

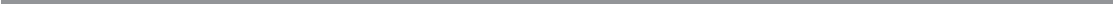
Felix Weber

Artificial Intelligence for Business Analytics

Algorithms, Platforms and Application
Scenarios



Springer



Artificial Intelligence for Business Analytics

Felix Weber

Artificial Intelligence for Business Analytics

Algorithms, Platforms and Application
Scenarios

Felix Weber
Chair of Business Informatics and Integrated Information Systems
University of Duisburg-Essen
Essen, Germany

ISBN 978-3-658-37598-0 ISBN 978-3-658-37599-7 (eBook)
<https://doi.org/10.1007/978-3-658-37599-7>

© Springer Fachmedien Wiesbaden GmbH, part of Springer Nature 2023

The translation was done with the help of artificial intelligence (machine translation by the service DeepL.com). A subsequent human revision was done primarily in terms of content.

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer Vieweg imprint is published by the registered company Springer Fachmedien Wiesbaden GmbH, part of Springer Nature.

The registered company address is: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Preface

With a purely academic background, it does surprise many newcomers what in practice is called business intelligence (BI) and business analytics (BA). The first naive thought of university graduates is certainly that really complex artificial intelligence (AI) and advanced machine learning (ML) models are applied in any larger companies. How else could one remain competitive if one does not already know before the consumer which color of the new car or cell phone model will be desired in the future and then play this out situation-dependently via all conceivable advertising channels?

I was very surprised to find that most of the “intelligence” done are just simple descriptive statistics.¹ One of Germany’s largest retailers only collects sales reports on a weekly basis and does not even aggregate them from the ERP system they use but rather have the individual local branches manually compile the numbers and enter them into their in-house information system. Another retail company is just starting with the most basic analyses of the store's business based on simple metrics, such as the promotional share of sales, that is, the number of advertised products relative to non-advertised ones, per receipt. This sudden interest is probably due to the emerging competition from online retailers, such as Amazon, which only entered the German grocery market in 2017 with Amazon Fresh.²

Much of the analysis in the business environment is descriptive analysis. They compute descriptive statistics (i.e., counts, totals, averages, percentages, minimums, maximums, and simple arithmetic) that summarize certain groupings or filtered versions of the data, which are typically simple counts of some events. These analyses are mostly based on standard aggregate functions in databases that require nothing more than elementary school math. Even basic statistics (e.g., standard deviations, variance, p-value, etc.) are

¹However, this observation holds true for many of the ideas, concepts, and recommended actions coming from the idealized world of research. The system architectures are more similar to monolithic mainframe architectures than to the state of the art of distributed service-oriented architectures – or the planning and execution of projects ignores the last decades of research in project management and instead uses, if at all, Microsoft Excel-based “planning.”

²To be honest, it must also be said that the success of German retailers in the past would not have necessitated a deeper examination of these issues.

quite rare. The purpose of descriptive analytics is to simply summarize and tell you what happened: sales, number of customers, percentage of total sales with items that were advertised, page views, etc. There are literally thousands of these metrics – it is pointless to list them – but they are all just simple event counts. Other descriptive analytics can be results of simple arithmetic operations, such as share of voice, average response time, percentage index, and the average number of responses per post. All of this takes place in a majority of companies today and is mostly referred to as business intelligence. Most often, the term advanced analytics is used to describe the extension of this reporting to include some filters on the data before the descriptive statistics are calculated. For example, if you apply a geo-filter first for social media analytics, you can get metrics like average post per week from Germany and average post per week from the Netherlands. And you can display that data on a fancy map for all the countries you are active in. Then all of a sudden you can call it advanced analytics.

However, this rudimentary analytics is not enough for a competitive advantage over competitors. Especially if you suddenly have to compete with digital natives like Google, Amazon, or Alibaba. In the age of digitalization, however, this is a real challenge for many industries. Amazon has turned book retailing, and then retail itself, upside down. Google is suddenly entering the automotive market with self-driving cars, Uber is demoting industry giants in the automotive industry (Volvo and Toyota) to mere suppliers, and Airbnb is taking over a large market share in the hotel industry without its real estate. As different as these examples are, they are based not only on software and platforms but, more importantly, also on sophisticated analytics. Uber has a huge database of drivers, so as soon as you request a car, Uber's algorithm is ready to go – in 15 seconds or less, it matches you with the driver closest to you. In the background, Uber stores data about every ride – even when the driver has no passengers. All of this data is stored and used to predict supply and demand and set fares. Uber also studies how transportation is handled between cities and tries to adjust for bottlenecks and other common problems.

Essen, Germany

Felix Weber

About the Aim of the Book

The aim of this book is not to train you as a data scientist or data analyst, nor will anyone be able to call themselves an expert in artificial intelligence or machine learning after reading it – even if some management consultants will do so. Instead, the book introduces the essential aspects of business analytics and the use of artificial intelligence methods in a condensed form. First of all, the basic terms and thought patterns of analytics from descriptive and predictive to prescriptive analytics are introduced in section “Categorisation of Analytical Methods and Models”. This is followed by the business analytics model for artificial intelligence (BAM.AI), a process model for the implementation of business analytics projects in section “Procedure Model: Business Analytics Model for Artificial intelligence (BAM.AI)”, and a technology framework, including the presentation of the most important frameworks, programming languages, and architectures, in Chap. 3. After an introduction to artificial intelligence in Chap. 2 and especially the subfield of machine learning, the most important problem categories are described, and the applicable algorithms are presented roughly but in an understandable way in section “Types of Problems in Artificial Intelligence and their Algorithms”. This is followed by a detailed overview of the common cloud platforms in section “Business Analytics and Machine Learning as a Service (Cloud Platforms)”, which enables quick implementation of a BA project. Here, the reader is provided with a guide that allows them to get an overview of the extensive offerings of the major providers. Finally, several application scenarios from different perspectives show the possible use of AI and BA in various industries as case studies section “Build or Buy?”.

Since the book definitely sees itself as an introduction and overview for decision-makers and implementers in IT and the related application domains, references to more in-depth literature are made in many places.

Essen, Germany

Felix Weber

Contents

1 Business Analytics and Intelligence	1
Need for Increasing Analytical Decision Support	1
Distinction Between Business Intelligence and Business Analytics	5
Categorisation of Analytical Methods and Models	7
Descriptive Analytics	7
Predictive Analytics	8
Prescriptive Analytics	10
Business Analytics Technology Framework (BA.TF)	11
Data Sources	13
Data Preparation	13
Data Storage	15
Analysis	16
Access and Use	17
(Big)-Data Management and Governance	18
Procedure Model: Business Analytics Model for Artificial Intelligence (BAM. AI)	19
Development Cycle	21
Business Understanding	21
Data Discovery	22
Data Wrangling	23
Analysis	24
Validation	25
New Data Acquisition	25
Deployment Cycle	25
Publish	26
Analytic Deployment	27
Application Integration	29
Test	29
Production/Operations	30
Continuous Improvement	30
References	30

2	Artificial Intelligence	33
	Machine Learning	35
	Supervised Learning	36
	Unsupervised Learning	37
	Reinforcement Learning	38
	Overview of the Types of Machine Learning	39
	Neural Networks	39
	Types of Problems in Artificial Intelligence and Their Algorithms	42
	Classification	42
	Dependencies and Associations	45
	Clustering	48
	Regression, Prediction, or Forecasting	50
	Optimization	53
	Detection of Anomalies (Outlier)	54
	Recommendation or Recommender Systems	56
	When to Use Which Algorithm?	59
	References	63
3	AI and BA Platforms	65
	Basic Concepts and Software Frameworks	65
	Data Management	65
	Data Warehouse	65
	Data Lake	66
	Data Streaming and Message Queuing	68
	Database Management System	70
	Apache Hadoop	71
	Data Analysis and Programming Languages	73
	Python	73
	R	74
	SQL	75
	Scala	75
	Julia	76
	AI Frameworks	77
	Tensorflow	77
	Theano	78
	Torch	79
	Scikit-Learn	79
	Jupyter Notebook	80
	Business Analytics and Machine Learning as a Service (Cloud Platforms)	81
	Amazon AWS	82
	Amazon AWS Data Services	83
	Amazon AWS ML Services	87

Google Cloud Platform	91
Data Services from Google	92
ML Services from Google	95
Google Prediction API and Cloud AutoML	96
Google Cloud Machine Learning Engine (Cloud Machine Learning Engine)	96
IBM Watson	96
Microsoft Azure	97
Data Services from Microsoft Azure	97
ML Services from Microsoft Azure	99
Overview of Other Microsoft Azure Services	100
SAP Business Technology Platform (SAP BTP)	100
Data Services from SAP	102
ML Services from SAP	105
SAP HANA Database Platform	109
Build or Buy?	110
References	112
4 Case Studies on the Use of AI-Based Business Analytics	113
Case Study: Analyzing Customer Sentiment in Real Time with Streaming Analytics	113
Customer Satisfaction in the Retail Sector	113
Technology Acceptance and Omnichannel for More Data	114
Customer Satisfaction Streaming Index (CSSI)	116
Implementation in a Retail Architecture	117
Results	119
Case Study: Market Segmentation and Automation in Retailing with Neural Networks	121
The Location Decision in Stationary Trade	122
Marketing Segmentation and Catchment Area	123
Classical Clustering Approaches and Growing Neural Gas	124
Project Structure	127
The Data and Sources	127
Implementation	130
Results	132
References	134