

Decision Analytics mit Heatmap-Visualisierung von mehrschrittigen Ensembledaten

Eine Anwendung von Unsicherheitsmodellierung für Historical Consistent Neural Network und andere Prognosetechniken

Die heutige Leistungsfähigkeit von Rechnern führt leicht zur Generierung von großen Datenmengen. Die wirkliche Herausforderung liegt in einer adäquaten Aggregation der Daten für den Entscheidungsprozess. Vorliegend fokussieren wir auf Ensemble-Daten, die typischerweise erzeugt werden, wenn Prognose-Verteilungen für mehrere Zeitschritte in die Zukunft erzeugt werden. Häufig wird eine Verteilung durch Mittelwert oder Median aggregiert. Dies führt zu einem einzelnen Pfad, der leicht zu interpretieren ist. Dieser einzelne Pfad kann jedoch zu ernsthaften Fehlinterpretationen führen, wenn die zugrunde liegende Verteilung sich tatsächlich auf mehrere Pfade aufteilt. Mittelwert und Median können auch in einem Bereich liegen, der nur wenige Ensemble-Mitglieder enthält. Um diesem entgegenzuwirken, schlagen wir vor, Heatmap-Visualisierung zu verwenden, um die Daten für Decision Analytics besser aufzubereiten. Heatmap-Visualisierung ist eine intuitive Möglichkeit, Bereich mit hoher und niedriger Aktivität darzustellen. Die Bereiche werden farbkodiert in Abhängigkeit der (gewichteten) Anzahl von Ensemble-Mitgliedern in einem bestimmten Bereich.

DOI 10.1007/s11576-014-0417-3

Die Autoren

Dipl.-Math. Cornelius Köpp
 Prof. Dr. Hans-Jörg
 von Mettenheim (✉)
 Prof. Dr. Michael H. Breitner
 Institut für Wirtschaftsinformatik
 Leibniz Universität Hannover
 Königsworther Platz 1
 30167 Hannover
 Deutschland
mettenheim@iwi.uni-hannover.de

Eingegangen: 2013-07-15
 Angenommen: 2014-01-31
 Angenommen nach zwei Überarbeiten durch die Herausgeber des Schwerpunktthemas.
 Online publiziert: 2014-04-29

This article is also available in English via <http://www.springerlink.com> and <http://www.bise-journal.org>: Köpp C, von Mettenheim H-J, Breitner MH (2014) Decision Analytics with Heatmap Visualization for Multi-step Ensemble Data. An Application of Uncertainty Modeling to Historical Consistent Neural Network and Other Forecasts. Bus Inf Syst Eng. doi: [10.1007/s12599-014-0326-4](https://doi.org/10.1007/s12599-014-0326-4).

Zusätzliche Information ist in der Online-Version dieses Beitrags (doi: [10.1007/s11576-014-0417-3](https://doi.org/10.1007/s11576-014-0417-3)) enthalten.

© Springer Fachmedien Wiesbaden 2014

1 Einführung und Motivation

Eine typische Aufgabe für ein Entscheidungsunterstützungssystem ist die Erzeugung von Prognosen und – wichtiger noch – eine Unterstützung bei der In-

terpretation derselben. Punktprognosen können jedoch irreführend sein. Aus diesem Grunde können Ensemble-Prognosen – basierend auf einer Sammlung von Einzelprognosen – verwendet werden, um die Prognosegenauigkeit zu verbessern; siehe z. B. (Zhang und Berardi 2001). Häufig werden auch bei Ensemble-Prognosen nur einzelne Werte (oder einige wenige Werte der Verteilung) verwendet. Dies ist für die automatisierte Verwendung in Informationssystemen nützlich, aber die wirkliche Form der Verteilung bleibt unklar; siehe z. B. (Welch 2001; Hansen 2008). Die Prognose-Verteilung ist nicht notwendigerweise unimodal. Im Allgemeinen kann die Verteilung multimodal sein. Dies bedeutet, dass die Prognose nicht durch einen einzelnen Wert (z. B. den Mittelwert) charakterisiert werden kann. Viele verschiedene Modelle erzeugen die Prognose und es kann nicht erwartet werden, dass die Prognose eingipflig ist. In solchen Fällen können Mittelwert oder Median irreführend sein. Eine Mehrschrittprognose bedeutet, dass nicht nur ein Wert für den kommenden Zeitschritt prognostiziert wird. Vielmehr erstrecken

sich die hier betrachteten Prognosen über mehrere Zeitschritte. Insgesamt können gemeinsam verwendete Aggregationsmethode Informationen verlieren oder skalieren nicht gut. In beiden Fällen wird die Nützlichkeit der Prognose reduziert.

Wir präsentieren erste Schritte zur Beantwortung folgender Forschungsfrage: „Wie ermöglicht eine adäquate Visualisierung Decision Analytics für heutige Ensembleprognosemethoden?“

Ensembles Künstlicher Neuronaler Netze (KNN) sind ein typisches Beispiel für Prognosen die aus mehreren hundert einzelnen Pfaden bestehen. Im folgenden Beitrag werden 20-Tages-Prognosen des Gaspreises (in US-Dollar) betrachtet. Die Prognosen werden aus einem Ensemble von 200 Netzen erzeugt. Es wird eine neue Klasse von KNN verwendet: das Historical Consistent Neural Network (HCNN), welches in Zimmermann et al. (2010) eingeführt wurde. Siehe auch von Mettenheim und Breitner (2010) für eine detaillierte Präsentation und Analyse der Performance. Für die Ziele des vorliegenden Beitrags reicht es, sich darüber im Klaren zu sein, dass HCNN eine einfache Zustandsübergangsgleichung verwenden, um den nachfolgenden Zustand aus dem unmittelbar vorgehenden Zustand zu bestimmen. Mehrschrittprognosen sind hierdurch einfach zu generieren. In der Regel werden HCNN verwendet, wenn mehrere Zeitreihen einschließlich ihrer Verteilung auf einmal modelliert werden sollen. Werden mehrere HCNNs mit zufällig initialisierten Gewichten trainiert, führt dies zu einem diversifizierten Ensemble.

Die genaue Ausgestaltung der Prognose ist für die folgende Diskussion nicht von zentraler Bedeutung und es ist auch nicht Ziel des Beitrags die Performance dieser KNN-Prognose zu evaluieren. Vielmehr liegt der Fokus des Beitrags darin, eine Prognose für den menschlichen Entscheider leichter interpretierbar zu machen. Der vorliegende Ansatz ist damit nicht auf KNN begrenzt. Andere Modelltypen, die auf Grundlage historischer Daten eine unbegrenzte Anzahl an Prognosen erzeugen, könnten in diesem Kontext ebenfalls verwendet werden. Hierunter fallen z. B. Support Vector Maschinen, Evolutionäre Programmierung oder auch Monte-Carlo-Simulation. Aus diesem Grund wird das KNN Modell nur kurz angerissen. Im Beitrag präsentieren wir eine Heatmap-Visualisierung der resultierenden Ensemble-Prognosen. Dies stellt einen Schritt zu einer visuellen Un-

terstützung des menschlichen Entscheiders dar: Heatmaps aggregieren Informationen, erhalten aber die wesentlichen Prognose-Informationen, auch wenn die Verteilung multimodal ist.

Wir können nun unsere Forschungsfrage konkretisieren: „Wie können wir die vollständigen Prognoseinformationen unter Erhaltung aller Verteilungsinformationen intuitiv für Entscheider darstellen?“

Hierzu schlagen wir eine Heatmap-Visualisierung vor. Eine Heatmap erlaubt es durch Farb-Kodierung, zwischen aktiven und weniger aktiven Bereichen einer Prognose zu differenzieren (s. **Abb. 1(d)–1(f)**) und die Beispiele auf den **Abb. 3, 4, 5**. Eine detaillierte Diskussion von **Abb. 1** folgt weiter unten, insb. in Abschn. 6.

Das einfache Plotten der aggregierten Werte in **Abb. 1(a)** und **1(b)** bewirkt einen Informationsverlust. Es ist auch nicht sinnvoll, schlicht jede Prognose einzeln zu plotten wie dies in **Abb. 1(c)** geschieht. Einzelne Prognosen können nicht mehr unterschieden werden und das Ergebnis ist nutzlos. **Abbildungen 1(d)** und **1(e)** stellen unseren Heatmap-Ansatz vor und **Abb. 1(f)** ist eine sogar noch nützlichere Darstellung. Deutlich können wir sich aufspaltende Pfade in **Abb. 1(d)** und **1(f)** erkennen. Dies ist ein Warnhinweis darauf, dass die Prognose mit einer hohen Unsicherheit behaftet ist.

Heatmaps unterstützen uns auch darin, die Qualität einer Prognose zu beurteilen. In Abhängigkeit der Breite des Prognosekorridors und der Anzahl an Pfaden, die wir in der Heatmap erkennen, können wir die Prognose als eher zuverlässig oder eher unzuverlässig einstufen. Dies ist eine Alternative zum üblichen binären Output mancher Prognosemethoden. Es ist gängig, dass ein Prognosemodell entweder die Prognose „zunehmend“ oder „abnehmend“ abgibt. Dies ist jedoch nicht ganz ehrlich. Es sollte eine dritte Möglichkeit, nämlich „unsicher“, geben. Heatmaps ermöglichen genau diese Art der Ausgabe: anstelle die Information hinter einer einzelnen Zahl, die ohnehin nicht genau treffend sein wird, zu verschleiern, stellen Heatmaps das gesamte Prognosespektrum vor. Problematische Abschnitte lassen sich leicht identifizieren.

2 Forschungsdesign

Zur Beantwortung der Forschungsfrage orientieren wir uns am Design-Science-Research-Ansatz von Hevner et al. (2004,

S. 83). **Abbildung 2** zeigt das von uns umgesetzte Forschungsdesign. Im Folgenden stellen wir unseren Beitrag im Kontext ausgewählter Forschungsleitlinien dar:

Die vorgestellte Forschung ist relevant (*Problem Relevance*), da unser Visualisierungsansatz Verteilungsprognosen behandelt, die zum Beispiel im Kontext von KNN Ensembles auftreten. Dieselbe Visualisierung könnte nützlich sein für alle anderen Prognosemodell die Ensembles erzeugen, oder Zeitreihen. In Abschn. 3 überprüfen wir existierende Konzepte im Bereich der visuellen Darstellung von Ensembles und Zeitreihen und zeigen die Grenzen gängiger Visualisierungstechniken auf. In Abschn. 4 präsentieren wir das formale Modell und unsere Methode zur Generierung der Heatmaps. Wir argumentieren, dass es sich bei diesem Ansatz um ein Artefakt handelt das „extends the boundaries of human problem solving“ (Hevner et al. 2004) weil es Entscheidungsträgern helfen kann bessere Entscheidungen zu treffen, indem zusätzliche Informationen bereitgestellt werden, die beim Einsatz statischer Aggregation unsichtbar bleiben (*Design as an Artifact*).

Der erste Schritt unserer Evaluation besteht in der Konstruktion eines Prototyps. Dieser wird in Abschn. 5 beschrieben. Wir nutzten den Prototyp zur Demonstration der Nützlichkeit unseres Ansatzes, in Form eines qualitativen Informationsgewinns. Ein Beispielszenario ist durch die oben genannte Gas-Preis-Prognose gegeben, ergänzend dazu dient ein künstliches Szenario zur deutlicheren Darstellung des Informationsgewinns durch Interpolation. Damit folgen wir dem Konzept der Deskriptiven *Design Evaluation* mit Szenarios (Hevner et al. 2004, S. 86).

Ein erster Ansatz des vorgestellten Konzepts wurde bereits Wissenschaftlern und Praktikern auf einer Konferenz (von Mettenheim et al. 2012) und einem Workshop vorgestellt (*Communication of Research*). Wir haben die Ergebnisse mit KNN-Experten diskutiert und deren Feedback berücksichtigt. Auf Basis dieser Diskussion entstand eine unabhängige Implementierung unseres Ansatzes durch erfahrene Anwender in einem großen internationalen Unternehmen. Hevner et al. (2004) fordern, dass „the objective of design-science research is to develop technology-based solutions to important and relevant business problems“. In der Existenz einer unabhängigen Implementierung sehen wir einen