

1	Thematische Einführung, Problemstellung und Zielsetzung	1
1.1	Thematische Einführung	1
1.2	Problemstellung	4
1.3	Zielsetzung	6
1.4	Übersicht über eine Auswahl bedeutender Literatur	8
2	Datensatz	11
2.1	Qualitative Datenbeschreibung	11
2.2	Datenaufbereitung	14
2.3	Quantitative Datenbeschreibung	16
2.4	Erklärende Variablen	18
3	Methodik	21
3.1	Entscheidungsbäume	23
3.1.1	Traditioneller Klassifizierungsbaum	23
3.1.2	Random Forest	25
3.2	Logistische Modelle	27
3.2.1	Binäres logistisches Modell	27
3.2.2	Ordinales logistisches Modell	31
3.3	Evaluierung der Modellgüte	32
3.3.1	Konfusionsmatrix	33
3.3.2	Korrektklassifikationsrate und Falschklassifikationsrate	35
3.3.3	Falsch-Positiv-Rate und Falsch-Negativ-Rate	38
3.3.4	Falsch-Negativ-Rate und Richtig-Positiv-Rate	40
3.3.5	Positiver Vorhersagewert und F1-Maß	42

3.3.6 ROC und AUC	43
3.3.7 Statistische Testverfahren	46
4 Modellierung, Schätzung und Analyse des IG-Ratings	47
4.1 Analyse des Modelldatensatzes (2020)	47
4.1.1 Entscheidungsbäume	47
4.1.1.1 Traditioneller Klassifizierungsbaum	48
4.1.1.2 Random Forest	56
4.1.2 Logistische Modelle	60
4.1.2.1 Logit-Modell mit Aufnahme aller Regressoren	60
4.1.2.2 Logit-Modell mit Aufnahme signifikanter Regressoren	64
4.1.3 Zusammenfassung der Ergebnisse des Modelldatensatzes	66
4.1.4 Validierung der Ergebnisse des Modelldatensatzes	68
4.2 Analyse des historischen Datensatzes (2019)	70
4.2.1 Entscheidungsbäume	70
4.2.1.1 Traditioneller Klassifizierungsbaum	70
4.2.1.2 Random Forest	77
4.2.2 Logistische Modelle	81
4.2.2.1 Logit-Modell mit Aufnahme aller Regressoren	81
4.2.2.2 Logit-Modell mit Aufnahme signifikanter Regressoren	84
4.2.3 Zusammenfassung der Ergebnisse des historischen Datensatzes	86
4.2.4 Validierung der Ergebnisse des historischen Datensatzes	88
4.3 Analyse des aktuellen Datensatzes (2021/2022)	90
4.3.1 Entscheidungsbäume	90
4.3.1.1 Traditioneller Klassifizierungsbaum	90
4.3.1.2 Random Forest	95
4.3.2 Logistische Modelle	99
4.3.2.1 Logit-Modell mit Aufnahme aller Regressoren	99
4.3.2.2 Logit-Modell mit Aufnahme signifikanter Regressoren	99
4.3.3 Zusammenfassung der Ergebnisse der Schätzung auf aktuellen Daten	101

5 Modellierung, Schätzung und Analyse der Ratingklasse	105
5.1 Analyse des Modelldatensatzes (2020)	105
5.1.1 Random Forest	105
5.1.2 Ordinal logistische Modelle	109
5.1.2.1 Ordinal logistisches Modell mit Aufnahme aller Regressoren	109
5.1.2.2 Ordinal logistisches Modell mit Aufnahme signifikanter Regressoren	111
5.1.3 Zusammenfassung der Ergebnisse des Modelldatensatzes	112
5.2 Analyse des historischen Datensatzes (2019)	114
5.2.1 Random Forest	114
5.2.2 Ordinal logistische Modelle	117
5.2.2.1 Ordinal logistisches Modell mit Aufnahme aller Regressoren	117
5.2.2.2 Ordinal logistisches Modell mit Aufnahme signifikanter Regressoren	119
5.2.3 Zusammenfassung der Ergebnisse des historischen Datensatzes	121
5.3 Analyse des aktuellen Datensatzes (2021/2022)	122
5.3.1 Random Forest	123
5.3.2 Ordinal logistische Modelle	125
5.3.2.1 Ordinal logistisches Modell mit Aufnahme aller Regressoren	125
5.3.2.2 Ordinal logistisches Modell mit Aufnahme signifikanter Regressoren	125
5.3.3 Zusammenfassung der Ergebnisse des aktuellen Datensatzes	127
6 Fazit und Ausblick	129
6.1 Fazit	129
6.2 Ausblick	131
Literaturverzeichnis	133