Chapter 84
Data Mining Applications in the Electrical Industry

Rubén Jaramillo Vacio
CFE – LAPEM & CIATEC – CONACYT, Mexico

Carlos Alberto Ochoa Ortiz Zezzatti
Juarez City University, México

Armando Rios
Institute Technologic of Celaya, Mexico

ABSTRACT

This chapter describes the experimental study partial discharges (PD) activities with artificial intelligent tools. The results present different patterns using a hybrid system with Self Organizing Maps (SOM) and Hierarchical clustering, this combination constitutes an excellent tool for exploration analysis of massive data such a partial discharge on underground power cables and electrical equipment. The SOM has been used for nonlinear feature extraction and the hierarchical clustering to visualization. The hybrid system is trained with different dataset using univariate phase-resolved distributions. The results show that the clustering method is fast, robust, and visually efficient.

INTRODUCTION

In the field of data analysis two terms are commonly encountered are supervised and unsupervised clustering methodologies. While supervised methods mostly deal with training classifiers for known symptoms, unsupervised clustering provides exploratory techniques for finding hidden patterns in data. What makes a system intelligent is its ability to analyze the data for efficient decision-making based on known or new cluster discovery, this is particularly important given the huge volumes of data being generated from the different systems every day.

Asset management (AM) is a concept used today for planning and operation of the electrical power system. The aim of AM is to handle physical assets in an optimal way in order to fulfill an organizations goal whilst considering risk where:

- The goal could be maximum asset value, maximum benefit or minimal life cycle cost

DOI: 10.4018/978-1-4666-2455-9.ch084

- The risk could be defined by the probability of failure occurrence and it consequence e.g., unavailability in power supply to customers

The partial discharge (PD) is a common phenomenon which occurs in insulation of high voltage, this definition is given in International Electrotechnical Commission (2000). In general, the partial discharges are in consequence of local stress in the insulation or on the surface of the insulation. This phenomenon has a damaging effect on the equipments, for example transformers, power cables, switchgears, and others. The first approach in a diagnosis is selecting the different features to classify measured PD activities into underlying insulation defects or source that generate PD's. The partial discharge measurement is a typical nondestructive test and it can be used to judge the insulation performance at the beginning of the service time taking into account the reduction of the performance during the service time by ageing, whereby the ageing depends on numerous parameters like electrical stress, thermal stress and mechanical stress. In particular for solid insulation like XLPE on power cables where a complete breakdown seriously damages the test object the partial discharge measurement is a tool for quality assessment (Wills, 1999).

Data Mining allows extraction of new information or knowledge and development of new methods for visualizing and analyzing the data held. Consequently, data mining should be seen as part of a well-considered information and asset management strategy. The database system industry has witnessed an evolutionary path in the development of the following functionalities: data collection and database creation, data management (including data storage and retrieval, and database transaction processing), and advanced data analysis (involving data warehousing and data mining). For instance, the early development of data collection and database creation mechanisms served as a prerequisite for later development of effective mechanisms for data storage and retrieval, and query and transaction processing. With numerous database systems offering query and transaction processing as common practice, advanced data analysis has naturally become the next target. This chapter outlines some work which has been done to achieve such optimal use of diagnostic data on underground transmission system. Increasing demands to fully exploit the capabilities of existing transmission equipment and systems require the best use from data assets in corporate databases and also require the development of new and comprehensive diagnostics for high voltage equipment.

In Allan, Birtwhistle, Blackburn, Groot, Gulski, and McGrail (2002) two variations of data mining approaches are discussed using examples from different diagnostic measurements and tests. The use of Kohonen mapping applied to existing and new data is illustrated using examples of dissolved gas analysis, tap-changer monitoring and insulator testing. In addition, two other examples are given illustrating the potential for data mining to produce useful information from comprehensive modern diagnostic monitoring of high voltage cables and of circuit breakers. A new task force within SC15 has been established to embark on a coordinated and focused approach to develop guidelines for data mining in practice.

In Strachan, Stephen, and McArthur (2007) proposes a data mining method for the analysis of condition monitoring data, and demonstrates this method in its discovery of useful knowledge from trip coil data captured from a population of in-service distribution circuit breakers and empirical UHF data captured from laboratory experiments simulating partial discharge defects typically found in HV transformers. This discovered knowledge then forms the basis of two separate decision support systems for the condition assessment/defect classification of these respective plant items.

Another very important article is Hernández-Mejia, Perkel, Harley, Begovic, Hampton, and