alfatraining Beratung
beratung@alfatraining.de
0800 3456-500
Mo. - Fr. von 8 bis 17 Uhr kostenfrei aus allen Netzen.

Kursstart alle 4 Wochen

Data Scientist

Data Scientists werden eingesetzt, um Firmen dabei zu unterstützen große Datenmengen zu handhaben und anhand dieser die bestehenden Prozesse zu optimieren. Sie wandeln Rohdaten in strukturierte Daten um, analysieren diese und liefern so eine Entscheidungsgrundlage für Unternehmen.

Abschlussart
Zertifikat „Data Scientist“

Abschlussprüfung
Praxisbezogene Projektarbeiten mit Abschlusspräsentationen

Dauer
16 Wochen

Unterrichtszeiten
Montag bis Freitag von 08:30 bis 15:35 Uhr
(in Wochen mit Feiertagen von 8:30 bis 17:10 Uhr)

Nächste Kursstarts
27.05.2024
24.06.2024
22.07.2024

LEHRGANGSZIEL

Du beherrschst Prozesse rund um die Zusammenführung, Aufbereitung, Anreicherung und Weitergabe von Daten sowie die Anwendung von Machine Learning. Ebenso sind dir die Einsatzbereiche von Deep Learning und die Funktionsweisen neuronaler Netzwerke vertraut.

ZIELGRUPPE

Der Lehrgang richtet sich an Personen mit abgeschlossenem Studium in der Informatik, Wirtschaftsinformatik, BWL, Mathematik oder vergleichbarer Qualifikation.

BERUFSAUSSICHTEN

Data Scientists kommen in Unternehmen zum Einsatz, die ihre Geschäftsprozesse auf Grundlage der Analyse und Modellierung großer Datenmengen optimieren möchten, wie in der Logistik, im Onlinehandel und Marketing, in der Energiebranche aber auch im Gesundheitssektor.

Dein aussagekräftiges Zertifikat gibt detaillierten Einblick in deine erworbenen Qualifikationen und verbessert deine beruflichen Chancen.

VORAUSSETZUNGEN

Programmierkenntnisse in Python und Erfahrungen mit Datenbanken (SQL) werden vorausgesetzt.

LEHRGANGSINHALTE

DATA ENGINEER

Grundlagen Business Intelligence (ca. 2 Tage)
Anwendungsfelder, Dimensionen einer BI Architektur
Grundlagen Business Intelligence, OLAP, OLTP, Aufgaben der Data Engineers
Data Warehousing (DWH): Umgang und Verarbeitung von strukturierten, semi-strukturierten und unstrukturierten Daten

Anforderungsmanagement (ca. 2 Tage)
Aufgaben, Ziele und Vorgehensweise in der Anforderungsanalyse
Datenmodellierung, Einführung/Modellierung mit ERM
Einführung/Modellierung in der UML
· Klassendiagramme
· Use-Case Analyse
· Aktivitätsdiagramme

Künstliche Intelligenz (KI) im Arbeitsprozess
Vorstellung von konkreten KI-Technologien im beruflichen Umfeld
Anwendungsmöglichkeiten und Praxis-Übungen

Datenbanken (ca. 3 Tage)
Grundlagen von Datenbanksystemen
Architektur von Datenbankmanagementsystemen
Anwendung RDBMS
Umsetzung Datenmodell in RDBMS, Normalformen
Praktische und theoretische Einführung in SQL
Grenzen von Relationalen Datenbanken, csv, json

Data Warehouse (ca. 4 Tage)
Star Schema
Datenmodellierung
Erstellung Star Schema in RDBMS
Snowflake Schema, Grundlagen, Datenmodellierung
Erstellung Snowflake Schema in RDBMS
Galaxy Schema: Grundlagen, Datenmodellierung
Slowly Changing Dimension Tables Typ 1 bis 5 – Restating, Stacking, Reorganizing, mini Dimension und Typ 5
Einführung in normal, causal, mini und monster, heterogeneous und sub Dimensions
Vergleich von state und transaction oriented
Faktentabellen, Density und Storage vom DWH

ETL (ca. 4 Tage)
Data Cleansing
· Null Values
· Aufbereitung von Daten
· Harmonisierung von Daten
· Anwendung von Regular Expressions
Data Understanding
· Datenvalidierung
· Statistische Datenanalyse
Datenschutz, Datensicherheit
Praktischer Aufbau von ETL-Strecken
Data Vault 2.0, Grundlagen, Hubs, Links, Satellites, Hash Key, Hash Diff.
Data Vault Datenmodellierung
Praktischer Aufbau eines Data Vault Modells – Raw Vault, Praktische Umsetzung von Hash-Verfahren

Projektarbeit (ca. 5 Tage)
Zur Vertiefung der gelernten Inhalte
Präsentation der Projektergebnisse

DATA ANALYTICS

Einführung Datenanalyse (ca. 1 Tag)
CRISP-DM Referenzmodell
Data Analytics Workflows
Begriffsabgrenzung Künstliche Intelligenz, Machine Learning, Deep Learning
Anforderungen und Rolle im Unternehmen der Data Engineers, Data Scientists und Data Analysts

Wiederholung Grundlagen Python (ca. 1 Tag)
Datentypen
Funktionen

Datenanalyse (ca. 3 Tage)
Zentrale Python-Module im Kontext Data Analytics (NumPy, Pandas)
Prozess der Datenaufbereitung
Data Mining Algorithmen in Python

Künstliche Intelligenz (KI) im Arbeitsprozess
Vorstellung von konkreten KI-Technologien im beruflichen Umfeld
Anwendungsmöglichkeiten und Praxis-Übungen

Datenvisualisierung (ca. 3 Tage)
Explorative Datenanalyse
Insights
Datenqualität
Nutzenanalyse
Visualisierung mit Python: Matplotlib, Seaborn, Plotly Express
Data Storytelling

Datenmanagement (ca. 2 Tage)
Big Data Architekturen
Relationale Datenbanken mit SQL
Vergleich von SQL- und NoSQL-Datenbanken
Business Intelligence
Datenschutz im Kontext der Datenanalyse

Datenanalyse im Big Data Kontext (ca. 1 Tag)
MapReduce-Ansatz
Spark
NoSQL

Dashboards (ca. 3 Tage)
Bibliothek: Dash
Aufbau von Dashboards – Dash Components
Customizing von Dashboards
Callbacks

Text Mining (ca. 1 Tag)
Data Preprocessing
Visualisierung
Bibliothek: SpaCy

Projektarbeit (ca. 5 Tage)
Zur Vertiefung der gelernten Inhalte
Präsentation der Projektergebnisse

MACHINE LEARNING

Einführung in Machine Learning (ca. 5 Tage)
Warum Machine Learning?
Anwendungsbeispiele
Überwachtes Lernen, Unüberwachtes Lernen, Teilüberwachtes Lernen, Reinforcement Lernen
Beispiele für Datenbestände
Daten kennenlernen
Trainings-, Validierungs- und Testdaten
Daten sichten
Vorhersagen treffen

Überwachtes Lernen (ca. 5 Tage)
Klassifikation und Regression
Verallgemeinerung, Overfitting und Underfitting
Größe des Datensatzes
Algorithmen zum überwachten Lernen
Lineare Modelle
Bayes-Klassifikatoren
Entscheidungsbäume
Random Forest
Gradient Boosting
k-nächste-Nachbarn
Support Vector Machines
Conditional Random Field
Neuronale Netze und Deep Learning
Wahrscheinlichkeiten

Unüberwachtes Lernen (ca. 5 Tage)
Arten unüberwachten Lernens
Vorverarbeiten und Skalieren
Datentransformationen
Trainings- und Testdaten skalieren
Dimensionsreduktion
Feature Engineering
Manifold Learning
Hauptkomponentenzerlegung (PCA)
Nicht-negative-Matrix-Faktorisierung (NMF)
Manifold Learning mit t-SNE
Clusteranalyse
k-Means-Clustering
Agglomeratives Clustering
Hierarchische Clusteranalyse
DBSCAN
Clusteralgorithmen

Evaluierung und Verbesserung (ca. 2 Tage)
Modellauswahl und Modellevaluation
Abstimmung der Hyperparameter eines Schätzers
Kreuzvalidierung
Gittersuche
Evaluationsmetriken
Klassifikation

Projektarbeit (ca. 3 Tage)
Zur Vertiefung der gelernten Inhalte
Präsentation der Projektergebnisse

DEEP LEARNING

Einführung Deep Learning (ca. 1 Tag)
Deep Learning als eine Art von Machine Learning

Grundlagen in neuronalen Netzen (ca. 4 Tage)
Perceptron
Berechnung neuronaler Netze
Optimierung der Modellparameter, Backpropagation
Deep-Learning-Bibliotheken
Regression vs. Klassifikation
Lernkurven, Überanpassung und Regularisierung
Hyperparameteroptimierung
Stochastischer Gradientenabstieg (SGD)
Momentum, Adam Optimizer
Lernrate

Convolutional Neural Network (CNN) (ca. 2 Tage)
Bildklassifizierung
Convolutional-Schichten, Pooling-Schichten
Reshaping-Schichten, Flatten, Global-Average-Pooling
CNN-Architekturen ImageNet-Competition
Tiefe neuronale Netze, Vanishing Gradients, Skip-Verbindungen, Batch-Normalization

Transfer Learning (ca. 1 Tag)
Anpassen von Modellen
Unüberwachtes Vortrainieren
Image-Data-Augmentation, Explainable AI

Regional CNN (ca. 1 Tag)
Objektlokalisierung
Regressionsprobleme
Verzweigte neuronale Netze

Methoden der kreativen Bilderzeugung (ca. 1 Tag)
Generative Adversarial Networks (GAN)
Deepfakes
Diffusionsmodelle

Recurrente neurale Netze (ca. 2 Tage)
Sequenzanalyse
Rekurrente Schichten
Backpropagation through time (BPTT)
Analyse von Zeitreihen
Exploding und Vanishing Gradient Probleme
LSTM (Long Short-Term Memory)
GRU (Gated Recurrent Unit)
Deep RNN
Deep LSTM

Textverarbeitung durch neuronale Netze (ca. 2 Tage)
Text-Preprocessing
Embedding-Schichten
Text-Klassifizierung
Sentimentanalyse
Transfer-Learning in NLP
Übersetzungen
Seqence-to-Sequence-Verfahren, Encoder-Decoder-Architektur

Sprachmodelle (ca. 1 Tag)
BERT, GPT
Attention-Schichten, Transformers
Textgeneration-Pipelines
Summarization
Chatbots

Deep Reinforcement Learning (ca. 1 Tag)
Steuerung dynamischer Systeme
Agentensysteme
Training durch Belohnungen
Policy Gradients
Deep-Q-Learning

Bayes'sche neuronale Netze (ca. 1 Tag)
Unsicherheiten in neuronalen Netzen
Statistische Bewertung von Prognosen
Konfidenz, Standardabweichung
Unbalancierte Daten
Sampling-Methoden

Projektarbeit (ca. 3 Tage)
Zur Vertiefung der gelernten Inhalte
Präsentation der Projektergebnisse

UNTERRICHTSKONZEPT

Didaktisches Konzept
Deine Dozierenden sind sowohl fachlich als auch didaktisch hoch qualifiziert und werden dich vom ersten bis zum letzten Tag unterrichten (kein Selbstlernsystem).
Du lernst in effektiven Kleingruppen. Die Kurse bestehen in der Regel aus 6 bis 25 Teilnehmenden. Der allgemeine Unterricht wird in allen Kursmodulen durch zahlreiche praxisbezogene Übungen ergänzt. Die Übungsphase ist ein wichtiger Bestandteil des Unterrichts, denn in dieser Zeit verarbeitest du das neu Erlernte und erlangst Sicherheit und Routine in der Anwendung. Im letzten Abschnitt des Lehrgangs findet eine Projektarbeit, eine Fallstudie oder eine Abschlussprüfung statt.

Virtueller Klassenraum alfaview®
Der Unterricht findet über die moderne Videotechnik alfaview® statt - entweder bequem von zu Hause oder bei uns im Bildungszentrum. Über alfaview® kann sich der gesamte Kurs face-to-face sehen, in lippensynchroner Sprachqualität miteinander kommunizieren und an gemeinsamen Projekten arbeiten. Du kannst selbstverständlich auch deine zugeschalteten Trainer:innen jederzeit live sehen, mit diesen sprechen und du wirst während der gesamten Kursdauer von deinen Dozierenden in Echtzeit unterrichtet. Der Unterricht ist kein E-Learning, sondern echter Live-Präsenzunterricht über Videotechnik.

FÖRDERMÖGLICHKEITEN

Die Lehrgänge bei alfatraining werden von der Agentur für Arbeit gefördert und sind nach der Zulassungsverordnung AZAV zertifiziert. Bei der Einreichung eines Bildungsgutscheines oder eines Aktivierungs- und Vermittlungsgutscheines werden in der Regel die gesamten Lehrgangskosten von deiner Förderstelle übernommen.
Eine Förderung ist auch über den Europäischen Sozialfonds (ESF), die Deutsche Rentenversicherung (DRV) oder über regionale Förderprogramme möglich. Als Zeitsoldat:in besteht die Möglichkeit, Weiterbildungen über den Berufsförderungsdienst (BFD) zu besuchen. Auch Firmen können ihre Mitarbeiter:innen über eine Förderung der Agentur für Arbeit (Qualifizierungschancengesetz) qualifizieren lassen.

ⓘ Änderungen möglich. Die Lehrgangsinhalte werden regelmäßig aktualisiert. Die aktuellen Lehrgangsinhalte findest Du immer unter www.alfatraining.de.