

Abstract

Ein sparsamer, zielgerichteter Einsatz von Werbung ist kosteneffizient und stärkt das Vertrauen der Kundinnen und Kunden ins Unternehmen, weil sie im Idealfall nur Werbung erhalten, die für sie von Interesse ist. Dazu kommen statistische Modelle, wie z. B. die logistische Regression, aber auch immer stärker Machine Learning-Verfahren, wie z. B. Random Forest oder auch künstliche neuronale Netzen zum Einsatz, um die ideale Werbemaßnahme für eine Kundin zu ermitteln. Dabei stellt sich im Speziellen im Online-Marketing nicht nur die Frage, ob man einem Kunden ein elektronisches Werbe-Signal, wie z.B. einen Werbebanner, präsentieren sollte, sondern auch wie häufig.

Die Arbeit beschäftigt sich mit der Frage, ob und wie man anhand von Kunden-Merkmalen unter Einsatz von Machine Learning-Verfahren vorhersagen kann, wie häufig eine Werbung angezeigt werden soll, um daraus Frequency-Capping-Regeln zu erstellen, die für jeden Kunden definieren, nach wie vielen Signalen eine Kappung erfolgt, d. h. kein weiteres Signal präsentiert wird. Dabei wird der Data Mining-Prozess anhand eines konkreten Anwendungsbeispiels beschrieben. Besonderes Augenmerk liegt dabei auf der Beschreibung eines selbst entwickelten Codes zur Erstellung geeigneter Datensätze mit SAS Base und des Modellierungsprozesses mit dem SAS Enterprise Miner, einem speziellen Data Mining-Tool von SAS. Letzteres geschieht anhand eines ausgewählten, mit dem zuvor beschriebenen Code erstellten Beispiel-Datensatzes. Die Arbeit stellt sowohl eine Machbarkeitsstudie dar und dient gleichzeitig als Einführung in Machine Learning an einem konkreten Beispiel. Der erstellte Code kann mit relativ wenigen Anpassungen in einen produktiven Einsatz überführt werden.

Abstract

An economical, targeted use of advertising is cost-efficient and strengthens customers' trust in the company, because ideally they only receive advertising that is of interest to them. Statistical models such as logistic regression, but also increasingly machine learning methods such as random forest or artificial neural networks are used to determine the ideal advertising for a customer. In online marketing in particular, the question arises not only whether a customer should present an electronic advertising signal, such as an advertising banner, but also how often.

This master thesis deals with the question whether and how one can predict how often an advertisement should be displayed based on customer characteristics using machine learning in order to create frequency capping rules, which define for each customer after how many signals a capping takes place, i.e. no further signal is presented. The data mining process is described using a concrete application example. Particular attention is paid to the description of a code for creating suitable data sets with SAS Base and the modeling process with the SAS Enterprise Miner, a special data mining tool from SAS. The latter is done on the basis of a selected sample data set created by the code described. The work is both a feasibility study and an introduction to machine learning using a concrete example. The generated code can be converted into productive use with relatively few adjustments.