

Akademie

# Machine learning mit Python

In diesem "Machine learning"-Kurs der Heinlein Akademie lernen Sie, wie Sie Machine Learning Probleme in Python realisieren, lösen und implementieren können.

● Profi 📅 3 Tage 🗨️ n.n. 📍 Berlin / Online

## Vorkenntnisse

Obgleich der Kurs die verwendeten Funktionen, Libraries, usw. abdeckt, wird nicht tiefer in die Details der Sprache eingegangen. Aus diesem Grund sollten die Teilnehmer/innen grundsätzliche Kenntnisse in Python haben.

## Kursinhalt:

### Tag 1: Installation und Einführung

- Einführung: Was ist Machine Learning?, Applikationen, Supervised Learning vs Unsupervised Learning, Andere Anwendungen, Exploratory Analysis
- Installation (Windows, Linux)
- Installation von PyCharm
- Einführung in Python für ML: numpy, pandas, scikit, matplotlib
- Überblick an PyCharm: Grundlegende Funktionen, Nützliche Shortcuts, Debug Mode, Data View

### Tag 2: Explorative Analyse

Richtlinien für die Vorgehensweise bei einer explorativen Analyse:

Welches Problem habe ich? Gibt es eine Zielvariable?

Wie sehen meine Daten aus? Welche Variablen sind korreliert? Sind alle Beziehungen wichtig? Mit welchem Kriterium stellt man fest, ob die gefundene Beziehung (Korrelation) wichtig ist oder nicht? Sind die Kriterien nur statistisch?

Welche statistischen Methoden gibt es, um eine explorativen Analyse durchzuführen? Korrelation-Tests, Korrelation: categorical-categorical, categorical-continuous, continuous-continuous, correlation matrices, PCA, Correspondence Analysis, etc.

Visualisierung als wichtige Komponente - Graphen als Teil der explorativen Analyse: Wie können Einblicke aus Graphen gezogen werden? Analyse von einfachen (gewöhnlichen) und komplexen Graphen. Die Antworten auf diese Fragen erarbeitet sich die Gruppe gemeinsam und lernt damit auch gleichzeitig, wie die zuvor angeführten Punkte in Python umzusetzen sind. In Bezug auf die technischen Aspekte wird der Kurs, bei der Ausführung der Aufgaben folgende Themen abdecken:

- Daten laden (CSV, SQL, JSON, etc.)
- Datenverarbeitung (Subsets, Aggregations, etc.)
- Laufende statistische Methoden
- Graphen

### Tag 3: Supervised Learning

- Hauptkonzepte in Supervised Learning: Training, Testing and Validation, Cross-validation, Overfitting and Underfitting, Generation von zusätzliche Variablen
- Supervised learning in Python: Regressions, Naive Bayes, Classification Trees & Random Forests, SVM, GBM, Neural Networks
- Leistungsmessungen: R2, Mean Square/Absolute Error, Confusion Matrix, ROC and AUC
- Technische Aspekte in Python: Train/test splitting, zusätzliche Variablen herstellen, Modelle trainieren, Vorhersagen, Leistungsmessungen

Gestartet wird der dritte Tag mit dem Kennenlernen der Hauptkonzepte, sowie verschiedenster Techniken in Supervised Learning. Aufgrund der starken Praxisorientierung des Kurses, werden die Teilnehmer/innen anschließend in kleine Gruppen geteilt, in denen sie unterschiedliche Supervised Learning Modelle trainieren werden. Jedes Team erhält die gleichen Ausgangsdaten und entscheidet individuell, welche zusätzliche Variablen aufzubauen sind. Die Erfahrungen und Denkprozesse werden hinterher mit dem Rest geteilt.

Die Performance der trainierten Modelle wird mit neuen, unbekanntem Daten getestet und nachfolgend gemeinsam analysiert und diskutiert. Der Dozent steht natürlich während des gesamten Prozesses für alle Fragen zur Verfügung. Durch das Training kommen die Teilnehmer/innen nicht nur in Kontakt mit einem "realen" Machine Learning Problem, sondern sie erlernen auch wie diese in Python zu lösen sind.

### Tag 4: Unsupervised Learning + Text Mining

- Unsupervised Learning: Kmeans, Hierarchical Clustering & Heatmap, Principal Component Analysis/ Correspondence Analysis
- Einführung in Text Mining und Information Retrieval: Applikationen (Sentiment Analysis, IR, etc.), Von Wörtern zu Nummern
  - Pre-processing (Punctuation, lowercase, etc)
  - Porting - Stemming
  - Lemmatization
  - N-grams
  - Numerische Matrizen aus Text
  - Supervised und Unsupervised problems in text mining
  - Information Retrieval (TF/IDF)

Am Tag 4 werden zwei Themen separat bearbeitet (Vormittag & Nachmittag). Jedes Modul wird wie folgt organisiert:

- theoretische Einführung
- Beispiele in Python (Verwendete Python Libraries werden erklärt, z.B. Scikit clustering, Nltk (Natural Language Toolkit))
- Analyse der Ergebnisse
- Selbstständiges Erarbeiten ähnlicher Aufgaben von den Teilnehmern/innen
- Gemeinsame Analyse und Interpretation der Ergebnisse

### Tag 5: Vom Labor zu Produktion

Vormittags wird an der Implementierung von Machine Learning in einem produktiven Umfeld gearbeitet.

- Online vs. Offline Predictions
- Datensammlungstrategien - Response Times
- APIs in Machine Learning
- Batch predictions
- Reports
- Dashboards

Für jedes Topic werden die entsprechende Python Werkzeuge erklärt, so z.B. : Flask, Jupyter, Dash

Am Nachmittag wird der Kurs mit einer Zusammenfassung der bearbeiteten Themen abgeschlossen:

- Überblick der Themen
- Diskussion: Welchen Bezug haben die Themen zueinander?
- Final Q&A

## Dozenten

**n.n.** Im Moment gibt es für diesen Kurs keinen festen Dozenten. Bei Interesse an dieser Schulung organisieren wir einen Profi zu diesem Thema aus unserem Netzwerk.

## Termine

Leider gibt es für diese Schulung momentan keinen festen Termin. Bei Interesse an einer Schulungsteilnahme oder an einer Inhouse-Schulung zu diesem Thema, wenden Sie sich bitte per **Mail** an uns. Vielen Dank.

## Preise

### **Komplett-Paket**

**2.580,00 EUR**

zuzüglich 19% Ust. (=3.070,20 EUR brutto)

inkl. Hotel, Abendessen und Abendprogramm

### **Standard-Paket / Online-Paket**

**2.400,00 EUR**

zuzüglich 19% Ust. (=2.856,00 EUR brutto)

ohne Hotel, Abendessen und Abendprogramm.

### **Komplett-Paket + Zusatznacht am Vortag**

**2.670,00 EUR**

zuzüglich 19% Ust. (=3.177,30 EUR brutto)

inkl. Hotel, Abendessen und Abendprogramm sowie einer zusätzlichen Hotelnacht am Vortag zum Preis von 90,00 EUR (zzgl. 19% USt = 107,10 EUR brutto).

Wenn Sie Fragen haben oder einen Platz reservieren möchten erreichen Sie uns telefonisch unter **030-405051-40** oder per Mail unter **mail@heinlein-akademie.de**.

Die Schulungen finden, sofern nicht anders angegeben, in den Räumlichkeiten der Heinlein Support GmbH, **Schwedter Str. 8/9B, 10119 Berlin** statt.

Anmeldung unter: <http://www.heinlein-support.de/schulung/machine-learning-mit-python>