Review on Customer Segmentation Technique on Ecommerce

1. Simple: use database query or statistic data
2. Business Rule (Baer 2012)

Link: <http://support.sas.com/resources/papers/proceedings12/103-2012.pdf>

* Group by predetermined class
* Data: Demographic, Purchase history
* Advantage: Easy to apply, use database query
* Disadvantage: Not focus on customer behavior

1. Customer profiling (Colica 2011)

* Using a query on the customer database or using the clustering algorithm when the data is huge
* Data: Demographic, Purchase history
* Advantage: Use database query if data is small
* Disadvantage: Not focus on customer behavior

1. Magento (2014)

* Perform an purchase history’s analysis to get best customer, unprofitable customer and potential customer profit
* Data: Demographic, Purchase history, Data Product, Data media, Data Marketing, Server log
* Advantage: Have a clear variable customer segmentation
* Disadvantage: there is no data processing for each variable

1. RFM: use RFM data
2. Quantile membership (Baer 2012)

* Divide RFM data into 5 groups from A to F 🡪 map 2 components table 🡪 result: good F (A/B), good M (A/B) but poor R (D/E)
* Data: Purchase history
* Advantage: Can process small data, can be used with other data
* Disadvantage: Good result obtained when determining a good classification

1. RFM cell classification grouping (Colica 2011)

* Use three dimensions to classify each customer in one cell after labeling each level of RFM (similar to Baer’s quantile membership)
* Data: Purchase history
* Advantage: Efficient 3D mapping
* Disadvantage: Same as quantile membership

1. Target: must have target to segment
2. Supervised clustering with decision tree (Baer 2012)

* Using decision tree with the target on their nodes
* Data: Demographic, Purchase history
* Advantage: Classify customers according to target
* Disadvantage: use 1 variable to cluster, show only 1 aspect behavior

1. Customer likeness clustering (Colica 2011)

* The same with Supervised clustering with decision tree
* Data: Demographic, Purchase history, Data Product
* Advantage: Same as supervised clustering with decision tree
* Disadvantage: Problem arises when there are different unit in record

1. Purchase affinity clustering (Colica 2011)

* Scoring on interesting in certain products then clusters customer database based on that score to get a similar group.
* Data: Purchase history, Data Product
* Advantage: Know the products most in demand
* Disadvantage: Specific to product segmentation

1. Unsupervised clustering ( Colica 2011)

* Measure similarity using Euclidean distance 🡪 cluster using Kmeans
* Data: Purchase history
* Advantage: Use any number of customer attributes
* Disadvantage: Speed of computation depends on k values