Abstract

Knowledge of customer behaviour helps organizations to continuously re-evaluate their strategies with the consumers and plan to improve and expand their application of the most effective strategies. Using expenditure data collected through daily mobile conversations with consumers in Kenya, this study sought to compare various clustering algorithms and establish one that best segments consumers, and subsequently providing profiles that provide a basis for marketing and brand strategy based on existing demographic data – age, gender, region and primary income source. K-Means, Hierarchical and Partitioning around Medoids (PAM) clustering algorithms were compared using internal and stability validation tests. Hierarchical clustering with four clusters had the best Connectivity (0.847) and Silhouette width (0.924) measures. Stability validation compares the results by removing a column, one at a time. Average Proportion of Non-overlap (APN), Average Distance (AD), Average Distance Between Means (AND) and Figure of Merit (FOM) were used to compare the algorithms. Again, Hierarchical clustering with four clusters was found to partition the data best. The study forms a
basis for the use of additional profile descriptors once available to provide a firmer understanding of the customer segments built on expenditure data in Kenya.

References

Index Terms

Computer Science  Data Mining

Keywords

Customer segmentation, clustering, clustering algorithms.