoden.
- De belangrijkste klassen van supervised algoritmes: beslisbomen, K-nearest neighbours, regressie, Naive Bayes.
- De belangrijkste klassen van algoritmen zonder toezicht: clustering, PCA, CAH, neurale netwerken.
- Text mining en andere families van algoritmen.

Eenvoudige analyses met R of Python om de technieken van supervised analysis (regressie en classificatie) en unsupervised analysis (clustering, segmentatie en anomaliedetectie) te illustreren.

3) Grafische weergave en ophalen van gegevens

- R- en Python-talen voor statistische analyse.
- Hun ontwikkelomgevingen (R-Studio, Anaconda, PyCharm) en bibliotheken (Pandas, machine learning).
- DataViz-tools (Power BI, Qlik, Tableau, enz.).
- Gegevensmodellering: weergave van processen, stromen, controles en voorwaarden.
- Resultaten communiceren via data storytelling: visuals organiseren (diagrammen, ranglijsten, kaarten).
- Gegevensmodellering: tools (Orange, Power BI).
- Resultaten communiceren met behulp van data storytelling: het belang van de resultaten overbrengen.

Oefeningen in grafische verkenning van gegevens, analyse van de positie en omvang van gegevens (wolken, histogrammen, enz.).

4) Een gegevenswetenschappelijk probleem modelleren

- Samenvatting van het proces.
- Casestudie 1: klantrelaties in de verzekeringsbranche.
- Marketingcampagnes richten. Inzicht in de oorzaken van klantuitval. Welke producten voor welke klanten?
- Casestudie 2: fraudedetectie.
- Vergelijk onderzoek met behulp van traditionele statistieken en datamining.
- Detectie met methode onder toezicht. Detectie zonder toezicht.
- Analyse van twee business cases, klantrelaties en fraudedetectie bijvoorbeeld, maar andere zijn ook mogelijk.

Praktische toepassing van de storytelling-methode op business cases.

DATA

Neem contact met ons op