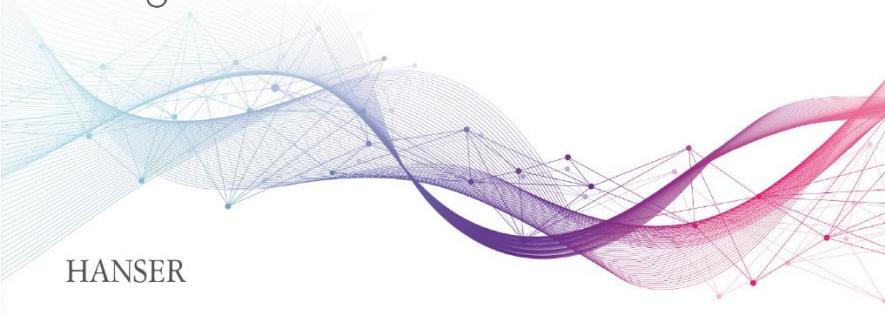


HANSER

Oliver Jöbstl · Jürgen Lipp · Manfred Strohrmann

WORKBOOK – Digitale Transformation des QM

Data-Science-Innovationen
erfolgreich umsetzen



HANSER

Leseprobe

zu

Workbook – Digitale Transformation des Qualitätsmanagements

von Oliver Jöbstl, Jürgen Lipp und Manfred Strohrmann

Print-ISBN: 978-3-446-47762-9

E-Book-ISBN: 978-3-446-48043-8

Weitere Informationen und Bestellungen unter

<https://www.hanser-kundencenter.de/fachbuch/artikel/9783446477629>

sowie im Buchhandel

© Carl Hanser Verlag, München

Inhalt

1 Warum dieses Workbook?	X
2 Einführung in Data Science	4
2.1 Arten von Use Cases	5
2.2 Machine Learning und künstliche Intelligenz	7
2.3 Typische Anwendungsfälle von Machine Learning	9
3 Vorgehensmodell	14
4 Die richtigen Use Cases finden und verstehen	18
4.1 Prozess identifizieren und abgrenzen	20
4.2 Stakeholder-Anforderungen verstehen	22
4.3 Prozessorientierte Use-Case-Ideen ableiten	24
4.4 Produkt- und serviceorientierte Use Cases finden	26
4.5 Use-Case-Ideen priorisieren	28
4.6 Use Cases aus Kundensicht detaillierter beschreiben	30
4.7 Business Case beschreiben	32
4.8 Use Cases final bewerten und auswählen	34
5 Business-Ziele ableiten und Projekt planen	36
5.1 Agile Projektorganisation aufsetzen	38
5.2 Use Case im Ablauf detailliert beschreiben	40
5.3 Wirkende Parameter erkennen ...	42
5.4 Geschäftsziele präzisieren und technische Ziele ableiten	44
5.5 Projektplan erstellen	46

6 Daten erheben und verstehen	48	8 Lösungen ausarbeiten und bewerten	80
6.1 Relevante Features finden	50	8.1 Visualisierungslösungen designen	82
6.2 Datenerfassung planen	52	8.2 Dashboard-Lösungen ausarbeiten	84
6.3 Messsystemfähigkeit sicherstellen	54	8.3 Test- und Tuning-Strategie festlegen	86
6.4 Datenbedarf abschätzen	56	8.4 Regressionsmodelle trainieren, tunen und testen	88
6.5 IT-Infrastruktur für die Datenspeicherung entwickeln und realisieren	58	8.5 Klassifizierungsmodelle trainieren, tunen und testen	90
6.6 Data Ingestion Pipelines entwerfen	60	8.6 Neuronale Netze trainieren, tunen und testen	92
6.7 Datensätze durch Versuche erzeugen	62	8.7 Cluster-Verfahren trainieren und tunen	94
6.8 Eindimensionale Daten verstehen	64	8.8 Reinforcement-Verfahren trainieren, tunen und bewerten ..	96
6.9 Mehrdimensionale Daten verstehen	66	8.9 Large Language Models feintunen	98
7 Daten präparieren	68	8.10 Zeitreihen zerlegen und visualisieren	100
7.1 Daten bereinigen	70	8.11 Ausreißer in Zeitreihen identifizieren	102
7.2 Relevante Features auswählen und konstruieren	72	8.12 Bestes Modell auswählen	104
7.3 Features codieren	74		
7.4 Daten komprimieren (Dimension reduzieren)	76		
7.5 Zeitreihen vorverarbeiten	78		

9 Lösung validieren	106
9.1 Interpretierbarkeit von Machine-Learning-Modellen sicherstellen ..	108
9.2 Lösung auf Plausibilität prüfen ...	110
9.3 Ergebnis aus der Business-Perspektive bewerten	112
9.4 Prozess reviewen	114
10 Lösung einführen und industrialisieren	116
10.1 Nicht-funktionale Anforderungen und Architekturtreiber identifizieren	118
10.2 Entscheidung treffen bezüglich On-Premises oder Cloud-Service-modellen	120
10.3 IT-Infrastruktur für die Produktiv-setzung erarbeiten	122
10.4 IT-Architekturvarianten bewerten und entscheiden	124
10.5 Data-Science-Lösung einführen und Vertrauen aufbauen	126
11 Lösung warten und aktualisieren	128
11.1 Alterungsverhalten des Modells beschreiben	130
11.2 Online-Überwachung für Modell-drift erarbeiten	132
11.3 Online-Überwachung für Daten-drift erarbeiten	134
11.4 Aktualisierungsstrategie festlegen	136
11.5 Training und Wartung von ML-Modellen automatisieren (MLOps)	138
12 Die Autoren	140
13 Index	144



01

**Warum dieses
Workbook?**

Unternehmen können sich immer weniger auf den vergangenen Erfolgen ausruhen. Die Mitbewerber, die Technologien und der Markt müssen ständig beobachtet werden, um den Innovationsbedarf und die damit verbundenen Chancen rechtzeitig zu erkennen und zu nutzen. Innovationen systematisch umzusetzen, ist mehr denn je ein essenzieller Erfolgsfaktor, und ein wesentliches Innovationspotenzial liegt in der Digitalisierung und den damit generierten Daten, die viel zu oft in Unternehmen ungenutzt bleiben.

In dem Buch „Die digitale Transformation des Qualitätsmanagements“ zeigen die Autoren Möglichkeiten auf, die Digitalisierung dazu zu nutzen, das Qualitätsmanagement auf ein deutlich höheres Niveau zu heben, beispielsweise durch digitale QM-Systeme oder durch die Verbesserung der Produkt- und Prozessqualität mithilfe von Daten und digitalen Technologien. In dem Buch wird auch ein Vorgehensmodell zur systematischen Identifikation und Umsetzung von Use-Case-Ideen zur Prozessverbesserung vorgestellt. Aufgrund der häufig gestellten

Fragen zu dem Thema und der hohen Praxisrelevanz haben sich die Autoren entschlossen, diese Vorgehenslogik im Detail zu erläutern und in Form des vorliegenden Workbooks umsetzungsorientiert zur Verfügung zu stellen.

Dieser Praxisleitfaden fokussiert sich nicht ausschließlich auf das Thema Qualitätsmanagement, sondern bezieht sich auf alle Arten von datengetriebenen Innovationen mit Fokus auf das industri-

elle Umfeld, wobei das Spektrum von einfachen Visualisierungslösungen bis hin zu Prognosemodellen mithilfe von Machine Learning und künstlicher Intelligenz reicht. Diese Innovationen im Bereich Data Science können sich einerseits darauf beziehen, neue digitale Lösungen und Services mit Mehrwert für Kunden zu entwickeln, und andererseits darauf, die bestehenden Prozesse systematisch und kontinuierlich hinsichtlich Effektivi-



Bild 1.1 Die Wirkung von Innovationen im Bereich Data Science

1 Warum dieses Workbook?

tät, Effizienz und Flexibilität zu verbessern. Darüber hinaus können neue Business-Modelle generiert werden, die eine völlig neue Art der Leistungserbringung und -verrechnung mit sich bringen (**Bild 1.1**).

Die Möglichkeiten hierzu sind in den letzten Jahren enorm gestiegen – nicht nur, weil immer mehr Daten zur Verfügung stehen, sondern weil aktuell auch die Möglichkeiten zur Verarbeitung von großen Datenmengen vorhanden sind und entsprechende intelligente Algorithmen in rasender Geschwindigkeit weiterentwickelt wurden, wie am jüngsten Hype um ChatGPT zu sehen ist.

Die Vorgehensweise vom Finden der erfolgversprechendsten Use Cases hin zur gewinnbringenden und nachhaltigen Umsetzung muss jedoch von Unternehmen beherrscht werden. Dieser Prozess ist unternehmensspezifisch durch die Definition von entsprechenden Verantwortlichkeiten und Rollen zu organisieren. Es handelt sich hierbei um einen hochgradig interdisziplinären Prozess, weil die unterschiedlichsten Kompetenzen, beispielsweise Fachexperte, IT-Spezialist, Data Analyst und Software Designer eine gute Gesprächs- und Arbeitsbasis finden und ein einheitliches Bild für die Vorgehensweise erarbeiten müs-

sen. Dies geschieht in einer sehr komplexen Ausgangssituation, wo der Erfolg der Innovation keineswegs garantiert werden kann.

Erfolgsentscheidend ist eine Vorgehenslogik, die Orientierung und Klarheit schafft, für effektive Teamarbeit sorgt und je nach Aufgabenstellung geeignete Techniken vorschlägt, die in zielführender Art und Weise zu verwenden sind. Genau diese Vorgehenslogik wird in diesem Workbook vorgestellt.



02

Einführung in Data Science

2.1 Arten von Use Cases

Data Science im industriellen Umfeld ist die Kunst, aus Daten Wissen zu generieren, um Mehrwert für Unternehmen zu schaffen. Data Science beschäftigt sich typischerweise mit der Analyse von großen Datenmengen, der Identifikation von Abnormalitäten in den Daten und mit der Vorhersage von zukünftigen Ereignissen. Use Cases im Bereich Data Science reichen von der Verbesserung eines Prozesses über die Entwicklung neuer Produkte und Dienstleistungen bis hin zur Realisierung von neuen Business-Modellen. Diese Use Cases können mithilfe der Statistik in vier Kategorien unterteilt werden (**Bild 2.1**).

Descriptive Analytics

Use Cases der beschreibenden Statistik erläutern dem Anwender durch die rechtzeitige Aufbereitung von Daten in anschaulicher Art und Weise, **was** passiert ist. Die dadurch gewonnenen Einsichten können dafür genutzt werden, bessere

Entscheidungen zu treffen. Beispielsweise ermöglicht die Messung von Prozessen in Echtzeit, Schwachstellen im Prozess rechtzeitig zu erkennen und diese zu beseitigen. Zusätzlich erfordert das Abrufen von Daten aus verschiedenen Quellen und Abteilungen im industriellen Umfeld oftmals viel Zeit, die durch eine automatische Aufbereitung eingespart werden

kann. Diese Use Cases sind zumeist einfach umzusetzen und haben daher in der Praxis sehr große Bedeutung. Es wird die natürliche Intelligenz der Menschen genutzt und erweitert („augmented“). Daraus werden diese Lösungen im industriellen Umfeld auch unter dem Stichwort „Augmented Workers“ thematisiert.

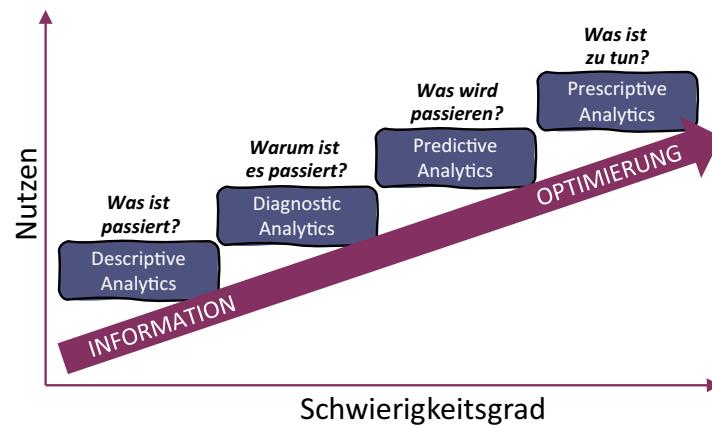


Bild 2.1 Unterteilung von Data Science Use Cases mithilfe der Statistik

Diagnostic Analytics

Hier werden bereits Aspekte mitaufbereitet, die Hinweise geben, **warum** bestimmte Dinge passiert sind. Beispielsweise werden Korrelationsdiagramme dargestellt, die Hinweise auf Ursachen und damit Ansatzpunkte geben.

Predictive Analytics

Im Bereich der prädiktiven Analyse werden Daten dazu verwendet, um zu prognostizieren, **was passieren wird**. Hierzu braucht es Lösungen im Bereich Machine Learning und künstlicher Intelligenz. Es wird aus Vergangenheitsdaten gelernt, indem Gesetzmäßigkeiten erkannt und als Funktion modelliert werden. Dieses Modell wird auf neue Daten angewandt, um Prognosen abzugeben.

Prescriptive Analytics

Präskriptive Lösungen haben zum Ziel, nicht nur Prognosen abzugeben, sondern auch konkrete Handlungsempfehlungen mitzuliefern, **was zu tun ist**. Dies erfordert eine eingehende Diskussion mit den Domänenexperten über die möglichen Optionen, um darauf basierend beispielsweise mithilfe von definierten Regeln oder Optimierungsverfahren die vielversprechendste Handlung abzuleiten. Angenommen, es wurde ein Modell trainiert, welches in der Lage ist, den Output von mehreren teilweise konkurrierenden Qualitätsmerkmalen in Abhängigkeit von verschiedenen Maschineneinstellungen zu prognostizieren. Wird festgestellt, dass sich ein Qualitätsmerkmal zu sehr einer Toleranzgrenze nähert und Gefahr be-

steht, Ausschuss zu produzieren, so gibt es mehrere Möglichkeiten, das Qualitätsmerkmal wieder näher dem Zielwert zu bringen. Mithilfe von Optimierungsverfahren kann eine Handlungsempfehlung abgeleitet werden, die beispielsweise lautet: „Wenn das betroffene Qualitätsmerkmal um $y\%$ verbessert und gleichzeitig sichergestellt werden soll, dass die anderen Qualitätsmerkmale innerhalb der Toleranzgrenzen bleiben, dann ist es am kostengünstigsten, den Maschinenparameter 1 um einen Betrag x zu verändern.“

2.2 Machine Learning und künstliche Intelligenz

Bei Machine Learning werden mithilfe von Algorithmen in einem Datensatz während einer Trainingsphase Muster und Gesetzmäßigkeiten erkannt und als Funktion modelliert. Diese Modelle werden danach in der sogenannten Inferenzphase auf neue Daten angewandt, um für diese Daten Prognosen zu erstellen und diese einem Anwender zur Verfügung zu stellen (**Bild 2.2**). Wichtig ist es zu verstehen, dass hier nach wie vor der Mensch eine zentrale Rolle spielt, da beispielsweise festgelegt werden muss, welche Daten für das Training herangezogen werden sollen, und auch das algorithmische Training entsprechend gestaltet werden muss.

Man kann bei Machine Learning zwischen drei wesentlichen Lernmethoden unterscheiden. Beim überwachten Lernen (Supervised Learning) lernt der Algorithmus anhand von gelabelten Daten: Das sind Daten, die bereits mit einer Antwort versehen sind, sozusagen von einem

Supervisor gelabelt wurden. Beim unüberwachten Lernen (Unsupervised Learning) sind die Lerndaten ungelabelt. Das Ziel ist es, Muster in den Daten zu erkennen und diese beispielsweise zu gruppieren. Beim bestärkenden Lernen (Reinforcement Learning) wird durch die Interaktion mit der Umgebung ge-

lernt. Die Aufgabe besteht darin, eine optimale Handlung zu finden, um eine Belohnung zu maximieren.

Künstliche Intelligenz

Weist ein Computer Fähigkeiten auf, die in der Gesellschaft zum aktuellen Zeitpunkt lediglich Menschen zugeordnet

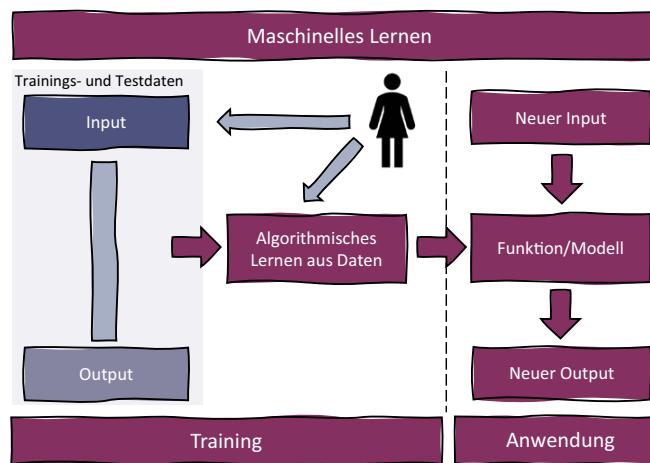


Bild 2.2 Lernen aus Daten durch Machine Learning

sind, dann kann man dies als künstliche Intelligenz bezeichnen. Bei derartigen Lösungen werden grundsätzlich als Eingangsgrößen Daten verarbeitet (Sense), darauf basierend wird eine Ausgangsgröße modelliert (Think) und schließlich eine Prognose oder Handlungsempfehlung erzeugt (Act). Man spricht von einer **Sense-Think-Act-Kette**, die sich durch

hohe Autonomie und durch die Möglichkeit des Lernens auszeichnen muss, um sie als künstliche Intelligenz zu bezeichnen (**Bild 2.3**). Der Begriff maschinelles Lernen ist eng verwandt mit künstlicher Intelligenz, weil der Think-Teil in Lösungen der künstlichen Intelligenz oftmals mit Methoden des Machine Learning realisiert wird.

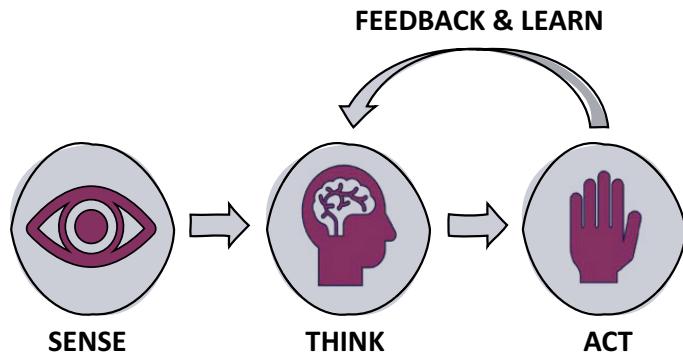


Bild 2.3 Sense-Think-Act-Kette als Erklärungsmodell für künstliche Intelligenz



13

Index

Symbol

- 2K-Workshops
 - Korrelation und Kausalität 110

A

- Abnormalitäten erkennen 10
- Abzinsfaktor 96
- Agile Artefakte 37
- Agile Projektorganisation aufsetzen 38
- Agiles Arbeiten 15
- Aktualisierungsstrategie
 - bei Prognosemodellen 136
- Alternativflüsse 40
- Alterungsverhalten 129
 - eines Modells beschreiben 130
- Amortisationsdauer 32
- Anforderungen
 - funktionale 118
 - nicht-funktionale 118
- Apache Kafka 60
- Architekturtreiber
 - festlegen 118
 - identifizieren 118
- Architekturvarianten 117
- ARMA-Verfahren 100

Ausreißer

- in Zeitreihen identifizieren 102
 - über Fehlergrenze klassifizieren 102
- Auto-Encoder-Verfahren 102
- Automatisierung 122
- Auto ML 81

B

- Back-of-the-Envelope-Berechnung 56
- Bilderkennung und Verarbeitung 11
- Birch-Algorithmus 94
- Boxplot
 - Daten 64
 - segmentierter 66
- Business Case
 - beschreiben 32
- Business Model Canvas 26
- Business-Perspektive
 - Lösungen bewerten 112
- Business Process Model and Notation (BPMN) 24, 40
- Business-Ziele
 - ableiten 36f.
 - präzisieren 32

C

- ChatGPT (Generative Pre-trained Transformer) 2, 11
- Cluster-Algorithmus
 - initialen C. auswählen 94
- Clustern
 - hierarchisches 94
 - spektrales 94
- Clusterung 9
- Cluster-Verfahren
 - trainieren und tunen 94
- Continuous Deployment (CD) 138
- Continuous Integration (CI) 138
- Continuous Testing (CT) 138
- Convolutional Neural Network (CNN) 92
- Crosser 60
- Cross Industry Standard Process for Data Science (CRISP-DS) 15

D

- Dashboard-Lösung 81, 126
 - ausarbeiten 84
 - Drill-downs 84
 - Hover-Effekte 84
 - Zoomen 84

Data Governance 58
Data Ingestion Pipeline 122
– entwerfen 60
Data Pipelines 117
Data Science
– Einführung 5
– Innovationen 1
– Use Cases 5
– Vorgehensmodell für Use Cases 15
Data-Science-Anwendungsfälle
– Datenbedarf abschätzen 56
Data-Science-Lösung
– einführen 126
– Validierung 107
– Verifikation 107
Data Science Use Case 38
– Art abklären 44
– Verbesserung der Datenqualität 69
Daten
– Ausreißer behandeln 70
– bereinigen 70
– Boxplot 64
– Dimension reduzieren 76
– eindimensionale 64
– erheben 48
– komprimieren 76
– mehrdimensionale 67

– präparieren 68
– verstehen 48
– Zeitreihen vorverarbeiten 78
Datenbedarf
– abschätzen 56
Datendrift 130
– Online-Überwachung 134
Datenerfassung
– planen 52
Datenmanagement 58
Datenqualität
– aktuelle 64
– kohärente 64
– korrekte 64
– repräsentative 64
– vollständige 64
Datensätze
– durch Versuche erzeugen 62
Datenspeicherung
– Infrastruktur entwickeln 58
Datenverarbeitungstechnologie 60
DBSCAN 94
– Verfahren 94
Deep Learning 92
Delphi-Methode 28
Deployment 122
Descriptive Analytics 5

Design of Experiments (DoE) 49
Diagnostic Analytics 6
Digitalisierung 1
Domänenverständnis 64
Downsampling 78
Drill-downs 84

E

Einträge
– Konsistenz prüfen und korrigieren 70
Entscheidung
– On-Premises oder Cloud 120
Entscheidungsbäume 90
Explainable AI (XAI) 108

F

Feature-Engineering-Workshop 49
Feature-Importance-Plot 108
Features
– codieren 74
– relevante auswählen und konstruieren 72
– relevante finden 50
– Vorgehensweise 50

Fehler

- akzeptable festlegen 44
- Fehlerkosten
 - erarbeiten 44
- Freigabeprozess
 - definieren 136

G

- Gateways 40
 - exklusive 40
 - parallele 40
- Geschäftsziele
 - präzisieren 44
- Gewichtungsfaktoren 92
- Grafana 84
- Greedy-Parameter 96
- Ground Truth 132
- Grundsatzentscheidung
 - On-Premises oder Cloud 117

H

- Häufigkeitsverteilung 64
- Hauptkomponenten (Principal Components) 76
 - bestimmen 76

Heatmap

- Korrelationen 66
- Hierarchisches Clustern 94
- Hover-Effekte 84
- Hybrid-Cloud-Lösung 120
- Hyperparameter 86

I

- Imputing
 - K-Nearest-Neighbours 70
- Individual SHAP Value Plot 108, 126
- Inferenzphase 122
- Inferenz-Pipeline 122
- Input-Daten 42
- Interpretierbarkeit
 - globale 108
 - lokale 108
- Ishikawa-Diagramm 50
- ISO 25010 118
- IT-Architekturvarianten
 - bewerten 124

K

- Kafka 122
- Kausalitätsprüfung 110
- Keep-Stop-Start-Methode 114
- k-fache Kreuzvalidierung 86
- K-Fold 86
- Klassifizierungsalgorithmus
 - initialen K. auswählen 90
- K-Means-Algorithmus 94
- K-Means-Clustering 94
- K-Nearest-Neighbours-Imputing 70
- Knime 60
- Koeffizientenquadrate 88
- Kollinearität 66
- Kolmogorow-Smirnow-Test 134
- Konzeptdrift 130
- Korrelationen
 - als Heatmap 66
- Korrelationsmatrix 66
- Kreuzvalidierung
 - k-fache 86
- Künstliche Intelligenz 8

L

- Lagekennwerte 64
- Large Language Model (LLM) 12, 98
 - ChatGPT 98
- Lasso-Regression 88
- Lieferant – Input – Prozess – Output –
 - Kunde (LIPOK) 20
- LIPOK-Analyse 22
- LIPOK-Methode 20
- Lösungen
 - ausarbeiten und bewerten 80
 - Auto ML 81
 - Dashboard-Lösungen 84
 - Ergebnis aus Business-Perspektive bewerten 112
 - Grafana 84
 - Machine-Learning-Lösung 81
 - Plausibilität prüfen 110
 - produktiv setzen 116
 - validieren 106
 - Visualisierungslösung 81
 - warten und aktualisieren 128
- Lösungsprozess
 - reviewen 114

M

- Machine Learning 7
 - Abnormalitäten erkennen 10
 - Anwendungsfälle 9
 - Bilderkennung und Verarbeitung 11
 - Clusterung 9
 - Lernmethoden 7
 - Modelliteration 129
 - Prädiktionsmodelle 9
 - Regression und Klassifizierung 9
 - Reinforcement Learning 7, 11
 - Retraining 129
 - Sprach- und Textgenerierung – Chatbots 11
 - Supervised Learning 7
 - Unsupervised Learning 7
 - Zeitreihenprognosen 10
- Machine-Learning-Lösung 81
- Machine-Learning-Modell
 - Aktualisierungsstrategie 136
 - Alterungsverhalten 129f.
 - Continuous Deployment (CD) 138
 - Continuous Integration (CI) 138
 - Continuous Testing (CT) 138
 - Datendrift 130
 - Datensätze erzeugen 62

- Erklärbarkeit 108
- Ground Truth 132
- Interpretierbarkeit 108
- Kausalitätsprüfung 110
- Konzeptdrift 130
- Modelldrift 130
- Prognose 50
- Training und Wartung automatisieren 138
- Machine Learning Operations (MLOps) 138
- Makigami-Analyse 24
- Mean Squared Error (MSE) 88
- Merkmale
 - kontinuierliche M. diskretisieren 74
 - One Hot Encoding 74
 - Ordinal Encoding 74
 - quantitative M. normieren 74
 - quantitative M. standardisieren 74
- Merkmalstypen 42
- Messgrößen
 - technische 44
- Messprozesse
 - bei Use Cases 54
- Messsystemanalyse (MSA) 54
- Messsystemfähigkeit
 - sicherstellen 54

Mindmap 50
 MLP-Netz 92
 Mock-up
 – testen und verbessern 82
 Modell
 – bestes M. auswählen 104
 – prädiktives 81
 Modelldrift 130
 – Online-Überwachung 132
 Modelliteration 129
 Modell-Tuning 86
 – Strategie festlegen 86
 MQTT 60, 122
 Multi-Layer-Perzepron (MLP) 92

N

Naive-Bayes-Klassifikation 90
 Neuronale Netze
 – trainieren, tunen, testen 92
 NoSQL-Datenbank 58

O

One Hot Encoding 74
 Orchestrierung 122
 Ordinal Encoding 74

Osterwalder, Alexander 26
 Output-Daten 42
 Overfitting 88

P

Partial Dependence Plot 108
 People + AI Guidebook 126
 Performance-Kriterien 81
 Personas 26
 Pipeline-Komponenten 60
 Population Stability Index (PSI) 134
 Prädiktionsmodell 9, 81
 – Training 86
 Precision 90
 Predictive Analytics 6
 Prescriptive Analytics 6
 Principal Component Analysis (PCA) 76
 Principal Components 76
 Priorisierungsmethoden 28
 Product Owner 38
 Produkt-Backlog 37f.
 Produktinkrement 37
 Produktivsetzung
 – IT-Struktur 122
 Produkt- und serviceorientierte Use Cases finden 26

Prognosemodell 40
 Projekt
 – planen 36
 Projektplan erstellen 46
 Prozess
 – abgrenzen 20
 – identifizieren 20
 Prozessorientierte Use-Case-Ideen
 ableiten 24
 Prozess reviewen 114
 Prozessverbesserung 1
 PSA-Methode 82
 Pugh-Matrix 124

Q

QM-System
 – digitales 1
 Qualitätsmanagement 1
 Question-Zero-Tabelle 30

R

Randhäufigkeiten 66
 Recall 90
 Receiver Operating Characteristic (ROC) 90

Regression
– logistische 90
– und Klassifizierung 9
Regressionsalgorithmus
– initialen R. definieren 88
Regressionskoeffizienten 88
Regressionsmodell
– trainieren und testen 88
Regularisierungsansatz 88
Regularisierungsfaktor 88
Reinforcement-Algorithmen 96
Reinforcement Learning 7, 11, 96
– Action 96
– Policy 96
– Reward 96
– State 96
Reinforcement Learning based on Human Feedback (RL-HF) 12
Reststreuung 100
Retraining 129
Retrospektive
– planen 38
Reviews 37
ROC-Kurve 90

S
Scrum Master 38
Sense-Think-Act-Kette 8
Shapley-Additive-Explanations-Methode (SHAP) 108
Signalverlauf
– Störungen 102
Skalierbarkeit 122
SMART-Regeln 32
Spektrales Clustern 94
Sprach- und Textgenerierung – Chatbots 11
Sprint-Backlog 37
Sprint-Planung 38
Stage-Review-Konzept 46
Stages
– definieren 46
Stakeholder-Anforderungen verstehen 22
Störfaktoren 42
Stratifizierungskonzept 86
Streudiagrammmatrix 66
Streuungskennwerte 64
Supervised Learning 7, 81
Support Vector Machines 90
Swim Lanes 40

T
Team 38
Technische Ziele
– ableiten 44
Teststrategie
– festlegen 86
Total Cost of Ownership (TOC) 120
Tuning-Strategie
– festlegen 86

U
Überschlagsberechnung 56
Unified Modeling Language (UML) 122
Unsupervised Learning 7
Upsampling 78
Use Case
– aus Kundensicht detaillierter beschreiben 30
– Business Case beschreiben und Business-Ziele präzisieren 32
– Data Science 5
– Descriptive Analytics 5
– Diagnostic Analytics 6
– final bewerten und auswählen 34
– finden 19

- Ideen priorisieren 28
- im Ablauf beschreiben 40
- Kategorien 5
- Messprozesse 54
- Predictive Analytics 6
- Prescriptive Analytics 6
- produkt- und serviceorientierte Use Cases finden 26
- Prozess identifizieren und abgrenzen 20
- prozessorientierte Ideen ableiten 24
- Schritte der Projektphase 19
- Stakeholder-Anforderungen verstehen 22
- Use-Case-Ideen 1
- priorisieren 28

V

- Validierung 107
- Value Proposition Canvas 26

- Varianzaufklärung
 - berechnen 76
- Verfügbarkeit der Daten 42
- Verifikation 107
- Visualisierungslösungen 81, 117
 - Alterungsverhalten 129
 - designen 82
- voestalpine High Performance Metals
DIGITAL SOLUTIONS GmbH 15

W

- Wirkende Parameter erkennen 42

X

- XAI (Explainable AI) 108

Z

- Zeitplan
 - ableiten 46
- Zeitreihen
 - Ausreißer identifizieren 102
 - vorverarbeiten 78
 - zerlegen und visualisieren 100
- Zeitreihenprognosen 10
- Zerlegung
 - additive 100
 - mit kleinerem Fehler 100
 - multiplikative 100
- Zielerreichung 112
- Zielwerte
 - ableiten 44
- Zoomen 84