

**Prof. Dr. Patrick Dienes**

**Professur für Controlling, Finanz- und Risikomanagement**  
Fakultät Betriebswirtschaft, Hochschule Landshut



# **Wissenschaftliche Arbeiten**

## **Vorgaben zur Dokumentation der Quellenarbeit**

Stand 03/2023  
(verbindlich für alle Anmeldungen ab SoSe 2023)

Alle in der Arbeit zitierten Quellen (außer Audio- und Videoquellen) müssen entweder in einem physischen ODER einem elektronischen Dokument (in pdf-Format) zusammengestellt werden. Dieses “Quellendokument” ist zusammen mit der Arbeit abzugeben. Das Quellendokument ist mit einem Titelblatt zu versehen (s. Anhang 1).

Zu jeder in einer Fußnote angegebenen Quelle ist die Seite bzw. sind die Seiten – und NUR diese Seite(n), welche die zitierte Stelle(n) enthält, fortlaufend in das Quellendokument aufzunehmen (z.B. Erstellen einer Kopie, eines Scans, eines Screenshots etc.) (s. Anhang 2).

Das Quellendokument folgt der Nummerierung der Fußnoten. Wird dieselbe Seite einer Quelle mehrmals in der Arbeit zitiert, so ist diese nur einmal in das Quellendokument unter der Fußnote aufgenommen werden, unter der die Quelle zum ersten Mal zitiert wird.

Bei Internetquellen, welche nicht in nummerierten Seiten strukturiert sind, ist statt der gesamten Seite ein Screenshot von der zitierten Stelle zu machen und in das Quellendokument aufzunehmen.

Im Anschluss an die Fußnoten sind in gleicher Weise die zitierten Quellen von Abbildungen, Tabellen und Anhang (in dieser Reihenfolge) ebenfalls einfach fortlaufend in das Quellendokument aufzunehmen (s. Anhang 3).

**Anhang 1: Quelldokument - Titelblatt**

**Quelldokument**

**zur Studien-/Bachelor-/Masterarbeit**

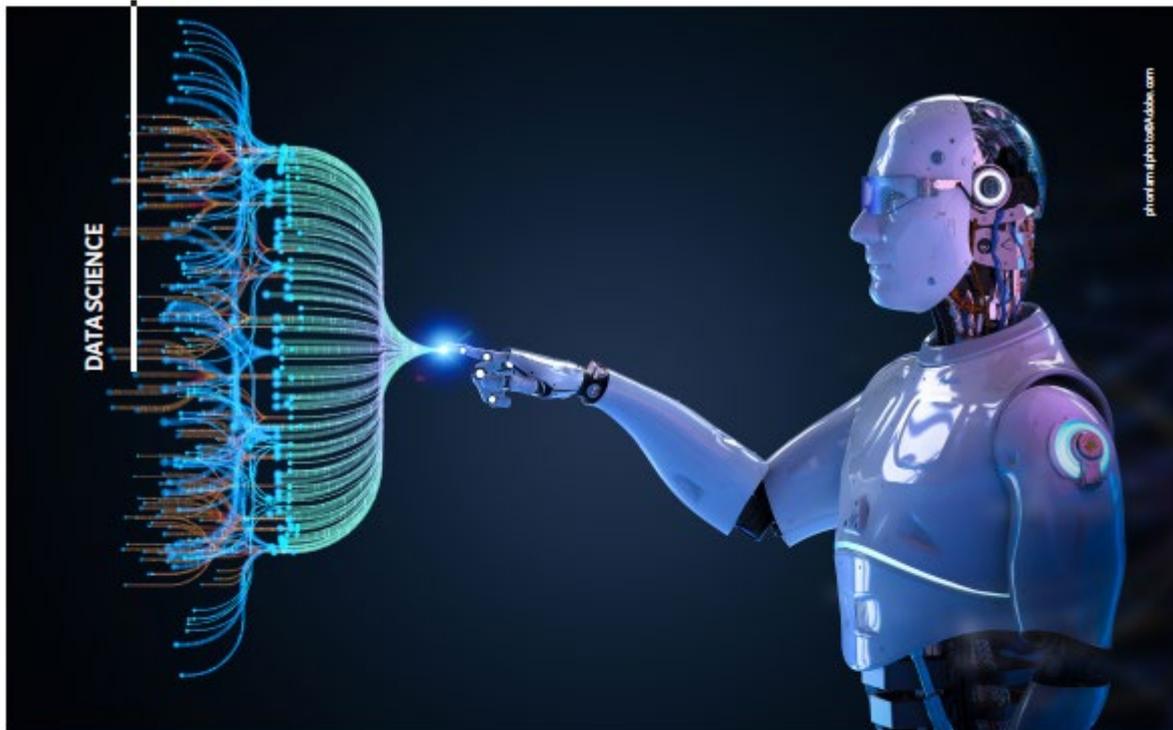
**[Titel der Arbeit]**

**vorgelegt von:**

**[Maxima Musterfrau]**

## Anhang 2: Quellen in Fußnoten - Beispiel

FN 1: Dieses/Helfrich (2022), S. 46–47.



# Data Science, auch für Controller!

Fallstudie zum Einsatz von Business Analytics – beispielhaft unter Verwendung von KNIME.

Patrick Dieses / Stefan Helfrich

Heutzutage sind Controller:innen ohne Excel-Kenntnisse kaum vorstellbar. Das gleiche könnte künftig in Bezug auf Kenntnisse in Business Analytics gelten. „Das Ziel von Business Analytics ist es, betriebswirtschaftliche Probleme im gesamten Managementzyklus von Planung, Steuerung und Kontrolle evidenzbasiert zu lösen.“<sup>1</sup> Der Einsatz von Business Analytics kann daher als essentiell für die Rationalitätssicherung, welche als Kernaufgabe des Controllings gilt, bewertet werden. In der WHU-Zukunftsstudie wird Business Analytics zwar erst seit 2017 als Zukunftsthema des Controllings angeführt, aber es wird hier der größte Bedeutungszuwachs (+1,8) unter allen Zukunftsthemen erwartet.<sup>2</sup>

In der Controlling-Praxis kommt Business Analytics bisher nur bei einer Minderheit der Unternehmen zum Einsatz, wie sich an dem besonders wichtigen Teilgebiet Predictive Analytics illustrieren lässt: Nur 5% der befragten Unternehmen haben solche Technologien erfolgreich im Controlling-Prozess implementiert, weitere 9% befinden sich in der Implementierungsphase.<sup>3</sup> Dabei gibt es leicht bedienbare Softwarelösungen, die – ohne das Schreiben von Programmcode zu erfordern – umfassende Data Science-Funktionalitäten bieten: Analysen werden mittels visueller Workflows zusammengestellt, so dass Controller:innen keine Data Scientists sein müssen, um Business Analytics einzusetzen.



### Summary

Intuitiv bedienbare Softwarelösungen wie KNIME Analytics Platform (kurz KNIME) erlauben Controller:innen das Potenzial von Data Science für ihre Zwecke zu erschließen, was anhand einer Fallstudie illustriert wird. Selbst wenn spezifische Data Science-Ressourcen im Unternehmen vorhanden sind, benötigen Controller:innen Basiskompetenzen, um erfolgreich ein synergetisches Zusammenwirken mit Data Scientists zu ermöglichen.

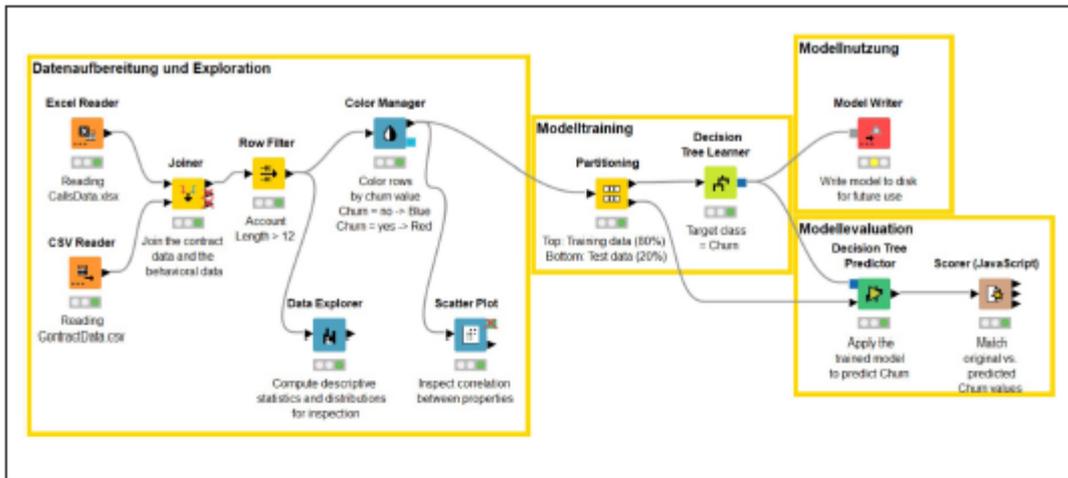


Abb. 1: KNIME Workflow zur Fallstudie

Dieser Artikel möchte – didaktisch eingebettet in eine Fallstudie – beitragen, Berührungspunkte der Controlling-Praktikerinnen mit Business Analytics abzubauen (beispielhaft anhand der kostenfreien und quell-offenen KNIME Analytics Plattform<sup>5</sup>).

Bei KNIME werden einfach sogenannte "Knoten" (als Icons dargestellt) aus einer Ablage ("Node Repository") per Drag & Drop in den Workflow gezogen und mit Pfeilen miteinander verbunden. Durch diese Verbindung einzelner Analyseschritte ("Knoten") wird der Datenfluss eines Prozesses rein visuell und ohne klassische Programmierung umgesetzt. Abb. 1 zeigt den KNIME Workflow zur Fallstudie. Das ganze Zahlenbeispiel sowie ein erläuternder Podcast zur Erstellung dieses Workflows sind online abrufbar (siehe Infobox), damit Sie die einzelnen Schritte selbständig reproduzieren können.<sup>5</sup>

### Fallstudie: Betriebswirtschaftliche Problemstellung

Die Abwanderungsvorhersage für Kunden (engl. "Churn Prediction") ist, insbesondere für Subskriptions-Geschäftsmodelle, ein höchst relevantes Anwendungsfeld von Machine Learning, um dem Verlust von Kunden entgegenzuwirken. Hierzu werden historische CRM-Daten benötigt, um Muster der Abwanderung bzw. Muster unter den Abwandernden zu erkennen: Zum Zeitpunkt einer Vertragsverlängerung in der Vergangenheit haben einige Kunden ihren Vertrag verlängert, andere nicht, d. h. sie sind abgewandert. Anhand der Unterschiede in den Historien zwischen abgewanderten und nicht abgewanderten Kunden wird ein Modell trainiert, das die Abwanderungswahrscheinlichkeiten für Kunden, bei denen Vertragsverlängerungen anstehen, vorhersagt.

Der CRM-Datensatz, den wir in diesem Artikel verwenden, besteht aus getrennt vor-

liegenden Vertragsdaten für 3.333 Telekommunikationskunden sowie deren Betriebsdaten. Die Vertragsdaten beinhalten u. a. die Vertragslaufzeit sowie Informationen, ob dem Vertrag eine Mailbox (Attribut VMail Message) zugeordnet ist. Weiterhin beinhaltet diese Tabelle das Attribut Churn, welches mit einem "no"-Eintrag einen erneuerten Vertrag enkodiert und mit "yes" einen gekündigten Vertrag. Die Betriebsdaten (gemittelte Werte aus Monatsdaten) umfassen z. B. die Anzahl an Anrufen und Telefoneminuten, aufgeteilt nach Tageszeiten, oder die Kosten für Anrufe tagsüber (Attribut Day Charge). Jeweils ein Auszug aus den beiden Tabellen ist in Abb. 2 ersichtlich.

### Auswahl des Machine Learning-Verfahrens

Machine Learning (ML) ist ein Teilgebiet der Künstlichen Intelligenz. ML bedient sich Algorithmen, mit denen aus Beispieldaten ein

Row ID	S Phone	I VMail Message	D Day Mins	D Eve Mins	I CustServ Calls	I Day Calls	D Day Charge
Row0	382-4657	25	265.1	197.4	1	110	45.07
Row1	371-7191	26	161.6	195.5	1	123	27.47
Row2	358-1921	0	243.4	121.2	0	114	41.38

Row ID	S Phone	I Account Length	S Churn	I Int'l Plan	I VMail Plan	S State	I Area Code
Row0	382-4657	128	no	0	1	KS	415
Row1	371-7191	107	no	0	1	OH	415
Row2	358-1921	137	no	0	0	NJ	415

Abb. 2: Auszug aus den Betriebsdaten (oben) sowie den Vertragsdaten (unten)

## Anhang 3: Quellen für Abbildungen/Tabellen/Anhang - Beispiel

Abb. 1: Dieses/Helfrich (2022), S. 50.

Scorer View  
Confusion Matrix

	no (Predicted)	yes (Predicted)	
no (Actual)	549	21	96.32%
yes (Actual)	26	71	73.20%
	95.48%	77.17%	

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa ( $\kappa$ )	Correctly Classified	Incorrectly Classified
92.95%	7.05%	0.710	620	47

Abb. 5: Konfusionsmatrix und Gesamtstatistik

gewandert sind oder nicht. Dabei werden Kunden, für die das Modell aufgrund deren Attributwerte eine Wahrscheinlichkeit für die Abwanderung von größer als 50 % ergibt, als Churn prognostiziert. Der Decision Tree Predictor erweitert schließlich die Tabelle der Testdaten um das Attribut Prediction (Churn) und weist jedem Kunden bzw. jeder Datenreihe entsprechend ein "yes" oder "no" als Wert hier zu.

Zum Schluss evaluieren wir die Modellgüte, indem wir den Prognosewert (Spalte Prediction (Churn)) mit dem Istwert (Spalte Churn) vergleichen. Hierzu bedienen wir uns des Knotens Scorer (JavaScript). Die Konfusionsmatrix (Abb. 5) ergibt für das Testdaten-Set, dass 77,17 % der als abwanderungsgefährdet klassifizierten Kunden tatsächlich abge-

wandert sind und 95,48 % der als nicht abwanderungsgefährdet klassifizierten Kunden auch tatsächlich Kunden geblieben sind, d. h. ihre Verträge verlängert haben. Insgesamt wird für 92,95 % der Kunden (im Testdaten-Set) ihr Verhalten korrekt prognostiziert. Wenn die Ergebnisse nicht zufriedenstellend sind, können Parameter in der Konfiguration des Decision Tree Learner optimiert oder andere ML-Algorithmen ausprobiert und die Modelle bzgl. ihrer Güte verglichen werden.

**Mehrwert: Fokussierte Kundenansprache**  
Wenn ein Modell die gewünschte Prognosequalität erreicht, kann dieses mit den optimierten Parametern gespeichert werden (Model Writer) und laufend zur Klassifikation der Kunden (abwanderungsgefährdet

vs. nicht abwanderungsgefährdet) eingesetzt werden. Abwanderungsgefährdete Kunden können so identifiziert, mit speziellen Marketing-Aktivitäten gezielt angesprochen und so hoffentlich zum Verbleib motiviert werden. Ressourcen (z. B. Rabatte, Vertriebsaktivitäten etc.) werden also nicht "gießkannenartig" auf alle Kunden verteilt, sondern auf die Kunden fokussiert, bei denen unter Einschluss von Opportunitätskosten (wie Verlust von Kundendeckungsbeitrag) die höchste Rendite erwartet wird. Umgekehrt werden auch Marketing-Aufwendungen für Kunden, die ohnehin nicht abwanderungsgefährdet sind, vermieden. Die Marketing-Effektivität und infolge die Rentabilität lassen sich so für Unternehmen erhöhen!

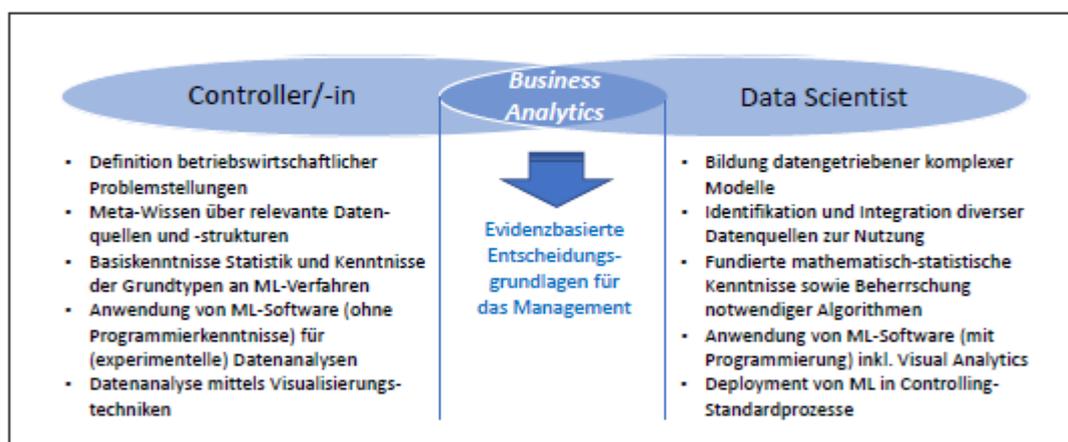


Abb. 6: Zielbild für synergetisches Zusammenwirken von Controller/in und Data Scientist<sup>38</sup>