



DATA SCIENCE UND MACHINE LEARNING EXPERTE ONLINE

Der amerikanische Zukunftsforscher John Naisbitt hat schon in den 1980er Jahren die heutige Situation unserer Informationsgesellschaft prognostiziert: "Wir ertrinken in Daten und hungern nach Wissen!".

Damit ist gemeint, dass viele Unternehmen zwar riesige Mengen an Daten wie beispielsweise Kunden-, Produktions-, Maschinen-, Zulieferer- und Verkaufsdaten sammeln, aber diese Daten an sich als Sammlung von großen Zahlenbergen noch keine Erkenntnisse und damit leider keinen Mehrwert fürs Unternehmen liefern. Zusammenhänge in den Daten zu erkennen und Erkenntnisse aus diesen zu gewinnen ist jedoch einer der wichtigsten Hebel, um die richtigen Entscheidungen im Unternehmen vorzubereiten bzw. daraus abzuleiten. Erst damit kann aus Daten ein Mehrwert fürs Unternehmen geschaffen werden.

Genau darum geht es bei Data Science. Data Science meint die im Unternehmen gesammelten Daten mit Verfahren wie beispielsweise dem Machine Learning zu analysieren, um Zusammenhänge aufzudecken und in sogenannten Machine Learning Modellen zu kondensieren. Diese Modelle erlauben es zum Beispiel das Kaufverhalten für neue Kunden vorherzusagen. Wie kann ein Machine Learning Modell aber beispielsweise das Verhalten meiner Kunden modellieren?

Antwort: Durch das automatische Erkennen von unentdeckten Zusammenhängen in großen Datenmengen, die das bisherige Kundenverhalten wieder spiegeln. Da es als Mensch sehr schwierig ist in der Flut von Zahlen diese Zusammenhänge zu erkennen, hilft der Computer, der für diese komplexe Aufgabe perfekt geeignet ist.

Dauer:
7 x 1 Tag
Vollzeit

Start:
19.04.2023

ONLINE

Abschluss:
Zertifikat

Preis:
2 495,00 € + MwSt.

Infoabende:
13.03.23/ 30.03.23
17.30 Uhr, Online

Anmeldung:
anmeldung@tcw-donau-ries.de

In diesem Zertifikatskurs werden Sie zum Experten für Data Science und Machine Learning ausgebildet und schließen den Kurs mit einem Zertifikat ab. Hierzu folgen wir den Phasen des OSEMN- bzw. CRISP-DM Data Science Prozessmodells – ein Prozessmodell, das die Datenanalyse systematisch in verschiedene Phasen unterteilt. Die Kursstruktur ergibt sich direkt durch dieses Standard-Prozessmodell und dient als roter Faden, der sich durch alle Kursmodule zieht.

Die Kursmodule sind hierbei so aufgebaut, dass Sie zuerst lernen in Python zu programmieren ("skripten"), um Daten aus Quellen wie Excel-Tabellen oder Datenbanken zu extrahieren, mit Python-Bibliotheken (Pandas) vorzuverarbeiten, um dann mit Statistiken und Visualisierungen (Matplotlib) erste Erkenntnisse zu gewinnen und dann schließlich mit Machine Learning (scikit-learn / TensorFlow / Keras) Ihre Daten automatisch zu modellieren und für Vorhersagen zu nutzen.

Der Kurs vermittelt nicht nur theoretische Hintergründe. Die Praxis steht im Vordergrund. Alle wichtigen Punkte bei der Datenanalyse werden wir uns daher in sogenannten Jupyter-Notebooks Schritt-für-Schritt gemeinsam erarbeiten. Sie arbeiten dabei alle praktischen Beispiele live mit durch, während der Dozent diese vormacht und auch Aufgaben stellt und wir Ihre Lösungen gemeinsam diskutieren. Konkret werden wir mehrere Fallstudien durcharbeiten, um beispielsweise

- Verkaufspreise für Immobilien und Autos vorherzusagen
- zu klassifizieren, ob Kunden in der Zukunft ein bestimmtes Produkt kaufen werden
- um Zusammenhänge zwischen Delikten in Berlin aufzudecken
- um auf Basis von Text zu verstehen, ob ein Kunde mit einem Produkt zufrieden ist

Zielgruppe:

Zielgruppe sind technisch oder kaufmännische Angestellte die selber Daten auswerten möchten bzw. Führungskräfte, die den Data Science Prozess und das Thema Machine Learning tiefgehend durchdringen wollen. Vorerfahrungen im Programmieren bzw. Mathematik sind nicht notwendig.

TERMINE:

Modul (1) + (2) Einleitung, Übersicht / Willkommen zu Python: 19.04.2023

Modul (3) Data Science Teil 1: 26.04.2023

Modul (4) Data Science Teil 2: 03.05.2023

Modul (5) Machine Learning mit scikit-learn Teil 1: 17.05.2023

Modul (6) Machine Learning mit scikit-learn Teil 2: 24.05.2023

Modul (7) Deep Learning mit TensorFlow/Keras Teil 1: 21.06.2023

Modul (8) Deep Learning mit TensorFlow/Keras Teil 2: 28.06.2023

Dauer: 7 x 1 Tag in Vollzeit / 8.30 bis 16.30 Uhr / jeweils mittwochs

Durchführungsformat: ONLINE

Abschluss: Zertifikat

Dozent: Prof. Dr. Jürgen Brauer

Preis: 2 495,00 € + MwSt.

Kontakt, weitere Informationen zum Kurs und Anmeldung:

Technologie Centrum Westbayern GmbH, Emil-Eigner-Str. 1, 86720 Nördlingen, Tel. 09081 8055-100

E-Mail: anmeldung@tcw-donau-ries.de; www.tcw-donau-ries.de

INHALTE Zertifikatskurs Data Science und Machine Learning Experte – Online

Modul 1: Einleitung / Übersicht

- Was ist eigentlich Künstliche Intelligenz / Machine Learning / Deep Learning / Data Science?
- Einige Uses Cases: Wie kann Data Science und Machine Learning helfen?
- Der OSEMN- und CRISP-DM Data Science Prozessmodell: Phasen, Vor- und Nachteile
- Was braucht ein Data Scientist / Machine Learning Engineer an Kompetenzen?

Modul 2: Willkommen zu Python

- Python und Python-Bibliotheken als DAS Werkzeug für Data Science / Machine Learning
- Historisches zu Python
- Python-Grundlagen: Einfache und komplexe Datentypen (Listen, Dictionaries, Tupel), Schleifen, Funktionen, Klassen
- Praxisbeispiel: Automatisch Aktienkurse mit Python abrufen, Diagramme erzeugen und speichern

Modul 3: Data Science, Teil 1

- Abrufen/Speichern von Daten: Tabellen und (SQLite) Datenbanken als Quellen
- Vorverarbeitung von Daten
- Wie umgehen mit "schmutzigen" Daten? (Falsche und unvollständige Daten filtern)
- Praxis - Fallstudie: Datenanalyse zur Kriminalität in Berlin

Modul 4: Data Science, Teil 2

- Fortsetzung der Praxis - Fallstudie: Datenanalyse zur Kriminalität in Berlin
- Zusammenhänge in Daten qualitativ und quantitativ analysieren
- Trends in Daten durch Visualisierungen erkennen
- Zusammenhänge zwischen Merkmalen automatisch erkennen und visualisieren

Modul 5: Machine Learning mit scikit-learn, Teil 1

- Einführung in das Maschinelle Lernen – Wichtige Begriffe (Terminologie)
- Theorie: Wie funktionieren Decision Trees zur Klassifikation und Regression?
- Theorie: Wie funktionieren Random Decision Forests?
- Praxis - Fallstudie: Vorhersage des Kaufverhaltens von Kunden (Klassifikation)

Modul 6: Machine Learning mit scikit-learn, Teil 2

- Theorie: Funktionsweise weiterer klassischer Machine Learning Modelle
- Overfitting – ein zentrales Problem im Machine Learning und Lösungen hierzu
- Vorverarbeitung für das Machine Learning (Input-/Output-Split, Train-/Test-Split, Skalierung)
- Praxis- Fallstudie: Vorhersage von Verkaufspreisen von Autos (Regression)

Modul 7: Deep Learning mit TensorFlow/Keras, Teil 1

- Theorie: Der Machine und Deep Learning Boom – Was ist passiert? Wieso erst jetzt?
- Funktionsweise Neuronaler Netze
- Wie werden Neuronale Netze eigentlich trainiert? Praxistipps
- Praxis - Fallstudie: Vorhersage von Immobilienpreisen (Regression)

Modul 8: Deep Learning mit TensorFlow/Keras, Teil 2

- Theorie: Bildverarbeitung mit Convolutional Neural Networks
- Theorie: Textverarbeitung mit Rekurrenten Neuronalen Netzen sowie Transformer-Netzen
- Praxis - Fallstudie: Klassifikation von Kundenrezensionen (Textklassifikation)
- Ausblick: Wohin entwickelt sich das Gebiet? (DALL-E 2, ChatGPT, etc.)