

Abstract (english)

Media Mix Models are used in marketing to measure the effectiveness of advertising and support marketers in making future budget allocation decisions. Key results in Media Mix Modeling (MMM) are metrics for marketing attribution such as Return on Advertising Spend (ROAS). The attribution metrics measure how advertising investments contribute to business-relevant Key Performance Indicators (KPIs) such as sales or brand awareness. Modeling non-linear advertising effects, such as lagged effects or saturation effects, is a challenge in Media Mix Modeling. The aim of this thesis is a comparative analysis of parametric, semi- and non-parametric regression methods and their evaluation regarding their applicability in Media Mix Modeling. The assessment of the examined methods is based on their predictive accuracy and their uncertainty in determining attribution metrics. The regression methods are applied on both simulated data and a real-world data set provided by the marketing department of R+V Allgemeine Versicherung. On simulated data, Bayesian methods show the best predictive accuracy and the lowest uncertainty in the measurement of attribution metrics. Informative prior distributions can reduce uncertainty in the model, but lead to biased estimates of attribution metrics, especially on small data sets. Semi-parametric and non-parametric regression methods show biased attribution metrics on the simulated data. On the real-world data set, a regularized parametric regression model with a MAPE of 7.07% achieves the best prediction accuracy.

Abstract (deutsch)

Media Mix Modelle werden im Marketing für die Messung der Wirksamkeit von Werbemaßnahmen verwendet und unterstützen Marketers bei Entscheidungen in der Allokation zukünftiger Mediabudgets. Zentrale Ergebnisse des Media Mix Modeling (MMM) sind Metriken zur Marketing-Attribution wie der Return on Advertising Spend (ROAS). Die Attributionsmetriken messen, wie Werbeinvestitionen zu unternehmensrelevanten Key Performance Indicators (KPIs) wie dem Umsatz oder der Markenbekanntheit beitragen. Die Modellierung nicht-linearer Werbeeffekte, wie zeitlich verzögerte Effekte oder Sättigungseffekte, stellt im Media Mix Modeling eine Herausforderung dar. Ziel dieser Arbeit ist eine Vergleichsanalyse von parametrischen, semi- und nicht parametrischen Regressionsmethoden und deren Bewertung hinsichtlich ihrer Anwendbarkeit im Media Mix Modeling. Die Evaluation der verwendeten Methoden erfolgt anhand ihrer Vorhersagegenauigkeit und ihrer Unsicherheit in der Messung von Attributionsmetriken. Die Regressionsmethoden werden sowohl auf simulierte Daten als auch auf einen echten Datensatz, bereitgestellt von der Marketingabteilung der R+V Allgemeine Versicherung, angewendet. Bayesianische Methoden zeigen auf simulierten Datensätzen die beste Vorhersagegenauigkeit und die geringste Unsicherheit bei der Bestimmung von Attributionsmetriken. Informative Priorverteilungen können die Unsicherheit im Modell verringern, aber insbesondere auf kleinen Datensätzen zu verzerrten Schätzern der Attributionsmetriken führen. Semi- und nicht parametrische Regressionsmethoden zeigen auf simulierten Daten verzerrte Attributionsmetriken. Auf dem echten Datensatz erreicht ein regularisiertes parametrisches Regressionsmodell mit einem MAPE von 7.07% die beste Vorhersagegenauigkeit.