MISQ Archivist

Mining Massive Fine-Grained Behavior Data to Improve Predictive Analytics

David Martens, Foster Provost, Jessica Clark, and Enric Junqué de Fortuny

Abstract

Organizations increasingly have access to massive, fine-grained data on consumer behavior. Despite the hype over big data, and the success of predictive analytics, only a few organizations have incorporated such fine-grained data in a non-aggregated manner into their predictive analytics. This paper examines the use of massive, fine-grained data on consumer behavior—specifically payments to a very large set of particular merchants—to improve predictive models for targeted marketing. The paper details how using this different sort of data can substantially improve predictive performance, even in an application for which predictive analytics has been applied for years. One of the most striking results has important implications for managers considering the value of big data. Using a real-life dataset of 21 million transactions by 1.2 million customers, as well as 289 other variables describing these customers, the results show that there is no appreciable improvement from moving to big data when using traditional structured data. However, in contrast, when using fine-grained behavior data, there continues to be substantial value to increasing the data size across the entire range of the analyses. This suggests that larger firms may have substantially more valuable data assets than smaller firms when using their transaction data for targeted marketing.

Keywords: Behavioral similarity, big data, response modeling, banking, payment data, customer analytics