


# Data Mining Techniken zur Restwertprognose in der Automobilindustrie: Eine Praktische Anwendung Künstlicher Neuronaler Netze

Bachelorarbeit

zur Erlangung des akademischen Grades „Bachelor of Science (B.Sc.)“ im  
Studiengang Wirtschaftswissenschaft der Wirtschaftswissenschaftlichen  
Fakultät der Leibniz Universität Hannover

vorgelegt von

Dennis Eilers



Prüfer: Prof. Dr. Hans-Jörg von Mettenheim

Hannover, den 06. Juli 2015

# Inhaltsverzeichnis

<b>Abbildungsverzeichnis</b>	<b>IV</b>
<b>Tabellenverzeichnis</b>	<b>V</b>
<b>Abkürzungsverzeichnis</b>	<b>VI</b>
<b>1 Einleitung</b>	<b>7</b>
<b>2 Grundlagen</b>	<b>8</b>
2.1 Ausgangslage und Problemstellung . . . . .	8
2.2 Literature Review . . . . .	11
<b>3 Daten und Methoden</b>	<b>13</b>
3.1 Datengrundlage . . . . .	13
3.2 Lineare Regression . . . . .	14
3.3 Künstliche Neuronale Netze . . . . .	15
<b>4 Ergebnisse</b>	<b>18</b>
4.1 Marktwert Analyse . . . . .	18
4.2 Saisonale Einflüsse . . . . .	21
<b>5 Prognosemodell – Vom Markt- zum Restwert</b>	<b>24</b>
5.1 Der Produktlebenszyklus – Lifecycle-Abschlag . . . . .	24
5.2 Signifikanztest . . . . .	24
5.3 Prognoseverfahren . . . . .	29
<b>6 Diskussion, Limitationen, Ausblick</b>	<b>32</b>
<b>Literatur</b>	<b>34</b>

# 1 Einleitung

*Prediction is very difficult, especially about the future.*

– Niels Bohr

Das Leasinggeschäft stellt für Automobilhersteller in Deutschland eine starke Säule in der Absatzentwicklung dar. 2012 wurden knapp 1,3 Millionen Leasing-Fahrzeuge in der Bundesrepublik zugelassen, was einem Anteil von 36% an den gesamten Neuzulassungen entspricht [14]. Aus Herstellersicht bietet dies jedoch nicht nur große Wettbewerbschancen, sondern vor allem auch schwer quantifizierbare Risiken.

Im Mittelpunkt dieser Arbeit steht das sogenannte Restwertrisiko. Ein Fahrzeug verliert dabei im Laufe der Leasingdauer, abhängig von Parametern wie der Anzahl gefahrener Kilometer, einen gewissen Prozentsatz seines ursprünglichen Wertes. Die während der Vertragslaufzeit zu zahlenden Leasingraten müssen daher genau kalkuliert werden, um die Entwertung inklusive einer geschäftsfähigen Marge abdecken zu können. Während die Leasingrate zum Zeitpunkt des Vertragsabschlusses festgeschrieben wird, realisiert sich der tatsächlich zu erzielende Wiederverkaufspreis des Fahrzeugs auf dem Gebrauchtwagenmarkt erst am Ende der Leasingdauer und ist damit der Höhe nach unsicher. Es wird deutlich, dass eine möglichst exakte Prognose der Restwerte einen erfolgskritischen Faktor für die Wettbewerbsfähigkeit in diesem Markt darstellt. Abweichungen der Prognose vom tatsächlich erzielten Preis, wirken sich in beide Richtungen negativ aus. Zu hoch prognostizierte Restwerte führen zwar durch geringere Leasingraten zu einem Wettbewerbsvorteil, jedoch kann damit die Entwertung nicht vollständig kompensiert werden, wodurch sich Verluste beim Wiederverkauf einstellen. Zu niedrig prognostizierte Restwerte hingegen lassen die Leasingraten ansteigen und machen das Geschäft für den Kunden unattraktiver.

Diese Arbeit stellt Methoden und Verarbeitungsprozesse vor, mit deren Hilfe diese Prognosen auf der Basis von Vergangenheitsdaten realisiert werden können. Zum Einsatz kommen hierfür klassische Modelle aus dem Bereich des Data Mining wie lineare Regressionen und Künstliche Neuronale Netze engl. Artificial Neural Networks (ANN). Das Ziel ist es aus den Daten von bereits durchgeführten Leasinggeschäften zu lernen und damit die Prognose für zukünftige Restwerte mit wissenschaftlich fundierten Methoden Schritt für Schritt zu verbessern.

Die Arbeit ist wie folgt strukturiert. In Abschnitt 2 wird die Problemstellung beschrieben und die aktuelle Literatur zu diesem Thema beleuchtet. Die Datenbasis und verwendete Methoden sind Gegenstand von Abschnitt 3. Analyseergebnisse und Vergleiche der Methoden werden in Abschnitt 4 vorgestellt. In Abschnitt 5 werden ein Prognosemodell eingeführt und Testergebnisse präsentiert. Die Arbeit schließt mit einer kritischen Beurteilung und einem Ausblick in Abschnitt 6.

## 6 Diskussion, Limitationen, Ausblick

In der vorliegenden Arbeit wurde das Restwertisiko im Leasinggeschäft anhand von Daten eines großen deutschen Automobilherstellers untersucht. Künstliche Neuronale Netze kommen für die Modellierung aktueller Marktwerte zum Einsatz. Erst die genaue Abbildung der aktuellen Preissituation ermöglicht die Aufstellung eines Prognosemodells für die zukünftige Marktwertentwicklung. Restwerte sind hier definiert als die Prognose zukünftiger Marktwerte. Ein lineares Regressionsmodell schätzt hierfür den Einfluss des Modellalters auf den Wiederverkaufspreis. Liegt ein statistisch signifikanter Einfluss vor, so werden die aktuellen Marktwerte entsprechend des geschätzten Parameters für die Restwertprognose korrigiert.

Die Herausforderung besteht somit darin zum Zeitpunkt des Abschlusses eines Leasingvertrages den Restwert des Fahrzeugs am Ende der Laufzeit so genau zu prognostizieren, dass die Leasingraten optimal festgelegt werden können, um gleichermaßen wettbewerbsfähig und profitabel zu sein. Das hier vorgestellte Prognoseverfahren bildet im ersten Schritt die aktuellen Marktwerte möglichst genau anhand von vergangenen Transaktionen ab und korrigiert diese Ergebnisse in einem zweiten Schritt um den Einfluss des fortschreitenden Modellalters. Die Analyse beruht aktuell nur auf Daten von Anfang 2010 bis Ende 2014, sodass keine Daten in wirtschaftlichen Krisenzeiten enthalten sind und keine externen Schocks (beispielsweise eine Umweltprämie für Neuwagen) untersucht werden können. Hierfür sind weitere Jahre der Datenerhebung notwendig.

Zudem ist die Anzahl erklärender Variablen noch sehr eingeschränkt. Möglicherweise entscheidende Merkmale wie die Farbe des Fahrzeugs, oder fahrzeugspezifische Ausstattungsmerkmale können zur Zeit nicht identifiziert werden. Ebenso könnte der Kunde selbst einen Einfluss auf den Restwert des Fahrzeugs haben, da der Einsatzzweck (beispielsweise Baustellenfahrzeug oder reiner Dienstwagen) Aufschluss über den zu erwartenden Verschleiß bieten kann. Somit fehlen wichtige Informationen für eine noch individuellere Einschätzung.

Weiterhin wurden in dieser Arbeit erste Erkenntnisse über saisonale/zyklische Einflüsse untersucht. Auch wenn die Datenbasis von fünf Jahren zu gering ist, um eine valide Aussage treffen zu können, zeigen sich Tendenzen, dass der Gebrauchtwagenmarkt innerhalb eines Jahres wiederkehrende Muster aufweist und auch langfristige Tendenzen erkennbar sind (siehe Abschnitt 4.2). Mit größerer Datenbasis können hieraus zukünftig Informationen generiert werden, die es ermöglichen die einfache lineare Korrektur der Marktwerte um diese saisonalen Einflüsse zu erweitern, sollten sich klare und wiederkehrende Zyklen feststellen lassen. Hierfür könnten auch makroökonomische Indikatoren eine wichtige Rolle spielen. Eine Analyse der Erklärungskraft dieser Indikatoren (wie beispielsweise Börsenkurse, Zinsen oder Inflation) ist ein spannendes Feld für zukünftige Untersuchungen.

Abschließend muss bei allen vorgestellten Analysen und technischen Möglichkeiten immer auch der Faktor Mensch beachtet werden. Ist die Akzeptanz der durch künstliche Intelligenz generierten Ergebnisse ausreichend, um das Vertrauen langjähriger Fachexperten zu gewinnen? Wie werden die Ergebnisse tatsächlich genutzt? Aktuell werden die Preise vom Menschen gemacht und datengetriebene Methoden können nur eine Entscheidungsunterstützung liefern. In welchem Maß auf die künstliche Intelligenz vertraut wird, bleibt den Menschen noch selbst überlassen.