

Jens Zeidler

Personalisierung im Web

*Algorithmen, Konzeption, Entwicklung und Bewertung
für einen Online-Shop*

Bibliografische Information der Deutschen Nationalbibliothek:

Bibliografische Information der Deutschen Nationalbibliothek: Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de/> abrufbar.

Dieses Werk sowie alle darin enthaltenen einzelnen Beiträge und Abbildungen sind urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsschutz zugelassen ist, bedarf der vorherigen Zustimmung des Verlanges. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen, Auswertungen durch Datenbanken und für die Einspeicherung und Verarbeitung in elektronische Systeme. Alle Rechte, auch die des auszugsweisen Nachdrucks, der fotomechanischen Wiedergabe (einschließlich Mikrokopie) sowie der Auswertung durch Datenbanken oder ähnliche Einrichtungen, vorbehalten.

Copyright © 2009 Diplomica Verlag GmbH
ISBN: 9783842812796

Jens Zeidler

Personalisierung im Web

Algorithmen, Konzeption, Entwicklung und Bewertung für einen Online-Shop

Jens Zeidler

Personalisierung im Web

*Algorithmen, Konzeption, Entwicklung und Bewertung
für einen Online-Shop*

Jens Zeidler

Personalisierung im Web

Algorithmen, Konzeption, Entwicklung und Bewertung für einen Online-Shop

ISBN: 978-3-8428-1279-6

Herstellung: Diplomica® Verlag GmbH, Hamburg, 2011

Zugl. Fachhochschule Erfurt, Erfurt, Deutschland, MA-Thesis / Master, 2009

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes der Bundesrepublik Deutschland in der jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechtes.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Die Informationen in diesem Werk wurden mit Sorgfalt erarbeitet. Dennoch können Fehler nicht vollständig ausgeschlossen werden und der Verlag, die Autoren oder Übersetzer übernehmen keine juristische Verantwortung oder irgendeine Haftung für evtl. verbliebene fehlerhafte Angaben und deren Folgen.

© Diplomica Verlag GmbH

<http://www.diplomica.de>, Hamburg 2011

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	1
1.1.1	Vorstellung des Unternehmens	3
1.1.2	Ziel der Softwarelösung	3
1.2	Ziel der Arbeit	4
2	Personalisierung im Web	5
2.1	Definition Personalisierung	5
2.2	Ebenen der Personalisierung	5
2.3	Personalisierungsprozess	6
2.3.1	Datenerfassung	7
2.3.2	Profilerstellung	8
2.3.3	Personalisierungsverfahren	9
2.4	Beispiele und Einsatzgebiete	10
2.4.1	Personalisierte E-Mail Kommunikation	10
2.4.2	Personalisierte Werbebanner / Seiteninhalte	11
2.4.3	Personalisierte Suchmaschinen	12
2.4.4	Personalisierte Produktempfehlungen	13
3	Anforderungsanalyse	15
3.1	Inhalt	15
3.2	Allgemeine Anforderungen	15
3.3	Grobspezifikation	15
3.3.1	Architekturentwurf	16
3.3.2	Darstellung der Use-Cases	17
3.4	Feinspezifikation	19
3.4.1	Interface	19
3.4.2	Datenbankschema	20
3.4.3	Schnittstelle zwischen beiden Projekten	21
3.4.4	Kommunikation zwischen Shop und Empfehlungssystem	22
4	Umsetzung	24
4.1	Auswahl Empfehlungsalgorithmen	24
4.2	Vorgehensweise Implementation	25
4.3	Softwaretests	27

4.3.1	Funktionsorientierte Tests	27
4.3.2	Kontrollflußorientierte Tests	28
4.3.3	Datenflußorientierte Tests	28
4.3.4	Modellbasierte Tests	28
4.3.5	Prüfstrategien	29
4.3.6	Umsetzung der Tests	30
5	Algorithmen	31
5.1	Grundlagen	31
5.1.1	Collective Intelligence	31
5.1.2	Funktionsweise von Empfehlungssystemen	31
5.1.3	Distanzberechnung in Empfehlungssystemen	32
5.1.4	Algorithmenklassen bei Produktempfehlungen	34
5.1.5	Speicherbasierte und modellbasierte Techniken	36
5.1.6	Naive-Bayes Modelle	38
5.1.7	Clusterbasierte Modelle	38
5.1.8	Herausforderungen	39
5.2	Überblick der Empfehlungsalgorithmen	40
5.2.1	Attributbasierte Empfehlung	40
5.2.2	Warenkorbanalyse	40
5.2.3	Apriori Algorithmus	41
5.2.4	Bewertungsbasierte Verfahren	44
5.2.4.1	Vergleich der Varianten	44
5.2.4.2	MovieLens - Filmempfehlungen mit Nutzerratings	44
5.2.5	Empfehlungen mit KNN Verfahren	45
5.3	Bewertungsmöglichkeiten	47
5.3.1	Prognosegüte	47
5.3.2	Receiver-Operating-Characteristics (ROC) Analyse	47
5.3.3	Mittlerer absoluter Fehler	50
5.3.4	Grenzen der Bewertungsverfahren	50
6	Evaluierung	51
6.1	Business understanding - Über das Unternehmen	51
6.2	Data understanding - Datenverständnis	52
6.2.1	Statistik	52
6.2.2	Data preparation - Datenaufbereitung	53
6.2.2.1	Import der Bestelldaten	54
6.2.2.2	Import der Ratings	54
6.2.2.3	Sicherung der Datenstände	55
6.2.2.4	Evaluierungsdaten	55
6.3	Evaluierung der Algorithmen	56
6.3.1	Evaluierungskriterien	56
6.3.2	Einzelbewertung der Verfahren	58

PERSONALISIERUNG IM WEB -

Algorithmen, Konzeption, Entwicklung und Bewertung für einen Online-Shop.

Inhaltsverzeichnis

6.3.2.1	Warenkorbanalyse	58
6.3.2.2	Apriori-Algorithmus	59
6.3.2.3	K-Nearest Neighbor (KNN)	62
6.3.3	Gesamtbewertung in ROC-Diagramm	64
7	Fazit und Ausblick	67
8	Glossar	69