



Data Science mit Python

Einführung in die Data Science-Programmiersprache Python zur Datenanalyse

bitkom
akademie

Data Science mit Python

Einführung in die Data Science-Programmiersprache Python zur Datenanalyse

Kurzbeschreibung

Python ist eine Programmiersprache, die auch im Bereich Data Science etabliert ist. Aufgrund ihrer klaren Syntax und leicht verständlichen Struktur eignet sie sich insbesondere für Anfänger. Bei richtiger Anwendung macht Python das Analysieren großer Datenmengen einfach, effizient und kann Unternehmen rasch neue Erkenntnisse aus dem eigenen Datenbestand verschaffen. Darüber hinaus ist Python plattformunabhängig – es läuft unter Linux, Mac und Windows ebenso wie unter Unix und vielen Unix-Systemen.

Inhalt

- Grundlagen von Python
- Einführung in das Data Science Paket pandas und dem DataFrame
- Daten aus flat files oder Datenbanken einlesen und schreiben
- Daten anpassen, konvertieren, modellieren
- Visualisierung von Daten
- Erste Algorithmen aus dem Machine Learning

Was lernen Sie in diesem Seminar?

Das dreitägige Seminar schafft einen kompakten Einstieg in die Data Science- und Programmiersprache Python. Sie sind anschließend in der Lage, Python selbstständig für einfache Datenanalysen im Unternehmen zu nutzen. Das Seminar umfasst hierbei eine Einführung in die Datenstruktur DataFrame vom Paket pandas (effizient Daten anpassen und verändern), Daten aus flat files (csv) oder einer Datenbank (SQLite) einlesen bzw. schreiben, Visualisierung mit matplotlib, Grundlagen von numpy, erste eigene Funktionen schreiben und enthält einen Einstieg in Machine Learning mit den Algorithmen Lineare Regression und Entscheidungsbaum.

An wen richtet sich das Seminar?

Das Seminar richtet sich an angehende data scientists und Datenanalysten und an der Programmierung in Python interessierte Fachkräfte bzw. Projektleiter, welche noch keine oder wenig Erfahrung mit Python haben und die Grundlagen der Programmiersprache Python für Data Science-Projekte erlernen möchten, um damit eigenständig an data mining Projekten mitzuwirken oder Python Code besser verstehen zu können.

Notwendig sind erste Erfahrungen mit Programmiersprachen (z.B. VBA, Java, C, R, etc.), um das Konzept einer Programmiersprache zu verstehen (Konzept einer Variable in einer Programmiersprache, Zuweisung von Werten zu einer Variablen, Aufrufen von Funktionen, Parameter einer Funktion).

Sehr hilfreich sind grundlegende Vorkenntnisse im Bereich der Statistik (Begriffsdefinitionen wie bspw. Mittelwert, Median, Standardabweichung, Quantil, Dichtefunktion, Normalverteilung), Kenntnisse grundlegender mathematischer Symbole und Begriff (Summenzeichen, Integral, Funktion, Ableitung, Menge der natürlichen und reellen Zahlen, Vektor, Matrix) und Kenntnis der booleschen Algebra mit den logischen Operatoren (UND, ODER, NICHT).

Didaktischer Aufbau des Seminars?

Dieses Seminar ist sehr praxisorientiert. Die Teilnehmer arbeiten direkt und selbstständig mit der Programmiersprache Python in der Entwicklungsumgebung Spyder, so dass das Erlernete direkt geübt und vertieft werden kann. Der Trainer moderiert dabei verschiedene Aufgaben und begleitet die Teilnehmer durch die einzelnen Lehreinheiten.

Was ist an Technik mitzubringen?

- Die Übungen finden auf Laptops statt. Es wird empfohlen, dass Teilnehmer einen eigenen Laptop mit der vorab installierten Software mitbringen. Die Bitkom Akademie kann jedoch auf Anfrage auch eigene Schulungslaptops bereitstellen.
- Die Teilnehmer sollten prüfen, ob firmeneigene Laptops Zugangsbeschränkungen beim Herunterladen von Dateien und Programmen haben, die der Referent bereitstellt. Die digitalen Unterlagen (Skript, Code, Dateien) werden über eine Cloud zur Verfügung gestellt. Die Einwahl in fremde WLAN-Netze sollte daher möglich sein.
- Im Idealfall ist der USB Port der Teilnehmer-Laptops freigeschaltet, so dass als Backup Pakete, verwendete Daten oder sonstige Unterlagen per USB-Stick übertragen werden können.



Zusatzinformationen

- Das Seminar findet in einer kleinen Gruppe mit mind. 5 und max. 10 Teilnehmern statt. Unser Referent kann dadurch auf individuelle Fragestellungen besser eingehen.
- Bitte stellen Sie sicher, dass Sie mit dem Betriebssystem (Windows, Mac OS, Linux), mit dem Sie arbeiten, vertraut sind.
- Die Bitkom Akademie ist [anerkannter Bildungsträger in Baden-Württemberg](#) und [Nordrhein-Westfalen](#). Teilnehmer haben im Rahmen des Bildungszeitgesetzes die Möglichkeit, Bildungsurlaub bzw. eine Bildungsfreistellung zu beantragen. Auf Anfrage erstellen wir auch Anträge auf Anerkennung unserer Veranstaltungen in anderen Bundesländern.
- Lunch und Getränke sind im Seminarpreis enthalten.
- Tipp: Nutzen Sie für Ihre Anreise zu unseren Akademie-Seminaren die [Sonderkonditionen unserer Partner](#).

Seminarprogramm

Data Science mit Python

TAG
1

09.30

Begrüßung durch den Seminarleiter

- Vorstellungsrunde & Erwartungshaltung der Teilnehmer

10.00-11.30

Hintergrund von Python

- Anwendungsmöglichkeiten von Python
- Installieren von Paketen
- Die Entwicklungsumgebung Spyder
- Aspekte von Python, welche von anderen Programmiersprachen abweichen

11.30-11.45

Kaffeepause mit Gelegenheit zum Erfahrungsaustausch und Networking

11.45-13.15

Datenstrukturen in Python

- Einführung in die grundlegenden Datenstrukturen (int, float, NaN)
- Der Wert NaN (not a number) und Besonderheiten
- Funktionen vs. Methoden
- Datenstrukturen in Python (number, tuple, list, dictionary, string)
- list vs. tuple
- Wann benutze ich welche Klammer [], (), {}?

13.15-14.15

Mittagspause

14.15-15.45

Einführung in das Data Science Paket pandas

- Elemente eines pandas data.frame
- Eine Zeile oder Spalte auswählen, hinzufügen und verändern
- Boolean indexing
- Ein leeres data.frame erstellen

15.45-16.00

Kaffeepause mit Gelegenheit zum Erfahrungsaustausch und Networking

16.00-17.30

Grundlegende Statistiken mit pandas

- Statistiken mit einem data.frame berechnen (Anzahl an Beobachtungen, Summe, Mittelwert, Median, Minimum, Maximum, Varianz,...)
- Methoden im data.frame, um einen Überblick der Daten zu erhalten
- Kreuztabelle (Kontingenztafel)
- Auswertungen nach einer Variable gruppieren
- Fehlende Werte löschen oder ergänzen

17.30

Ende des ersten Seminartages

Seminarprogramm

Data Science von Python

TAG
2

09.00

Begrüßung durch den Seminarleiter und Rückblick auf Tag 1

09.30-10.45

Daten einlesen

- Das Arbeitsverzeichnis setzen
- Eine CSV Datei einlesen / schreiben
- Überblick über nützliche Parameter beim Lesen und Schreiben
- Große Dateien mit dem Paket pandas einlesen
- Das pickle modul zum Speichern von Python Objekten

10.45-11:00

Kaffeepause mit Gelegenheit zum Erfahrungsaustausch und Networking

11.00-12.30

Daten aus Datenbanken einlesen

- Eine Verbindung zu einer Datenbank aus Python herstellen
- Eine Tabelle aus der Datenbank abfragen bzw. hinzufügen
- Eine bestehende Tabelle in der Datenbank erweitern
- SQL Abfragen über Python

12.30-13.30

Mittagspause

13.30-15.15

Visualisierung mit matplotlib und pandas

- Die Hauptelemente beim Plotten
- Einen Plot anpassen (x- und y-Achse ändern, Beschriftungen, Legende und Titel) und speichern
- Auswahl von Farben
- Subplots in einer Grafik erstellen
- Erstellen von Scatterplot, Linienplot, Barplot, (gruppiertes) Histogram, Boxplot

15.15-15.30

Kaffeepause mit Gelegenheit zum Erfahrungsaustausch und Networking

15.00-17.00

Control Flows

- Die range() Funktion
- Eine eigene Funktion schreiben und default Parameter setzen
- For Schleifen, If-Else Bedingungen
- List comprehension
- Logische Vergleichsoperatoren

17.00

Ende des zweiten Seminartages

Seminarprogramm

Data Science mit Python

TAG
3

09.00

Begrüßung durch den Seminarleiter und Rückblick auf Tag 2

09.30-10.30

Grundlagen in Numpy

- Attribute eines arrays (Dimension, shape, Größe)
- Einzelne Elemente oder Teilarrays abfragen
- Arrays zusammenfügen und teilen
- Statistische Funktionen in numpy

10.30-11.00

Kaffeepause mit Gelegenheit zum Erfahrungsaustausch und Networking

11.00-12.30

Data Preprocessing

- Daten normalisieren
- Kategorisches Enkodieren von Variablen
- One-Hot Encoding

12.30-13.30

Mittagspause

13.30-15.00

Algorithmen in Machine Learning

- Überblick zu Linearer Regression und Entscheidungsbaum
- Die Algorithmen in Python umsetzen
- Erste Interpretation der Ergebnisse

15.00-15.30

Kaffeepause mit Gelegenheit zum Erfahrungsaustausch und Networking

15.30-17.00

Datenbeispiel

- Vorstellung eines Datenbeispiels
- Eigenständiges Arbeiten, um das bisher Gelernte anzuwenden und zu wiederholen

17.00

Ende des Seminars

Ihr Referent



Jan Köhler

**Technischer Projektleiter und research engineer
Bosch Center for Artificial Intelligence, Robert Bosch GmbH**

Jan Köhler arbeitet und forscht derzeit an den neuesten Technologien im Bereich Data Science, Machine Learning und Deep Learning im Bosch Center for Artificial Intelligence in Stuttgart. Zuvor unterstützte er über viele Jahre als Data Scientist unterschiedliche Geschäftsbereiche bei den verschiedenen Stufen eines Data Mining Projekts von der Beratung bis zur Implementierung. Die technische Umsetzung erfolgte mit R und Python. Seit 2012 ist er als Trainer im Bereich Data Science tätig und schult verschiedene Ebenen, von Projektmitarbeitern bis zum Vorstand.

Shortfacts



Preise

1.650 €* Regulär

1.450 €* für Bitkom-Mitglieder

**Die angegebenen Preise sind in Netto-Beträgen ausgewiesen.*



Termine & Veranstaltungsorte

Die Termine & Veranstaltungsorte entnehmen Sie der Website der Bitkom Akademie. [hier ↗](#)

Buchung des Seminars in Kombination mit „[Supervised & Unsupervised Machine Learning](#)“ (optional)

Die Bitkom Akademie bietet ein zweitägiges Vertiefungsseminar zu Machine Learning an, welches inhaltlich und didaktisch auf dem Seminar Data Science mit Python aufbaut.

Kontaktieren Sie uns – wir beraten Sie gern.

Bitkom Akademie | Albrechtstraße 10 | 10117 Berlin
T 030 27576-540 | info@bitkom-akademie.de
Weitere Seminare finden Sie unter www.bitkom-akademie.de

bitkom
akademie