

(Thema)

Predictive Analytics im Controlling - Vergleich von Methoden im Automobilssektor

---

---

---

## Masterarbeit

zur Erlangung des akademischen Grades „Master of Science (M.Sc.)“ im  
Studiengang Wirtschaftsingenieur der Fakultät für Elektrotechnik und  
Informatik, Fakultät für Maschinenbau und der  
Wirtschaftswissenschaftlichen Fakultät der Leibniz Universität Hannover

vorgelegt von

Name: Kost Vorname: Leonard

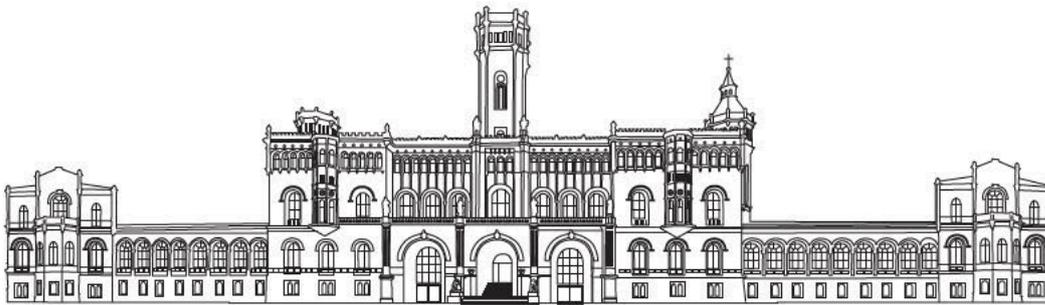
Geb. am:                      in:                     

Prüfer: Prof. Dr. M. H. Breitner

Hannover, 02.04.2024

Ort, Datum\*

\*(Abgabedatum)



# Predictive Analytics im Controlling - Vergleich von Methoden im Automobilsektor

Masterarbeit

Leonard Kost



02. April 2024

Betreuer:



Prüfer:

Prof. Dr. M. H. Breitner

Institute für Wirtschaftsinformatik

# Inhaltsverzeichnis

Abkürzungsverzeichnis	iii
Abbildungsverzeichnis	v
1 Einleitung	1
2 Grundlagen	4
2.1 Predictive Analytics . . . . .	4
2.1.1 Big Data . . . . .	7
2.1.2 Statistische Modelle . . . . .	9
2.1.3 Künstliche Intelligenz . . . . .	21
2.2 Bosch . . . . .	34
2.2.1 Unternehmen . . . . .	34
2.2.2 Controlling . . . . .	36
2.2.3 Masterthesis . . . . .	36
3 Stand der Technik	39
3.1 ARIMA . . . . .	39
3.2 LightGBM . . . . .	41
3.3 LSTM . . . . .	45
4 Forschungsdesign und Methodik	50
4.1 Forschungsfragen . . . . .	50
4.2 Literaturanalyse . . . . .	52
4.2.1 Methodik der Literaturanalyse . . . . .	52
4.2.2 Durchführung und Evaluation der Literaturanalyse . . . . .	53
4.3 Methodik . . . . .	56
4.3.1 Pugh-Matrix . . . . .	57
4.3.2 Experteninterview . . . . .	58
4.4 Durchführung . . . . .	60
4.4.1 Einordnung der theoretischen Methoden . . . . .	61
4.4.2 Identifizierung der praktischen Methoden . . . . .	72
4.4.3 Vergleich der Methoden . . . . .	76
5 Ergebnisse	80
5.1 Diskussion . . . . .	83
5.2 Limitation . . . . .	86

6	Fazit und Ausblick	88
A	Appendix	I

## Abkürzungsverzeichnis

AI	Artificial Intelligence
AD	Autonomous Driving
ANN	Artificial Neuronal Network
AR	Auto-Regressive
ARIMA	Auto-Regressive Integrated Moving Average
AUC	Area under the Curve
BACI	Bosch Center for Artificial Intelligence
BBM	Bosch Business Sector Mobility
CEC	Constant Error Carousel
CC	Chassis Systems Control
CNN	Convolutional Neural Network
CPE	Controlling Process Excellence
DL	Deep Learning
EFB	Exclusive Feature Bundling
GOSS	Gradient-based One-Side Sampling
HMI	Human Machine Interface
IoT	Internet of Things
JMP	Junior Manager Program
KI	Künstliche Intelligenz
KPIs	Key Performance Indicators
LightGBM	Light Gradient-Boosting Machine
LSTM	Long Short Term Memory
MA	Mean Average
ML	Machine Learning
MLP	Multiple Linear Regression
MSE	Mean Square Error
N-BEATS	Neural Basis Expansion Analysis
OEM	Original Equipment Manufacturer
PA	Predictive Analytics
PS	Power Solutions
ROC	Receiver Operating Characteristic

RNN	Recurrent Neuronal Network
xBI	Cross Business Intelligence
XC	Cross Domain Computing Solutions
XGBR	Extrem Gradient Booster for Regression
z.B.	zum Beispiel

---

# 1 Einleitung

Zu den wichtigsten Bausteinen einer sinnvollen Managemententscheidung gehörte es schon immer die Zukunft der Geschäfte zu interpretieren und damit einhergehend die Vorhersage der Geschäftszahlen. Somit ist es erforderlich, dass innerhalb eines Unternehmens der Informationsfluss alle Daten gewährleistet ist und die Möglichkeit besteht eine adäquate Vorhersagen der Geschäfte zu haben, um daraus eine Veränderungen des Ecosystems des Unternehmens erkennen zu können. Während ein solcher Vorhersage der zukünftigen Geschäfte in der Vergangenheit durch eine hohes Personalaufkommen bewerkstelligt wurde, können durch die voranschreitende Technik heute schon andere Methoden eingesetzt werden. Durch den technologisch schnellen Wandel und einer Vielzahl von Innovationen sind die Möglichkeiten des Blicks in die Zukunft nicht nur begrenzt auf Experten innerhalb der Firma, sondern werden immer mehr durch Computersysteme unterstützt. Um heutzutage mit dieser Innovation mithalten und ein Unternehmen zukunftssicher aufzustellen zu können, benötigt es nicht nur ein Management, welches Informationen in einer Vielzahl von Unternehmenstätigkeiten besitzt, sondern auch Abteilungen, welche die Umweltfaktoren des Unternehmens deuten können und analysieren welchen Einfluss diese auf die Geschäfte des jeweiligen Unternehmens besitzt. Hier ist die Abteilung des Finanzcontrollings angesiedelt und hält Überblick über alle finanziellen Flüsse, Prozesse und die Zukunft des Unternehmens. Sowohl in der entscheidungstragenden Ebene als auch im Controlling selbst wird immer mehr Personal für die Analyse von Daten eingesetzt, ebenso mehren sich die Auswertungen von Statistiken und saisonalen Erscheinungen in Datensätzen[41]. Weiter wächst der Berg an Daten eines Unternehmens, welche zur Vorhersage herangezogen werden, kontinuierlich an und bedeuten einen enormen Mehraufwand für die Abteilung des Controllings.

Um diesem Trend entgegen zu wirken, den Controllern wieder die Möglichkeit sich strategischen Aufgaben zu widmen und eine zuverlässige Voraussage der Geschäfte in der Zukunft zu gewährleisten wird der Einsatz von Predictive Analytics Anwendungen immer wichtiger.

Durch die Analyse von großen Datensätzen mithilfe von künstlicher Intelligenz können Unternehmen eine verbesserte Vorhersage der relevanten Key Performance Indikator vornehmen und somit, wie oben beschrieben, zukunftssichere Entscheidungen als bisher treffen. Eine solche Vorhersage durch KI-gestützte Modelle lässt sich in nahezu jedem Markt anwenden[15] und ist sowohl auf reine Geschäftszahlen in Form von Geldfluss anwendbar, als auch auf Zahlen wie Absatz und Korrelationen zwischen Produkten und Indikatoren. Das Predictive Analytic Tools sich bereits großer Beliebtheit erfreuen dürfen zeigen einige prominente Beispiele. Amazon, das weltweit führende Unternehmen im Bereich Online Consumer Goods, setzt KI Tools ein, welche den Zusammenhang zwischen gekauften Pro-

---

dukten analysieren und somit voraussagen können, in welcher Korrelation andere Produkte stehen. Diese Erkenntnisse nutzt Amazon wiederum um dem Verbraucher gezielt Produkte anzuzeigen, welche durch die Vorhersage von Predictive Analytic Tools wahrscheinlich auch dem Interesse des Käufers entsprechen. So steigt die Wahrscheinlichkeit, dass der Kunde mehr kauft, als er ursprünglich geplant hatte. Hier werden die Korrelationen zwischen den Produkten vorhergesagt.

Ein weiteres ebenso prominentes Beispiel, welches sich auf den Verkauf von Consumer Goods in physischen Geschäften bezieht, ist Walmart. Diese konnten durch Verwendung von KI-Tools erkennen, dass in den Geschäften welche in einer Hurrikan Region liegen, nicht nur überdurchschnittlich viele Taschenlampen zur Hurrikan Saison verkauft wurden, sondern parallel dazu auch „Strawberry Pop Tats“. Durch diese Erkenntnis konnte der Lagerbestand und somit auch der Absatz von eben solchen Gegenständen erhöht werden. Hier werden die Vorhersagen auf Grund einer Korrelation zwischen einem Indikator in Form des Wetters und einer Produktreihe getätigt. Weiterhin nutzt Walmart Predictive Analytics um die Anzahl der Kunden zu einer bestimmten Tages- und Wochenzeit vorauszusagen und so die optimale Anzahl an Mitarbeitern zur jeweiligen Zeit zur Verfügung zu stellen [44].

Alle diese Beispiele zeigen, wie nützlich Vorhersagen von elementaren KPIs eines Unternehmens sein kann. Dabei helfen sie dem Unternehmen dabei Prozesse zu optimieren, Kosten einzusparen, ein besseres Verständnis der Kunden zu erlangen oder sogar Probleme und Möglichkeiten zu identifizieren bevor diese überhaupt eintreten können.

In dieser Arbeit soll sich dabei mit der Fragestellung auseinandergesetzt werden, welche Methode für das Controlling eines Unternehmens im Automobilssektor am besten geeignet ist. Der Hauptfokus liegt dabei auf der Vorhersage von KPIs durch univariabel Modelle, welche mittels historischer Daten des selben KPIs antrainiert werden. Dies beschreibt ein klassisches Zeitreihenproblem. Bei der Betrachtung der Vorhersage der KPIs soll eine Vielzahl von Modellen untersucht werden um eine lückenlose Identifizierung des besten Modells vorzunehmen. In einer weiteren Fragestellung soll dargelegt werden, welche Methoden gegenwärtig in eben dieser Abteilung genutzt werden. Dabei soll auch die bisherige Entwicklung, als auch die Qualität der Vorhersage untersucht werden. Aber auch der Trend, welcher in dem Unternehmen vorherrscht soll betrachtet werden, um diesen mit Forschungstrends abgleichen zu können. Die abschließende Fragestellung der Thesis soll sich mit dem Vergleich der Modelle aus den beiden vorangegangenen Fragestellungen beschäftigen. Hierbei soll untersucht werden, ob State-of-the-Art und industrieller Ansatz von einander abweichen und falls dies der Fall sein sollte, darzulegen, wie es zu erklären ist und in welchen Kriterien sich die beiden identifizierten Modelle unterscheiden.

Um diese Forschungsfragen zu beantworten, werden verschiedene Methoden genutzt. Die

---

Arbeit wird dabei wie folgt strukturiert aufgebaut. Das Thema AI wird anfänglich durch eine Literaturrecherche erklärt. Dabei wird besonders auf die mathematischen Grundlagen von statistischen, Machine Learning und Deep Learning Modellen eingegangen. Es werden gezielt jene mathematischen Grundlagen erklärt, welche für eben diese Modelle relevant sind, die in dem Hauptteil der Arbeit besonders beleuchtet werden. Aber auch das zu untersuchte Unternehmen, Robert Bosch Cross Domain Computing Solutions, wird kurz vorgestellt. Ebenso die Abteilung, in der diese Methoden angewandt werden, dem Finanzcontrolling. Darauf folgend soll der aktuelle Stand der Technik analysiert werden. Hierbei werden verschiedene Modelle genauer betrachtet, welche für den Verlauf der Arbeit essentiell sind. Es soll deutlich gemacht werden, auf welcher mathematischen Grundlage und speziellem Aufbau sich diese unterscheiden. Dabei sollen sowohl die Vorteile unterstrichen werden, als auch erklärt werden wie diese Methoden funktionieren. Der Hauptteil beschäftigt sich mit verschiedenen Methodiken um die Forschungsfragen zu beantworten. Während die Identifizierung des theoretischen Modells durch eine Pugh-Matrix durchgeführt wird, werden die praktischen Modelle durch Experteninterviews innerhalb der Abteilung, aber auch abteilungsübergreifend und geschäftsbereichübergreifend durchgeführt. Schlussendlich sollen die beiden Ergebnisse dieser Methodiken verglichen werden. Nach der theoretischen Betrachtung der Methodiken und deren Durchführungen, endet die Thesis mit der Darstellung der Ergebnisse aller drei Forschungsfragen und der Interpretation und Diskussion der Ergebnisse. Hier soll vor allem darauf eingegangen werden, welche Rückschlüsse zwischen Forschungsstand und Industriestandart gezogen werden können. Es folgt ein Fazit und ein Ausblick zu einem möglichen Weiterführen des Themas.

Das Thema Predictive Analytics ist schnelllebig, die Methoden und Modelle werden kontinuierlich weiterentwickelt und die Ergebnisse übertreffen sich regelmäßig. Verschiedene Literatur, Konferenzen und Wettbewerbe ziehen immer wieder neue Schlüsse, die Technologie macht einen Sprung nach dem anderen und setzt immer neue Maßstäbe. Gerade deswegen ist diese Sektor gefragt wie noch nie, Konzerne optimieren sich durch AI-Themen, Softwareentwickler und Know-How sind gefragter den je. Gleichzeitig stellt sich die Frage, wie Unternehmen mit diesem Tempo umgehen. Können neue Erkenntnisse schnell genug implementiert werden, oder sind diese bereits wieder überholt, wenn diese ausgerollt werden?

---

## 6 Fazit und Ausblick

In dieser Arbeit konnte sowohl der beste theoretische Ansatz für den Use Case der Vorhersage von Geschäftszahlen im Controlling identifiziert werden, als auch der praktisch genutzte. Dabei wurde sowohl Literatur genutzt, um verschiedene Ansätze zu identifizieren, als auch die Anforderungen an das Modell durch die Aufgabenstellung innerhalb des Bosch dargestellt. Durch diese Kombination konnten die Modelle bewertet werden und eine Dominanz der Machine Learning Modelle in Form des LightGBMs identifiziert werden. In der praktischen Betrachtung konnte durch die Befragung mehrerer Experten aus unterschiedlichen Bereichen und Abteilungen sowohl subjektive Wahrnehmungen, als auch objektive Daten erhoben werden. Diese führen nicht nur zu einem besseren Verständnis des Nutzen von Predictive Analytics innerhalb des Boschs, sondern zeigen schlussendlich auch ein klares Bild über die genutzten Methoden in diesem Unternehmen auf. So wird nach der Analyse der Interviewergebnisse und der Datenlage klar, dass Bosch vor allem statistische Modelle wie ARIMA für die Vorhersage der Geschäftszahlen nutzt. Nur selten haben Modelle aus der Kategorie des Machine Learnings einen Einfluss auf das Ergebnis des im Bosch genutzten Metalearners. Deep Learning Modelle kommen bis dato nicht zum Einsatz. In der letzten Fragestellung der Arbeit, dem Vergleich zwischen praktischer und theoretischer Methode, wird deutlich, dass eine Zeitversetzung der Methoden zwischen Industrieanwendung und Forschung besteht. So ist bei Bosch der selbe Trend hin zum Machine Learning und weg von statistischen Methoden zu erkennen, welche vor wenigen Jahren auch in der Forschung vollzogen wurde. Somit lässt sich interpretieren, dass die Forschungsergebnisse auch für die Industrie richtungsweisend sind. Die Implementierung der Forschungs-Insights benötigen allerdings einige Jahre, um auf eine Unternehmensgröße wie von Bosch skaliert zu werden. Es ist davon auszugehen, dass dieser zeitversetzte Trend auch in Zukunft fortgesetzt wird. So ist in der Forschung bereits zu erkennen, dass Deep Learning Modelle immer mehr Relevanz besitzen und allmählich in den Rankings der Benchmarkwettbewerbe auftauchen. Auch Bosch unternimmt erste Versuche Deep Learning Modelle innerhalb des Metalearners zu implementieren. Diese besitzen noch keinen regelmäßigen Einfluss auf das Ergebnis des Metalearners. Eine weitere Erkenntnis ist der Vergleich der Qualität und dem aktuellen Wissensstand der drei unterschiedlichen Modellart innerhalb der künstlichen Intelligenz. So wird in dieser Arbeit immer wieder zwischen den verschiedenen Modellarten unterschieden. Vor allem der Vergleich zwischen statistischen Modellen und Machine Learning Modellen wird genauer betrachtet, was allem voran an der qualitativen Nähe und dem Wechsel des Best-Practice-Ansatzes liegt. Aber auch Deep Learning Modelle werden betrachtet, erklärt und schlussendlich in Hinblick auf die Zukunft eingeordnet. Weiter wird an diesen fest gemacht, wie sich die Forschung und der State-of-the-Art-Ansatz entwickeln. Eine deutliche Abgrenzung der verschiedenen Ansät-

---

zen der künstlichen Intelligenz kann durch diese Arbeit vorgenommen werden. Insgesamt lässt sich sagen, dass die theoretischen Ansätze nicht augenblicklich in der Industrie umgesetzt werden können. Somit entsteht immer ein Lücke zwischen dem Ansatz, welcher als bestes Einzelmodell oder Kombinationsmodell identifiziert wird und eben dem, der in der Industrie genutzt wird. Es ist zu unterstreichen, dass diese sich nur zeitlich unterscheiden und nicht vollständig andere Ansätze verfolgen. So ist zu wiederholen, dass in dieser Arbeit zu erkennen ist, dass Bosch dem Zug des Forschungsstands mit einem Abstand von wenigen Jahren folgt und diesen auch aktiv beobachtet.

Aus dieser Arbeit haben sich einige interessante Ansätze herauskristallisiert, die in folgenden Arbeiten gut zu betrachten wären. So könnte ein Ausblick auf die Vergleichbarkeit der unterschiedlichen Wissensstände der führenden Unternehmen sein. Sowohl um ein besseres Verständnis zubekommen, welche Unternehmen bereits Predictive Analytics anwenden, als auch um die im Abschnitt 5.2 geäußerten Zweifel zu bestätigen oder zu eliminieren. Eine weitere mögliche Fragestellung aus dieser Arbeit, wäre es zu untersuchen, was für den Zeitversatz zwischen Forschung und Industrie sorgt. Hier stellt sich die Frage, wieso Forschungsstände nicht schneller auf reale Use Cases anwendbar sind und zum Teil mehrere Jahre vergehen. Weiter ist fraglich, ob dieser Versatz mit einem fehlenden Tracking des Forschungsstandes zu Stande kommt oder ob die Implementierung der neusten Modelle einen Großteil der Zeit konsumieren. Eine letzte Folgeaufgabe könnte die Überprüfung der Qualität der Modelle auf die gegebenen Daten und die Aufgabenstellung sein. Da das Kriterium der Qualität in der Pugh-Matrix am höchsten bewertet wurde, wäre eine Überprüfung der Literaturergebnisse durch Anwendung der sechs Modelle auf einen Controllingdatensatz sinnvoll. So könnte das Ergebnis aus der Pugh-Matrix, soweit identisch, unterstrichen werden.