





Data Science mit Python

Seminardauer: 3 Tage, 09:00 Uhr bis 17:00 Uhr Schulungsunterlagen: nach Absprache

Im **Kurs Data Science mit Python** lernen Sie, Python im Kontext von Data Science einzusetzen. Sie kennen die grundlegenden Funktionen von Algorithmen im python-gestützten Machine Learning und wissen, wie Sie weitergehende Algorithmen entwickeln. Einlesen, Aufbereiten und Visualisieren von Daten verschiedener Herkunft ist Ihnen nach dieser Schulung ebenfalls vertraut.

Python ist neben R die Programmiersprache, die am häufigsten im Zusammenhang mit Data Science, Datenanalyse und Machine Learning genutzt wird

Zielgruppe

Kurs Voraussetzungen

Fach- und Führungskräfte, Projektleiter für Data Science Projekte

Grundkenntnisse in Python, Grundwissen Statistik



Python-Entwicklungsumgebung

- Installieren neuer Pakete
- Hands-On: Erste Schritte mit Python
- Virtuelle Umgebungen mit venv für projekt-spezifische Software-Entwicklung

Einführung in NumPy

- Listen, Arrays und NumPy-Datentypen
- Berechnungen in Arrays
- · Broadcasting
- · Fancy Indexing

Einführung in Pandas

- Datenaufbereitung
- Indexierung und Slicing eines DataFrames
- Hands-On: Eigenständiges Arbeiten mit Datenstrukturen

Daten einlesen

- · Unterschiedliche Datenformate
- · Konvertieren von Daten
- Objekte in Pandas (Series Object, DataFrame Object, Index Object)
- · Universal Function (ufunc) mit Pandas
- · Hierarchical Indexing
- Arbeiten mit Datasets (Merge, Join)
- · Aggregation von Datasets (Grouping)
- Hands-On: Eigenständiges Arbeiten, um Daten verschiedener Formate einzulesen und zu konvertieren

Grundlegende Statistiken (pandas)

- Erstellen von grundlegenden, deskriptiven Statistiken
- Kontingenztafeln erstellen
- · Hands-On: Eigenständiges Arbeiten, um Daten verschiedener Formate einzulesen, zu konvertieren und erste Statistiken zu erstellen

Datenvisualisierung (matplotlib)

- · Typische Datenvisualisierungen
- · Datenvisualisierung in Python





- · Datenanalyse mit matplotlib
- Histogramme
- Anpassen von Legenden, Farben, Anmerkungen
- Hands-On: Eigenständiges Arbeiten mit Visualisierungen in Python

Machine Learning Algorithmen

- · Feature Engineering
- Technischer Überblick über Machine Learning Algorithmen (supervised, unsupervised)
- · Supervised Learning
 - ^o Regression
 - ⁰ Decision Trees und Random Forests
 - O Support Vector Machines
- · Unsupervised Learning
 - ⁰ K-Means-Clustering
 - ⁰ Principal Component Analysis
 - ^o Gaussian Mixture Models
- Vertiefung einzelner ML Algorithmen
- Validierungsmöglichkeiten: Wie bestimme ich die Generalisierung meiner Lösung

Hands-On: Algorithmen in Python (scikit-learn und tensorflow)

- Einführung: wie Python Algorithmen verwendet
- Hands-On: Eigenständiges Arbeiten, um verschiedene Algorithmen zu testen und zu validieren
- Beispiele in Frameworks scikit-learn und tensorflow für supervised und unsupervised learning