Expert Systems with Applications 36 (2009) 11808–11813

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Data mining approach for supply unbalance detection in induction motor

Abdülkadir Çakır, Hakan Çalış *, Ecir U. Küçüksille

Suleyman Demirel University, Faculty of Technical Education, Department of Electronics and Computer Education, E14 Blok Batı Kampusu, 32260 Isparta, Turkey

ARTICLE INFO

Keywords:
Fault detection
Induction motor
Voltage unbalance
Current zero crossing
Data mining

ABSTRACT

This paper describes an approach for detection of the supply unbalance condition in induction motors by using data mining process. Simulation results have shown that a good indicator of the fault is the amplitude of the second harmonic of the supply frequency component (2*f*) in the signal obtained by the differences in supply current zero crossing instants. In the study, linear regression (LR), pace regression (PR), sequential minimal optimization (SMO), M5 model tree, M5'Rules, KStar, additive regression and back propagation neural network (BPNN) models are applied within the data mining process for determining the condition of the motor supply voltage. All data mining algorithms were applied using WEKA software. The best result for the determination of the fault related dominant parameter was obtained by using the M5P algorithm model.

1. Introduction

The squirrel cage induction motor has been widely used in all kinds of industry because of its reliability, robustness and simplicity of its construction. However condition monitoring of the motor becomes a necessity to prevent any unplanned stops and breakdowns.

The effects of unbalanced supply voltage on the induction motors are stated as reduction on efficiency, temperature rises, and increments on the rotor losses. Voltage unbalance generates negative sequence component in the voltage due to a reverse rotating air gap field in the opposite direction of rotor rotation. The reduction in efficiency of a three phase induction motor supplied by unbalanced voltages was studied in Williams (1954). This leads to a temperature rise and shorter life time of the machine (Gafford, Duestenhoef, Mosher, 1959). Supplying the motor with unbalanced voltages also decreases the rating of the motor (Berndt & Schmitz, 1964). A basic method was applied in order to study the impact of unbalanced voltages on the losses and its negative effect on the insulation of the motor (Woll, 1975). Protection of the motor against these risks and adjusting the relays were proposed in reference (Cummings, Dunki-Jacobs, & Kerr, 1985).

In Cummings et al. (1985), Kersting and Philips (1997) and Kersting (2000), the impact of unbalanced voltage on the losses of the motor has been investigated. Kersting and Philips have presented a work about discussion of the effect of 0–5% unbalanced factors (Kersting & Philips, 1997). Unbalances on the terminal voltages cause a considerable effect on the stator and rotor copper losses. An increase in the unbalanced voltage at the terminals can increase rotor losses more than stator losses. The reason is that the rotor currents have a larger deviation than the stator currents. In addition, the resistance of the rotor bars is higher for unbalanced voltage (Gillbert, 1980). Even 5% unbalanced voltage increases the losses considerably, and lead to temperature rises that could damage the motor. Therefore, use of a motor connected to a large unbalanced voltage is not normally allowed. In reference (Lee, 1999), the temperature rise due to unbalanced voltages has been studied experimentally.

The results of this study showed that monitoring amplitude of the only single component (2*f*) is enough for the detection of unbalanced voltage in induction machines. It is implemented by measuring the amplitude of fault indicator component appeared in the spectrum of stator current zero crossing instants (ZCT). The frequency component is found to be function of the level of voltage unbalance. In this system, for the classification purpose and determination of the most dominant parameter in the detection of the supply unbalance data mining process is used.

Data mining applied in a wide range of applications such as in the field of prediction, production control, and in the area of medical. However, in the available literature, there are very limited numbers of studies on the fault detection in the electrical motors with using data mining approach.

The unbalance voltage is similar caused by extensive motor failure or even catastrophic phenomena, differing only in degree of unbalance. In order to detect unbalanced phase voltages or currents are readily identified, by the abnormality in induction motors several approaches are looking for the presence of negative phase sequence component. A small voltage unbalance produces a large negative sequence current flow in induction motor that will produce excess heating. The 5% voltage unbalance produces a stator negative sequence currents of 30% of full load of current. With this extra current, the motor experiences a 40% to 50% increase in temperature rise.

* Corresponding author. Tel./fax: +90 2462111479.
E-mail addresses: cakir@tef.sdu.edu.tr (A. Çakır), hcalis@tef.sdu.edu.tr, hcalis32@hotmail.com (H. Çalış), ecir@sdu.edu.tr (E.U. Küçüksille).

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.04.006

In this paper, the authors propose the data mining approach to investigate the level of unbalanced voltages. This approach was successfully used to predict the dominant parameter. In order to analyze the effects of the supply voltage unbalance, the three phase induction machine was simulated with 5% of unbalanced supply with and without any other fault.

2. Data mining process

2.1. Data mining

Data mining (DM) is the process of selecting, exploring and modeling large amounts of data in order to discover unknown patterns or relationships which provide a clear and useful result to the data analyst. Coined in the mid-1990s, the term data mining has today become a synonym for 'Knowledge Discovery in Databases' which, as proposed by Fayyad et al., emphasized the data analysis process rather than the use of specific analysis methods (Belazzi & Zupan, 2008; Fayyad, Piatetsky-Shapiro, & Smyth, 1996). Data mining problems are often solved by using a mosaic of different approaches drawn from computer science, including multi-dimensional databases, machine learning, soft computing and data visualization, and from statistics, including hypothesis testing, clustering, classification and regression techniques. The craft of data mining lies in the appropriate choice and combination of these techniques to efficiently and reliably solve a given problem (Belazzi & Zupan, 2008).

DM has several functions such as association analysis, classification and prediction, clustering analysis, outlier analysis and so on. Each of them may have several alternative data mining algorithms (Sun & Li, 2008). Some of these algorithms are linear regression, multi layer perceptron, KStar, decision trees, K-means. Data mining is accepted as a process and one of these processes is called as CRISP-DM.

2.2. CRISP-DM data mining process

To be successful in DM studies, systematic approach is required. Many DM software developers are provided different process models for the users. One of the well known products is developed by a consortium who sell DM software product and first users who adapted it in their business areas. This consortium consists of NCR System Engineering (USA-Danimarka), DaimlerChrysler AG (Germany), SPSS Inc. (USA) and OHRA Verzekeringen en Bank Groep B.V (Holland). They called it as CRISP-DM (Chapman et al., 2000; Fernandez, Zanakis, & Walczak, 2002).

CRISP-DM, is a process that defines basic steps in data mining, and it is shown as in Fig. 1.

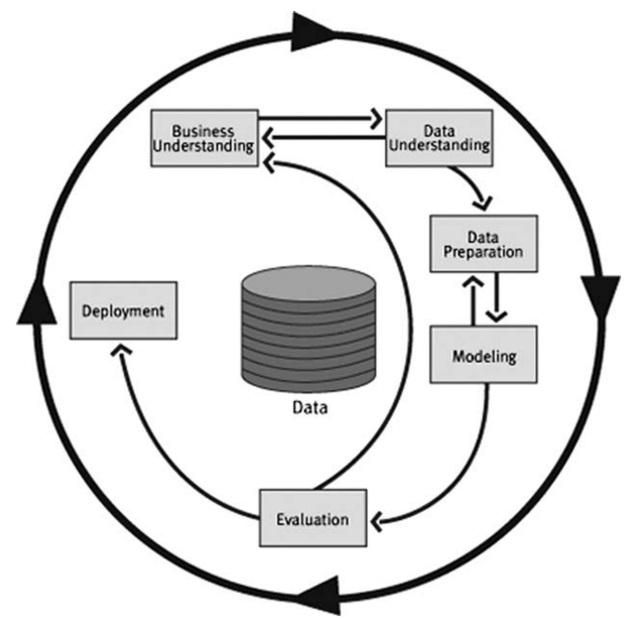

Fig. 1. CRISP-DM Data mining process (Chapman et al., 2000).

2.2.1. Business understanding

This initial phase focuses on understanding the project objectives and requirements from a business perspective, and then converting this knowledge into a data mining problem definition, and a preliminary project plan designed to achieve the objectives (Fernandez et al., 2002; Wirth & Hipp, 2000).

2.2.2. Data understanding

Second stage is called as the understanding of the data content. Data stored in a various environments such as Microsoft stores data in a hundred different database and 70 different data warehouses. First step is getting the meaningful data from those database or data warehouses for the selected application. In the meantime, in this phase data quality and discovering first insights into the data are seen as two important aspects.

2.2.3. Data preparation

The data preparation phase covers all activities to construct the final data set or the data that will be fed into the modeling tool from the initial raw data. Tasks include table, record, and attribute selection, as well as transformation and cleaning of data for modeling tools. The five steps in data preparation are the selection of data, the cleaning of data, the construction of data, the integration of data, and the formatting of data (Shearar, 2000).

The purpose of the cleaning stage is selecting unsuitable or incorrectly entered data in the data. For example, filling the mean value for the instead of the incomplete data or erasing abnormal data records outside of the normal dispersion area assuming the meaningful data is in the normal distribution (Fernandez et al., 2002).

Data conversion is required for recording data in different formats or values since some data mining algorithms work only with data in digital format. In this case it needs to convert data in text format to the digital one. Purpose of the feature selection is determination of the most dominant parameters in forecasting a value. It might be assigned many features to estimate a value. However, it is not always easy to collect the determined data. For this case, by finding the effective properties data acquisition can be fast and simple. In addition, data are divided into two groups as training and testing data.

2.2.4. Modeling

In this phase, various modeling techniques are selected and applied and their parameters are calibrated to optimal values. Typically, several techniques exist for the same data mining problem type. Some techniques have specific requirements on the form of data. Therefore, stepping back to the data preparation phase may be necessary. Modeling steps include the selection of the modeling technique, the generation of test design, the creation of models, and the assessment of models (Shearar, 2000). If the task is fully accomplished, in this case, selection of the correct algorithm is much easier. Each task requires different algorithms and it is not known which one gives the best result without constructing the model. It may be only possible to guess according to the condition of the data in hand. If there is a linear relationship between whole input and estimation variables, choosing the decision tree algorithm can be good choice. If there is a complex relation among the variables, in this situation neural network algorithm can be selected (Tang & MacLennan, 2005).