

# **Automatisierte Intraday Prognosen mit Neurosimulation**

## **Diplomarbeit**

zur Erlangung des Grades eines  
Diplom-Ökonomen der Wirtschaftswissenschaftlichen Fakultät  
der Leibniz Universität Hannover

vorgelegt von:

Christoph Polus,



Erstprüfer: Prof. Dr. Michael H. Breitner

Hannover, 11. November 2010

# Inhaltsverzeichnis

Inhaltsverzeichnis.....	I
Tabellenverzeichnis.....	IV
Abbildungsverzeichnis.....	V
Abkürzungsverzeichnis.....	VII
1 Einführung.....	1
1.1 Motivation und Zielsetzung.....	1
1.2 Aufbau der Arbeit.....	3
2 Grundlagen und Definitionen.....	4
2.1 Kapitalmarkttheorien.....	4
2.1.1 Stochastische Prozesse.....	4
2.1.2 Markov-Ketten.....	5
2.1.3 Effizienzmarkt-Hypothese.....	6
2.1.4 Random-Walk-Hypothese.....	7
2.1.5 Behavioral Finance.....	8
2.2 Formen der Wertpapieranalyse.....	9
2.2.1 Fundamentalanalyse.....	10
2.2.2 Technische Analyse.....	11
2.2.3 Quantitative Analyse.....	13
2.3 Implikationen der Kapitalmarkttheorien für die Wertpapieranalyse.....	13
2.4 Grundlegende Eigenschaften und Analyse von Finanzzeitreihen.....	16
2.5 Künstliche Neuronale Netze.....	19
2.5.1 Grundlegende Eigenschaften von KNN.....	19
2.5.2 Das Multi-Layer Perceptron.....	22
2.6 Der Neurosimulator FAUN 1.1.....	23
2.7 Zeitreihenprognose mittels Künstlicher Neuronaler Netze.....	25
2.8 Genetische Algorithmen.....	26
3 Beschreibung des verwendeten Genetischen Algorithmus.....	29
3.1 Aufbau und Funktionsweise des Genetischen Algorithmus.....	31
3.2 Binärcodierung.....	33
3.3 Crossover.....	36
3.4 Mutation.....	38

4 Konzeption und Entwicklung des Prototyps zur automatisierten Prognose von Finanzzeitreihen .....	39
4.1 Einleitung .....	39
4.2 Konzeptionierung der Problemstellung.....	40
4.3 Bestimmung der zur Realisierung benötigten Grundelemente .....	41
4.4 Die Programmiersprache Java.....	42
4.5 Softwaretechnische Umsetzung .....	43
4.5.1 Analyse des zu erstellenden automatisierten Prognosesystems .....	44
4.5.2 System Design und Implementierung .....	45
4.5.2.1 Anbindung an die Interactive Brokers API.....	45
4.5.2.2 Prognosekomponente.....	48
4.5.2.3 Optimierungskomponente.....	49
5 Methodologie zur Prognose von ausgewählten Finanzzeitreihen.....	54
5.1 Auswahl zu prognostizierender Finanzzeitreihen .....	55
5.1.1 Der Währungsmarkt .....	58
5.1.2 Fixed Income.....	58
5.1.3 Aktien-Märkte .....	59
5.2 Auswahl der Netzinputfaktoren.....	61
5.2.1 Renditen .....	63
5.2.2 Indikatoren .....	63
5.2.2.1 Einfacher und exponentieller Moving Average.....	64
5.2.2.2 Momentum.....	65
5.2.2.3 Relative Strenth Index .....	65
5.2.3 Modellbildung .....	66
5.3 Datenvorverarbeitung.....	73
5.3.1 Besondere Eigenschaften von Intraday-Daten .....	73
5.3.2 Statistische Analyse der Zeitreihen-Momente und Eigeschaften.....	75
5.3.3 Ausreißersuche und deren Eliminierung.....	84
5.3.4 Stationarisierung und Skalierung der Daten.....	88
5.4 Bestimmung der Trainings- und Testzeiträume .....	89
5.5 Training der neuronalen Netze .....	93
5.5.1 Wahl der Netztopologie .....	93
5.5.2 FAUN-Parametereinstellung.....	95
5.5.3 Fehlermaße und Risikokennzahlen .....	97
5.6 Evaluierung der Trainingsergebnisse .....	100

5.7	Prognosedurchführung .....	103
5.7.1	Prognosegegenstand und Generalisierungszeitraum .....	103
5.7.2	Evaluierung der Prognoseergebnisse.....	107
5.7.3	Backtest des Prognosesystems anhand der out of Sample-Daten ..	119
5.7.4	Validierung des Backtests durch Live-Test des Prognosesystems ..	124
6	Fazit und Ausblick .....	126
	Literaturverzeichnis.....	129
	Anhang .....	i

# 1. Einführung

## 1.1 Motivation und Zielsetzung

Das Bestreben den zukünftigen Verlauf einer Zeitreihe zu prognostizieren hat vor allem in der Kapitalmarktforschung sowie der praktischen Finanzwirtschaft die Entwicklung vielfältiger Methoden aus unterschiedlichen Disziplinen begünstigt.

Die Motivation zur Erstellung von Prognosemodellen ist die Gewinnung eines Informationsvorsprungs gegenüber anderen Marktakteuren um eine positive Rendite zu erwirtschaften. Die Verbreitungsgeschwindigkeit und die Menge an global verfügbaren Informationen nehmen jedoch aufgrund des technischen Fortschritts stetig zu, so dass Informationsasymmetrien sowie Vorteile für den Einzelnen stetig abgebaut werden und somit die Markteffizienz steigt. Schließlich nimmt auch die Genauigkeit von Vorhersagen mit steigendem Prognosehorizont ab, da es zum einen immer schwieriger wird aus der Fülle der Informationen diejenigen zu filtern, welche den Verlauf einer Zeitreihe beeinflussen. Zum anderen werden die Informationen mit steigender Geschwindigkeit eingepreist, so dass bei unerwarteten Ereignissen, welche nicht für die Prognose berücksichtigt wurden, eine abweichende Korrekturbewegung erfolgt.

Die zunehmende Geschwindigkeit mit der Informationen bereitgestellt und verarbeitet werden sowie der kontinuierliche technische Fortschritt wirken sich zudem auf die Frequenz des Wertpapierhandels an den Kapitalmärkten aus. Beispielsweise ist die durchschnittliche Haltedauer einer Aktie in den letzten 25 Jahren um 90 Prozent gesunken.<sup>1</sup>

Gleichzeitig nimmt der Anteil des computergesteuerten Wertpapierhandels vornehmlich durch Großbanken im hochfrequenten Intraday-Bereich stetig zu.<sup>2</sup> Eine Konsequenz ist unter anderem, dass die Zeitfenster in denen die Marktsituation analysiert und Handelsentscheidungen getroffen werden können zunehmend kürzer werden.

Je mehr der beschriebene Entwicklungsprozess fortschreitet, desto nützlicher erscheinen computergestützte kurzfristige Prognosesysteme, deren Ziel es ist die Entscheidungsfindung im Wertpapierhandel erleichtern.

---

<sup>1</sup> Durchschnittliche Haltedauer einer Aktie nach Fama und French, vgl. World Federation of Exchanges, <http://www.world-exchanges.org/>, Stand: 02.11.2010, sowie Eisenhofer 2008 S.18

<sup>2</sup> Vgl. Jahresbericht NYSE, <http://www.nyse.com>, Stand: 02.11.2010, sowie Eisenhofer 2008

Neben mathematisch statistischen Zeitreihen- und Volatilitätsmodellen sowie Methoden der Populärwissenschaft wie der chart- und markttechnischen Analyse, haben sich für die Prognose von Finanzzeitreihen in den vergangenen Jahren vor allem die Neuronale Netze aus dem Forschungsbereich der Künstlichen Intelligenz hervorgetan.

Die Idee der Neuronale Netze leitet sich aus der Struktur des menschlichen Gehirns und des Nervensystems ab.<sup>3</sup> Sie bilden des Weiteren eine Klasse von nichtlinearen sowie nichtparametrischen Modellen. Im Gegensatz zu anderen statistischen Verfahren können sie komplexe nichtlineare Beziehungen, wie sie in Zeitreihen vorkommen, ohne spezifische Kenntnisse über die Verteilungseigenschaften der Daten abbilden.<sup>4</sup> Auch weisen Neuronale Netze ein weitgehend robustes Verhalten gegenüber verrauschten Datensätzen auf.<sup>5</sup>

Diese und weitere Faktoren sind nur einige Argumente für den Gebrauch von Neuronale Netzen zur Prognose von Finanzzeitreihen. Aufgrund dessen dienen diese in der vorliegenden Arbeit als Basis für die Entwicklung eines Prototyps zur automatisierten Intraday-Prognose von Finanzzeitreihen.

Ziel ist es, mit Hilfe der Neuronale Netze für ausgewählte Basiswerte Prognosen auf Minutenbasis zu erstellen, welche für die Entscheidungsunterstützung im Wertpapierhandel genutzt werden können. Weiterhin soll der Prototyp alle Schritte der Prognose von der Informationsbeschaffung und Verarbeitung bis hin zur Ausgabe des Prognosewertes automatisiert durchführen. Weiterhin kommen Genetische Algorithmen zum Einsatz um die Anpassungsgüte der Neuronale Netze sowie die Prognoseergebnisse zu optimieren. Zudem werden für die Validierung von Prognosestrategien Maximierungsprobleme in Bezug auf die optimale Gestaltung von Handelsregeln mit Genetischen Algorithmen approximiert.

Bevor eine Prognose von Zeitreihen durchgeführt werden kann, sind jedoch einige grundlegende Fragestellungen zu beantworten. Beispielsweise ist im Hinblick auf die Erkenntnisse der etablierten Kapitalmarkttheorien zu klären, inwiefern eine Prognose von Finanzzeitreihen möglich ist. Weiterhin ist zu analysieren welche Zeitreihen sich überhaupt für eine Prognose eignen und schließlich müssen geeignete Netztopologien sowie brauchbare Netzinputstrukturen bestimmt werden. Nach zufriedenstellen-

---

<sup>3</sup> Vgl. Borgelt, Klawonn 2003, S. 5

<sup>4</sup> Vgl. Erdmann, Zweifel 2008, S. 76

<sup>5</sup> Vgl. Lange 2003, S. 54

der Beantwortung dieser und einiger weiterer Fragen, kann die Entwicklung des Prognosesystems beginnen.

## **1.2 Aufbau der Arbeit**

Die Arbeit gliedert sich wie folgt: Kapitel zwei erörtert grundlegende Fragestellungen und Begriffsdefinitionen. Zu Beginn werden die Positionen der etablierten Kapitalmarkttheorien in Bezug auf die Prognostizierbarkeit von Zeitreihen evaluiert sowie die gängigen Formen der Wertpapieranalyse vorgestellt und diskutiert, inwiefern diese für die Zeitreihenprognose zielführend sind.

Nachstehend werden die Charakteristika von Zeitreihen präsentiert und Eigenschaften erläutert, welche für eine Zeitreihenprognose von Bedeutung sind. Die folgenden Abschnitte beschäftigen sich mit den technischen und mathematischen Grundlagen Neuronaler Netze, die anschließend für die Eignung zur Prognose von Finanzzeitreihen evaluiert werden. Der Grundlagenteil beschließt mit der Erläuterung des Grundprinzips und der Funktionsweise von Genetischen Algorithmen in Kapitel 2.8.

Die praktische Anwendung sowie Beschreibung des in dieser Arbeit verwendeten Genetischen Algorithmus ist anschließend Gegenstand von Kapitel drei. Mit der Erarbeitung, Entwicklung sowie der Beschreibung jener Komponenten des zu konstruierenden Prognosesystem-Prototypen befasst sich Kapitel vier.

In Abschnitt fünf erfolgt die Ausarbeitung und Gestaltung eines Prognoseansatzes, welcher schließlich in dem zu entwickelnden System zum Einsatz kommen soll. Dabei orientiert sich die Vorgehensweise an einer systematischen Methodologie, welche sich in acht Prozessstufen unterteilt. Nachstehend sind entsprechende Finanzzeitreihen zu bestimmen, die prognostiziert werden sollen. Anhand der Eigenschaften und statistischen Merkmale jener Zeitreihen, erfolgt die Selektion von zweckmäßigen Netzinputfaktoren sowie eine angemessene Datenvorverarbeitung. Schließlich müssen die vorliegenden Daten in geeignete Trainings-, Validierungs- und Testzeiträume aufgeteilt und im gleichen Schritt ein geeigneter Prognosehorizont sowie Zeitraum festgelegt werden.

Im Anschluss erfolgt, hinsichtlich der Charakteristika der Trainingsdaten, die Wahl einer Netztopologie sowie die Einstellung wichtiger Trainingsparameter. Nach erfolgreichem Training der Neuronalen Netze können diese evaluiert und die besten für eine anschließende Prognose ausgewählt werden. Der letzte Schritt beurteilt schließlich die Prognose anhand verschiedener Performance-Maße. Dennoch kann nur unter

realen Bedingungen getestet werden inwiefern eine Strategie für ein Prognosesystem geeignet ist. Dieser Thematik widmet sich Kapitel sechs.

Die Arbeit schließt mit einem Resümee der gewonnenen Ergebnisse und gibt einen Ausblick über weiterführende Modifikationen des automatisierten Prognoseansatzes.

## **2. Grundlagen und Definitionen**

Bevor die zentrale Fragestellung der Arbeit in den Mittelpunkt der Betrachtung rückt, ist ein Fundament von grundlegenden technischen sowie fachlichen Begriffen und Definitionen zu erarbeiten. Es schafft die Voraussetzung für das weitere Verständnis der zu behandelnden und den sich daraus resultierenden Problemstellungen.

### **2.1 Kapitalmarkttheorien**

In diesem Abschnitt werden die verschiedenen Kapitalmarkttheorien und deren Implikationen bezüglich der Prognostizierbarkeit von zukünftigen Ereignissen erörtert. Der grundlegende Einblick in die Kapitalmarkttheorien soll das wissenschaftlich vorherrschende Verständnis für die Preisbildungsprozesse und die Informationsverarbeitung der Märkte verdeutlichen. Es ist die Frage zu klären, inwiefern die Prognose von zukünftigen Preisbewegungen durch die Theorien gestützt und legitimiert wird. Des Weiteren, bilden die Theorien das Fundament für die verschiedenen Analysemethoden die im Anschluss erläutert und dafür genutzt werden den Verlauf von Wertpapierkursen und somit von Finanzzeitreihen zu beschreiben und möglicherweise zu prognostizieren.

Das Fundament einiger theoretischer Ansätze bildet das Verständnis zur Entstehung von Zeitreihen. Diese wird einleitend anhand der stochastischen Prozesse näher erläutert.

#### **2.1.1 Stochastische Prozesse**

Die Theorie der stochastischen Prozesse liefert einen Ansatz zur Entstehung und Beschreibung von Zeitreihen. Grob formuliert bezeichnet eine Zeitreihe eine zeitlich



Live Prognose	Handelszeitraum	Kum. Log. Rendite	Durchschnittliche Rendite pro Periode
Fünf Minuten ES-Future	2,66 Stunden 32 Perioden a fünf Minuten	0,9%	$(0,9\%)/(32) = 0,026\%$

Tabelle 26: ES-Future Life-Test, Quelle: Selbst erstellt

Live Prognose Fehlermaße	Correct Direction	MAPE	RMSE	MAE
Fünf Minuten ES-Future	65,52%	84,24%	0,000146	0,000107

Tabelle 27: Fehlermetriken ES-Future Life-Test, Quelle: Selbst erstellt

Da es sich um einen vergleichsweise kurzen Zeitraum handelt, ist die erzielte Rendite schwierig im Vergleich zu den Backtest-Ergebnissen zu beurteilen. Es ist jedoch hervorzuheben, dass überhaupt eine positive Rendite erzielt werden konnte.<sup>149</sup> Bei einem Blick auf die Fehlermaße der Prognose fällt auf, dass die Werte etwas schlechter ausfallen. Das Maß der Correct Direction ist jedoch noch deutlich über 50 Prozent, so dass mehrheitlich eine korrekte Richtungsangabe erfolgt. Das Abweichen der Kennzahlen ist auch mit der Stichprobengröße zu erklären. Es ist sehr wahrscheinlich, dass diese zu klein ist um ein repräsentatives Ergebnis bereitzustellen. Für eine intensive Betrachtung und eine valide Beurteilung der Backtest-Ergebnisse sind weitere Test-Läufe notwendig, die vor allem einen längeren Zeitraum umfassen.

## 6. Fazit und Ausblick

Die vorliegende Arbeit befasst sich mit der Erarbeitung eines Prognoseansatzes für Finanzzeitreihen auf Basis von Neuronalen Netzen, welcher zudem für die Entwicklung eines Prototyps zur automatisierten Intraday-Prognose genutzt wird. Des Weiteren kommen Genetische Algorithmen für die Lösung verschiedener Optimierungs-

<sup>149</sup> Die Dokumentation des Prognoselaufs findet sich auf dem beigegeführten Datenträger

probleme, die im Zusammenhang mit Prognosen mittels Neuronaler Netze auftreten können, zum Einsatz.

Bevor die zentrale Aufgabenstellung der Arbeit in den Fokus der Betrachtung rückt, werden einige grundlegende Fragestellungen bezüglich der Prognostizierbarkeit von Zeitreihen näher erläutert und diskutiert. Beispielsweise ist im Hinblick auf die Erkenntnisse der etablierten Kapitalmarkttheorien zu klären, inwiefern eine Prognose von Finanzzeitreihen möglich ist. Weiterhin ist zu analysieren welche Zeitreihen sich überhaupt für eine Prognose eignen. Vor allem im Hinblick auf Intraday-Daten existieren spezielle Charakteristika die es für die Bildung eines Prognosemodells zu beachten gilt. Schließlich müssen geeignete Netztopologien sowie brauchbare Netzinputstrukturen bestimmt werden. Nach zufriedenstellender Beantwortung dieser und einiger weiterer Fragen beginnt die Entwicklung der Prognosestrategie sowie des automatisierten Prototypen.

Der System-Prototyp wurde schließlich so konzipiert, dass es historische sowie real-time Daten automatisiert in beliebigen Frequenzen beziehen kann. Zudem werden die bezogenen Daten aufbereitet und transformiert, so dass sie für eine Prognose mit Hilfe der trainierten Neuronalen Netze verwendet werden können. Im Anschluss wird der Netz-Output evaluiert und anhand der Ergebnisse über die Anbindung an den Broker das Ausführen bestimmter Aktivitäten ermöglicht. Dies kann zum Beispiel eine Orderplatzierung sein. Desweiteren werden die Aktivitäten dokumentiert und gespeichert, damit eine Rekonstruktion des Systemablaufs sowie eine mögliche Fehlerbehebung erleichtert werden. Zuletzt wird die Qualität der Prognosen anhand einer Fehlermetrik in regelmäßigen Abständen beurteilt. Bei Verschlechterung der Prognosegüte können die Neuronalen Netze mit Hilfe eines Genetischen Algorithmus manuell nachtrainiert werden.

Um den Prognoseansatz zu testen, wurden schließlich mit Hilfe der Neuronalen Netze für ausgewählte Basiswerte Prognosen auf Minutenbasis erstellt. Die betrachteten Zeithorizonte umfassen sowohl eine Periodizität von drei als auch fünf Minuten. Die robuste Prognosestrategie konnte anhand eines out-of Sample-Backtests schließlich bestätigt werden. Dabei richtet sich der Fokus ebenfalls auf die Generierung von geeigneten Handelssignalen, die mit Hilfe eines Genetischen Algorithmus bestimmt werden. Eine Validierung der Backtest-Ergebnisse anhand eines Live-Tests der Prognosestrategien konnte aufgrund der begrenzten Bearbeitungszeit nicht in repräsentativer Form erfolgen. Dieser Punkt muss Gegenstand weiterer Betrachtungen

sein, um die Prognosestrategie auf ein solides Verhalten im realen Marktumfeld zu testen.

Der vorgestellte Prognoseansatz sowie das automatisierte System stellen nur einen grundlegenden Ansatz dar. Vor allem im Bereich der zu implementierenden Handelsstrategien sind unzählige Modifikationen und somit eventuelle Renditesteigerungen möglich. Desweiteren ist die Bestimmung der Handelssignale mit Hilfe des Genetischen Algorithmus nur eine Variante von vielen. Auch hier könnte eine Neudefinition der einzusetzenden Schwellwerte bessere Ergebnisse liefern. Zudem ist die Anpassungsgüte der Neuronalen Netze ein wichtiger Punkt mit Modifikationspotenzial. Je genauer die Zielgröße approximiert wird, desto rentabler ist die Prognosestrategie. Zusätzliche oder stärkere Vorverarbeitungsmaßnahmen der Rohdaten könnten auch hier die Qualität der Prognose verbessern. Wie zu erkennen ist, ergeben sich eine Reihe von Optimierungsproblemen. Wie intensiv jede dieser Aufgaben bearbeitet und gelöst wird, bleibt letztlich dem Anwender überlassen.

*„Der Kampf gegen Gipfel vermag ein Menschenherz auszufüllen. Wir müssen uns Sisyphus als einen glücklichen Menschen vorstellen.“*

*Albert Camus, 1942 (Le mythe de Sisyphe.)*