Abstract No: SO-15   **Software Intensive Systems**

Applicability of unsupervised learning algorithms for setting profiles for consumer buying behaviour

Iromi R Paranavithana and Thashika D Rupasinghe[*]

Department of Industrial Management, Faculty of Science,
University of Kelaniya, Sri Lanka
thashika@kln.ac.lk

The Consumer Buying Behaviour consists of a summation of attitudes, preferences, intentions and, decisions taken by them. The process that customer buys a product or service varies for each individual and each category of products they may purchase. With the development of Information Technology, the products and the behaviour of purchasing those products have drastically changed and become more unique to individuals. With respect to these changes, the data collection and analysis have become more dynamic and customer data has become larger and nosier in terms of volume and complexity. As a result of that, handling, analysing, and interpreting customer Point of Sale (POS) data has become a challenge for Retail Supply Chains (RSC) who wish to segregate customers into specific niche markets. Furthermore, it makes increasingly difficult for the retailer to find out when a person comes and buys the products from their outlets and to predict his/her behaviour for the subsequent purchases. As a solution for the aforementioned problems faced by the retailers, a novel a consumer buying behaviour profile mechanism is proposed. The profiles are created with respect to the frequency, time-stamp, and product category using a large POS dataset. The Unsupervised learning techniques were utilized in categorizing consumers in determining similar purchasing behaviour using *K-means, Expectation Maximization, and Hierarchical Agglomerative Clustering (HAC)*. Along with the above clustering techniques, text mining techniques were used in categorizing the product descriptions to create the desired product categories. The study has used data from the UCI machine learning repository with 541,909 POS type records and has applied the aforementioned unsupervised learning techniques to setup the profiles. It has unveiled product related and non-product related charateristics for the given POS data and has laid a novel foundation to construct the profiles to determine buying behaviour. Furthermore, these profiles can be used in segmentation of consumers, RSC specific promotions, and to predict future possibilities to minimize inventory related problems.

Keywords: Cluster analysis, Consumer buying behaviour, Unsupervised learning algorithms