

R tools and tricks for marketing inference in a big internet company

Paweł Ładyżyński^{1*}

1. Naspers Classifieds

*Contact author: pawelladyz@wp.pl

Keywords: marketing inference; clustering; big data; predictive modeling; feature engineering; feature selection

R packages: data.table; dplyr; ggplot; h2o; rmarkdown; shiny

Data is becoming an integral part of digital marketing, as businesses realize the power of information to create successful campaigns and see real-time results. In 2016, big data continues its growth as an important part of supporting business decisions. Armed with information on customer behaviors and purchases, we are now able to build profiles or user personas which ensure each marketing effort is geared toward a specific type of customer. However, the crucial part of business reasoning is to examine large data sets to uncover hidden patterns, unknown correlations, market trends, customer preferences and other useful business information. This step is usually connected with looking for new libraries or packages which performance is good enough to deal with large amounts of data. Fortunately, there are many **R** packages, like *data.table* or *h2o*, which enables **R** to be involved in the big data reasoning process.