

Recommendation Engines in Customer Relationship Management Systems

POPULÄRVETENSKAPLIG SAMMANFATTNING **Felix Glas**

Results show that software sales tools can predict potential new customers using well-known machine-learning algorithms such as: k -nearest neighbors, C4.5 decision tree induction, and k -means clustering, provided that there is enough data available.

Today, there is a demand for automated procedures for predicting future customers using recommendation engines in the customer relationship management (CRM) market. Imagine you are in sales and you are getting suggestions on possible new customers in the same way that Netflix recommends movies and Spotify helps you find new songs and genres. Machine-learning techniques are behind these high quality recommendations that derive deep insights from large amounts of past customer data. Such techniques can benefit most users of CRM systems, provided that they have gathered enough data to train models so that they can make accurate predictions on their customers. Our results show that relevant prospects can be extracted using off-the-shelf algorithms such as decision trees and k -means clustering.

The purpose of CRM is to help companies better manage their customers, present and future. It is an overall strategy for finding promising prospects and keeping existing customers by collecting and refining information about individual companies and learning their behaviors and needs. In today's expanding CRM market, there is a demand for automated procedures that can be used for customer prospecting. There are already functions commonly available for finding "twins", *i.e.*, possible customers that are similar to existing customers, and for browsing through lists of customers partitioned into categories such as locations or lines of business. In the near future, computer algorithms will enable systems to automatically suggest prospects that have a high potential

for becoming profitable customers. Such systems need a recommendation engine for making predictions on which companies are relevant.

Current recommendation engines are typically built using machine-learning techniques and the results provided by this project show that relevant customer predictions can be made using common algorithms. During the project, an array of machine-learning algorithms were evaluated using measures of accuracy and relevancy of the data predicted. Using a decision tree classifier based on the C4.5 algorithm we showed that a manageable number of similar customers can be predicted, provided there is a sufficiently large number of existing customers to train on. However, observations show that there is a trade-off effect between getting high precision of relevancy and getting a large number of predictions. This trade-off can nevertheless be controlled by adjusting the complexity of the model. By decreasing the number of existing customers in relation to the proportion of non-customers in the training set, the model will obtain a higher bias, thus shifting the predictions toward being greater in number and of having less relevance. This can sometimes be desirable when targeting new customers more aggressively. Also, by performing k -means cluster analysis on the set of existing customers, homogeneous groups of customers can be identified. By storing the relationships between commonly occurring clusters of existing customers, new companies can be given recommendations based on these relationships.