The Pricing of Individual and Aggregate Consumer Information in Today’s Marketplace

Tania Mohd Nor • Kaira Gui
Faculty Advisor: Jonathan M. Smith

Abstract

Determine the price of consumer information by analyzing the relationship between the revenue generated by a firm and the amount of consumer data collected.

Methodology

We assumed that by analyzing churn in the telecommunications industry, we can observe a general trend across industries.

The primary data mining software used for analyzing customer data was TreeNet. We ran TreeNet 29 times, using a range of sample sizes in order to determine the model correctness trend resulting from an increase in sample size. We then used Verizon’s statistics to estimate from our trend line the marginal revenue generated by an additional unit of customer information.

We also observed the importance of several predictor variables by examining the relevant output of TreeNet.

Results

Effect of the Sample Size of Customer Information on the Predictive Success of the Model

![Chart showing the effect of sample size on model correctness.](image)

This graph shows that subsequent increases in sample size after 300 yields no significant improvement in model correctness.

Interpretation of Results

- Thus, the additional value of a single consumer’s information to the firm is approximately $1.86 to $6.19 million if the firm has collected less than 300 separate data points. However, the additional value of a consumer’s information over 300 is negligible to none.

Implications

The results from the churn dataset have more generalized implications for the divide between a consumer and firm’s valuation of privacy. Whereas a firm could place a price within the range of $1.86 to $6.19 million on a single consumer’s profile, this consumer may only need a $30.49 to $44.62 incentive to give it up.

Firms who possess vast databases of consumer information such as Google and other online retail sites, have the incentive to compile profiles of consumer characteristics to be used by their own firms, or to be farmed out to business partners, telemarketers, or direct-mail solicitors. By aggregating a large and broad database, Google and similar sites may have access to enough data to be relevant to as wide a variety of consumer information users as possible.