



# Machine Learning (ML) Praxisworkshop für Innovationstreiber, Six Sigma und Analyseexperten

Leitfaden für Mustererkennung: Anwendung von Methoden- Analysemethoden zur Effizienzsteigerung in Prozessen.

**Termin: Donnerstag 16.10. – 18.10.2024 ONLINE**

Referenten:

Prof. Dr. Maik Kschischo, Hochschule Koblenz -University of applied sciences

Björn Noreik, BNB-Qualitätsstatistik und Training

**Motivation:** Unternehmen treffen regelmäßig datenbasierte Entscheidungen. Bei diesen Daten handelt es sich häufig um Prozess-, Produkt- oder Felddaten. Durch die Digitalisierung ergeben sich hier vollkommen neue Möglichkeiten, um Zusammenhänge zu erkennen, Abläufe zu verstehen und zu steuern und um gezielt Nutzen zu generieren. Dabei werden Experten jedoch immer häufiger mit komplexeren Datenstrukturen konfrontiert, bei denen klassische Vorgehensweisen zur Datenaufbereitung und statistische Standardverfahren häufig ineffizient sind oder sogar versagen.

**Zielgruppe:** Das Angebot richtet sich speziell an Six-Sigma Belts und Methodenexperten zum Beispiel aus dem Qualitätsmanagement, Prozess- und Produktentwicklung, die sich einen Überblick über Begriffe, Softwarepakete, Möglichkeiten sowie Grenzen rund um die Themen Big Data und Data-Mining verschaffen möchten.

Während dieses dreitägigen Workshops werden methodische- und strukturelle Ansätze für die Analyse komplexerer Datenstrukturen vorgestellt, verglichen und diskutiert. Teilnehmerinnen und Teilnehmer sollten in grafischer und statistischer Datenanalyse geübt sein, um den vollen Nutzen aus diesem Workshop zu ziehen.

## **Warum wir diesen Workshop anbieten:**

In der heutigen datengetriebenen Geschäftswelt sind datenbasierte Entscheidungen unerlässlich für den Unternehmenserfolg. Unternehmen verlassen sich dabei zunehmend auf vielfältige Datenquellen wie Prozess-, Produkt- und Felddaten. Die fortschreitende Digitalisierung eröffnet neue Horizonte für das Erkennen von Zusammenhängen, das Verstehen und Steuern von Abläufen sowie für die zielgerichtete Nutzengenerierung. Hier stoßen traditionelle Methoden der Datenaufbereitung und statistische Standardverfahren oft an ihre Grenzen, was die Effizienz beeinträchtigt oder zu Fehlschlägen führt.

Dieser Workshop bietet eine umfassende Einführung in die Anwendung innovativer Analysemethoden und -werkzeuge, die speziell für diese komplexen Datenlandschaften entwickelt wurden aber auch für einfache Datenstrukturen oft sehr gut anwendbar sind!

Wir richten uns an Six Sigma Belts, Methodenexperten und alle, die sich mit der Optimierung von Geschäftsprozessen durch fortschrittliche Datenanalyse beschäftigen wollen. Mit einem Fokus auf praktischer Anwendung lernen Sie, wie man moderne Data-Mining-Techniken effektiv einsetzt, um tiefere Einblicke zu gewinnen, Prozesseffizienz zu steigern und messbaren Nutzen aus Ihren Daten zu ziehen.

## **Ihr Nutzen aus diesem Workshop:**

Im Zentrum dieses Kurses steht die praktische Anwendung ausgewählter Methoden für fortschrittliche Datenanalysen. Sie werden mit Tools wie Minitab®, KNIME® und Python in Kombination mit Scikit-Learn arbeiten, um deren Anwendungsmöglichkeiten und Effizienz im direkten Vergleich zu erleben. Die im Workshop erworbenen Kenntnisse sind universell und lassen sich mit verschiedenen Werkzeugen und Softwarepaketen umsetzen. Sie werden befähigt, das Gelernte in Ihre berufliche Praxis zu integrieren, um datengetriebene Entscheidungsprozesse in Ihrem Unternehmen effektiv zu unterstützen und zu optimieren.

- Erlernen Sie grundlegende Methoden und essentielle Begrifflichkeiten im Bereich des Machine Learnings
- Entwickeln Sie Fähigkeiten zur Planung und Durchführung eigener Datenanalysen, mit Fokus auf dem gesamten Prozess von Datenerfassung bis Lösungsfindung
- Bringen Sie verschiedene statistische Ansätze für einen flexiblen Zugang zur Datenanalyse zur Anwendung
- Erwerben Sie universell einsetzbare Kenntnisse, die mit unterschiedlichen Werkzeugen und Softwarepaketen umgesetzt werden können.

## **Inhalte** (09:00 Uhr bis ca. 17:00 Uhr)

- Erfahrungsaustausch, Erwartungen und Ziele
- Explorative Datenanalyse und Datenvorbereitung
- Modelle entwickeln mit Minitab (Korrelation, Regression und ANOVA) - Auffrischung (Minitab®)
- Bewerten der Modell- und Prognosequalität mit Vertrauensbereichen (Minitab®)
- Absichern von Modellen mit Hilfe der Kreuzvalidierung (Minitab®)
- Einblick Reduktion der Datenkomplexität anhand Multivariater Analysen (Minitab®)
  
- Industrie 4.0 in Produktionsprozessen – Transformation von Daten in Nutzen
- Typische Datenstrukturen und die sich daraus ergebenden Herausforderungen
- Typische Data-Mining Verfahren, Ansätze und Begriffe
- Datenaufbereitung, Umgang mit Ausreißern und fehlenden Werten (KNIME®)
- Anreichern, Aggregieren und Pivotieren, Joinen von Daten (KNIME®)

## Inhalte (09:00 Uhr bis ca. 17:00 Uhr)

- Vorstellung typischer Data-Mining Ansätze
- Gerichtete und ungerichtete Analysen sowie Feature Selection (KNIME®)
- Entwickeln und bewerten von Trainings- und Testdatensatz
- Entwickeln von Modellen und Bewerten der Prognosequalität (KNIME®, Minitab®)
  
- Verifizierung und Validierung von Ergebnissen aus dem Data-Mining (KNIME®, Minitab®)
  
- Gruppenübung(en)
- Präsentation der Gruppenergebnisse - Vom Problem bis hin zur Empfehlung
- Python und ChatGPT
- Besprechung offener Punkte, Erfahrungsaustausch und nächste Schritte

## **Inhalte** (09:00 Uhr bis ca. 17:00 Uhr)

- **Einführung in Python für Data Science**
  - Grundlagen von Python: Syntax, Variablen, Datentypen, Kontrollstrukturen
  - Einführung in wichtige Python-Bibliotheken für Data Science: NumPy und pandas
  - Übungsbeispiele zur Datenmanipulation und -analyse
- **Grundlagen von Machine Learning und Scikit-Learn**
  - Einführung in Machine Learning: Überwachtes vs. unüberwachtes Lernen
  - Erklärung der Scikit-Learn Bibliothek und ihrer Kernfunktionalitäten
  - Laden und Vorbereiten von Daten mit Scikit-Learn
- **Prozessoptimierung mit Machine Learning**
  - Aufbau eines einfachen Vorhersagemodells (z.B. lineare Regression, Klassifizierung)
  - Trainieren des Modells und Vorhersagen treffen
  - Anwendungsfälle von Machine Learning in der Prozessoptimierung
  - Die Teilnehmenden arbeiten in kleinen Gruppen an einem abschließenden Projekt, das die Anwendung der gelernten Konzepte umfasst.
  - Kurze Präsentation der Projektergebnisse vor der Gruppe

# Interesse?



## Anmeldung und weitere Informationen

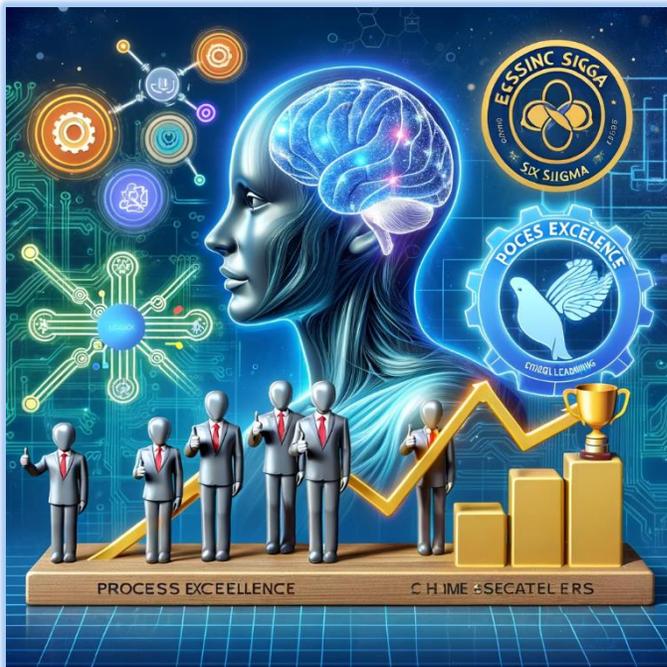
- Über [Webseite](#) oder E-Mail an [essc@sixsigmaclub.de](mailto:essc@sixsigmaclub.de)
- Verbindliche Anmeldungen werden erbeten bis 3 Wochen vor dem jeweiligen Workshoptermin
- *Workshoppauschale inklusive elektronischer Unterlagen(\*)*
- *Für alle drei Tage: 1.490.- € (für ESSC-D Mitglieder: 1290.- €).*
- *(\*) Wir arbeiten ressourcenschonend. Gegen einen Aufpreis von 80€ stellen wir Ihnen die Unterlagen zusätzlich auch gerne gedruckt und als Ordner zur Verfügung.*

Wir empfehlen folgende Softwarepakete für den Workshop:

- Office®
- Statistiksoftware Minitab® - (*Testversion ausreichend*)
- KNIME®
- Python 3 Release, Jupyter Notebook, Scikit-Learn, Anaconda (Optional, jedoch empfohlen)

Bitte installieren Sie diese Pakete vorab als lizenzierte- oder Testversionen.

**Hinweis:** Die Installation der oben genannten Pakete ist optional. Teilnehmerinnen und Teilnehmer sind auch eingeladen, bei den Analysen zuzuschauen und aktiv mitzudenken. Wer auch etwas Erfahrung in der Anwendung der unterschiedlichen Softwarepakete sammeln möchte, sollte diese vor dem Workshop jedoch installieren.



Wir freuen uns auf Ihre Teilnahme