


# 1. Motivation

Unternehmen treffen unzählige Entscheidungen jeden Tag. Wertschöpfende Entscheidungen tragen, wie ihr Name schon sagt, zur Wertschöpfung bei und beeinflussen somit unmittelbar den wirtschaftlichen Erfolg eines Unternehmens. Eine Bank beispielsweise muss einen Kreditnehmer auf dessen Fähigkeit, Raten zurückzahlen zu können, prüfen. Somit soll sichergestellt werden, dass die Bank den Betrag des Kredites sowie auch die ihr zustehenden Zinsen zurück bekommt. Dadurch steht und fällt der Erfolg einer Bank mit den Krediten, die sie vergibt [4, vgl. S. 1]. Prominentestes Beispiel hierfür ist die ehemalige US-amerikanische Investmentbank Lehmann Brothers, die im Zuge der Finanzkrise 2008 Insolvenz anmelden musste [7, vgl. S. 3]. Boomende Immobilien-Preise resultierten in höheren Sicherheiten von Privathaushalten, wodurch auch Kredite an Darlehnsnehmer mit geringer Bonität vergeben wurden (Subprime-Kredit) [3, vgl. S. 82-83]. Nach dem Platzen der Immobilienblase kam es zu Zahlungsausfällen und -störungen im US-amerikanischen Hypothekenmarkt [1, vgl. S. 6 ff.], was erst in einem Finanzcrash und dann in einer Weltwirtschaftskrise resultierte. [5, vgl. S. 109]





~~Online~~-Banken werben seit Längerem mit Kreditentscheidungen innerhalb von wenigen Sekunden. Das Ausfüllen eines Formulars ist ausreichend, um wenige Sekunden später eine Kreditentscheidung zu erhalten. Die Prüfung der Bonität geschieht automatisiert im Hintergrund durch eine zuvor definierte Entscheidungslogik. ~~Die~~ Entscheidungslogik wird von Fach-Experten erstellt, wobei die Fach-Experten im Wesentlichen zwei essentielle Dinge beachten müssen:

1. Verhinderung der Vergabe von Krediten an Personen, die diese nicht zurückzahlen können (“default rate“)
2. Verhinderung der Nicht-Vergabe von Krediten an Personen, die eigentlich in der Lage sind, diese zurück zuzahlen, wodurch Einnahmen an Konkurrenten verloren gehen (“acceptance rate“)

Banken sind somit angehalten eine gute Balance zwischen Sicherheit und Gewinn zu finden. Dazu muss die fachliche Entscheidungslogik drohende Zahlungsausfälle so genau wie möglich vorhersagen. Studien wie [10] zeigen, dass sieben Prozent der vergebenen Kredite im Euroraum mehr als 90 Tage im Zahlungsverzug sind. Das entspricht einem Volumen von einer Billionen Euro. Bei einem Zahlungsverzug von über 90 Tagen wird generell davon ausgegangen, dass ein Kredit nicht mehr in voller Höhe zurückgezahlt wird und somit zu einem Verlust-Geschäft führt [2, vgl. S. 8]. Könnte man die Entscheidungslogiken bei den Banken dahingehend optimieren, ein Prozent weniger faule Kredite zu vergeben, könnten große Summen eingespart werden, die sonst im Zuge einer Banken-Rettung dem Steuerzahler auferlegt werden. 

## 1.1. Problemstellung


Das Einsparen der genannten Summen wäre erstrebenswert, allerdings gibt es unzählige Gründe, die den Finanzcrash ab 2007 verschuldet haben. Als Hauptgrund wird in der Literatur die umfangreiche Vergabe von Subprime-Krediten genannt [3] [7]. Bei der Vergabe dieser Kredite wurde das Risiko nicht korrekt identifiziert [7, vgl. S. 3], was zu unzähligen falschen fachlichen Entscheidungen führte (~~Problemstellung~~). Die Menge der falsch getroffenen fachlichen Entscheidungen, trug somit immens zur Verschlechterung der Situation bei. 

Ein Grund für die vielen falschen Entscheidungen, sind fehlerhafte Entscheidungslogiken und falsche Annahmen (Grund G1 - siehe Abbildung 1.1). Durch die unrealistisch hohen Immobilien-Preise, waren die Kredite mit Immobilien gesichert, die real weitaus weniger wert waren [6]. Bei der Vergabe der Kredite wurden die Entscheidungen verfälscht, da die Entscheidungslogiken mit Eingabe-Daten befüllt wurden, die nicht der Realität sprachen.

Ein weiterer Grund ist, dass vergangene Entscheidungen nicht mehr reflektiert wurden. (Grund G2 - siehe Abbildung 1.1). Die Kredite wurden gebündelt unter den Banken weiterverkauft [7, vgl. S. 2], ein Verlust-Risiko bestand somit nur für die kurze Zeit, in der ein Kredit-Bündel noch nicht weiterverkauft wurde. Saunders et al sagen "The underwriters [...] had little incentive to screen and monitor the activities of borrowers for whom they originated loans" [9]. Nach der ursprünglichen Vergabe des Kredites, wurden also keine weitere Aktivitäten unternommen, um die damaligen Kredit-Entscheidungen

weiterhin zu prüfen oder zu überwachen. Ein Kontrollverlust der Situation war nur eine Frage der Zeit.

Im Falle der Finanzkrise waren allerdings nicht nur fehlerhafte Entscheidungslogik, falsche Annahmen und mangelnde Reflektion schuld am Zusammenbruch der Finanzmärkte. Absichtlich falsch getroffene Entscheidungen sind ebenso als Ursache der vielen falsch getroffenen Entscheidungen anzuführen (Grund G3 - siehe Abbildung 1.1). In einer Pressemitteilung des Vorstandes der US-Notenbank, wurden Bonus-Zahlungen an Bank-Mitarbeiter als weiterer Grund der Finanzkrise aufgezählt: "Flaws in incentive compensation practices were one of many factors contributing to the financial crisis. Inappropriate bonus or other compensation practices can incent senior executives or lower level employees, such as traders or mortgage officers, to take imprudent risks that significantly and adversely affect the firm" [8]. Mitarbeiter übertraten bewusst Risiken, nutzten mangelnde Regulationen zu ihrem Vorteil und ließen keine Möglichkeit aus ihren Gewinn auf Kosten von Stabilität und Sicherheit zu maximieren.

Zusammenfassend soll Abbildung 1.1 nochmals die Gründe aufzeigen die verantwortlich sind, dass zu viele falsche fachliche Entscheidungen getroffen werden ( Problem).

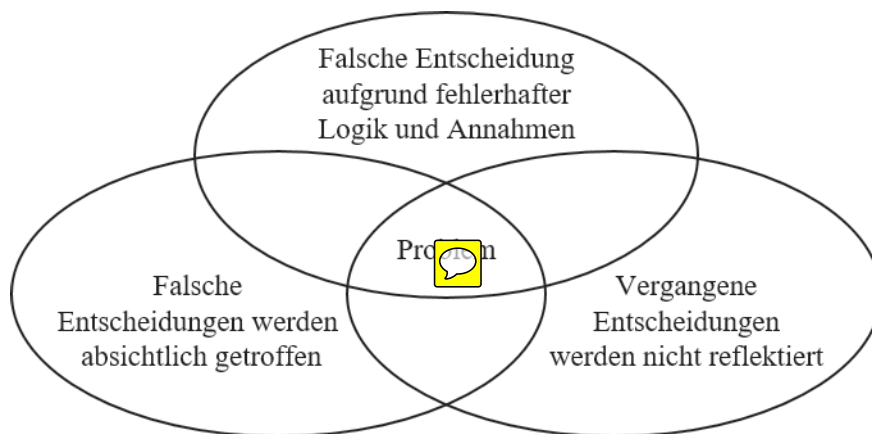


Abbildung 1.1.: Gründe für falsch getroffene fachliche Entscheidungen

## 1.2. Zielsetzung

Das übergeordnete Ziel ist die in 1.1 genannte Problemstellung, dass zu viele falsche fachliche Entscheidungen getroffen werden, zu verbessern, indem man die Anzahl falscher

fachlichen Entscheidungen reduziert. Diese Arbeit soll hierzu ein Machine Learning Verfahren evaluieren, analysieren und implementieren, um fachliche Entscheidungen zu verbessern. Aus dieser Zielsetzung können die folgenden zwei Teilziele abgeleitet werden.

Das *erste Teilziel* ist, das Erstellen und die kontinuierliche Optimierung von Modellen, anhand von vergangen fachlichen Entscheidungen und deren Korrektheit. Dazu soll im ersten Schritt ein Machine Learning Verfahren implementiert werden, das Modelle aus den Daten vergangener Entscheidungen generiert. Die Datensätze dieser Entscheidungen werden um ein weiteres Attribut ergänzt, die Performance. Die Performance beschreibt den Fall, der sich in der realen Welt als der richtige erwiesen hat. Bei der Kredit-Entscheidungslogik wird beispielsweise entschieden, ob ein Kunde ein Kredit bekommt oder nicht. Oder in anderen Worten, ob der Kunde seinen Kredit zurückzahlen wird oder nicht. Sollte ein Kunde in Zahlungsverzug geraten oder sogar insolvent gehen, wäre dessen Performance mit "false" zu bewerten. Das Machine Learning Verfahren, soll die Kombination von Input-Daten und Performance auf Gesetzmäßigkeiten prüfen und diese anschließend als Modell zur Verfügung stellen. Ein Modell beschreibt den Output eines Machine Learning Verfahrens und bietet zwei fundamentale Mechanismen. Erstens soll das Modell in der Lage sein, weiter optimiert werden zu können ("lernen"), zweitens soll für einen gegebenen Input-Datensatz ein dazugehöriger Output vorhergesagt werden ("evaluieren").




Das *zweite Teilziel* besteht darin, Modelle heranzuziehen um Entscheidungslogiken so zu verbessern, dass zukünftige fachliche Entscheidungen korrekter getroffen werden. Das erlernte Modell wird dazu verwendet, mehrere Datensätze zu evaluieren und den Ausgang deren Entscheidungen vorherzusagen. Währenddessen berechnet die zu verbessernde Entscheidungslogik dieselben Datensätze. Anschließend werden dann die Ergebnisse des Models mit den Ergebnissen der Entscheidungslogik sowie der tatsächlichen Performance verglichen. Die daraus resultierenden Ergebnisse sollen dann in den Entscheidungsfindungsprozess mit einfließen und somit zu mehr richtig getroffenen Entscheidungen verhelfen.



Auf Basis der Teilziele eins und zwei, ergeben sich eine Reihe von Forschungsfragen:

- **Forschungsfrage 1:** Können ML-Verfahren verwendet werden, um fachliche Entscheidungen zu verbessern?
  - Was gibt es für einsetzbare ML-Verfahren?
  - Welche ML Verfahren eignen sich am besten?


- **Forschungsfrage 2:** Wie können einsetzbare ML-Verfahren anhand eines fachlichen Use Cases angewendet werden?   
 - Was für Ergebnisse liefern diese?  
 - Wie können die Ergebnisse interpretiert werden?
- **Forschungsfrage 3:** Wie kann die Funktionsweise des ML-Verfahrens in ein Modell überführt werden, ohne dass dabei Performance-Daten offen gelegt oder zur Verfügung gestellt werden müssen?
- **Forschungsfrage 4:** Wie kann sichergestellt werden, dass die fachliche Entscheidung verbessert werden ~~konnte~~?

Diese vier Forschungsfragen werden in Abbildung 1.2 zu einem oder mehreren Kapiteln dieser Thesis zu geordnet.

	<i>Kapitel 1</i>	<i>Kapitel 2</i>	<i>Kapitel 3</i>	<i>Kapitel 4</i>	<i>Kapitel 5</i>	<i>Kapitel 6</i>	<i>Kapitel 7</i>
<i>Forschungsfrage 1</i>		x	x	x	x	x	x
<i>Forschungsfrage 2</i>		x	x			x	
<i>Forschungsfrage 3</i>				x	x		
<i>Forschungsfrage 4</i>						x	x

Abbildung 1.2.: Forschungsfragen entlang der Kapitel

## 1.3. Abgrenzung

Einhergehend mit den zu beantwortenden Forschungsfragen findet auch folgende Ab-  
grenzung statt: Es ist kein Ziel dieser Arbeit, den besten ML-Algorithmus für einen spezifischen Anwendungsfall zu entwickeln. Es gibt unzählige Algorithmen im Bereich des ML und jeder Algorithmus hat seine Vor- und Nachteile. Allerdings ist eine Detailbetrachtung und Auswertung dieser Algorithmen, kein Bestandteil dieser Arbeit. Diese Betrachtung wäre im Rahmen dieser Bachelorthesis zu aufwendig und würde nicht zur Erreichung des übergeordneten Forschungsziel dieser Arbeit beitragen. Vielmehr sollen bereits bestehende Lösungen, die sich in der Praxis als erfolgreich bewiesen haben, intelligent verknüpft werden, um einen Prototypen zu entwickeln, mit dessen Hilfe sich die in Kapitel 1.2 genannten Forschungsfragen beantworten lassen.