

Machinaal leren, de stand van de techniek

Seminar van 2 dagen - 14u

Ref : MLE - Prijs 2024 : € 2 090 excl. BTW

Dit seminar gaat dieper in op de problemen bij het verwerken van gegevens met behulp van kunstmatige intelligentie en Machine Learning-algoritmen in het bijzonder. Het toont besluitvormers de belangrijkste algoritmen op dit gebied, praktische oplossingen en de projectaanpak die moet worden toegepast afhankelijk van de use case in het bedrijf.

PEDAGOGISCHE DOELSTELLINGEN

Na afloop van de opleiding kan de cursist:

Machine Learning positioneren in de gegevensverwerkingsketen

Onderscheid maken tussen de vereiste vaardigheden en de aan te werven profielen

De sleutels tot een succesvol Machine Learning-project identificeren

Begrip van concepten voor machinaal leren en de evolutie van Big Data naar machinaal leren

De uitdagingen van het gebruik van Machine Learning begrijpen, inclusief de verwachte voordelen en voorbeelden van gebruik

PEDAGOGISCHE METHODEN

Geïllustreerd door casestudies. Presentatie van de belangrijkste use cases per bedrijfssector (auto-industrie, industrie, consumentengoederen, financiën, gezondheid, energie, landbouw, transport, telecommunicatie, enz.)

HET PROGRAMMA

laatste update: 09/2024

1) Geschiedenis van Machine Learning en de Big Data-context

- Zet de concepten van Kunstmatige Intelligentie, machinaal leren, enz. weer in perspectief.
- De link met wiskunde, (inferentiële) statistiek, datamining en data science.
- Van beschrijvende analyse naar voorspellende en vervolgens prescriptieve analyse.
- Machine Learning-toepassingen (zoekmachines, spamdetectie, lezen van cheques).
- De typologie van algoritmen van Dominique CARDON.
- De Data Science-gemeenschap en Kaggle-uitdagingen (bijv. Netflix).

Studies van concrete toepassingen van Machine Learning (zoekmachines, spamdetectie, het lezen van cheques).

2) Beschikbare gegevens: verzameling en voorbereiding

- Gestructureerde, halfgestructureerde en ongestructureerde gegevens.
- Statistische aard van de gegevens (kwalitatief of kwantitatief).
- Aangesloten objecten (IoT) en streaming.
- Mogelijkheden en grenzen van Open Data.
- Correlaties identificeren, het probleem van multicollineariteit.
- Dimensievermindering door principale componentenanalyse.
- Detectie en correctie van uitschieters.
- ETL (Extract Transform Load).
- Schrapen van het web.

ETL-demonstratie (Extract Transform Load). Gegevensverzameling via het web.

DEELNEMERS

Bedrijfsdirecteuren (CEO, COO, CFO, SG, HRD, enz.), CIO's, CDO's, IT-managers, consultants, Big Data-projectmanagers.

VOORAFGAANDE VEREISTEN

Algemene computervaardigheden en kennis van waarschijnlijkheid en statistiek worden aanbevolen.

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 MINDERVERVALDEN

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.

3) Markttools voor gegevensverwerking en Machine Learning

- Traditionele software (SAS, SPSS, Stata, etc.) en de openheid naar Open Source.
- Kies tussen de twee Open Source leiders: Python en R.
- Cloudplatforms (Azure, AWS, Google Cloud Platform) en SaaS-oplossingen (IBM Watson, Dataiku).
- Nieuwe banen in bedrijven: data engineer, data scientist, data analyst, enz.
- De juiste vaardigheden afstemmen op deze verschillende tools.
- Online API's (IBM Watson, Microsoft Cortana Intelligence, enz.).
- Chatbots (gespreksagenten).

Demonstratie van een chatbot (conversational agent) en Azure Machine Learning.

4) De verschillende soorten Machine Learning

- Leren onder toezicht: een voorbeeld herhalen.
- Leren zonder toezicht: de gegevens ontdekken.
- Online (Machine) Leren in tegenstelling tot batch technieken.
- Versterkingsleren: een beloning optimaliseren.
- Andere vormen van leren (overdracht, sequentieel, actief, enz.).
- Illustraties (aanbevelingsmotoren, enz.).

Demonstraties van de verschillende soorten Machine Learning die mogelijk zijn.

5) Machine-learning algoritmen

- Eenvoudige en meervoudige lineaire regressie. Beperkingen van lineaire benaderingen.
- Polynomiale regressie (LASSO). Tijdreeksen.
- Logistische regressie en scoretoepassingen.
- Hiërarchische en niet-hiërarchische classificatie (KMeans).
- Classificatie met behulp van beslisbomen of de Naïve Bayes-benadering.
- Random Forest (ontwikkeling van beslisbomen).
- Gradient boosting. Neurale netwerken. Vector-ondersteuningsmachine.
- Deep Learning: voorbeelden en redenen voor het huidige succes.
- Text mining: analyse van corpora van tekstuele gegevens.

Demonstratie van de verschillende basisalgoritmen in R of Python.

6) Algoritmetraining en evaluatieprocedure

- Scheiding van datasets: trainen, testen en valideren.
- Bootstrap (bagging) technieken.
- Voorbeeld van kruisvalidatie.
- Definitie van een prestatiemeting.
- Stochastische gradiëntdaling (metrische minimalisatie).
- ROC- en liftcurven om algoritmen te evalueren en te vergelijken.
- Verwarringsmatrix: fout-positieven en fout-negatieven.

De keuze van het beste algoritme aantonen.

7) Productie van een Machine Learning-algoritme

- Beschrijving van een Big Data-platform.
- Hoe PLC's werken.
- Van ontwikkeling tot productie.
- Correctieve en evolutionaire onderhoudsstrategie.
- Evaluatie van operationele productiekosten.

Demonstratie van API's voor geolocatie en sentimentanalyse.

8) Ethische en juridische aspecten van kunstmatige intelligentie

- De opdracht van de CNIL en toekomstige ontwikkelingen.
- Recht op toegang tot persoonlijke gegevens.
- De kwestie van het intellectuele eigendom van algoritmen.
- Nieuwe functies in het bedrijf: Chief Data Officer en Data Protection Officer.
- De vraag over de onpartijdigheid van algoritmen.
- Pas op voor confirmation bias.

- Sectoren en beroepen beïnvloed door automatisering.
Brainstormen om de sleutels tot succes te identificeren.

DATA

Neem contact met ons op