



Supervised & Unsupervised Machine Learning

Machine Learning-Algorithmen in Python mit scikit-learn

Supervised & Unsupervised Machine Learning

Machine Learning-Algorithmen in Python mit scikit-learn

Kurzbeschreibung

Machine Learning-Algorithmen sind ein elementares Element von Künstlicher Intelligenz und somit zu einem Erfolgsfaktor im Zeitalter der Digitalisierung. Avanciert. Im Bereich Machine Learning (ML) unterscheidet man zwischen überwachtem Lernen (supervised) und unüberwachtem Lernen (unsupervised). Supervised Learning beschreibt die Fähigkeit einer künstlichen Intelligenz, eigenständig Gesetzmäßigkeiten nachzubilden (z.B. durch Erlernen von Parametern in einem Algorithmus). Unsupervised Learning bedeutet hingegen, dass keine Zielwerte bekannt sind und der Algorithmus versucht, bestimmte Muster zu erkennen.

Die Programmiersprache Python ist im Machine Learning (und auch Deep Learning) weit genutzt und bietet die Bibliothek scikit-learn, in welcher zahlreiche Machine Learning-Algorithmen enthalten sind. Python eignet sich somit optimal für die Umsetzung praxisnaher ML-Vorhaben.

Inhalt

- Grundlagen von Python
- Einführung in Machine Learning
- Supervised vs. unsupervised learning
- Vermeidung von overfitting
- Lineare Regression, Entscheidungsbaum, Logistische Regression, Ensemble Methods, Clustering, Hyperparameter Tuning
- Umsetzung von Algorithmen in scikit-learn
- Interpretation der Ergebnisse

Was lernen Sie in diesem Seminar?

Der Fokus liegt auf dem Python Paket scikit-learn und zusätzlich dem Paket statsmodels. Die Konzepte verschiedener Machine Learning-Algorithmen aus dem supervised und unsupervised learning werden behandelt und deren Umsetzung in Python erörtert. Konzepte des Train-Test-Split bzw. der Cross Validation zur Vermeidung von overfitting und die Interpretation der wichtigsten scores zur Bestimmung der Algorithmusgüte sind Bestandteil wie auch die Hyperparametersuche.

Sie kennen anschließend das scikit-learn Paket und dessen Aufbau, um selbstständig erste Machine Learning Algorithmen im Unternehmen umzusetzen und Ihre Pythonfähigkeiten selbstständig erweitern zu können. Das Seminar beinhaltet Algorithmen für die Regression, Klassifikation und dem Clustering, u.a. Lineare Regression, Logistische Regression, Entscheidungsbaum, Random Forest, k-means clustering, SVM, Neuronales Netz.

An wen richtet sich das Seminar?

Das Seminar richtet sich an data scientists, Datenanalysten, angehende Machine Learning engineers o.ä. mit erster Erfahrung in der Programmierung in Python. Das Seminar findet im Anschluss an das Bitkom Seminar „[Data Science mit Python](#)“ statt, welches die grundlegenden Vorkenntnisse vermittelt. Beide Seminare bauen aufeinander auf, können jedoch unabhängig voneinander gebucht werden.

Notwendig ist, dass Teilnehmer mit Python Daten einlesen können, die Datenstruktur DataFrame aus dem Paket pandas kennen, eigene Funktionen schreiben können und das Konzept der Algorithmen Lineare Regression bzw. Entscheidungsbaum kennen.

Notwendig 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 Erlernte 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.
- Die Beispiele werden mit dem Windows Betriebssystem umgesetzt. Wird ein anderes Betriebssystem verwendet (z.B. MacOS), sollte der Umgang vertraut sein.



Zusatzinformationen

- Bitte stellen Sie sicher, dass Sie mit dem Betriebssystem (Windows, Mac OS, Linux), mit dem Sie arbeiten, vertraut sind.
- Das Seminar findet in einer kleinen Gruppe mit mind. 5 und max. 11 Teilnehmern statt. Unser Referent kann dadurch auf individuelle Fragestellungen besser eingehen.
- 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

Supervised & Unsupervised Machine Learning

TAG
1

10:00-10.30

Begrüßung durch den Seminarleiter

- Vorstellungsrunde & Erwartungshaltung der Teilnehmer

10.30-12.00

Überblick über Machine Learning

- Einführung in Machine Learning
- Unterschied Supervised – Unsupervised Learning (überwachtes – unüberwachtes Lernen)
- Overfitting, Train-Test-Split

12.00-12.15

Kaffeepause mit Gelegenheit zum Erfahrungsaustausch und Networking

12.15-13.45

Lineare Regression

- Einführung in den Algorithmus
- Aufteilung der Daten in Train und Test Daten
- Ein lineares Regressionsmodell in Python umsetzen
- Ergebnisse validieren

13.45-14.45

Mittagspause

14.45-16.15

Entscheidungsbaum

- Einführung in den Algorithmus
- Aufteilung der Daten in Train und Test Daten
- Einen Entscheidungsbaum in Python umsetzen
- Hyperparameter anpassen
- Ergebnisse validieren (confusion matrix, accuracy)

16.15-16.30

Kaffeepause mit Gelegenheit zum Erfahrungsaustausch und Networking

16.30-18.00

Logistische Regression (mit statsmodels)

- Einführung in den Algorithmus (von linearer zur logistischen Regression)
- Aufteilung der Daten in Train und Test Daten
- Logistische Regression in Python umsetzen
- Hyperparameter anpassen
- Ergebnisse und Modell validieren (AIC, BIC, confusion matrix)
- Interpretation der Koeffizienten (odds-ratio)
- ROC Curve und AUC

18.00

Ende des ersten Seminartages

Seminarprogramm

Supervised & Unsupervised Machine Learning

TAG
2

09.00-09:30

Begrüßung durch den Seminarleiter und Rückblick auf Tag 1

09.30-10.45

Ensemble Methods

- Random Forest zur Klassifizierung und Regression
- Ada Boost
- Die Algorithmen in Python umsetzen und Ergebnisse validieren

10.45-11:00

Kaffeepause mit Gelegenheit zum Erfahrungsaustausch und Networking

11.00-12.30

Weitere bekannte Algorithmen

- Support Vector Machine (SVM) zur Klassifizierung und Regression
- Neuronales Netz
- Die Algorithmen in Python umsetzen und Ergebnisse validieren

12.30-13.30

Mittagspause

13.30-15.00

Cross-Validation und Hyperparameter Tuning

- Automatisierte Anpassen von Hyperparametern in den Algorithmen
- Cross-Validation (Kreuzvalidierung)
- Erklärung der Konzepte und Umsetzung in sklearn

15.00-15.30

Kaffeepause mit Gelegenheit zum Erfahrungsaustausch und Networking

15.30-17.00

Clustering

- Grundlagen von K-means Clustering und DBSCAN
- Güte eines Cluster-Ergebnisses bestimmen
- Vergleich von Cluster-Ergebnissen verschiedener Algorithmen

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

1.550 €* Regulär

1.350 €* 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 „[Data Science mit Python](#)“ (optional)

Die Bitkom Akademie bietet ein Grundlagenseminar zum Thema Data Science mit Python (dreitägig) an, welches inhaltlich auf dieses Seminar abgestimmt ist und unmittelbar davor stattfindet. Weitere Informationen finden Sie [hier](#).

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