
Nutzung von Data-Mining-Verfahren zur Indexprognose

Jonas Rommelspacher

Institut für Wirtschaftsinformatik, Philipps-Universität Marburg, Germany
rommelspacher@wiwi.uni-marburg.de

1 Problemstellung

Finanzmarktakteure müssen ihren Investitionsentscheidungen Erwartungen bzgl. der zukünftigen Marktentwicklung zugrunde legen. Sie stehen vor einer Entscheidung unter Unsicherheit und sind bestrebt mittels Prognoseverfahren die zukünftige Kursentwicklung möglichst gut vorherzusagen.

Die klassische Kapitalmarkttheorie basiert im Kern auf der Theorie der rationalen Erwartungen (Muth) und der Hypothese der effizienten Informationsverarbeitung von Finanzmärkten (Fama). In ihrer schwachen Form postuliert letztere, dass die gegenwärtigen Preise alle Informationen der Vergangenheit widerspiegeln und die zukünftigen Preise einem Random Walk folgen [3]. Weiter besagt diese Hypothese, dass es unmöglich ist, mit vergangenen Kursen zukünftige Kurse vorherzusagen [7]. Somit steht die Hypothese der effizienten Informationsverarbeitung im Gegensatz zu der o.g. Anforderung der Finanzmarktakteure. Es besteht allerdings die Vermutung, dass die klassischen linearen Modelle aufgrund von Nichtlinearitäten in Finanzmarktreihen versagen [11].

Neben der Tatsache, dass einige Korrelationen an Finanzmärkten kaum zu leugnen sind, gilt es mittlerweile als relativ gesichert, dass aufgrund von Nichtlinearitäten durchaus bis zu einem gewissen Grad erfolgreiche Prognosen erstellt werden können [4].

Letztendlich handelt es sich bei der Prognose zukünftiger Kurse, gleich ob dafür mathematische Modelle oder heuristischer Verfahren (Gleitende Durchschnitte, Technische Analyse) eingesetzt werden, um ein Mustererkennungsproblem [5]. Zur Erkennung von Datenmustern werden üblicherweise Data-Mining-Verfahren aus dem Forschungsgebiet der Künstlichen Intelligenz eingesetzt.

Ausgehend von der geschilderten Problematik soll in diesem Beitrag diskutiert werden, inwiefern Data-Mining-Verfahren für die Prognose von Aktienindizes eingesetzt werden können.

Der Beitrag ist folgendermaßen aufgebaut: In (2) wird das Framework der Untersuchung dargestellt, bevor in (3) mittels Künstlichen Neuronalen Netzen und Nearest-Neighbour-Verfahren eine Indexprognose des Deutschen Aktienindexes (DAX) durchgeführt wird. Da es sich bei den dargestellten Prognosen um Single-Chart-Analysen handelt, wird in (4) diskutiert, wie die Prognose aus (3) durch

Nutzung von Genetischen Algorithmen erweitert werden kann und wie etwa Entscheidungsbäume zur Entwicklung einer sinnvollen Anlagestrategie eingesetzt werden können. Der Beitrag schließt mit einem Fazit (5).

2 Lösungsansatz zur Indexprognose

Im Folgenden soll kurz das Framework der Untersuchung dargestellt werden. Empirisch gibt es Hinweise darauf, dass die Prognosegüte massiv davon abhängt, in welcher Phase (z.B. Bull- oder Bear-Markt) sich der zu prognostizierende Markt befindet [5]. Grundsätzlich liegt die Vermutung nahe, dass Finanzzeitreihen nicht-stationär und deshalb nicht sinnvoll durch ein Globalmodell zu beschreiben sind. Da es in dem Beitrag um die grundsätzliche Eignung von Data-Mining-Verfahren zur Prognose geht wurde zur Umgehung des Problems ein relativ kleiner Datensatz ausgewählt.

Die durch die Algorithmen zu lösende Lernaufgabe ist es, basierend auf den historischen Daten des Deutschen Aktienindexes vom 2004-08-06 bis 2006-07-20 (500 Handelstage), so genannte steigt/fällt-Prognosen durchzuführen.¹ Es soll also lediglich die Richtung des Kurses prognostiziert werden.

Zur Messung der Performancequalität soll die Trefferquote verwendet werden. Da Aktienkurse im Allgemeinen schwierig vorherzusagen sind werden in hohem Maße Fehlurteile bei der Prognose akzeptiert. Um die Problematik zu verdeutlichen sei betont, dass es bei steigt/fällt-Prognosen keinesfalls darum geht, eine möglichst 100%ige Genauigkeit zu erreichen, sondern möglichst weit oberhalb der Zufallsgrenze (Trefferquote von 50 %) zu liegen. Problematisch an der Kennzahl Trefferquote ist, dass große und kleine Gewinne/Verluste gleich gewichtet werden. Zur Umgehung dieses Problems gibt es in der Literatur verschiedene Vorschläge [10][5].

Neben der Performancequalität müssen die Ergebnisse auf Signifikanz getestet werden. Dies ist mittels eines χ^2 -Unabhängigkeitstests zu beurteilen. (Vgl. zur ausführlichen Beschreibung des Tests: [2]).

3 Indexprognose mit Data-Mining-Verfahren

3.1 Künstliche Neuronale Netze

Künstliche Neuronale Netze ahmen die Funktionsweise des menschlichen Gehirns nach und gelten als ein flexibles und robustes nichtlineares statistisches Verfahren. Die Aufgabe jedes Neurons besteht darin, die Informationen von seinem vorgelegerten Nachbarn zu empfangen, zu verarbeiten und das Ergebnis an seinen jeweils nächsten Nachbarn weiterzugeben. Die Menge der an einem neuronalen Netz beteiligten Neuronen lässt sich in disjunkte Teilmengen (Schichten) aufteilen. Zwischen Eingabeschicht und Ausgabeschicht liegen in der Regel verborgene Schichten von Zwischenneuronen. Sind Rückkopplungen zwischen den nicht Schichten zugelassen, spricht man von einem Feedforward-Netz, andernfalls handelt es sich um ein Feedback-Netz. Die Ausgestaltung des Künstlichen Neuronalen Netzes wird auch Topologie genannt.

¹ Datenquelle: <http://finance.yahoo.com>