

DATA Science im Live-Betrieb

Eine Online-Seminarreihe der ÖFdv GmbH
mit Informationen von Praktikern für Praktiker

Auftaktveranstaltung am Di. 06.10.2020 im Actuarial Modelling Club

Die Online-Seminare finden am Di. 20.10.2020, Di. 03.11.2020, Di. 17.11.2020,
Di. 01.12.2020 und Di. 15.12.2020, jeweils von 16:30 bis 18:00 Uhr statt

Kosten für diese Seminarreihe (alle fünf Seminare): **390,- EURO (inkl. 20 % USt.)**
CPD-Punkte: 8 (bei Teilnahme an allen fünf Seminaren)

Auch wenn der Begriff „Data Science“ schon allgegenwärtig ist, so fehlt es oft an einer konkreten Vorstellung, was sich dahinter verbirgt. Der Arbeitskreis Data Science der Aktuarvereinigung Österreichs (AVÖ) hat es sich deshalb zum Ziel gesetzt, Interessierten im Zuge von Vorträgen und Online-Seminaren einen leichten Einstieg in dieses Thema zu ermöglichen. Der Fokus soll dabei auf der praktischen Anwendbarkeit der vorgestellten Methoden liegen.

Zum Auftakt einer fünfteiligen Seminarreihe zeigen die Vortragenden Fabian Pribahnsnik und Dietmar Hareter im Actuarial Modelling Club (AMC) am 06.10.2020, wie Data-Science-Methoden im Arbeitsalltag angewendet werden können. Dabei beleuchten die Vortragenden anhand eines Use Cases die verschiedenen Aspekte eines Data-Science-Projekts. In einer daran anschließenden modularen Online-Seminarreihe (fünf Termine à 1,5 Stunden) erhalten Interessierte dann vertiefende Einblicke in die im AMC-Vortrag vorgestellten Themengebiete. Die Auftaktveranstaltung ist kostenlos.

Fünfteilige Online-Seminarreihe mit Praxisbezug

Die fünfteilige Seminarreihe ist wie folgt geplant: Die Themen der ersten drei Vorträge sind vorgegeben und geben eine Einführung in die wichtigsten Methoden und deren praktische Anwendung. Die letzten zwei Vorträge orientieren sich am Interesse des Publikums und werden in Abstimmung mit den Teilnehmerinnen und Teilnehmern im ersten Seminar festgelegt. Jedes Seminar besteht aus einer theoretischen Einführung in die verwendeten Methoden, praktischen Übungsbeispielen und einem Ausblick auf deren praktische Umsetzungsmöglichkeiten.

1. Seminar (Di. 20.10.2020): **Data Science Toolbox – Eine Einführung**

In diesem Vortrag geben die Vortragenden den Teilnehmerinnen und Teilnehmern einen Überblick über die wichtigsten Open-Source-Anwendungen und deren Einsatzgebiete in einem Data-Science-Projekt. Dabei besprechen sie insbesondere die Themen:

- Source Code Management
- Software-Entwicklungsumgebungen und
- Reproduzierbare Reports

Viele der hier vorgestellten Softwarelösungen finden dann im Laufe der Seminarreihe Anwendung.

2. Seminar (Di. 03.11.2020): **Datenvisualisierung mit R**

Die richtige Art und Weise, Daten zu visualisieren, hilft nicht nur, die Daten besser zu verstehen und mögliche Muster zu erkennen, sondern ermöglicht es auch, Ergebnisse einem breiten Publikum klar und verständlich zu vermitteln. Dieser Vortrag gibt einen Einblick in die Welt von ggplot2, einem R-Paket zur grafischen Darstellung von Daten, sowie deren interaktive Visualisierung mittel plotly.

3. Seminar (Di. 17.11.2020): **Regressions- und Clustermethoden**

Regression und Clustering sind grundlegende Werkzeuge in der Analyse. In diesem Vortrag wird die lineare und nichtlineare Regression vorgestellt und auf die Interpretation der Ergebnisse eingegangen. Im zweiten Teil werden die Grundlagen der Clusteranalyse beschrieben.

4. Seminar (Di. 01.12.2020): **Nach Abstimmung mit den Teilnehmerinnen und Teilnehmern**

- Methoden des überwachten und unüberwachten Lernens
- Neuronale Netze (Aufbau, Funktionsweise, Training, Anwendung)
- Dimensionsreduktion
- Klassifizierungsmethoden
- Entscheidungsbäume
- Methoden zur Modellvalidierung
- Textanalyse in R
- Analyse von Webseiten mittels Datenextraktion

5. Seminar (Di. 15.12.2020): **Nach Abstimmung mit den Teilnehmerinnen und Teilnehmern**

- Methoden des überwachten und unüberwachten Lernens
- Neuronale Netze (Aufbau, Funktionsweise, Training, Anwendung)
- Dimensionsreduktion
- Klassifizierungsmethoden
- Entscheidungsbäume
- Methoden zur Modellvalidierung
- Textanalyse in R
- Analyse von Webseiten mittels Datenextraktion