

Feyza Altinay

Anwendung von Data Mining Verfahren bei politischen Wahlkämpfen

Bachelorarbeit

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 Verlags. 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 © 2017 Diplomica Verlag GmbH
ISBN: 9783961161911

Anwendung von Data Mining Verfahren bei politischen Wahlkämpfen

Inhalt

Abbildungsverzeichnis	ii
Tabellenverzeichnis	iii
Diagrammverzeichnis.....	iv
1 Einleitung	1
2 Wählerbeeinflussung.....	2
2.1 Problemstellung	2
2.2 Geschichtliche Entwicklung.....	4
2.3 Zielsetzung der Bachelorarbeit.....	9
3 Data Mining	10
4 Microtargeting	21
5 Anwendungsszenario.....	25
5.1 Datengrundlage.....	25
5.2 Clusterbildung mit RapidMiner.....	32
5.3 Ergebnispräsentation und Evaluation	46
6 Fazit / Ausblick	53
Anhang.....	54
Literaturverzeichnis	58

Abbildungsverzeichnis

Abbildung 1: Das CRISP-DM Modell	11
Abbildung 2: Abstimmung im US-Senat	13
Abbildung 3: Schlechtes und optimales Clustering.....	16
Abbildung 4: Datenpunkte und 1 Iteration.....	18
Abbildung 5: 2 und 3 Iterationen.....	18
Abbildung 6: 4 und 5 Iterationen.....	18
Abbildung 7: Operator 'Read Excel'	33
Abbildung 8: Import Configuration Wizard.....	33
Abbildung 9: Import Configuration Wizard (2)	33
Abbildung 10: Operator 'Generate ID'	34
Abbildung 11: Operator 'Select Attributes'	34
Abbildung 12: Parametereinstellung des 'Select Attributes'	35
Abbildung 13: Operator 'Set Role'	35
Abbildung 14: Parametereinstellung 'Set Role'	36
Abbildung 15: Operator 'Shuffle'	36
Abbildung 16: Operator 'Clustering (k-Means)'	38
Abbildung 17: Parametereinstellung des 'Clustering (k-Means)'	38
Abbildung 18: Erstellte Cluster (1)	39
Abbildung 19: Erstellte Cluster (2)	39
Abbildung 20: Parametereinstellung des 'Data to Similarity'	45
Abbildung 21: 'Data to Similarity' und 'Performance'	45
Abbildung 22: Silhouetten-Messwerte	46
Abbildung 23: Dendrogram US Senat 2003.....	54
Abbildung 24: Symmetrische Unähnlichkeitsmatrix	55

Tabellenverzeichnis

Tabelle 1: Basisdaten für das k-Means Verfahren.....	19
Tabelle 2: Numerische Werte für das k-Means Verfahren.....	20
Tabelle 3: Beispieldatensätze	30

Diagrammverzeichnis

Diagramm 1: Centroiden zum 'Alter'	40
Diagramm 2: Centroiden zu 'Westen_jn'	40
Diagramm 3: Centroiden zu 'Süden_jn'	40
Diagramm 4: Centroiden zu 'Akademiker'	40
Diagramm 5: Centroiden zum 'Einkommen'	41
Diagramm 6: Centroiden zu 'verheiratet'	41
Diagramm 7: Centroiden zu 'Sport_jn'	41
Diagramm 8: Centroiden zu 'Kunst_jn'	41
Diagramm 9: Centroiden zu 'Reality-Shows_jn'	42
Diagramm 10: Centroiden zu 'Serien_jn'	42
Diagramm 11: Centroiden zu 'Current-affairs_jn'	42
Diagramm 12: Centroiden zu 'Science-Fiction_jn'	42
Diagramm 13: Centroiden zu 'Drama_jn'	43
Diagramm 14: Centroiden zu 'Pop_jn'	43
Diagramm 15: Centroiden zu 'Country_jn'	43
Diagramm 16: Cluster 0 - Demografie und Interessen.....	46
Diagramm 17: Cluster 0 - Präferenzen.....	47
Diagramm 18: Cluster 1 - Demografie und Interessen.....	48
Diagramm 19: Cluster 1 - Präferenzen.....	48
Diagramm 20: Demografien und Interessen der Konservativen und Liberalen im Datenbestand.....	49
Diagramm 21: Präferenzen der Konservativen und Liberalen im Datenbestand	50
Diagramm 22: Altersverteilung in den Clustern.....	51
Diagramm 23: Altersverteilung im Datenbestand.....	51
Diagramm 24: Einkommensverteilung in den Clustern	52
Diagramm 25: Einkommensverteilung im Datenbestand.....	52