

# Workshop: Statistik, Data Mining und Machine Learning in der Praxis mit Minitab

In diesem Workshop werden Verfahren und Methoden der Industriestatistik (Basis: Six Sigma) tiefgreifend im Hinblick auf weiterführende Analysen mit Bezug zu KI, Data Mining und Maschinellem Lernen vorgestellt und durchgeführt.

Anhand von Prozess- und Entwicklungsdaten werden die Vorgehensweisen von der Datenquelle über die Datenhaltung bis zur analytischen Auswertung und Schlussfolgerung in Fallbeispielen sehr praxisnah erarbeitet, interpretiert und bewertet. Eine Datengewinnung für den Workshop kann mit Hilfe von Simulationen direkt auf Basis der Industrie-4.0-Ansätze geschehen. Alternativ wird angeboten, auch (anonymisierte) Echtdaten der Teilnehmer zu verwenden.

Die Bedeutung der Gewinnung der Stichproben, deren Aussagekraft und Qualität die Analysen entscheidend beeinflusst, stellt den Beginn des Kurses dar. Ein Datenerhebungsplan wird angefertigt, die Datenqualität beurteilt und anhand von deskriptiver und explorativer Statistik die Datenlage beschrieben. Dabei sind von besonderem Interesse Ausreißer und „schiefe“ Stichproben (abhängige, unabhängige und seltene Ereignisse).

Die Verfahren der Komplexitätsreduktion, insbesondere die Feature Selection (Prinzip der Sparsamkeit), werden anhand von überwachtem und unüberwachtem Lernen diskutiert. Trainings- und Testdatensätze werden so erstellt, dass die diesbezüglich zu beachtenden Prinzipien erlernt werden. Beispieldatenbanken (Cloud) und Zugriffssoftware wird zudem so zur Verfügung gestellt, dass bis hin zu diversen Minitab-Auswertungen der Analyseprozess vollständig eigenständig bearbeitet werden kann.

Dieser Workshop ist somit ideal für Anwender, die regelmäßig mit komplexen Datenanalysen konfrontiert sind. Angesprochen werden insbesondere Personen, die über die 6Sigma-Analytik hinaus Verfahren der erweiterten Analytik erlernen wollen.

## Kursinhalte

### *Tag 1: Regressionsansätze als Zusammenhangsanalysen*

- Einstieg mittels klassischer statistischer Analysen: lineare und nichtlineare Regression, logistische Regression, Varianzanalyse
- Anmerkungen zu Merkmalen und Zusammenhangsmaßen
- Statistische Modellbildung anhand Versuchsplanung und einfacher Lebensdaueranalyse
- Zusammenhangsmaße und Modellgüte
- Modellbildung mit starken Abhängigkeiten
- Validierung von Regressionsmodellen
- Best-Subset-Regression, Partial Least Square-Regression und Hauptkomponentenanalyse
- Cluster- und Diskriminanzanalyse

### *Tag 2: Modellierung für komplexe Zusammenhänge*

- Die Verfahren des Minitab Predictive Analytic Moduls (PAM)
- Bewertung der Verfahrensgüte: ROC und AUC, Gain, Lift
- Reduktion der Merkmale: das Prinzip der Sparsamkeit
- Die Ansätze der Predictive Modules in Minitab und die typische Projekt-Arbeitsweise
- Modell Deployment (Minitab Model Ops) und Optimierung

### Tag 3: Praxisanwendung

- Erzeugung von Datensätzen von Sensoren und Simulatoren, z.B. einer Klimastation und eines DoE-Simulators (Web-Technik)
- Erfassen der Daten in einer Datenbank (z.B. Azure Cloud)
- Erzeugung einer Live-Regelkarte für die Messdaten mit Minitab Automatisierung
- Gewinnung der Daten über Cloud-Abfragen und Transfer der Daten nach Minitab
- Auswertung der Daten mit Hilfe der Verfahren des PAM
- Projektierung typischer Analyseprojekte: Team und Ressourcen, Fachgebiete und Skills, Technische Ansätze und Prozesse

**Beginn:**

Mittwoch, 25. September 2024, 09:00 Uhr

**Ende:**

Freitag, 27. September 2024, 17:00 Uhr

**Veranstaltungsort:**

61381 Friedrichsdorf  
Deutschland

**Website & Anmeldung:**

Email [academy@additive-net.de](mailto:academy@additive-net.de)

<https://www.additive-net.de/de/training/schulungen/minitab/workshop-statistik-data-mining-und-machine-learning>