



# Predictive Analytics im Praxistest

Predictive Analytics ist seit Langem fester Bestandteil in Fachartikeln und Diskussionen rund um ein modernes Controlling. Dennoch stellen wir in den Seminaren und Beratungsprojekten der CA controller akademie immer wieder fest, dass viele Organisationen noch am Anfang dieser Entwicklung stehen. Dabei schwingt vor allem viel Unsicherheit mit, wie man sich dem Thema am besten nähert und was mögliche und sinnvolle Anwendungsgebiete (sogenannte Use Cases) im eigenen Unternehmen sein könnten.

## Predictive Analytics Community

Aus diesem Grund haben wir im vergangenen Jahr eine CA Predictive Analytics Community ins Leben gerufen, die in der Zwischenzeit mehr als 220 Teilnehmende aus unterschiedlichsten Branchen und Unternehmensgrößen umfasst. Wir bieten eine kostenfreie Plattform, auf der die Community-Mitglieder von Best Practices, Erfah-

rungen und dem Wissen aller lernen können und miteinander ins Gespräch kommen. Einige der Mitglieder haben bereits umfangreiche Erfahrungen mit der Einführung und Nutzung von Predictive Analytics gesammelt, während andere noch am Anfang stehen und zunächst an der Vorgehensweise und tiefergehendem Wissen interessiert sind.

Im vergangenen Jahr haben bereits 4 Online-Events stattgefunden, in denen die CA controller akademie ebenso wie die Community-Mitglieder spannende Vorträge gehalten und wertvolle Diskussionen miteinander geführt haben. In diesem Jahr sind 3 weitere Online-Termine und eine Präsenzveranstaltung in Hamburg geplant.

## Praxisbeispiel aus der Lebensmittel-industrie

Eines der praktischen Anwendungsbeispiele aus der Community stellt an dieser Stelle Simon Rossi vor, bei dem es um die Einführung von Predictive Analytics im Produktions-Controlling der Mila Bergmilch Südtirol geht. Mit dem Aufbau eines Vorhersagemodells für die Nachfrage nach Joghurt ist es ihm gelungen, die Produktionsplanung hinsichtlich der notwendigen Rohstoffe als auch der benötigten Fertigungskapazitäten zu optimieren und somit Engpässe zu vermeiden. Der automatisierte Forecast auf Basis eines ARIMA-Modells liefert in diesem Fall eine Forecast-Qualität mit einer Abweichung von nur +/- 2%.



**Simon Rossi**

ist Produktionscontroller bei der Mila Bergmilch Südtirol.

Er untersucht Daten der Fertigung und wertet diese anhand wissenschaftlicher Methoden für die Unternehmenssteuerung aus.

## Decomposition

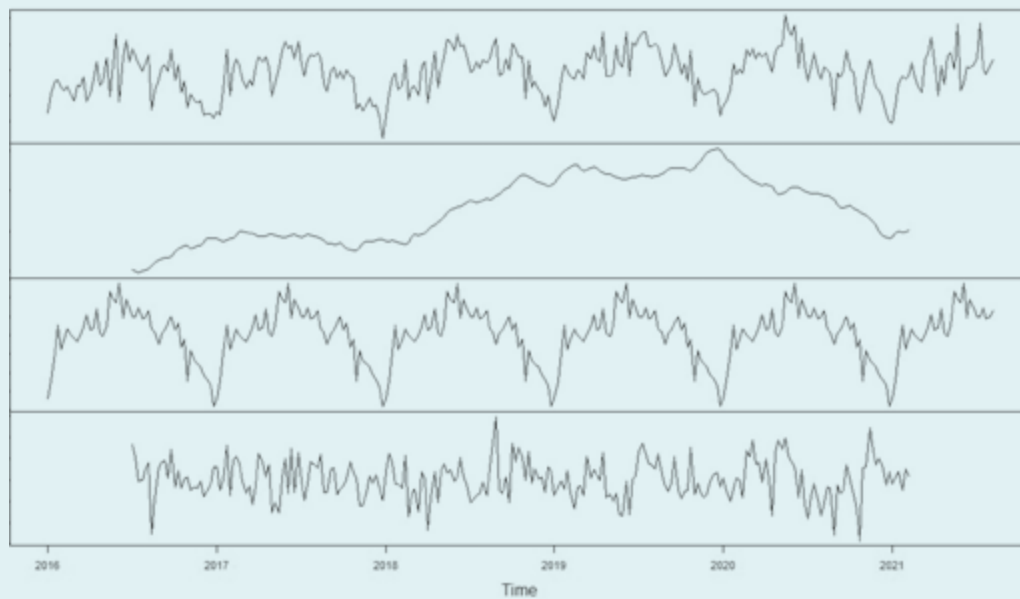


Abb. 1: Grafische Darstellung der Zerlegung der Zeitreihe in ihre Bestandteile

## Forecast ARIMA – Modell

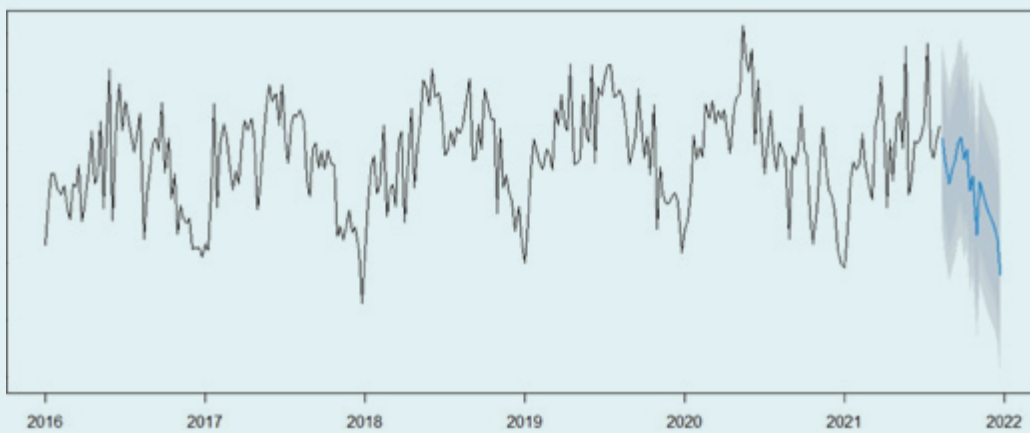


Abb. 2: Ursprüngliche Zeitreihe (grau) mit maschinellem Forecast (blau)

## Zeitreihenanalyse mit der Statistiksoftware R

In fast jedem Unternehmen kann man eine der schönsten Erscheinungsformen der Datenwelt finden: die Zeitreihe. Unter einer Zeitreihe kann man sich Datenpunkte oder Ereignisse vorstellen, die in regelmäßigen Abständen über einen bestimmten Zeitraum auftreten. Beispiele dafür sind Börsenkurse, Umsätze, Wetterbeobachtungen oder auch Fertigungsmengen.

Zeitreihen liefern Informationen über Trends, Saisonalitäten und Abhängigkeiten. Darüber hinaus besteht die Möglichkeit, auf Basis von Zeitreihenanalysen ein Prognosemodell zu erstellen. Anhand des Praxisbeispiels aus der Joghurtproduktion lässt sich die Zeitreihenanalyse anschaulich darstellen. Als Basis für die Analyse wurden die wöchentlich produzierten Joghurtmengen über die

letzten 5 Jahre herangezogen. Mit dem daraus erstellten Modell soll die notwendige Fertigungsmenge bis Ende des laufenden Jahres vorhergesagt werden. Die Datenaufbereitung und die finale Modellierung erfolgten mit der Statistiksoftware R.

Wenn es um die Beschreibung und Analyse von Zeitreihen geht kommt man an sogenannten ARIMA-Modellen nicht vorbei. Die Abkürzung steht für Auto-Regressive Integrated Moving Average. Dabei handelt es sich um eine Fortführung des ARMA-Modells. Das ARIMA-Modell wird zusätzlich zum autoregressiven Teil und dem gleitenden Mittelwertbeitrag um die Differenzierung und Integration zur Trendbeseitigung und Herstellung der Stationarität erweitert. Eine Zeitreihe ist stationär, wenn sie zu jeder Zeit den gleichen Mittelwert und die gleiche Varianz aufweist. Saisonale Effekte und Trends sind demnach nicht zu erkennen. Leider ist solch eine Begebenheit

in der Realität eher selten, denn ständig stoßen wir auf Beobachtungen, die klaren Trends und Saisonalitäten folgen. Der erste Schritt beinhaltet die Zerlegung der Zeitreihe, um deren Eigenschaften zu analysieren. Mittels einer speziellen Methode (Decomposition) können wir diesen Schritt grafisch abbilden, siehe **Abb. 1**.

Die Zeitreihe wird zunächst in ihre Bestandteile zerlegt. Die Grafik zeigt dabei von oben nach unten:

- 1) Originale Datenpunkte
- 2) Trend
- 3) Saisonalität
- 4) White Noise Punkte (Datenpunkte, die weder einem Trend noch einer Saisonalität zugeordnet werden können)

Wie zuvor angedeutet, ist die Voraussetzung für die Modellierung von Zeitreihen deren Stationarität; Im vorliegenden Beispiel ist dies leider nicht der Fall. Genau aus diesem Grund sind ARIMA-Modelle die passende Lösung, denn sie ermöglichen die Modellierung unregelmäßiger Zeitreihen. Trend und Nicht-Stationarität werden vom Algorithmus anhand statistischer Methoden (Transformation) sofort erkannt und im Modell mitberücksichtigt, wobei Mittelwert und Varianz über den Zeitraum stabilisiert werden. Dies erfolgt entweder durch Logarithmieren oder Bildung der First-Order-Difference, d.h. man bildet die Differenz von Periode zu Periode (hier Woche zu Woche).

**„Im präsentierten Beispiel betrug die durchschnittliche Abweichung ca. +/- 2%. Dies stellt ein sehr gutes Ergebnis dar, mit dem man exzellent arbeiten kann.“**

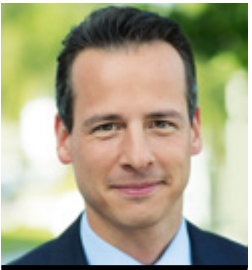
Mit Hilfe eines Correlograms kann man visuell herauszufinden, ob ein Auto Regressive (AR) oder ein Moving Average (MA)-Prozess vorliegt. Stellt man beispielsweise eine Autokorrelation über mehrere vergangene Datenpunkte fest (sog. „LAGS“), dann sprechen wir von einem autoregressiven Prozess, der Abhängigkeiten über die gesamte Zeitreihe berücksichtigen sollte. Sind die Abhängigkeiten jedoch nur für eine kurze Dauer zu beobachten, handelt es sich eher um ein MA-Prozess.

Nach erfolgter Analyse der Zeitreiheneigenschaften können wir das finale Modell laufen lassen, um Schätzungen und Validierungen vornehmen zu können, siehe **Abb. 2**. Zu sehen ist die ursprüngliche Zeitreihe mit einem maschinellen Forecast in blau. Die grauen Bereiche definieren die Konfidenzintervalle. Sie stellen einen statistisch berechneten Bereich dar, mit dem man besser einschätzen kann, wo der wahre Mittelwert des vorhergesagten Datenpunktes liegt.

Der Forecast lässt sofort erkennen, wie das Modell den Trend und die Saisonalität mitberücksichtigt hat. Wichtig ist jetzt, noch das Modell zu validieren. Dies erfolgt durch die genauere Analyse der sogenannten Residuen oder Abweichungen. Die wichtigsten Punkte sind dabei dessen Stationarität und Normalverteilung. Hier gibt es im Statistikprogramm R eine entsprechende Funktion, die die Analyse grafisch aufbereitet.

Das Vorhersagemodell ist nun aufgebaut. Zu guter Letzt ist die Qualität des Modells zu testen, indem die vorhergesagten Werte mit den tatsächlich eingetretenen Werten verglichen werden (sogenanntes Backtesting). Im präsentierten Beispiel betrug die durchschnittliche Abweichung ca. +/- 2%. Dies stellt ein sehr gutes Ergebnis dar, mit dem man exzellent arbeiten kann.

Abschließend noch zwei Hinweise: Je mehr Datenpunkte zur Verfügung stehen, desto besser kann der Algorithmus arbeiten. Als Faustregel ist von mindestens 50 Datenpunkten auszugehen. Im Modell kann der Parameter des vorhergesagten Zeitraums manuell definiert werden (in Beispiel sind es 20 Wochen). Je höher dieser Wert ist, desto ungenauer wird die Vorhersage. Man kann das mit einer Wettervorhersage vergleichen. Wie das Wetter in den nächsten 3 Tagen aussehen wird ist recht gut prognostizierbar. Eine Vorhersage der nächsten 2 Wochen hingegen ist ziemlich schwierig. ■



Danny Szajnowicz

ist Partner der CA Akademie AG und verantwortet den Bereich CA Consulting. Er verfügt über umfangreiche Erfahrung in der Durchführung und Leitung von Projekten in den Bereichen Strategieentwicklung, Unternehmenssteuerung/Controlling, Organisationsentwicklung und Prozessoptimierung.

## Predictive Analytics Community

### Nächste Termine

- 13.05.2022: Vorgehen und Erfahrungen bei der Einführung von Predictive Analytics bei bofrost
- 16.09.2022: Excelbasierte Prognose von Marktparametern
- 18.11.2022: Risikomanagement mit „R“

[www.controllerakademie.de/predictive-analytics/](http://www.controllerakademie.de/predictive-analytics/)

