

ABSTRACT

At company X, most prices are changed every year. During the year, prices are also adjusted for better alignments. Before a new price is approved, the impact it could have on sales will be calculated based on a simple calculation method, which assumes that demand will remain constant after the change. Therefore, this thesis attempts to examine the effects of list price changes on demand in order to provide guidance for pricing partners at company X on price changes.

For this purpose, two main approaches are employed: a parametric method (linear regressions) and a non-parametric method (regression trees and random forests). Different data visualizations such as histograms, box plots, correlation matrices and scatter plots suggest that apart from list prices, there are other driving factors which affect demand, including: price lists (country level), discount ranges, customer groups and time. Therefore, the models are built on these controlling variables. Before modeling, the data is cleaned, pre-aggregated, filtered on the chosen Stock Keeping Unit (SKU) and partitioned to training set and test set. Different model specifications suggest that the SKU of interest has an inelastic demand. Two linear regression specifications and a random forest are replicated for 1000 SKUs. However, the models sometimes provide different insights into the sensitiveness of demand to prices. Each model has its own advantages. One must not choose only one model to work with. A guidance tree is created as an example how the models could be utilized in parallel, where pricing partners will receive either a green light or a warning about a possible decrease in demand for his price change proposals. An alternative would also be averaging the estimates of the three models. The estimates of price elasticity on SKU level could then be used to predict the impact of price changes on sales.

The study has several limitations, including a lack of potential variables, model replication for multiple SKUs based on whole samples, impossibility to replicate for the whole portfolio due to an out of memory error. In the future, with more data coming in, the models can incorporate more driving factors to achieve a better accuracy.

ZUSAMMENFASSUNG

Bei der Firma X werden die meisten Preise jedes Jahr geändert. Im Laufe des Jahres werden die Preise auch angepasst, um eine bessere Angleichung zu erreichen. Bevor ein neuer Preis genehmigt wird, werden die Auswirkungen, die er auf den Umsatz haben könnte, auf der Grundlage einer einfachen Berechnungsmethode berechnet, die davon ausgeht, dass die Nachfrage nach der Änderung konstant bleibt. In dieser Arbeit wird daher versucht, die Auswirkungen von Listenpreisänderungen auf die Nachfrage zu untersuchen, um den Pricing Partners bei der Firma X eine Orientierungshilfe für Preisänderungen zu geben.

Zu diesem Zweck werden im Wesentlichen zwei Ansätze verwendet: eine parametrische Methode (lineare Regressionen) und eine nicht-parametrische Methode (Regressionsbäume und Random Forests). Verschiedene Datenvisualisierungen zeigen, dass es neben den Listenpreisen noch andere treibende Faktoren gibt, die die Nachfrage beeinflussen, darunter: Preislisten (auf Länderebene), Rabattbereiche, Kundengruppen und Zeit. Daher bauen die Modelle auf diesen Kontrollvariablen auf. Vor der Modellierung werden die Daten bereinigt, voraggregiert, nach der gewählten Stock Keeping Unit (SKU) gefiltert und in Trainings- und Testsätze aufgeteilt. Verschiedene Modellspezifikationen deuten darauf hin, dass die interessierende SKU eine unelastische Nachfrage hat. Zwei lineare Regressionen und ein Random Forest werden für 1000 SKUs repliziert. Die Modelle liefern jedoch manchmal unterschiedliche Einblicke in die Empfindlichkeit der Nachfrage gegenüber Preisen. Jedes Modell hat seine eigenen Vorteile. Als Beispiel dafür, wie die Modelle parallel eingesetzt werden könnten, wird eine Orientierungshilfe erstellt, in dem die Pricing Partners entweder grünes Licht oder eine Warnung vor einem möglichen Nachfragerückgang für ihre Preisänderungsvorschläge erhalten. Eine Alternative wäre auch die Mittelwertbildung der Schätzungen der drei Modelle. Die Schätzungen der Preiselastizität auf SKU-Ebene könnten dann verwendet werden, um die Auswirkungen von Preisänderungen auf den Verkauf vorherzusagen.

Diese Arbeit hat mehrere Einschränkungen, darunter das Fehlen potenzieller Variablen, die Modellreplikation für mehrere SKUs auf der Grundlage ganzer Stichproben und die Unmöglichkeit der Replikation für das gesamte Portfolio aufgrund eines Out-of-Memory-Fehlers. In Zukunft, wenn mehr Daten eingehen, können die Modelle mehr treibende Faktoren einbeziehen, um eine bessere Genauigkeit zu erreichen.