

Time Series Analysis and Forecasting

By Bastian Berle & Ron Holzapfel

Idee

Bei der Zeitreihenanalyse werden Vorhersagen beruhend auf der zeitlichen Reihenfolge von historischen Daten getroffen. Dabei wird versucht, den Wert einer Variablen zu einem bestimmten Zeitpunkt in der Zukunft abzubilden. Damit ist diese eine Spezialform zur Regressionsanalyse.

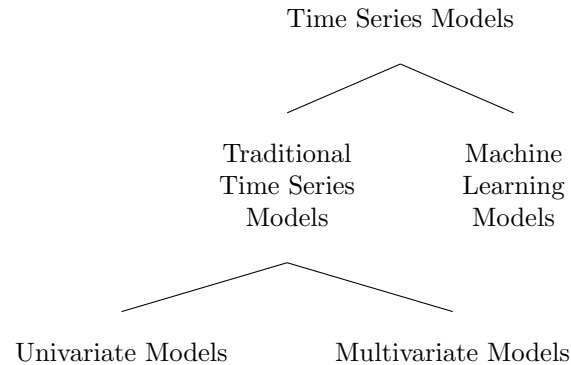
Additives Komponentenmodell

Eine Zeitreihe kann in mehrere Komponenten aufgeteilt werden. Diese sind

- Trendkomponente (g_t),
- Saisonkomponente (s_t), und
- Restkomponente (z_t).

Die Trendkomponente gibt die langfristige Veränderung des mittleren Niveaus der Zeitreihe an. Die Saisonkomponente bildet hingegen zyklische Schwankungen ab, die relativ unverändert in regelmäßigen Abständen wiederholt auftreten. Die Restkomponente spiegelt den Anteil der Zeitreihe wider, der nicht durch die Trend- oder Saisonkomponente abgebildet wird. Zusammen ergibt sich $x_t = g_t + s_t + z_t$. $x_t - g_t$ ist dabei die trendbereinigte und $x_t - s_t$ die saisonbereinigte Zeitreihe.

Traditional vs. ML Models



Univariate Models

Beispiel Output: Anzahl der verkauften Produkte für morgen

- ARIMA
- SARIMAX
- Prophet
- Neural Prophet

Multivariate Models

Beispiel Output: Anzahl der verkauften Produkte für morgen & Anzahl der Rücksendung

- Vector Autoregression

Machine Learning

- RNNs
- LSTM (Long Short-Term Memory)

Traditional Models	Machine Learning
Recursive	Direct
Easy to Extend	Tough to Exent
Can't add time varying features	Can add time varying features

Rekursiv bedeutet, dass die Vorhersagen basierend auf den vorherigen Daten getroffen werden. Es wird also für eine Vorhersage über den nächsten Tag die Daten des heutigen benötigt. Dafür können auch die bereits vorhergesagten Daten genutzt werden. Bei der *direkten* Methode können Vorhersagen über Zeitpunkte ohne vorherigen getroffen werden.

Unterschied zur Regression

Bei Time Series Analysis and Forecasting handelt es sich um eine Unterform der Regressionen. Dennoch kann man zwischen der traditionellen Regression und der zeitlichen Vorhersage unterscheiden.

Bei der klassischen Regression werden die Zielwerte basierend auf vorliegenden Eigenschaften getroffen. Dabei spielt der Zeitpunkt eine eher untergeordnete Rolle. Möchte man den Verkaufspreis von Häusern, die 2020 gebaut wurden, vorhersagen, ist der exakte Bauzeitpunkt weniger entscheidend, als Attribute, wie Wohnfläche, Anzahl der Bäder etc.

Die Analyse/Vorhersage von Zeitreihen werden dann eingesetzt, wenn man einen bestimmten Zielwert für einen bestimmten Zeitpunkt berechnen will. Möchte man wissen, wie viele Produkte an einem bestimmten Zeitpunkt verkauft werden sollen, sind die Produkteigenschaften weniger relevant als z.B. die Verkäufe von vor zwei und drei Tagen. Insgesamt ist dadurch die Auswahl passender Attribute bei der Vorhersage zeitlicher Reihenfolgen schwieriger.