

Analyse eines modularen Ansatzes zur automatisierten Erstellung eines Prognosemodells unter Verwendung der SAP HANA-Technologie

Kenan Atas

Technische Hochschule
Mittelhessen

Fachbereich Mathematik,
Naturwissenschaften und
Datenverarbeitung
Wilhelm-Leuschner-Str. 13
61169 Friedberg (Hessen)
E-Mail:
kenan.atas@mnd.thm.de

Prof. Dr. Harald Ritz

Technische Hochschule
Mittelhessen

Fachbereich Mathematik,
Naturwissenschaften und
Informatik
Wiesenstraße 14
35390 Gießen
E-Mail:
harald.ritz@mni.thm.de

Prof. Dr. Frank Kammer

Technische Hochschule
Mittelhessen

Fachbereich Mathematik,
Naturwissenschaften und
Informatik
Wiesenstraße 14
35390 Gießen
E-Mail:
frank.kammer@mni.thm.de

Kategorie

Bachelorarbeit

Schlüsselwörter

Predictive Analytics, Business Intelligence, CRISP-DM, SAP HANA, SAP Predictive Analytics, R

Zusammenfassung

Tag für Tag werden durch das Anrufen mit dem Handy, dem Versenden von Nachrichten, dem Bestellen von Onlinewaren oder anderen digitalen Tätigkeiten digitale Spuren hinterlassen. Jede dieser Tätigkeiten generiert personenbezogene Daten und diese Informationen sind für Unternehmen goldwert. Durch die enorme Ausbreitung solcher Daten verwenden Unternehmen Business Intelligence Lösungen. Auf diese Weise können sie aus ihren Daten Informationen gewinnen und ihre operativen und strategischen Entscheidungen besser vorbereiten. Jedoch ist es für Unternehmen inzwischen nicht mehr ausreichend zu wissen, was gegenwärtig passiert. Sie wollen antizipieren, was in der Zukunft passieren wird, wodurch ihre entscheidungsrelevanten Informationen gestärkt werden, um anschließend optimale Ergebnisse zu erzielen.

Um diese Aufgabe zu überwinden, kommt Predictive Analytics zum Einsatz, welches als Teilbereich von Business Intelligence angesehen wird. Predictive Analytics unterstützt Entscheider nicht ausschließlich in der Analyse der vergangenen Daten, sondern gibt auch zukunftsorientierte Unterstützung. Das Ziel ist es, durch eine effiziente Nutzung der angefallenen Daten nützliche Informationen zu gewinnen und neues Wissen zu erhalten.

Um Informationen aus den Daten zu gewinnen, muss an erster Stelle eine leistungsstarke Datenbank implementiert

beziehungsweise angebunden werden, welche zudem einen schnellen Zugriff auf die Daten ermöglicht und somit die Abfragezeiten verkürzt.

Das Thema der Bachelorarbeit ist die exemplarische Umsetzung eines modular aufgebauten Prognosemodells mithilfe der SAP HANA-Technologie unter Einbezug der Anwendung SAP Predictive Analytics in Kombination mit R. Die Untersuchung erfolgt anhand von Daten, die von einer Open-Source-Datenbank mit Informationen zu Fußballspielern und deren Bewertungen im Jahre 2019/2020 zur Verfügung gestellt worden sind, welche auf realen Daten und Mannschaften beruhen. Um eine theoretische Basis zu verschaffen, wird zunächst ein detaillierter Überblick über die eingesetzten Technologien gegeben. Diesem theoretisch ausgerichteten Teil der Arbeit folgt schließlich die Entwicklung des Prognosemodells. Das Vorgehen der Arbeit ist angelehnt an das ‚Cross-Industry Standard Process Model for Data Mining‘. Mithilfe der eingesetzten Technologien wird innerhalb dieses Vorgehensmodells der Business Case definiert, die Daten analysiert und vorbereitet, das Modell gebaut und die Ergebnisse evaluiert. Schlussendlich kann gezeigt werden, dass die Möglichkeit besteht, mit den eingesetzten Technologien ein Prognosemodell zu bauen.