Getting your clustering right (Part I)

It takes not only sound technical knowledge, but also good understanding of business. We have split this topic into two articles because of the complexity of the topic. As the technique is very subjective in nature, getting the basics right is very critical.

*This article will take you through the basics of clustering. The next article will get into finer details of the technique and identify certain scenarios where the technique fails. The article will also introduce to a simple method to counter such scenarios*

*is the task of grouping*

**in the same group (called a cluster) are more similar (in some sense or another) to each other than to those in other groups (clusters). Following**

**one of the subjective modelling technique widely used in the industry. One of the examples of common Clustering usage is segmenting**

Analytics industry

It is one of the subjective modelling technique widely used in the industry. One of the examples of common Clustering usage is segmenting customer portfolio based on demographics, transaction behavior or other behavioral attributes.

*of the subjective modelling*

Analytics industry is dominated by objective modelling like decision tree and regression. If decision tree is capable of doing segmentation, do we even need such an open ended technique? The answer to this question is in one of the advantages of using clustering technique. Clustering generates natural clusters and is not dependent on any driving objective function. Hence such a cluster can be used to analyze the portfolio on different target attributes. For instance, say a decision tree is built on customer profitability in next 3 months. This segmentation cannot be used for making retention strategy for each segment. If segmentation were developed through clustering, both retention and profitability strategy can be built on these segments

REFERENCE

It takes not only sound technical knowledge, but also good understanding of business. We have split this topic into two articles because of the