

Ein agiles Vorgehensmodell zur Einführung von Predictive Analytics in Unternehmen

Jule Abmann, Joachim Sauer und Michael Schulz

NORDAKADEMIE Hochschule der Wirtschaft, Elmshorn, Deutschland
assmann.jule@guj.de,
{joachim.sauer, michael.schulz}@nordakademie.de

Zusammenfassung. Predictive Analytics dient dazu, unter Zuhilfenahme von statistischen Verfahren und auf Basis historischer und aktueller Daten Vorhersagen zu zukünftigen Ereignissen zu treffen. Anwendungsfälle für eine solche Form der Analyse sind in einigen Branchen bzw. Unternehmen zurzeit noch schwer zu finden. In diesem Artikel wird ein agiles Modell vorgestellt, das die Einführung von Predictive Analytics unterstützt und dabei die Unsicherheit minimiert. Damit wird eine Lücke zu etablierten Vorgehensmodellen geschlossen, die erst ansetzen, wenn die Entscheidung, eine Datenanalyse auf einer spezifischen Problemstellung durchzuführen, bereits getroffen wurde.

Der Entstehung des Modells liegen Erfahrungen zugrunde, die bei der Gruner + Jahr GmbH in einem Geschäftsbereich gesammelt wurden, der dem Bereich des Digital Marketings & Advertisings zuzuordnen ist. Eine Übertragbarkeit auf andere Branchen wird berücksichtigt.

Abstract. Predictive analytics is used to get insights about future events on the basis of historical and current data with the help of statistical methods. In some industries or companies such use cases are still hardly available. In this article an agile model is presented that supports the introduction of predictive analytics in companies to minimize prediction uncertainty. This closes a gap of established process models, which only start once the decision to perform a data analysis on a specific problem has already been made.

The development of the model is based on the experience gained at Gruner + Jahr GmbH in a business unit that is part of the Digital Marketing & Advertising industry. The application to other domains is considered.

Keywords: *Predictive Analytics, Anwendungsfälle, Digital Marketing & Advertising*

1 Einleitung

Predictive Analytics (PA) bietet die Möglichkeit, historische und aktuelle Daten zur Vorhersage zukünftiger Ereignisse zu nutzen. Obwohl PA bereits seit vielen Jahren als relevante Disziplin für Unternehmen allgemein akzeptiert ist, ist dies in einigen domänenspezifischen Anwendungen nicht immer der Fall. Im Gartner Hype Cycle für das Themengebiet „Business Intelligence & Analytics“ ist PA seit 2011 vertreten und hat den Gipfel der überzogenen Erwartungen bereits passiert [1]. Im Hype Cycle für das Themengebiet „Digital Marketing & Advertising“ (DMA) wird PA dagegen erst seit 2015 geführt und aktuell an der Spitze der Aufmerksamkeit eingeordnet; die Phasen der Erleuchtung und anschließender Produktivität haben noch nicht stattgefunden [2]. Die Entscheidung für den Einsatz neuer Technologien wie PA erfordert von Unternehmen Innovationsbereitschaft. Weniger risikoaffine Firmen suchen daher vor einer Durchführungsentscheidung häufig nach Use Cases (im Deutschen: Anwendungsfälle) mit einem erkennbaren wirtschaftlichen Nutzen (vgl. z. B. [3]). Diese sind unter anderem im Bereich des DMA jedoch noch nicht ohne größeren Aufwand zu identifizieren.

In dieser Arbeit wird ein Vorgehensmodell vorgestellt, das Firmen bei der Einführung von PA unterstützt und die Unsicherheit in dieser Anfangsphase minimiert. Dem Modell liegen dabei Erfahrungen zugrunde, die in einem Geschäftsbereich der Gruner + Jahr GmbH, der dem DMA-Bereich zuzuordnen ist und zu Beginn der Untersuchung noch keine Anwendung für PA besaß, gesammelt wurden. Damit besteht ein spezieller Branchenfokus; eine Übertragbarkeit des erarbeiteten Modells auf weitere Branchen und Unternehmen, die den Einsatz von PA prüfen, wurde jedoch berücksichtigt.

Es wurden möglichst viele Verfahren gewählt, um für den Einsatz von PA relevante Einsatzszenarien im betrachteten Unternehmen zu identifizieren: Nach einer Betrachtung theoretischer Grundlagen wird aufbauend auf einer strukturierten Literaturanalyse eine empirische Untersuchung in Form qualitativer Befragungen mit ausgewählten Personen aus der Vermarktung, sowohl innerhalb der Gruner + Jahr GmbH, als auch mit unternehmensfremden Vermarktungsakteuren, beschrieben. Diese Interviews zielten darauf ab, (Teil-) Automatisierungspotenziale und Ineffizienzen in den Tätigkeiten der Befragten zu identifizieren. Basierend auf den erzielten Ergebnissen konnten Anwendungsfälle für PA-Vorhaben erarbeitet werden, die im Kontext der Vermarktungsorganisation Einsatz finden können. Für diese Use Cases wurden Faktoren erhoben, die Anhaltspunkte zum Aufwand einer Implementierung darstellen. Ergänzt wurde die empirische Untersuchung durch eine Gruppendiskussion mit den Führungskräften aus der Gruner + Jahr GmbH, die darauf abzielte, ein im Sinne des Unternehmens übergeordnetes Verständnis zu den aus den qualitativen Befragungen hervorgehenden Problemen zu gewinnen, sowie die Relevanz der Use Cases einzuordnen und ihren Nutzen zu bewerten.

Aus den gewonnenen Erkenntnissen wurde anschließend ein generelles Vorgehensmodell für die Einführung von PA in Unternehmen abgeleitet, das in Abschnitt 4 vorgestellt und abstrakt beschrieben wird. Danach werden Erfahrungen bei der Umsetzung dargestellt und diskutiert. Die Arbeit schließt mit einer Bewertung der Ergebnisse und einem Ausblick ab.

2 Theoretische Grundlagen

In diesem Abschnitt werden Grundlagen von Predictive Analytics und etablierte Vorgehensmodelle zu deren Anwendung erläutert.

2.1 Predictive Analytics

Die Aufbereitung und Verwendung von Daten zur Entscheidungsunterstützung ist bis in die 1970er Jahre zurückzuführen [4]. Durch immer weiter zunehmende Datenmengen und die Verbesserung der zugrundeliegenden Technologien hat die Datenanalyse im letzten Jahrzehnt einen deutlich höheren Stellenwert erhalten als dies zuvor der Fall war [5]. In der Folge entstanden bzw. verbreiteten sich zahlreiche Methoden und Disziplinen, zu denen auch Predictive Analytics zählt.

Der PA-Begriff ist nicht einheitlich definiert. Da diese Herausforderung jedoch nicht den Fokus der vorliegenden Arbeit darstellt, wird an dieser Stelle auf die Literatur verwiesen (vgl. dazu z. B. [5–11]). In Anlehnung an bestehende Definitionen soll in der vorliegenden Arbeit unter Predictive Analytics eine Form der Datenanalyse verstanden werden, mit der unter Zuhilfenahme statistischer Verfahren auf Basis historischer und aktueller Daten Vorhersagen zu zukünftigen Ereignissen getroffen werden können.

PA basiert auf unterschiedlichen Methoden zur Aufdeckung von Ursache-Wirkungs-Beziehungen, deren Ursprung im Data Mining zu verorten ist [11]. In Abbildung 1 sind Methodengruppen und deren Eigenschaften dargestellt, in die ein Großteil der verwendeten Analyseverfahren eingeordnet werden können. Charakterisierende Merkmale einzelner Methoden, wie beispielsweise Einfachheit und Nachvollziehbarkeit, werden in der Matrix bewusst nicht als Kriterien erfasst. Die Ursache hierfür liegt in der Vielfalt statistischer Methoden, die je Gruppierung angewendet werden können und eine übergreifende Bewertung unmöglich macht.

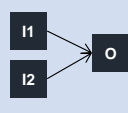
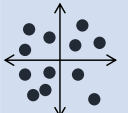
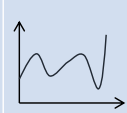

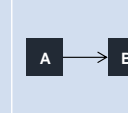
Eigenschaft	Prognose	Segmentierung	Zeitreihenanalyse	Klassifikation	Assoziation
Visualisierung					
Art	überwacht	unüberwacht	überwacht	überwacht	unüberwacht
Anwendungsargument	Vorhersage Output-Variable über Ursache-Wirkungs-Beziehung	Unterteilung Objekte notwendig, aber Klassen nicht bekannt	Zeitabhängigkeit der Daten (z.B. saisonale Schwankung)	Zuordnung Objekt zu festen Klassen erforderlich	(Unbekannte) Wenn-Dann-Beziehungen identifizieren
Schwächen	Vielzahl Input-Variablen → Kenntnis erforderlich	Einfluss von Ausreißern	Nur bei Zeitabhängigkeiten nutzbar	Klassifikationsfehler (falsch positiv oder falsch negativ)	Offensichtliche Assoziationen
Beispiel der Anwendung (marketing-spezifisch)	Vorhersage von Angebotsannahme / -ablehnung	Marktanalysen Kundenclustering	Umsatzforecasts	Identifikation Kunden / Nicht-Kunden	Identifikation Cross-Selling-Potenziale

Abbildung 1. Predictive-Analytics-Methodengruppen, in Anlehnung an [8, 11, 12]

In der wissenschaftlichen Literatur existieren Veröffentlichungen zu PA unter anderem in den Bereichen Luftfahrt, Gesundheitswesen, Einzelhandel, bei der Produkteinführung und Preisgestaltung [5, 12–15]. Darüber hinaus sind Ansätze zur Nutzung für marketingtreibende Unternehmen vorhanden [6]. Einsatzpotenziale speziell für die Medienbranche / Vermarktung sind in der wissenschaftlichen Diskussion dagegen bisher kaum thematisiert. Die Herausforderungen durch Unsicherheiten in der Entscheidung für oder gegen den PA-Einsatz sind für Unternehmen aus diesem Bereich daher besonders hoch, weshalb sich der Bereich gut für eine beispielhafte Betrachtung eignet.

2.2 Vorgehensmodelle zur Anwendung von Predictive Analytics

Predictive Analytics basiert in großen Teilen auf Techniken des Data Minings [7], weshalb sich in diesen beiden Bereichen verwendete Vorgehensmodelle nicht unterscheiden. Genannt werden können hier, bezogen auf die weite Verbreitung, vor allem das Modell *Knowledge Discovery in Databases* (KDD) [16], der *Cross Industry Standard Process for Data Mining* (CRISP-DM) [17] und der Prozess des *Sampling, Exploring, Modifying* und *Assessing* (SEMMA) [18]. All diese Vorgehensmodelle sind bewusst einfach gestaltet, sodass sie nicht nur von PA-Spezialisten, sondern auch von Domänenexperten verstanden werden und so ein Bearbeiten von Problemstellungen über unterschiedliche Interessengruppen und verschiedene Phasen, wie etwa der Datenaufbereitung und -analyse, hinweg ermöglichen.

Im Gegensatz zu KDD und SEMMA beinhaltet das CRISP-DM (vgl. Abbildung 2) mit dem *Business Understanding* eine Phase, die dazu dient, das Ziel und die Anforderungen eines Vorhabens aus Geschäftsperspektive zu verstehen, in eine für die Analyse geeignete Fragestellung zu übersetzen und einen Projektplan zur Umsetzung abzuleiten [17]. Ein umfangreiches Geschäftsverständnis ist essenziell, um eine Datenanalyse erfolgreich durchführen zu können [19]. Ebenfalls wird im CRISP-DM die Agilität der Datenanalyse deutlicher hervorgehoben, als dies in den anderen beiden genannten Vorgehensmodellen der Fall ist.

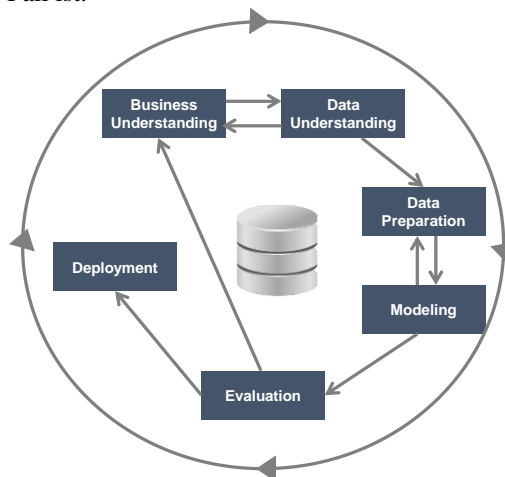


Abbildung 2. CRISP-DM Vorgehensmodell, aus [17]

Auch wenn alle genannten Vorgehensmodelle und vor allem das CRISP-DM viele relevante Aspekte eines PA-Projektes adressieren, setzen sie doch erst dort an, wo die Entscheidung für den Einsatz einer Datenanalyse auf eine spezifische Problemstellung bereits getroffen wurde. In dieser Arbeit soll ein Modell entwickelt werden, das früher, nämlich bei der Auswahl der am besten geeigneten Fragestellung für PA, ansetzt. Dabei sollen die wichtigsten Eigenschaften des CRISP-DM, wie Einfachheit, Agilität und die Möglichkeit des Einbezugs sämtlicher Interessengruppen, als Vorbild für die Entwicklung dienen. Dadurch wird auch eine kombinierte Anwendung dieser beiden Modelle möglich.

3 Forschungsdesign

Das Vorgehensmodell wurde iterativ bei der Gruner + Jahr GmbH entwickelt, da eine einfache Übertragbarkeit aus anderen Unternehmen nicht möglich ist und deshalb die Expertise der Mitarbeiterinnen und Mitarbeiter des spezifischen Unternehmens besonders berücksichtigt werden muss. Dabei wurden ein umfangreiches Literatur-Review, eine qualitative Befragung sowie eine abschließende Gruppendiskussion mit Experten genutzt (vgl. Abbildung 3). Dieses Vorgehen ermöglicht die umfangreiche Berücksichtigung verschiedener Perspektiven auf die Tauglichkeit von PA im spezifischen Fall und sichert durch die Einbeziehung von Interessengruppen zusätzlich die Akzeptanz von Entscheidungen.

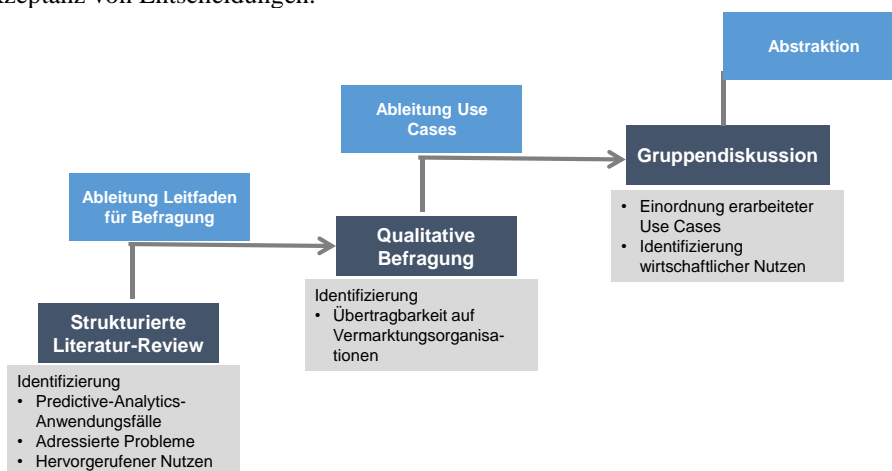


Abbildung 3. Forschungsvorgehen

Mit einem Literatur-Review wurden über die Datenbanken ScienceDirect, Beluga (Katalog der Hamburger Bibliothek), Springer-Link, EBSCO und ACM Digital Library Use Cases mithilfe der in Abbildung 4 dargestellten, aus jeweils drei Elementen bestehenden Suchbegriffe identifiziert.

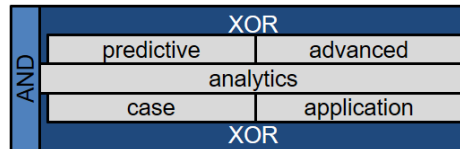


Abbildung 4. Muster der bei der Literaturrecherche verwendeten Suchbegriffe

Im ersten Prozessschritt konnten 30 Use Cases aus unterschiedlichen Branchen erkannt und auf ihren Ursprung und hervorgerufenen Nutzen durch eine PA-Anwendung hin untersucht werden. Eine komprimierte Übersicht der Use Cases befindet sich im Anhang dieses Artikels. Die identifizierten Effekte konnten zu den Kategorien Prozessoptimierung, Verkaufunterstützung, Wissen, Sicherheit und Kundenerlebnis gruppiert werden, die verschiedene Tätigkeiten eines Jobprofils ansprechen können.

Die anschließende qualitative Befragung wurde unter sechs Vermarktungsmitarbeitern durchgeführt, welche basierend auf den zuvor gebildeten Kategorien den Anteil der manuellen Arbeit und des Bauchgefühls für ihren Aufgabenbereich bewerteten. Als Resultat konnten fünf potenzielle PA-Use Cases für die Vermarktung ermittelt werden: Die Erstellung von Umsatzprognosen für verschiedene Vermarktungsblickwinkel, die optimale Bestimmung von Konditionen, die Identifikation von (Neu-)Kundenpotenzialen, die Identifikation von Cross-Selling-Potenzialen sowie die Unterstützung einer gezielten Kundenansprache.

Eine abschließende Gruppendiskussion unter Führungskräften der Vermarktung ordnete die erarbeiteten Use Cases übergreifend ein. Mithilfe einer Nutzwertanalyse wurden Aufwandstreiber (Anzahl Datenquellen, Art und Unabhängigkeit der Daten, Reifegrad, notwendige Kombination analytisches und fachliches Know-how) und Nutzenindikatoren (Effizienzsteigerung, Kostenreduktion, Erlössteigerung, Wettbewerbsvorteil, Risikominimierung, Entscheidungsrelevanz, Zukunftsrelevanz und Investitionsbereitschaft) für die Use Cases gegenübergestellt und diese vergleichbar gemacht. Für die erste Umsetzung von PA in DMA-Unternehmen wurden daraus resultierend die Use Cases Umsatzprognosen, Cross-Selling-Potenziale und eine Anwendung zur optimalen Bestimmung von Konditionen empfohlen.

Im Anschluss erfolgte für das Unternehmen Gruner + Jahr GmbH die Umsetzung zweier PA-Use Cases. Konform der Empfehlung der vorausgehenden Erarbeitung wurden zum einen Umsatzforecasts für verschiedene Vermarktungsblickwinkel erstellt. Zum anderen wurde die Identifizierung von Neukundenpotenzialen angewandt, die in der Gesamtbewertung aus Aufwand und Nutzen zwar nicht das höchste Ranking erhielt, jedoch mit einer zentralen Datenquelle auskommen kann und daher für die schnelle Ergebniserzielung besonders geeignet erschien. Auf Basis der realen Datengrundlage des Unternehmens konnten innerhalb eines Entwicklungszyklus aus wenigen Tagen schnelle Ergebnisse erzielt werden, die eine Evaluation der Machbarkeit der Use Cases zuließen. Es zeigte sich, dass für beide Use Cases zwar Ergebnisse erzielt werden konnten, das Vorhandensein von Daten in ausreichender Menge jedoch nicht für alle betrachteten Vermarktungsblickwinkel gegeben war. Dies wird als ein allgemeines Hindernis von PA-Vorhaben gesehen [5].

Nach einer kritischen Prüfung der Ergebnisse wurde über den weiteren Umgang entschieden. Use Cases können verworfen, in einem weiteren Zyklus angepasst, oder erweitert oder sofort in den Betrieb überführt werden. Für das Unternehmen Gruner + Jahr GmbH wurde entschieden, die Use Cases zunächst anzupassen und zu erweitern. Dafür wurden zum einen eine weitere Datenquellen angebunden, welche die Fehlermetriken der Prognose verringern konnte, sowie zum anderen das Prognoseergebnis um zusätzliche Informationen verfeinert. Das Ziel bestand darin, die Akzeptanz der Predictive-Analytics-Lösung durch inhaltliche Verbesserungen für eine nachfolgende Überführung in den Betrieb zu gewährleisten. Diese Überführung ist als Folgeaktivität geplant, jedoch zu diesem Zeitpunkt noch ausstehend. Vor der Weiterentwicklung der Use Cases wurde diese Überführung gegen die Umsetzung weiterer PA-Use Cases in Hinblick auf den zu erzielenden Nutzen und Aufwand erneut bewertet und gegeneinander abgewogen.

4 Beschreibung des agilen Modells

Aus den Erfahrungen des Einsatzes im Unternehmen wurde ein allgemeines Vorgehensmodell zur Einführung von Predictive Analytics in Unternehmen abstrahiert, das in Abbildung 5 dargestellt ist. Es ist nicht nur iterativ-inkrementell, sondern auch agil, da aus Erfahrungen aus der Nutzung gelernt werden und das Modell flexibel angepasst werden kann.

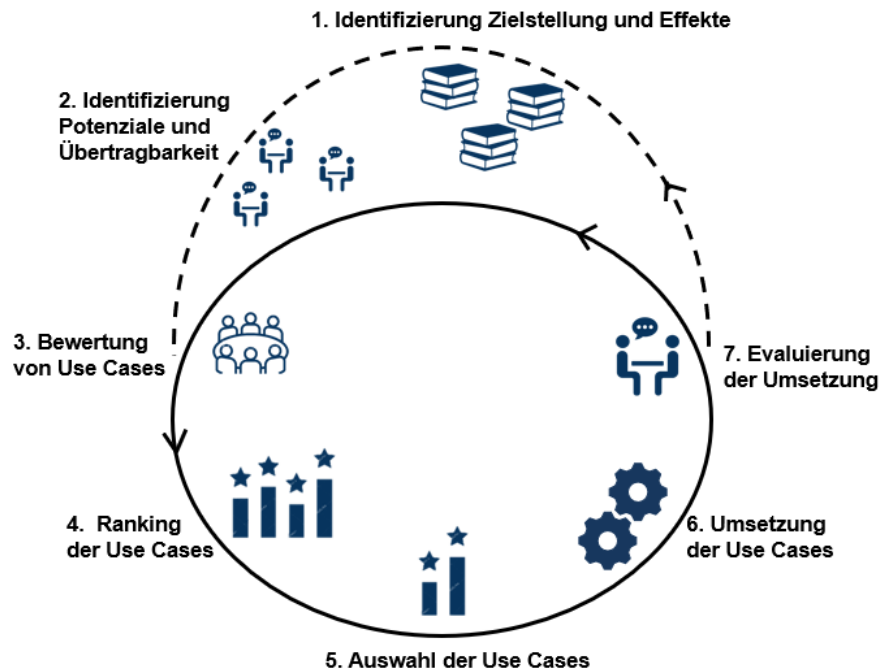


Abbildung 5. Vorgehensmodell zur Einführung von Predictive Analytics

Das Modell besteht aus sieben Schritten, die in zwei Zyklen angelegt sind. Anfangs werden die Schritte eins bis sieben durchlaufen. Abhängig von der Evaluierung im siebten Schritt wird der vollständige Zyklus erneut durchlaufen oder die vorhandenen Use Cases werden neu bewertet, was den Zyklus um zwei Schritte kürzt.

Schritt 1: Identifizierung Zielstellung und Effekte

Predictive Analytics sollte erst dann eingesetzt werden, wenn sinnvolle Einsatzpotenziale für das Unternehmen bzw. die Branche vorliegen. Um diese Beurteilung zu erreichen, sollten Verantwortliche zunächst die direkten Effekte und Zielstellungen von PA betrachten.

Schritt 2: Identifizierung Potenziale und Prüfung der Übertragbarkeit

Im zweiten Schritt sollte eine Prüfung der Übertragbarkeit der Zielstellungen auf das eigene Unternehmen bzw. die Branche hin erfolgen. Sofern es im Unternehmen Aufgaben und Prozessschritte gibt, welche die Effekte der PA optimieren können, sollten die Effekte mit Hilfe der folgenden Leitfragen konkretisiert werden. Das Potenzial des PA-Einsatzes kann dann, aufgeteilt auf zwei Dimensionen, ermittelt werden. Die erste Dimension betrifft die manuelle Arbeit: Ist die Aufgabe wiederkehrend, erfordert aber trotzdem viele manuelle Prozessschritte? Fehlen Informationen (= Wissen / Daten), welche die Prozesse beschleunigen könnten? Die zweite Dimension der Potenzial-

mittlung betrachtet das Bauchgefühl: Erfordern die Aufgaben viel Bauchgefühl, da Informationen (Wissen oder Daten) den Mitarbeitern nicht vorliegen? Ist die subjektive Einschätzung ein elementarer Bestandteil der Tätigkeit?

Die identifizierten Effekte können als Orientierungshilfe zur Ableitung von Handlungsempfehlungen in eine Matrix eingeordnet werden. Ein Beispiel, das bei Gruner + Jahr GmbH Anwendung findet und keinen Anspruch auf Übertragbarkeit besitzt, zeigt Abbildung 6.

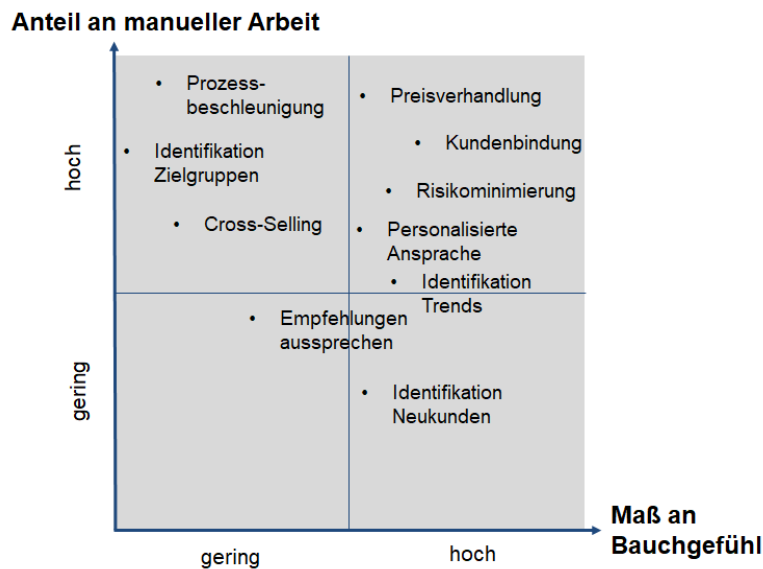


Abbildung 6. Einordnung Predictive-Analytics-Potenziale

In der Matrix spannen das Maß an Bauchgefühl und der Anteil an manueller Arbeit die Dimensionen auf. Wenn beides hoch ist, kann der Einsatz von PA besonders großen Nutzen schaffen (Beispiel Preisverhandlung). Für diesen Bereich sollten als erstes Use Cases erstellt werden. Wenn beides gering ist, kann auf PA verzichtet werden. Wenn nur der Anteil manueller Arbeit hoch ist, kann PA zur Effizienzsteigerung verwendet werden (Beispiel Prozessbeschleunigung); wenn nur das Bauchgefühl hoch ist, kann mit PA die Erfolgswahrscheinlichkeit erhöht werden (Beispiel Identifikation neuer Kunden). Die letzten beiden Bereiche sollten in Betracht gezogen werden, wenn Kapazität und Unternehmensfokus es zulassen.

Schritt 3: Bewertung von Use Cases

Die entstandenen Use Cases sind im Anschluss zu bewerten. Die Kriterien zur Bewertung von Nutzen und Aufwandstreibern sollten unternehmensindividuell identifiziert und gewichtet werden. Mögliche Nutzentreiber sind z. B. Erzielung von Effizienzsteigerung, Ermöglichung von Kostenreduktionen oder Minimierung von Risiken. Mögliche Aufwandstreiber sind z. B. die Anzahl Datenquellen, die Unabhängigkeit der

Daten (intern / extern) sowie der Reifegrad des Use Cases. Auch Chancen (neue Erlösmöglichkeiten, verbesserte Planungssicherheit, Steigerung der Profitabilität u. Ä.) und Risiken (Limitationen durch Datenschutz, Verfügbarkeit der notwendigen Fachkräfte, Auswirkung von Fehlern u. Ä.) sollten Kriterien zur Bewertung bilden.

Schritt 4: Ranking der Use Cases

Abschließend sind die entstandenen Use Cases auf Basis dieser Bewertung miteinander zu vergleichen. Dazu kann das Instrument der Nutzwertanalyse hilfreich sein, wobei die Bewertungskriterien in diesem Fall unternehmensindividuell gewichtet werden müssen. Je konkreter die Umsetzung und Rahmenbedingungen durch ein Unternehmen bekannt sind, desto eher kann zudem eine Quantifizierung von Aufwand und Nutzen erfolgen.

Schritt 5: Auswahl der Use Cases

Im fünften Schritt werden die Use Cases ausgewählt, die umgesetzt werden sollen. Da Abhängigkeiten zwischen den Use Cases bestehen können, müssen das nicht die am besten bewerteten Use Cases sein. Eine wichtige Abhängigkeit besteht beispielsweise im Personalbedarf zur Umsetzung.

Je nach Aufwand der einzelnen Use Cases muss auch geplant werden, wie viele Use Cases in diesem Zyklus umgesetzt werden sollen. Das kann insbesondere dann schwierig sein, wenn das Unternehmen noch nicht über ausreichende Erfahrungen mit der Implementierung von Use Cases verfügt.

Schritt 6: Umsetzung der Use Cases

Die ausgewählten Use Cases werden umgesetzt. Für Details sei hier auf etablierte Vorgehensmodelle wie das CRISP-DM verwiesen.

Schritt 7: Evaluierung der Use Cases

Im letzten Schritt werden die Use Cases unter Berücksichtigung der ursprünglich erstellten Bewertung kritisch evaluiert. Dabei sollten nicht nur alle Use Cases einzeln betrachtet werden; auch ein Vergleich verschiedener Use Cases ist notwendig, um daraus Schlüsse für das weitere Vorgehen zu ziehen.

Wenn die Umsetzung vom Unternehmen als erfolgreich angesehen wird und noch nicht alle identifizierten Potenziale genutzt wurden, erfolgt ein Übergang zu Schritt 3, bei dem die verbleibenden Use Cases im Licht der gesammelten Erfahrungen neu bewertet werden. Es ist auch möglich, schon (teilweise) implementierte Use Cases erneut aufzugreifen und zu verbessern.

Ansonsten beginnt die Bearbeitung wieder bei Schritt 1 mit der Identifizierung weiterer Zielstellungen und Effekte. Die Bearbeitung endet, wenn keine neuen Zielstellungen identifiziert werden können.

5 Erfahrungen mit der Umsetzung

Durch den Einsatz des Vorgehensmodells bei der Gruner + Jahr GmbH konnten fünf nutzenstiftende Use Cases identifiziert werden. Der mehrstufige Prozess und das Einbeziehen verschiedener Expertengruppen aus externer Literatur, internen Mitarbeitern und internen Führungskräften erhöhte dabei die Relevanz und Akzeptanz der entstandenen Use Cases deutlich und vermied so eine PA-Umsetzung, die hohen Aufwand bei geringem Nutzen bedeutet hätte. In Kombination mit der prototypischen Umsetzung wird diese Gefahr auf ein Minimum reduziert.

Die Anwendung des Modells im Unternehmenskontext zeigte, dass sich für Schritt 6 (Umsetzung der Use Cases) die Erstellung eines Prototyps anbietet, wenn die Unsicherheit hoch ist, ob die Daten ausreichend und in geeigneter Form vorhanden sind. In diesem Fall erlaubt eine prototypische Umsetzung auf Basis der Echt-Daten eine Einschätzung darüber. Im Anschluss kann eine finale Umsetzung, die höheren Aufwand bedeutet, mit einer Integration in das Business-Intelligence-System sowie einer regelmäßigen und automatischen Neu-Berechnung des Modells erfolgen.

Bereits bei der Umsetzung sowie Bewertung der Prototypen konnte die Relevanz einer Kombination aus analytischem und Domänenwissen deutlich erkannt werden, welche als eine der größten Herausforderungen von PA-Vorhaben gilt [20]. Ebenso zeigte sich, dass typische PA-Probleme, wie zu geringe Datenmengen oder siloartige Daten, die bisher keine Verknüpfung zulassen [5, 20, 21], frühzeitig identifiziert werden konnten.

Des Weiteren führt die praktische Umsetzung des Modells bei der Gruner + Jahr GmbH zu der Empfehlung, die verwendeten Datenquellen für den PA-Prototypen iterativ zu steigern, um schneller Zwischenergebnisse zu generieren. Nach Einbeziehen erster Datenquellen sollte ein weiterer Zyklus durchlaufen und eine Erweiterung um zusätzliche Datenquellen gegen die Umsetzung anderer Use Cases erwogen werden. In jedem Zyklus werden die Zwischenergebnisse zusammen mit Fachexperten ausgewertet, um ganz im agilen Sinne direktes Feedback zu sammeln und aus den Erfahrungen zu lernen.

Ein weiterer Vorteil, den die praktischen Erfahrungen aufdecken konnten, sind die verschiedenen Berührungspunkte, welche die relevanten Führungskräfte vor einer Durchführungsentscheidung mit PA haben. Die für die Unternehmensvertreter neuartige Thematik wurde so greifbarer – Unsicherheiten konnten reduziert und die Investitionsbereitschaft erhöht werden.

Gleichzeitig sollte ein besonderer Fokus darauf gelegt werden, eine Überführung in den Betrieb zeitnah zu realisieren – ebenfalls den agilen Prinzipien entsprechend. Andernfalls besteht die Gefahr vieler gute Prototypen ohne eines nutzbaren PA-Produktes. Für die Überführung in den Betrieb muss insbesondere dem Prozess und Change Management eine besondere Bedeutung zugeschrieben werden, da die praktische Umsetzung zeigt, dass einer PA-Lösung, die mit Datenwissen ein Bauchgefühl ersetzen oder ergänzen soll, Skepsis entgegengebracht wird.

6 Fazit

Die erste Erfahrung mit dem praktischen Einsatz des Modells zeigt, dass es geeignet ist, um mit den Herausforderungen eines Prozesses zur Einführung von Predictive Analytics in geeigneter Form umzugehen. Die Lücke etablierter Vorgehensmodelle wie KDD, SEMMA und CRISP-DM, die erst ansetzen, wenn die Entscheidung einer Datenanalyse auf eine spezifische Problemstellung bereits getroffen wurde, kann somit geschlossen werden. Zukünftig sollten weitere Umsetzungen auch in anderen Unternehmen bzw. mit anderen Analyseproblemstellungen erfolgen, um das Modell zu überprüfen und weiterzuentwickeln.

Das Modell gibt keine Aussage darüber, ob bzw. nach wie vielen Jahren der Nutzen der ermittelten PA-Use Cases für DMA-Unternehmen den Aufwand übersteigen wird. Diese Aussage ist nur unternehmensindividuell zu treffen, da organisatorische und technische Gegebenheiten Basisfaktoren für den Aufwand darstellen.

Für zukünftige Forschungen bietet es sich daher an, auf Basis einer Implementierung den tatsächlichen Nutzen messbar zu machen sowie einen Vergleich zum erwarteten Nutzen herzustellen. Darauf aufbauend kann schließlich eine valide Aussage zur Wirtschaftlichkeit getroffen werden, da der Aufwand konkret beziffert werden kann. Auch können Risikofaktoren, wie beispielsweise die Akzeptanz unter potenziellen Anwendern, differenzierter bewertet werden.

Das Vorgehen zur Einführung von PA, das in dieser Arbeit praktisch für die Gruner + Jahr GmbH angewendet wurde, enthält ein Abstraktionsniveau, das es zulässt, dieses Vorgehen auch auf andere Branchenbetrachtungen zu übertragen. Ebenso kann das Vorgehen als Orientierungshilfe für die Prüfung der Einsatzpotenziale weiterer Disziplinen neben Predictive Analytics dienen. Ein denkbare Szenario ist die Prüfung der Einsetzbarkeit von Künstlicher Intelligenz, da dies aktuell für immer mehr Unternehmen relevant wird.

Insbesondere für DMA-Unternehmen wird PA voraussichtlich in Zukunft ein entscheidender Baustein werden, um langfristig wettbewerbsfähig zu bleiben. Die gezielte Anwendung eines Vorgehensmodells zur erfolgreichen Einführung von PA wird daher in Zukunft essenziell sein, um die relevanten Use Cases zu identifizieren und zu bewerten sowie die Entscheidung für aber auch gegen den Einsatz von PA zu unterstützen. Nach erfolgreicher Durchführung erster Analysevorhaben muss das Vorgehensmodell allerdings auch um eine projektübergreifende, strategische Struktur ergänzt werden, die eine dauerhafte Etablierung der Disziplin in die Organisation ermöglicht.

References

1. Gartner, I.: Hype Cycle for Business Intelligence, 2011, <https://www.gartner.com/doc/1766215/hype-cycle-business-intelligence->
2. Gartner, I.: Hype Cycle for Digital Marketing and Advertising, 2018, <https://www.gartner.com/doc/3884103/hype-cycle-digital-marketing-advertising>
3. Next Generation Predictive Analytics. Using Forward-Looking Insights to Gain Competitive Advantage (2015)

4. Gluchowski, P.: Business Analytics – Grundlagen, Methoden und Einsatzpotenziale. HMD 53, 273–286 (2016)
5. Attaran, M., Attaran, S.: Opportunities and Challenges of Implementing Predictive Analytics for Competitive Advantage. *International Journal of Business Intelligence Research* 9, 1–26 (2018)
6. Leventhal, B.: Predictive analytics for marketers. Using data mining for business advantage. KoganPage, London, New York, NY, New Delhi (2018)
7. Kridel, D., Dolk, D.: Automated self-service modeling: predictive analytics as a service. *Inf Syst E-Bus Manage* 11, 119–140 (2013)
8. Halper, F.: Predictive Analytics for Business Advantage. Best Practices Report (2014)
9. Gandomi, A., Haider, M.: Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management* 35, 137–144 (2015)
10. Eckerson, W.: Predictive Analytics, <https://tdwi.org/Articles/2007/05/10/Predictive-Analytics.aspx>
11. Chamoni, P.: Advanced Analytics: Eine Annäherung. *BI Spektrum* 12., 8–10 (2017)
12. Siegel, E.: Predictive analytics. The power to predict who will click, buy, lie, or die. John Wiley & Sons, Hoboken, New Jersey (2016)
13. Janke, A.T., Overbeek, D.L., Kocher, K.E., Levy, P.D.: Exploring the Potential of Predictive Analytics and Big Data in Emergency Care. *Annals of emergency medicine* 67, 227–236 (2016)
14. Granovsky, L., Kamienchick, R., Yacovzada, N., Viswanathan, P., Cao, S., Alevras, D., Tamir, R., Grossman, I., Ferro, T., Chary, D., et al.: Using Predictive Analytics to Identify Risk of Clinical Asthma Exacerbations. *Journal of Allergy and Clinical Immunology* 141, AB222 (2018)
15. Bradlow, E.T., Gangwar, M., Kopalle, P., Voleti, S.: The Role of Big Data and Predictive Analytics in Retailing. *Journal of Retailing* 93, 79–95 (2017)
16. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37–37, 37 (1996)
17. Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. and Wirth, R.: CRISP-DM 1.0, <https://www.the-modeling-agency.com/crisp-dm.pdf>
18. SAS: Introduction to SEMMA, <https://documentation.sas.com/?docsetId=emref&docsetTarget=n061bzurmej4j3n1jnj8bbjmm1a2.htm&docsetVersion=14.3&locale=en>
19. Shearer, C.: The CRISP-DM Model: The New Blueprint for Data Mining. *Journal of Data Warehousing*, 13–22 (2000)
20. Iffert, L., Bange, C., Mack, M., Vitsenko, J.: Advanced & Predictive Analytics. Schlüssel zur zukünftigen Wettbewerbsfähigkeit (2016)
21. TDWI (ed.): Ten Mistakes to Avoid in Predictive Analytics Efforts (2015)

Anhang 1: Auswertung Literatur-Review

Quelle	Art	Branche	Ursprung	Effekt	Cluster	Use Case	Nutzen	Methoden	Praktisch
Abbar et al. 2018 Altaran und Altaran 2018, S. 22f	Journal	Medien	Hoher Konkurrenzdruck führt zu gesteigertem Interesse, die Performance von journalistischem Online-Inhalten vor Veröffentlichung zu kennen	Reichweiteroptimierung	Prozessoptimierung	Vorhersage der Reichweite eines Artikels vor Veröffentlichung	Entscheidungsunterstützung Promoten von bestimmten Artikeln Optimierung des Artikels zur Erhöhung einer höheren Reichweite	Clustering Zeitreihenanalyse	x
Altaran und Altaran 2018, S. 22f	Journal	Einzelhandel	Merk- und Kundenstruktur unklar Es existieren Produkte, die häufig zusammengekauft werden	Optimierung Targeting Crossselling zur Umsatzsteigerung	Verkaufsunterstützung Verkaufsunterstützung	Markt- und Kundensegmentierung Warenkorbanalyse Trends im Einkaufsverhalten erkennen	Angebote bedarfsgerechter steuern durch Positionierung im Markt Crosssellingmöglichkeiten erkennen und anbieten Neuproduktempfehlungen steuern Cross- und Upsellingmöglichkeiten vorbereiten und steuern	Assoziation	
Altaran und Altaran 2018, S. 22f	Journal	Einzelhandel	Bestimmte Faktoren (z.B. Weiter) beeinflussen das Einkaufsverhalten Hohe Auslastung in Flugzeügen erstrebenswert, höher	Identifikation von Trends Preisfindung	Wissen	Optimalen Verkaufspreis zu verschiedenen Zeitpunkten ermitteln Erkennen von Bedürfnissen der Kunden und dazugehörigen Kunden-Loyaltäts-Angeboten	Erzielung des höchstmöglichen Umsatz durch Auslastung und Preisgestaltung Erfolgreiche Kunden-Loyaltätsprogramme für den zu Umsatzsteigerung und Kundenbindung		
Altaran und Altaran 2018, S. 22f	Journal	Luftfahrt	Hoher Konkurrenzdruck der Branche erhöht Relevanz der Kundenbindung Loyalitätsprogrammen führen zur Schwerebelastung	Kundenbindung	Verkaufsunterstützung	Erkennung der optimalen Route und Lieferstrukturen Erfolgreiche Auswahl neuer Mitarbeiter	Optimale Lieferwege reduzieren Kosten und beschleunigen den Transport Optimale Ressourcennutzung führt zu Umsatz- und Effizienzsteigerungen		
Altaran und Altaran 2018, S. 22f	Journal	Logistik	Auswahl geeigneter neuer Mitarbeiter meist schwierig	Effizientere Ressourcennutzung	Prozessoptimierung	Mitarbeiterboni optimal steuern Optimierung der vorgeschlagenen Personen zum Vernetzen	Soziale Ausschüttung der Boni mit dem meisten Nutzen zu den geringsten Kosten Je genauer die vorgeschlagenen Personen, desto höher die Bindung des Users mit dem sozialen Netzwerk		
Altaran und Altaran 2018, S. 22f	Journal	Human Resources	Mitarbeiterbindung und -bindung in sozialen Netzwerken werden vorgeschlagen Teils Vergabe von Krediten an kreditunwürdige Kunden Fehlentscheidung ruft hohen Schaden hervor	Mitarbeiterbindung Empfehlungsoptimierung	Verkaufsunterstützung Kundenerlebnis	Optimierung der vorgeschlagenen Personen zum Vernetzen	Soziale Ausschüttung der Boni mit dem meisten Nutzen zu den geringsten Kosten Je genauer die vorgeschlagenen Personen, desto höher die Bindung des Users mit dem sozialen Netzwerk		
Altaran und Altaran 2018, S. 22f	Journal	Online	Personen zum Vernetzen vorgeschlagen Teils Vergabe von Krediten an kreditunwürdige Kunden Fehlentscheidung ruft hohen Schaden hervor	Empfehlungsoptimierung	Verkaufsunterstützung Kundenerlebnis	Optimierung der vorgeschlagenen Personen zum Vernetzen	Soziale Ausschüttung der Boni mit dem meisten Nutzen zu den geringsten Kosten Je genauer die vorgeschlagenen Personen, desto höher die Bindung des Users mit dem sozialen Netzwerk		
Altaran und Altaran 2018, S. 22f	Journal	Finanzwesen	Personen zum Vernetzen vorgeschlagen Teils Vergabe von Krediten an kreditunwürdige Kunden Fehlentscheidung ruft hohen Schaden hervor	Empfehlungsoptimierung	Verkaufsunterstützung Kundenerlebnis	Optimierung der vorgeschlagenen Personen zum Vernetzen	Soziale Ausschüttung der Boni mit dem meisten Nutzen zu den geringsten Kosten Je genauer die vorgeschlagenen Personen, desto höher die Bindung des Users mit dem sozialen Netzwerk		
Bauckhege et al. 2017	Journal	Medien	Volatile und saisonale Schwankungen bei Newsblogs bisher unverständlich Ziel von Werbetreibenden ist das möglichst umfassende Erreichen einer Zielgruppe über verschiedene Kanäle zu geringen Kosten Preise der Produkte bestimmen Profitabilität eines Stores	Reichweiteroptimierung Reichweiteroptimierung Effizientere Ressourcennutzung	Prozessoptimierung Prozessoptimierung	Vorhersage der Kreditwürdigkeit inaktiven bei Newsblogs Forecasting-Modell zu User-Engagement Optimierung des Media-Mix aus TV, Print und Internet für Werbausspielung Anpassung des Preises von Einzelhandelsprodukten	Risikominimierung durch verbesserte Einschätzung der Kreditwürdigkeit gezielte Marketingstrategien neue oder loyale Anwender Entscheidungsunterstützung	Zeitreihenanalyse	x
Bhaduri 2016, S. 47ff	Monographie	Medien	Volatile und saisonale Schwankungen bei Newsblogs bisher unverständlich Ziel von Werbetreibenden ist das möglichst umfassende Erreichen einer Zielgruppe über verschiedene Kanäle zu geringen Kosten Preise der Produkte bestimmen Profitabilität eines Stores	Reichweiteroptimierung Reichweiteroptimierung Effizientere Ressourcennutzung	Prozessoptimierung Prozessoptimierung	Vorhersage der Kreditwürdigkeit inaktiven bei Newsblogs Forecasting-Modell zu User-Engagement Optimierung des Media-Mix aus TV, Print und Internet für Werbausspielung Anpassung des Preises von Einzelhandelsprodukten	Risikominimierung durch verbesserte Einschätzung der Kreditwürdigkeit gezielte Marketingstrategien neue oder loyale Anwender Entscheidungsunterstützung	Zeitreihenanalyse	x
Bradlow et al. 2017	Journal	Einzelhandel	Geringe Überwachung von Asthma-Patienten und somit späte/keine Identifikation von deutlichen Verschlechterungen des Krankheitsstadiums	Preisfindung	Verkaufsunterstützung	Frühzeitige Erkennung von Asthma-Patienten mit Risiko einer deutlichen Verschlechterung des Krankheitsstadiums	Gezieltere Werbausspielung reduziert Kosten bei gleichzeitiger Erhöhung der Reichweite und somit Wirksamkeit Optimale Preisgestaltung	Klassifikation	Bayes x
Granovsky 2018	Journal	Medizin	Geringe Überwachung von Asthma-Patienten und somit späte/keine Identifikation von deutlichen Verschlechterungen des Krankheitsstadiums	Optimierung Targeting	Verkaufsunterstützung	Frühzeitige Erkennung von Asthma-Patienten mit Risiko einer deutlichen Verschlechterung des Krankheitsstadiums	Gezieltere Werbausspielung reduziert Kosten bei gleichzeitiger Erhöhung der Reichweite und somit Wirksamkeit Optimale Preisgestaltung	Klassifikation	Klassifikation x

Quelle	Art	Branche	Ursprung	Effekt	Cluster	Use Case	Nutzen	Data-Mining-Methode	Details	Praktisch
Jan. 2016	Journal	Handel	Geringe Erfolgsquote von neu eingeführten Produkten trotz hohen Umsatzpotenzials	Verständnisgewinn	Wissen	Schnelles Erkennen der Ursache für geringe Nachfrage neuer Produkte	Gezielte und schnelle Reaktion möglich: Veränderung des Produkts oder der Kommunikation	Social-Media-Analyse	Text-Analyse	x
Litsey 2017	Journal	Bibliothek	Lange Wartezeiten in Bibliotheken bei Ausleihe von Büchern und viele Interlibrary-Transaktionen	Beschleunigung	Prozessoptimierung	Automatische Erkennung zukünftiger Bedarfe	Bestand wird frühzeitig auf Bedürfnisse angepasst	Clustering	k-means	x
Munoz 2018	Journal	Luftfahrt	Viele Stakeholder an Luftfahrt beteiligt; ungenaue Ankunftszeiten führen zu hohen Kosten	Effizientere Ressourcennutzung Sicherheitsenhöhung	Prozessoptimierung Sicherheit	Live-System zur Schätzung von Online Kaufverhalten in Echtzeit Flugankürfen in Echtzeit	Reduzierung Sicherheitsrisiken Bessere Ressourcennutzung	Regression	Neuronale Netze	x
Niklaus et al. 2017	Journal	Handel	Hoher Konkurrenzdruck bei Online-Shops	Erhöhung Verkaufschance	Verkaufsunterstützung	Kundenverhalten ableiten	Proaktive Ansprache von Online-Kunden	Prognose		x
Prill et al. 2017	Journal	Industrie	Maschinenfälle kostenintensiv und risikoreich	Effizientere Ressourcennutzung Risikominimierung	Prozessoptimierung Sicherheit	Erkennen von zu wartenden Komponenten vor Ausfall	Verhinderung von Downtimes durch rechtzeitige und gezielte Wartung führt zu Kostenreduktion	Klassifikation		x
Siegel 2016, S.296f	Monographie	Öffentlicher Dienst	Polizei oft zu spät an Tatorten; Streifenwagen absichern	Sicherheitsenhöhung	Sicherheit	Potenziell nächste Orte für Kriminaldelikte erkennen	Streifenwagen gezielt gesteuert			
Siegel 2016, S.296f	Monographie	Versicherung Finanzwesen	Nicht alle Transaktionen können durch Banken und Versicherungen gleichermaßen überprüft werden	Effizientere Ressourcennutzung	Prozessoptimierung Sicherheit	Betrugserkennung von Transaktionen / Anwendungen	Gezielte Überprüfung der Transaktionen mit höherer Betrugswahrscheinlichkeit			
Siegel 2016, S.296f	Monographie	Online	Spam-E-Mails stellen Sicherheitsrisiko für den Anwender da	Sicherheitsenhöhung Betrugsminderung	Sicherheit	Erkennung von Spam-E-Mails	Verschieben der potenziellen Spam-E-Mails in einen separaten Ordner			
Siegel 2016, S.296f	Monographie	Human Resources	Firmen wollen (überraschende) Kündigungen von guten Mitarbeitern verhindern	Mitarbeiterbindung	Wissen	Erkennung der Mitarbeiter, die potenzielle Kündiger darstellen	Berücksichtigung der potenzielle Kündiger durch Manager			
Siegel 2016, S.296f	Monographie	Übergeordnet / Vertrieb	Umsatzverluste durch kündigende Kunden können nicht immer durch Neukunden kompensiert werden	Identifikation von Trends	Verkaufsunterstützung	Churn-Analyse: Vorhersage der Kunden, die kündigen / abspringen werden	Gezielte Aussteuerung von Marketing-/ Kundenbindungsmaßnahmen für die gefährdeten Kunden			
Siegel 2016, S.296f	Monographie	Online	Bestimmte Plattformen bieten Filmvorschläge für Nutzer an	Kundenbindung	Verkaufsunterstützung	Vorschlag von Filmen auf Basis der angeschauten und bewerteten Filme	Hoher Trefferquote in den Vorschlagswerten erhöht Kundenbindung			
Siegel 2016, S.296f	Monographie	Übergeordnet / Marketing	Ansprache vieler Marketingkontakte kosten- und ressourcenintensiv	Effizientere Ressourcennutzung Optimierung Targeting	Prozessoptimierung Verkaufsunterstützung	Erkennen der Marketingkontakte mit der höchsten Response-Wahrscheinlichkeit	Das gezielte Kontaktieren der Kontakte mit der höchsten Wahrscheinlichkeit einer Reaktion spart Kosten und Ressourcen			
Siegel 2016, S.296f	Monographie	Medien	Ausspielung von Werbung wirkungslos, wenn keine Interaktion erfolgt	Effizientere Ressourcennutzung Optimierung Targeting	Prozessoptimierung Verkaufsunterstützung	Ausspielung von Online-Werbung mit der höchsten Interaktionswahrscheinlichkeit	Hohe Interaktion mit Werbung verspricht Marketing Erfolg			
Wedel und Kamann 2016	Journal	Online	Kunden haben individuelle Bedürfnisse, die über generische Angebote nicht befriedigt werden können	Personalisierung	Kundenerlebnis	Individuelle Angebote für Kunden erstellen	Kundenindividuelle Angebote versprechen höhere Abschlussraten			
Yasmin 2013	Journal	Öffentlicher Dienst	Unvollständigen wollen hohe Abruchquote, insbesondere bei Fernstudiengängen, vermeiden	Studentenbindung	Verkaufsunterstützung	Studienabbruch bedingten und darüber Vorhersage von potenziellen Abbruchstadien	Potenzielle Abbruchstadien können frühzeitig erkannt werden und mit Maßnahmen unterstützt werden	Klassifikation		x