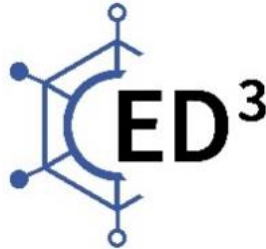


Anleitung zur Steigerung der Akzeptanz mittels transparenter Erklärungen



Diese Anleitung ist ein Ergebnis des Forschungsprojekts ED³ (IGF 20694 N) des IPRI - International Performance Research Institute und des Lehrstuhls für Strategie und Organisation der Technischen Universität München.



**IPRI – International Performance
Research Institute gGmbH**

Königstraße 5, 70173 Stuttgart



**Lehrstuhl für Strategie und
Organisation der Technischen
Universität München**

Arcisstraße 21, 80333 München

Das IGF-Vorhaben 20694 N der Forschungsvereinigung Institut für Energie- und Umwelttechnik e.V. – IUTA wird über die AiF im Rahmen des Programms zur Förderung der industriellen Gemeinschaftsforschung (IGF) vom Bundesministerium für Wirtschaft und Energie (BMWi) aufgrund eines Beschlusses des Deutschen Bundestages gefördert.

Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages

Einleitung

Die Verständlichkeit statistischer Modelle und Modelle des maschinellen Lernens wird als ein wichtiger Faktor für die Akzeptanz von Business Analytics (BA) erachtet (Martens und Provost 2014). Während einige BA-Systeme dem Nutzer die Möglichkeit bieten, die dem BA-System und seinen Analysen zugrundeliegenden Annahmen oder Einschränkungen zu verstehen (Shollo und Galliers 2016), sind komplexere Systeme oftmals unverständlich (Breiman 2001). Die oft als „Black-Box“-Algorithmen bezeichneten Algorithmen sind leistungsfähige Prognosewerkzeuge. Jedoch müssen Entscheidungsträger auch in der Lage sein, diese „Black-Box“-Algorithmen zu verstehen, um Vertrauen zu entwickeln und anschließend die Systeme zur Entscheidungsfindung einzusetzen (Martens und Provost 2014).

In der Literatur sind Anhaltspunkte für Erklärungen bekannt, mit denen die Akzeptanz von Algorithmen bzw. von datengetriebenen Ergebnissen gesteigert werden kann. Z. B. wurden in der Literatur zu Entscheidungsunterstützungssystemen (engl. Decision Support Systems [DSS]) Erklärungen schon frühzeitig als ein notwendiges Mittel identifiziert, um die Akzeptanz der DSS-Anwender zu fördern (Lilien et al. 2004). Auch die Forschung im Bereich der wissensbasierten Systeme (engl. Knowledge Based Systems [KBS]) hat gezeigt, dass Erklärungen, welche zusätzlich zu den Empfehlungen des Systems gegeben werden, dazu beitragen können, Vorbehalte zu überwinden, die Nutzer gegenüber der Zuverlässigkeit eines Systems und seiner Empfehlungen haben (Dhaliwal und Benbasat 1996; Gregor 1999). Indem die Entscheidungsprozesse eines Systems erläutert werden, wird auch seine wahrgenommene Zuverlässigkeit erhöht. Dies wiederum gibt dem Nutzer die Möglichkeit, die Verarbeitung der Daten zu verstehen, was zu einer erhöhten Akzeptanz (Ye und Johnson 1995) und einem höheren Vertrauen führen kann (Moa und Benbasat 2000). Darüber hinaus können Erklärungen auch zu einer effektiveren Beeinflussung der Überzeugungen des Benutzers durch die Empfehlungen von KBS führen (Ye und Johnson 1995). In späteren Untersuchungen über Empfehlungssysteme zeigten die Autoren Wang und Benbasat (2007), dass verschiedene Arten von Erklärungen das Vertrauen in die Empfehlung eines Systems erhöhen können.

Diese Anleitung fokussiert die Wirkungsweise von Erklärungsmethoden zur Steigerung der Akzeptanz BA-Systemen und datenbasierten Erkenntnissen. Praktiker können diese Anleitung nutzen, um die Akzeptanz der Nutzer solcher Analysen im eigenen Unternehmen zu steigern.

Wirkungsweise von Erklärungsmethoden

Die Erklärungsmethoden werden ferner hinsichtlich ihrer Auswirkungen auf das Verhalten und die Wahrnehmung des Systems durch den Benutzer analysiert und in einem morphologischen Kasten zusammengefasst. Erklärungsmethoden unterscheiden sich dabei bezüglich ihres positiven Einfluss auf die die Akzeptanz der Ergebnisse des Systems und auf die Wahrnehmung des Systems, die Verständlichkeit des Systems, das Vertrauen in das System, die Problemlösungsfähigkeit die durch das System gegeben wird, und die wahrgenommene Qualität der Entscheidung des Benutzers sowie die Zufriedenheit des Nutzers mit der Entscheidung.

Erklärungsmethode	Erklärungsmethode hat einen positiven Einfluss auf...				
	... die Akzeptanz der Ergebnisse des Systems und auf die Wahrnehmung des Systems	... die Verständlichkeit des Systems	... das Vertrauen in das System	... die Problemlösungsfähigkeit	... die wahrgenommene Qualität der Entscheidung sowie Zufriedenheit mit der Entscheidung
Justification	x	x	x	x	x
Rule Trace	x	x	x	x	
Strategic	x	x			
Tradeoff	x	x	x		x
Neighbor Style	x	x	x		
Definition	x			x	
Confidence		x		x	
Social	x		x		x
Feature	x		x	x	x
Tag		x		x	x
Influence		x	x	x	x

Erklärungsmethode	Erklärungsmethode hat einen positiven Einfluss auf...				
	... die Akzeptanz der Ergebnisse des Systems und auf die Wahrnehmung des Systems	... die Verständlichkeit des Systems	... das Vertrauen in das System	... die Problemlösungsfähigkeit	... die wahrgenommene Qualität der Entscheidung sowie Zufriedenheit mit der Entscheidung
Algorithm		x	x		
Procedural				x	
Directive				x	
Why not		x			

Katalog an Erklärungsmethoden

Im Folgenden wird ein Auszug der ermittelten Erklärungsmethoden genauer beschrieben. In Abstimmung mit den Industrievertretern des Forschungsprojekts wurden fünf Erklärungsmethoden als besonders interessant befunden und nachfolgend im Detail beschrieben: **Justification**, **Strategic**, **Rule Trace**, **Influence** sowie **Algorithm**. Dementsprechend werden nachfolgend ausschließlich die ausgewählten Erklärungsmethoden vorgestellt und beispielhaft dargestellt. Auch aus der Literatur entnommene Beispiele für die Umsetzung der einzelnen Erklärungsmethoden sind dargestellt.

Unter allen analysierten Aufsätzen ist **Justification** die am häufigsten verwendete Erklärungsmethode. Justification Erklärungen wurden durch Swartout's (1983) Expertensystem "XPLAIN" eingeführt und rechtfertigen die "zugrundeliegenden Gründe für einen Zustand oder eine Handlung" eines Systems (Dhaliwal & Benbasat, 1996). Sie werden als Feedbackerkklärungen verwendet und rechtfertigen die Bedeutung des Outputs des Systems (Dhaliwal & Benbasat, 1996), z.B. indem sie ihn mit dem Domänenwissen verknüpfen, welches im Entscheidungsprozess des Systems verwendet wurde (Chandrasekaran, Tanner, & Josephson, 1989). Darüber hinaus werden sie verwendet, um mögliche Konsequenzen des ermittelten Ergebnisses zu erklären (Dhaliwal & Benbasat, 1996; Mao & Benbasat, 2001). Diese Erläuterungen

können auch verwendet werden, um die Eingabeaufforderung der Systeme gegenüber dem Benutzer zu begründen, indem die Bedeutung der angeforderten Daten für die folgenden Auswertungen oder Prozesse aufgezeigt werden (Dhaliwal & Benbasat, 1996). Eines der Hauptmerkmale dieser Erklärungsmethode besteht darin, dass sie auf der Grundlage vorhandener Informationen in der Wissensbasis zusammengestellt werden kann, manchmal sogar, bevor das System irgendeine Prozedur durchgeführt hat (Chandrasekaran et al., 1989). Vereinfacht ausgedrückt, liefern Justification Erklärungen eine Antwort auf die Frage "Warum [...]?" (Gönül, Önkal, & Lawrence, 2006). Alle analysierten Rechtfertigungserklärungen werden in einer natürlichen Sprache präsentiert und sind textbasiert. Sie werden entweder automatisch bereitgestellt, indem sie unmittelbar neben dem Ergebnis angezeigt werden (Gönül et al., 2006; Li & Gregor, 2011), oder auf Wunsch des Benutzers. Dieser Benutzeraufruf erfolgt über Schaltflächen, zum Beispiel mit der Beschriftung "Warum" (Mao & Benbasat, 2001) oder in Form von Hyperlinks (Arnold, Clark, Collier, Leech, & Sutton, 2004). Dies kann bspw. auf der Nutzeroberfläche des Ausgabebereichs einer Analytics Anwendung realisiert werden.

Erklärungsmethode	Erklärungstext
Justification	<p>Der durchschnittliche Zufriedenheitswert von Mitarbeitern, welche das Unternehmen verlassen haben, war niedriger als von Mitarbeitern, welche im Unternehmen geblieben sind.</p> <p>Mitarbeiter, welche das Unternehmen verlassen haben, waren durchschnittlich an etwas mehr Projekten beteiligt, als Mitarbeiter, welche das Unternehmen nicht verlassen haben.</p> <p>Mitarbeiter, welche schon länger im Unternehmen sind, haben eine höhere Wahrscheinlichkeit das Unternehmen zu verlassen, als Mitarbeiter, welche noch nicht so lang im Unternehmen sind.</p> <p>Mitarbeiter, welche das Unternehmen verlassen haben, arbeiteten im Monat durchschnittlich mehr als Mitarbeiter, welche im Unternehmen geblieben sind.</p> <p>Bei Mitarbeitern, welche das Unternehmen verlassen haben, war durchschnittlich etwas mehr Zeit seit der letzten Evaluation vergangen als bei Mitarbeitern, welche im Unternehmen geblieben sind.</p> <p>Sollte dieser Mitarbeiter das Unternehmen verlassen, fallen zusätzliche Kosten an, welche für die Neufindung und die Einarbeitung des Ersatzes aufgebracht werden müssen.</p> <p>Außerdem würde die Produktivität in der Abteilung aufgrund von Unterbesetzung sinken und die Abteilungsziele könnten nicht erreicht werden.</p>

Rule Trace ist die am zweithäufigsten verwendete Erklärungsmethode. Sie wurde wie Justification durch das XPLAIN-System von (Swartout, 1983) eingeführt. Rule Trace oder Line of Reasoning-Erklärungen liefern eine Beschreibung, wie das System sein Ergebnis erzielt hat, indem sie aufzeigen, auf welchen Daten das Ergebnis beruht oder welche Regeln verwendet oder befolgt wurden (Chandrasekaran et al., 1989). Die Eingabewerte können im Rahmen der Erklärung bspw. in Bezug zu den Perzentilen des gesamten Datensatzes für den bestimmten Parameter gesetzt werden und somit eine Indikation liefern, wie das System entschieden hat. Dies gibt wiederum Aufschluss über die Bewertungen und Entscheidungen, die das System getroffen hat, um das Ergebnis zu erreichen (Mao & Benbasat, 2001). Als Feedforward-Erklärungen können

sie angeben, wie der Input vom System verwendet wird und welche Regeln angewendet werden (Dhaliwal & Benbasat, 1996). Da Rule Trace-Erklärungen hauptsächlich auf dem Vorgehen des Systems basieren, können sie nicht immer vorab kompiliert werden (Chandrasekaran et al., 1989). In anderen Worten: Rule Trace-Erklärungen beantworten die Frage "Wie [...]?" (Mao & Benbasat, 2001). In den meisten Studien wurden Rule Trace-Erklärungen in Textform und in natürlicher Sprache formuliert. Nur wenige Wissenschaftler haben eine multimediale Form der Erklärungsmethode untersucht. Ein System zur Analyse und dem Monitoring der Cyber Security Vorfällen von Ehrlich, Kirk, Patterson, Rasmussen, Ross, and Gruen (2011) zum Beispiel nutzte Grafiken, mit deren Hilfe das System einen Sicherheitsvorfall kommunizieren konnte. Das Empfehlungssystem von Tan, Tan, and Teo (2012) präsentierte neben dem empfohlenen Element auch Sternbewertungen, die angeben, wie gut die Attribute des Elements den Präferenzen des Benutzers entsprechen. Rule Trace-Erklärungen werden entweder automatisch neben den Ergebnissen des Systems dargestellt (Bussone, Stumpf, & O'Sullivan, 2015) oder auf Wunsch des Benutzers angezeigt. Der Aufruf durch den Benutzer kann z.B. durch einen Klick auf einen Button mit der Aufschrift "Wie" (Wang, Qiu, Kim, & Benbasat, 2016), durch Hyperlinks (Arnold, Clark, Collier, Leech, & Sutton, 2006) oder durch die Auswahl eines bestimmten Klassifikationsergebnisses (Ehrlich et al., 2011) erfolgen.

Erklärungsmethode	Erklärungstext
Rule Trace	<p>Das System basiert sein Vorhersage-Modell auf Daten von 14.999 Mitarbeitern, mit jeweils 9 Attributen, von denen 23,8% das Unternehmen verlassen haben.</p> <p>Der Zufriedenheitswert dieses Mitarbeiters liegt über dem Durchschnitt von Mitarbeitern, die das Unternehmen nicht verlassen haben (0,66).</p> <p>Die Anzahl der Projekte dieses Mitarbeiters ist höher, als die durchschnittliche Anzahl von Projekten von Mitarbeitern, die das Unternehmen nicht verlassen haben (3,78).</p> <p>Dieser Mitarbeiter ist länger im Unternehmen, als ein durchschnittlicher Mitarbeiter, der das Unternehmen nicht verlassen hat (3,38 Jahre).</p> <p>Dieser Mitarbeiter hat durchschnittlich mehr gearbeitet, als ein durchschnittlicher Mitarbeiter, der das Unternehmen nicht verlassen hat (199,06 Stunden).</p> <p>Seit der letzten Evaluation dieses Mitarbeiters ist mehr Zeit vergangen als bei einem durchschnittlichen Mitarbeiter, der das Unternehmen nicht verlassen hat (0,71 Jahre).</p>

Strategic Erklärungen stellen die zugrundeliegende Strategie dar, indem sie den Problemlösungsansatz des Systems beschreiben (Chandrasekaran et al., 1989; Dhaliwal & Benbasat, 1996) und zudem beschreiben, wie dieses spezielle Ergebnis oder der angeforderte Input, wenn er als Feedforward-Erklärung verwendet wird, in das Gesamtziel des Systems passt (Dhaliwal & Benbasat, 1996). Strategische Erklärungen wurden erstmals mit dem Expertensystem "NEOMYCIN" von Clancey (1983) eingeführt. Dabei entschieden sich fast alle Wissenschaftler für eine textuelle Darstellungsform, d. h. die Erklärungen wurden in natürlicher Sprache verfasst. Ausgenommen sind hier die Autoren (Tan et al., 2012). Die Autoren verwenden eine Sternenbewertung, um die mögliche Übereinstimmung des empfohlenen Elements anzuzeigen. Sie begründen diese Art der Darstellung damit, dass das Hauptziel des Systems darin besteht, das am besten übereinstimmende Element für den Benutzer zu empfehlen und deshalb das System die voraussichtliche Übereinstimmung anzeigt. Kaptein, Broekens, Hindriks, and Neerincx (2017) haben einem humanoiden Roboter so programmiert, dass dieser sich verbal erklären kann, um Kindern ohne

Lesefähigkeit die Möglichkeit zu geben seinen Ausführungen zu folgen. Das Empfehlungssystem von Tan et al. (2012) und der Roboter von Kaptein et al. (2017) waren die einzigen Fälle, in denen die strategischen Erklärungen automatisch neben der Empfehlung präsentiert wurden. In allen anderen Systemen werden die strategischen Erklärungen erst dann präsentiert, wenn ein Benutzer sie angefordert hat, z.B. durch Anklicken eines Hyperlinks mit dem Text "Strategie des Problemlösungsansatz" (Arnold et al., 2004).

Erklärungsmethode	Erklärungstext
Strategic	<p>Das BI&A-System nutzt historische Daten des Unternehmens, um eine Vorhersage über eine mögliche Kündigung durch den Mitarbeiter zu treffen.</p> <p>Das Ziel ist es zuverlässig vorherzusagen, ob der Mitarbeiter das Unternehmen verlassen wird. Das gibt dem Unternehmen die Möglichkeit, frühzeitig zu reagieren, einer möglichen Kündigung entgegenzuwirken und so die Mitarbeiterfluktuation zu verringern.</p> <p>Dieser Mitarbeiter könnte zum Beispiel an einem Mitarbeiterbindungs-Programm teilnehmen und so von einem Verbleib überzeugt werden. Dies würde verhindern, dass zusätzliche Kosten entstehen welche z.B. für die Suche, Einstellung und Einarbeitung eines Ersatzes anfallen würden.</p>

Influence Erklärungen zeigen den Einfluss von Attributen, Modellmerkmalen oder bereits bewerteten Elementen auf das aktuelle Ergebnis oder die Empfehlung an (Bilgic & Mooney, 2005). Alle gefundenen Artikel zu dieser Erklärungsmethode lieferten die Erklärung automatisch zusammen mit dem Ergebnis des Systems. Bilgic and Mooney (2005a) verwendeten sie, um ihre Empfehlung für Bücher zu erläutern. Neben der Empfehlung zeigten sie eine Tabelle, in der sie frühere Bücher aufführten, die vom Benutzer bewertet wurden und die die Empfehlung des Systems verursachten. Neben dem Namen der Bücher zeigten sie die Bewertung des Benutzers und den Einfluss dieser Bewertung auf das aktuelle Ergebnis durch eine Zahl zwischen 0 und 100 an, wobei 100 den höchsten Einfluss darstellt. In der Studie von Lai and Tan (2019) hatten die Teilnehmer die Aufgabe, zu beurteilen, ob eine

Hotelbewertung echt oder irreführend war. Zusätzlich zur Klassifizierung hebt das System zehn Wörter in der Bewertung mit blauer Farbe hervor, um zu veranschaulichen, welche Wörter den größten Einfluss auf das Machine Learning Modell hatten. Die Farbstärke der Hervorhebung wurde nach und nach entsprechend einer Heatmap verteilt, die über der Rezension angezeigt wurde. Je dunkler das Blau, desto größer die Bedeutung für das Modell. Darüber hinaus erklärt ein kurzer Text oberhalb der Heatmap die Bedeutung der Hervorhebung. In einer neueren Studie von Ribeiro, Singh, and Guestrin (2016) hat ein Machine Learning Modell die Aufgabe, das Tier in einem Bild als Hund oder Wolf zu klassifizieren. Durch das Hervorheben bestimmter Bereiche des betrachteten Bildes als Heatmap, veranschaulicht das System, welche Bereiche bei der Klassifizierung am wichtigsten waren.

Erklärungsmethode	Erklärungstext
Influence	<p>Die folgenden Mitarbeiter-Attribute haben den größten prozentualen Einfluss auf das Ergebnis des Modells:</p> <p>Das Zufriedenheitslevel des Mitarbeiters beeinflusst das Ergebnis mit 30,33 Prozent am stärksten.</p> <p>Die Anzahl an Projekten, an denen der Mitarbeiter beteiligt ist, hat den zweitgrößten Einfluss mit 20,98 Prozent.</p> <p>Mit 17,42 Prozent hat die Anzahl an Jahren, die der Mitarbeiter bereits im Unternehmen arbeitet, den drittgrößten Einfluss auf das Ergebnis.</p> <p>Das Ergebnis wird von der durchschnittlichen monatlichen Arbeitszeit des Mitarbeiters zu 14,43 Prozent beeinflusst.</p> <p>Der Zeitpunkt der letzten Evaluation des Mitarbeiters beeinflusst 12,59 Prozent des Ergebnisses.</p> <p>Alle weiteren Attribute haben weniger als 2 Prozent Einfluss auf das Ergebnis.</p>

Die **Algorithm** Erklärungen liefern eine vereinfachte Erklärung des Algorithmus oder des Modells, welches dem System zugrunde liegt (Kulesza, Stumpf, Burnett, Yang, Kwan, and Wong 2013) (Nur ein Paper in dieser Analyse verwendete diese Methode. Kulesza et al. (2013) schufen verschiedene Ebenen von vereinfachten Erklärungen

ihres Musikempfehlungsalgorithmus, um die Klangqualität ihrer Erklärungsschnittstelle zu manipulieren. Die detaillierteste Erklärung beschrieb den tatsächlichen Entscheidungsbaum, welches der Algorithmus des Systems verwendet hat. Die Erklärung mit mittlerem Detaillierungsgrad beschrieb einen einfachen Entscheidungsbaum und die Erklärung mit einem niedrigen Detaillierungsgrad basierte auf einem Entscheidungsbaum mit nur einem Merkmal. Alle Erklärungen wurden zusammen mit der Empfehlung und in Textform gegeben.

Erklärungsmethode	Erklärungstext
Algorithm	<p>Das BI&A-System basiert seine Vorhersage auf einem Random-Forest Machine Learning Modell.</p> <p>Das Random-Forest Modell besteht aus einer Vielzahl von einzelnen Entscheidungsbäumen (in diesem Fall 10), die als Ensemble funktionieren.</p> <p>Ein Entscheidungsbaum ist eine Flussdiagramm-ähnliche Baumstruktur, bei der jeder interne Knoten einen Test auf ein Attribut (z.B. Zufriedenheitslevel >0.5) abbildet. Desweiteren stellt jeder Zweig ein Ergebnis des Tests dar (z.B. Wahr – Test trifft zu) und jeder Zweig endet in einem weiteren Test oder in einer Klassenvorhersage (Mitarbeiter verlässt das Unternehmen oder Mitarbeiter bleibt).</p> <p>Bei jedem Test wird der Datensatz in immer kleinere Teilmengen zerlegt.</p> <p>Die Entscheidungsbäume werden vom System während des Trainings des Modells auf Basis von zufälligen Mitarbeitern und Attributen erstellt. Für das Training wurden 10.500 historische Datensätze von Mitarbeitern verwendet. Mit weiteren 4.499 Datensätzen wurde das Modell anschließend getestet.</p> <p>Während der Klassifikation gibt jeder einzelne Baum im „Random Forest“ eine Klassenvorhersage für den Mitarbeiter ab. Die Klasse mit den meisten Stimmen wird abschließen zu der Vorhersage dieses Modells benutzt.</p>

Vorgehen zur Steigerung der Akzeptanz an einem Beispiel

Das Vorgehen um die Akzeptanz eine Business Analytics Anwendung mit Erklärungen zu steigern, lässt sich in drei Schritte gliedern: (1) **Analyse der Business Analytics Anwendung**, (2) **Auswahl einer passenden Erklärung**, (3) **Umsetzung der Erklärung in der Business Analytics Anwendung**.

Im Folgenden werden diese 3 Schritte dargestellt und anhand eines Beispiels veranschaulicht.

1. Analyse der Business Analytics Anwendung

Zunächst gilt es die Situation zu beschreiben und zu erkennen wer Nutzer der BA-Anwendung ist. Insbesondere die Kenntnisse und Charakteristika der Nutzergruppe haben einen Einfluss auf die Wahl der Erklärungsmethode.

Beispiel für eine Situation:

Sie sind Human Resources Manager (HR-Manager) in einem großen Unternehmen. Ihre heutige Aufgabe ist es, zu entscheiden bei welchen Mitarbeitern das Risiko besteht, dass sie das Unternehmen verlassen.

Die Business-Intelligence & Analytics (BI&A) Abteilung Ihres Unternehmens hat Ihnen zur Unterstützung für diese Aufgabe ein **BI&A-System** zur Verfügung gestellt, welches anhand von verschiedenen Mitarbeiterattributen und mit Hilfe eines **Machine Learning Modells** vorhersagt, ob ein bestimmter Mitarbeiter das Unternehmen verlassen wird oder nicht.

2. Auswahl einer passenden Erklärung

Es stehen nun verschiedene Erklärungsmethoden mit verschiedenen Wirkungsweisen zur Verfügung. Im Folgenden ist eine Auswahl verbreiteter Erklärungsmethoden mit unterschiedlichen Wirkungsweisen aufgeführt und jeweils mit einem Beispiel versehen. In diesem Fall hat die Nutzergruppe eine mäßige IT-Expertise. Dementsprechend ist eine technische Erklärung nicht zielführend. Eine strategische Erklärung bietet zwar einen guten Überblick, jedoch handelt es sich um einen sensiblen Bereich. Der Nutzer der BA-Anwendung hat Zeit sich mit der Anwendung genau auseinanderzusetzen und will verstehen, wie das System funktioniert. Deshalb bieten sich in diesem Fall die Erklärungsmethoden Rule Trace oder Justification an.

3. Umsetzung der Erklärung in der Business Analytics Anwendung

Die Umsetzung der Erklärungsmethode kann analog zu den Beispielen im Katalog erfolgen. Hierbei ist zu beachten, dass die Benutzeroberfläche nicht durch zu viel Text unübersichtlich wird. Hier haben sich optionale Einblendungen bewährt. Die Kombination verschiedener Erklärungsmethoden kann zielführend sein, insbesondere wenn eine heterogene Benutzergruppe vorliegt.

Literaturverzeichnis

- Arnold, Clark, Collier, Leech, & Sutton. 2006. The Differential Use and Effect of Knowledge-Based System Explanations in Novice and Expert Judgment Decisions. *MIS Quarterly*, 30(1): 79.
- Arnold, V., Clark, N., Collier, P. A., Leech, S. A., & Sutton, S. G. 2004. Explanation provision and use in an intelligent decision aid. *Intelligent Systems in Accounting, Finance & Management*, 12(1): 5–27.
- Bilgic, M., & Mooney, R. J. 2005. Explaining Recommendations: Satisfaction vs. Promotion. *Beyond Personalization Workshop*, (IUI): 1–6.
- Breiman, L. 2001. Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). In: *Statist. Sci.* 16 (3), S. 199–231.
- Bussone, A., Stumpf, S., & O'Sullivan, D. 2015. The Role of Explanations on Trust and Reliance in Clinical Decision Support Systems. In *International Conference on Healthcare 21.10.2015 - 23.10.2015*: 160–169.
- Chandrasekaran, B., Tanner, M. C., & Josephson, J. R. 1989. Explaining control strategies in problem solving. *IEEE Expert*, 4(1): 9–15.
- Dhaliwal, J. S., & Benbasat, I. 1996. The Use and Effects of Knowledge-Based System Explanations: Theoretical Foundations and a Framework for Empirical Evaluation. *Information Systems Research*, 7(3): 342–362.
- Ehrlich, K., Kirk, S. E., Patterson, J., Rasmussen, J. C., Ross, S. I., & Gruen, D. M. 2011. Taking advice from intelligent systems. the 15th international conference. 13.02.2011 - 16.02.2011: 125. Palo Alto, CA, USA: ACM Press.

- Gönül, M. S., Önkal, D., & Lawrence, M. 2006. The effects of structural characteristics of explanations on use of a DSS. *Decision Support Systems*, 42(3): 1481–1493.
- Gregor, S. 1999. Explanations from Intelligent Systems: Theoretical Foundations and Implications for Practice. In: *MIS Quarterly* 23 (4), S. 497.
- Kaptein, F., Broekens, J., Hindriks, K., & Neerincx, M. 2017 - 2017. Personalised self-explanation by robots: The role of goals versus beliefs in robot-action explanation for children and adults. 2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN). 28.08.2017 - 01.09.2017: 676–682. Lisbon: IEEE.
- Kulesza, T., Stumpf, S., Burnett, M., Yang, S., Kwan, I., & Wong, W.-K. 2013 - 2013. Too much, too little, or just right? Ways explanations impact end users' mental models. 2013 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC). 15.09.2013 - 19.09.2013: 3–10. San Jose, CA, USA: IEEE.
- Lai, V., & Tan, C. 2019. On Human Predictions with Explanations and Predictions of Machine Learning Models. the Conference. 29.01.2019 - 31.01.2019: 29–38. Atlanta, GA, USA: ACM Press.
- Li, M., & Gregor, S. 2011. Outcomes of effective explanations: Empowering citizens through online advice. *Decision Support Systems*, 52(1): 119–132.
- Lilien, G. L., Rangaswamy, A., van Bruggen, G. H., Starke, K. 2004. DSS Effectiveness in Marketing Resource Allocation Decisions: Reality vs. Perception. In: *Information Systems Research* 15 (3), S. 216–235.
- Mao, J., & Benbasat, I. 2001. The effects of contextualized access to knowledge on judgement. *International Journal of Human-Computer Studies*, 55(5): 787–814.
- Martens, D., Provost, F. 2014. Explaining Data-Driven Document Classifications. In: *MIS Quarterly* 38 (1), S. 73–99.
- Ribeiro, M. T., Singh, S., & Guestrin, C. 08132016. "Why Should I Trust You?". KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 13 08 2016 17 08 2016: 1135–1144. San Francisco California USA: ACM.
- Shollo, A., Galliers, R. D. 2016. Towards an understanding of the role of business intelligence systems in organisational knowing. In: *Info Systems J* 26 (4), S. 339–367.

- Swartout, W. R. 1983. XPLAIN: a system for creating and explaining expert consulting programs. *Artificial Intelligence*, 21(3): 285–325.
- Wang, W., Benbasat, I. 2007. Recommendation Agents for Electronic Commerce: Effects of Explanation Facilities on Trusting Beliefs. In: *Journal of Management Information Systems* 23 (4), S. 217–246.
- Ye, L. R., Johnson, P. E. 1995. The Impact of Explanation Facilities on User Acceptance of Expert Systems Advice. In: *MIS Quarterly* 19 (2), S. 157.