

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN
PROFESSUR FÜR ANWENDUNGSSYSTEME UND E-BUSINESS

PROJEKTSTUDIUM
SOMMERSEMESTER 2022

Konzeption und Implementierung eines Dashboards zur Darstellung
KI-gestützter Insolvenzprognose mithilfe mehrerer und differenzierter
Kennzahlen über die Unternehmenslage

Micha Alexander Teweleit
Matr.-Nr.: 21676005
Wirtschaftsinformatik

Joachim von Alvensleben
Matr.-Nr.: 11780318
Wirtschaftsinformatik

Inhaltsverzeichnis

Abbildungsverzeichnis.....	V
Tabellenverzeichnis.....	VII
Abkürzungsverzeichnis.....	VIII
1 Einleitung.....	1
2 Theoretischer Hintergrund	3
2.1 Betriebswirtschaftliche Bestandteile der Insolvenzprognose	3
2.2 KI-gestützte Insolvenzprognose.....	4
2.3 Explainable Artificial Intelligence.....	6
2.3.1 Konfusionsmatrix	9
2.3.2 Kennzahlen zur Messung der Schätzgüte	9
2.3.3 SHAP - Shapley Additive Explanations	11
2.3.4 XGBoost Feature Importance	12
3 Methodisches Vorgehen.....	13
3.1 Literaturreview zur Anforderungsanalyse für erklärbares ML in der Insolvenzprognose	13
3.2 Hybrider Prototyping-Ansatz zur Entwicklung eines Dashboards	17
4 Anforderungsanalyse für erklärbares ML in der Insolvenzprognose	20
4.1 Anforderungsanalyse zur Förderung der Erklärbarkeit ML-basierter Insolvenzprognose	20
4.2 Anforderungen an Datenerhebung	27
5 Konzeption eines Dashboards zur KI-gestützten Insolvenzprognose	29
6 Praktische Entwicklung eines Dashboards zur KI-gestützten Insolvenzprognose ...	37
6.1 Datenvorverarbeitung	37
6.1.1 Datenbereinigung	38
6.1.2 Feature Engineering.....	39

6.1.3 Segmentierung und Modellvorbereitung der Daten	39
6.2 Modellentwicklung	43
6.3 Implementierung eines Dashboards zur Visualisierung ML-basierter Insolvenzprognose	51
6.3.1 Technologien und Frameworks	51
6.3.2 Prozess Workflow des Dashboards.....	51
6.3.3 Landing Page, Prozessdokumentation und persönliche Übersicht	54
6.3.4 Daten-Upload und Daten-Übersicht.....	55
6.3.5 Datenbereinigung	58
6.3.6 Feature Engineering.....	59
6.3.7 Daten-Segmentierung.....	60
6.3.8 Modell-Setup und Optimierung.....	62
6.3.9 Modell-Beurteilung.....	65
7 Diskussion der Ergebnisse.....	73
7.1 Die Rolle der Erklärbarkeit bei der Verbesserung KI-gestützter Insolvenzprognosen	73
7.1.1 Anforderungsanalyse und Implementierung von Erklärbarkeitsmethoden .	74
7.1.2 Anwendung und Nutzbarkeit der Erklärbarkeitsansätze in der Praxis.....	75
7.1.3 Erreichte Prognoseergebnisse und Resultate der Interpretierbarkeitsmethoden.....	76
7.2 Das Dashboard als Instrument: Konfigurierbarkeit und Erklärbarkeit in der KI- gestützten Insolvenzprognose	81
7.2.1 Unterstützung der Konfigurierbarkeit und Erklärbarkeit durch das Dashboard	81
7.2.2 Verbesserung der Konfigurierbarkeit und Erklärbarkeit durch das Dashboard	82
8 Fazit	84
Anhang	87

Literaturverzeichnis	IX
-----------------------------------	-----------

Abbildungsverzeichnis

Abbildung 1: Hybrider Prototyping-Ansatz.....	18
Abbildung 2: Mockup - Datenbank für Datensätze und Modelle.....	30
Abbildung 3: Mockup - Feature-Auswahl	31
Abbildung 4: Mockup - Segmentierung der Daten	32
Abbildung 5: Mockup - Deskriptive Auswertung der segmentierten Daten	32
Abbildung 6: Mockup - Auswahl von Oversampling und Undersampling Techniken	33
Abbildung 7: Mockup - Auswahl KI-Modell	34
Abbildung 8: Mockup - Hyperparameter-Auswahl und Splittung des Datensatzes	34
Abbildung 9: Mockup - Kennzahlen zur Schätzgüte	35
Abbildung 10: Mockup - Feature Importance	35
Abbildung 11: Prozessworkflow der gesamten Applikation.....	53
Abbildung 12: Landing-Page	54
Abbildung 13: Prozessdokumentation.....	54
Abbildung 14: Persönliche Übersicht der Datensätze und KI-Modelle	55
Abbildung 15: Daten-Upload	56
Abbildung 16: Datensatz nach Nutzer	56
Abbildung 17: Deskriptive Auswertung des rohen Datensatzes	58
Abbildung 18: Hinweise zur Datenbereinigung	59
Abbildung 19: Selektion der Feature	60
Abbildung 20: Auswahl der Segmente.....	60
Abbildung 21: Auswahl der Gruppengröße	61
Abbildung 22: Auswahl der Grenzen für die Gruppen	61
Abbildung 23: Modell-Setup.....	63
Abbildung 24: Parameterauswahl	65
Abbildung 25: Metrik zur Beurteilung der Güte des Modells.....	66
Abbildung 26: Ausschnitt der Hyperparameter je Modell.....	67
Abbildung 27: Ausschnitt der XGBoost Feature Importance.....	68
Abbildung 28: SHAP - Bar Summary Plot	69
Abbildung 29: SHAP - Scatter Summary Plot	69
Abbildung 30: SHAP - Dependence Plot	70
Abbildung 31: Konfusionsmatrix.....	70

Abbildung 32: Filterfunktion für einzelne Vorhersagen	72
Abbildung 33: Homepage der Applikation	87
Abbildung 34: Datenupload	87
Abbildung 35: Hyperparameter Einstellungen für Neuronales Netzwerk	88
Abbildung 36: Auswahl der Erklärbarkeitsmethoden.....	88

Tabellenverzeichnis

Tabelle 1: Suchstrings nach Themenfeldern geordnet	14
Tabelle 2: Initiale Suchergebnisse nach Datenbanken	15
Tabelle 3: Vorauswahl der Suchergebnisse nach Datenbanken.....	16
Tabelle 4: Finale Auswahl der Literatur	16
Tabelle 5: Auflistung der Anforderungen	27
Tabelle 6: Abbildungen der Mockups nach Anforderungen sortiert.....	30
Tabelle 7: Metriken zur Insolvenzprognose	39
Tabelle 8: Differenzierung der Unternehmen nach der Bilanzsumme.....	40
Tabelle 9: Beispiel von Paneldaten	41
Tabelle 10: Anzahl der Insolvenzen pro Gruppe und Anzahl der berücksichtigten Zeiträume im Wide-Format	42
Tabelle 11: Bewertungsergebnisse der einzelnen XGBoost-Modelle mit der Oversampling-Technik SMOTE.....	45
Tabelle 12: Bewertungsergebnisse der einzelnen XGBoost-Modelle mit der Oversampling-Technik SMOTE-Tomek	46
Tabelle 13: Bewertungsergebnisse der einzelnen XGBoost-Modelle mit der Oversampling-Technik SMOTE-ENN	46
Tabelle 14: Anzahl der Insolvenzen pro Gruppe und Anzahl der berücksichtigten Zeiträume im Long-Format angepasst für LSTM	49
Tabelle 15: Bewertungsergebnisse der einzelnen LSTM-Modelle	50
Tabelle 16: Schätzgüte von Kapitalgesellschaften beispielhaft mit Bayesian Optimization	77
Tabelle 17: Merkmalswichtigkeit von großen und mittelgroßen Kapitalgesellschaften beispielhaft mit Bayesian Optimization	79
Tabelle 18: Merkmalswichtigkeit von mittelgroßen Kapitalgesellschaften beispielhaft in zwei Modellen.....	80

Abkürzungsverzeichnis

Anlagendeckungsgrad II/Asset Coverage Ratio	-	ACR
Artificial Intelligence	-	AI
Artificial Intelligence Act	-	AIA
Benutzeroberfläche/User Interface	-	UI
Bundesanstalt für Finanzdienstleistungsaufsicht	-	BaFin
Edited Nearest Neighbor	-	ENN
Eigenkapitalquote/Equity Ratio	-	ER
Eigenkapitalrentabilität/Return on Equity	-	ROE
Explainable Artificial Intelligence	-	XAI
Extreme Gradient Boosting	-	XGBoost
False Negative	-	FN
False Positive	-	FP
Gated Recurrent Unit	-	GRU
Gesamtkapitalrentabilität/Return on Total Assets	-	RTA
Insolvenzordnung	-	InsO
Künstliche Intelligenz	-	KI
Kurzfristige Verschuldungsquote/Short-term Debt Ratio	-	STDR
Layer-wise Relevance Propagation	-	LRP
Liquidität zweiten Grades/Second Degree Liquidity	-	L2
Local Interpretable Model-Agnostic Explanations	-	LIME
Long Short-Term Memory	-	LSTM
Maschinelles Lernen / Machine Learning	-	ML
Partielle Abhängigkeitsanalyse/Partial Dependence Plot	-	PDP
Receiver Operating Characteristic – Area Under the Curve	-	ROC-AUC
Rekurrentes neuronales Netzwerk/Recurrent Neural Network	-	RNN
Shapley Additive Explanations	-	SHAP
Support Vector Machine	-	SVM
Synthetic Minority Oversampling Technique	-	SMOTE
True Negative	-	TN
True Positive	-	TP
Working Capital Ratio	-	WCR

1 Einleitung

In der sich ständig weiterentwickelnden Welt der Finanztechnologie ist die Fähigkeit, genaue und frühzeitige Insolvenzprognosen zu erstellen, von entscheidender Bedeutung (Ashraf et al., 2019). Künstliche Intelligenz (KI) hat sich in diesem Bereich als leistungsfähiges Werkzeug erwiesen (Lahmiri & Bekiros, 2019). Eine mögliche Zukunftsvision ist, dass automatisierte Entscheidungen auf der Grundlage von KI-gestützten Insolvenzprognosen getroffen werden können (McKinsey & Company, 2021). Allerdings bestehen regulatorische Anforderungen, die sicherstellen sollen, dass die Entscheidungsfindung in automatisierten Prozessen nachvollziehbar bleibt. Nach der 2018 in Kraft getretenen EU-Datenschutz-Grundverordnung (Artikel 22 EU-DSGVO) hat eine Person ein "Recht auf Erklärung" bei algorithmischen Entscheidungen, die sie betreffen (Europäische Union, 2016). Sollte das Szenario der automatisierten individuellen Entscheidungsfindung durch KI-basierte Insolvenzprognosen Realität werden, wird die Erklärbarkeit der Entscheidungsfindung daher von zentraler Bedeutung sein. Daraus leitet sich die folgende Forschungsfrage ab:

1. Welche Ansätze zur Förderung der Erklärbarkeit von Machine Learning Modellen können die Ergebnisse einer KI-gestützten Insolvenzprognose verbessern?

Da die Interaktion zwischen Nutzern und KI-Systemen eine wesentliche Rolle für die Akzeptanz und Nutzung dieser Technologien spielt, ist eine transparente und nutzerfreundliche Darstellung des Prognoseprozesses und der Prognoseergebnisse unerlässlich (Shin, 2021). Dabei kommt der Konzeption und Gestaltung eines Dashboards eine entscheidende Rolle zu (Viégas & Wattenberg, 2023). Es stellt eine Schnittstelle dar, über die die KI-Modelle transparent erstellt und ihre Ausgaben visualisiert und erläutert werden können. Die zweite Forschungsfrage lautet daher:

2. Wie muss ein Dashboard aufgebaut sein, um eine KI-gestützte Insolvenzprognose möglichst konfigurierbar und erklärbar zu machen?

Die Relevanz dieser Forschung liegt in der wachsenden Bedeutung von KI in der Finanzindustrie und der Notwendigkeit, diese Technologie zugänglicher und verständlicher zu machen. Durch die Verbesserung der Erklärbarkeit von KI-Modellen

und die Gestaltung benutzerfreundlicher Dashboards kann die Akzeptanz und Anwendung dieser Technologien in der Praxis gefördert werden. Ziel dieser Arbeit ist es, einen substantiellen Beitrag zur aktuellen Forschung im Bereich der KI-gestützten Insolvenzprognose zu leisten und praktische Lösungen für die bestehenden Herausforderungen in diesem Forschungsschwerpunkt zu entwickeln. Es wird angestrebt, sowohl die theoretische als auch die praktische Perspektive zu beleuchten, um ein umfassendes Verständnis des Themas zu ermöglichen und die Anwendung von KI in der Finanzbranche zu fördern.

Die Arbeit gliedert sich wie folgt: Nach der Einleitung folgt ein Literaturüberblick, der den aktuellen Stand der Forschung wiedergibt. Danach wird die Methodik der Arbeit beschrieben. Basierend auf einer Anforderungsanalyse für erklärbares Machine Learning (ML) in der Insolvenzprognose folgt die Konzeption des Dashboards. Im Hauptteil der Arbeit werden eigene KI-gestützte Insolvenzprognosen vorgestellt und deren Ergebnisse evaluiert. Im Weiteren wird die praktische Umsetzung des Dashboards beschrieben. Abschließend folgt eine Diskussion der Ergebnisse und ein Fazit, in dem die Forschungsfragen beantwortet und mögliche zukünftige Forschungsrichtungen aufgezeigt werden.

2 Theoretischer Hintergrund

Um die Bedeutung von Insolvenzprognosen und insbesondere KI-gestützten Insolvenzprognosen zu verstehen, ist es notwendig, die folgenden grundlegenden Begriffe theoretisch abzuleiten.

2.1 Betriebswirtschaftliche Bestandteile der Insolvenzprognose

Die Insolvenzordnung (InsO) in Deutschland befasst sich mit der Abwicklung von Insolvenzen bei Unternehmen, die aufgrund von Zahlungsunfähigkeit oder Überschuldung entstehen (Insolvenzordnung, 2022, §§ 16-19). Dabei wird ein Insolvenzverwalter ernannt (Insolvenzordnung, 2022, § 56), dessen Hauptaufgabe darin besteht, die Vermögenswerte des betroffenen Unternehmens zu verwerten und den Forderungen der Gläubiger bestmöglich Folge zu leisten (Insolvenzordnung, 2022, § 148). Der Zweck des Insolvenzverfahrens ist die gemeinschaftliche Befriedigung der Gläubiger (Insolvenzordnung, 2022, § 1).

Nach HARHOFF ET AL. (Harhoff et al., 1998) haben Unternehmen mit beschränkter Haftung über Branchen hinweg höhere Wachstums- und Insolvenzraten. Dies ist für den in Kapitel 4.2 beschriebenen Datensatz relevant, da dieser auf öffentlich zugänglichen Unternehmensdaten basiert, die vermehrt für Unternehmen mit beschränkter Haftung verfügbar sind.

Die Insolvenzprognose, definiert als Verfahren zur Bestimmung der Insolvenzwahrscheinlichkeit von Schuldnern innerhalb eines bestimmten Zeitraums (z.B. ein Jahr), ist im Geschäftsbereich mit Firmenkunden von Bedeutung. Sie dient zur Prüfung der Kreditwürdigkeit und Beurteilung von Kreditrisiken (Hölscher & Helms, 2018). Sowohl, wie folgend weiter erläutert, künstliche Intelligenz, als auch klassische empirisch-statistische Methoden, können für die Insolvenzprognose verwendet werden. Simulationsgestützte Verfahren ergänzen die Prognose und unterstützen dabei, die erwartete Unternehmensentwicklung und Insolvenzwahrscheinlichkeit aus verschiedenen Zukunftsszenarien abzuleiten. Ratingagenturen nutzen derartige Methoden beispielsweise, um Ratingklassifizierungen für Unternehmen und Staaten zu erstellen (Hölscher & Helms, 2018).

2.2 KI-gestützte Insolvenzprognose

Im Risikomanagement und bei der Bewertung der Kreditwürdigkeit von Unternehmen ist, wie bereits ausgeführt, die Insolvenzprognose von entscheidender Bedeutung. ML-Ansätze werden in den letzten Jahren immer wichtiger, um die Genauigkeit solcher Vorhersagen zu verbessern und die Forschungsergebnisse voranzutreiben. Das Forschungsfeld umfasst die Untersuchung verschiedener ML-Techniken, die sich insbesondere durch ihre Leistung in Bezug auf die Insolvenzvorhersage auszeichnen. Dazu gehören unter anderem Support-Vektor-Maschinen, XGBoost, Logistische Regression, Entscheidungsbäume und Neuronale Netze (Kim & Kang, 2010; Min et al., 2006; Pamuk et al., 2021; Shin et al., 2005). Die Autoren diskutieren jedoch ebenfalls Bedenken hinsichtlich der Interpretierbarkeit von Ensemble-Methoden und betonen die Bedeutung der Standardisierung bei ML-Ansätzen, insbesondere im Fall unausgewogener Datensätze.

LE ET AL. (2018) untersuchen verschiedene Oversampling-Techniken für die Insolvenzprognose anhand eines koreanischen Datensatzes aus den Jahren 2016-2017. Ihre Experimente zeigen, dass durch ausgewogene Trainingsdaten in Folge von Oversampling die Leistung der Insolvenzprognose verbessert werden kann. Laut den Autoren werden die besten Ergebnisse, basierend auf dem verwendeten Datensatz, mit einer Kombination aus Synthetic Minority Oversampling Technique (SMOTE) und Edited Nearest Neighbor (ENN) für das Oversampling und einem Random Forest als Klassifikator erzielt (Le et al., 2018).

CHI UND CHU (2021) betrachten den Einsatz von Long Short-Term Memory (LSTM-) und Gated Recurrent Unit (GRU-) Modellen für die Insolvenzprognose taiwanesischer Unternehmen. Die Autoren betonen, dass tiefe neuronale Netze wie LSTM und GRU in der Lage sind, große Datenmengen effizient zu verarbeiten und genaue Vorhersagen zu treffen. Gleichzeitig betrachten sie kleine Stichproben und die Abhängigkeit von Vergangenheitsdaten als Limitationen (Chi & Chu, 2021).

CHOU (2019) untersucht einen Hybrid-Ansatz, der Entscheidungsbäume und Deep Neural Networks kombiniert, um ML-Methoden zur Insolvenzprognose zu untersuchen. Die vorgestellte Methode zielt darauf ab, die Vorhersagegenauigkeit zu erhöhen und gleichzeitig die Interpretierbarkeit der Entscheidungsbäume zu erhalten. CHOU (2019)

betont jedoch, dass bei einer Erhöhung der Vorhersagegenauigkeit der mögliche Verlust der Interpretierbarkeit berücksichtigt werden muss.

In einer bibliometrischen Analyse untersuchen SHI UND LI (2019) 413 englischsprachige Publikationen zum Thema Insolvenzprognose mit KI-Methoden aus den Jahren 1968-2018. Sie zeigen auf, dass neuronale Netze und multivariate Diskriminanzanalyse bereits lange in diesem Bereich genutzt werden. Seit 2010 sind laut der Autoren zunehmend Support Vector Machines (SVMs) und fallbasierte Argumentationsmethoden in Anwendung. Darüber hinaus weisen sie darauf hin, dass Bayesian Networks noch nicht ausreichend erforscht sind, was eine Forschungslücke darstellt, auf die im Zusammenhang mit der KI-gestützten Insolvenzprognose eingegangen werden kann (Shi & Li, 2019).

GAO, GAO UND ZOU (2022) untersuchen die Anwendung von ML-Algorithmen für die Vorhersage von Insolvenzen und betonen die Herausforderung der ungleichen Verteilung von positiven und negativen Beispielen. Um eine kostensensitive Insolvenzprognose auf Basis des verwendeten, unausgeglichene Datensatzes aus China zu ermöglichen, stellen sie das XGBoost-Modell mit einem gewichteten Kostenansatz und verschiedenen Resampling-Strategien vor. In diesem Kontext bedeutet "kostensensitiv," dass das XGBoost-Modell so angepasst ist, dass es die unterschiedlichen Kosten für falsche Klassifikationen berücksichtigt (Zou et al., 2022).

ALAKA ET AL. (2018) schlussfolgern, dass hybride Modelle, die die Stärken mehrerer Ansätze kombinieren, insgesamt die beste Leistung erzielen. Sie betonen, dass auf eine Balance zwischen Genauigkeit, Interpretierbarkeit und praktischer Anwendbarkeit geachtet werden soll.

Zusammenfassend zeigen die Studien, dass ML-basierte Insolvenzprognosen ein wirksames Risikomanagementinstrument sein können, das jedoch Herausforderungen in Bezug auf Interpretierbarkeit und Ungleichgewicht der Klassen beinhaltet. Die Wahl eines geeigneten ML-Modells und entsprechender Bewertungsmetriken, wie dem Brier Score (Pamuk et al., 2021), sollte daher sorgfältig abgewogen werden, um den aktuellen Stand der Forschung angemessen zu reflektieren und zukünftige Entwicklungen zu fördern.

2.3 Explainable Artificial Intelligence

Im Hinblick auf die Verständlichkeit wird die Bedeutung der Erklärbarkeit von KI- und Machine-Learning-Modellen bei der Vorhersage von Unternehmensinsolvenzen untersucht. Dabei werden auch Methoden der Explainable Artificial Intelligence (XAI) verwendet. Die Erklärbarkeit solcher Modelle ist entscheidend für ihre Akzeptanz und Vertrauenswürdigkeit in verschiedenen Anwendungsbereichen wie Finanzen, Recht und Medizin (Chou, 2019). KIM UND ROUTLEDGE (2022) argumentieren, dass ein Recht auf Erklärung bei algorithmischen Entscheidungen notwendig ist, um das Vertrauen in künstliche Intelligenz zu stärken. Die Autoren betonen die Bedeutung von Transparenz und Nachvollziehbarkeit gerade in sensiblen Bereichen wie dem Gesundheits- oder Finanzwesen (Kim & Routledge, 2022). RAI (2020) betont ebenfalls die Notwendigkeit der Erklärbarkeit von KI-Systemen sowohl aus ethischer Sicht als auch aus Sicht der Nutzerakzeptanz. Er verweist jedoch auch darauf, dass es eine Balance zwischen Erklärbarkeit und Leistungsfähigkeit geben muss. Deshalb bemerkt der Autor, dass es keine allgemeingültige Lösung zur Erklärbarkeit von KI-Systemen gibt, da dies von spezifischen Anforderungen und Kontexten abhängt. CARVALHO ET AL. (2019) betonen zusätzlich, dass XAI-Methoden auch die Entwickler unterstützen können, da sie dazu beitragen können, Modelle zu verbessern und zu optimieren. Die Autoren merken an, dass die genutzten Methoden nicht nur auf die Erklärung von Entscheidungen beschränkt sind, sondern auch auf die Erklärung von Vorhersagen und Empfehlungen ausgeweitet werden sollten. Außerdem weisen CARVALHO ET AL. (2019) auf die Bedeutung des inhaltlichen Kontexts bei der Bewertung genutzter Erklärbarkeitsmethoden hin. Dies verdeutlicht, dass XAI-Methoden nicht kontextunabhängig entwickelt und beurteilt werden sollten. LOYOLA-GONZALEZ (2019) beschreibt White-Box-Modelle als solche, deren interne Logik und Funktionsweise für den Menschen nachvollziehbar und interpretierbar ist, während Black-Box-Modelle aufgrund ihrer Komplexität weniger transparent in Bezug auf ihre interne Logik sind. Als Beispiele werden Entscheidungsbäume als Form des White-Box-Modells und neuronale Netze als Form des Black-Box-Modells genannt. Es wird spezifiziert, dass zwar sowohl Black-Box als auch White-Box-Ansätze valide Optionen sein können, um Ergebnisse zu generieren, jedoch die Notwendigkeit besteht, die Eingangsdaten, das eigentlich zu lösende Problem und die Methodik zur Darstellung der Ausgangsdaten zu verstehen, bevor ein Modell

angewandt wird. Der Autor betont, dass Erklärungen basierend auf einem White-Box-Modell von Experten im Anwendungsbereich validiert werden, um sicherzustellen, dass sie für praktische Probleme geeignet sind. Gleichzeitig verweist er darauf, dass die Anwendung von Black-Box-Modellen ebenfalls berechtigt sein kann, dies jedoch vom spezifischen Anwendungsfall abhängig gemacht werden muss. Die Möglichkeit, eine Kombination aus Black- und White-Box-Modellen zu verwenden, besteht ebenfalls (Loyola-Gonzalez, 2019). Verschiedene Studien haben sich aus diesen Gründen mit Methoden zur Verbesserung der Erklärbarkeit von Black-Box-Modellen beschäftigt.

ISLAM ET AL. (2019) schlagen vor, Domänenwissen aus der Literatur zu sammeln und zu nutzen, um die Interpretierbarkeit und Erklärbarkeit von Modellen zu verbessern. Die Autoren diskutieren auch verschiedene Erklärbarkeitsmethoden wie Layer-wise Relevance Propagation (LRP) und Tree Interpreter und vergleichen die Leistung von Modellen, die auf den ursprünglichen Merkmalen und den häufigen Merkmalssets basieren (Islam et al., 2019). Dies ist besonders relevant, da im Rahmen des in dieser Arbeit entwickelten Dashboards auf relative Finanzkennzahlen gesetzt wird, was als Form des Domänenwissens gesehen werden kann. Das folgend genutzte Modell XGBoost agiert als Tree Interpreter, indem es auf einem Ensemble von Entscheidungsbäumen basiert und bei der Vorhersage den Beitrag jedes Merkmals im Baum quantifiziert, wodurch eine detaillierte Interpretation und Erklärbarkeit des Modells gewährleistet wird (Chen & Guestrin, 2016).

CHOU (2019) schlägt einen hybriden Ansatz vor, der Entscheidungsbaum- und Deep-Learning-Modelle kombiniert, um eine ausgewogene Lösung zwischen Vorhersagegenauigkeit und Erklärbarkeit zu bieten. Dieser Ansatz beinhaltet die Verknüpfung eines Deep-Learning-Modells mit ausgewählten Zweigen eines Entscheidungsbaums, die während des Modelltrainings eine schlechte Vorhersagegenauigkeit aufweisen.

In einer weiteren Studie werden statistische Methoden und künstliche Intelligenz-Algorithmen zur Vorhersage von Unternehmensinsolvenzen verwendet. Um die Erklärbarkeit der Vorhersagemodelle zu erhöhen, werden verschiedene Erklärbarkeitsmethoden angewendet, wie zum Beispiel die partielle Abhängigkeitsanalyse (PDP) und Shapley Additive Explanations (SHAP) (Park et al.,

2021). PARK ET AL. (2021) untersuchen ebenfalls die Erklärbarkeit von Machine-Learning-Modellen zur Vorhersage von Unternehmensinsolvenzen und wenden verschiedene Erklärbarkeitsmethoden wie LIME und SHAP an. Die Ergebnisse zeigen, dass die Erklärbarkeitsmethoden dazu beitragen können, die Interpretierbarkeit von Black-Box-Modellen zu verbessern und wichtige Merkmale für die Vorhersage von Insolvenzen zu identifizieren. Die Autoren betonen die Bedeutung ethischer Gesichtspunkte, sowie von Fairness und Transparenz bei Entscheidungen im Zusammenhang mit Kreditvergabe und Insolvenzvorhersage.

XAI-Methoden können laut LI ET AL. (2022) lokale Ersatzmodelle generieren, Shapley-Werte nutzen oder Gradienten integrieren, um Merkmalswichtigkeit zu attribuieren. Die benannten Shapley-Werte werden folgend in Unterkapitel 2.3.3 kurz aufgegriffen, um zu verdeutlichen, welche Rolle sie bei der Interpretierbarkeit von Insolvenzvorhersagemodellen spielen können.

Bei der Entwicklung von KI-basierten Insolvenzprognosemodellen ergibt sich folglich die Herausforderung, wie die Funktionsweise und Entscheidungen dieser Modelle für den Endnutzer nachvollziehbar und erklärbar gemacht werden können. Hierfür lassen sich zwei wesentliche Ziele definieren: Zum einen soll transparent aufgezeigt werden, wie zuverlässig und treffsicher das Modell zwischen solventen und insolventen Unternehmen unterscheiden kann. Zum anderen soll erläutert werden, auf Basis welcher Faktoren und Merkmale das Modell seine Prognosen trifft.

Zur Beurteilung der Modellqualität und Modelleistung eignen sich beispielsweise Kennzahlen und Darstellungsformen, die die Vorhersagegenauigkeit quantitativ erfassen. Hierzu zählen Konfusionsmatrizen (vgl. Kapitel 2.3.1), welche die richtigen und falschen Prognosen tabellarisch gegenüberstellen. Daraus lassen sich wichtige Kennzahlen zur Messung der Schätzgüte ableiten, wie die Genauigkeit, die Präzision, die Sensitivität und der F1-Score. Ergänzend gibt der Brier Score Aufschluss über die Kalibrierung von Wahrscheinlichkeitsaussagen (vgl. Kapitel 2.3.2).

Methoden wie SHAP und XGBoost Feature Importance sind geeignet, um zu erklären, auf welchen Faktoren und Merkmalen das Modell seine Vorhersagen basiert. SHAP (vgl. Kapitel 2.3.3) berechnet den Einfluss jedes Merkmals auf die Differenz zwischen vorhergesagter und erwarteter Modellausgabe. Die Visualisierung der SHAP-Werte für

jedes Merkmal ermöglicht es, die Auswirkungen der Merkmalswerte auf die Modellentscheidung nachzuvollziehen. XGBoost Feature Importance (vgl. Kapitel 2.3.4) misst den Beitrag der einzelnen Merkmale zur Genauigkeit des Modells, indem es die Verbesserung der Genauigkeit durch jedes Merkmal quantifiziert. Dies hilft zu verstehen, welche Merkmale für die Insolvenzvorhersage am wichtigsten sind.

2.3.1 Konfusionsmatrix

Eine Konfusionsmatrix ist eine Tabelle, die zur Bewertung der Leistung eines ML-Algorithmus verwendet wird. Sie ist eine Matrix aus True Positives (TP), True Negatives (TN), False Positives (FP) und False Negatives (FN). Bei der binären Klassifikation ist die Konfusionsmatrix eine 2x2-Tabelle, die die Anzahl der richtigen und falschen Vorhersagen des Classifiers angibt. Die Zeilen stellen die tatsächlichen Klassenbezeichnungen und die Spalten die vorhergesagten Klassenbezeichnungen dar. True Positives sind die Fälle, in denen die tatsächliche Klasse positiv ist und die vorhergesagte Klasse ebenfalls positiv ist. True Negatives sind die Fälle, in denen die tatsächliche Klasse negativ ist und die vorhergesagte Klasse ebenfalls negativ ist. False Positives sind die Fälle, in denen die tatsächliche Klasse negativ ist, die vorhergesagte Klasse jedoch positiv ist. False Negatives sind die Fälle, in denen die tatsächliche Klasse positiv ist, die vorhergesagte Klasse jedoch negativ ist. Konfusionsmatrizen sind nützlich für die Bewertung der Leistung eines Classifiers, da sie eine detaillierte Aufschlüsselung von dessen Leistung für verschiedene Klassen bieten. Aus der Konfusionsmatrix können verschiedene Leistungsmaße wie Genauigkeit (Accuracy), Präzision (Precision), Sensitivität (Recall) und F1-Score berechnet werden (Maria Navin & Pankaja, 2016).

2.3.2 Kennzahlen zur Messung der Schätzgüte

Wie bereits in den Ausführungen zur Konfusionsmatrix erläutert, lassen sich aus den TPs, TNs, FPs und FNs verschiedene Kennzahlen berechnen. Diese Leistungsmaße beschreiben unterschiedliche Metriken zur Einordnung der Schätzgüte.

Die Accuracy, also Genauigkeit, ist die Anzahl an korrekten Vorhersagen, die ein ML-Modell getroffen hat, gemessen an der Gesamtzahl an Vorhersagen (Shalev-Shwartz & Ben-David, 2014). Formell ausgedrückt bedeutet das:

$$\text{Genauigkeit} = \frac{\text{Anzahl korrekter Vorhersagen}}{\text{Gesamtanzahl an Vorhersagen}}$$

Daraus ergibt sich bei einem ungleich verteilten Datensatz (wie dem, welcher im Folgenden für die Durchführung der Insolvenzprognose verwendet wird) jedoch, dass diese Kennzahl nur bedingt aussagekräftig ist, da die Genauigkeit dennoch sehr hoch wäre, wenn kein einziges Unternehmen als insolvent klassifiziert wird.

Die Precision, also Präzision, betrachtet die Anzahl an korrekt vorhergesagten positiven Fällen gemessen an den insgesamt vorhergesagten positiven Fällen (Shalev-Shwartz & Ben-David, 2014). Formell ausgedrückt bedeutet das:

$$\text{Präzision} = \frac{\text{Korrekte positive Vorhersagen}}{\text{Korrekte positive Vorhersagen} + \text{Inkorrekte positive Vorhersagen}}$$

Dieses Leistungsmaß befasst sich also damit, wie viele der als insolvent vorhergesagten Unternehmen tatsächlich insolvent sind.

Der Recall, also die Sensitivität, beschreibt die Anzahl an korrekt vorhergesagten positiven Ergebnissen, in diesem Zusammenhang also die Anzahl an korrekt vorhergesagten Insolvenzen gemessen an der Gesamtzahl an Insolvenzen (Shalev-Shwartz & Ben-David, 2014). Formell ausgedrückt bedeutet das:

$$\text{Sensitivität} = \frac{\text{Korrekte positive Vorhersagen}}{\text{Korrekte positive Vorhersagen} + \text{Inkorrekte negative Vorhersagen}}$$

Daraus lässt sich schließen, dass diese Kennzahl dafür geeignet ist, nachzuvollziehen, wie viele tatsächlich insolvente Unternehmen vom trainierten Modell als solche klassifiziert wurden.

Um ein harmonisches Mittel zwischen der Sensitivität und der Präzision zu finden, kann der F1-Score genutzt werden (Powers, 2007; Shalev-Shwartz & Ben-David, 2014). Dieser enthält einen gleichgewichteten relativen Beitrag von Sensitivität und Präzision und berechnet sich wie folgt:

$$F1 - \text{Score} = \frac{2 * (\text{Sensitivität} * \text{Präzision})}{\text{Sensitivität} + \text{Präzision}}$$

Dies schafft ein Gleichgewicht zwischen der Erkennung von Fehlalarmen und übersehenen Insolvenzfällen, um sicherzustellen, dass das Insolvenzprognose-Modell eine verwertbare Insolvenzprognose bereitstellt.

Der Brier Score wiederum befasst sich mit dem durchschnittlichen quadrierten Unterschied zwischen vorhergesagter Wahrscheinlichkeit und tatsächlichem Ergebnis. Er kann eine Zahl zwischen 0 und 1 annehmen, wobei ein niedrigerer Brier Score für eine bessere Vorhersage spricht, da er einen kleineren durchschnittlichen (quadrierten) Unterschied abbildet (Sawyer et al., 1950). Dies ist im Falle binärer Klassifikation, wie sie zur Insolvenzprognose genutzt wird, von Vorteil, da an dieser Stelle eine Entscheidung getroffen werden muss, um zu definieren, ab welcher Wahrscheinlichkeit eine Vorhersage als insolvent klassifiziert werden sollte.

2.3.3 SHAP - Shapley Additive Explanations

SHAP (Shapley Additive Explanations) ist ein vereinheitlichtes Framework zur Interpretation von Vorhersagen, die von komplexen Modellen wie Ensemble- oder Deep-Learning-Modellen getroffen werden. Es gibt jedem Merkmal einen Wert für eine bestimmte Vorhersage und ermöglicht es den Benutzern zu verstehen, warum ein Modell eine bestimmte Vorhersage trifft. Im Vergleich zu früheren Methoden bietet das Framework neue Ansätze, die eine verbesserte Rechenleistung und/oder eine bessere Übereinstimmung mit der menschlichen Intuition aufweisen (Lundberg et al., 2017).

SHAP vereint sechs bestehende Methoden, die sogenannte additive Merkmalsattributionsmethoden nutzen. Diese basieren auf linearen Funktionen, die aus binären Variablen bestehen. Zu diesen Methoden gehören LIME, DeepLIFT und Shapley Sampling-Werte (Lundberg et al., 2017).

SHAP Summary Plots werden verwendet, um die allgemeine Bedeutung jedes Merkmals im Ergebnis eines Modells zu visualisieren. Sie zeigen die SHAP-Werte für jedes Merkmal, die den Beitrag dieses Merkmals zur Differenz zwischen der vorhergesagten Ausgabe und der erwarteten Ausgabe darstellen. Die Merkmale werden nach ihrer Bedeutung sortiert, wobei das wichtigste Merkmal oben steht. Das Diagramm zeigt auch die Verteilung der Merkmalswerte, wobei die Farbe die Korrelation zwischen dem Merkmalswert und dem SHAP-Wert anzeigt (Lundberg et al., 2017).

SHAP Dependence Plots werden verwendet, um die Beziehung zwischen zwei Merkmalen und der vorhergesagten Ausgabe eines Modells zu visualisieren. Die Plots zeigen die SHAP-Werte für ein Merkmal A in Abhängigkeit von den Werten eines anderen Merkmals B. Das Diagramm veranschaulicht, wie sich die vorhergesagte

Modellausgabe ändert, wenn sich der Wert von Merkmal A bei unterschiedlichen Werten von Merkmal B verändert. Die SHAP-Werte geben dabei den Beitrag von Merkmal A zur Differenz zwischen der vorhergesagten Ausgabe und der erwarteten Ausgabe an. Zudem zeigt das Diagramm die Verteilung der Werte von Merkmal B, wobei die Farbe die Korrelation zwischen dem Wert von Merkmal B und dem SHAP-Wert von Merkmal A widerspiegelt. Insgesamt verdeutlichen die SHAP Dependence Plots, wie sich die Vorhersagekraft eines Merkmals A in Abhängigkeit der Ausprägungen eines anderen Merkmals B verändert (Lundberg et al., 2017).

2.3.4 XGBoost Feature Importance

XGBoost Feature Importance kann verwendet werden, um Wissen über einen Datensatz und die Ergebnisse eines Modells im Kontext der Insolvenzprognose zu gewinnen. Bei der Insolvenzprognose liegt der Schwerpunkt darauf, XGBoost Feature Importance zu verwenden, um den Zusammenhang zwischen den Merkmalen des Datensatzes und der Wahrscheinlichkeit einer vorhergesagten Insolvenz hervorzuheben. Die Feature-Importance in XGBoost kann anhand von „Gain“ gemessen werden. Gain ist die Verbesserung der Genauigkeit, die ein Merkmal für die Zweige, auf denen es sich befindet, mit sich bringt. Auf diese Weise misst Gain den Beitrag jedes Merkmals zur Genauigkeit des Modells in der Insolvenzprognose. Die Darstellung der Feature-Importance kann helfen, die wichtigsten Merkmale im Datensatz für Insolvenzprognosen zu visualisieren. Durch die Identifizierung dieser entscheidenden Merkmale kann das Modell besser verstanden und interpretiert werden (Chen & Guestrin, 2016).

3 Methodisches Vorgehen

Im Rahmen des methodischen Vorgehens wird in Kapitel 3.1 zunächst eine strukturierte Literaturrecherche basierend auf VOM BROCKE ET AL. (2015) durchgeführt. Diese hat zum Ziel, ein repräsentatives Bild der relevanten Literatur in Bezug auf die Nutzung künstlicher Intelligenz im Rahmen der Insolvenzprognose und die damit zusammenhängende Verwendung von Erklärbarkeitsmethoden zu ermöglichen, um so eine umfassende Anforderungsanalyse gestalten zu können. Kapitel 3.2 befasst sich mit dem genutzten Ansatz zur Entwicklung des Dashboards und beschreibt, welche Praktiken der Softwareentwicklung bei der Entwicklung des in dieser Arbeit vorgestellten Prototyps Anwendung finden.

3.1 Literaturreview zur Anforderungsanalyse für erklärbares ML in der Insolvenzprognose

Um eine wissenschaftlich fundierte Grundlage für die Anforderungsanalyse zur Erstellung eines Dashboards zur KI-gestützten Insolvenzprognose zu erlangen, wurde eine systematische Literaturrecherche gemäß VOM BROCKE ET AL. (2015) durchgeführt. Konkret wurde sich im Rahmen dieser Arbeit für eine repräsentative Literaturrecherche entschieden, da der Kern der Arbeit die Erstellung des Dashboards und die Implementierung geeigneter KI-Methoden, sowie ihre Erklärbarkeit sein soll und somit eine vollumfängliche Literaturrecherche den Rahmen der Arbeit deutlich ausweiten würde.

In Hinblick auf die beiden Forschungsfragen wurde sich dazu entschieden, drei zentrale Themenfelder zu berücksichtigen. Dabei handelt es sich um die **KI-gestützte Insolvenzprognose**, um **XAI** und um den **Umgang mit unausgeglichene Datensätzen**. Ziel ist es, den aktuellen Stand der Forschung in diesen Feldern hinreichend abzubilden, um auf Basis dessen eine Anforderungsanalyse durchführen zu können.

In einem ersten Schritt wurden auf Basis der Forschungsfragen und Themenfelder insgesamt 13 englischsprachige Suchstrings entwickelt. Diese zielen jeweils gezielt auf die benannten Themenfelder ab und lassen sich folgend in Tabelle 1 einsehen.

Themenfeld	Suchstring
Explainable Artificial Intelligence	Explainable Artificial Intelligence OR XAI
KI-gestützte Insolvenzprognose	Artificial Intelligence AND Insolvency Prediction
	Artificial Intelligence AND Bankruptcy Prediction
	Machine Learning AND Bankruptcy Prediction
	ML AND Bankruptcy Prediction
	ML AND Bankruptcy Forecast*
	Machine Learning AND Bankruptcy Forecast*
	Artificial Intelligence AND Bankruptcy Forecast*
	Artificial Intelligence AND Insolvency Forecast*
Umgang mit unausgeglichene Daten	Machine Learning AND Imbalanced Data
	ML AND Imbalanced Data
	Artificial Intelligence AND Imbalanced Data
	AI AND Imbalanced Data

Tabelle 1: Suchstrings nach Themenfeldern geordnet

Als Datenbanken für die Recherche wurden Web of Science und EBSCOHost ausgewählt. Diese wissenschaftlichen Publikationsdatenbanken decken ein breites Spektrum an technischen und wirtschaftswissenschaftlichen Veröffentlichungen ab und ermöglichen so eine umfassende Recherche der einschlägigen Literatur. Um die Treffermenge bereits bei der initialen Recherche einzugrenzen, wurde eine Einschränkung auf englischsprachige Open Access Veröffentlichungen der letzten 10 Jahre vorgenommen, basierend auf dem 31.05.2012 als Start- und dem 31.05.2022 als Enddatum. In der EBSCOHost Datenbank wurde zusätzlich das Kriterium gewählt, dass es sich um wissenschaftliche Journale mit Volltextzugriff handeln soll (vgl. Tabelle 2).

Suchstring	Treffer Web of Science	Treffer EBSCOHost
Explainable Artificial Intelligence OR XAI	903	90
Artificial Intelligence AND Insolvency Prediction	4	-
Artificial Intelligence AND Bankruptcy Prediction	23	14
Machine Learning AND Bankruptcy Prediction	63	23
ML AND Bankruptcy Prediction	3	8
ML AND Bankruptcy Forecast*	2	2
Machine Learning AND Bankruptcy Forecast*	11	4
Artificial Intelligence AND Bankruptcy Forecast*	2	4
Artificial Intelligence AND Insolvency Forecast*	-	-
Machine Learning AND Imbalanced Data	947	125
ML AND Imbalanced Data	126	80
Artificial Intelligence AND Imbalanced Data	192	39
AI AND Imbalanced Data	88	11

Tabelle 2: Initiale Suchergebnisse nach Datenbanken

Die große Anzahl an Treffern wurde in einem nächsten Schritt zunächst über eine Filterung in Web of Science nach Highly Cited Papers reduziert, für den Fall, dass mehr als 50 Ergebnisse und zugleich mindestens 10 Highly Cited Papers bei einem Suchbegriff vorhanden waren. Im Anschluss daran wurde eine manuelle Selektion gemäß der Titel der resultierenden Literatur vorgenommen. Die sich daraus ergebenden Trefferzahlen lassen sich in der folgenden Tabelle einsehen (vgl. Tabelle 3).

Suchstring	Treffer Web of Science	Treffer EBSCOHost
Explainable Artificial Intelligence OR XAI	45	10
Artificial Intelligence AND Insolvency Prediction	2	-
Artificial Intelligence AND Bankruptcy Prediction	20	10
Machine Learning AND Bankruptcy Prediction	44	14
ML AND Bankruptcy Prediction	1	6
ML AND Bankruptcy Forecast*	-	1
Machine Learning AND Bankruptcy Forecast*	6	2
Artificial Intelligence AND Bankruptcy Forecast*	1	2
Artificial Intelligence AND Insolvency Forecast*	-	-
Machine Learning AND Imbalanced Data	20	21
ML AND Imbalanced Data	12	10
Artificial Intelligence AND Imbalanced Data	4	6
AI AND Imbalanced Data	17	3

Tabelle 3: Vorauswahl der Suchergebnisse nach Datenbanken

Im Anschluss an die Vorselektion wurde sich dazu entschieden, die übrig gebliebene Literatur wieder basierend auf den drei Themenfeldern zu kategorisieren. An dieser Stelle wurde auf Dubletten geprüft und folgend eine Prüfung der Abstracts auf Applikabilität vorgenommen. Anhand dieser Prüfung konnte die inhaltliche Relevanz für die angestrebte Anforderungsanalyse genauer eingeschätzt werden. Die daraus resultierende repräsentative Gesamtmenge an Treffern je Themenfeld findet sich in der folgenden Tabelle (vgl. Tabelle 4).

Themenfeld	Anzahl an wissenschaftlichen Arbeiten
Explainable Artificial Intelligence	27
KI-gestützte Insolvenzprognose	53
Umgang mit unausgeglichenen Daten	52

Tabelle 4: Finale Auswahl der Literatur

Nach der inhaltlichen Analyse der Abstracts und Selektion nach Relevanz verbleiben 132 wissenschaftliche Arbeiten, aufgeteilt auf die drei vorab definierten Themenfelder, die als Basis für die folgende Anforderungsanalyse in Frage kommen. Somit konnte für jeden der drei zentralen Themenbereiche eine angemessene Anzahl relevanter Publikationen ermittelt werden. Die Recherche bietet also eine repräsentative Auswahl an relevanter Literatur. Diese bildet eine solide Basis für die weitere Inhaltsanalyse zur Ableitung

konkreter Anforderungen an das zu entwickelnde Dashboard. Durch die Beschränkung auf insgesamt 132 wissenschaftliche Quellen wurde eine dem Umfang dieser Arbeit angemessene Menge für die tiefergehende Analyse erreicht.

3.2 Hybrider Prototyping-Ansatz zur Entwicklung eines Dashboards

In Folge der Literaturrecherche zur Anforderungsanalyse sowie der folgend näher beschriebenen tatsächlich durchgeführten Anforderungsanalyse wurde ein systematischer, hybrider Prozess entworfen, um die Entwicklung eines Dashboards zur KI-gestützten Insolvenzprognose durchzuführen. Dieser Prozess kombiniert Elemente des Prototyping, eine initiale wasserfallartige Konzeptionsphase und auch iterative agile Methoden bei der Implementierung des Dashboards. Diese hybride Herangehensweise ergibt sich aus der Struktur und den Besonderheiten des Projekts selbst. Eine visuelle Darstellung des entworfenen und implementierten Prozesses findet sich in Abbildung 1.

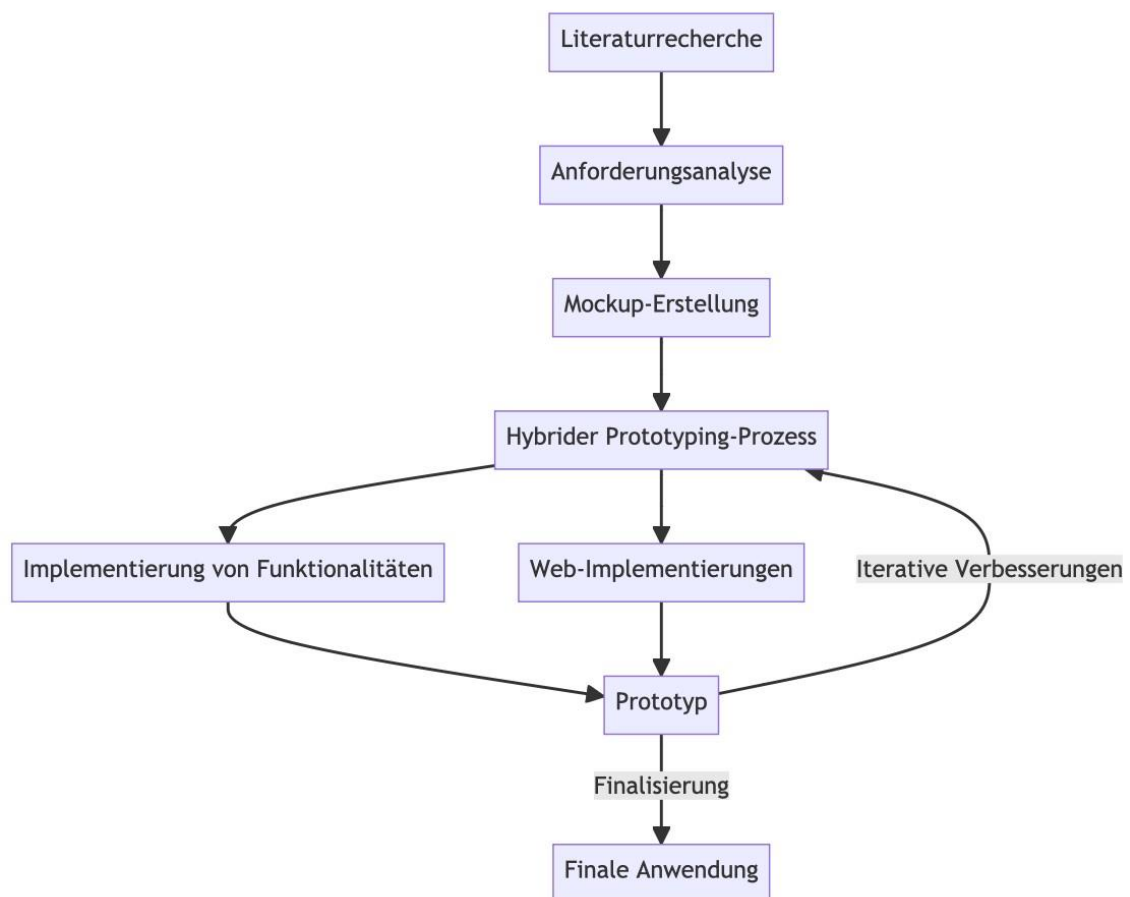


Abbildung 1: Hybrider Prototyping-Ansatz

Zunächst wurden, um auf die theoretischen Hintergründe des Dashboard-Designs einzugehen und einen visuellen Zugang zu den Anforderungen zu erhalten, Mockups erstellt. Der Prozess der Mockup-Erstellung wurde in *uizard* durchgeführt und iterativ in gegenseitiger Rücksprache verfeinert, um die gewünschten Funktionalitäten abbilden zu können. Mockups sind im Entwicklungsprozess von Bedeutung, da sie eine visuelle Darstellung der Benutzeroberfläche, des User Interface (UI), liefern, die von Nutzern und Entwicklern gleichermaßen verstanden werden kann. Durch die Verwendung von Mockups können Nutzer und Entwickler frühzeitig im Entwicklungsprozess zusammenarbeiten, um sicherzustellen, dass die Anforderungen der Nutzer erfüllt werden und dass das Endprodukt den Erwartungen entspricht. Mockups können auch dazu beitragen, Missverständnisse zwischen Endanwendern und Entwicklern zu vermeiden, da sie eine klare und einheitliche Vorstellung davon vermitteln, wie das Endprodukt aussehen wird. Darüber hinaus können Mockups als Grundlage für die Erstellung von UI-Modellen dienen, die dann in konkrete Web-Implementierungen

umgesetzt werden können (Rivero et al., 2011). Eine vertiefte Darstellung der Konzeptionsphase anhand der Mockups erfolgt in Kapitel 5.

Bei der tatsächlichen Softwareentwicklung wurde sich ebenfalls auf eine iterative Vorgehensweise in regelmäßiger Rücksprache geeinigt. Dennoch wurde die eigentliche Konzeption des Projekts eher nach einem klassischen Wasserfall-Schema aufgebaut. Somit wird von einem hybriden Ansatz ausgegangen, der für das gegebene Projekt und die Interaktion der Beteiligten nützlich erschien.

Hybride Softwareentwicklungsansätze gewinnen bereits seit Jahren an Bedeutung. Diese Ansätze kombinieren verschiedene Entwicklungsprozesse, um einen maßgeschneiderten Ansatz zu schaffen, der den spezifischen Anforderungen eines Projekts entspricht. Ein prototypbasiertes, iteratives Entwicklungsmodell, wie in dieser Arbeit verwendet, kann entsprechend für ein spezifisches Projekt als schlüssig erachtet werden (Kuhrmann et al., 2017). Im Rahmen dieser Arbeit erschien diese Vorgehensweise aufgrund der umfangreichen inhaltlichen Vorarbeiten, von der Literaturrecherche über die Anforderungsanalyse und das Design des Dashboards hinweg, vielversprechend. Die anfänglich wasserfallartige Struktur ermöglicht dabei eine solide inhaltliche Basis, bevor in der weiteren Durchführung agile Prozessschritte und die Erstellung iterativer Prototypen folgen.

Die Vorgehensweise im Anschluss an die inhaltlichen Vorarbeiten besteht in der Erstellung eines Prototyps und dessen iterativer Verbesserung, bis er die Anforderungen auf Basis der Anforderungsanalyse hinreichend erfüllt. Jede Iteration – sowohl auf Ebene einzelner Funktionalitäten als auch auf Ebene des Gesamt-Prototyps – kann als eigene Phase innerhalb des Gesamtprojekts betrachtet werden. Dies hat es ermöglicht, Probleme frühzeitig zu identifizieren und Änderungen vorzunehmen. Daraus ergeben sich mehrere Vorteile. Kurze Kommunikationswege und Feedback-Loops haben für iterative Anpassung und Verbesserung gesorgt, bevor das Projekt zu weit fortgeschritten war. Darüber hinaus konnten auch interne Anmerkungen schnell umgesetzt werden.

4 Anforderungsanalyse für erklärbares ML in der Insolvenzprognose

In diesem Kapitel werden die Anforderungen an erklärbare ML in der Insolvenzprognose analysiert und diskutiert. Dabei werden sowohl technische als auch dokumentarische Anforderungen berücksichtigt, um die Nachvollziehbarkeit und Transparenz von ML-Modellen zu gewährleisten. Darüber hinaus werden die Anforderungen an die Datenerhebung zur Entwicklung und Validierung von ML-Modellen in der Insolvenzprognose behandelt.

4.1 Anforderungsanalyse zur Förderung der Erklärbarkeit ML-basierter Insolvenzprognose

In dieser Anforderungsanalyse werden die identifizierten Anforderungen an ein Modellbildungs- und Validierungssystem mittels künstlicher Intelligenz dargestellt und diskutiert. Dabei wird insbesondere auf die relevanten rechtlichen und institutionellen Anforderungen eingegangen.

Die erste **Anforderung (A.1)** betrifft die Erstellung eines visuell-interaktiven Dashboards. Gemäß Art. 14 des noch nicht in Kraft getretenen Artificial Intelligence Act (AIA) der Europäischen Kommission und dem KI-Prüfkatalog des Fraunhofer IAIS soll das System so gestaltet sein, dass der gesamte Prozess des Trainings und der Validierung eines Modells über eine Benutzeroberfläche gesteuert werden kann. Dies soll es den Nutzern ermöglichen, den Prozess vollständig zu definieren und zu überwachen, wobei die Dateneingabe über einen Datenupload erfolgen soll. Wie auch von ISLAM ET AL. (2022) beschrieben, sind Erklärbarkeit und Interpretierbarkeit von KI-Systemen von entscheidender Bedeutung, um das Vertrauen der Nutzer in solche Systeme zu stärken. Insbesondere interaktive Dashboards, die den Trainingsprozess visualisieren und erklären, könnten dies unterstützen (Islam et al., 2022). Laut LINARDATOS ET AL. (2021) ist es wichtig, Erklärungen für die jeweiligen Nutzer anzupassen, da das Verständnis von Erklärungen vom Expertenwissen der Nutzer abhängt. RAI (2020) verweist zusätzlich darauf, dass das Verständnis der Erklärbarkeit und Transparenz mit den Bedürfnissen verschiedener Nutzer übereinstimmen muss und auch KIM UND ROUTLEDGE (2022) spezifizieren, dass Dashboards durch unterschiedliche Erklärungsstufen helfen,

Bedürfnisse verschiedener Benutzergruppen zu berücksichtigen. Daher sollte ein Dashboard so gestaltet werden, dass es Erklärungen auf unterschiedlichen Komplexitätsstufen bietet. Laut ABEDIN (2022) sollten Erklärungen kontrastiv, selektiv und wahrscheinlichkeitsunabhängig sein, was zusätzlich im Dashboarddesign berücksichtigt werden sollte. Die europäische Kommission verweist darauf, dass KI-Systeme sowohl eigenständig als auch als Bestandteil einer Software genutzt werden können. Sie priorisiert eine sichere Erprobungsumgebung und geeignete Schutzvorkehrungen und Risikominderungsmaßnahmen. Dies verdeutlicht, dass ein visuell-interaktives Dashboard sowohl als eigenständiges System konzipiert werden kann, als auch sicher und rechtskonform ausgestaltet werden muss. Ebenso wird die Datenqualität bei der Verwendung von KI-Systemen hervorgehoben, was verdeutlicht, dass in Form eines standardisierten Datenuploads die Datenstruktur konsistent implementiert werden und eine entsprechende Datenbereinigung hervorgehoben werden muss (European Commission, 2021).

Die zweite **Anforderung (A.2)** bezieht sich auf die Auswahl und das Training der KI-Methode. Gemäß den Anforderungen der BaFin von 2021, dem AIA Art. 10 Nr. 1 und dem KI-Prüfkatalog des Fraunhofer IAIS muss die Modellauswahl umfassend begründet und dokumentiert werden. Dabei sollen sowohl intrinsisch interpretierbare Modelle als auch Black-Box-Modelle verwendet werden können. Bei der Modellauswahl sollten Ansätze priorisiert werden, die das Datenungleichgewicht mindern. Beispielsweise werden von KAUR ET AL. (2019) hybride Ansätze aus Under- und Oversampling vorgeschlagen. Wie in der Arbeit von ISLAM ET AL. (2022) diskutiert, gibt es unterschiedliche mögliche Ansätze, um die Erklärbarkeit von KI-Modellen in den Vordergrund zu rücken. Dazu gehören sowohl ante-hoc als auch post-hoc Methoden. Bei ante-hoc-Methoden werden die Erklärungen während des Trainings generiert, während bei post-hoc-Methoden die Erklärungen im Nachhinein für ein bereits trainiertes Modell generiert werden. Post-hoc Methoden werden häufig für komplexe Modelle wie neurale Netze verwendet. Dies sollte bei der Modellauswahl berücksichtigt werden. Laut LINARDATOS ET AL. (2021) gibt es ein breites Spektrum an Methoden zur Förderung der Erklärbarkeit von KI-Modellen, die bei der Modellauswahl in Betracht gezogen werden sollten. RAI (2020) betont, dass die Wahl zwischen inhärent interpretierbaren und Blackbox-Modellen mit Post-hoc-Erklärungen von der

Anwendung abhängt. Gerade bei der Verwendung letzterer ist es wichtig, sie mit Post-hoc-Erklärungen zu versehen, um ihre Entscheidungen und Vorhersagen erklären zu können. Gleichzeitig verweist RAI (2020) darauf, dass inhärent interpretierbare Modelle je nach Anwendungsfall bevorzugt werden können, da sie von Natur aus leichter zu verstehen sind. KIM ET AL. (2020) betonen zusätzlich, dass die Interpretierbarkeit der Modelle berücksichtigt werden muss, nicht nur die Genauigkeit. Die Modelle müssen mit einem geeigneten Datensatz trainiert und getestet werden und die Speicherung von Modell- und Trainingsparametern muss zur besseren Nachvollziehbarkeit ermöglicht werden. Gemäß den Vorgaben des europäischen Parlaments sollen für sogenannte Hochrisiko-KI-Systeme Transparenzvorschriften definiert werden, um Undurchsichtigkeit entgegenzuwirken. Nutzer sollen dazu in der Lage sein, die Ergebnisse des Systems zu interpretieren und es entsprechend angemessen verwenden zu können. Dies bedingt die Notwendigkeit, eindeutige Informationen beizufügen, welche die umfassende Dokumentation der Modellauswahl und die Speicherung von Modell- und Trainingsparametern zur besseren Nachvollziehbarkeit beinhalten können (European Commission, 2021). Die theoretische Grundlage des genutzten KI-Modells muss gemäß KI-Prüfkatalog ebenfalls betont werden, da sichergestellt werden soll, dass die Algorithmen und Modelle für den vorhergesehenen Zweck geeignet sind. Unterschiedliche Modelle sollten verglichen werden, um sicherzustellen, dass das beste Modell für den entsprechenden Zweck ausgewählt wurde. Ebenso soll die Modellauswahl begründet werden (Fraunhofer IAIS, 2021).

Die dritte **Anforderung (A.3)** bezieht sich auf das Testen der KI-Algorithmen. Gemäß AIA Art. 10 Nr. 1 und Art. 13 Nr. 3 lit. b) (ii) sind die trainierten Modelle mit einem separaten Testdatensatz auf Genauigkeit und Konsistenz zu prüfen. Gemäß KAURE ET AL. (2019) ist es vorzuziehen, eine Kreuzvalidierung durchzuführen. Da das europäische Parlament ebenfalls darauf verweist, dass die einschlägige Dokumentation des Systems gegebenenfalls in Bezug auf mögliche Risiken für Grundrechte und Diskriminierung ausgestaltet werden sollte, ist die Dokumentation und Transparentmachung von Testergebnissen der KI-Algorithmen von Bedeutung. LINARDATOS ET AL. (2021) verweisen in ihrer Arbeit darauf, dass die Evaluierung von KI-Systemen ein wichtiges aber noch wenig erforschtes Gebiet darstellt. Die Autoren merken an, dass ein Mangel an etablierten Metriken zur Bewertung der Erklärbarkeit besteht. Dies gilt es bei Wahl der

Testverfahren zu beachten. Gemäß KIM UND ROUTLEDGE (2022) sind Testdaten wichtig, um Überanpassung zu vermeiden. KIM ET AL. (2020) betonen, dass im Testdatensatz möglichst viele Insolvenzfälle enthalten und dieser so repräsentativ wie möglich sein sollte. Weiterhin wird an dieser Stelle aufgegriffen, dass die genutzten Trainings, Validierungs- und Testdatensätze in Hinblick auf die Zweckbestimmung des Systems hinreichend relevant, repräsentativ, fehlerfrei und vollständig sein sollen. Dies legt nahe, dass die entsprechenden Datensätze so gestaltet werden müssen, dass sie repräsentativ für die Zweckbestimmung des Systems sind. Darüber hinaus wird die Bedeutung hoher Datenqualität hervorgehoben (European Commission, 2021). Hierbei muss laut Fraunhofer IAIS darauf geachtet werden, dass die Testdaten unabhängig von den Trainingsdaten sind, um Überanpassungen des Modells zu vermeiden. Die Testergebnisse sollen dokumentiert werden (Fraunhofer IAIS, 2021).

Die vierte **Anforderung (A.4)** betrifft die Protokollierung und Speicherung der Ergebnisse der Modelle und Transparenzmethoden. Gemäß dem KI-Prüfkatalog des Fraunhofer IAIS sollen Modellergebnisse und Transparenzmethoden zur besseren Nachvollziehbarkeit und Reproduzierbarkeit gespeichert werden können, um sie bei Bedarf später überprüfen oder auch reproduzieren zu können. ISLAM ET AL. (2022) betonen, dass die Evaluierung von KI-Systemen und deren Erklärungen ein wichtiger Aspekt im Zusammenhang mit XAI sind. Hierfür muss sichergestellt werden, dass auch dokumentiert wird, wie genau die Ergebnisse erzielt wurden. Dies legt eine Speicherung der zugehörigen Parameter nahe (Fraunhofer IAIS, 2021). Wie in Hinblick auf die vorangegangenen Anforderungen bereits thematisiert, inkludiert auch das europäische Parlament durch seine Anforderungen an die Dokumentation des Systems die Speicherung von Modell- und Trainingsparametern zur besseren Nachvollziehbarkeit und befasst sich ebenfalls mit der Notwendigkeit für Transparenz, welche diese Speicherung und Protokollierung essenziell macht (European Commission, 2021). Auch RAI (2020) und ABEDIN (2022) betonen, dass die Speicherung von Ergebnissen und Parametern bei der Entwicklung erklärbarer KI ein hohes Maß an Bedeutung einnimmt, da so die Ergebnisse und Erklärungen verlässlich reproduziert werden können. RAI (2020) benennt hier auch explizit regulatorische und ethische Zwecke als Grundlage für die Bedeutung der Ergebnis- und Parameterspeicherung (Rai, 2020).

Die fünfte **Anforderung (A.5)** umfasst differenzierte Bewertungsgrößen für die Schätzgüte der Ergebnisse. Gemäß AIA Art. 10, Art. 13 Nr. 3 lit. b), Art. 15 Nr. 1-2 ist der Anwender über die relevanten Eigenschaften und Leistungswerte der Ergebnisse zu informieren, darunter z.B. Genauigkeit (Accuracy), Brier Score, ROC-AUC, F1-Score, Sensitivität (Recall). Auch ISLAM ET AL. (2022) verweisen darauf, dass die objektive Bewertung von KI-Systemen und ihren Erklärungen beispielsweise durch Stabilitäts- und Treffgenauigkeitsmaße ermöglicht werden kann. LINARDATOS ET AL. (2021) schreiben weiterhin, dass es einen Mangel an etablierten Metriken in diesem Zusammenhang gibt. RAI (2020) thematisiert in diesem Zusammenhang, dass Metriken zur Erklärbarkeit anhand ihrer Verständlichkeit und Nützlichkeit bewertet werden sollten. Hierzu sollen verbindliche Anforderungen an Robustheit und Genauigkeit an das System gestellt werden. Dies deutet auf die Notwendigkeit hin, Bewertungsgrößen für die Schätzgüte des Systems differenziert abzubilden. Darüber hinaus sollten diese Kenngrößen zu Zwecken der Dokumentation und Transparenz ebenfalls in einem derartigen System dargelegt werden (European Commission, 2021). Der Prüfkatalog des Fraunhofer IAIS verdeutlicht darüber hinaus, dass bei der Entwicklung eines Systems sichergestellt werden sollte, dass es über Bewertungsgrößen verfügt, die eine differenzierte Einschätzung zur Schätzgüte des Ergebnisses ermöglichen. Hierzu sollen die Kenngrößen verwendet werden, die am besten geeignet sind, um die Leistung des Systems in Bezug auf die zugrundeliegende Domäne, hier also die Insolvenzprognose, zu bewerten. Diese sollten dokumentiert und gespeichert werden können und auch ihre Berechnung und Bedeutung sollten dargelegt werden (Fraunhofer IAIS, 2021).

Die sechste **Anforderung (A.6)** betrifft die Dokumentation der Systemprozesse. Gemäß BaFin 2021, AIA Art. 11 und 13 sowie dem KI-Prüfkatalog des Fraunhofer IAIS ist eine umfassende Dokumentation erforderlich, um eine Überprüfung der Algorithmen und der zugrundeliegenden Modelle durch die externe Aufsicht zu ermöglichen. Wie in dem Paper von ISLAM ET AL. (2022) betont, ist die Evaluierung und Dokumentation von KI-Systemen und deren Erklärungen ein wichtiger Aspekt für die Nachvollziehbarkeit. Insbesondere die Datenauswahl und -aufbereitung, die Modellauswahl, -kalibrierung und -validierung sowie die Wahl der Transparenzverfahren sind detailliert zu dokumentieren und zu begründen. Der gesamte Prognoseprozess sollte im Dashboard nachvollziehbar dargestellt werden. Dies bedeutet, dass die Nutzer in der Lage sein

sollten, den Prozess und die verwendeten Methoden zu verstehen, um das Vertrauen in das System zu stärken und mögliche Fehler oder Missverständnisse zu vermeiden. Darüber hinaus sollte die Dokumentation regelmäßig aktualisiert werden (Fraunhofer IAIS, 2021). Laut europäischer Kommission soll die technische Dokumentation stets auf dem neusten Stand gehalten werden und die Merkmale, Fähigkeiten und Grenzen des Systems sowie die verwendeten Algorithmen und Methoden inkludiert werden sollen. Auch die Dokumentation des entsprechenden Systems soll hervorgehoben und aktuell gehalten werden. Dies legt eine umfassende Dokumentation der Systemprozesse nahe, die eine Überprüfung der Algorithmen und Modelle durch externe Aufsichten zu ermöglichen. Ebenso sollen auch die Endnutzer des Systems auf hinreichende Art und Weise über dessen Transparenz aufgeklärt werden, um seine Ergebnisse angemessen interpretieren und verwenden zu können. Es soll also eine Dokumentation geschaffen werden, die sowohl externen Aufsichtsinstitutionen als auch Nutzern angemessene und möglichst detaillierte Einblicke in die Funktionalitäten des Systems bietet (European Commission, 2021). RAI (2020) betont ebenfalls, dass die Dokumentation von KI-Systemen Details zu Daten, Modellen und Erklärungen enthalten soll. So sollen einerseits die Datenquellen, andererseits aber auch die Art der Daten, die durchgeführten Verarbeitungsschritte und die Vorverarbeitung dokumentiert werden. Die Hyperparameter und Modelldetails sowie die Art der Erklärungsmethoden und ihre Verständlichkeit werden an dieser Stelle zusätzlich betont. Dies soll Entwicklern und Nutzern ermöglichen, Modelle besser zu verstehen und Erklärungen besser interpretieren zu können (Rai, 2020). KAUR ET AL. (2019) gehen darauf ein, dass die Dokumentation eine Beschreibung der Datenaufbereitung beinhalten sollte, um die Nachvollziehbarkeit des Prozesses zu erhöhen.

Schließlich konzentriert sich die siebte **Anforderung (A.7)** auf die Implementierung von Methoden zur Förderung der Interpretierbarkeit von Black-Box-Modellen. Gemäß AIA Art. 14 Nr. 4 lit. c) und Art. 13 Nr. 3 lit. d) sowie dem KI-Prüfkatalog des Fraunhofer IAIS sind Methoden zu implementieren, die dazu dienen, Erklärungen zu generieren oder die Interpretierbarkeit der Modelle zu verbessern. Sowohl LINARDATOS ET AL. (2021) als auch RAI (2020) verweisen darauf, dass es verschiedenste Methoden zur Erhöhung der Interpretierbarkeit von KI-Modellen gibt (Linardatos et al., 2021; Rai, 2020). Es wird empfohlen, dass diese Methoden modellagnostisch, d.h. unabhängig vom spezifischen

Modelltyp anwendbar sind. Die Modelle sollen so interpretierbar wie möglich gemacht werden, um die Entscheidungen des Systems nachvollziehbar zu gestalten (Fraunhofer IAIS, 2021). Ebenso wie auch die Dokumentation die Interpretationsmöglichkeiten und Anwendungsfelder des Nutzers fördern soll, sind also auch solche Methoden von Bedeutung, die es ermöglichen, Vorhersageergebnisse einfacher zu interpretieren und erklärbar zu machen (European Commission, 2021). Schließlich verweist das Fraunhofer IAIS noch darauf, dass die Dokumentation der Methoden zur Förderung der Erklärbarkeit so verständlich und umfassend ist, dass sie die Nutzerakzeptanz erhöht und die Überprüfung durch externe Aufsichten ermöglicht (Fraunhofer IAIS, 2021).

Zusammenfassend lassen sich die Anforderungen in zwei Hauptbereiche unterteilen: Zum einen die technischen Anforderungen, die das Design und die Implementierung des Systems betreffen und zum anderen die dokumentarischen Anforderungen, die auf eine hohe Nachvollziehbarkeit und Transparenz des Systems abzielen. Dabei ist zu beachten, dass die Einhaltung der genannten Anforderungen nicht nur der Erfüllung rechtlicher und institutioneller Vorgaben dient, sondern auch das Vertrauen der Nutzer in das System stärkt und die Qualität und Verlässlichkeit der Modellergebnisse sicherstellt (BaFin, 2021; European Commission, 2021; Fraunhofer IAIS, 2021).

Eine kompakte Auflistung der Anforderungen zum Zweck der besseren Übersichtlichkeit findet sich in Tabelle 5.

Nr.	Anforderung	Quellen
A.1	Erstellung eines visuell-interaktiven Dashboards	AIA Art. 14, KI-Prüfkatalog des Fraunhofer IAIS, Islam et al. (2022), Linardatos et al. (2021), Rai (2020), Kim und Routledge (2022), Abedin (2022)
A.2	Auswahl und Training der KI-Methode	BaFin 2021, AIA Art. 10 Nr. 1, KI-Prüfkatalog des Fraunhofer IAIS, Kaur et al. (2019), Islam et al. (2022), Linardatos et al. (2021), Rai (2020), Kim et al. (2020)
A.3	Testen der KI-Algorithmen	AIA Art. 10 Nr. 1, Art. 13 Nr. 3 Bstb. b) - (ii), Kaur et al. (2019), Linardatos et al. (2021), Kim und Routledge (2022) Kim et al. (2020) KI-Prüfkatalog des Fraunhofer IAIS
A.4	Protokollierung und Speicherung von Ergebnissen der Modelle und Transparenzmethoden	KI-Prüfkatalog des Fraunhofer IAIS, Islam et al. (2022), Rai (2020), Abedin (2022)
A.5	Differenzierte Beurteilungsgrößen für die Schätzgüte des Ergebnisses	AIA Art. 10 Art. 13 Nr. 3 Bstb. b), Art. 15, Nr. 1-2, Islam et al. (2022), Linardatos et al. (2021), Rai (2020), KI-Prüfkatalog des Fraunhofer IAIS
A.6	Dokumentation der Systemprozesse	BaFin 2021, AIA Art. 11 u. 13, KI-Prüfkatalog des Fraunhofer IAIS, Islam et al. (2022), Rai (2020), Kaur et al. (2019)
A.7	Implementierung von Interpretierbarkeitsmethoden die die Insolvenzprognose nachvollziehbar für ihren Benutzer macht	AIA Art. 14 Nr. 4 Bstb. c), Art 13 Nr. 3 Bstb d), KI-Prüfkatalog des Fraunhofer IAIS, Linardatos et al. (2021), Rai (2020)

Tabelle 5: Auflistung der Anforderungen

4.2 Anforderungen an Datenerhebung

Hinsichtlich der zu verwendenden Daten haben wir uns auf deutsche Unternehmensdaten konzentriert. Um Zugang zu einem großen Datensatz zu erhalten, wurde die Datenbank Amadeus verwendet (Amadeus, 2022). Diese enthält große Mengen an Finanzdaten von Millionen deutscher Unternehmen unterschiedlicher Größe. Um ein genaues Bild der Faktoren zu erhalten, die Insolvenzen beeinflussen, haben wir uns entschieden, Daten für die Jahre 2012 bis 2021 zu extrahieren. Als

Informationen zu den Unternehmen wurden folgende Daten spezifisch abgefragt: BvD ID-Nummer, Status, Status Datum, Anlagevermögen, Umlaufvermögen, Bilanzsumme, Eigenkapital, kurzfristige Verbindlichkeiten und Jahresüberschuss bzw. -fehlbetrag. „BvD ID Nummer“ steht für eine ID-Nummer, die jedes Unternehmen eindeutig identifiziert. „Status“ und „Status Datum“ wurden gewählt, um mögliche Insolvenzverfahren und deren Datum zuzuordnen.

5 Konzeption eines Dashboards zur KI-gestützten Insolvenzprognose

Basierend auf dem in Abschnitt 3.2 vorgestellten Ansatz zur Entwicklung eines Dashboards wurde ein konzeptionelles Modell der Benutzeroberfläche entwickelt. Durch die Verwendung von Mockups konnten die Projektanforderungen visuell dargestellt und gleichzeitig eine Vorstellung davon vermittelt werden, wie die endgültige Benutzeroberfläche aussehen würde.

Dies hat uns geholfen, eine bessere Vorstellung davon zu bekommen, wie die Benutzerschnittstelle gestaltet sein sollte, um die spezifischen Anforderungen und Erwartungen der Benutzer zu erfüllen. Prototyping ermöglicht es im späteren Prozess, die Benutzererfahrung in einem frühen Stadium des Entwicklungsprozesses zu optimieren und sicherzustellen, dass das Endprodukt den Anforderungen der Benutzer entspricht.

Die erstellten Mockup-Seiten decken verschiedene Aspekte der KI-gestützten Insolvenzprognose ab. Sie enthalten unter anderem Seiten zur Darstellung der Datenbasis von Modellen und Datensätzen, zur Auswahl von Merkmalen und zur Segmentierung. Weitere Seiten bieten deskriptive Informationen zu den Gruppen, Optionen zur Selektion des Oversamplings, zur Wahl der KI-Methode und zur Konfiguration der Hyperparameter. Schließlich werden auch Seiten zur Darstellung von Kennzahlen zur Messung der Schätzgüte und der Bedeutung von Merkmalen bereitgestellt, die die Erklärbarkeit der Modelle unterstützen sollen.

Jede dieser Seiten wurde speziell entworfen, um den Nutzern klare, leicht verständliche und zugängliche Informationen zu bieten. Sie sind so gestaltet, dass die Komplexität der zugrunde liegenden Daten und Modelle intuitiv und benutzerfreundlich dargestellt wird.

Die genauen Darstellungen und Layouts dieser Seiten werden weiter unten in diesem Abschnitt beschrieben und aus den Anforderungen des vorhergehenden Kapitels abgeleitet (vgl. Tabelle 6). Es ist zu beachten, dass die tatsächlichen Darstellungen im finalen Dashboard von den in den Mockups gezeigten Versionen abweichen können, da der iterative Prozess des Prototypings eine ständige Anpassung und Optimierung

erlaubt. Die für dieses Kapitel als nicht relevant eingestuft übrigen Mockup-Seiten befinden sich im Anhang.

Nr.	Anforderung	Abbildungs-Nr.
A.1	Erstellung eines visuell-interaktiven Dashboards	2-10
A.2	Auswahl und Training der KI-Methode	7
A.3	Testen der KI-Algorithmen	8
A.4	Protokollierung und Speicherung von Ergebnissen der Modelle und Transparenzmethoden	2
A.5	Differenzierte Beurteilungsgrößen für die Schätzgüte des Ergebnisses	9
A.6	Dokumentation der Systemprozesse	2-10
A.7	Implementierung von Interpretierbarkeitmethoden die die Insolvenzprognose nachvollziehbar für ihren Benutzer macht	10

Tabelle 6: Abbildungen der Mockups nach Anforderungen sortiert

In Anlehnung an A.4 ist es entscheidend, dem Nutzer die Möglichkeit zu geben, seine bereits hochgeladenen Datensätze und Modelle einzusehen und auf diese übersichtlich zuzugreifen, so dass eine weitere Bearbeitung der Datensätze oder Modelle jederzeit möglich ist. Darüber hinaus sollen auch statistische Informationen zu den Datensätzen abrufbar sein (vgl. Abbildung 2).

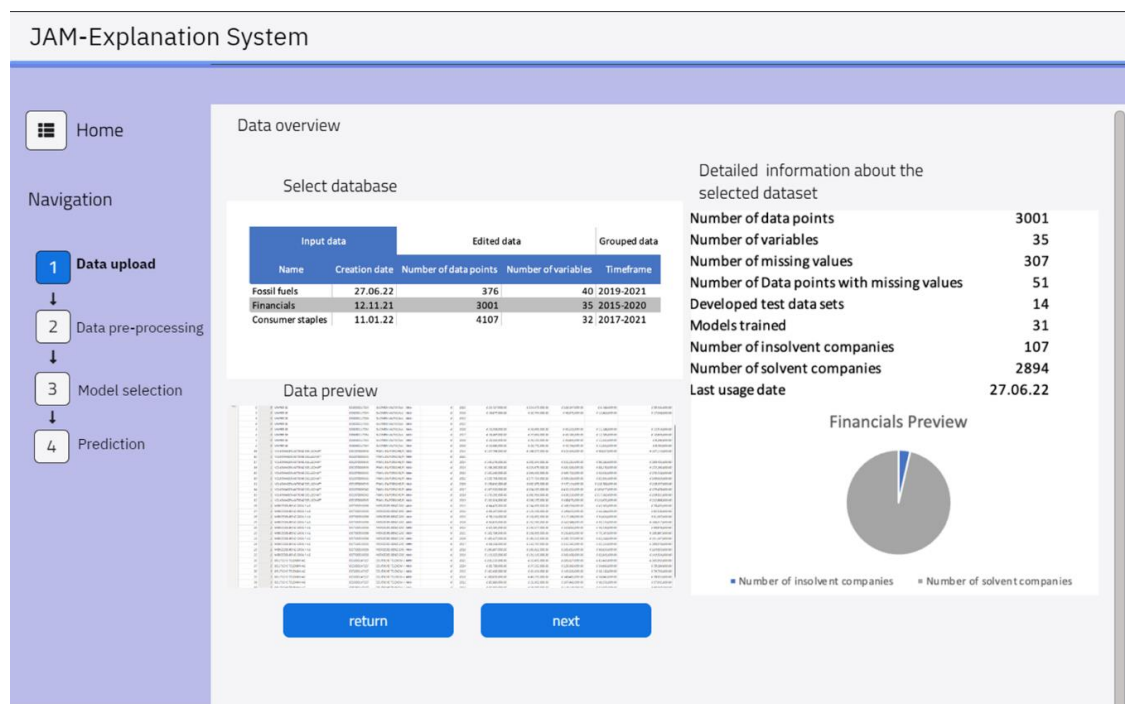


Abbildung 2: Mockup - Datenbank für Datensätze und Modelle

Der Benutzer soll in der Lage sein, relevante Merkmale, die als Grundlage für die Prognose dienen, manuell auszuwählen und auch eigene Merkmale zu erstellen (vgl. Abbildung 3).

The mockup shows the 'JAM-Explanation System' interface. On the left is a navigation sidebar with a 'Home' icon and a vertical list of steps: 1 Data upload, 2 Data pre-processing (highlighted), 3 Model selection, and 4 Prediction. The main area is titled 'Data pre-processing' and is divided into two panels. The left panel, 'Select relevant feature', contains eight input fields labeled 'Feature 1' through 'Feature 8'. 'Feature 4', 'Feature 7', and 'Feature 8' are highlighted in blue. A '+' button is at the bottom right of this panel. The right panel, 'Create own features', has a radio button for 'Add a new measure' which is selected. Below it is a 'Formula:' input field containing $5 + Var1 * Var2 - Var3$. There are three 'Set column' buttons for 'Var1', 'Var2', and 'Var3'. Below these is a 'Store and name measure' button. At the bottom of this panel are 'Add measure' and 'Remove measure' buttons. To the right of these panels is a section titled 'Additional features' with a text input field labeled 'PEG-Ratio'. At the bottom of the main area are 'return' and 'next' buttons.

Abbildung 3: Mockup - Feature-Auswahl

Eine optionale Segmentierung der Daten in Gruppen ist ebenfalls Teil des Konzepts. Hier sollen Grenzwerte für Gruppen gesetzt werden können, die sich z.B. an der Bilanzsumme der Unternehmen orientieren (vgl. Abbildung 4).

Abbildung 4: Mockup - Segmentierung der Daten

Darüber hinaus werden statistische Informationen zu den einzelnen segmentierten Datensätzen integriert (vgl. Abbildung 5).

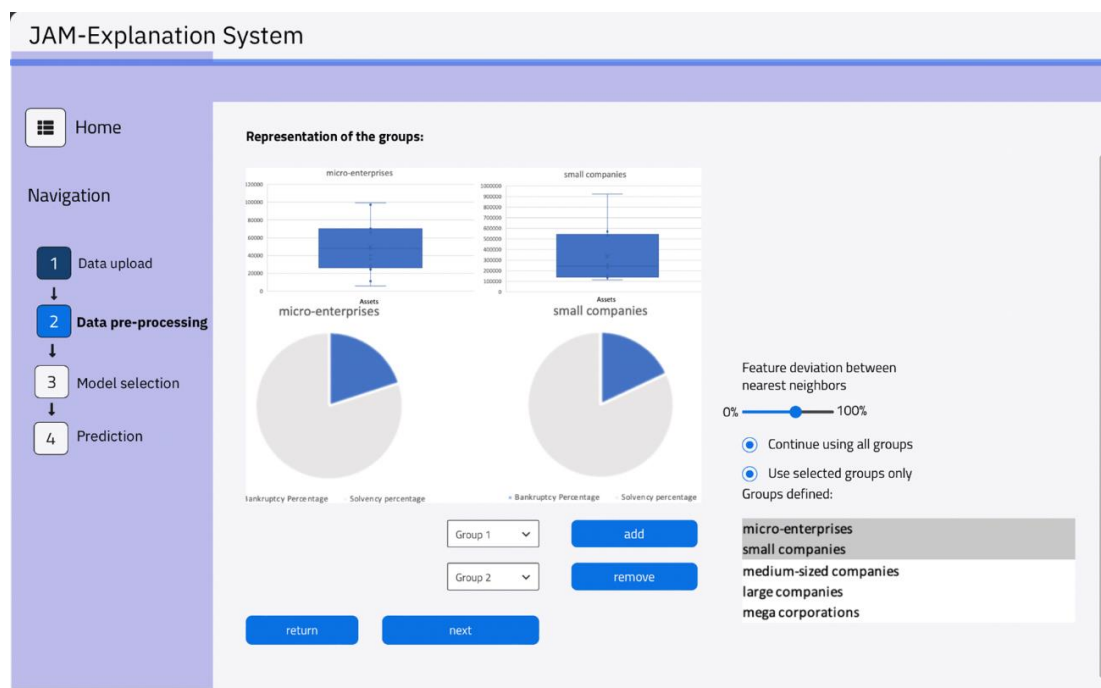


Abbildung 5: Mockup - Deskriptive Auswertung der segmentierten Daten

Im weiteren Verlauf soll der Nutzer die Möglichkeit haben, Oversampling oder Undersampling Techniken einzusetzen, um möglichen Ungleichgewichten im Datensatz

entgegenzuwirken und somit eine nutzbare Datenbasis für die KI-Modelle zu generieren (vgl. Abbildung 6).

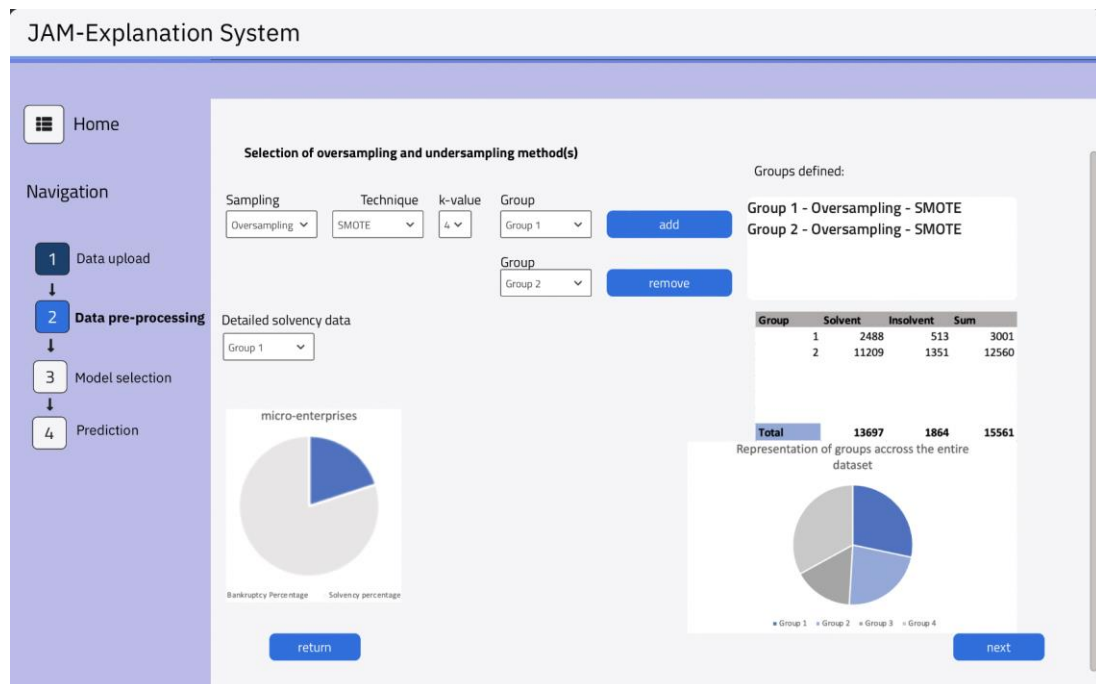


Abbildung 6: Mockup - Auswahl von Oversampling und Undersampling Techniken

Im nächsten Schritt soll, wie in A.2 formuliert, ein KI-Modell ausgewählt werden können. Bei vorheriger Segmentierung soll auch eine Auswahl pro Gruppe möglich sein. Bestehende Modelle sollen ebenfalls einsehbar sein (vgl. Abbildung 7).

JAM-Explanation System

Home

Navigation

- 1 Data upload
- 2 Data pre-processing
- 3 Model selection
- 4 Prediction

Selection of the AI method

ML approach: Neural network
Group: All groups
add
Group 2
remove

Current selection:
Group 1 - Neural Network
Group 2 - Neural Network

Ensemble method: Bagging

Model selection

☒ Create new model
☒ Select existing model

ML-Approach: Neural network
Dataset: Germany

Existing models:

ML-Approach	Training data	Creation date	Additional information
Neural network	Fossil fuels	27.06.22	Details
Neural network	Financials	12.11.21	Details
Neural network	Consumer staples	11.01.22	Details
XGBoost	Financials	13.11.21	Details

return next

Abbildung 7: Mockup - Auswahl KI-Modell

Außerdem sollen Hyperparameter der Modelle manuell gesetzt werden können. Der Benutzer kann auch die Aufteilung des Datensatzes in Trainings-, Test- und Validierungsset bestimmen (vgl. Abbildung 8). In Anlehnung an A.3 kann somit die Größe des Testsets bestimmt werden.

JAM-Explanation System

Home

Navigation

- 1 Data upload
- 2 Data pre-processing
- 3 Model selection
- 4 Prediction

Creation of new AI-model

XGBoost

General parameters

Booster: gbmtree
Verbosity: 0 (silent)
nthread: Default

Learning task parameters

objective: binary:logistic
eval_metric: error
seed: 0

Booster parameters

learning rate: 0.3
Gamma: Default
max_depth: 6
min_child_weight: 10
max_delta_step: 0
subsample: 1

colsample_bytree: 1
lambda: 1
alpha: 1
tree_method: hist
scale_pos_weight: 1
max_leaves: 0

Separation into training-/test-/validation-dataset:

Training data: 60%
Test data: 20%
Validation data: 20%

return next

New model

■ Training data ■ Test data ■ Validation data

Abbildung 8: Mockup - Hyperparameter-Auswahl und Splitting des Datensatzes

Nach der Modellgenerierung sollen analog zu A.5 verschiedene Kennzahlen die Schätzgüte der Modelle übersichtlich darstellen (vgl. Abbildung 9).

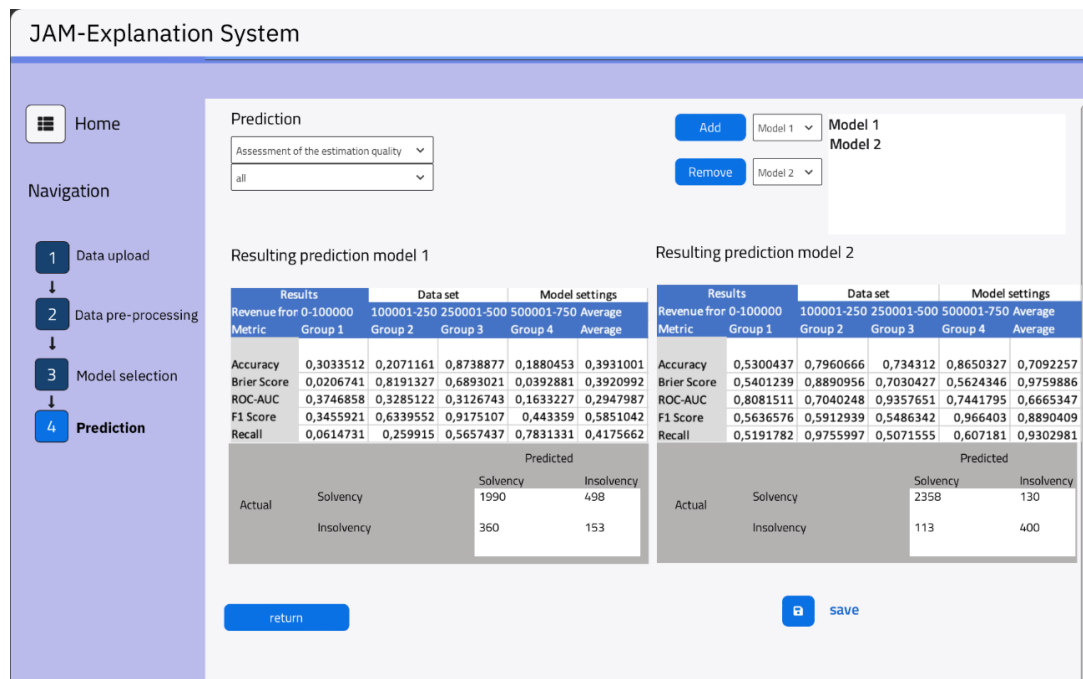


Abbildung 9: Mockup - Kennzahlen zur Schätzgüte

Schließlich sollen Erklärungsmethoden, wie sie in A.7 formuliert sind, dem Anwender die Bedeutung der einzelnen Merkmale für die Prognose transparent machen (vgl. Abbildung 10).

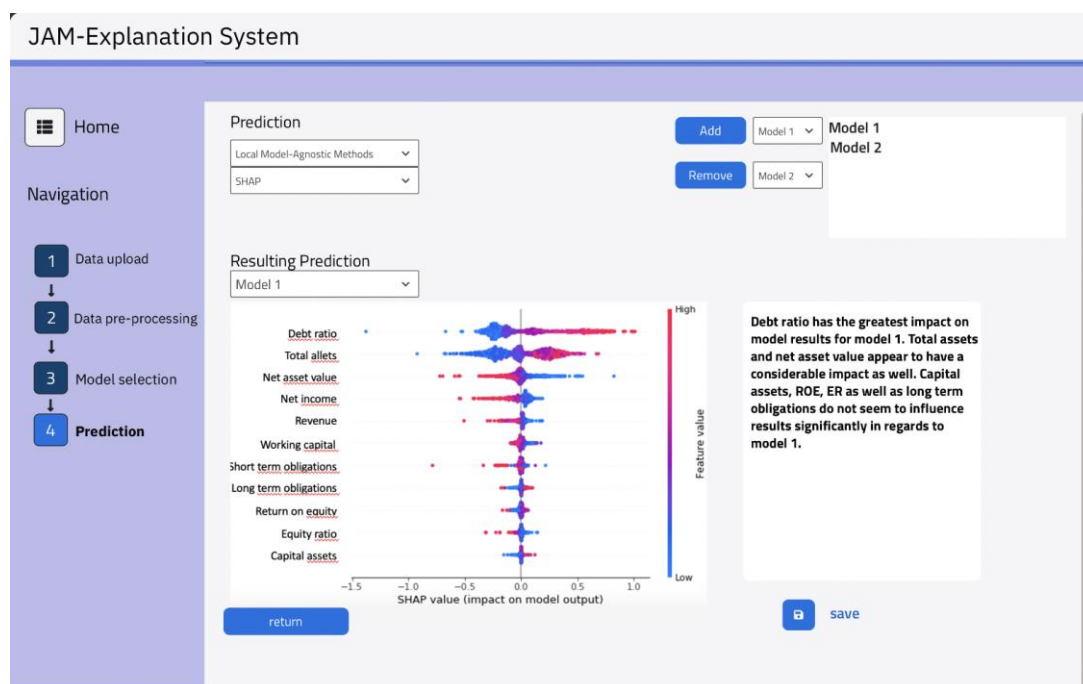


Abbildung 10: Mockup - Feature Importance

Die durchgehende Navigationsleiste dient der Orientierung des Anwenders und der Dokumentation des Bearbeitungsstandes des Anwenders in Anlehnung an A.6.

6 Praktische Entwicklung eines Dashboards zur KI-gestützten Insolvenzprognose

Die folgenden Abschnitte befassen sich mit der praktischen Umsetzung eines Dashboards zur KI-gestützten Insolvenzprognose. Dabei werden alle wichtigen Schritte erläutert, die für die Umsetzung des Projekts notwendig sind. Der erste Abschnitt beschäftigt sich mit der Datenvorverarbeitung, bei der der Prozess der Aufbereitung von Finanzdaten für eine Insolvenzprognose dargestellt wird. Dabei werden Schritte wie Datenbereinigung, Feature Engineering und Datensegmentierung erläutert. Der zweite Abschnitt widmet sich der Entwicklung des Prognosemodells, das im Rahmen des Projekts eingesetzt wird. Schließlich wird im dritten Abschnitt die Entwicklung eines Dashboards zur Visualisierung und Unterstützung der Erklärbarkeit von Insolvenzprognosen beschrieben, das den Nutzern eine Reihe von Funktionen zur Verwaltung und Analyse der Daten und Prognosen bietet.

6.1 Datenvorverarbeitung

In diesem Abschnitt wird der Prozess der Datenvorverarbeitung für eine auf Finanzdaten basierende Insolvenzprognose beschrieben. Es werden die Schritte der Datenbereinigung, des Feature Engineerings, der Segmentierung und der Modellvorbereitung der Daten erläutert. Bei der Datenbereinigung wird der Datensatz in ein Long-Format umgewandelt, das Beobachtungen über verschiedene Querschnittseinheiten im Zeitverlauf ermöglicht. Beobachtungen mit fehlenden oder inkonsistenten Werten werden entfernt. Nach der Datentransformation werden Finanzkennzahlen für die Kapitalstruktur, die Vermögensstruktur und die Liquidität berechnet, die als Grundlage für die Prognose dienen. Um die Analyse zu erleichtern, werden die Unternehmen nach ihrer Bilanzsumme in homogene Gruppen eingeteilt. Schließlich wird der Datensatz in Trainings- und Testdatensätze aufgeteilt, wobei die Unausgewogenheit des Datensatzes durch Oversampling-Techniken reduziert wird. Der gesamte Prozess der Datenvorverarbeitung ist notwendig, um einen geeigneten und repräsentativen Datensatz für die Erstellung und Überprüfung von Insolvenzprognosemodellen zu erhalten.

6.1.1 Datenbereinigung

Nach der Extraktion lagen die Daten in einem Wide-Format mit einer Zeile pro Unternehmen vor. Die Finanzdaten waren in Spalten pro Jahr angeordnet. Die Zielvariable wurde aus den Informationen der Spalten „Status“ und „Status Datum“ ebenfalls in eine Spalte „Insolvenz“ pro Jahr angeordnet. Die Zielvariable ist binär, wobei der Status 0 als solvent und 1 als insolvent definiert ist. Anschließend wurden die Daten in ein Long-Format umgewandelt, in dem jedes Unternehmen 10 Zeilen für 10 Jahre hatte, um die Daten vergleichbar zu machen. In diesem Format werden die Daten als Paneldaten bezeichnet, da sie Beobachtungen über verschiedene Querschnittseinheiten im Zeitverlauf ermöglichen. Zu diesem Zeitpunkt umfasste der Datensatz 14,53 Mio. Datenpunkte, die 1,45 Mio. Einzelunternehmen und 8721 Insolvenzen, d.h. 0,6% der Einzelunternehmen, entsprachen. Bezüglich des Verhältnisses von Insolvenzen und Solvenzen kann somit eindeutig nachvollzogen werden, dass der Datensatz unausgewogen ist, da die Minderheitsklasse der Insolvenzen deutlich kleiner ist als die Mehrheitsklasse der Solvenzen. Die Imbalance innerhalb des Datensatzes verdeutlicht die Notwendigkeit, Parameteroptimierung oder Over-/Undersampling-Techniken zu verwenden. Zu diesem Zweck müssen Zeilen mit fehlenden Werten aus dem Datensatz entfernt werden.

Damit eine realitätsnahe Insolvenzprognose möglich ist, wurde bei den insolventen Unternehmen das Insolvenzjahr um ein Jahr nach hinten verschoben, da bei einer Insolvenzprognose mit neuen Unternehmensdaten die Finanzdaten im Insolvenzjahr nicht zur Verfügung stehen. Darüber hinaus wurden Daten, die nach der Insolvenz entstanden sind, gelöscht. Für die weitere Datenbereinigung wurden die Kennzahlen Gesamtvermögen, Anlagenintensität, Umlaufintensität und kurzfristige Verschuldungsquote aus den bisherigen Kennzahlen berechnet. Alle Beobachtungen mit fehlenden Werten in einer der folgenden Spalten wurden gelöscht: Anlagevermögen, Umlaufvermögen, Bilanzsumme, Eigenkapital, kurzfristige Verbindlichkeiten und Jahresüberschuss bzw. -fehlbetrag. Darüber hinaus wurden Beobachtungen mit negativen Werten für eindeutig positive Kennzahlen (Bilanzsumme, Anlagenintensität, Umlaufintensität, kurzfristiger Verschuldungsgrad) entfernt.

6.1.2 Feature Engineering

Nach der Transformation der Daten wurden Finanzkennzahlen für die Kapitalstruktur, die Vermögensstruktur und die Liquidität als Grundlage für die Prognose berechnet (vgl. Tabelle 7). Die Auswahl der Finanzkennzahlen sowie die Unterscheidung zwischen Kennzahlen zur Kapitalstruktur, solchen zur Vermögensstruktur und Kennzahlen zur Liquiditätsanalyse sind an Coenenberg et al. (2014) angelehnt.

Metrik-Typ	Abkürzung	Name	Formel
Kapitalstruktur	ER (1)	Equity Ratio – Eigenkapitalquote	$\frac{\text{Eigenkapital}}{\text{Bilanzsumme}}$
	STDR (2)	Short-term Debt Ratio – Kurzfristige Verschuldungsquote	$\frac{\text{Kurzfristige Verbindlichkeiten}}{\text{Bilanzsumme}}$
	WCR (3)	Working Capital Ratio – Working Capital Anteil	$\frac{\text{Umlaufvermögen}}{\text{Kurzfristige Verbindlichkeiten}}$
Vermögensstruktur	RTA (4)	Return on Total Assets – Gesamtkapitalrentabilität	$\frac{\text{Jahresüberschuss bzw. –fehlbetrag}}{\text{Bilanzsumme}}$
	ROE (5)	Return on Equity – Eigenkapitalrentabilität	$\frac{\text{Jahresüberschuss bzw. –fehlbetrag}}{\text{Eigenkapital}}$
Liquidität	ACR (6)	Asset Coverage Ratio – Anlagendeckungsgrad II	$\frac{\text{Eigenkapital}}{\text{Anlagevermögen}}$
	L2 (7)	Second Degree Liquidity – Liquidität zweiten Grades	$\frac{\text{Umlaufvermögen}}{\text{Kurzfristige Verbindlichkeiten}}$

Tabelle 7: Metriken zur Insolvenzprognose

6.1.3 Segmentierung und Modellvorbereitung der Daten

Darüber hinaus wurden die Unternehmen nach der Größe der Bilanzsumme unterteilt, um homogene Gruppen zu untersuchen (vgl. Tabelle 8) (Ciampi, 2015; El Kalak & Hudson, 2016).

Gruppe	Differenzierung der Unternehmen	Bilanzsumme
1	Kleinstkapitalgesellschaften	< 350,000€
2	kleine Kapitalgesellschaften	350,000€ - 6,000,000€
3	mittelgroße Kapitalgesellschaften	6,000,000€ - 20,000,000€
4	große Kapitalgesellschaften	> €20,000,000

Tabelle 8: Differenzierung der Unternehmen nach der Bilanzsumme

Das Jahr 2021 wurde bewusst ausgelassen, da sich herausstellte, dass die meisten Unternehmen zum Zeitpunkt der Datenerhebung noch keine verlässlichen Daten für 2021 geliefert hatten. Nach der Datenbereinigung sind von den über 14 Millionen Datenpunkten bloß 324.620 übriggeblieben. Dies ist darauf zurückzuführen, dass in den Amadeus-Daten häufig Informationen/Metriken für einige oder alle Jahre fehlen. Über alle Gruppen hinweg enthielt der bereinigte Datensatz 330 Insolvenzen, was das Ungleichgewicht im verwendeten Datensatz verdeutlicht.

Im nächsten Schritt wurde der Datensatz in einen Trainings- und einen Testdatensatz separiert. Um sowohl im Trainings- als auch im Testdatensatz eine hinreichende Anzahl an Insolvenzen zu gewährleisten, wurde eine Aufteilung von 70% Trainings- und 30% Testdatensatz gewählt. Um eine Überanpassung im Zusammenhang mit den verwendeten Paneldaten zu vermeiden, erfolgte die Aufteilung in Trainings- und Testdaten nach Unternehmenseinheiten und nicht nach Datenpunkten. Paneldaten sind mehrdimensionale Daten, die über einen bestimmten Zeitraum hinweg Informationen von denselben Individuen, in unserem Fall Unternehmen, erfassen (vgl. Tabelle 9) (Baur & Blasius, 2014).

Name	Company_ID	Jahr	Insolvenz	ER	WCR	RTA	ROE	ACR	L2	STDR
Unternehmen 1	58587	2016	0	0,95	57,7	0,04	0,04	5,88	57,7	0,01
Unternehmen 1	58587	2015	0	0,89	6,5	0,04	0,04	2,89	6,5	0,11
Unternehmen 1	58587	2012	0	0,76	2,73	0,09	0,12	1,78	2,73	0,21
Unternehmen 2	70546	2020	0	0,01	0,43	-0,2	-14	0,02	0,43	0,41
Unternehmen 2	70546	2019	0	0,26	0,59	0,03	0,13	0,35	0,59	0,45
Unternehmen 2	70546	2018	0	0,28	0,72	0,05	0,19	0,41	0,72	0,44
Unternehmen 2	70546	2017	0	0,25	0,66	0,05	0,22	0,36	0,66	0,46
Unternehmen 2	70546	2016	0	0,22	0,74	0,08	0,35	0,36	0,74	0,5
Unternehmen 2	70546	2015	0	0,17	0,58	0,03	0,16	0,25	0,58	0,55
Unternehmen 3	7058600	2020	0	0,94	1,49	0,31	0,33	1,02	1,49	0,05
Unternehmen 3	7058600	2019	0	0,63	15,7	0,08	0,12	0,64	15,7	0
Unternehmen 3	7058600	2018	0	0,56	10	0,07	0,13	0,56	10	0
Unternehmen 3	7058600	2017	0	0,48	0	0,06	0,13	0,48	0	0
Unternehmen 3	7058600	2016	0	0,42	0	0,06	0,13	0,42	0	0
Unternehmen 3	7058600	2015	0	0,36	0	0,05	0,13	0,36	0	0,37
Unternehmen 3	7058600	2014	0	0,31	0,02	0,04	0,13	0,32	0,02	0,37
Unternehmen 3	7058600	2013	0	0,33	0,02	0,06	0,18	0,33	0,02	0,45
Unternehmen 3	7058600	2012	0	0,27	0,02	0,05	0,2	0,27	0,02	0,45

Tabelle 9: Beispiel von Paneldaten

Darüber hinaus wurde der Panel-Datensatz wieder in ein Wide-Format transformiert, um die Ergebnisse des verwendeten Modells in Abhängigkeit von der Datenform vergleichen zu können. Für die Transformation des Datensatzes in ein Wide-Format mussten alle Unternehmen die gleiche Anzahl an Perioden aufweisen. Um dies zu gewährleisten, wurde nur die Anzahl der Perioden der insolventen Unternehmen betrachtet und der Durchschnitt dieser Perioden berechnet. Diese Periodenzahl wurde dann auf alle Unternehmen angewendet. Unternehmen mit weniger Perioden wurden ausgeschlossen. Bei Unternehmen mit mehr Perioden wurden die aktuellsten n Perioden vor und inklusive der Insolvenzperiode ausgewählt. Dieses Verfahren wurde für jede Gruppe (Unternehmensgröße) einzeln durchgeführt (vgl. Tabelle 10).

	Anzahl an Insolvenzen		Anzahl der Perioden
Gruppe	Long-Format	Wide-Format	Wide-Format
1	40	3	2
2	138	40	2
3	99	60	4
4	53	31	4
Summe	330	134	

Tabelle 10: Anzahl der Insolvenzen pro Gruppe und Anzahl der berücksichtigten Zeiträume im Wide-Format

Auch in diesem Fall wurde eine Aufteilung von 70% zu 30% für den Trainings- und Testdatensatz vorgenommen. Die zehnfach geschichtete Kreuzvalidierung wurde in einer zweiten Variante für das Trainieren und Testen des Modells verwendet (Olson et al., 2012). Jeder der drei Datensätze wurde in zehn disjunkte Test- und Trainingsdatensätze aufgeteilt (Liang et al., 2015). Die zehnfache geschichtete Kreuzvalidierung wurde genutzt, da sie sicherstellt, dass die Klassenhäufigkeiten in respektiven Datensätzen ausgeglichener sind. Die geschichtete Kreuzvalidierung kann hier angewendet werden, da die Datenpunkte im Wide-Format unkorreliert sind.

Da der zuvor beschriebene Datensatz unausgewogen war, wurden Oversampling-Techniken eingesetzt. Konkret wurde die Oversampling-Methode Synthetic Minority Oversampling Technique (SMOTE) und deren Kombination mit den Algorithmen Tomek Links (SMOTE-Tomek) und Edited Nearest Neighbor (SMOTE-ENN) verwendet.

SMOTE ist eine Technik, die synthetische Beispiele zur Minderheitsklasse hinzufügt, indem die Merkmalsräume zwischen den vorhandenen Beispielen interpoliert werden. Dies hilft, das Problem des unausgeglichene Datensatzes zu mildern, indem die Anzahl der Minderheitsklassenbeispiele erhöht wird. SMOTE-Tomek kombiniert SMOTE mit Tomek Links, einer Methode, die Paare von Beispielen identifiziert und entfernt, die als Rauschen oder Grenzbeispiele betrachtet werden. Dies führt zu einer klareren Trennung zwischen den Klassen und verbessert die Klassifikationsleistung. SMOTE-ENN kombiniert SMOTE mit dem Edited Nearest Neighbor Algorithmus. Dieser Algorithmus entfernt jedes Beispiel, dessen Klasse nicht mit der Klasse der Mehrheit seiner k nächsten

Nachbarn übereinstimmt. Dies führt zu einer glatteren Entscheidungsgrenze und hilft, das Rauschen zu reduzieren (Batista et al., 2004; Chawla et al., 2002).

Die Kombination von SMOTE mit Tomek Links und Edited Nearest Neighbor ist sinnvoll, da sie nicht nur die Minderheitsklasse durch synthetische Beispiele aufwertet, sondern auch die Qualität des Datensatzes durch die Reduzierung von Rauschen und Grenzbeispielen verbessert. Gerade in Szenarien, in denen es entscheidend ist, zwischen zahlungsfähigen und zahlungsunfähigen Firmen zu unterscheiden, bietet dieser Ansatz verbesserte Robustheit (Batista et al., 2004; Chawla et al., 2002).

6.2 Modellentwicklung

Im Zuge der Modellentwicklung war das Ziel, ein Modell zu finden, welches bestmöglich Insolvenzen prognostizieren kann. Wir verwenden einen gelabelten Datensatz, um unsere Daten zu klassifizieren, was als überwachtes Lernen bezeichnet wird (Zhang et al., 2021). Die Variable, die unser Modell vorhersagen soll, ist die Insolvenz selbst. Das Modell soll in der Lage sein, Unternehmen danach zu klassifizieren, ob sie solvent erscheinen oder nicht. Es handelt sich also um eine binäre Klassifikation.

Die Literaturrecherche zu KI-gestützter Insolvenzprognose hat ergeben, dass XGBoost belastbare Ergebnisse bei hoher Skalierbarkeit liefert. XGBoost ermöglicht überwachtes Lernen mit einem Boosted Tree Algorithmus (Chen & Guestrin, 2016; Pamuk et al., 2021; Zięba et al., 2016).

Da nun sowohl Daten im Wide- als auch im Long-Format vorlagen, wurde sich dazu entschieden, beide Formate im Rahmen einer XGBoost-Klassifikation zu verwenden. Nach der Anwendung der XGBoost-Klassifikation auf die verschiedenen Unternehmensgruppen wurden folgende Ergebnisse erzielt:

Eine hohe Genauigkeit (Accuracy) ist für alle Datenformate und Oversampling-Techniken zu beobachten. Dies ist hier jedoch weniger relevant, da es sich um das Verhältnis der Anzahl der richtigen Vorhersagen zur Gesamtzahl der Vorhersagen für einen Datensatz handelt. Die Ergebnisse der Scores beziehen sich auf die Fähigkeit des Modells, den Testdatensatz vorherzusagen. Der Testdatensatz ist jedoch nicht oversampled und daher ungleich gewichtet. Somit ist die Aussagekraft eines Modells, das viele solvente und wenige insolvente Unternehmen prognostiziert beschränkt. Wie

bereits in Kapitel 2.3.2 erläutert, sind im Kontext der Insolvenzprognose mit einem unausgewogenen Datensatz die Präzision (Precision), die Sensitivität (Recall) und der F1-Score relevante Messgrößen. Um die Leistung des trainierten Modells zu beurteilen, bezieht sich diese Arbeit vornehmlich auf den F1-Score, da er das harmonische Mittel zwischen Präzision und Sensitivität abbildet (Goyal, 2021).

Die Ergebnisse des Extreme Gradient Boostings mit dem Paneldatensatz zeigen eine sehr geringe Präzision (vgl. Tabelle 11, 12, 13). So wurde kaum ein insolventes Unternehmen korrekt vorhergesagt. Gruppe 3 lieferte bessere Ergebnisse für die verschiedenen Oversampling-Techniken. Für Gruppe 1 konnte mit dem Wide-Format keine Prognose erstellt werden, da die Anzahl der Insolvenzen für die Oversampling-Techniken nicht ausreichend war.

Insgesamt sind die Ergebnisse mit dem Wide-Format-Datensatz deutlich besser. Der F1-Score für Gruppe 3 liegt bei allen implementierten Oversampling-Techniken zwischen 20 und 30 %. Die Verwendung eines 70/30-Splits oder einer geschichteten Kreuzvalidierung für den Trainings- und Testdatensatz führt zu keinem signifikanten Unterschied. Es fällt auf, dass nur die Ergebnisse für Gruppe 3 vergleichbare Werte aufweisen. Betrachtet man die Sensitivität, so fällt auf, dass diese auch bei Gruppe 3 nur bei ca. 35% liegt. Dies bedeutet, dass nur 35% der insolventen Unternehmen im Testdatensatz als solche identifiziert werden. Mit SMOTE-ENN und geschichteter Kreuzvalidierung liegt der Wert bei 40% (vgl. Tabelle 13).

Oversampling	SMOTE											
ML Model	XGBoost											
Daten Format	Panel Data				Wide Data							
Train/Test Split	70/30				70/30				SKF			
Gruppe	1	2	3	4	1	2	3	4	1	2	3	4
Accuracy	97.33%	96.04%	98.03%	98.85%	99.19%	99.07%	99.23%		99.12%	99.17%	99.43%	
Brier Score	2.10%	2.99%	1.51%	0.91%	0.70%	0.79%	0.56%		0.72%	0.69%	0.48%	
ROC-AUC	42.98%	59.53%	63.51%	55.94%	60.50%	93.19%	74.54%		58.14%	92.06%	77.38%	
F1-Score	2.11%	0.67%	2.20%	0.00%	3.77%	23.08%	0.00%		0.83%	25.30%	9.00%	
Recall	9.09%	6.38%	25.81%	0.00%	7.69%	27.27%	0.00%		2.50%	33.30%	6.67%	

Precision	1.19%	0.35%	1.15%	0.00%	2.50%	20.00%	0.00%	0.50%	21.28%	15.00%
Log Loss	92.28%	136.64%	68.17%	39.88%	28.00%	32.06%	26.42%	30.33%	28.55%	19.59%

Tabelle 11: Bewertungsergebnisse der einzelnen XGBoost-Modelle mit der Oversampling-Technik SMOTE

Oversampling	SMOTE-Tomek											
ML Model	XGBoost											
Daten Format	Panel Data				Wide Data							
Train/Test Split	70/30				70/30				SKF			
Gruppe	1	2	3	4	1	2	3	4	1	2	3	4
Accuracy	96.93%	95.73%	98.02%	98.93%	99.13%	98.98%	99.26%		99.10%	99.12%	99.45%	
Brier Score	2.21%	3.14%	1.55%	0.85%	0.70%	0.81%	0.56%		0.71%	0.72%	0.49%	
ROC-AUC	46.37%	59.48%	64.50%	49.34%	67.80%	94.00%	74.91%		57.88%	92.21%	74.60%	
F1-Score	1.83%	1.02%	1.92%	0.00%	3.51%	24.14%	0.00%		0.83%	23.49%	9.00%	
Recall	9.09%	10.64%	22.58%	0.00%	7.69%	31.82%	0.00%		2.50%	33.33%	6.67%	
Precision	1.02%	0.54%	1.00%	0.00%	2.27%	19.44%	0.00%		0.50%	18.68%	15.00%	
Log Loss	106.17%	147.64%	68.46%	37.07%	30.20%	35.27%	25.60%		31.02%	30.45%	19.10%	

Tabelle 12: Bewertungsergebnisse der einzelnen XGBoost-Modelle mit der Oversampling-Technik SMOTE-Tomek

Oversampling	SMOTE-ENN											
ML Model	XGBoost											
Daten Format	Panel Data				Wide Data							
Train/Test Split	70/30				70/30				SKF			
Gruppe	1	2	3	4	1	2	3	4	1	2	3	4
Accuracy	94.40%	94.90%	97.64%	98.53%	98.71%	98.70%	99.16%		98.60%	98.76%	99.22%	
Brier Score	4.26%	3.84%	1.87%	1.16%	1.08%	1.06%	0.71%		1.10%	0.98%	0.63%	
ROC-AUC	44.82%	59.26%	62.82%	55.77%	66.56%	94.37%	76.86%		63.71%	90.68%	77.08%	
F1-Score	1.02%	0.69%	2.08%	0.00%	2.41%	22.22%	0.00%		1.07%	21.08%	11.67%	
Recall	9.09%	8.51%	29.03%	0.00%	7.69%	36.36%	0.00%		5.00%	40.00%	10.00%	
Precision	0.54%	0.36%	1.08%	0.00%	1.43%	16.00%	0.00%		0.60%	14.62%	16.67%	
Log Loss	193.48%	176.22%	81.40%	50.93%	44.47%	44.89%	28.90%		48.26%	42.82%	26.94%	

Tabelle 13: Bewertungsergebnisse der einzelnen XGBoost-Modelle mit der Oversampling-Technik SMOTE-ENN

Wie die Ergebnisse des Modells zeigen, liefern die Metriken, die als aussagekräftige Kenngrößen definiert wurden, nämlich Präzision, Sensitivität und F1-Score, keine

zufriedenstellenden Ergebnisse in Bezug auf die Fähigkeit des Modells, Insolvenzen vorherzusagen. In diesem Fall lässt sich aufgrund der Metriken davon ausgehen, dass die Vorhersagen in Gruppe 3 ("mittelgroße Kapitalgesellschaften") robuster sind als in den anderen Gruppen. Die Analyse von Tabelle 11, 12, 13 und Tabelle 10 deutet darauf hin, dass der Grund für die höhere Prognosegenauigkeit der Gruppe 3 in der im Vergleich zu den anderen Gruppen größeren Gesamtzahl der Insolvenzen liegt, insbesondere bei den Modellen, die auf Wide-Daten zurückgreifen. Gruppe 3 bezieht auch Finanzkennzahlen über einen längeren Zeitraum von vier Perioden ein, während Gruppe 1 und 2 jeweils nur zwei Perioden betrachten. Dies lässt vermuten, dass die Veränderung der Finanzkennzahlen über einen längeren Zeitraum sowie eine ausreichende Anzahl tatsächlicher Insolvenzen notwendig sind, um die Zuverlässigkeit der entwickelten Schätzungen und die damit verbundene Signifikanz der Merkmale zu gewährleisten.

Die Verwendung eines einzigen Algorithmus für maschinelles Lernen, XGBoost, kann jedoch die Anwendbarkeit der resultierenden Vorhersagemodelle einschränken. Alternative Methoden, die auf Zeitreihen basieren, könnten sich als nützlicherer Ansatz für die Schätzung von Insolvenzen erweisen (Åkerblom & Johan Lyhagen, 2022; H. Kim et al., 2022).

Daher wurde ebenfalls LSTM als alternative Prognosemethode getestet. LSTM ist eine Art rekurrentes neuronales Netzwerk (RNN), das speziell für die Verarbeitung von Datensequenzen entwickelt wurde. Aus der Literatur geht hervor, dass solche Modelle bereits erfolgreich in der Insolvenzprognose eingesetzt wurden (Åkerblom & Johan Lyhagen, 2022; H. Kim et al., 2022).

Auf Basis der Paneldaten aus 4.1.3. wurde sichergestellt, dass nur konsekutive Daten und Daten bis zur möglichen Insolvenz verwendet werden. Außerdem mussten aufgrund der Modellanforderungen die Periodenlängen pro Gruppe gleich sein.

Um die Paneldaten für LSTM nutzbar zu machen, wurden die Eingabedaten anhand der ID-Nummer gruppiert. Anschließend wurden die gruppierten Daten in separate Arrays aufgeteilt, wobei jede ID-Nummer ein Array darstellt. Jede Tabelle hat n Zeilen (entsprechend der Anzahl der aufeinander folgenden Jahresdaten) und 7 Spalten (entsprechend der Anzahl der Merkmale). Die gruppierten Arrays sind die endgültigen

Eingabedaten, die in das LSTM-Modell eingegeben werden. Als Zielvariable wurde die zeitlich letzte Beobachtung des Insolvenzstatus gewählt.

Nach der weiteren Aufbereitung der Daten für das LSTM-Modell sind im Datensatz basierend auf dem Long-Format nun weniger insolvente Unternehmen als in dem auf dem Wide-Format basierenden. Besonders hervorzuheben ist hier die Gruppe 1 mit nur 3 insolventen Unternehmen. Die Gruppe mit den meisten insolventen Unternehmen ist Gruppe 3 mit 42 Insolvenzen (vgl. Tabelle 14).

	Anzahl an Insolvenzen	Anzahl der Perioden
Gruppe	Long-Format angepasst für LSTM	
1	3	2
2	30	2
3	42	4
4	29	5
Summe	103	

Tabelle 14: Anzahl der Insolvenzen pro Gruppe und Anzahl der berücksichtigten Zeiträume im Long-Format angepasst für LSTM

Die LSTM-Modelle wurden über 30 Epochen trainiert. Da die SMOTE-Techniken bei diesem Datenformat nicht angewendet werden konnten, wurde sich für zwei verschiedene Varianten entschieden: Im ersten Fall wurde Cost-sensitive Learning verwendet. Bei dieser Methode wird das Modell während des Trainings dazu angehalten, die Minderheitsklasse besser zu erkennen, indem eine falsche Klassifizierung der Minderheitsklasse stärker bestraft wird. Dies verbessert die Leistung des Modells, insbesondere in Bezug auf die Erkennung der Minderheitsklasse. Die zweite Variante nutzt Random Oversampling. Random Oversampling funktioniert, indem Beobachtungen der Minderheitsklasse zufällig ausgewählt und im Trainingsdatensatz mehrfach repliziert werden (Zhang et al., 2021).

Oversampling	Kein Oversampling, aber Cost-sensitive learning				Random Oversampler			
ML Model	LSTM				LSTM			
Daten Format	Panel Data				Panel Data			
Train/Test Split	70/30				70/30			
Gruppe	1	2	3	4	1	2	3	4
Accuracy	97.54%	93.36%	95.47%	98.07%	55.03%	6.91%	80.87%	59.21%
Brier Score	2.46%	6.64%	4.53%	1.93%	44.97%	93.09%	19.13%	40.79%
ROC-AUC	48.82%	46.73%	75.95%	54.71%	77.49%	53.41%	84.17%	68.47%
F1-Score	0.00%	0.00%	10.40%	3.12%	0.50%	0.22%	4.11%	1.06%
Recall	0.00%	0.00%	56.25%	11.11%	100.00%	100.00%	87.50%	77.78%
Precision	0.00%	0.00%	5.73%	1.82%	0.25%	0.11%	2.10%	0.53%
Log Loss	85.00%	229.21%	156.58%	66.63%	1553.12%	3215.41%	660.90%	1408.88%

Tabelle 15: Bewertungsergebnisse der einzelnen LSTM-Modelle

Wie aus Tabelle 1 ersichtlich, sind die Ergebnisse ohne Oversampling in Bezug auf die Genauigkeit und den F1-Score besser als die Prognose mit Random Oversampling. Die Gruppe 1 kann hier aufgrund der wenigen Insolvenzen generell ausgeklammert werden. Für den Datensatz mit Oversampling kann ein hoher Prozentsatz bezüglich der Sensitivität, beziehungsweise des Recalls, festgestellt werden, jedoch ist die Genauigkeit und somit der F1-Score niedrig. Insgesamt hat sich die Modellgüte mit LSTM im Vergleich zu XGBoost nicht verbessert. Gruppe 3 liefert, wie schon bei den vorherigen Prognosen, die besten Ergebnisse, die aber für die Insolvenzprognose immer noch als unbefriedigend eingestuft werden können.

Die Grenzen des erstellten Prototyps eines auf maschinellem Lernen basierenden Insolvenzprognosemodells liegen in der Qualität der Daten, mit denen das Modell trainiert wird. Amadeus liefert viele einzelne Datenpunkte, wobei sich die Stichprobengröße nach Bereinigung der Daten erheblich verkleinert. Darüber hinaus erscheint die Gesamtzahl der Insolvenzen nach der Datenbereinigung zu gering, um ein genaues Bild der Insolvenzursachen zu liefern. Selbst bei einer Oversampling-Stichprobe führt die geringe Stichprobengröße zu Daten von geringer Qualität, da die Variation der zuverlässigen Prädiktoren für Insolvenzen nicht repräsentativ ist.

Da die Erfassung eines zusätzlichen Datensatzes im Rahmen dieser Arbeit nicht mehr umsetzbar war, wurde sich dazu entschieden, das in Bezug auf die Insolvenzprognosegenauigkeit genauere/performantere Modell zu verwenden, welches auf XGBoost basiert.

6.3 Implementierung eines Dashboards zur Visualisierung ML-basierter Insolvenzprognose

Abschnitt 6.3 beschreibt die Entwicklung eines Dashboards zur Visualisierung von maschinenlernbasierten Insolvenzprognosen. Das Dashboard basiert auf Python und dem Django Webframework und gliedert sich in mehrere Prozessphasen. Am Anfang steht die Datenverarbeitung, gefolgt von der Modellauswahl und -spezifikation, dem Training und schließlich der Prognose inklusive Erklärbarkeit des Modells und Exportmöglichkeiten der Ergebnisse für den Anwender. Im Rahmen der Erklärbarkeit des Modells werden die dafür identifizierten Ansätze beschrieben. Das System bietet dem Anwender eine Reihe von Funktionen wie Daten-Upload, Daten-Übersicht, Daten-Bereinigung, Feature Engineering, Daten-Segmentierung, Modell-Setup und -Optimierung sowie Modell-Evaluierung.

6.3.1 Technologien und Frameworks

Als Programmiersprache, um Daten zu bereinigen und KI-Modelle zu kreieren, wurde Python verwendet, da es eine benutzerfreundliche Umgebung sowie eine Vielfalt an Bibliotheken und Tools für maschinelles Lernen anbietet (Mckinney, 2010; Pedregosa et al., 2011). Darauf aufbauend wurde das Django-Framework genutzt, um das KI-Dashboard zu entwickeln. Django ist ein Web-Framework, welches auf Python aufbaut und sich als robuste, skalierbare und sichere Umgebung gut für die Erstellung von KI-gesteuerten Webanwendungen eignet (Holovaty & Kaplan-Moss, 2009; Rubio, 2017).

6.3.2 Prozess Workflow des Dashboards

Insgesamt kann der Prozessablauf der Anwendung in die folgenden Phasen unterteilt werden: Datenverarbeitung, Modellauswahl und Verarbeitung, Modellspezifikation und Training sowie Vorhersage inklusive Erklärbarkeit des Modells & Exportmöglichkeiten der Ergebnisse für den Anwender.

In der Datenverarbeitungsphase werden die Rohdaten nach dem Datenupload bereinigt. Anschließend werden auf Basis der bereinigten Daten Merkmale erstellt, die im weiteren Verlauf als Prognosegrundlage dienen. Zusätzlich wird dem Nutzer eine Segmentierungsmöglichkeit der Daten angeboten, die sich an der Bilanzsumme der Unternehmen orientiert.

Im nächsten Schritt erfolgt die Modellauswahl. Hier kann ein Hyperparameter-Tuning oder alternativ ein Oversampling-Verfahren vom Anwender ausgewählt werden. Weiterhin wird ein Train-Test-Split-Verhältnis bestimmt. Im Hintergrund wird außerdem der Datensatz von einem Long-Format in ein Wide-Format konvertiert. In einem weiteren Schritt werden die Hyperparameter ausgewählt, was im Falle einer Bayesian Optimization nicht notwendig ist. Anschließend werden das oder die Modelle trainiert.

Schließlich kann der Benutzer verschiedene Tabellen und Diagramme einsehen, die die Erklärbarkeit des Modells unterstützen. Eine Filteroption der Prognoseergebnisse sowie verschiedene Exportmöglichkeiten der generierten Informationen werden ebenfalls zur Verfügung gestellt. Die Modelle werden gespeichert (vgl. Abbildung 11).

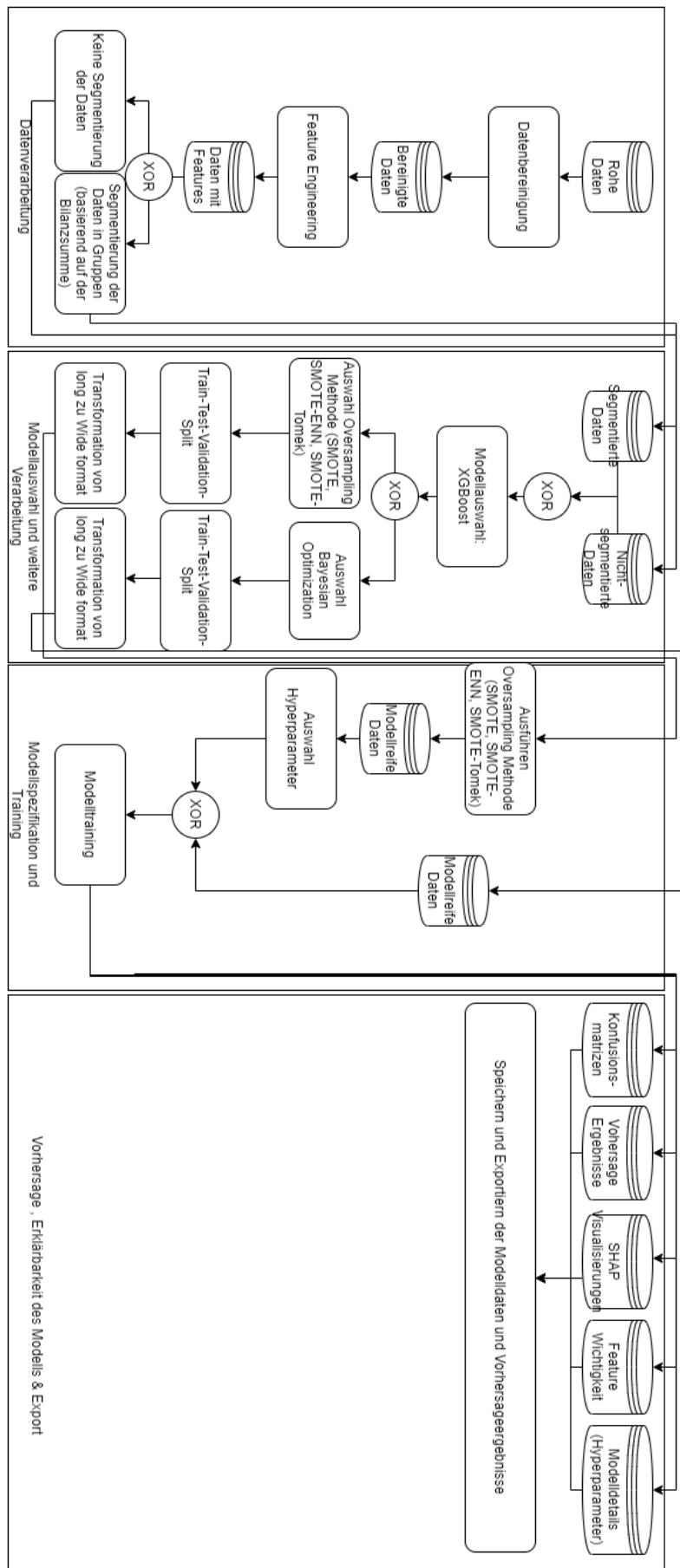


Abbildung 11: Prozessworkflow der gesamten Applikation

6.3.3 Landing Page, Prozessdokumentation und persönliche Übersicht

Die Landing Page ist der Startpunkt der Anwendung für den Benutzer. Hier wird der Nutzen der Applikation erläutert und der weitere Prozess kann von dieser Seite aus gestartet werden (vgl. Abbildung 12).

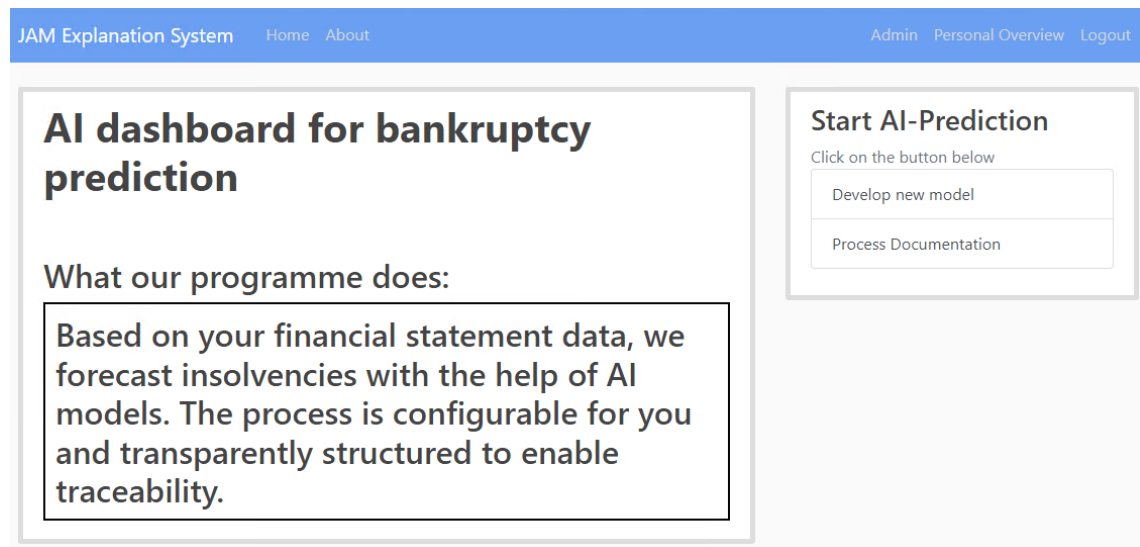


Abbildung 12: Landing-Page

Neben dem Start des Modellerstellungsprozesses kann eine Prozessdokumentation eingesehen werden, die den Anwendungsprozess Schritt für Schritt beschreibt. Auch von hier aus hat der Benutzer die Möglichkeit, den Prozess zu starten (vgl. Abbildung 13).

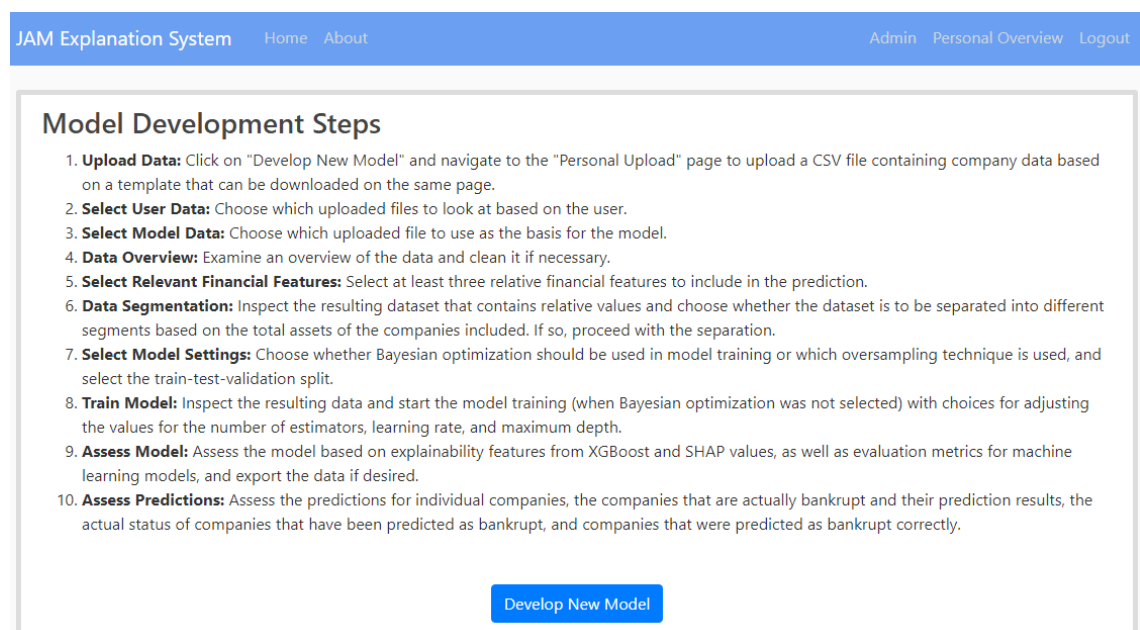


Abbildung 13: Prozessdokumentation

Auf allen Seiten der Applikation gibt es zudem eine Navigationsleiste mit Informationen zu den Programmierern der Applikation, einen Administrationsbereich, den das Django Framework vorkonfiguriert anbietet und einer persönlichen Übersichtsseite, auf der der Nutzer bereits hochgeladene Datensätze und erstellte Modelle einsehen kann (vgl. Abbildung 14). Die persönliche Übersicht basiert auf dem im Rahmen der Konzeption entwickelten Mockup zur Übersicht über die Datensätze (vgl. Abbildung 2).

The screenshot shows a web interface for the 'JAM Explanation System'. At the top is a navigation bar with links for 'Home', 'About', 'Admin', 'Personal Overview', and 'Logout'. Below the navigation bar, there are two main sections. The first section, titled 'Datasets used by user:', contains a table with columns 'Superuser', 'User', 'Uploaded at', and 'File'. It shows one entry where Joachim (Superuser) and User1 (User) uploaded a file named 'Insolvenzen_Django.csv' on April 3, 2023, at 12:09 p.m. The second section, titled 'AI-Models used by user:', contains a similar table with columns 'Superuser', 'User', 'Uploaded at', and 'File'. It shows three entries where Joachim (Superuser) and User1 (User) uploaded files named 'xg_model/Small.pkl', 'xg_model/Medium.pkl', and 'xg_model/Big.pkl' on April 3, 2023, at 12:17 p.m.

Superuser	User	Uploaded at	File
Joachim	User1	April 3, 2023, 12:09 p.m.	Insolvenzen_Django.csv

Superuser	User	Uploaded at	File
Joachim	User1	April 3, 2023, 12:17 p.m.	xg_model/Small.pkl
Joachim	User1	April 3, 2023, 12:17 p.m.	xg_model/Medium.pkl
Joachim	User1	April 3, 2023, 12:17 p.m.	xg_model/Big.pkl

Abbildung 14: Persönliche Übersicht der Datensätze und KI-Modelle

6.3.4 Daten-Upload und Daten-Übersicht

Im nächsten Prozessschritt wird dem Benutzer ein Datei-Upload in Form einer CSV-Datei angeboten. Es wird genau beschrieben, welche Spaltennamen und welche Spaltenreihenfolge zu verwenden sind. Außerdem werden die Datentypen für jede Spalte angegeben. Das gewünschte Datenformat sind Paneldaten, da für dasselbe Unternehmen Daten für mehrere Jahre angegeben werden können. Der Benutzer wird aufgefordert, für jedes Unternehmen mindestens zwei aufeinander folgende Jahre einzugeben. Eine Template-Datei im CSV-Format, in der die Spaltennamen und die Reihenfolge der Spalten bereits implementiert sind, kann vom Benutzer heruntergeladen werden. Ein Info-Button erläutert die einzelnen Spalten. Außerdem muss der Anwender einen Benutzernamen eingeben, der den Datensatz einem bestimmten Nutzer zuordnet. Anschließend kann er den Datensatz zur weiteren Bearbeitung hochladen (vgl. Abbildung 15).

JAM Explanation System Home About Admin Personal Overview Logout

Upload a file

[Datei auswählen](#) Keine ausgewählt

ONLY ACCEPTS CSV files

Order of the CSV should be:

1. Company_ID (Integer) e.g. 8350117034
2. Jahr (Integer) e.g. 2021
3. Insolvenz (Integer) e.g. 0
4. Anlagevermoegen (Float) e.g. 13727000000.0
5. Umlaufvermoegen (Float) e.g. 114670000000.0
6. Bilanzsumme (Float) e.g. 128397000000.0
7. Eigenkapital (Float) e.g. 6788000000.0
8. kurzfristigeVerbindlichkeiten (Float) e.g. 91153000000.0
9. JahresueberschussFehlbetrag (Float) e.g. -4106000000.0

Please provide at least two consecutive years per company.

Insolvenz = 0 --> keine Insolvenz

Insolvenz = 1 --> Insolvenz

[Info](#)

Please download the template and fill it out:

[Download Template File](#)

Please enter the name of the user uploading the data:

Please upload the completed file:

[Upload](#)

You will be automatically forwarded to the next page once the file has been successfully uploaded

Abbildung 15: Daten-Upload

Anschließend kann der Benutzer seine gesammelten Uploads inklusive ihres Upload-Datums einsehen und den gewünschten Datensatz auswählen. Ein Fortschrittsbalken zeigt dem Nutzer den Fortschritt des Prozess-Workflows an (vgl. Abbildung 16).

JAM Explanation System Home About Admin Personal Overview Logout

Datasets used by user:

Superuser	User	Uploaded at	File
Joachim	Test	May 16, 2023, 3:44 p.m.	Insolvenzen_Django_o51o3c1.csv

[Change User](#)

Abbildung 16: Datensatz nach Nutzer

Wenn kein Upload gewünscht wird, kann der Anwender über die persönliche Übersicht die bereits hochgeladenen Datensätze einsehen und zur weiteren Bearbeitung auswählen.

Im nächsten Schritt kann der Benutzer deskriptive Informationen zum Datensatz einsehen. Diese beinhalten die ersten 5 Zeilen der Datei, eine statistische Zusammenfassung der Daten, die Anzahl der Solvenzen/Insolvenzen, die Anzahl der spezifischen Unternehmen, Jahre und Insolvenzen sowie die Anzahl der fehlenden Werte je Spalte (vgl. Abbildung 17).

JAM Explanation System
Home
About
Admin
Personal Overview
Logout

Datasets used by user:

Superuser	User	Uploaded at	File
Joachim	Test	May 16, 2023, 3:44 p.m.	Insolvenzen_Django_o51o3c1.csv

File Information

Superuser	User	File
Joachim	Test	Insolvenzen_Django_o51o3c1.csv

Dataframe

Only showing first five rows

	Company_ID	Jahr	Insolvenz	Anlagevermoege	Umlaufvermoege	Bilanzsumme	Eigenkapital	kurzfristigeVerbind
0	8350117034	2021	0	13727000000.00	114670000000.00	128397000000.00	6788000000.00	91153000000.00
1	8350117034	2020	0	13559000000.00	26663000000.00	40222000000.00	11188000000.00	16520000000.00
2	8350117034	2019	0	13985000000.00	29771000000.00	43756000000.00	11942000000.00	23320000000.00
3	8350117034	2018	0	14502000000.00	36103000000.00	50605000000.00	11445000000.00	30960000000.00
4	8350117034	2017	0	15467000000.00	27694000000.00	43161000000.00	12789000000.00	15733000000.00
5	8350117034	2016	0	18077000000.00	30794000000.00	48871000000.00	12803000000.00	19030000000.00

Statistics

Statistics of the whole Dataframe

	Company_ID	Jahr	Insolvenz	Anlagevermoegen	Umlaufvermoegen	Bilanzsumme	Eigenkapital	k
count	14529810.00	14529810.00	14529810.00	3421027.00	3466309.00	3418895.00	3424571.00	3
mean	5056985485.65	2016.50	0.00	10822238.82	8416692.16	19351010.49	7241408.67	4
std	2274847261.74	2.87	0.03	503467524.74	561604416.03	1017975954.78	295396841.51	3
min	56643.00	2012.00	0.00	-1307000.00	-862000.00	-24489000.00	-7438200000.00	-4
25%	2390502896.00	2014.00	0.00	26000.00	211000.00	377000.00	46000.00	0
50%	5130021831.00	2016.50	0.00	136000.00	583000.00	951000.00	251000.00	6
75%	7250018234.00	2019.00	0.00	765000.00	1618000.00	2968000.00	929000.00	4
max	8350402083.00	2021.00	1.00	22613200000.00	30207300000.00	49711400000.00	128783000000.00	1

Insolvenz Statistics

	Insolvenz
0	14520577
1	9233

Number of unique values

Company_ID	1452970
Jahr	10
Insolvenz	2

Number of null values

Company_ID	0
Jahr	0
Insolvenz	0
Anlagevermoegen	11108783
Umlaufvermoegen	11063501
Bilanzsumme	11110915
Eigenkapital	11105239
kurzfristigeVerbindlichkeiten	11074902
JahresueberschussFehlbetrag	14145138

Change User

Continue

Abbildung 17: Deskriptive Auswertung des rohen Datensatzes

6.3.5 Datenbereinigung

Die Datenbereinigung entspricht den in Abschnitt 6.1.1 beschriebenen Schritten, nachdem der Datensatz in ein Panelformat übertragen wurde. Dem Benutzer wird eine

Seite zur Verfügung gestellt, auf der Schritt für Schritt Erklärungen zum Bereinigungsprozess zu finden sind (vgl. Abbildung 18).

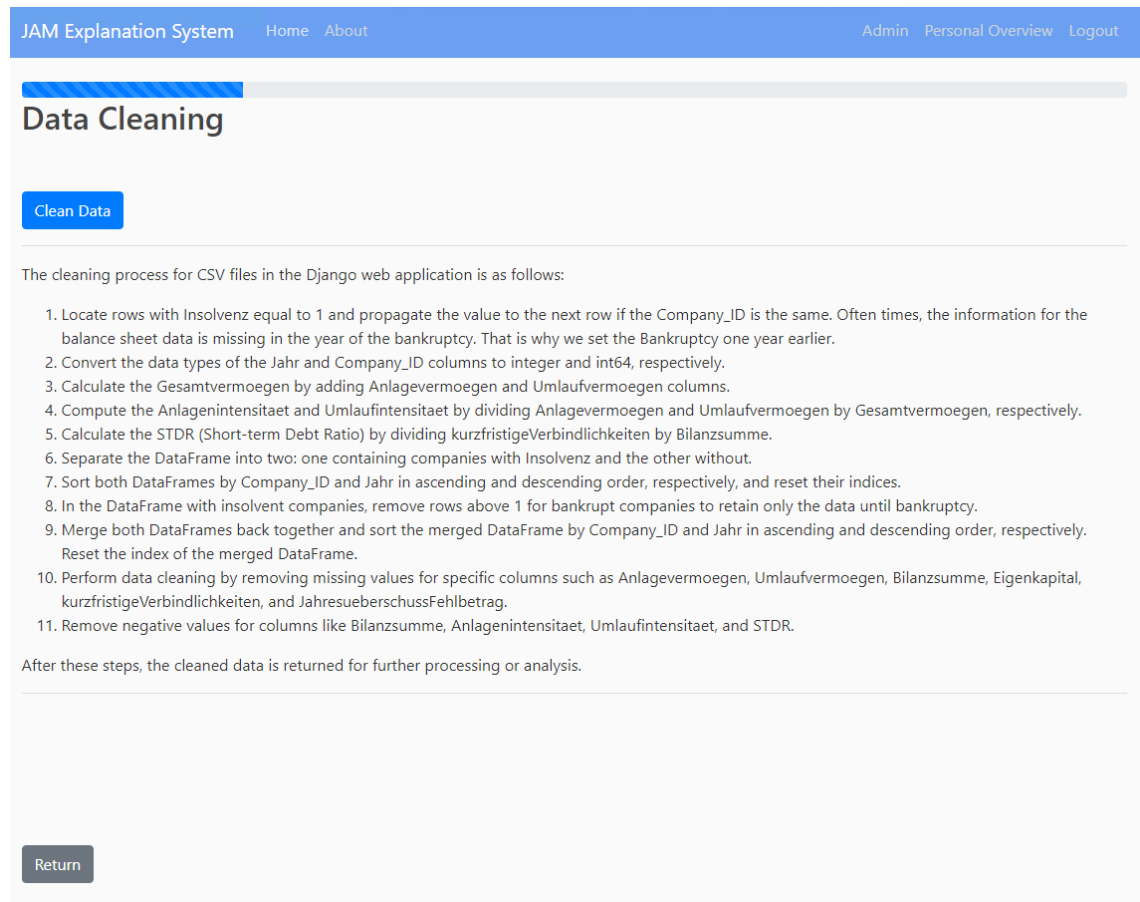


Abbildung 18: Hinweise zur Datenbereinigung

Für die bereinigten Daten wird dem Nutzer im nächsten Schritt die gleiche deskriptive Auswertung wie in 6.4.3, Abbildung 17, angezeigt.

6.3.6 Feature Engineering

Im Rahmen des Feature Engineerings wählt der Anwender die Merkmale aus, die als Grundlage für die Prognose verwendet werden sollen. Hier kann in Anlehnung an Tabelle 7 zwischen Eigenkapitalquote, kurzfristiger Verschuldungsquote, Working Capital Quote, Gesamtkapitalrentabilität, Eigenkapitalrentabilität, Anlagendeckungsgrad II und der Liquidität 2. Grades gewählt werden, wobei mindestens drei Merkmale ausgewählt werden müssen. Zusätzlich werden Info-Buttons eingeblendet, die dem Nutzer die Kennzahlen und deren Berechnung erläutern (vgl. Abbildung 10). Die Auswahl der Merkmale bezieht sich auf die im Mockup erstellte Merkmalsauswahl (vgl. Abbildung 3).

JAM Explanation System Home About Admin Personal Overview Logout

Select Features

Select All Deselect All

- ☐ Equity Ratio
- ☐ Working Capital Ratio
- ☐ Return on Total Assets
- ☐ Return on Equity
- ☐ Asset Coverage Ratio
- ☐ Second-Degree Liquidity
- ☐ Short-term Debt Ratio

Information about the equity ratio

Information about the working capital ratio

Information about the return on total assets

Information about the return on equity

Information about the asset coverage ratio

Information about the second-degree liquidity

Information about the short-term debt ratio

Return Continue

Abbildung 19: Selektion der Feature

Nach der Berechnung der Kenngrößen werden dem Benutzer wieder statistische Informationen über die Datenbasis wie in 6.4.3, Abbildung 17, angezeigt.

6.3.7 Daten-Segmentierung

Im nächsten Prozessschritt hat der Anwender die Möglichkeit, die Daten in Anlehnung an Abbildung 5 aus dem konzeptionellen Teil in Gruppen zu segmentieren oder die Daten ohne Segmentierung zu verwenden (vgl. Abbildung 20).

JAM Explanation System Home About Admin Personal Overview Logout

Group Processing

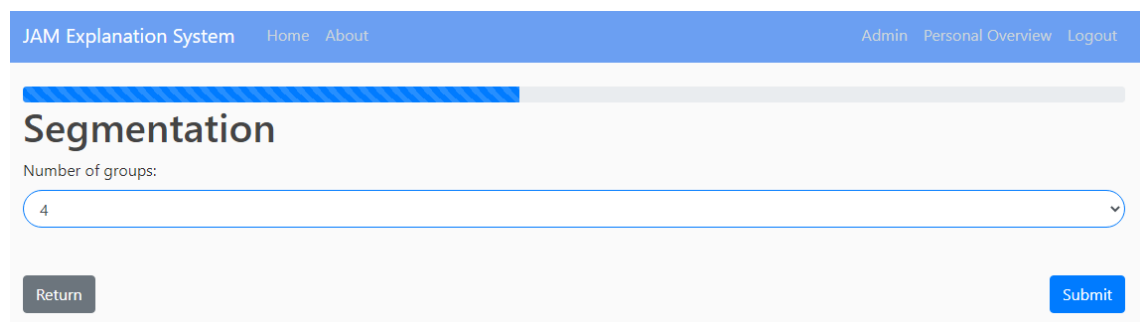
Segmentation of Data No Segmentation of Data

Return

Abbildung 20: Auswahl der Segmente

Bei gewählter Segmentierung ist festzulegen, in wie viele Gruppen der Datensatz unterteilt wird. Es können 2 bis 4 Gruppen ausgewählt werden (vgl. Abbildung 21).

Außerdem wählt der Benutzer einen Gruppennamen, eine Untergrenze und eine Obergrenze für jede Gruppe. Die Grenzen beziehen sich auf die Bilanzsumme der einzelnen Unternehmensdaten. Die Unter- und Obergrenze definieren somit den Bereich der Bilanzsumme, in dem ein Unternehmen liegen muss, um einer bestimmten Gruppe zugeordnet zu werden. Infobuttons erläutern dem Nutzer das Kriterium für die Einteilung des Datensatzes und verweisen auf die Klassifizierung nach § 267 HGB, die ebenfalls anhand der Bilanzsumme zwischen kleinen, mittelgroßen und großen Kapitalgesellschaften unterscheidet (vgl. Abbildung 22) (Handelsgesetzbuch, n.d.).



JAM Explanation System Home About Admin Personal Overview Logout

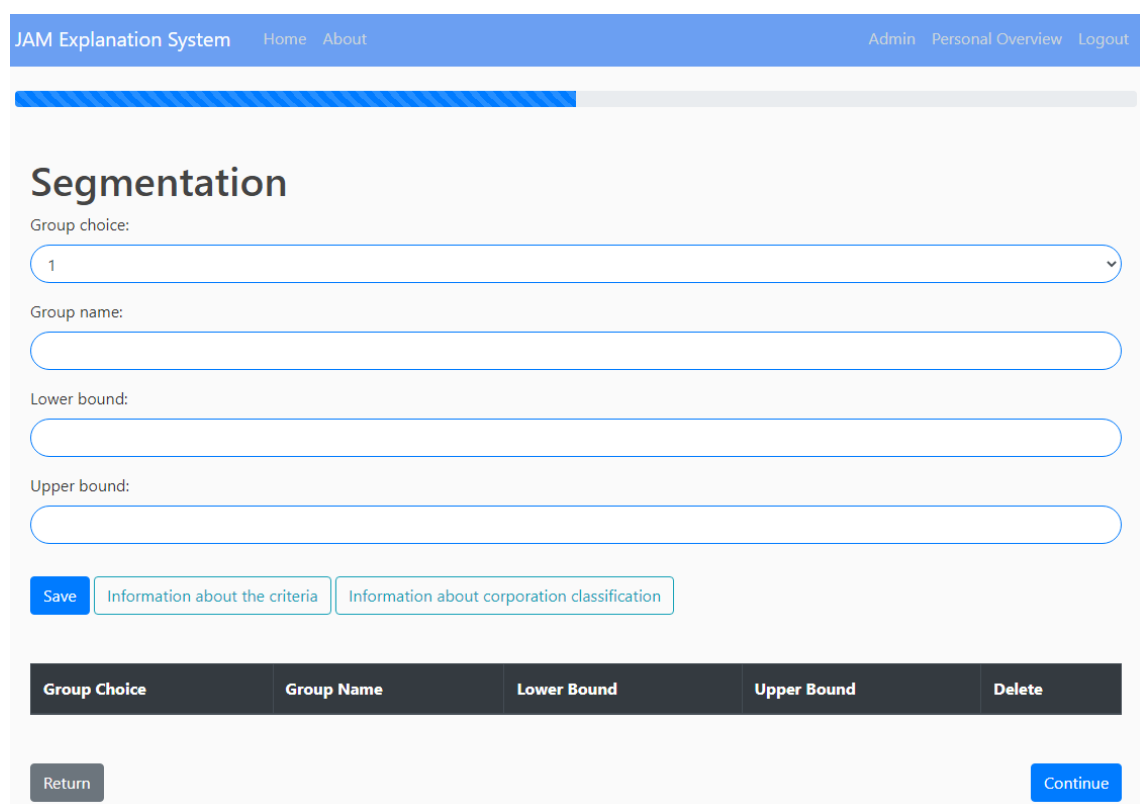
Segmentation

Number of groups:

4

Return Submit

Abbildung 21: Auswahl der Gruppengröße



JAM Explanation System Home About Admin Personal Overview Logout

Segmentation

Group choice:

1

Group name:

Lower bound:

Upper bound:

Save Information about the criteria Information about corporation classification

Group Choice	Group Name	Lower Bound	Upper Bound	Delete
--------------	------------	-------------	-------------	--------

Return Continue

Abbildung 22: Auswahl der Grenzen für die Gruppen

In Anlehnung an Abschnitt 6.3.3, Abbildung 17, findet der Benutzer anschließend statistische Informationen zu den Daten für die einzelnen Gruppen.

6.3.8 Modell-Setup und Optimierung

Nachdem die Segmentierung der Daten durchgeführt wurde, bietet die Anwendung die Auswahl der KI-Methode an. Wurde keine Datensegmentierung ausgewählt, überspringt man die Abbildungen 21 und 22 und landet direkt bei Abbildung 23. Für jede Gruppe wird zusätzlich der Einsatz von Hyperparameter-Tuning, das Verhältnis von Train-Test-Split und die Sampling-Methode vom Benutzer festgelegt. Mit Bezug auf 6.2.2. wird im Rahmen dieses Prototyps lediglich XGBoost als KI-Methode zur Verfügung gestellt. Im Backend könnten jedoch weitere KI-Methoden hinzugefügt werden, da die Anwendung modular aufgebaut ist. Als Hyperparameter Tuning wird Bayesian Optimization angeboten. Diese Funktion bietet eine automatisierte Methode, um die besten Hyperparameter für das spezifische Modell zu finden. Als Train-Test-Split-Verhältnisse werden 80/20, 70/30 und 60/40 zur Wahl gestellt. Zusätzlich kann zwischen den Sampling-Methoden SMOTE, SMOTE-Tomek und SMOTE-ENN selektiert werden, wenn keine Bayesian Optimization ausgewählt wurde. Neben der Ausführung der gewählten Optionen wird in diesem Schritt der Panel-Datensatz in ein Wide-Format übertragen. Dies bezieht sich auf Erkenntnisse aus Abschnitt 6.2, in dem festgestellt wurde, dass im Rahmen der Insolvenzprognose mit XGBoost das Wide-Format zu besseren Ergebnissen führt. Infobuttons erläutern die einzelnen Auswahlmöglichkeiten und die weitere Datenverarbeitung (vgl. Abbildung 23). Zur besseren Übersicht wurden die Auswahl der Oversampling-Techniken und die Auswahl der KI-Methode im Vergleich zu den konzeptionellen Abbildungen 6 und 7 auf einer Seite zusammengefasst.

JAM Explanation System

Home About

Admin Personal Overview Logout

Information about the XGBoost Machine Learning Algorithm

Information about Train Test Validation Split

Information about Bayesian Optimization

Information about Sampling Methods

Information about the data transformation from long to wide

Selection of the AI Algorithm

Train-Test-Validation Split

Select AI Method

XG-Boost

Large

Bayesian Optimization

No

Train-Test-Validation Split

60-40-0

Sampling Method

SMOTE

Medium

Bayesian Optimization

No

Train-Test-Validation Split

60-40-0

Sampling Method

SMOTE

Small

Bayesian Optimization

No

Train-Test-Validation Split

60-40-0

Sampling Method

SMOTE

Return

Continue

Abbildung 23: Modell-Setup

Die statistischen Informationen über die Trainings- und Testdatensätze für jede Gruppe werden ähnlich wie in Abbildung 7 dargestellt. Wenn keine Hyperparameteroptimierung

ausgewählt wurde, hat der Benutzer die Möglichkeit, ausgewählte Hyperparameter pro Gruppe für die Modelle manuell zu definieren. Infobuttons erläutern diese Parameter. Die Aufspaltung der Datensätze wird hier nicht wie in der konzeptionellen Abbildung 8 integriert, da sie dem Benutzer bereits im Rahmen des Modell-Setups zur Auswahl steht (vgl. Abbildung 23). Anschließend beginnt die Generierung des Modells bzw. der Modelle je nach Segmentierung (vgl. Abbildung 24).

JAM Explanation System
Home
About
Admin
Personal Overview
Logout

XGBoost Parameters

Small

Learning rate	What is a learning rate?	0,300000012
Max depth	What is a max depth?	6
Number of estimators	What is the number of estimators?	100
Min Child Weight	What is the min child weight?	0,5
Gamma	What is gamma?	1
Subsample	What is a subsample?	0,8
Colsample bytree	What is a col sample by tree?	0,8
Scale Pos Weight	What is the scale pos weight?	1

Medium

Learning rate	What is a learning rate?	0,300000012
Max depth	What is a max depth?	6
Number of estimators	What is the number of estimators?	100
Min Child Weight	What is the min child weight?	0,5
Gamma	What is gamma?	1
Subsample	What is a subsample?	0,8
Colsample bytree	What is a col sample by tree?	0,8
Scale Pos Weight	What is the scale pos weight?	1

Parameter	Question	Value
Learning rate	What is a learning rate?	0,300000012
Max depth	What is a max depth?	6
Number of estimators	What is the number of estimators?	100
Min Child Weight	What is the min child weight?	0,5
Gamma	What is gamma?	1
Subsample	What is a subsample?	0,8
Colsample bytree	What is a col sample by tree?	0,8
Scale Pos Weight	What is the scale pos weight?	1

Return Save

Abbildung 24: Parameterauswahl

6.3.9 Modell-Beurteilung

Nach der Modellgenerierung werden dem Benutzer die Modellergebnisse entsprechend den konzeptionellen Entwürfen in Abbildung 9 und 10 präsentiert. Die Metriken zur Bewertung der Modellqualität werden in tabellarischer Form und nach Gruppen geordnet angezeigt. Die Metriken, die gemäß Abschnitt 2.3.2 ausgewählt wurden, sind Genauigkeit (Accuracy), F1-Score, Präzision (Precision), Brier Score, Sensitivität (Recall) und ROC-AUC. Diese Ergebnisse können für den Benutzer auch im CSV-Format exportiert werden (vgl. Abbildung 25). Die Quantifizierung der Schätzgüte des Modells ist von hoher Bedeutung, um die in dieser Arbeit gestellten Forschungsfragen beantworten zu können und die Anforderungen aus der Anforderungsanalyse zu erfüllen. Die Implementierung stellt daher die definierten Kennzahlen zur Schätzgüte in einer Tabelle dar, welche dem Nutzer ermöglicht für jede der definierten Gruppen - sofern der Datensatz segmentiert wurde - einzusehen, wie hoch die Schätzgüte des Prognoseergebnisses ist. Damit kann der Nutzer die Genauigkeit der Modellvorhersagen bewerten.

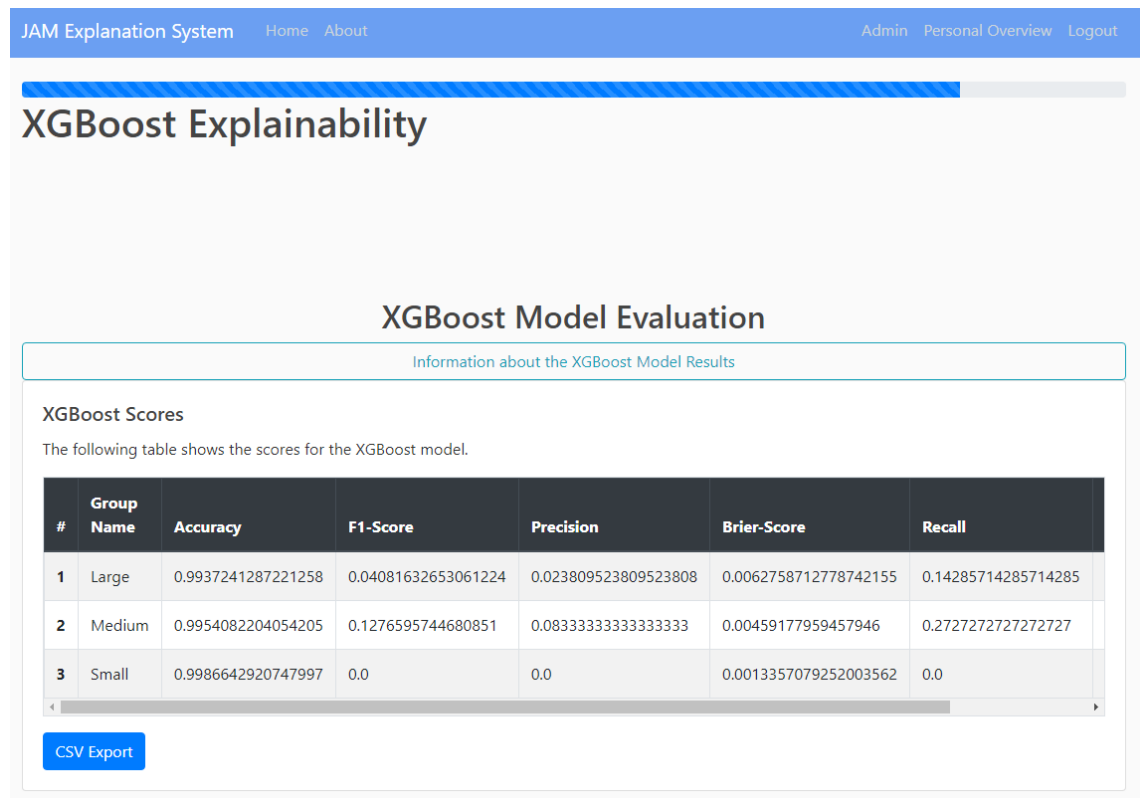


Abbildung 25: Metrik zur Beurteilung der Güte des Modells

Informationen über die Hyperparameter der Modelle pro Gruppe werden dem Benutzer zur Verfügung gestellt. Hier finden sich auch Parameter, die in der Parameterauswahl in Abbildung 24 nicht zur Auswahl standen, da der Benutzer so auch die Standardwerte der Hyperparameter bei der XGBoost-Modellgenerierung einsehen und somit das Modell vollkommen transparent nachvollzogen werden kann. Abbildung 26 zeigt einen Ausschnitt der angezeigten Hyperparameter.

XGBoost Model Details

Information about the XGBoost Model Specification

Model Details - Group Large

Hyperparameter	Value
objective	binary:logistic
base_score	0.5
booster	gbtree
colsample_bylevel	1
colsample_bynode	1
colsample_bytree	0.8
eval_metric	aucpr
gamma	1
gpu_id	-1
grow_policy	lossguide
interaction_constraints	
learning_rate	0.300000012

Model Details - Group Medium

Hyperparameter	Value
objective	binary:logistic
base_score	0.5
booster	gbtree
colsample_bylevel	1
colsample_bynode	1
colsample_bytree	0.8
eval_metric	aucpr
gamma	1
gpu_id	-1
grow_policy	lossguide
interaction_constraints	
learning_rate	0.300000012

Model Details - Group Small

Hyperparameter	Value
objective	binary:logistic
base_score	0.5
booster	gbtree
colsample_bylevel	1
colsample_bynode	1
colsample_bytree	0.8
eval_metric	aucpr
gamma	1
gpu_id	-1
grow_policy	lossguide
interaction_constraints	
learning_rate	0.300000012

Abbildung 26: Ausschnitt der Hyperparameter je Modell

Um den Einfluss der einzelnen Merkmale auf die Prognoseergebnisse transparent darzustellen, wurde zunächst die XGBoost Feature Importance Methode genutzt. Da das Modelltraining bereits mit XGBoost durchgeführt wurde, erschien die Anwendbarkeit der Feature Importance Methode als besonders sinnvoll. Eine Eigenschaft dieser Methode ist, dass sie irrelevante Features ignoriert, was bedeutet, dass die Anzahl an Merkmalen, die für eine Gruppe innerhalb einer Modellinstanz dargestellt werden, variieren kann. Die Darstellung dieser Methode und ihrer Ergebnisse kann exemplarisch in Abbildung 27 nachvollzogen werden.

XGBoost Feature Importance

Information about the XGBoost Feature Importance

XGBoost Feature Importance - Group Large

Feature	Importance
1. RTA(t0)	0.25424435952811825
2. ROE(t0)	0.23054754975848926
3. RTA(t-1)	0.1286870898964941
4. STDR(t0)	0.07976000689313308
5. WCR(t-1)	0.07353447672986572
6. ACR(t0)	0.07104598145558613
7. L2(t-1)	0.03917206512708136
8. WCR(t0)	0.027739666848204864
9. ER(t-1)	0.025195954113486674

XGBoost Feature Importance - Group Medium

Feature	Importance
1. ACR(t0)	0.3807291461968256
2. RTA(t0)	0.20061673220639975
3. WCR(t0)	0.07801428679262631
4. STDR(t-1)	0.0636529860962187
5. STDR(t0)	0.05837973963987933
6. ER(t0)	0.04371672749055373
7. ROE(t0)	0.040520148097196375
8. RTA(t-1)	0.026130016277317637
9. ROE(t-1)	0.024354996966994186

XGBoost Feature Importance - Group Small

Feature	Importance
1. RTA(t0)	0.20024666536625102
2. ER(t-1)	0.1809723075570357
3. ACR(t0)	0.15552767606218007
4. STDR(t0)	0.12376565971754618
5. L2(t-1)	0.12004842127842204
6. RTA(t-1)	0.08704413700850507
7. ER(t0)	0.03156318038998663
8. ROE(t-1)	0.027649989527334787
9. ACR(t-1)	0.027054115483035227

Abbildung 27: Ausschnitt der XGBoost Feature Importance

Darüber hinaus wurde die SHAP-Bibliothek im Rahmen der Erklärbarkeitsmethoden eingesetzt, um die KI-basierte Insolvenzprognose verständlicher zu gestalten. Mit Hilfe von Shapley Additive Explanations wird dem Nutzer veranschaulicht, welche Merkmale einen besonders hohen Einfluss auf die Gesamtprognose haben. Ebenso wird dargelegt, auf welche Weise – im Gegensatz zur Darstellung durch die XGBoost Feature Importance – höhere und niedrigere Merkmalsausprägungen einen positiven oder negativen Einfluss auf die Prognose nehmen.

Um den Benutzer bei der Auswertung der Vorhersageergebnisse visuell zu unterstützen, wurden drei verschiedene Diagramme verwendet. Der Bar Summary Plot, der die Auswirkungen der Merkmale auf die Prognose als Balkendiagramm aufzeigt. Die Reihenfolge der Merkmale auf der vertikalen Achse zeigt an, welche Merkmale am wichtigsten für die Vorhersage sind (vgl. Abbildung 28). Dieser und der Scatter Summary Plot sind Implementierungen der genannten Summary Plots. Der Scatter Summary Plot stellt die Beziehungen zwischen den Merkmalswerten und der Vorhersage dar. Die SHAP-Werte, die in diesem Zusammenhang den Einfluss der Merkmale auf die Prognose messen, wurden entsprechend der binären Klassifikation in unserem Modell zwischen 0 und 1 skaliert. Es kann davon ausgegangen werden, dass höhere Werte eine stärkere

Beziehung zur positiven Klasse (insolvent = 1) signalisieren. Die Farbe der Punkte in der Grafik kann somit Informationen über die Richtung des Effekts liefern. Beispielsweise kann eine rote Einfärbung bedeuten, dass höhere Werte dieses Merkmals die Wahrscheinlichkeit einer Insolvenz erhöhen (vgl. Abbildung 29). Der Dependence Plot untersucht den Einfluss eines bestimmten Merkmals auf die Vorhersage unter Berücksichtigung der Werte eines anderen Merkmals (vgl. Abbildung 30). In der Implementierung wurde sich dazu entschieden, dass das betrachtete Merkmal mit dem einflussreichsten Merkmal laut XGBoost Feature Importance übereinstimmt.

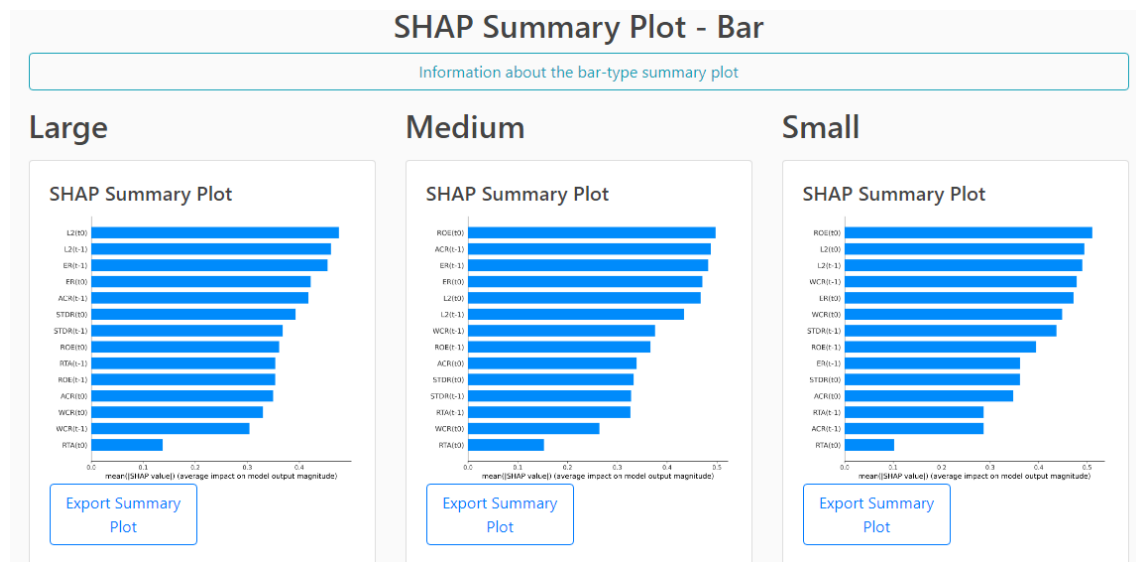


Abbildung 28: SHAP - Bar Summary Plot

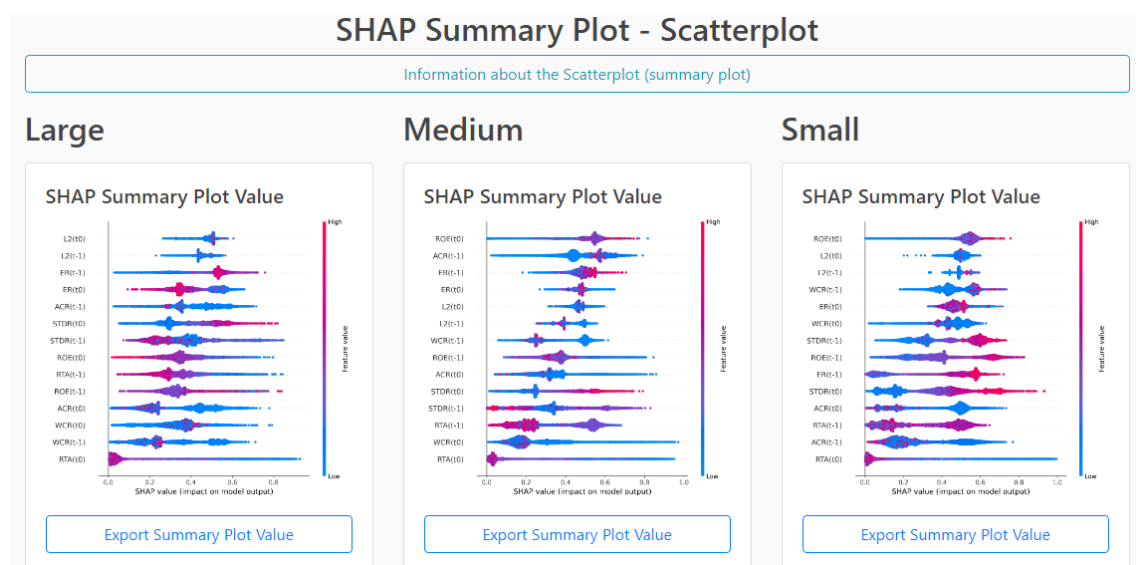


Abbildung 29: SHAP - Scatter Summary Plot

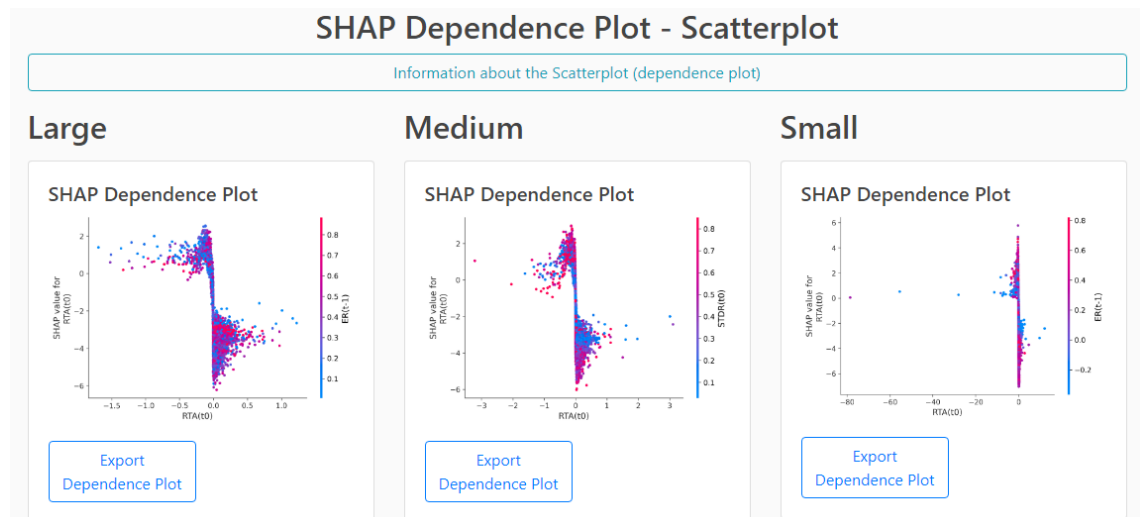


Abbildung 30: SHAP - Dependence Plot

Die Prognoseergebnisse werden im Dashboard zusätzlich mittels Konfusionsmatrizen visualisiert. Somit kann der Nutzer des Dashboards intuitiv beurteilen, wie aussagekräftig das Modell für seine Zielsetzung ist. In einer Konfusionsmatrix werden die vorhergesagten Klassen den tatsächlichen Klassen gegenübergestellt. Der Benutzer sieht auf einen Blick die Anzahl der True Positives, True Negatives, False Positives und False Negatives (vgl. Abbildung 31). Es wird also ermöglicht, den Anteil korrekter Vorhersagen visuell nachzuvollziehen.

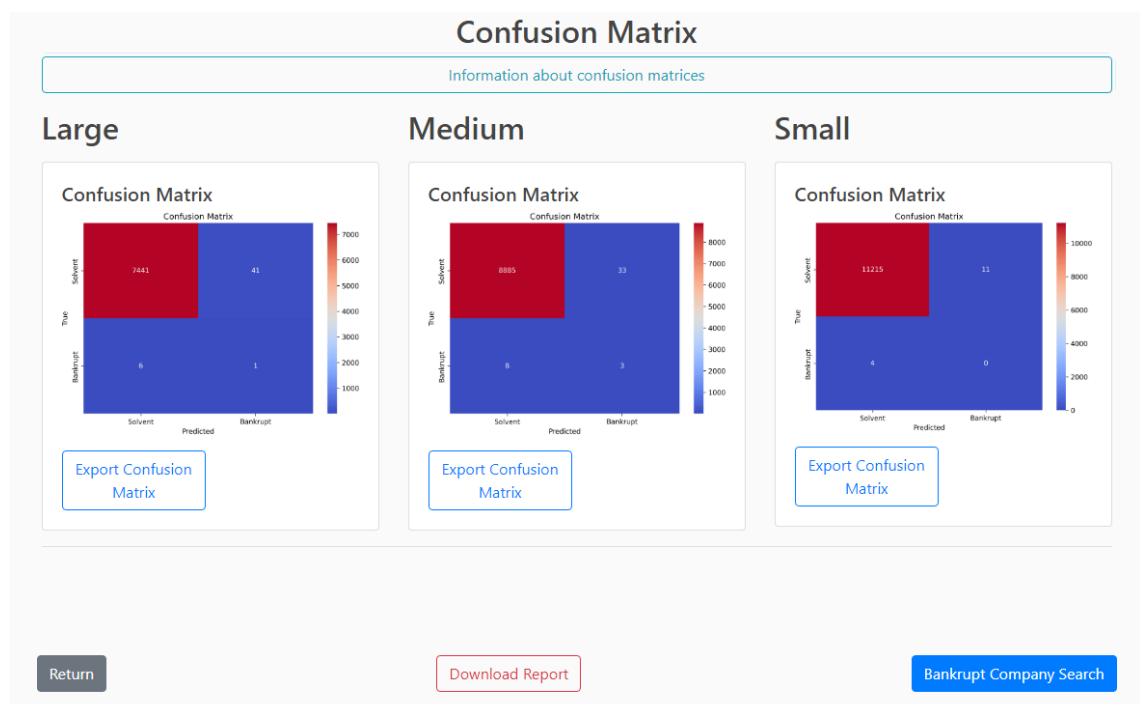


Abbildung 31:Konfusionsmatrix

Info-Buttons sind für jede Tabelle und Visualisierung verfügbar und erläutern diese. SHAP-Diagramme und Konfusionsmatrizen können als Bilder exportiert werden. Zusätzlich kann der Benutzer einen Bericht im PDF-Format herunterladen, der alle oben genannten Informationen zur Modellbewertung enthält.

Abschließend werden dem Anwender die Prognoseergebnisse des Testdatensatzes im Dashboard präsentiert, wobei die einzelnen Prognosen mit dem tatsächlichen Insolvenzstatus verglichen werden. Diese Visualisierung erleichtert den Nutzern und Regulatoren die Nachvollziehbarkeit der Ergebnisse. Für jedes Unternehmen sind darüber hinaus Informationen zu Metriken, Unternehmens-ID und Gruppenzugehörigkeit abrufbar.

Diese Filter-/Suchfunktion erweitert die Nutzbarkeit des Dashboards erheblich. Sie ermöglicht es den Anwendern, für jede im Testdatensatz enthaltene Unternehmens-ID sowohl das Prognoseergebnis als auch den tatsächlichen Insolvenzstatus und die relativen finanziellen Kennzahlen einzusehen. Durch diese Erklärbarkeitsmethoden können Nutzer prüfen, welchen Einfluss bestimmte Kennzahlen auf das Prognoseergebnis eines Unternehmens hatten.

Zur weiteren Verwendung oder für eine Weitergabe an Dritte ermöglicht das Dashboard, diese Informationen und Ergebnisse als CSV-Datei zu exportieren (vgl. Abbildung 32).

The screenshot displays the 'JAM Explanation System' dashboard. At the top, a blue navigation bar contains links for 'Home', 'About', 'Admin', 'Personal Overview', and 'Logout'. The main content area has a light gray background with a blue header bar. Below this, the title 'Filter bankruptcy predictions for any company out of your test data set' is centered. A text input field prompts the user to 'Enter a company ID to filter the bankruptcy predictions table for that company based on the test data set.' Below the input field is a blue 'Filter' button. At the bottom of the main content area, there are three buttons: 'Show Bankrupt Companies', 'Show Correct Bankruptcy Predictions', and 'Show Predicted Bankrupt Companies'. A dark gray bar at the very bottom contains a 'Remove Filters' button.

Bankruptcy Predictions:

Group	Company ID	Predicted Bankruptcy	Actual Bankruptcy	ER(t-1)	WCR(t-1)	RTA(t-1)	ROE(t-1)
Medium	2250010722	yes	yes	-0.1620604310611529	1.1489685726152568	0.0047952975146952	-0.029589564110
Medium	4090047419	yes	yes	0.5670573907748516	1.5273039799021553	-0.2308316841731364	-0.407069351230
Medium	5350066682	yes	yes	-0.0969633780789782	0.7821114369501466	-0.1728137625439854	1.782258064516
Large	3030028173	yes	yes	0.3275238879146834	1.2493021772071138	-0.0548616827112335	-0.167504370629

Return Download CSV Save to Personal Overview

Abbildung 32: Filterfunktion für einzelne Vorhersagen

Im letzten Schritt gelangt der Benutzer, wie bereits in Abbildung 14 beschrieben, zur persönlichen Übersicht der Datensätze und KI-Modelle. Hier befinden sich die gespeicherten KI-Modelle und Datensätze, die der Benutzer jederzeit wieder aufrufen kann. Bei Auswahl eines KI-Modells wird er wieder auf die Seite mit den Informationen zur Modellbewertung weitergeleitet. Es ist jedoch auch möglich, zu vorherigen Schritten des Prozessworkflows zurückzukehren und Änderungen am Modell oder Datensatz vorzunehmen.

7 Diskussion der Ergebnisse

Kapitel 7 befasst sich mit dem Diskussionsteil unserer Arbeit und erläutert die Ergebnisse unserer Forschungsfragen im Zusammenhang mit der KI-basierten Insolvenzprognose. Abschnitt 7.1 behandelt die Bedeutung von Erklärbarkeit bei der Verbesserung dieser Prognosen. Hierbei werden theoretische und praktische Aspekte sowie konkrete Ergebnisse analysiert. Der nächste Teil (7.2) beschäftigt sich mit der Analyse des Dashboards als Instrument der Konfigurierbarkeit und Erklärbarkeit, indem seine grundlegenden und erweiterten Funktionen beleuchtet werden. Diese Abschnitte bieten eine detaillierte Diskussion und Analyse der Forschungsergebnisse und zeigen, wie Erklärbarkeit und Konfigurierbarkeit dazu beitragen können, die Zuverlässigkeit und Qualität von KI-gestützten Insolvenzprognosen zu verbessern.

7.1 Die Rolle der Erklärbarkeit bei der Verbesserung KI-gestützter Insolvenzprognosen

In diesem Abschnitt wird die erste Forschungsfrage adressiert, die sich auf die Verbesserung der Erklärbarkeit von Machine Learning Modellen im Rahmen der KI-gestützten Insolvenzprognose bezieht. Das Kapitel ist in drei Unterkapitel unterteilt, die sich jeweils auf unterschiedliche Aspekte der Erklärbarkeit konzentrieren.

Das Unterkapitel 7.1.1 setzt sich damit auseinander, wie theoretische Aspekte der Erklärbarkeit sowie die Methoden SHAP, Konfusionsmatrizen, Kennzahlen zur Schätzgüte und die XGBoost Feature Importance eingebunden werden.

Das Unterkapitel 7.1.2 hingegen konzentriert sich auf die praktische Anwendung und Nutzbarkeit der implementierten Erklärbarkeitsmethoden. Hier werden Aspekte wie die Such- bzw. Filterfunktion, die transparente Kommunikation der Modellparameter und die Möglichkeit, Informationen, Prognoseergebnisse und Visualisierungen zu exportieren, gesondert hervorgehoben.

Das Unterkapitel 7.1.3 befasst sich schlussendlich mit den konkreten Prognoseergebnissen und Resultaten der Interpretierbarkeitsmethoden am Beispiel des in dieser Arbeit verwendeten Datensatzes. In diesem Unterkapitel wird auf konkrete Resultate unterschiedlicher Modellinstanzen bezüglich ihrer Vorhersagegenauigkeit, der Gewichtung und Bedeutung einzelner Einflussfaktoren und auf die Robustheit dieser

beobachtbaren Ergebnisse eingegangen. Dies ermöglicht es, die Leistungsfähigkeit der eingesetzten Methoden umfassend zu beurteilen.

Im Zusammenspiel bilden die drei Unterkapitel eine ganzheitliche Betrachtung, die erforscht, wie die Erklärbarkeit von Machine Learning Modellen zur Steigerung der Zuverlässigkeit KI-gestützter Insolvenzprognosen führen kann und wie sich diese basierend auf der Implementierung in dieser Arbeit tatsächlich ausprägt. Sie zeigen, dass sowohl auf theoretischer als auch auf praktischer Ebene eine erhebliche Verbesserung erzielt werden kann, die die Qualität und Akzeptanz der KI-gestützten Insolvenzprognose verbessern. Gleichzeitig verweist das dritte Unterkapitel darauf, dass die Prognoseergebnisse anhand der in dieser Arbeit genutzten Daten in ihrer Aussagekraft begrenzt sind und stellt damit deutlich die Bedeutung einer hohen Datenqualität sowie einer ausreichend hohen Anzahl an Insolvenzen für eine verlässliche Prognose und damit zusammenhängend auch zielführende Interpretierbarkeit heraus (H. Kim et al., 2020).

7.1.1 Anforderungsanalyse und Implementierung von Erklärbarkeitsmethoden

Die Ergebnisse dieser Arbeit zeigen, dass verschiedene Verfahren zur Verbesserung der Erklärbarkeit KI-gestützter Insolvenzprognose erfolgreich implementiert wurden und dazu beitragen können, die Ergebnisse und ihre Interpretierbarkeit maßgeblich zu optimieren.

Eines der implementierten Verfahren ist die Verwendung von SHAP-Plots (Lundberg et al., 2017). Diese ermöglichen eine intuitive Visualisierung der Beiträge einzelner Features zur Modellvorhersage. Sie tragen dazu bei, den Einfluss bestimmter Merkmale auf die Vorhersage der Insolvenz von Unternehmen im Datensatz nachvollziehbar zu machen.

Ein weiteres Verfahren zur Verbesserung der Erklärbarkeit sind Konfusionsmatrizen. Sie visualisieren, wie viele Vorhersagen korrekt bzw. inkorrekt getroffen wurden und auf welche Weise. Darüber hinaus ermöglichen sie die Berechnung verschiedener Kennzahlen zur Schätzgüte, wie beispielsweise der Präzision (Precision), der Sensitivität (Recall) und des F1-Scores. Zusätzlich werden auch die Metriken ROC-AUC und Brier Score eingebunden, um die Schätzgüte des Modells detailliert beurteilen zu können.

Die XGBoost Feature Importance (Chen & Guestrin, 2016) stellt eine Rangliste der Features hinsichtlich ihres Beitrags zur Modellvorhersage dar. Dies verbessert die Erklärbarkeit des Modells, indem hervorgehoben wird, welche Merkmale besonders relevant für die Insolvenzprognose sind.

Zusätzlich trägt die transparente Kommunikation sämtlicher dem Modell zu Grunde liegenden Parameter zur Erklärbarkeit bei. Die Offenlegung dieser Parameter ermöglicht es Nutzern, zu verstehen, wie das Modell konfiguriert ist und welche Einstellungen zur erreichten Prognoseleistung beitragen.

Zusammenfassend tragen also der Einsatz der benannten Erklärbarkeitsmethoden und die transparente Darstellung der Modellkonfiguration zur Erklärbarkeit des respektiven Modells und damit zur Qualität der KI-gestützten Insolvenzprognose bei.

7.1.2 Anwendung und Nutzbarkeit der Erklärbarkeitsansätze in der Praxis

Neben der theoretischen Verbesserung der Erklärbarkeit von Machine Learning Modellen durch die in Kapitel 7.1.1 aufgegriffenen Verfahren, ist es weiterhin wichtig, diese Ansätze in der praktischen Anwendung nutzbar zu machen. Die implementierten Verfahren zur Verbesserung der Erklärbarkeit der Ergebnisse KI-gestützter Insolvenzprognose lassen sich anhand verschiedener Aspekte des erarbeiteten Prototyps veranschaulichen.

Ein wichtiger Aspekt, der zu dieser Verbesserung beiträgt, ist die implementierte Such- bzw. Filterfunktion nach einzelnen Unternehmen sowie nach insolventen, insolvent vorhergesagten und korrekt insolvent vorhergesagten Unternehmen. Diese Funktion ermöglicht es, Prognosen und ihre Genauigkeit für bestimmte Unternehmen gezielt zu überprüfen und zu analysieren. Darüber hinaus kann der Nutzer sämtliche Merkmalsausprägungen, die Gruppenzugehörigkeit sowie die Unternehmens-ID nachvollziehen, was die Transparenz und Verständlichkeit der Prognoseergebnisse weiter erhöht.

Die überprüfbare und transparente Kommunikation sämtlicher Parameter, die bereits zuvor aufgegriffen wurde, trägt auch hier zur Erklärbarkeit für den Nutzer bei. Durch deren Offenlegung ist der Anwender dazu in der Lage, zu verstehen, welche Einstellungen zur erreichten Prognoseleistung des Modells geführt haben.

Darüber hinaus ist die Möglichkeit entscheidend, sämtliche Informationen, Prognoseergebnisse und Visualisierungen zu exportieren. Dies erhöht die Nutzbarkeit der Informationen sowohl für den Nutzer als auch hinsichtlich regulatorischer Vorgaben und ermöglicht die Verwendung und Weitergabe in unterschiedlichen Kontexten.

Somit lässt sich sagen, dass sowohl auf theoretischer als auch auf praktischer, anwenderbezogener Seite eine Verbesserung der KI-gestützten Insolvenzprognose bzw. ihrer Erklärbarkeit erreicht wird. Die Nutzbarkeit und Verständlichkeit der Prognoseergebnisse wird erhöht, was eine gezielte und fundierte Entscheidungsfindung auf Basis der bereitgestellten Informationen ermöglicht.

7.1.3 Erreichte Prognoseergebnisse und Resultate der Interpretierbarkeitsmethoden

Ziel dieses Kapitels ist es, die Leistungsfähigkeit der eingesetzten Methoden anhand der erzielten Prognosegüte sowie der Interpretierbarkeit der Modelle zu analysieren. Hierfür werden zunächst die Kennzahlen zur Messung der Vorhersagegüte der einzelnen Modelle betrachtet. Anschließend wird analysiert, inwiefern die Gewichtung der Eingabemerkmale aussagekräftige Erkenntnisse über die Bedeutung einzelner Faktoren liefert.

Die Bewertung der Prognosegüte erfolgt anhand der bereits in Kapitel 2.3.2 beschriebenen Kennzahlen zur Messung der Schätzgüte. Die Ergebnisse für die mit Bayesian Optimization und einem 70/30-Split optimierten Modelle sind in Tabelle 16 ersichtlich. Die mittelgroßen Kapitalgesellschaften weisen in Hinblick auf den F1-Score die höchste Prognosegüte auf, wenngleich die Werte insgesamt nicht zufriedenstellend sind. Wie schon in Kapitel 6.2 diskutiert, scheint die geringe Gesamtzahl an Insolvenzfällen (mit verfügbaren Finanzinformationen) eine Limitierung für bessere Prognoseergebnisse darzustellen, was sich auch in den Konfusionsmatrizen in der Tabelle widerspiegelt. Nichtsdestotrotz geben die Kennzahlen sowie die Konfusionsmatrizen eindeutige Hinweise auf die Vorhersagequalität der verschiedenen Modelle. Mithilfe der in Kapitel 6.3.9 vorgestellten Filterfunktion für einzelne Prognosen können die korrekten Vorhersagen bzw. prognostizierten Insolvenzen unter Berücksichtigung der Kennwerte näher analysiert werden. Somit ist eine Nachvollziehbarkeit spezifischer Prognosen durch den Nutzer möglich.

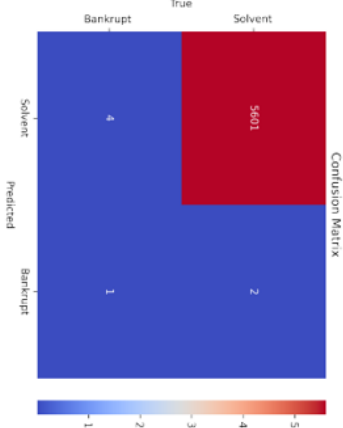
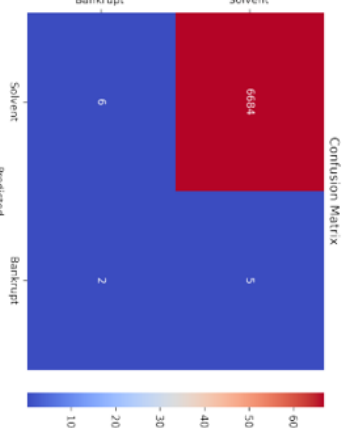
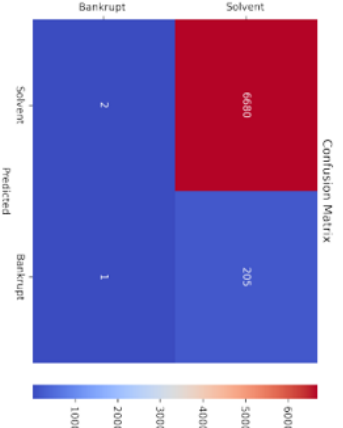
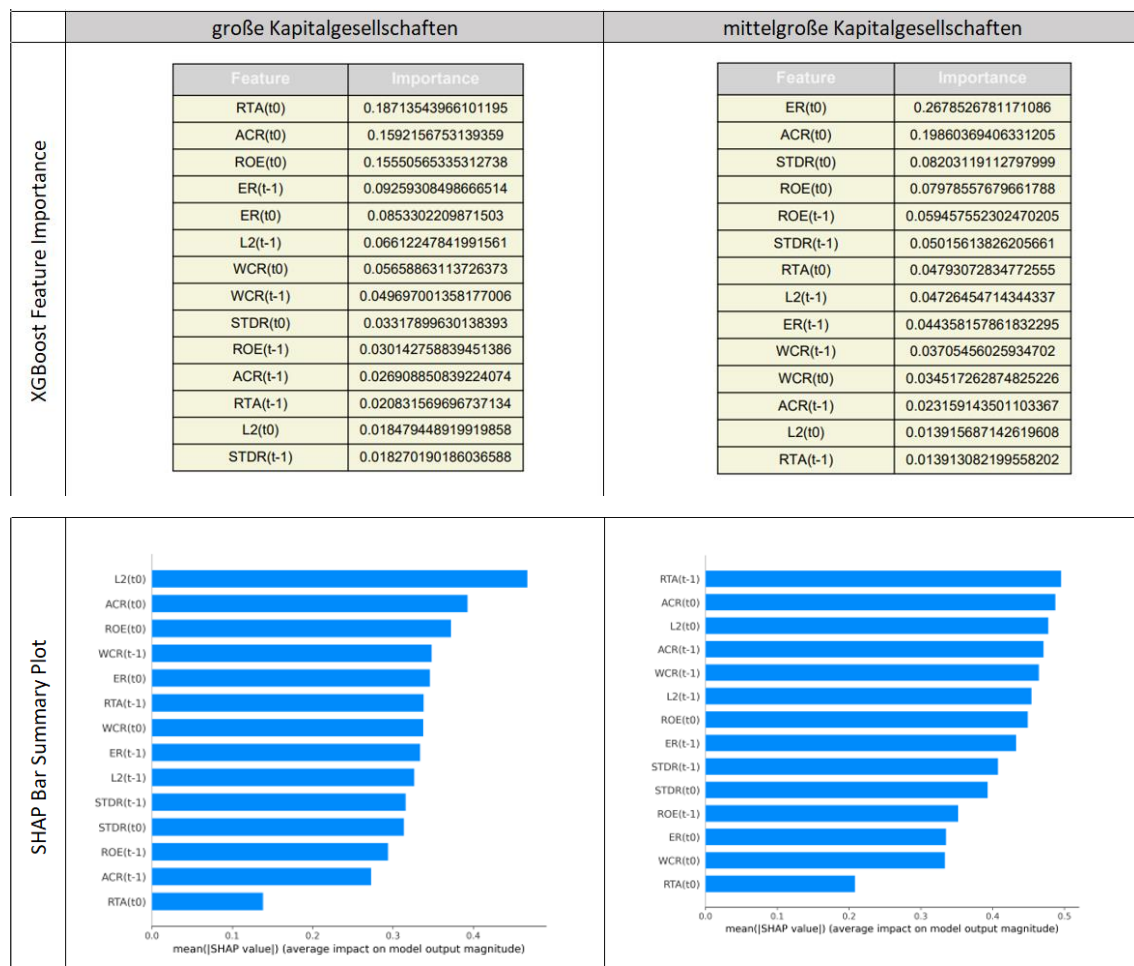
große Kapitalgesellschaften	mittelgroße Kapitalgesellschaften	kleine Kapitalgesellschaften																																										
<table><tr><th>Metric</th><th>Value</th></tr><tr><td>Accuracy</td><td>0.9989300998573466</td></tr><tr><td>F1-score</td><td>0.25</td></tr><tr><td>Precision</td><td>0.3333333333333333</td></tr><tr><td>Recall</td><td>0.2</td></tr><tr><td>ROC AUC</td><td>0.5998215241834731</td></tr><tr><td>Brier Score</td><td>0.001069900142653524</td></tr></table> 	Metric	Value	Accuracy	0.9989300998573466	F1-score	0.25	Precision	0.3333333333333333	Recall	0.2	ROC AUC	0.5998215241834731	Brier Score	0.001069900142653524	<table><tr><th>Metric</th><th>Value</th></tr><tr><td>Accuracy</td><td>0.998357473495595</td></tr><tr><td>F1-score</td><td>0.2666666666666666</td></tr><tr><td>Precision</td><td>0.2857142857142857</td></tr><tr><td>Recall</td><td>0.25</td></tr><tr><td>ROC AUC</td><td>0.6246262520556136</td></tr><tr><td>Brier Score</td><td>0.0016425265044049575</td></tr></table> 	Metric	Value	Accuracy	0.998357473495595	F1-score	0.2666666666666666	Precision	0.2857142857142857	Recall	0.25	ROC AUC	0.6246262520556136	Brier Score	0.0016425265044049575	<table><tr><th>Metric</th><th>Value</th></tr><tr><td>Accuracy</td><td>0.9699477351916377</td></tr><tr><td>F1-score</td><td>0.009569377990430622</td></tr><tr><td>Precision</td><td>0.0048543689320388345</td></tr><tr><td>Recall</td><td>0.3333333333333333</td></tr><tr><td>ROC AUC</td><td>0.6517792302106027</td></tr><tr><td>Brier Score</td><td>0.03005226480836237</td></tr></table> 	Metric	Value	Accuracy	0.9699477351916377	F1-score	0.009569377990430622	Precision	0.0048543689320388345	Recall	0.3333333333333333	ROC AUC	0.6517792302106027	Brier Score	0.03005226480836237
Metric	Value																																											
Accuracy	0.9989300998573466																																											
F1-score	0.25																																											
Precision	0.3333333333333333																																											
Recall	0.2																																											
ROC AUC	0.5998215241834731																																											
Brier Score	0.001069900142653524																																											
Metric	Value																																											
Accuracy	0.998357473495595																																											
F1-score	0.2666666666666666																																											
Precision	0.2857142857142857																																											
Recall	0.25																																											
ROC AUC	0.6246262520556136																																											
Brier Score	0.0016425265044049575																																											
Metric	Value																																											
Accuracy	0.9699477351916377																																											
F1-score	0.009569377990430622																																											
Precision	0.0048543689320388345																																											
Recall	0.3333333333333333																																											
ROC AUC	0.6517792302106027																																											
Brier Score	0.03005226480836237																																											

Tabelle 16: Schätzgüte von Kapitalgesellschaften beispielhaft mit Bayesian Optimization

Bei der Betrachtung der Merkmalswichtigkeiten fällt auf, dass sich die Ergebnisse der beiden Ansätze SHAP und XGBoost Feature Importance bei gleicher Modellinstanz stark unterscheiden. Dies lässt sich auf die unterschiedlichen Berechnungsweisen der Statistiken der jeweiligen Ansätze zurückführen (siehe Abschnitte 2.3.3 und 2.3.4).

Darüber hinaus ändert sich die Gewichtung der Merkmale auch in Abhängigkeit von der gewählten Untergruppe des Datensatzes (vgl. Tabelle 17). So wirken sich verschiedene Finanzkennzahlen je nach Unternehmensgröße unterschiedlich stark auf die Insolvenzwahrscheinlichkeit aus. Dabei ist zu berücksichtigen, dass für jede Gruppe ein eigenes Modell erstellt wird. Die Konfigurationsparameter sind jedoch ähnlich, da jeweils ein 70/30 Split und eine Bayesian Optimization gewählt wurden. Bei der Betrachtung der Scatter Summary Plots, die neben der Bedeutung der Merkmale für die Prognose auch die Merkmalsausprägungen in Relation setzen, sind jedoch Zusammenhänge zwischen den Unternehmensgruppen erkennbar. Diese sind bei einzelnen Merkmalen wie hier $L2(t_0)$, $STDR(t_0)$, $ROE(t-1)$, $ER(t_0)$ und $RTA(t_0)$ durch ähnliche Verteilungen und Farbverläufe erkennbar.



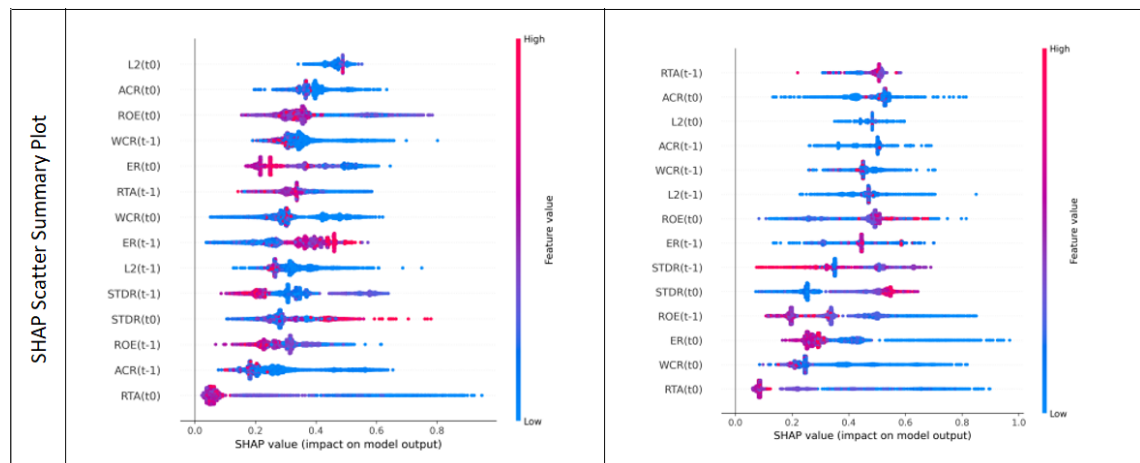


Tabelle 17: Merkmalswichtigkeit von großen und mittelgroßen Kapitalgesellschaften beispielhaft mit Bayesian Optimization

Die Gewichte der Merkmale je Modellinstanz unterscheiden sich auch bei gleicher Unternehmensgruppe und ansonsten gleichen Einstellungen der Modellparameter. In der folgenden Tabelle 18 wurden jeweils mittelgroße Kapitalgesellschaften untersucht. Als Oversampling-Technik wurde SMOTE verwendet und für beide Modellinstanzen eine 70/30-Aufteilung vorgenommen. Bestimmte Zusammenhänge zwischen den Modellinstanzen sind wiederum nur im Scatter Summary-Plot für die Merkmale ER(t-1), L2(t-1), ROE(t0), ER(t0), WCR(t0), WCR(t-1), STDR(t-1), STDR(t0), RTA(t-1) und RTA(t0) erkennbar.

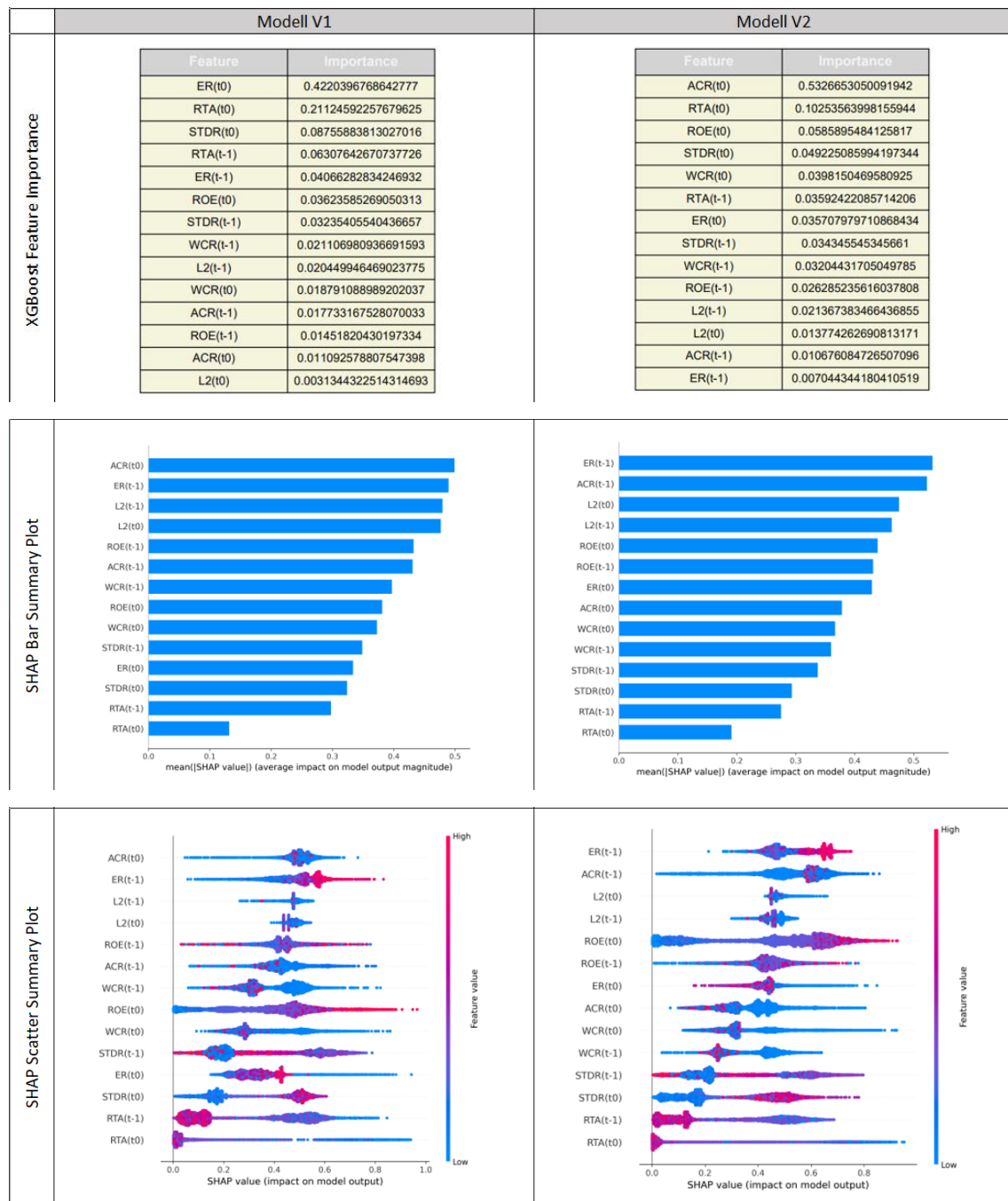


Tabelle 18: Merkmalswichtigkeit von mittelgroßen Kapitalgesellschaften beispielhaft in zwei Modellen

Zusammenfassend lässt sich sagen, dass die Ergebnisse der verwendeten Transparenzmethoden mit dem spezifischen Amadeus-Datensatz hinsichtlich ihrer Erklärbarkeit begrenzt sind. Die Kennzahlen zur Schätzgüte liefern eindeutige Ergebnisse, die es dem Anwender ermöglichen, die Güte der Insolvenzprognosen je Modell zu beurteilen. Die Filterfunktion für einzelne Prognosen ermöglicht darüber hinaus eine detaillierte Darstellung der spezifischen Prognosen. Allerdings können keine verlässlichen Informationen über die Gewichtung der Merkmale für den Anwender

abgeleitet werden. Lediglich die SHAP Scatter Summary Plots, die neben der Merkmalswichtigkeit auch die Höhe der Merkmalsausprägungen betrachten, scheinen Zusammenhänge zwischen den Modellinstanzen bei gleichem Datensatz zu ermöglichen. Es ist zu vermuten, dass die Gesamtmenge der Insolvenzen nicht ausreicht, um die entscheidenden relativen Kennzahlen zu ermitteln.

7.2 Das Dashboard als Instrument: Konfigurierbarkeit und Erklärbarkeit in der KI-gestützten Insolvenzprognose

Die Beantwortung der zweiten Forschungsfrage befasst sich mit dem Aufbau eines Dashboards zur Konfigurierbarkeit und Erklärbarkeit einer KI-gestützten Insolvenzprognose. Das Kapitel ist erneut in zwei Unterkapitel unterteilt, die sich jeweils auf unterschiedliche Aspekte der Konfigurierbarkeit und Erklärbarkeit konzentrieren.

Das Unterkapitel 7.2.1 konzentriert sich auf die grundlegenden Funktionen des Dashboards, die zur Unterstützung der Konfigurierbarkeit und Erklärbarkeit des Modells beitragen. Hier werden Aspekte wie die detaillierte Prozessdokumentation, der Upload eigener Daten und die Segmentierung des Datensatzes diskutiert.

Das Unterkapitel 7.2.2 befasst sich wiederum mit den erweiterten Funktionen des Dashboards, die zur Verbesserung der Konfigurierbarkeit und Erklärbarkeit beitragen. Hier werden Aspekte wie die Wahl zwischen Oversampling-Methoden, die Bayesian Optimization, die Einstellung individueller Hyperparameter, die Auswahl zwischen verschiedenen Train-Test-Validation Verhältnissen und die Visualisierung und Speicherung der Ergebnisse diskutiert.

Beide Unterkapitel ergänzen sich gegenseitig und bieten zusammen einen umfassenden Überblick darüber, wie das Dashboard zur Konfigurierbarkeit und Erklärbarkeit einer KI-gestützten Insolvenzprognose beiträgt.

7.2.1 Unterstützung der Konfigurierbarkeit und Erklärbarkeit durch das Dashboard

Die detaillierte Prozessdokumentation, die sämtliche Prozesse auflistet, ist essenziell, um den gesamten Modellerstellungsprozess nachvollziehbar zu machen. Eine Fortschrittsanzeige über einen Progress Bar auf jeder Seite des

Modellerstellungsprozesses trägt zusätzlich dazu bei, den aktuellen Stand, an dem sich der Nutzer im Prozess befindet, mit einem Blick zu erfassen.

Das Dashboard ermöglicht den Upload und die Verwendung eigener Daten. Dabei ist die Datenbereinigung transparent und direkt im Prozess dokumentiert, um sie so nachvollziehbar wie möglich zu gestalten. Detaillierte Erklärungen der einzelnen Schritte zur Datenbereinigung helfen dem Nutzer, zu verstehen, was mit den von ihm hochgeladenen Daten passiert und wie sie für das Modelltraining vorbereitet werden.

Die Möglichkeit, die zu nutzenden Features, im Rahmen der bereitgestellten Alternativen und unter den vorgegebenen Bedingungen, selbst zu selektieren, erhöht die Konfigurierbarkeit und ermöglicht es dem Nutzer, Einfluss auf die Modellentstehung und -ausgestaltung zu nehmen.

Die Option, den Datensatz entweder unsegmentiert weiter zu verwenden, oder ihn in zwei bis vier Untergruppen zu segmentieren, bietet eine weitere Möglichkeit zur Kontrolle über diesen Prozess und kann, abhängig von der Heterogenität des Datensatzes, von entscheidender Bedeutung in Hinblick auf die Modellqualität sein.

Zusammenfassend lässt sich sagen, dass das Dashboard durch seine vielfältigen Funktionen und Optionen die Konfigurierbarkeit und Erklärbarkeit des Modells erheblich unterstützt und verbessert. Es bietet dem Nutzer ein hohes Maß an Flexibilität und Kontrolle über den Modellerstellungsprozess und trägt dazu bei, die Qualität und Akzeptanz der KI-gestützten Insolvenzprognose zu verbessern.

7.2.2 Verbesserung der Konfigurierbarkeit und Erklärbarkeit durch das Dashboard

Die Wahl zwischen verschiedenen Oversampling-Methoden ermöglicht es dem Nutzer, die Balance zwischen den Klassen in den Trainingsdaten zu steuern. Dies ist besonders in der Insolvenzprognose von Bedeutung, da die Klassen sehr unausgewogen ausgeprägt sind. Durch die Auswahl der geeigneten Oversampling-Methode kann der Nutzer die Vorhersageleistung des Modells verbessern.

Die Bayesian Optimization ist eine Methode zur Optimierung von Modellparametern. Durch die Bereitstellung dieser Funktion im Dashboard kann der Nutzer die

Modellleistung weiter verbessern, ohne tiefes Wissen über die zu Grunde liegenden mathematischen Prozesse zu benötigen.

Die Einstellung individueller Hyperparameter ermöglicht es dem Nutzer, das Modell an die spezifischen Anforderungen der Insolvenzprognose anzupassen. Dies kann dazu beitragen, die Vorhersagegenauigkeit zu erhöhen.

Die Auswahl zwischen verschiedenen Verhältnissen von Train-Test-Validation-Splits ermöglicht es dem Nutzer, das Modell auf verschiedene Weisen zu validieren. Dies kann dazu beitragen, ein besseres Verständnis für die Leistung des Modells zu erlangen und mögliche Überanpassungen zu vermeiden.

Die Visualisierung und Speicherung der Ergebnisse ist ein weiterer wichtiger Aspekt, der zur Erklärbarkeit der Modelle beiträgt. Durch die Visualisierung der Modellergebnisse kann der Nutzer die Leistung des Modells besser verstehen und interpretieren. Die Speicherung der Ergebnisse ermöglicht es dem Nutzer, diese Leistung über Zeit hinweg zu verfolgen und zu protokollieren und bei Bedarf Anpassungen vorzunehmen.

Zusammenfassend lässt sich also sagen, dass die vielfältigen Funktionen zur Verbesserung der Konfigurierbarkeit und Erklärbarkeit des Modells einen nennenswerten Mehrwert stiften. Auch an dieser Stelle werden dem Nutzer sowohl ein hohes Maß an Flexibilität als auch Kontrolle geboten.

8 Fazit

Diese Arbeit hat sich intensiv mit der Verbesserung der Erklärbarkeit von KI-Modellen in der Insolvenzprognose und der Gestaltung benutzerfreundlicher Dashboards zu diesem Zweck beschäftigt. Die Ergebnisse zeigen, dass die Nutzung und Verständlichkeit einer KI-gestützten Insolvenzprognose durch den Einsatz verschiedener Ansätze zur Förderung der Modell-Erklärbarkeit verbessert werden kann. Darüber hinaus kann ein übersichtliches Dashboard die Konfigurierbarkeit und Erklärbarkeit der Prognose unterstützen.

Bezüglich der ersten Forschungsfrage, die sich auf die Verbesserung der Erklärbarkeit von maschinellen Lernmodellen bezieht, wurde festgestellt, dass verschiedene Verfahren zur Verbesserung der Erklärbarkeit von KI-gestützten Insolvenzprognosen erfolgreich implementiert wurden. Dazu gehören die Verwendung von SHAP-Plots, Konfusionsmatrizen, Schätzgütekennzahlen und XGBoost Feature Importance. Darüber hinaus trägt die transparente Kommunikation aller Modellparameter zur Erklärbarkeit bei. Durch die Offenlegung dieser Parameter kann der Nutzer nachvollziehen, wie das Modell konfiguriert ist und welche Einstellungen zur erzielten Prognosegüte beitragen. Eine Filterfunktion nach einzelnen Unternehmen bietet zudem die Möglichkeit, Prognosen und deren Genauigkeit gezielt für einzelne Unternehmen zu betrachten und zu analysieren. Auf der praktischen Ebene wurde die Nutzbarkeit und Verständlichkeit der Prognoseergebnisse erhöht, was eine zielgerichtete und fundierte Entscheidungsfindung auf Basis der bereitgestellten Informationen ermöglicht. Betrachtet man die erzielten Prognoseergebnisse und die Ergebnisse der Interpretierbarkeitsmethoden mit dem spezifischen Amadeus-Datensatz, so zeigt sich, dass die angewandten Methoden zur Verbesserung der Erklärbarkeit von KI-gestützten Insolvenzprognosen zwar erfolgreich umgesetzt wurden. Aufgrund der eingeschränkten Datenlage sind die Ergebnisse allerdings nur bedingt aussagekräftig. So konnten trotz der detaillierten Darstellung der einzelnen Prognosen über eine Filterfunktion keine verlässlichen Aussagen über die Gewichtung der Merkmale für den Anwender abgeleitet werden. Lediglich die SHAP Scatter Summary Plots, die neben der Merkmalsgewichtung auch die Höhe der Merkmalsausprägungen berücksichtigen, scheinen Zusammenhänge zwischen den Modellinstanzen bei gleichem Datensatz zu ermöglichen. Es ist zu

vermuten, dass die Gesamtzahl der Insolvenzen nicht ausreicht, um die entscheidenden relativen Kennzahlen zu ermitteln. Die Kennzahlen zur Schätzgüte liefern jedoch eindeutige Ergebnisse, die es dem Anwender ermöglichen, die Güte der Insolvenzprognose je Modell zu beurteilen.

Die zweite Forschungsfrage bezog sich auf die Gestaltung eines Dashboards zur KI-gestützten Insolvenzprognose. Unsere Ergebnisse zeigen, dass ein effektives Dashboard klar strukturiert sein sollte, interaktive Elemente zur Anpassung von Modellparametern enthalten und Erklärungen und Visualisierungen zur Unterstützung der Modellinterpretation bereitstellen sollte. Das Dashboard unterstützt die Konfigurierbarkeit und Erklärbarkeit des Modells wesentlich durch seine vielfältigen Funktionen und Auswahlmöglichkeiten. Es bietet dem Benutzer ein hohes Maß an Flexibilität und Kontrolle über den Modellerstellungsprozess. Dazu gehören Funktionen wie die detaillierte Prozessdokumentation, der Upload eigener Daten, die Segmentierung des Datensatzes, die Auswahl zwischen Oversampling-Methoden, die Bayesian Optimization, die Einstellung individueller Hyperparameter, die Auswahl zwischen verschiedenen Train-Test-Validierungsverhältnissen sowie die Visualisierung und Speicherung der Ergebnisse.

Diese Ergebnisse tragen zur aktuellen Forschung über KI-gestützte Insolvenzprognosen bei, indem sie praktische Lösungen für die Herausforderungen in diesem Bereich bieten. Sie zeigen, dass die Verbesserung der Erklärbarkeit von KI-Modellen und die Gestaltung benutzerfreundlicher Dashboards möglich sind. Gleichzeitig zeigen sie auf, dass eine größere Anzahl von Insolvenzfällen notwendig ist, um robustere Prognosen zu erhalten.

In zukünftigen Forschungsarbeiten könnte es interessant sein, die in dieser Arbeit vorgestellten Ansätze zur Verbesserung der Erklärbarkeit und Konfigurierbarkeit von KI-Modellen auf andere Anwendungsfälle in der Finanztechnologie auszuweiten. Dies könnte z.B. die Anwendung auf Kreditrisikobewertung, Betrugserkennung oder algorithmischen Handel umfassen.

Darüber hinaus könnten Untersuchungen zur Verbesserung der Benutzererfahrung und zur weiteren Personalisierung des Dashboard-Designs durchgeführt werden. Dabei könnte auch untersucht werden, wie Benutzerinteraktion und Feedback in das Design

integriert werden können, um eine noch intuitivere und benutzerfreundlichere Oberfläche zu schaffen.

Schließlich wäre es auch interessant, die Auswirkungen der Implementierung solcher Modelle und Dashboards auf die Entscheidungsfindung in der Praxis zu untersuchen. Dies könnte z.B. durch Fallstudien oder Nutzerbefragungen geschehen, um zu verstehen, wie diese Werkzeuge die Effizienz und Genauigkeit von Entscheidungen im Finanzsektor verbessern können.

Anhang

Im folgenden Anhang befinden sich weiterführende Inhalte aus der vorliegenden Arbeit. Die Auflistung erfolgt chronologisch nach dem Ausarbeitungsaufbau:

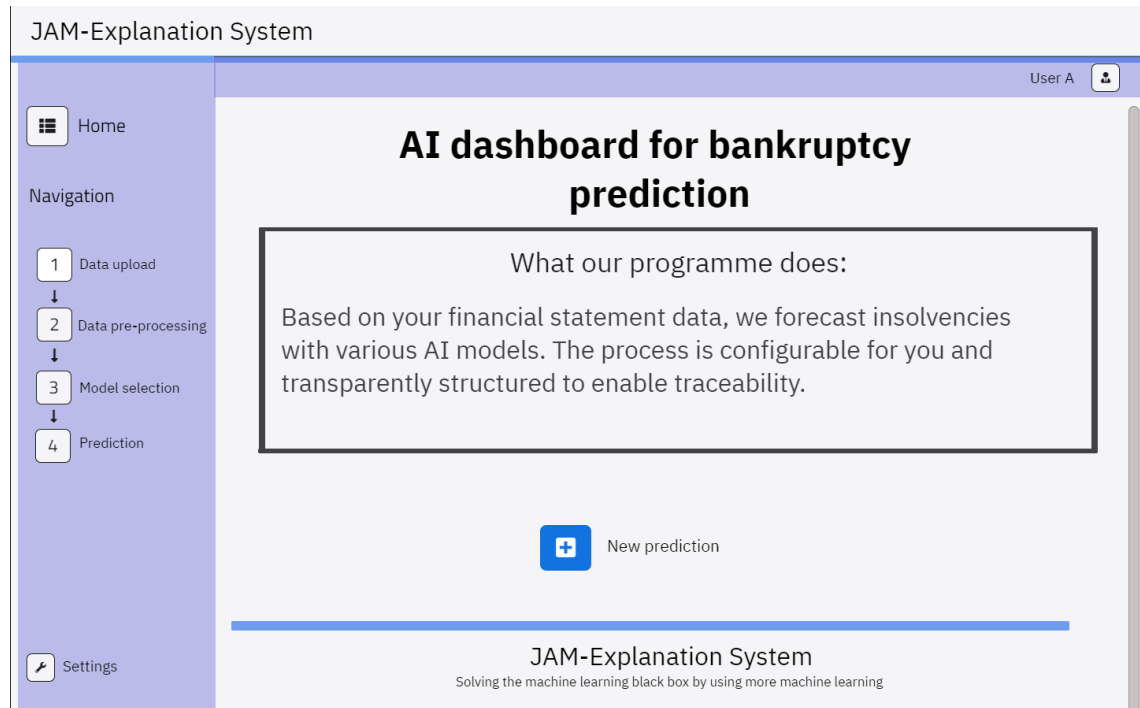


Abbildung 33: Homepage der Applikation

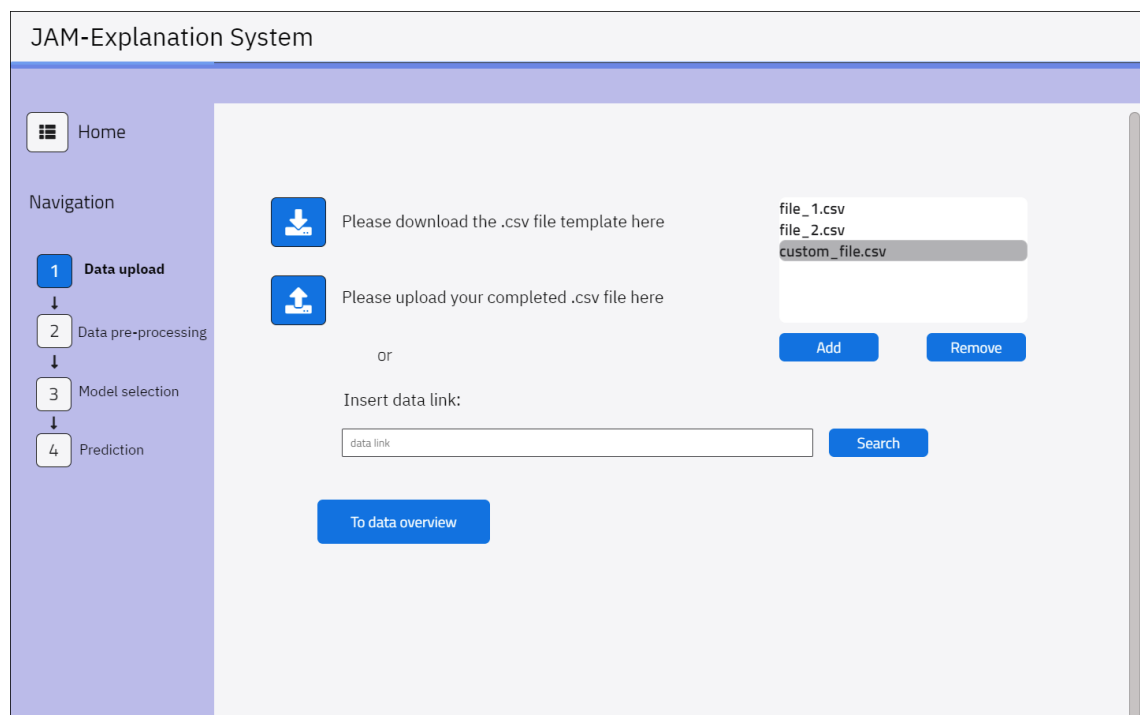


Abbildung 34: Datenupload

JAM-Explanation System

Home

Navigation

- 1 Data upload
- 2 Data pre-processing
- 3 **Model selection**
- 4 Prediction

Creation of new AI-model

Neural Network

Set parameters

Choose parameters Initialise parameters

Number of iterations 5

Layer	Number of neurons	Activation function
Layer 1	20	relu
Layer 2	20	relu
Layer 3	10	sigmoid

Further settings

Optimizer RMSProp

Learning rate 0.001

Normalization of variables Yes

Training approach

k-fold cross-validation: k-value 5

Separation into training-/test-/validation-dataset:

Training data: 60%

Test data: 20%

Validation data: 20%

return next

New model

■ Training data ■ Test data ■ Validation data

Abbildung 35: Hyperparameter Einstellungen für Neuronales Netzwerk

JAM-Explanation System

Home

Navigation

- 1 Data upload
- 2 Data pre-processing
- 3 **Model selection**
- 4 Prediction

Output models

Global Model-Agnostic Methods

- ☒ TREPAN-Algorithmus
- ☒ Partial Dependency Plots
- ☒ Accumulated Local Effects
- ☒ Permutation Feature Importance
- ☒ Feature Interaction

Local Model-Agnostic Methods

- ☒ LIME-Algorithmus
- ☒ SHAP-Diagramm

Assessment of the estimation quality

- ☒ Accuracy
- ☒ Brier-Score
- ☒ ROC-AUC
- ☒ F1 Score
- ☒ Recall

return start prediction

Abbildung 36: Auswahl der Erklärbarkeitsmethoden

Literaturverzeichnis

- Abedin, B. (2022). Managing the tension between opposing effects of explainability of artificial intelligence: a contingency theory perspective. *Internet Research*, 32(2), 425–453. <https://doi.org/10.1108/INTR-05-2020-0300>
- Åkerblom, T., & Johan Lyhagen, S. (2022). *Predicting Bankruptcy with Machine Learning Models*.
- Alaka, H. A., Oyedele, L. O., Owolabi, H. A., Kumar, V., Ajayi, S. O., Akinade, O. O., & Bilal, M. (2018). Systematic review of bankruptcy prediction models: Towards a framework for tool selection. In *Expert Systems with Applications* (Vol. 94, pp. 164–184). Elsevier Ltd. <https://doi.org/10.1016/j.eswa.2017.10.040>
- Amadeus. (2022, May 22). *Home*. https://amadeus-2r1-1bvdinfo-1com-112ntt1y110fd.han.sub.uni-goettingen.de/version-20230309-1-3/Search.ContactPersonalTab.serv?_CID=25&databaseContext=Contacts&EditSearchStep=false&product=amadeusneo&restart=0&SearchStepId=Current.NeoAdvancedSearch
- Ashraf, S., G. S. Félix, E., & Serrasqueiro, Z. (2019). Do Traditional Financial Distress Prediction Models Predict the Early Warning Signs of Financial Distress? *Journal of Risk and Financial Management*, 12(2). <https://doi.org/10.3390/jrfm12020055>
- BaFin. (2021). *Big Data und künstliche Intelligenz: Prinzipien für den Einsatz von Algorithmen in Entscheidungsprozessen*. https://www.bafin.de/SharedDocs/Downloads/DE/Aufsichtsrecht/dl_Prinzipienpapier_BDAI.html;jsessionid=9BEFE75310E9362406B2FA3F82DF9541.1_cid500?nn=9021442
- Batista, G. E. A. P. A., Prati, R. C., & Monard, M. C. (2004). A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explorations Newsletter*, 6(1), 20–29. <https://doi.org/10.1145/1007730.1007735>
- Baur, N., & Blasius, J. (2014). Methoden der empirischen Sozialforschung. *Handbuch Methoden Der Empirischen Sozialforschung*, 925. https://doi.org/10.1007/978-3-531-18939-0_1

- Carvalho, D. V., Pereira, E. M., & Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics. In *Electronics (Switzerland)* (Vol. 8, Issue 8). MDPI AG. <https://doi.org/10.3390/electronics8080832>
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/JAIR.953>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Chi, D. J., & Chu, C. C. (2021). Artificial intelligence in corporate sustainability: Using lstm and gru for going concern prediction. *Sustainability (Switzerland)*, 13(21). <https://doi.org/10.3390/su132111631>
- Chou, T.-N. (2019). An Explainable Hybrid Model for Bankruptcy Prediction Based on the Decision Tree and Deep Neural Network. *2019 IEEE 2nd International Conference on Knowledge Innovation and Invention (ICKII)*, 122–125. <https://doi.org/10.1109/ICKII46306.2019.9042639>
- Ciampi, F. (2015). Corporate governance characteristics and default prediction modeling for small enterprises. An empirical analysis of Italian firms. *Journal of Business Research*, 68(5), 1012–1025. <https://doi.org/10.1016/J.JBUSRES.2014.10.003>
- El Kalak, I., & Hudson, R. (2016). The effect of size on the failure probabilities of SMEs: An empirical study on the US market using discrete hazard model. *International Review of Financial Analysis*, 43, 135–145. <https://doi.org/10.1016/J.IRFA.2015.11.009>
- Europäische Union. (2016). Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung). In *Amtsblatt der Europäischen Union: Vol. Artikel 22*. <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=celex%3A32016R0679>

- European Commission. (2021). *Proposal for the The Artificial Intelligence Act*.
<https://artificialintelligenceact.eu/the-act/>
- Fraunhofer IAIS. (2021). *Leitfaden zur Gestaltung vertrauenswürdiger Künstlicher Intelligenz: KI-Prüfkatalog*. www.iais.fraunhofer.de/ki-pruefkatalog
- Goyal, S. (2021). *Evaluation Metrics for Classification Models*.
<https://medium.com/analytics-vidhya/evaluation-metrics-for-classification-models-e2f0d8009d69>
- Handelsgesetzbuch. (n.d.). § 267 Umschreibung der Größenklassen. Retrieved May 17, 2023, from https://www.gesetze-im-internet.de/hgb/_267.html
- Harhoff, D., Stahl, K., & Woywode, M. (1998). Legal form, growth and exit of west German firms - Empirical results for manufacturing, construction, trade and service industries. *Journal of Industrial Economics*, 46(4), 453–488.
<https://doi.org/10.1111/1467-6451.00083>
- Holovaty, A., & Kaplan-Moss, J. (2009). The Definitive Guide to Django. *The Definitive Guide to Django*. <https://doi.org/10.1007/978-1-4302-1937-8>
- Hölscher, R., & Helms, N. (2018). Insolvenzprognose Gabler Wirtschaftslexikon. In *Gabler Wirtschaftslexikon*. Springer gabler.
- Insolvenzordnung, (2022). <https://www.gesetze-im-internet.de/inso/BJNR286600994.html>
- Islam, M. R., Ahmed, M. U., Barua, S., & Begum, S. (2022). A Systematic Review of Explainable Artificial Intelligence in Terms of Different Application Domains and Tasks. *Applied Sciences (Switzerland)*, 12(3). <https://doi.org/10.3390/app12031353>
- Islam, S. R., Eberle, W., Bundy, S., & Ghafoor, S. K. (2019). *Infusing domain knowledge in AI-based “black box” models for better explainability with application in bankruptcy prediction*. <http://arxiv.org/abs/1905.11474>
- Kaur, H., Pannu, H. S., & Mali, A. K. (2019). A Systematic Review on Unbalanced Data Challenges in Machine Learning: Applications and Solutions. *ACM Comput. Surv.*, 52(4).

- Kim, H., Cho, H., & Ryu, D. (2020). Corporate default predictions using machine learning: Literature review. In *Sustainability (Switzerland)* (Vol. 12, Issue 16). MDPI. <https://doi.org/10.3390/SU12166325>
- Kim, H., Cho, H., & Ryu, D. (2022). Corporate Bankruptcy Prediction Using Machine Learning Methodologies with a Focus on Sequential Data. *Computational Economics*, 59(3), 1231–1249. <https://doi.org/10.1007/S10614-021-10126-5/TABLES/4>
- Kim, T. W., & Routledge, B. R. (2022). Why a Right to an Explanation of Algorithmic Decision-Making Should Exist: A Trust-Based Approach. *Business Ethics Quarterly*, 32(1), 75–102. <https://doi.org/10.1017/beq.2021.3>
- Kuhrmann, M., Diebold, P., Münch, J., Tell, P., Garousi, V., Federer, M., Trektore, K., McCaffery, F., Linssen, O., Hanser, E., & Prause, C. R. (2017). Hybrid Software and System Development in Practice: Waterfall, Scrum, and Beyond. *Proceedings of International Conference on Software System Process, 2022-March*.
- Lahmiri, S., & Bekiros, S. (2019). Can machine learning approaches predict corporate bankruptcy? Evidence from a qualitative experimental design. *Quantitative Finance*, 19(9), 1569–1577. <https://doi.org/10.1080/14697688.2019.1588468>
- Le, T., Lee, M. Y., Park, J. R., & Baik, S. W. (2018). Oversampling techniques for bankruptcy prediction: Novel features from a transaction dataset. *Symmetry*, 10(4). <https://doi.org/10.3390/sym10040079>
- Li, X. H., Cao, C. C., Shi, Y., Bai, W., Gao, H., Qiu, L., Wang, C., Gao, Y., Zhang, S., Xue, X., & Chen, L. (2022). A Survey of Data-Driven and Knowledge-Aware eXplainable AI. *IEEE Transactions on Knowledge and Data Engineering*, 34(1), 29–49. <https://doi.org/10.1109/TKDE.2020.2983930>
- Liang, D., Tsai, C., Systems, H. W.-K.-B., & 2015, undefined. (2015). The effect of feature selection on financial distress prediction. *Elsevier*. <https://doi.org/10.1016/j.knosys.2014.10.010>

- Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2021). Explainable ai: A review of machine learning interpretability methods. In *Entropy* (Vol. 23, Issue 1, pp. 1–45). MDPI AG. <https://doi.org/10.3390/e23010018>
- Loyola-Gonzalez, O. (2019). Black-box vs. White-Box: Understanding their advantages and weaknesses from a practical point of view. In *IEEE Access* (Vol. 7, pp. 154096–154113). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ACCESS.2019.2949286>
- Lundberg, S. M., Allen, P. G., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *31st Conference on Neural Information Processing Systems*. <https://github.com/slundberg/shap>
- Maria Navin, J. R., & Pankaja, R. (2016). Performance Analysis of Text Classification Algorithms using Confusion Matrix. *International Journal of Engineering and Technical Research*, 75–78. www.erpublication.org
- McKinney, W. (2010). *Data Structures for Statistical Computing in Python*.
- McKinsey & Company. (2021). *Building the AI bank of the future*. <https://www.mckinsey.com/industries/financial-services/our-insights/building-the-ai-bank-of-the-future>
- Olson, D. L., Delen, D., & Meng, Y. (2012). Comparative analysis of data mining methods for bankruptcy prediction. *Decision Support Systems*, 52(2), 464–473. <https://doi.org/10.1016/J.DSS.2011.10.007>
- Pamuk, M., Grendel, R. O., & Schumann, M. (2021). *Towards ML-based Platforms in the Finance Industry-An ML Approach to Generate Corporate Bankruptcy Probabilities based on Annual Financial Statements*.
- Park, M. S., Son, H., Hyun, C., & Hwang, H. J. (2021). Explainability of Machine Learning Models for Bankruptcy Prediction. *IEEE Access*, 9, 124887–124899. <https://doi.org/10.1109/ACCESS.2021.3110270>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., & Thirion, B. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, 12, 2825–2830.

- <https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf?ref=htps://>
- Powers, D. M. W. (2007). *Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation*.
- Rai, A. (2020). Explainable AI: from black box to glass box. In *Journal of the Academy of Marketing Science* (Vol. 48, Issue 1, pp. 137–141). Springer. <https://doi.org/10.1007/s11747-019-00710-5>
- Rivero, J. M., Rossi, G., Grigera, J., Luna, E. R., & Navarro, A. (2011). From Interface Mockups to Web Application Models. *Web Information System Engineering*.
- Rubio, D. (2017). Beginning Django. *Beginning Django*. <https://doi.org/10.1007/978-1-4842-2787-9>
- Sawyer, C., Reichelderfer, F. W., Editor, J. E., Caskey, J. R., & Brie, G. W. (1950). *VERIFICATION OF FORECASTS EXPRESSED IN TERMS OF PROBABILITY*.
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press. <http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning>
- Shi, Y., & Li, X. (2019). A bibliometric study on intelligent techniques of bankruptcy prediction for corporate firms. In *Heliyon* (Vol. 5, Issue 12). Elsevier Ltd. <https://doi.org/10.1016/j.heliyon.2019.e02997>
- Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, 146, 102551. <https://doi.org/10.1016/J.IJHCS.2020.102551>
- Viégas, F., & Wattenberg, M. (2023). *The System Model and the User Model: Exploring AI Dashboard Design*. <https://arxiv.org/abs/2305.02469v1>
- vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Brocke, vom. (2015). Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. In *Communications of the Association for Information Systems* (Vol. 37). <http://aisel.aisnet.org/cais/vol37/iss1/9>

- Zhang, A., Lipton, Z. C., Li, M. U., & Smola, A. J. (2021). *Dive into Deep Learning*.
- Zięba, M., Tomczak, S. K., & Tomczak, J. M. (2016). Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction. *Expert Systems with Applications*, 58, 93–101.
<https://doi.org/10.1016/j.eswa.2016.04.001>
- Zou, Y., Gao, C., & Gao, H. (2022). Business Failure Prediction Based on a Cost-Sensitive Extreme Gradient Boosting Machine. *IEEE Access*, 10, 42623–42639.
<https://doi.org/10.1109/ACCESS.2022.3168857>

Ehrenwörtliche Erklärung

Wir versichern, dass wir die Arbeit „Konzeption und Implementierung eines Dashboards zur Darstellung KI-gestützter Insolvenzprognose mithilfe mehrerer und differenzierter Kennzahlen über die Unternehmenslage“ selbständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt haben. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen oder anderen Quellen entnommen sind, sind als solche kenntlich gemacht. Die schriftliche und die elektronische Form der Arbeit stimmen überein. Wir stimmen der Überprüfung der Arbeit durch eine Plagiatssoftware zu.

Micha Alexander Teweleit

Datum

Unterschrift

Göttingen, den 25.09.2023



Joachim von Alvensleben

Datum

Unterschrift

Göttingen, den 25.09.2023

