

Regression mit Python und scikit-learn

Kontinuierliche Zielgrößen modellieren und Vorhersagen fundiert bewerten

01 INTRO

Regressionsverfahren werden eingesetzt, wenn numerische Größen vorhergesagt oder Zusammenhänge zwischen Variablen quantitativ beschrieben werden sollen. Von einfachen linearen Modellen bis zu regularisierten und nichtlinearen Verfahren stehen unterschiedliche Ansätze zur Verfügung.

Der Kurs behandelt Regression als vollständige Modellierungsaufgabe: Problemformulierung, Datenvorbereitung, Modellbildung, Diagnose, Bewertung und Verbesserung. Python und scikit-learn dienen dabei als praktische Werkzeuge.

DAUER

2 Tage

FORMAT

Präsenz · Live Online · Inhouse

VORKENNTNISSE

Python-Grundkenntnisse und grundlegendes Verständnis von Machine Learning

02 TRAININGSPROGRAMM

Inhalte

Regressionsprobleme verstehen

- kontinuierliche Zielgrößen
- Regression versus Klassifikation
- Zusammenhang und Vorhersage
- Features und Target
- typische Anwendungsfälle

Lineare Regression

- lineares Modell
- Koeffizienten
- Intercept
- Modellanpassung
- Vorhersage
- Interpretation

Voraussetzungen und Grenzen linearer Modelle

- Linearität
- Residuen
- Varianz
- Multikollinearität
- Ausreißer
- Extrapolation
- Modellannahmen

Multiple Regression

- mehrere Features
- Einfluss einzelner Merkmale
- kategoriale Variablen
- Interaktionen
- Skalierung
- Interpretation komplexerer Modelle

Nichtlineare Zusammenhänge

- Transformationen
- Polynomial Features
- nichtlineare Effekte
- Modellkomplexität
- Grenzen polynomialer Modelle

Regularisierung

- Overfitting
- Ridge
- Lasso
- Elastic Net
- Regularisierungsparameter
- Feature-Selektion durch Regularisierung

Datenvorbereitung

- fehlende Werte
- Skalierung
- kategoriale Features
- Transformationen
- Ausreißer
- Datenleckage vermeiden

Regressionsmodelle mit scikit-learn

- Estimator API
- Training
- Prediction
- Parameter
- Preprocessing
- Pipelines
- reproduzierbare Modellierung

Regressionsmetriken

- MAE
- MSE
- RMSE
- R^2
- relative Fehler
- Auswahl geeigneter Metriken
- fachliche Interpretation

Residuen und Fehleranalyse

- Residuenplots
- systematische Fehler
- Heteroskedastizität
- problematische Beobachtungen
- Modellschwächen erkennen

Validierung und Modellvergleich

- Train-Test-Split
- Cross Validation
- Baselines
- Modelle vergleichen
- Hyperparameter
- Generalisierung

Praktisches Regressionsprojekt

- Problem formulieren
- Daten vorbereiten
- Modelle entwickeln
- Varianten vergleichen
- Fehler analysieren
- Ergebnisse interpretieren

03 DURCHFÜHRUNG

Rahmen und Organisation

Dauer	2 Tage
Durchführung	Präsenz oder Live Online, vorzugsweise als Inhouse-Schulung
Schwerpunkt	Inhouse-Schulungen für Unternehmen und Teams
Sprache	Deutsch, englische Fachbegriffe und Dokumentation
Unterlagen	Kursunterlagen, Beispiele und Übungsaufgaben
Technik	Eigener Rechner mit Python, Jupyter und scikit-learn

Inhouse-Schulungen

Eigene Datenbestände und Vorhersageaufgaben können in die Übungen einbezogen werden.

04 ZIELGRUPPE

Für wen ist der Kurs gedacht?

Data Analysts, Entwickler und angehende Data Scientists, die Regressionsmodelle mit Python entwickeln und fundiert beurteilen möchten.

Voraussetzungen

Grundkenntnisse in Python, Pandas und den Grundlagen des Machine Learning.

05 LERNZIELE

Was Sie nach dem Kurs können

Die Teilnehmer können Regressionsprobleme formulieren, lineare und regularisierte Modelle entwickeln, geeignete Metriken auswählen und Modellfehler systematisch untersuchen.

06 FAQ

Häufige Fragen

Geht es nur um lineare Regression?

Nein. Lineare Regression bildet den Ausgangspunkt; Regularisierung und nichtlineare Erweiterungen werden ebenfalls behandelt.

Werden Baumverfahren verwendet?

Sie können zum Vergleich herangezogen werden. Ihr Schwerpunkt liegt jedoch im Kurs zu Entscheidungsbäumen und Ensemble-Verfahren.

Wie viel Statistik wird benötigt?

Die für Modellverständnis und Diagnose erforderlichen statistischen Grundlagen werden im Kurs behandelt.

UC IT Service · Ulrich Cuber · kontakt@uc-it.de
<https://www.uc-it.de/coaching/regression-scikit-learn/>