A fault diagnosis methodology for an external gear pump with the use of Machine Learning classification algorithms: Support Vector Machine and Multilayer Perceptron

Kayalvizhi Lakshmanan[1,2,*], Antonio J Gil[2], Ferdinando Auricchio[1], Fabrizio Tessicini[3]

[1] Department of Civil Engineering and Architecture, University of Pavia, Pavia, Italy.
kayalvizhi.lakshmanan@unipv.it, 937129@swansea.ac.uk
[2] Zienkiewicz Centre for Computational Engineering, College of Engineering, Swansea University, Swansea, UK.
[3]F-Lab, 20094, Corsico, Milano, Italy.
auricchi@unipv.it, a.j.gil@swansea.ac.uk, f.tessicini@f-lab.org

Abstract

This paper presents a fault diagnosis Machine Learning (ML) computational strategy for an external gear pump. The method uses supervised learning of descriptive features. The focus is on two types of ML nonlinear multi-class classification algorithms: Support Vector Machine (SVM) and Multilayer Perceptron (MLP) algorithms. Although significant work has been reported by previous authors, it is still difficult to optimise *ab initio* the choice of the hyper parameters (ML method dependent) for each specific application. For instance, the type of SVM kernel function or the selection of the MLP activation function and the optimum number of hidden layers (and neurons). As it is well known, reliability of ML algorithms is strongly dependent upon the existence of a sufficiently large quantity of high-quality training data. In our case, and in the absence of experimental data, high-fidelity in-silico data (generated via the software PumpLinx®) have been used for the training of the underlying ML metamodel. A variety of working conditions are recreated, ranging from healthy to various kinds of faulty scenarios (i.e. clogging, radial gap variations, viscosity variations). In addition, noise perturbation has been considered in order to increase the sample data available for ML training.
This paper explores and compares the use of SVM and MLP algorithms for predictive maintenance. To reduce the high computational cost during the training stage in the MLP algorithm, some predefined network architectures, like 2^n neurons per hidden layer, are used to speed up the identification of the precise number of neurons (shown to be useful when the sample data set is sufficiently large). A series of benchmark tests are presented, enabling to conclude that the use of wavelet features and SVM or MLP algorithms can provide the best accuracy for classification.

Key words: *Fault diagnosis; Wavelet feature extraction; Noise perturbation; Support Vector Machine; MultiLayer Perceptron*

1 Introduction

Gear pumps are comprised of sophisticated gear arrangements used to effectively pump a fluid under a variety of working conditions. Due to their sophisticated and compact design, they can be easily fit inside complex industry equipment reducing the manufacturing time and associated cost. The sudden failure of any industry component can have negative consequences to industry in terms of time and workflow. As a result, fault diagnosis has emerged as a very useful tool in order to predict failure and minimise their industrial impact. Recent developments in fault diagnosis are mostly based on vibration signals extracted from sensors distributed within the equipment [3, 4]. Along these lines, fault diagnosis of a centrifugal pump was recently carried out using MLP with a genetic algorithm and SVM with continuous wavelet transform[2].
In this work, we propose a fault diagnosis methodology using supervised learning ML algorithms, specifically, SVM and MLP algorithms, where the offline learning phase uses three different kinds of descriptive features: real time features, frequency features and wavelet features. The method was developed following three steps; (1) generation of high-fidelity in silico data using a Computational Fluid Dynamics (CFD) model of a gear pump, (2) use of feature extraction techniques to extract descriptive features from real time series sample data and (3) training of ML classification algorithms using a subset sample data and enhancement of the classification accuracy by optimisation of hyper-parameters.

2 Methodology

The CFD model has been built using the commercial code 'PumpLinx®'. The main objective is the computational simulation of the gear geometry shown in Figure 1. The conservation of mass and momentum are solved to obtain the relevant flow field. The vapor fraction equation is added to account for cavitation phenomena, in order to accurately model the pressure ripple across space and time. These quasi compressible equations are modelled numerically using a cell centred Finite Volume method [1]. Using the CFD model a variety of working conditions including faulty scenarios (i.e. clogging, radial gap variations, viscosity variations) are recreated. Figure 2 shows the faulty scenarios implemented to examine the machine's health under different working conditions. Pressure, flow rate and torque are monitored for various working conditions of the pump, which are used as raw real time data for subsequent feature extraction.

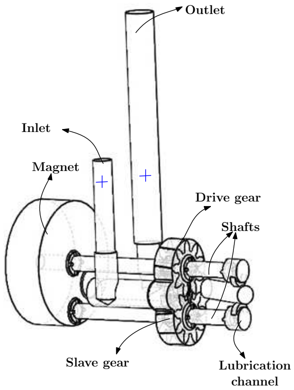

Figure 1: Geometry of the gear pump

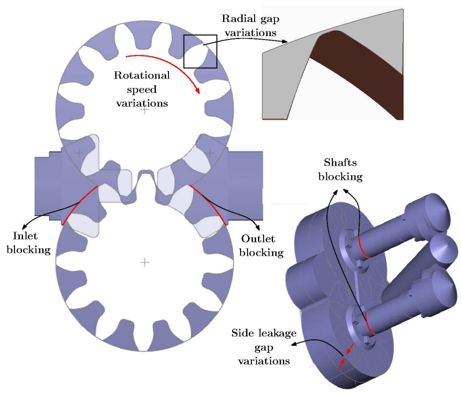

Figure 2: Implemented faulty scenarios

SVM is a type of ML algorithm that provides a globally optimal solution for any binary/multi class classification problems. For the case of a linear binary classification problem, a decision hypothesis $\mathcal{H}$ is sought which separates n sample data points $\boldsymbol{x}_i = [x_1, x_2, \dots x_m]_i$ into two distinct categories,

$$\mathcal{H} = \{\boldsymbol{x}_i \mapsto \text{sign}\,(\boldsymbol{w} \cdot \boldsymbol{x}_i + b) : \boldsymbol{x}_i \in \mathbb{R}^m, \boldsymbol{w} \in \mathbb{R}^m, b \in \mathbb{R}\}, \tag{1}$$

where $\boldsymbol{w}$ is the so-called weight vector, b is the bias and $\boldsymbol{w} \cdot \boldsymbol{x} + b = 0$ is the equation of the hyperplane separating the population data. For a more generic non-linear classification problem, kernel functions (Gaussian, Sigmoid and polynomial) are used to map the training sample data into a higher dimensional space, where the classification problem transforms into a linear one.
MLP algorithms are inspired by biological neural networks, with an architecture comprised of layers of interconnected neurons. The MLP consists of an input layer, hidden layers and an output layer of neurons, all interconnected via a feed-forward mechanism. During the training process, the input data is allocated to the input layer, the information is passed through hidden layers and directed to the output layer. Back propagation is then used to adapt the weights of the neurons, starting from the output layer and propagating backwards by adapting the weights, layer by layer, depending on an estimation error. Moreover, MLP includes activation functions that introduce further non-linear properties to the network. Typically used activation functions are Sigmoid and hyperbolic tangent function and, only for classification algorithms, the softmax function is used in the output layer. In general, the output of a neuron i within a layer h is calculated as

$$y_i^{(h)} = f^{(h)}\left(\sum_{j=1}^{n_{(h-1)}} W_{ij}^{(h)} y_j^{(h-1)} + {b_i}^{(h)}\right); \quad i = 1 \dots n_{(h)}, \tag{2}$$

where $W_{ij}^{(h)}$ is the weight matrix coefficient, $b_i^{(h)}$ is the bias and $f^{(h)}$ is the activation function.

For classification problems (SVM and MLP), the metric used for training (or testing) during the offline phase is the accuracy of the predicted target values y_i against the true target values t_i, namely

$$\text{Accuracy} = \frac{1}{n}\sum_{i=1}^{n}\left(\begin{cases}1 & \text{if} \quad y_i = t_i \\ 0 & \text{otherwise}\end{cases}\right), \tag{3}$$

where n is the number of samples in the training (or test) set.

3 Numerical results

The results extracted from the CFD model have been used as a data set for the two ML algorithms described above. To improve the learning phase of the ML algorithms, further in-silico data have been manufactured using a noise perturbation method. Figure 3 shows the process of data manufacturing using the noise perturbation method. The addition of noise is carried out by perturbing the frequency content by an error function constructed in the following manner

$$\hat{x}^{\text{mod}}(\tilde{\omega}_j) = \hat{x}(\tilde{\omega}_j)[1 + \alpha\epsilon(\tilde{\omega}_j)], \tag{4}$$

where $\hat{x}(\tilde{\omega}_j)$ is the original frequency content (for a selected frequency $\tilde{\omega}_j$) obtained from a discrete Fourier transform, α is a scalar parameter tuned depending on the amount of noise added and $\epsilon(\omega)$ is a monotonically increasing function of the frequency ω (larger noise added in larger frequencies). From the manufactured sample dataset, data features have been extracted, such as time features (mean, standard deviation, skewness), frequency features (spectral density, spectral kurtosis) and wavelet transform features (wavelet packet transform or sub-band tree). The extracted features are standardised and divided into training data (70%) and test data (30%). The training data set contains n input standardised variables $\boldsymbol{x}_i$ and response categorical variables t_i. After the training, the test data set has been classified with the use of the two trained ML models. The accuracy of classification has been computed using Equation (3). Table 1 shows the accuracy of SVM and MLP for pressure at outlet, flow rate at outlet and torque of the gear. Table 2 shows the analysis of quantification of noise addition by increasing the noise and calculating the accuracy using SVM. Figures 4, 5 show the accuracy plot for different hidden neurons and hidden layers, which is used to optimise the number of hidden neurons and hidden layers and avoid over fitting.

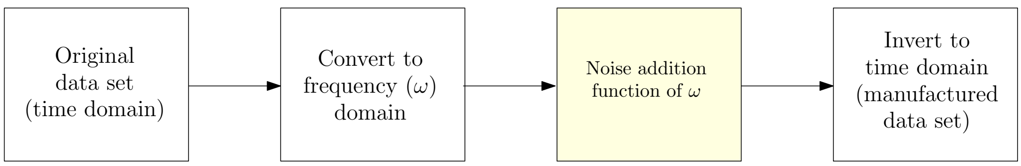

Figure 3: Data manufacturing using noise perturbation method

		Time features	Frequency features	Wavelet
Pressure	SVM	92.1 %	93 %	99.1 %
	MLP	95.5 (1) %	97.3 (1) 99.1 (2) %	100 (1) %
Flow rate	SVM	90.4 %	87.7 %	100 %
	MLP	96.4 (1) %	98.23 (1) 99.1 (2) %	100 (1) %
Torque	SVM	82.5 %	83.3 %	93 %
	MLP	84.07 (2) 90.2 (3) %	91.15 (2) 92.03 (3) %	100(1) %

Table 1: Classification accuracy of SVM and MLP of based on measured variables; (1), (2), (3)-number of hidden layers in MLP

4 Conclusions

Feature extraction and classification of various gear pump conditions using SVM and MLP algorithm is performed, where the use of wavelet transform for feature extraction (with either SVM or

		Time features	Frequency features	Wavelet features
Pressure	$\alpha = 1$	92.1 %	93 %	99.1 %
	$\alpha = 2$	86 %	81.6 %	98.2 %
	$\alpha = 3$	82.5 %	80.7 %	98.2 %
Flow rate	$\alpha = 1$	90.4 %	87.7 %	100 %
	$\alpha = 2$	89.5 %	86.8 %	10 0%
	$\alpha = 3$	89.5 %	83.3 %	99.1 %
Torque	$\alpha = 1$	82.5 %	83.3 %	93 %
	$\alpha = 2$	77.2 %	79.8 %	87.7 %
	$\alpha = 3$	76.3 %	68.4 %	87.7 %

Table 2: SVM classification accuracy with respect to increment of the noise added (α in Equation (4)) during data manufacturing

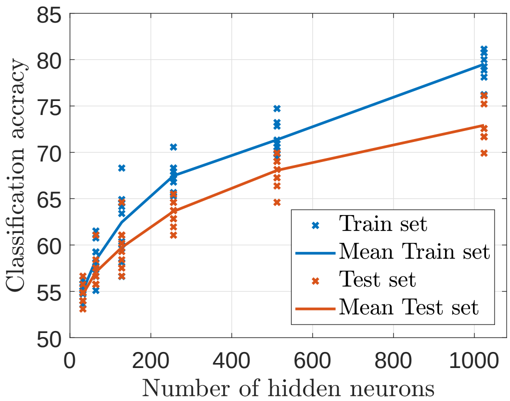

Figure 4: Classification accuracy of torque (time features) with respect to number of hidden neurons (32,64,128,256,512,1024) for 1 hidden layer using MLP

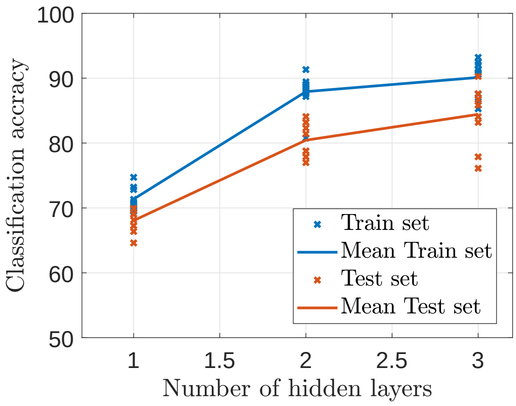

Figure 5: Classification accuracy of torque (time features) with respect to number of hidden layers (1, 2, 3) using MLP

MLP) gives the highest accuracy. Moreover, wavelet features with MLP provides the highest classification accuracy. The use of torque data (a magnitude preferred by industry) shows relatively similar accuracy to that obtained using other input data such as pressure and flow rate values.

References

[1] H. Ding, F. C. Visser, Y. Jiang, M. Furmanczyk (2011): *Demonstration and validation of a 3D CFD simulation tool predicting pump performance and cavitation for industrial applications.* Journal of fluids engineering, 133(1).

[2] M. A. S. ALTobi, G. Bevan, P. Wallace, D. Harrison, K.P. Ramachandran (2019): *Fault diagnosis of a centrifugal pump using MLP-GABP and SVM with CWT.* Engineering Science and Technology, an International Journal, vol. 22, no. 3, pp. 854-861.

[3] Z. Liu, Y. Liu, H. Shan, B. Cai, Q. Huang (2015): *A fault diagnosis methodology for gear pump based on EEMD and Bayesian network.* PloS one, 10(5).

[4] T. Wang, Q. Han, F. Chu, Z. Feng (2019): *Vibration based condition monitoring and fault diagnosis of wind turbine planetary gearbox: A review.* Mechanical Systems and Signal Processing, 126, 662-685.