

Muhammed Fatih Kaya

# Automated Pattern Recognition of Communication Behaviour in Electronic Business Negotiations



Springer Gabler

---

# **Gabler Theses**

In der Schriftenreihe „Gabler Theses“ erscheinen ausgewählte, englischsprachige Doktorarbeiten, die an renommierten Hochschulen in Deutschland, Österreich und der Schweiz entstanden sind. Die Arbeiten behandeln aktuelle Themen der Wirtschaftswissenschaften und vermitteln innovative Beiträge für Wissenschaft und Praxis. Informationen zum Einreichungsvorgang und eine Übersicht unserer Publikationsangebote finden Sie hier.

---

Muhammed Fatih Kaya

# Automated Pattern Recognition of Communication Behaviour in Electronic Business Negotiations

 Springer Gabler

Muhammed Fatih Kaya  
Information Systems 1  
University of Hohenheim  
Stuttgart-Hohenheim, Germany

Dissertation University of Hohenheim, Germany, 2022  
D100

ISSN 2731-3220

ISSN 2731-3239 (electronic)

Gabler Theses

ISBN 978-3-658-40533-5

ISBN 978-3-658-40534-2 (eBook)

<https://doi.org/10.1007/978-3-658-40534-2>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer  
Fachmedien Wiesbaden GmbH, part of Springer Nature 2023

This work is subject to copyright. All rights are solely and exclusively licensed 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 Gabler 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

---

## Foreword

Electronic negotiations are nowadays a status quo in business negotiations. Using digital media makes communication even more important than in face-to-face interactions. Communication in negotiations takes place on a functional level, on a process level, and on a relational level and is one of the key factors for success or failure of negotiation processes.

Existing analyses of negotiation communication aim at transforming qualitative communication data into quantitative data. This is mostly done by manual coding which is not only expensive, laborious, and time-consuming but can only be applied to small data sets. An alternative is to use dictionary-based approaches that are application-specific and need to be created before using them for analysis. There is not yet a dictionary that has been designed for negotiations.

An automated approach based on machine learning approaches could overcome such limitations. Large negotiation data could be handled and communication patterns could be automatically detected. This is the goal of the research described in the present book. It analyses how descriptive and predictive communication patterns in electronic negotiation data can be detected and which algorithms for quantification and dimension reduction work best to maintain data richness. Clustering methods are analysed next to show their performance regarding pattern recognition in high-dimensional negotiation communication data. Finally, the research shows which patterns are detected using the novel approach developed by the author.

Whilst most analysis approaches take the structured and quantitative data from the decision side of negotiations, the present book solves the much more challenging topic of a structured automatic analysis of unstructured and multi-dimensional communication data. Every step was evaluated in detail concerning various algorithms and AI approaches that the author implemented and applied

to large negotiation data sets. This has not been done before in such depth and to such large extent.

The topic of automated communication analysis for negotiation data is highly relevant for negotiation researchers and AI researchers alike who work on the digitalisation of negotiations or work on other unstructured high-dimensional data. The present work paves the way for predictive approaches that offer in-process support of electronic negotiations to predict success or failure of negotiations which is one of the grand challenges of negotiation research.

This excellent work deserves wide-spread dissemination and will find its way into new research approaches without any doubt.

Stuttgart  
December 2022

Professor Mareike Schoop, PhD

---

## Preface

Developments in the field of Artificial Intelligence (AI) are evolving rapidly and reveal numerous potentials for the economy and society. Especially because of the data age in which mankind finds itself, the use of AI technology manifests numerous analytical possibilities. The application of AI can provide valuable benefits in many areas, beginning with e.g., the use for the detection of cancer cells up to predictive maintenance tasks in the automotive industry. The diverse application and the ability of these methods to derive valuable knowledge from large data sets led me to focus on this technological field during my university studies in Information Systems. During this time, I got to know the application area of interorganisational electronic negotiations through my scientific work at the chair of Information Systems 1 at the University of Hohenheim. My will to extend this methodical AI knowledge and investigate its application potentials increased steadily.

My PhD journey began at the beginning of the year 2018 to explore the pattern recognition potential of Machine Learning (ML) in business communication data. This PhD thesis analytically evaluates and applies several ML methods by considering the context of electronic negotiations. As central results, real-world data representations of high-dimensional communication data are generated which are used for the investigation of the automated pattern recognition potential of ML models in natural communication language. The findings provide valuable answers to numerous questions and make an important contribution to researchers and practitioners who are concerned with the digital transformation of electronic negotiations and want to learn more about the potential of AI.

The successful completion of this PhD thesis involved several challenges and would not have been possible without the support of many people. Therefore,



I would like to thank all those who have supported me in this journey, both in terms of content and mental way!

A special thank goes to my supervisor Prof. Mareike Schoop, PhD who supported me with valuable discussions and recommendations during the whole PhD process. She gave me the opportunity to conduct my research in this topic area very early on, supported me in developing my research profile and facilitated exchanges at various international conferences. All these experiences made me the researcher I am today.

I would like to thank Prof. Dr. Henner Gimpel for co-supervising my thesis and Prof. Dr. Caroline Ruiner for chairing the board of examiners. Furthermore, I would like to thank all my colleagues and ex-colleagues from the chair of Information Systems 1—Dr. Bernd Schneider, Dr. Andreas Schmid, Dr. Michael Körner, Dr. Philipp Melzer, Dr. Annika Lenz, Marlene Meyer, Josepha Witt and Stefan Ullmann. The collegial exchange, the exciting discussions and especially the harmonious team spirit have always provided a pleasant research atmosphere.

I would also like to thank my parents Seval Kaya and Ali Kaya and my sisters Nur Gülbahar Kaya and Ayse Betül Kaya in particular. I am grateful for their patience, prayers, and their mental support. Another special thank goes to my wife Sema Kaya who accompanied me throughout the whole PhD phase by showing thoughtfulness, love and motivated me with her kind words. Also special thanks to my cute daughter Feyza Kaya who was born at the end of 2021. Her enchanting smile gave me a lot of power during the difficult final phase of the doctoral work. Thank you all for believing in me. I am glad to have you all as part of my family!

Denkendorf  
December 2022

Muhammed Fatih Kaya

---

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	The Importance of Communication in Electronic Business Negotiations	2
1.2	The Need for the Application of Machine Learning	4
1.3	Overall Research Objective and Research Questions	7
1.4	Structure of the Thesis	10
<b>2</b>	<b>Application of Data Mining Methods for Pattern Recognition in Negotiation Support Systems</b>	<b>13</b>
2.1	Introduction	13
2.2	The Application Field of Negotiation Support Systems	14
2.3	The Implementation of KDD Using Data Mining	15
2.4	Data Processing	17
2.4.1	Processing Textual Negotiation Messages Using LIWC	18
2.4.2	Consistent Utility Values	20
2.5	Results	21
2.5.1	Association Rule Discovery	21
2.5.2	Decision Tree	24
2.6	Discussion and Outlook	25
<b>3</b>	<b>Advanced Maintenance of Data Richness in Business Communication Data—An Evaluation of Dimensionality Reduction Techniques</b>	<b>29</b>
3.1	Introduction	29
3.2	Theoretical Background	31
3.2.1	Application Field: Electronic Business Negotiations	31

3.2.2	Methodology of Knowledge Discovery in Databases . . . .	32
3.3	Research Framework . . . . .	33
3.4	Results . . . . .	42
3.5	Discussion . . . . .	49
3.6	Conclusion and Outlook . . . . .	54
<b>4</b>	<b>Analytical Comparison of Clustering Techniques for the Recognition of Communication Patterns . . . . .</b>	<b>57</b>
4.1	Motivation . . . . .	57
4.2	Theoretical Background . . . . .	59
4.2.1	The Importance of Communicative Interactions in Electronic Negotiations . . . . .	59
4.2.2	Clustering of High-dimensional Negotiation Messages . . . . .	61
4.3	Research Approach . . . . .	64
4.3.1	Dimensionality Reduction . . . . .	65
4.3.2	Calculation of Similarity Measure . . . . .	66
4.3.3	Evaluation of Optimal Cluster Number . . . . .	66
4.3.4	Clustering Techniques . . . . .	68
4.3.5	Performance Evaluation . . . . .	69
4.4	Results . . . . .	71
4.5	Discussion . . . . .	82
4.6	Conclusion and Outlook . . . . .	87
4.7	Appendix . . . . .	90
<b>5</b>	<b>Pattern Labelling of Business Communication Data . . . . .</b>	<b>91</b>
5.1	Introduction . . . . .	91
5.2	Theoretical Background . . . . .	94
5.2.1	Support of E-Negotiations . . . . .	94
5.2.2	Negotiation Behaviour . . . . .	94
5.2.3	Ways of Pattern Labelling in High-Dimensional Business Communication Data . . . . .	96
5.3	Research Procedure . . . . .	101
5.4	Results . . . . .	103
5.5	Discussion . . . . .	112
5.6	Summary and Outlook . . . . .	117
5.7	Appendix . . . . .	119
5.7.1	Functionality of used ML-methods . . . . .	119
5.7.2	ML-based Process Pipeline of the Research Procedure . . . . .	122

---

<b>6 Discussion and Outlook</b> .....	125
6.1 Summary .....	125
6.2 Discussion .....	127
6.3 Limitations .....	131
6.4 Research Contribution .....	133
6.5 Future Research .....	135
<b>References</b> .....	137

---

# Abbreviations

ARD	Association Rule Discovery
AUC	Area Under the ROC Curve
DB	Davies Bouldin
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DM	Data Mining
DT	Decision Tree
FE	Feature Extraction
FPR	False Positive Rate
FS	Feature Selection
GBDT	Gradient Boosting Decision Tree
KDD	Knowledge Discovery Process
LDA	Latent Dirichlet Allocation
LIWC	Linguistic Inquiry and Word Count
LR	Logistic Regression
MAUT	Multi Attribute Utility Function
ML	Machine Learning
NB	Naive Bayes
NNET	Neural Network
NSS	Negotiation Support System
OS	Optimise Selection
PCA	Principal Component Analysis
RF	Random Forest
ROC	Receiver Operating Characteristics
RPC	Rate of Perplexity Change
SVD	Singular Value Decomposition
SVM	Support Vector Machine

TDM	Term Document Matrix
TF	Term Frequency
TF-IDF	Term Frequency Inverse Document Frequency
TM	Topic Modelling
TPR	True Positive Rate

---

# List of Figures

Figure 1.1	Structure of the Thesis .....	11
Figure 2.1	The Knowledge Discovery Process according to Fayyad et al. (1996) .....	16
Figure 2.2	Association Patterns—Balanced datasets .....	22
Figure 2.3	Association Patterns—Final Reject .....	23
Figure 2.4	Association Patterns—Final Accept .....	23
Figure 2.5	Performance Vector of the Decision Tree .....	24
Figure 3.1	Research Framework (cf. Fayyad et al. 1996; Kumar and Bhatia 2013; Vidya and Aghila 2010) .....	34
Figure 3.2	The Term Document Matrix (TDM) .....	35
Figure 3.3	The Evaluation of Classification Models via ROC and AUC .....	41
Figure 4.1	Overview of Clustering Techniques (cf. Santhisree and Damodaram 2011; Saxena et al. 2017) .....	62
Figure 4.2	Research Approach (cf. Kadhim et al. 2014; Rana et al. 2019) .....	64
Figure 4.3	Elbow Method for Cluster Number Determination based on Feature Selected Dataset ( <i>OS-dataset</i> ) .....	72
Figure 4.4	Elbow Method for Cluster Number Determination based on Feature Extracted Dataset ( <i>PCA-dataset</i> ) .....	73
Figure 4.5	Davies Bouldin evaluation based on Feature Selected dataset ( <i>OS-dataset</i> ) .....	74
Figure 4.6	Silhouette Evaluation based on Feature Extracted dataset ( <i>PCA-dataset</i> ) .....	75
Figure 4.7	Silhouette Evaluation based on Feature Selected dataset ( <i>OS-dataset</i> ) .....	75

---

Figure 4.8	Cluster partitioning for k-Means Clustering .....	76
Figure 4.9	Cluster partitioning for x-Means Clustering .....	77
Figure 4.10	Cluster partitioning for DBSCAN Clustering .....	78
Figure 4.11	Internal Performance Evaluation Results .....	80
Figure 4.12	External Performance Evaluation Results .....	81
Figure 5.1	Decision Tree Example .....	98
Figure 5.2	Topic Model Example .....	100
Figure 5.3	Overall Research Procedure .....	102
Figure 5.4	Topics of OS-dataset with k-Means ( $k = 5$ ) .....	106
Figure 5.5	Topics of PCA-dataset with k-Means ( $k = 5$ ) .....	108
Figure 5.6	Topics of PCA-dataset with k-Means ( $k = 7$ ) .....	109
Figure 5.7	Topics of PCA-dataset with DBSCAN .....	110
Figure 5.8	ML-based Process Pipeline .....	122
Figure 6.1	Meta Research Process of the Doctoral Thesis .....	126



---

## List of Tables

Table 3.1	Number of reduced dimensions for Correlation Removal . . . .	43
Table 3.2	Number of reduced dimensions for Feature Selection variations . . . . .	43
Table 3.3	Number of reduced dimensions for Feature Extraction variations . . . . .	45
Table 3.4	AUC-Results for Optimize Selection . . . . .	46
Table 3.5	AUC-Results for Singular Value Decomposition . . . . .	47
Table 3.6	AUC-Results for Principal Component Analysis . . . . .	48
Table 4.1	Average of the largest sum of Cluster shares between the Clustering methods (at least 50% share on both respective cluster partitions) . . . . .	90
Table 5.1	Extracted Association Rules of OS-dataset with k-Means (k = 5) . . . . .	105
Table 5.2	Extracted Association Rules of PCA-dataset with k-Means (k = 5) . . . . .	107
Table 5.3	Extracted Association Rules of PCA-dataset with k-Means (k = 7) . . . . .	108
Table 5.4	Extracted Association Rules of PCA-dataset with DBSCAN . . . . .	109
Table 5.5	Overview of derived Label Indicators . . . . .	111