DEVELOPMENT OF RECOMMENDATION ENGINE BY AUTHENTICATING ASSOCIATION RULES ON SALES DATA USING LOGISTIC REGRESSION

Sonu Verma, Indian Institute of Foreign Trade, New Delhi, India
Kavita Pabreja, Maharaja Surajmal Institute, GGSIP University, New Delhi, India
Arpan Kumar Patra, Indian Institute of Foreign Trade, New Delhi, India

ABSTRACT

In modern era of stiff competition, organizations are facing a lot of challenges to offer best deal to a customer so as to retain him and get good profit too at the same time. Numerous data mining and data analytics techniques have been used in the past to uncover hidden patterns from the buying behavior of customers in order to develop recommendation engine. Though association rule mining algorithms like Apriori, ECLAT (Equivalence Class Transformation), FP-growth (Frequent Pattern growth) find immense applications in the discovery of association rules, but one needs to assess the reliability of rules generated by these techniques. In this paper, an attempt has been made to validate the rules generated by Apriori algorithm by evaluating the Goodness of fit test in logistic regression. We have also analyzed pseudo R squared parameter to determine how much variation in dependent variable is explained by independent variables. The study is based on real dataset collected from the retail outlets in South Delhi area. With the application of Logistic regression algorithm, we are able to uncover certain novel facts about the association rules generated by Apriori algorithm. which will help retailers to optimize store layout, target precise campaigns, plan inventory to customer's need and to offer more profitable advertising and promotions.

Keywords: *Association rule mining, Apriori algorithm, Logistic regression, Pseudo R squared, goodness of fit.*