

Enhanced Six Sigma

Integration von Data Mining Verfahren in den DMAIC Zyklus

.....

Durch die zunehmende Komplexität und Variabilität in heutigen Geschäftsprozessen stoßen Verbesserungsansätze wie der Six Sigma DMAIC bei der Optimierung von Prozessen an ihre Grenzen. Durch die Erweiterung klassischer Vorgehensweisen um datenbasierte Konzepte können auch komplexe Prozesse im Kontext von Industrie 4.0 und Big Data zielgerichtet optimiert werden.

.....

Jacqueline Schmitt, Dietmar Stemann und Jochen Deuse



Die strukturierte Vorgehensweise von Six Sigma bietet zahlreiche Vorteile bei der nachhaltigen Verbesserung von Prozessen. Zum einen kann durch die erfolgreiche Umsetzung von Projekten nach der Six-Sigma-Methode ein nachhaltiger Erfolg für das Unternehmen generiert werden, welcher wesentlich zum Wachstum und Bestand in heutzutage schnelllebigen Märkten beiträgt. Zum anderen bildet Six Sigma die Basis für eine neue Unternehmenskultur und fortwährende Erneuerung, sodass ein geschlossener Regelkreis entsteht. Des Weiteren ermöglicht Six Sigma eine Kaskadierung der Unternehmensziele bis >>>

INFORMATION & SERVICE

LITERATUR

- Brecht, A.; Fromm, R.; Mildner, F.; Rösner, C.; Sommerhoff, B.; Sroka, F.: Der Missing Link – Mix Sigma macht die Qualitätssicherung fit für Industrie 4.0. QZ 4/2018, S. 26–31
- Müller, L.; Loderer, A.; Schwulera, E.; Hausotte, T.: DMAIC powered by TRIZ – Einsatzweiterung von Six Sigma durch bewährte Kreativ-Tools. QZ 3/2017, S. 30–33
- Jung, B.; Wappis, J.: Null-Fehler-Management – Umsetzung von Six Sigma. Carl Hanser Verlag, München 2016

AUTOREN

- Jacqueline Schmitt** ist wissenschaftliche Mitarbeiterin und Doktorandin am Institut für Produktionssysteme, TU Dortmund und Six Sigma Green Belt.
- Dr. Dietmar Stemann** ist Geschäftsführer von mts Consulting Partner und Six Sigma Master Black Belt.
- Prof. Dr.-Ing. Jochen Deuse** ist Professor an der TU Dortmund und Leiter des Instituts für Produktionssysteme (IPS).

KONTAKT

Jacqueline Schmitt
T 0231 755 5701
jacqueline.schmitt@ips.tu-dortmund.de

Dr. Dietmar Stemann
T 0174 2471807
stemann@mts-consultingpartner.com

QZ-ARCHIV

Diesen Beitrag finden Sie online:
www.qz-online.de/6885677

auf die operative Ebene, sodass jeder Geschäftsbereich, jede Abteilung und jeder Mitarbeiter auf die tatsächlich benötigten Ziele hinarbeitet. Im Zuge der zunehmenden Individualisierung von Kundenwünschen reichen gute oder fehlerfreie Produkte nicht mehr aus, um dauerhaften Erfolg zu garantieren. Vielmehr müssen Erwartungshaltungen und Vorstellungen des Kunden erkannt und effizient und wirtschaftlich erfüllt werden. Hierbei ermöglicht die Six-Sigma Methode durch die explizite Berücksichtigung der Kundenanforderungen, den Wert für den Kunden zu steigern. Six Sigma bietet weiterhin einen Ansatz für immerwährendes Lernen im Rahmen einer lernenden Organisation, indem die Entwicklung und Verbreitung von Wissen innerhalb einer Organisation vorangetrieben wird.

Grenzen klassischer Methoden

Im Kontext der stetig wachsenden Datenmengen im Zuge der Digitalisierung stoßen klassische Vorgehensweisen zur Qualitätssicherung und Prozessverbesserung immer mehr an ihre Grenzen. Zwar lassen sich Verbesserungsprojekte mit dem DMAIC-Zyklus strukturiert umsetzen, allerdings lässt sich mit seinen klassischen, meist manuellen Methoden keine vollumfängliche Transparenz mehr über die Geschäftsprozesse erzielen, da diese hierfür meist zu komplex sind. Große Datenmengen und komplexe Strukturen in den Prozessen erfordern spezielle Tools zur Speicherung, Verarbeitung und zielgerichteten Analyse der Daten.

Durch die zunehmende Variantenvielfalt und eine Vielzahl eingesetzter Sensorik zur Erfassung von Daten führt die Heterogenität dieser Daten zu neuen Herausforderungen in der Datenaggregation und -selekti-

on. Die klassische Vorgehensweise ist vor diesem Hintergrund häufig nur mit erheblichem Ressourcenaufwand möglich und daher unter wirtschaftlichen Aspekten nicht sinnvoll.

Integration von Data Mining in DMAIC

Durch die Integration von Data-Mining-Methoden in den Six Sigma DMAIC Zyklus lassen sich die genannten Grenzen der Six Sigma Methode überwinden und auch große Datenmengen aus heterogenen Datenquellen wirtschaftlich analysieren.

Insbesondere die Analyse Phase bietet zahlreiche Möglichkeiten, bestehende Werkzeuge der klassischen Statistik durch Data Mining Verfahren zu ergänzen oder zu ersetzen. In der klassischen Six Sigma DMAIC Vorgehensweise werden in der Analyse Phase die möglichen, oft meinungsbasierten Ursachen analysiert. Diese können im ersten Schritt durch Brainstorming, Flow Charts, Ishikawa-Diagramme, FMEA oder XY-Matrix ermittelt, gesammelt und strukturiert werden (siehe Bild 1 oben). Auf Basis der meinungsbasierten Ursachen werden im nächsten Schritt Hypothesen formuliert und getestet. Hierzu können Hypothesentests und grafische Tools, z.B. Streudiagramm, Regressionsanalyse, Chi-Quadrat-Test, ANOVA, zum Einsatz kommen (siehe Bild 1 unten). Das Ergebnis sind verifizierte, d.h. bewiesene, Ursachen sowie verworfene, als widerlegt gekennzeichnete, Ursachen, die von der weiteren Betrachtung ausgeschlossen werden.

Beispiel: Prozesskette in der Nudelherstellung

Am Beispiel der Prozesskette der Nudelherstellung, welche analog zu typischen

Materialdaten	Prozessdaten			Qualitätsinformation
	Mischen/Kneten	Extrudieren	Tempern	
Ei	Drehzahl	Dauer	Dauer	Qualitätsindikator
Öl	Temperatur	Druck		
Mehl	Dauer	Temperatur		
Salz	Drehmoment			
Wasser				

Tabelle 1. Verfügbare Daten des Anwendungsbeispiels Nudelherstellung (© mts Consulting Partner)

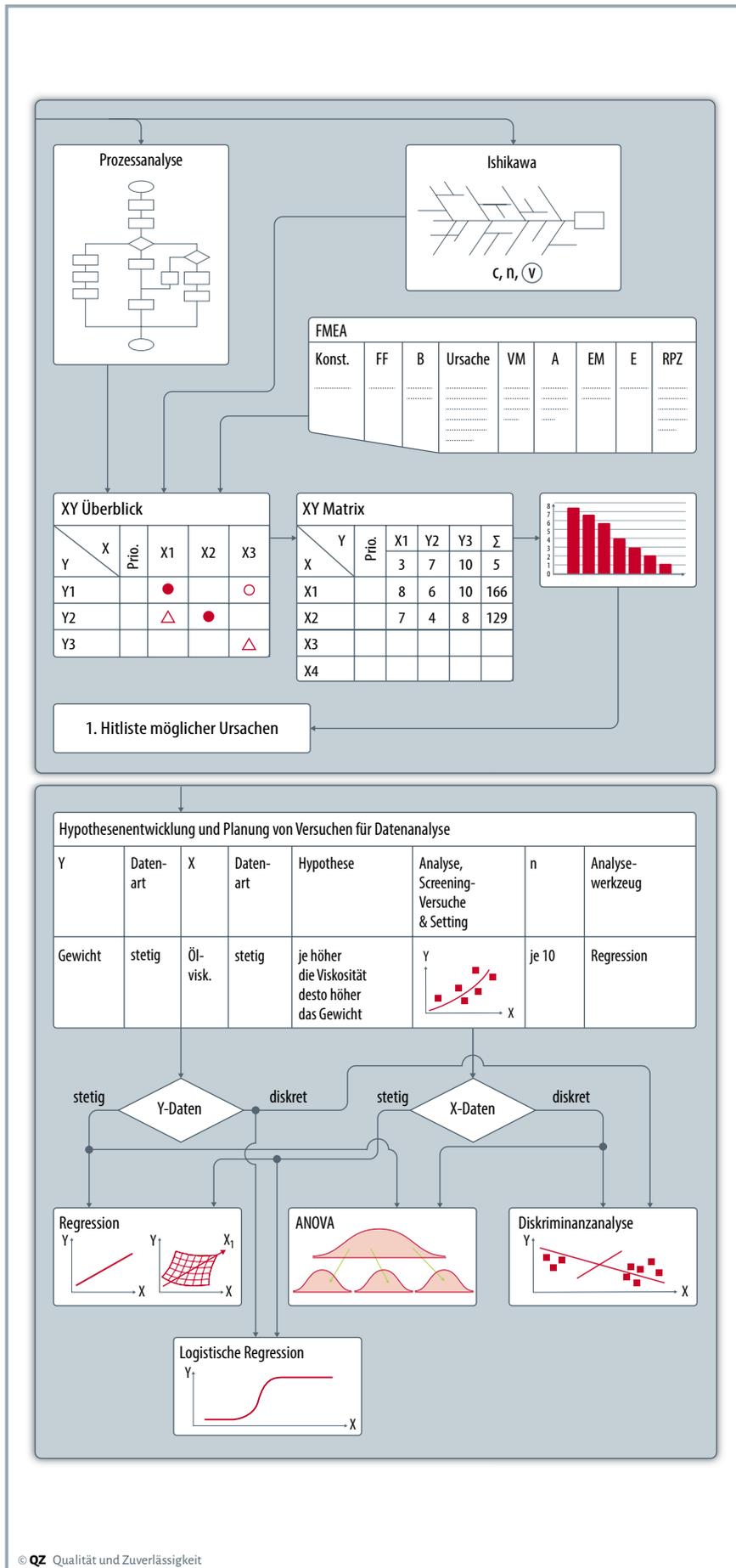


Bild 1. Roadmap der klassischen Vorgehensweise in der Analyse-Phase des DMAIC (Quelle: www.six-sigma.de)

Batch-Prozessen in der Dichtungsproduktion behandelt und analysiert werden kann, lässt sich veranschaulichen, wie sich Data Mining in die Analyse Phase des DMAIC integrieren lässt. Der für die Analyse bereitgestellte Datenauszug umfasst 1.000 Datensätze mit den in Tabelle 1 dargestellten Variablen.

Im Rahmen der Measure-Phase wurde festgestellt, dass 9,31 % der Nudeln außerhalb der unteren Spezifikationsgrenze liegen (erwartete Prozessleistung) und die aktuelle Prozesslage mit $C_{pk} = 0,533$ und $C_p = 0,901$ als nicht fähig eingestuft werden kann.

Klassische Vorgehensweise

Für die klassische Six Sigma DMAIC Vorgehensweise in der Analyse Phase wurden mit dem Ziel der Verifizierung gesammelter, meinungsbasierter Ursachen, zunächst eine Korrelationsanalyse durchgeführt und je ein Regressionsmodell auf den Gesamtdaten und auf den vier Haupteinflussfaktoren gebildet. Das Ergebnis ist ein lineares Regressionsmodell mit der Gleichung:

$$\text{Qualitätsindikator} = -163,31 - 124,83 \cdot \text{Salz} + 10,01 \cdot \text{Druck}_E - 0,04 \cdot \text{Temperatur}_M - 422,72 \cdot \text{Flüssigkeitsanteil}$$

Dabei wurde die Variable Flüssigkeitsanteil aufgrund der negativen linearen Korrelation der Variablen Wasser und Öl als Flüssigkeitsanteil = (Wasser + Öl) / Mehl neu gebildet; die ursprünglichen Variablen wurden entfernt.

Als Gütemaß wird das Bestimmtheitsmaß R^2 eingesetzt. Für das Modell ergibt sich $R^2 = 0,97$. Das bedeutet, dass 97 % der Streuung des Qualitätsindikators durch das Modell beschrieben werden können. Die Analyse der ermittelten Regressionskoeffizienten erfolgt mit Hilfe der zugehörigen P-Werte. Diese geben an, ob sich ein Koeffizient signifikant von 0 unterscheidet, d.h. der Einfluss als signifikant gewertet werden kann. Im Anwendungsfall wurde der Einfluss der Variablen Druck_E, Salz, Flüssigkeitsanteil und Temperatur_M mit $P \approx 0,00000$ als signifikant identifiziert.

Integration von Data Mining

Parallel zum klassischen Vorgehen wurde für diese Anwendung auf Data Mining zurückgegriffen. Aus der Vielzahl von Lernverfahren wurden zunächst Entscheidungsbäume ausgewählt, da sich diese einfach >>>

anwenden und gut interpretieren lassen. Durch die grafische Visualisierung von Entscheidungsregeln eignen sich Entscheidungsbäume gut, um Hauptinflussfaktoren zu identifizieren und grundlegende Strukturen in den Daten zu entdecken.

Abhängig vom Skalenniveau der Zielgröße können Entscheidungsbäume sowohl zur Klassifikation als auch zur Regression (als Regressionsbäume) eingesetzt werden. Meist ist es jedoch sinnvoll, auch stetige Zielgrößen zu diskretisieren, um die Ergebnisse des Baums besser interpretieren und in Management- und Expertenkreisen diskutieren zu können. Im vorliegenden Anwendungsbeispiel wurde daher der stetige Qualitätsindikator mit Hilfe der Spezifikationsgrenzen in ein binäres Label überführt, sodass der Entscheidungsbaum zur Klassifikation von Gut- und Schlechttteilen eingesetzt werden konnte (Bild 2).

Um ein sog. Overfitting des Baumes, d.h. einen speziell auf den verwendeten Datensatz angepassten Baum, zu verhindern, wird eine Teilung des Datensatzes in Trainings- und Testmengen vorgenommen, um so allgemeingültigere Strukturen erkennen zu können. Bild 3 zeigt den Entscheidungsbaum des Anwendungsbeispiels nach An-

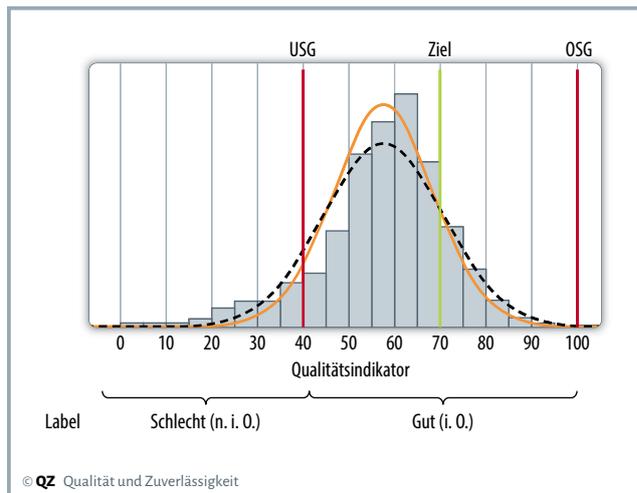


Bild 2. Bildung eines binären Qualitätslabels anhand der Spezifikationsgrenzen (Quelle: IPS, TU Dortmund)

wendung einer 10-fachen Kreuzvalidierung.

Die Gesamtklassifikationsgüte des Modells von 98,2 % gibt an, dass 98,2 % der Beobachtungen durch das Modell richtig klassifiziert, d.h. aufgrund der im Entscheidungsbaum verwendeten Entscheidungsregeln erklärt werden können. Durch die Analyse des Entscheidungsbaums ergibt sich analog zur Regressionsanalyse die Wertigkeit der Einflussfaktoren. Je weiter oben eine Variable im Entscheidungsbaum zu finden ist, desto größer ist ihr Einfluss bei der Klassifikation. Wie bei der klassischen Vor-

gehensweise ergeben sich hier die signifikanten Einflussgrößen Druck_E, Salz, Flüssigkeitsanteil und Temperatur_M (siehe Bild 3) mit einer Klassifikationsgenauigkeit von 98,2 %.

Die gewonnenen Erkenntnisse ermöglichen es, in der Improve Phase die Prozessparameter entsprechend einzustellen, um den Anteil fehlerhafter Produkte zu reduzieren. Während bei der klassischen Vorgehensweise optimale Parametereinstellungen zunächst durch Simulation hergeleitet werden müssen, ermöglicht der Einsatz von Ent-

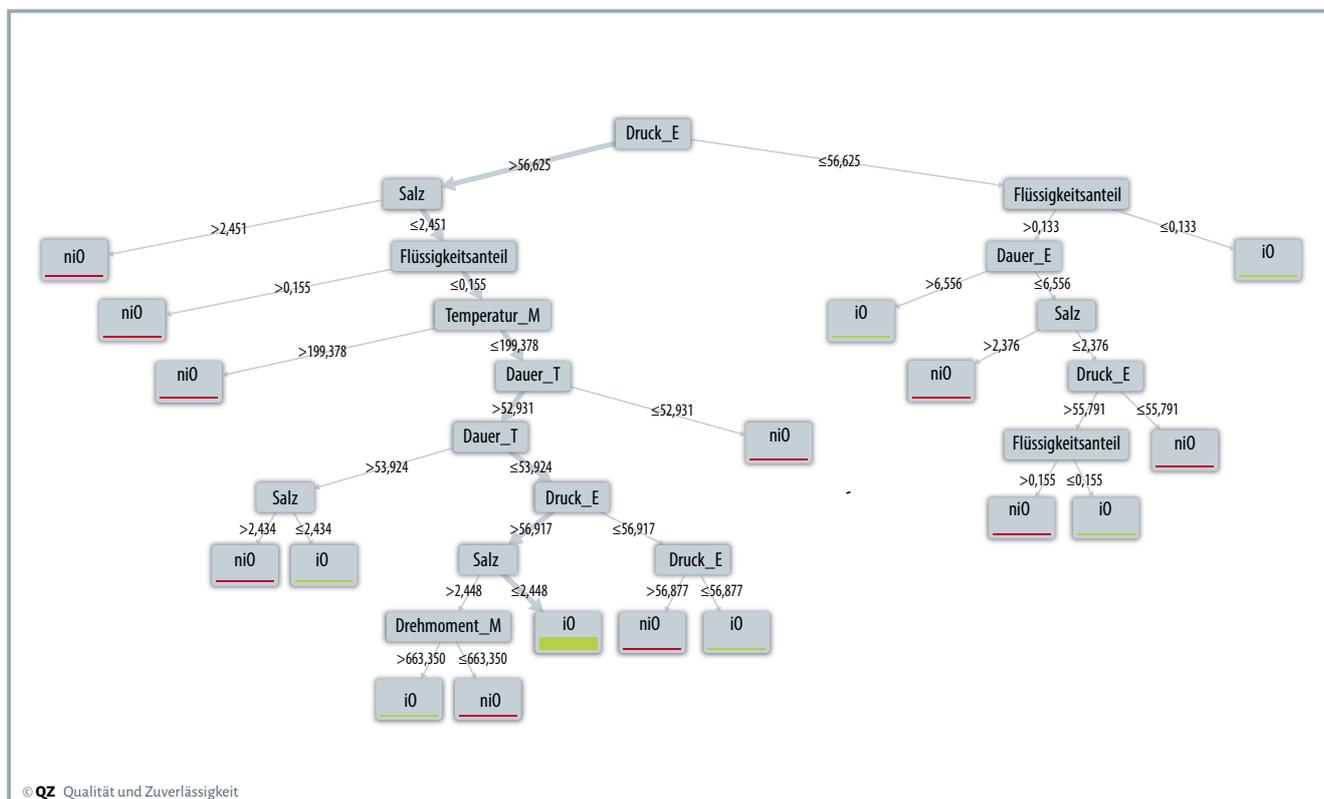


Bild 3. Entscheidungsbaum (Quelle: IPS, TU Dortmund)

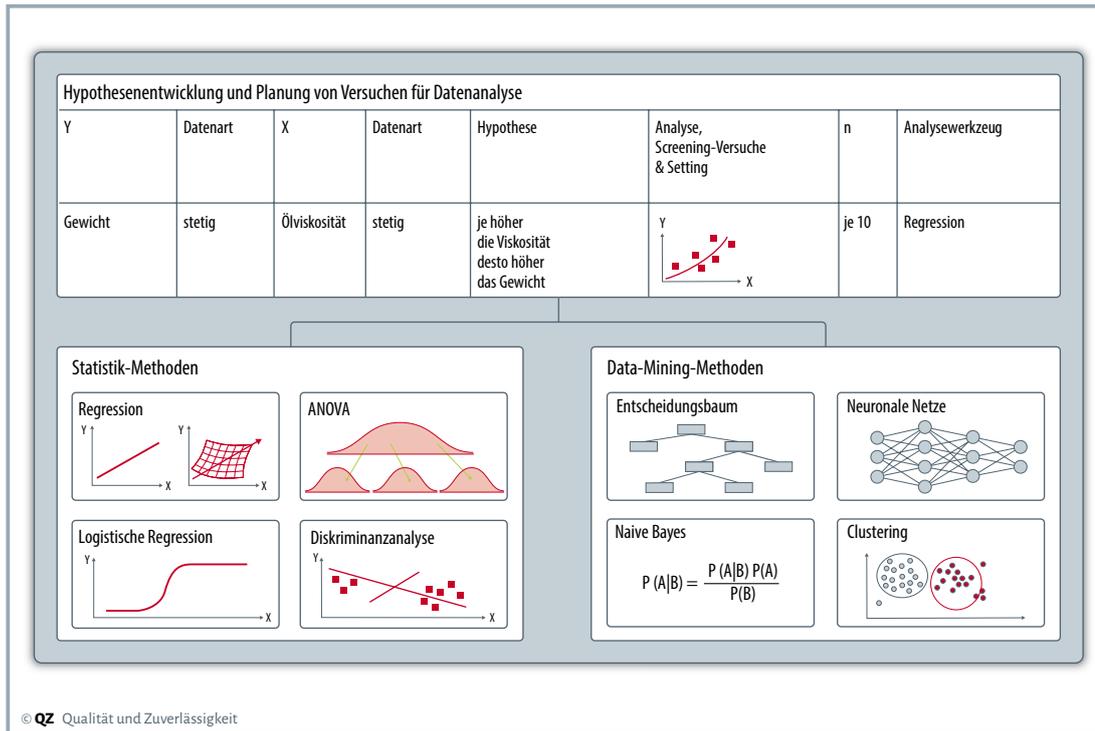


Bild 4. Erweiterung der DMAIC-Roadmap um Data-Mining-Methoden

(Quelle: mts Consulting Partner)

scheidungsbaumen das direkte Ablesen der einzustellenden Parameterwerte an den Verzweigungen des Baumes. Mit den neu eingestellten Prozessparametern wird im nächsten Schritt ein Verifizierungslauf durchgeführt und die Ergebnisse mit der Ausgangssituation der Measure Phase anhand von Fähigkeitsindizes verglichen.

Vorteile der Integration

Durch die Integration von Data Mining Verfahren in den DMAIC Zyklus wird eine Anwendung auch auf großen Datenmengen ohne manuellen Mehraufwand ermöglicht, sodass auch Zusammenhänge und Muster in großen Datenmengen komplexer Prozesse erkannt werden können. Weiterhin kann das explizite Einbeziehen von Datenverständnis und -vorverarbeitung Vorteile bei der Analyse von Ursache-Wirkungsbeziehungen und der Ursachenanalyse bringen. So können weitere Einflussgrößen und Ursachen erkannt werden, die bspw. im Ishikawa-Diagramm als Datensicht ergänzt werden können.

Der interdisziplinäre Ansatz erfordert ferner das Zusammenbringen der Experten aus verschiedenen Bereichen (Statistik, Datenanalyse, maschinelles Lernen, künstliche Intelligenz, Prozessexperten). Auf diese Weise wird die unternehmensinterne, abteilungsübergreifende Zusammenarbeit in Verbesserungsprojekten gefördert. Weite-

re Vorteile bestehen in der Visualisierung und Interpretierbarkeit der Verfahren. Während funktionale Zusammenhänge der Regressionsanalyse nur schwer interpretiert werden können, bieten beispielsweise Entscheidungsbaume eine einfache, gut verständliche Repräsentation von Strukturen in den Daten und entdeckten Entscheidungsregeln.

Zusammenfassung und Ausblick

Vor dem Hintergrund immer komplexerer Prozesse und großer Datenmengen bietet die Wissensgewinnung aus Daten durch Einsatz von Data Mining Methoden vielversprechende Potenziale, um die klassischen Methoden auch in Zukunft wirtschaftlich und effizient durch gezielte methodische und technologische Ergänzungen für Verbesserungsprojekte im Kontext von Industrie 4.0 und Big Data einsetzen zu können. Neben der methodischen Erweiterung der

Analyse Phase bieten sich auch in den weiteren Phasen des DMAIC Zyklus Erweiterungspotenziale durch den gezielten Einsatz von Big Data Technologien, bspw. zur automatisierten Erhebung und Vorverarbeitung von Daten, und Data Mining Tools, z.B. zur Plausibilitätsprüfung in der Measure Phase. Ebenso kann die Umsetzung und Kontrolle von Maßnahmen durch die fortschreitende Digitalisierung und Vernetzung beschleunigt und langfristig verbessert werden. Unterstützend wirkt hierbei auch eine Integration automatisierter Data Mining Verfahren zur Vorverarbeitung und Analyse der Daten, sodass eine fortlaufende Kontrolle ohne manuellen Analyseaufwand ermöglicht wird.

Ferner können Data Mining Verfahren zur Entdeckung von Anomalien und Prognose, bspw. von Qualitätsmerkmalen oder Prozessverhalten, eingesetzt werden, um präventive Maßnahmen ergreifen oder entsprechende Prozesseingriffe rechtzeitig ergreifen zu können. ■



Jetzt Karriere machen!

www.qz-online.de/stellenmarkt

HANSER



Portal für Qualitätsmanagement