Harmonic Data Preparation for Clustering and Classification

Authors : Ali Asheibi

Abstract : The rapid increase in the size of databases required to store power quality monitoring data has demanded new techniques for analysing and understanding the data. One suggested technique to assist in analysis is data mining. Preparing raw data to be ready for data mining exploration take up most of the effort and time spent in the whole data mining process. Clustering is an important technique in data mining and machine learning in which underlying and meaningful groups of data are discovered. Large amounts of harmonic data have been collected from an actual harmonic monitoring system in a distribution system in Australia for three years. This amount of acquired data makes it difficult to identify operational events that significantly impact the harmonics generated on the system. In this paper, harmonic data preparation processes to better understanding of the data have been presented. Underlying classes in this data has then been identified using clustering technique based on the Minimum Message Length (MML) method. The underlying operational information contained within the clusters can be rapidly visualised by the engineers. The C5.0 algorithm was used for classification and interpretation of the generated clusters.

Keywords : data mining, harmonic data, clustering, classification
Conference Title : ICEPES 2018 : International Conference on Electrical Power and Energy Systems
Conference Location : Sydney, Australia
Conference Dates : January 29-30, 2018